UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

# NOTICE OF ALLOWANCE AND FEE(S) DUE

36601 7590 03/18/2025
Wasserbauer Law LLC
PO BOX 382
Collinsville, CT 06022

| EXAMINER | |
|---|---|
| OSINSKI, MICHAEL S | |
| ART UNIT | PAPER NUMBER |
| 2674 | |

DATE MAILED: 03/18/2025

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 17/609,911 | 11/09/2021 | Lucinda Lewis | P1409US01 | 1651 |

TITLE OF INVENTION: METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

| APPLN. TYPE | ENTITY STATUS | ISSUE FEE DUE | PUBLICATION FEE DUE | PREV. PAID ISSUE FEE | TOTAL FEE(S) DUE | DATE DUE |
|---|---|---|---|---|---|---|
| nonprovisional | SMALL | $516 | $0.00 | $0.00 | $516 | 06/18/2025 |

**THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.**

**THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE DOES NOT REFLECT A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE IN THIS APPLICATION. IF AN ISSUE FEE HAS PREVIOUSLY BEEN PAID IN THIS APPLICATION (AS SHOWN ABOVE), THE RETURN OF PART B OF THIS FORM WILL BE CONSIDERED A REQUEST TO REAPPLY THE PREVIOUSLY PAID ISSUE FEE TOWARD THE ISSUE FEE NOW DUE.**

**HOW TO REPLY TO THIS NOTICE:**

I. Review the ENTITY STATUS shown above. If the ENTITY STATUS is shown as SMALL or MICRO, verify whether entitlement to that entity status still applies.

If the ENTITY STATUS is the same as shown above, pay the TOTAL FEE(S) DUE shown above.

If the ENTITY STATUS is changed from that shown above, on PART B - FEE(S) TRANSMITTAL, complete section number 5 titled "Change in Entity Status (from status indicated above)".

For purposes of this notice, small entity fees are 40% the amount of undiscounted fees, and micro entity fees are 20% the amount of undiscounted fees.

II. PART B - FEE(S) TRANSMITTAL, or its equivalent, must be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed. If an equivalent of Part B is filed, a request to reapply a previously paid issue fee must be clearly made, and delays in processing may occur due to the difficulty in recognizing the paper as an equivalent of Part B.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Mail Stop ISSUE FEE unless advised to the contrary.

**IMPORTANT REMINDER: Maintenance fees are due in utility patents issuing on applications filed on or after Dec. 12, 1980. It is patentee's responsibility to ensure timely payment of maintenance fees when due. More information is available at www.uspto.gov/PatentMaintenanceFees.**

# PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), by mail or fax, or via the USPTO patent electronic filing system.

By mail, send to: Mail Stop ISSUE FEE
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

By fax, send to: (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications. **Because electronic patent issuance may occur shortly after issue fee payment, any desired continuing application should preferably be filed prior to payment of this issue fee in order not to jeopardize copendency.**

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

36601 7590 03/18/2025

Wasserbauer Law LLC
PO BOX 382
Collinsville, CT 06022

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

**Certificate of Mailing or Transmission**

I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being transmitted to the USPTO via the USPTO patent electronic filing system or by facsimile to (571) 273-2885, on the date below.

(Typed or printed name)

(Signature)

(Date)

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 17/609,911 | 11/09/2021 | Lucinda Lewis | P1409US01 | 1651 |

TITLE OF INVENTION: METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

| APPLN. TYPE | ENTITY STATUS | ISSUE FEE DUE | PUBLICATION FEE DUE | PREV. PAID ISSUE FEE | TOTAL FEE(S) DUE | DATE DUE |
|---|---|---|---|---|---|---|
| nonprovisional | SMALL | $516 | $0.00 | $0.00 | $516 | 06/18/2025 |

| EXAMINER | ART UNIT | CLASS-SUBCLASS |
|---|---|---|
| OSINSKI, MICHAEL S | 2674 | 382-224000 |

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).

❑ Change of correspondence address (or Change of Correspondence Address form PTO/AIA/122 or PTO/SB/122) attached.

❑ "Fee Address" indication (or "Fee Address" Indication form PTO/AIA/47 or PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) The names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) The name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

1____________________

2____________________

3____________________

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document must have been previously recorded, or filed for recordation, as set forth in 37 CFR 3.11 and 37 CFR 3.81(a). Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE

(B) RESIDENCE: (CITY and STATE OR COUNTRY)

Please check the appropriate assignee category or categories (will not be printed on the patent) : ❑ Individual ❑ Corporation or other private group entity ❑ Government

4a. Fees submitted: ❑Issue Fee ❑Publication Fee (if required)

4b. Method of Payment: *(Please first reapply any previously paid fee shown above)*

❑ Electronic Payment via the USPTO patent electronic filing system ❑ Enclosed check ❑ Non-electronic payment by credit card (Attach form PTO-2038)

❑ The Director is hereby authorized to charge the required fee(s), any deficiency, or credit any overpayment to Deposit Account No. __________

5. **Change in Entity Status** (from status indicated above)

❑ Applicant certifying micro entity status. See 37 CFR 1.29

NOTE: Absent a valid certification of Micro Entity Status (see forms PTO/SB/15A and 15B), issue fee payment in the micro entity amount will not be accepted at the risk of application abandonment.

❑ Applicant asserting small entity status. See 37 CFR 1.27

NOTE: If the application was previously under micro entity status, checking this box will be taken to be a notification of loss of entitlement to micro entity status.

❑ Applicant changing to regular undiscounted fee status.

NOTE: Checking this box will be taken to be a notification of loss of entitlement to small or micro entity status, as applicable.

NOTE: This form must be signed in accordance with 37 CFR 1.31 and 1.33. See 37 CFR 1.4 for signature requirements and certifications.

Authorized Signature ____________________ Date ____________________

Typed or printed name ____________________ Registration No. ____________________

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 17/609,911 | 11/09/2021 | Lucinda Lewis | P1409US01 | 1651 |

36601 7590 03/18/2025

Wasserbauer Law LLC
PO BOX 382
Collinsville, CT 06022

| EXAMINER | |
|---|---|
| OSINSKI, MICHAEL S | |
| ART UNIT | PAPER NUMBER |
| 2674 | |

DATE MAILED: 03/18/2025

## Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)

(Applications filed on or after May 29, 2000)

The Office has discontinued providing a Patent Term Adjustment (PTA) calculation with the Notice of Allowance.

Section 1(h)(2) of the AIA Technical Corrections Act amended 35 U.S.C. 154(b)(3)(B)(i) to eliminate the requirement that the Office provide a patent term adjustment determination with the notice of allowance. See Revisions to Patent Term Adjustment, 78 Fed. Reg. 19416, 19417 (Apr. 1, 2013). Therefore, the Office is no longer providing an initial patent term adjustment determination with the notice of allowance. The Office will continue to provide a patent term adjustment determination with the Issue Notification Letter that is mailed to applicant approximately three weeks prior to the issue date of the patent, and will include the patent term adjustment on the patent. Any request for reconsideration of the patent term adjustment determination (or reinstatement of patent term adjustment) should follow the process outlined in 37 CFR 1.705.

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at 1-(888)-786-0101 or (571)-272-4200.

## OMB Clearance and PRA Burden Statement for PTOL-85 Part B

The Paperwork Reduction Act (PRA) of 1995 requires Federal agencies to obtain Office of Management and Budget approval before requesting most types of information from the public. When OMB approves an agency request to collect information from the public, OMB (i) provides a valid OMB Control Number and expiration date for the agency to display on the instrument that will be used to collect the information and (ii) requires the agency to inform the public about the OMB Control Number's legal significance in accordance with 5 CFR 1320.5(b).

The information collected by PTOL-85 Part B is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 30 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450. Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

## Privacy Act Statement

**The Privacy Act of 1974 (P.L. 93-579)** requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. The United States Patent and Trademark Office (USPTO) collects the information in this record under authority of 35 U.S.C. 2. The USPTO's system of records is used to manage all applicant and owner information including name, citizenship, residence, post office address, and other information with respect to inventors and their legal representatives pertaining to the applicant's/owner's activities in connection with the invention for which a patent is sought or has been granted. The applicable Privacy Act System of Records Notice for the information collected in this form is COMMERCE/PAT-TM-7 Patent Application Files, available in the Federal Register at 78 FR 19243 (March 29, 2013).
https://www.govinfo.gov/content/pkg/FR-2013-03-29/pdf/2013-07341.pdf

Routine uses of the information in this record may include disclosure to:

1) law enforcement, in the event that the system of records indicates a violation or potential violation of law;
2) a federal, state, local, or international agency, in response to its request;
3) a contractor of the USPTO having need for the information in order to perform a contract;
4) the Department of Justice for determination of whether the Freedom of Information Act (FOIA) requires disclosure of the record;
5) a Member of Congress submitting a request involving an individual to whom the record pertains, when the individual has requested the Member's assistance with respect to the subject matter of the record;
6) a court, magistrate, or administrative tribunal, in the course of presenting evidence, including disclosures to opposing counsel in the course of settlement negotiations;
7) the Administrator, General Services Administration (GSA), or their designee, during an inspection of records conducted by GSA under authority of 44 U.S.C. 2904 and 2906, in accordance with the GSA regulations and any other relevant (i.e., GSA or Commerce) directive, where such disclosure shall not be used to make determinations about individuals;
8) another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c));
9) the Office of Personnel Management (OPM) for personnel research purposes; and
10) the Office of Management and Budget (OMB) for legislative coordination and clearance.

If you do not furnish the information requested on this form, the USPTO may not be able to process and/or examine your submission, which may result in termination of proceedings, abandonment of the application, and/or expiration of the patent.

| *Notice of Allowability* | **Application No.**<br>17/609,911 | **Applicant(s)**<br>Lewis, Lucinda | |
|---|---|---|---|
| | **Examiner**<br>MICHAEL S OSINSKI | **Art Unit**<br>2674 | **AIA (FITF) Status**<br>Yes |

***-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--***

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1.☑ This communication is responsive to communications filed on 1/23/2025.
☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on ______.

2.☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on ______; the restriction requirement and election have been incorporated into this action.

3.☑ The allowed claim(s) is/are 1-20 . As a result of the allowed claim(s), you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see **http://www.uspto.gov/patents/init_events/pph/index.jsp** or send an inquiry to **PPHfeedback@uspto.gov.**

4.☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

**Certified copies:**

a) ☐All b) ☐ Some* c) ☐ None of the:

1. ☐ Certified copies of the priority documents have been received.
2. ☐ Certified copies of the priority documents have been received in Application No. ______ .
3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).

* Certified copies not received: ______ .

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
**THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.**

5.☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date ______ .

**Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).**

6.☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

**Attachment(s)**

1.☑ Notice of References Cited (PTO-892)
2.☐ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date ______.
3.☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material ______.
4.☐ Interview Summary (PTO-413), Paper No./Mail Date. ______.
5. ☐ Examiner's Amendment/Comment
6. ☑ Examiner's Statement of Reasons for Allowance
7. ☐ Other ______.

/MICHAEL S OSINSKI/
Primary Examiner, Art Unit 2674

### *Notice of Pre-AIA or AIA Status*

The present application, filed on or after March 16, 2013, is being examined under the first inventor to file provisions of the AIA.

## DETAILED ACTION

1. The following Office action is in response to communications filed on 1/23/2025. Claims 1-20 are currently pending within this application.

### *Claim Rejections*

2. The previous rejections of the claims under 35 USC 101, 35 USC 102, and 35 USC 103, are withdrawn in response to amended claims filed on 1/23/2025.

### *Allowable Subject Matter*

3. Claims 1-20 are allowed.

4. The following is an Examiner’s statement for the reasons of allowance:

5. Independent claims 1, 18, and 20 are directed towards a method, non-transitory CRM, and system that include/perform the operations of at least “processing multimodal input data based upon multimodal and/or multiclass training data, the method comprising: training, by a processor, a convolutional neural network (CNN) using the multimodal and/or multiclass training data, which includes authenticated data and a taxonomy, to classify at least one data object in a media asset, authenticate the media

asset received by the CNN, and verify the data object received by the CNN; training, by a processor, a recurrent neural network (RNN) using the multimodal and/or multiclass training data; receiving, by the processor, a query comprising input data; classifying, by the trained CNN, the input data with respect to the authenticated data and elements of one or more classes of the taxonomy; generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of one or more classes of the taxonomy comprising a closest match to the input data; and displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content".

The cited and considered prior art, specifically those cited and relied upon in the previous Office action(s) as well as the instant Office action, fails to anticipate or render obvious either alone or in combination with proper motivation the above mentioned claimed limitations of the instant application in conjunction with the other claimed limitations as claimed in the instant application, and therefore independent claims 1, 18, and 20 are allowed.

Claims 2-17 and 19 are allowed for being dependent upon allowed base claims 1 and 18.

## ***Conclusion***

6. Any inquiry concerning this communication or earlier communications from the examiner should be directed to MICHAEL S OSINSKI whose telephone number is (571) 270-3949. The examiner can normally be reached on Monday - Thursday, 10:00am - 6:00pm. If attempts to reach the examiner by telephone are unsuccessful, the

examiner’s supervisor, Oneal Mistry can be reached on (313) 446-4912. The fax phone number for the organization where this application or proceeding is assigned is (571)-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

MO
/MICHAEL S OSINSKI/
Primary Examiner, Art Unit 2664

3/5/2024

| ***Notice of References Cited*** | Application/Control No. 17/609,911 | Applicant(s)/Patent Under Reexamination Lewis, Lucinda | |
|---|---|---|---|
| | Examiner MICHAEL S OSINSKI | Art Unit 2674 | Page 1 of 2 |

**U.S. PATENT DOCUMENTS**

| * | | Document Number Country Code-Number-Kind Code | Date YYYY-MM-DD | Name | CPC Classification | US Classification |
|---|---|---|---|---|---|---|
| * | A | US-20170098153-A1 | 2017-04-06 | Mao; Junhua | G06N3/045 | 1/1 |
| * | B | US-20170061250-A1 | 2017-03-02 | Gao; Jianfeng | G06V20/70 | 1/1 |
| * | C | US-20180204120-A1 | 2018-07-19 | REI; Marek | G06N3/04 | 1/1 |
| * | D | US-10423874-B2 | 2019-09-24 | Mao; Junhua | G06N3/044 | 1/1 |
| * | E | US-10198671-B1 | 2019-02-05 | Yang; Linjie | G06V10/768 | 1/1 |
| * | F | US-10691899-B2 | 2020-06-23 | Lubbers; Niels | G06N3/044 | 1/1 |
| * | G | US-20200302234-A1 | 2020-09-24 | Walters; Austin | G06F16/9035 | 1/1 |
| * | H | US-11055557-B2 | 2021-07-06 | Chatterjee; Anirban | G06F16/3334 | 1/1 |
| * | I | US-10540378-B1 | 2020-01-21 | Hsiao; Edward | G06F18/24 | 1/1 |
| * | J | US-11257217-B2 | 2022-02-22 | Shlens; Jonathon | G06N3/044 | 1/1 |
| * | K | US-11636330-B2 | 2023-04-25 | Krishnan; Abhinandan | G06N3/045 | 706/20 |
| * | L | US-20190311188-A1 | 2019-10-10 | Qing; Linbo | G06V10/82 | 1/1 |
| * | M | US-20210256365-A1 | 2021-08-19 | Wang; Wenmin | G06F16/583 | 1/1 |

**FOREIGN PATENT DOCUMENTS**

| * | | Document Number Country Code-Number-Kind Code | Date YYYY-MM-DD | Country | Name | CPC Classification |
|---|---|---|---|---|---|---|
| | N | | | | | |
| | O | | | | | |
| | P | | | | | |
| | Q | | | | | |
| | R | | | | | |
| | S | | | | | |
| | T | | | | | |

**NON-PATENT DOCUMENTS**

| * | | Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages) |
|---|---|---|
| | U | |
| | V | |
| | W | |
| | X | |

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in YYYY-MM-DD format are publication dates. Classifications may be US or foreign.

| *Notice of References Cited* | Application/Control No. 17/609,911 | Applicant(s)/Patent Under Reexamination Lewis, Lucinda | |
|---|---|---|---|
| | Examiner MICHAEL S OSINSKI | Art Unit 2674 | Page 2 of 2 |

**U.S. PATENT DOCUMENTS**

| * | | Document Number Country Code-Number-Kind Code | Date YYYY-MM-DD | Name | CPC Classification | US Classification |
|---|---|---|---|---|---|---|
| * | A | US-20220233129-A1 | 2022-07-28 | LIU; HUAFENG | G06N3/044 | 1/1 |
| * | B | US-20210248400-A1 | 2021-08-12 | Cronje; Jaco | G06V20/597 | 1/1 |
| | C | | | | | |
| | D | | | | | |
| | E | | | | | |
| | F | | | | | |
| | G | | | | | |
| | H | | | | | |
| | I | | | | | |
| | J | | | | | |
| | K | | | | | |
| | L | | | | | |
| | M | | | | | |

**FOREIGN PATENT DOCUMENTS**

| * | | Document Number Country Code-Number-Kind Code | Date YYYY-MM-DD | Country | Name | CPC Classification |
|---|---|---|---|---|---|---|
| | N | | | | | |
| | O | | | | | |
| | P | | | | | |
| | Q | | | | | |
| | R | | | | | |
| | S | | | | | |
| | T | | | | | |

**NON-PATENT DOCUMENTS**

| * | | Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages) |
|---|---|---|
| | U | |
| | V | |
| | W | |
| | X | |

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in YYYY-MM-DD format are publication dates. Classifications may be US or foreign.

| *Index of Claims* | Application/Control No. 17/609,911 | Applicant(s)/Patent Under Reexamination Lewis, Lucinda |
|---|---|---|
| | Examiner MICHAEL S OSINSKI | Art Unit 2674 |

| ✓ | Rejected | - | Cancelled | N | Non-Elected | A | Appeal |
|---|---|---|---|---|---|---|---|
| = | Allowed | ÷ | Restricted | I | Interference | O | Objected |

**CLAIMS**

☑ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

| CLAIM | | DATE | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Final | Original | 07/18/2024 | 03/05/2025 | | | | | | | |
| | 1 | ✓ | = | | | | | | | |
| | 2 | ✓ | = | | | | | | | |
| | 3 | ✓ | = | | | | | | | |
| | 4 | ✓ | = | | | | | | | |
| | 5 | ✓ | = | | | | | | | |
| | 6 | ✓ | = | | | | | | | |
| | 7 | ✓ | = | | | | | | | |
| | 8 | ✓ | = | | | | | | | |
| | 9 | ✓ | = | | | | | | | |
| | 10 | ✓ | = | | | | | | | |
| | 11 | ✓ | = | | | | | | | |
| | 12 | ✓ | = | | | | | | | |
| | 13 | ✓ | = | | | | | | | |
| | 14 | ✓ | = | | | | | | | |
| | 15 | ✓ | = | | | | | | | |
| | 16 | ✓ | = | | | | | | | |
| | 17 | ✓ | = | | | | | | | |
| | 18 | ✓ | = | | | | | | | |
| | 19 | ✓ | = | | | | | | | |
| | 20 | ✓ | = | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |

# Search Notes

| Application/Control No. | Applicant(s)/Patent Under Reexamination |
|---|---|
| 17/609,911 | Lewis, Lucinda |
| **Examiner** | **Art Unit** |
| MICHAEL S OSINSKI | 2674 |

| CPC - Searched* | | |
|---|---|---|
| **Symbol** | **Date** | **Examiner** |
| G06N3/08 G06V10/454 G06V10/764 G06V10/82 G06V20/20 G06V20/56 G06N3/042 G06N3/044 G06N3/048 G06F16/433 G06F16/434 G06F16/438 G06F16/90332 G06F18/24143 G06N3/045 G06Q30/018 G06Q30/0241 G06N5/02 G06V2201/08 H04L63/02 | 07/17/2024 | MO |

| CPC Combination Sets - Searched* | | |
|---|---|---|
| **Symbol** | **Date** | **Examiner** |
| | | |

| US Classification - Searched* | | | |
|---|---|---|---|
| **Class** | **Subclass** | **Date** | **Examiner** |
| | | | |

* See search history printout included with this form or the SEARCH NOTES box below to determine the scope of the search.

| Search Notes | | |
|---|---|---|
| **Search Notes** | **Date** | **Examiner** |
| PE2E Search | 07/16/2024 | MO |
| Google Patents Search | 07/16/2024 | MO |
| Consider references cited in IDS | 07/17/2024 | MO |
| Assignee Search | 07/18/2024 | MO |
| Inventor Search | 07/18/2024 | MO |
| Updated PE2E Search | 03/05/2025 | MO |
| IP.com NPL Search | 03/05/2025 | MO |
| Assignee Search | 03/05/2025 | MO |
| Inventor Search | 03/05/2025 | MO |

/MICHAEL S OSINSKI/
Primary Examiner, Art Unit 2664

| Search Notes | Application/Control No. 17/609,911 | Applicant(s)/Patent Under Reexamination Lewis, Lucinda |
|---|---|---|
| | Examiner MICHAEL S OSINSKI | Art Unit 2674 |

| Interference Search | | | |
|---|---|---|---|
| **US Class/CPC Symbol** | **US Subclass/CPC Group** | **Date** | **Examiner** |
| All | See PE2E Search | 03/05/2025 | MO |

| | |
|---|---|
| /MICHAEL S OSINSKI/<br>Primary Examiner, Art Unit 2664 | |

| *Issue Classification* | Application/Control No. 17/609,911 | Applicant(s)/Patent Under Reexamination Lewis, Lucinda |
|---|---|---|
| | Examiner MICHAEL S OSINSKI | Art Unit 2674 |

| CPC | | | | | | |
|---|---|---|---|---|---|---|
| Symbol | | | | | Type | Version |
| G06N | / | 3 | / | 08 | F | 2013-01-01 |
| G06V | / | 20 | / | 20 | I | 2022-01-01 |
| G06V | / | 10 | / | 82 | I | 2022-01-01 |
| G06V | / | 10 | / | 764 | I | 2022-01-01 |
| G06V | / | 10 | / | 454 | I | 2022-01-01 |
| G06V | / | 20 | / | 56 | I | 2022-01-01 |

| CPC Combination Sets | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Symbol | | | | | Type | Set | Ranking | Version |
| | / | | / | | | | | |

| NONE (Assistant Examiner) (Date) | Total Claims Allowed: 20 | |
|---|---|---|
| /MICHAEL S OSINSKI/ Primary Examiner, Art Unit 2674 (Primary Examiner) 05 March 2025 (Date) | O.G. Print Claim(s) 1 | O.G. Print Figure 1, 5 |

| Issue Classification | Application/Control No. | Applicant(s)/Patent Under Reexamination |
|---|---|---|
| | 17/609,911 | Lewis, Lucinda |
| | **Examiner** | **Art Unit** |
| | MICHAEL S OSINSKI | 2674 |

| INTERNATIONAL CLASSIFICATION | | | | |
|---|---|---|---|---|
| **CLAIMED** | | | | |
| G06N3/08 | / | 3 | / | 08 |
| G06V20/20 | / | 20 | / | 20 |
| G06V10/82 | / | 10 | / | 82 |
| G06V10/764 | / | 10 | / | 764 |
| G06V10/44 | / | 10 | / | 44 |
| G06V20/56 | / | 20 | / | 56 |

| NON-CLAIMED | | | | |
|---|---|---|---|---|
| | / | | / | |

| US ORIGINAL CLASSIFICATION | |
|---|---|
| **CLASS** | **SUBCLASS** |
| | |

| CROSS REFERENCES(S) | | | | | | |
|---|---|---|---|---|---|---|
| **CLASS** | **SUBCLASS (ONE SUBCLASS PER BLOCK)** | | | | | |
| | | | | | | |

| | | |
|---|---|---|
| NONE<br><br>(Assistant Examiner) (Date) | **Total Claims Allowed:**<br>20 | |
| /MICHAEL S OSINSKI/ 05 March 2025<br>Primary Examiner, Art Unit 2674<br>(Primary Examiner) (Date) | O.G. Print Claim(s)<br>1 | O.G. Print Figure<br>1, 5 |

| *Issue Classification* | Application/Control No. 17/609,911 | Applicant(s)/Patent Under Reexamination Lewis, Lucinda |
|---|---|---|
| | Examiner MICHAEL S OSINSKI | Art Unit 2674 |

☑ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

CLAIMS

| Final | Original | Final | Original | Final | Original | Final | Original | Final | Original | Final | Original | Final | Original | Final | Original |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | | 10 | | 19 | | | | | | | | | | |
| | 2 | | 11 | | 20 | | | | | | | | | | |
| | 3 | | 12 | | | | | | | | | | | | |
| | 4 | | 13 | | | | | | | | | | | | |
| | 5 | | 14 | | | | | | | | | | | | |
| | 6 | | 15 | | | | | | | | | | | | |
| | 7 | | 16 | | | | | | | | | | | | |
| | 8 | | 17 | | | | | | | | | | | | |
| | 9 | | 18 | | | | | | | | | | | | |

| | | | |
|---|---|---|---|
| NONE<br>(Assistant Examiner) | (Date) | Total Claims Allowed:<br>20 | |
| /MICHAEL S OSINSKI/<br>Primary Examiner, Art Unit 2674<br>(Primary Examiner) | 05 March 2025<br>(Date) | O.G. Print Claim(s)<br>1 | O.G. Print Figure<br>1, 5 |

U.S. Patent and Trademark Office Part of Paper No.: 20250305

Discover IP-Analyze multimodal multiclass training CNN ne...

Visuals

1 - 50

1. Report of 2017 NSF Workshop on Multimedia Challenges, Opportunities and Research Roadmaps

With the transformative technologies and the rapidly changing global R&D landscape, the multimedia and multimodal community is now faced with many new opportuni
dissemination platform and pervasive computing resources, new research results are being...

190802308arXiv.org2019-08-064

2. **Towards Deep Learning Prospects: Insights for Social Media Analytics**

Deep learning (DL) has attracted increasing attention on account of its significant processing power in tasks, such as speech, image, or text processing. In order to the exp
digital social media (SM), analyzing these data using traditional tools and...

IEEE Xplore-IEEE Periodicals2019-01-014IEEE Access

3. Multimodal Sentiment Analysis: A GRU Approach for Improved Accuracy

In the real world, data often comes in different forms, such as text, audio, and video, making it multi-modal. This poses a challenge for machine learning models, which ne
data. One way to address this challenge is to use models that can handle...

IEEE XploreIEEE Conferences2024-08-2942024 International Conference on Electrical Electronics and Computing Technologies (ICEECT)

4. **Deep Learning in Information Security**

Machine learning has a long tradition of helping to solve complex information security problems that are difficult to solve manually. Machine learning techniques learn mo
data representations are hand-crafted by domain experts. Deep Learning is a...

180904332arXiv.org2018-09-124

5. **Semantics for Robotic Mapping, Perception and Interaction: A Survey**

For robots to navigate and interact more richly with the world around them, they will likely require a deeper understanding of the world in which they operate. In robotics
is often referred to as semantics, which dictates what does the world...

21010044arXiv.org2021-01-024Foundations and Trends in Robotics: Vol. 8, No. 1-2, pp 1-224 (2020)

6. **Video Summarization Using Deep Neural Networks: A Survey**

Video summarization technologies aim to create a concise and complete synopsis by selecting the most informative parts of the video content. Several approaches have b
current state of the art is represented by methods that rely on modern deep...

IEEE Xplore-IEEE Periodicals2021-10-294Proceedings of the IEEE

7. **A Survey on State-of-the-art Deep Learning Applications and Challenges**

Deep learning, a branch of artificial intelligence, is a data-driven method that uses multiple layers of interconnected units (neurons) to learn intricate patterns and represe
learning capability, it has become a powerful tool for solving...

240317561arXiv.org2024-03-264

Keyboard Shortcuts Result 1

# PE2E SEARCH - Search History (Prior Art)

| Ref # | Hits | Search Query | DBs | Default Operator | Plurals | British Equivalents | Time Stamp |
|---|---|---|---|---|---|---|---|
| L75 | 7703 | ((train training trained learn learning learnt learned propagate propagating propagation propagated backpropagate backpropagated backpropagation backpropagating tune tuning tuned optimize optimizing optimization optimized strengthen strengthening strengthened improve improving improvement improved maximize maximizing maximized bolster bolstering bolstered enhance enhancing enhanced enhancement boost boosting boosted revise revising revision revised upgrade upgrading upgraded update updating updated) WITH (CNN GAN FFNN DNN SNN ((convolution convolutional deep Siamese feed-forward (feed NEAR3 forward)) NEAR3 (network subnetwork sub-network neural model layer tree)) Convnet) WITH (plural two plural plurality set pair dual multi multiple numerous many multitude multi bunch many assorted assort assorting several two three third four five six seven eight nine ten plenty group collection bunch various variety vary varying varied) WITH (image media imagery still static picture photo photograph video frame movie JPEG MPEG stream live-view | (US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2025/03/05 01:20 AM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | through-image preview-image scene info information data story sketch article media song movie video show book website site story log entry record account listing register passage list listing item registration memo memorandum minute note planogram plano-gram plan-o-gram POG realogram realo-gram real-o-gram ROG diagram model image picture map list log plan portrayal chart graph list text evidence source publication article book literature story passage script journal page proof website site paragraph passage sentence verse clause editorial essay disclosure exposition document chapter report) WITH (profile type characteristic distribution classified classify classifying categorize categorizing categorized categorization describe described describing category class evaluate evaluating evaluation evaluated classification group grouping grouped category classifier type class grade league tier descriptor description brand branding branded taxonomy label labeled labeling category categorical hierarchy hierarchical ID identity identification identifying identified pointer context contextual metadata meta-data fringe tag tagging tagged indicator marker pointer header footer)) AND ((train training trained learn learning learnt leamed | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | propagate propagating propagation propagated backpropagate backpropagated backpropagation backpropagating tune tuning tuned optimize optimizing optimization optimized strengthen strengthening strengthened improve improving improvement improved maximize maximizing maximized bolster bolstering bolstered enhance enhancing enhanced enhancement boost boosting boosted revise revising revision revised upgrade upgrading upgraded update updating updated) WITH (RNN LSTM GRU LFTM ((recur recurring recurrent (long NEAR3 short)) NEAR3 (network subnetwork sub-network neural model layer tree)) Convnet) WITH (plural two plural plurality set pair dual multi multiple numerous many multitude multi bunch many assorted assort assorting several two three third four five six seven eight nine ten plenty group collection bunch various variety vary varying varied) WITH (image media imagery still static picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene info information data story sketch article media song movie video show book website site story log entry record account listing register passage list listing item registration memo | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | memorandum minute note planogram plano-gram plan-o-gram POG realogram realo-gram real-o-gram ROG diagram model image picture map list log plan portrayal chart graph list text evidence source publication article book literature story passage script journal page proof website site paragraph passage sentence verse clause editorial essay disclosure exposition document chapter report) WITH (profile type characteristic distribution classified classify classifying categorize categorizing categorized categorization describe described describing category class evaluate evaluating evaluation evaluated classification group grouping grouped category classifier type class grade league tier descriptor description brand branding branded taxonomy label labeled labeling category categorical hierarchy hierarchical ID identity identification identifying identified pointer context contextual metadata meta-data fringe tag tagging tagged indicator marker pointer header footer)) | | | | | |
| L76 | 602 | L75 AND ((CNN FFNN DNN SNN ((convolution convolutional deep Siamese feed-forward (feed NEAR3 forward)) NEAR3 (network subnetwork sub-network neural model layer tree)) Convnet) WITH (authentic authenticated authenticating | (US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2025/03/05 01:21 AM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | authentication copyright copyrighted copyrighting truth publish published publishing credit credited crediting citation cited citing accredited accrediting accredit validate validating validation validated confirm confirmation confirmed confirming ratify ratified ratifying prove proof proved proving affirm affirming affirmed certify certifying certified certification corroborate corroborating corroborated corroboration substantiate substantiation substantiating substantiated verify verifying verified verification attest attesting attested endorse endorsing endorsed endorsement vouch voucher vouching) WITH (image media imagery still static picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene info information data story sketch article media song movie video show book website site story log entry record account listing register passage list listing item registration memo memorandum minute note planogram plano-gram plan-o-gram POG realogram realo-gram real-o-gram ROG diagram model image picture map list log plan portrayal chart graph list text evidence source publication article book literature story passage | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | script journal page proof website site paragraph passage sentence verse clause editorial essay disclosure exposition document chapter report)) AND ((CNN FFNN DNN SNN ((convolution convolutional deep Siamese feed-forward (feed NEAR3 forward)) NEAR3 (network subnetwork sub-network neural model layer tree)) Convnet) WITH (determine determining determination determined decide deciding decision decided discover discovering discovery discovered conclude concluding conclusion concluded ascertain ascertaining ascertained ascertainment check checking checked detect detecting detection detected find finding found obtain obtaining obtainment obtained attain attaining attainment attained acquire acquiring acquired acquirement measure measuring measurement measured derive derivation derived deriving recognize recognizing recognized recognition sense sensing sensed ID identify identification identified identifying capture capturing captured photographing photographed imaging imaged obtain obtaining obtainment obtained receive receiving reception received search searching searched) WITH (object subject person article | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | body item shoe shirt clothing pants thing region interest ROI region-of-interest product car automobile vehicle cart dolly crane trolley truck SUV van transport train airplane plane aircraft ambulance autocycle auto-cycle bike boat bus cab camper caravan carriage jeep kart limo limousine locomotive motorcycle robot drone UAV railcar tank tractor transportation motorbike home house article object entity structure) WITH (image media imagery still static picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene info information data story sketch article media song movie video show book website site story log entry record account listing register passage list listing item registration memo memorandum minute note planogram plano-gram plan-o-gram POG realogram realo-gram real-o-gram ROG diagram model image picture map list log plan portrayal chart graph list text evidence source publication article book literature story passage script journal page proof website site paragraph passage sentence verse clause editorial essay disclosure exposition document chapter report)) | | | | | |
| L77 | 2 | ("20170061250" \| "20180204120").pn. OR ("10423874").urpn. | (US-PGPUB; USPAT; USOCR) | OR | ON | ON | 2025/03/05 02:01 PM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | AND (PGPB \| USPT \| USOC).dbnm. | | | | | |
| L78 | 24 | "20170098153" | (US-PGPUB; USPAT; USOCR; FIT (AU, AP, AT, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, OA, RU, SU, WO, MC, MD, MY, NL, NO, NZ, PH, PL, PT, RO, RS, SE, SG, SI, SK, TH, TN, TR, TW, UA, VN); FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2025/03/05 02:03 PM |
| L81 | 17789 | (G06N3/08 G06V10/454 G06V10/764 G06V10/82 G06V20/20 G06V20/56 G06N3/042 G06N3/044 G06N3/048 G06F16/433 G06F16/434 G06F16/438 G06F16/90332 G06F18/24143 G06N3/045 G06Q30/018 G06Q30/0241 G06N5/02 G06V2201/08 H04L63/02).cpc. AND ((CNN FFNN DNN SNN ((convolution convolutional deep Siamese feed-forward (feed NEAR3 forward)) NEAR3 (network subnetwork sub-network neural model layer tree)) Convnet) WITH (authentic authenticated authenticating authentication copyright copyrighted copyrighting truth publish published publishing credit credited crediting citation cited citing accredited accrediting accredit validate validating validation | (US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2025/03/05 08:27 PM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | validated confirm confirmation confirmed confirming ratify ratified ratifying prove proof proved proving affirm affirming affirmed certify certifying certified certification corroborate corroborating corroborated corroboration substantiate substantiation substantiating substantiated verify verifying verified verification attest attesting attested endorse endorsing endorsed endorsement vouch voucher vouching) WITH (image media imagery still static picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene info information data story sketch article media song movie video show book website site story log entry record account listing register passage list listing item registration memo memorandum minute note planogram plano-gram plan-o-gram POG realogram realo-gram real-o-gram ROG diagram model image picture map list log plan portrayal chart graph list text evidence source publication article book literature story passage script journal page proof website site paragraph passage sentence verse clause editorial essay disclosure exposition document chapter report)) | | | | | |
| L82 | 706 | L81 AND ((train training trained learn learning | (US-PGPUB; USPAT; USOCR; FPRS; EPO; | OR | ON | ON | 2025/03/05 08:27 PM |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | learnt learned propagate propagating propagation propagated backpropagate backpropagated backpropagation backpropagating tune tuning tuned optimize optimizing optimization optimized strengthen strengthening strengthened improve improving improvement improved maximize maximizing maximized bolster bolstering bolstered enhance enhancing enhanced enhancement boost boosting boosted revise revising revision revised upgrade upgrading upgraded update updating updated) WITH (CNN GAN FFNN DNN SNN ((convolution convolutional deep Siamese feed-forward (feed NEAR3 forward)) NEAR3 (network subnetwork sub-network neural model layer tree)) Convnet) WITH (plural two plural plurality set pair dual multi multiple numerous many multitude multi bunch many assorted assort assorting several two three third four five six seven eight nine ten plenty group collection bunch various variety vary varying varied) WITH (image media imagery still static picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene info information data story sketch article media song movie video show book website site story log entry record account listing register passage | JPO; DERWENT; IBM_TDB) | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | list listing item registration memo memorandum minute note planogram plano-gram plan-o-gram POG realogram realo-gram real-o-gram ROG diagram model image picture map list log plan portrayal chart graph list text evidence source publication article book literature story passage script journal page proof website site paragraph passage sentence verse clause editorial essay disclosure exposition document chapter report) WITH (profile type characteristic distribution classified classify classifying categorize categorizing categorized categorization describe described describing category class evaluate evaluating evaluation evaluated classification group grouping grouped category classifier type class grade league tier descriptor description brand branding branded taxonomy label labeled labeling category categorical hierarchy hierarchical ID identity identification identifying identified pointer context contextual metadata meta-data fringe tag tagging tagged indicator marker pointer header footer)) AND ((train training trained learn learning learnt learned propagate propagating propagation propagated backpropagate backpropagated backpropagation backpropagating tune tuning tuned optimize optimizing optimization | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | optimized strengthen strengthening strengthened improve improving improvement improved maximize maximizing maximized bolster bolstering bolstered enhance enhancing enhanced enhancement boost boosting boosted revise revising revision revised upgrade upgrading upgraded update updating updated) WITH (RNN LSTM GRU LFTM ((recur recurring recurrent (long NEAR3 short)) NEAR3 (network subnetwork sub-network neural model layer tree)) Convnet) WITH (plural two plural plurality set pair dual multi multiple numerous many multitude multi bunch many assorted assort assorting several two three third four five six seven eight nine ten plenty group collection bunch various variety vary varying varied) WITH (image media imagery still static picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene info information data story sketch article media song movie video show book website site story log entry record account listing register passage list listing item registration memo memorandum minute note planogram plano-gram plan-o-gram POG realogram realo-gram real-o-gram ROG diagram model image picture map list log plan portrayal chart graph list | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | text evidence source publication article book literature story passage script journal page proof website site paragraph passage sentence verse clause editorial essay disclosure exposition document chapter report) WITH (profile type characteristic distribution classified classify classifying categorize categorizing categorized categorization describe described describing category class evaluate evaluating evaluation evaluated classification group grouping grouped category classifier type class grade league tier descriptor description brand branding branded taxonomy label labeled labeling category categorical hierarchy hierarchical ID identity identification identifying identified pointer context contextual metadata meta-data fringe tag tagging tagged indicator marker pointer header footer)) | | | | | |
| L83 | 2 | ((("LEWIS") near3 ("Lucinda"))).INV. | (US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT) | OR | ON | ON | 2025/03/05 09:54 PM |
| L84 | 2 | ((("AUTOMOBILIA") near3 ("II") near3 ("LLC"))).AS,AANM. | (US-PGPUB; USPAT) | OR | ON | ON | 2025/03/05 09:55 PM |

## PE2E SEARCH - Search History (Interference)

| Ref # | Hits | Search Query | DBs | Default Operator | Plurals | British Equivalents | Time Stamp |
|---|---|---|---|---|---|---|---|
| N1 | 299 | (((train training trained learn learning learnt learned propagate propagating propagation propagated backpropagate backpropagated backpropagation | (US-PGPUB; USPAT) | OR | ON | ON | 2025/03/05 09:56 PM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | backpropagating tune tuning tuned optimize optimizing optimization optimized strengthen strengthening strengthened improve improving improvement improved maximize maximizing maximized bolster bolstering bolstered enhance enhancing enhanced enhancement boost boosting boosted revise revising revision revised upgrade upgrading upgraded update updating updated) WITH (CNN GAN FFNN DNN SNN ((convolution convolutional deep Siamese feed-forward (feed NEAR3 forward)) NEAR3 (network subnetwork sub-network neural model layer tree)) Convnet) WITH (plural two plural plurality set pair dual multi multiple numerous many multitude multi bunch many assorted assort assorting several two three third four five six seven eight nine ten plenty group collection bunch various variety vary varying varied) WITH (image media imagery still static picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene info information data story sketch article media song movie video show book website site story log entry record account listing register passage list listing item registration memo memorandum minute note planogram plano-gram plan-o-gram POG realogram realo-gram | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | real-o-gram ROG diagram model image picture map list log plan portrayal chart graph list text evidence source publication article book literature story passage script journal page proof website site paragraph passage sentence verse clause editorial essay disclosure exposition document chapter report) WITH (profile type characteristic distribution classified classify classifying categorize categorizing categorized categorization describe described describing category class evaluate evaluating evaluation evaluated classification group grouping grouped category classifier type class grade league tier descriptor description brand branding branded taxonomy label labeled labeling category categorical hierarchy hierarchical ID identity identification identifying identified pointer context contextual metadata meta-data fringe tag tagging tagged indicator marker pointer header footer)) AND ((train training trained learn learning learnt learned propagate propagating propagation propagated backpropagate backpropagated backpropagation backpropagating tune tuning tuned optimize optimizing optimization optimized strengthen strengthening strengthened improve improving improvement improved maximize maximizing maximized | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | bolster bolstering bolstered enhance enhancing enhanced enhancement boost boosting boosted revise revising revision revised upgrade upgrading upgraded update updating updated) WITH (RNN LSTM GRU LFTM ((recur recurring recurrent (long NEAR3 short)) NEAR3 (network subnetwork sub-network neural model layer tree)) Convnet) WITH (plural two plural plurality set pair dual multi multiple numerous many multitude multi bunch many assorted assort assorting several two three third four five six seven eight nine ten plenty group collection bunch various variety vary varying varied) WITH (image media imagery still static picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene info information data story sketch article media song movie video show book website site story log entry record account listing register passage list listing item registration memo memorandum minute note planogram plano-gram plan-o-gram POG realogram realo-gram real-o-gram ROG diagram model image picture map list log plan portrayal chart graph list text evidence source publication article book literature story passage script journal page proof website site paragraph passage sentence | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | verse clause editorial essay disclosure exposition document chapter report) WITH (profile type characteristic distribution classified classify classifying categorize categorizing categorized categorization describe described describing category class evaluate evaluating evaluation evaluated classification group grouping grouped category classifier type class grade league tier descriptor description brand branding branded taxonomy label labeled labeling category categorical hierarchy hierarchical ID identity identification identifying identified pointer context contextual metadata meta-data fringe tag tagging tagged indicator marker pointer header footer))).clm. | | | | | |
| N2 | 11 | N1 AND (((CNN FFNN DNN SNN ((convolution convolutional deep Siamese feed-forward (feed NEAR3 forward)) NEAR3 (network subnetwork sub-network neural model layer tree)) Convnet) WITH (authentic authenticated authenticating authentication copyright copyrighted copyrighting truth publish published publishing credit credited crediting citation cited citing accredited accrediting accredit validate validating validation validated confirm confirmation confirmed confirming ratify ratified | (US-PGPUB; USPAT) | OR | ON | ON | 2025/03/05 09:56 PM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | ratifying prove proof proved proving affirm affirming affirmed certify certifying certified certification corroborate corroborating corroborated corroboration substantiate substantiation substantiating substantiated verify verifying verified verification attest attesting attested endorse endorsing endorsed endorsement vouch voucher vouching) WITH (image media imagery still static picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene info information data story sketch article media song movie video show book website site story log entry record account listing register passage list listing item registration memo memorandum minute note planogram plano-gram plan-o-gram POG realogram realo-gram real-o-gram ROG diagram model image picture map list log plan portrayal chart graph list text evidence source publication article book literature story passage script journal page proof website site paragraph passage sentence verse clause editorial essay disclosure exposition document chapter report)) AND ((CNN FFNN DNN SNN ((convolution convolutional deep Siamese feed-forward (feed NEAR3 forward)) NEAR3 (network | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | subnetwork sub-network neural model layer tree)) Convnet) WITH (determine determining determination determined decide deciding decision decided discover discovering discovery discovered conclude concluding conclusion concluded ascertain ascertaining ascertained ascertainment check checking checked detect detecting detection detected find finding found obtain obtaining obtainment obtained attain attaining attainment attained acquire acquiring acquired acquirement measure measuring measurement measured derive derivation derived deriving recognize recognizing recognized recognition sense sensing sensed ID identify identification identified identifying capture capturing captured photographing photographed imaging imaged obtain obtaining obtainment obtained receive receiving reception received search searching searched) WITH (object subject person article body item shoe shirt clothing pants thing region interest ROI region-of-interest product car automobile vehicle cart dolly crane trolley truck SUV van transport train airplane plane aircraft ambulance autocycle auto-cycle bike boat bus cab camper caravan carriage jeep | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | kart limo limousine locomotive motorcycle robot drone UAV railcar tank tractor transportation motorbike home house article object entity structure) WITH (image media imagery still static picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene info information data story sketch article media song movie video show book website site story log entry record account listing register passage list listing item registration memo memorandum minute note planogram plano-gram plan-o-gram POG realogram realo-gram real-o-gram ROG diagram model image picture map list log plan portrayal chart graph list text evidence source publication article book literature story passage script journal page proof website site paragraph passage sentence verse clause editorial essay disclosure exposition document chapter report))).clm. | | | | | |

UNITED STATES
PATENT AND TRADEMARK OFFICE

P.O. Box 1450
Alexandria, VA 22313 - 1450
www.uspto.gov

# ELECTRONIC PAYMENT RECEIPT

| APPLICATION # | RECEIPT DATE / TIME | ATTORNEY DOCKET # |
|---|---|---|
| **17/609,911** | **01/23/2025 08:14:56 PM Z ET** | **P1409US01** |

## Title of Invention

METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

## Application Information

| | | | |
|---|---|---|---|
| APPLICATION TYPE | Utility - U.S. National Stage under 35 USC 371 | PATENT # | - |
| CONFIRMATION # | 1651 | FILED BY | Nicholas Blanton |
| PATENT CENTER # | 68889065 | AUTHORIZED BY | - |
| CUSTOMER # | 36601 | FILING DATE | 11/09/2021 |
| INTL. APPLICATION # | - | INTL. FILING DATE | - |
| CORRESPONDENCE ADDRESS | - | FIRST NAMED INVENTOR | Lucinda Lewis |

## Payment Information

| PAYMENT METHOD | PAYMENT TRANSACTION ID | PAYMENT AUTHORIZED BY |
|---|---|---|
| **CARD / 4007** | **E20251MK15138402** | **Nicholas Blanton** |

| FEE CODE | DESCRIPTION | ITEM PRICE($) | QUANTITY | ITEM TOTAL($) |
|---|---|---|---|---|
| 3253 | EXTENSION FOR RESPONSE WITHIN THIRD MONTH, EXCEPT PROVISIONAL APPLICATIONS | 318.00 | 1 | 318.00 |
| | | | **TOTAL AMOUNT:** | **$318.00** |

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

**New Applications Under 35 U.S.C. 111**

If a new application is being filed and the application includes the necessary components for filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement

Receipt will establish the filing date of the application

**National Stage of an International Application under 35 U.S.C. 371**

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

**New International Application Filed with the USPTO as a Receiving Office**

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

Doc Code: MES.GIB
Document Description: Certification of Micro Entity Status (Gross Income Basis)

PTO/SB/15A (07-14)

# CERTIFICATION OF MICRO ENTITY STATUS
## (GROSS INCOME BASIS)

| Application Number or Control Number (if applicable): | Patent Number (if applicable): |
|---|---|
| 17/609,911 | N/A |
| First Named Inventor: | Title of Invention: |
| Lucinda LEWIS | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY |

The applicant hereby certifies the following—

(1) **SMALL ENTITY REQUIREMENT –** The applicant qualifies as a small entity as defined in 37 CFR 1.27.

(2) **APPLICATION FILING LIMIT –** Neither the applicant nor the inventor nor a joint inventor has been named as the inventor or a joint inventor on more than four previously filed U.S. patent applications, excluding provisional applications and international applications under the Patent Cooperation Treaty (PCT) for which the basic national fee under 37 CFR 1.492(a) was not paid, and also excluding patent applications for which the applicant has assigned all ownership rights, or is obligated to assign all ownership rights, as a result of the applicant's previous employment.

(3) **GROSS INCOME LIMIT ON APPLICANTS AND INVENTORS –** Neither the applicant nor the inventor nor a joint inventor, in the calendar year preceding the calendar year in which the applicable fee is being paid, had a gross income, as defined in section 61(a) of the Internal Revenue Code of 1986 (26 U.S.C. 61(a)), exceeding the "Maximum Qualifying Gross Income" reported on the USPTO Web site at http://www.uspto.gov/patents/law/micro_entity.jsp which is equal to three times the median household income for that preceding calendar year, as most recently reported by the Bureau of the Census.

(4) **GROSS INCOME LIMIT ON PARTIES WITH AN "OWNERSHIP INTEREST" –** Neither the applicant nor the inventor nor a joint inventor has assigned, granted, or conveyed, nor is under an obligation by contract or law to assign, grant, or convey, a license or other ownership interest in the application concerned to an entity that, in the calendar year preceding the calendar year in which the applicable fee is being paid, had a gross income, as defined in section 61(a) of the Internal Revenue Code of 1986, exceeding the "Maximum Qualifying Gross Income" reported on the USPTO Web site at http://www.uspto.gov/patents/law/micro_entity.jsp which is equal to three times the median household income for that preceding calendar year, as most recently reported by the Bureau of the Census.

**SIGNATURE by an authorized party set forth in 37 CFR 1.33(b)**

| | | | | | |
|---|---|---|---|---|---|
| Signature | /Nicholas Blanton/ | | | | |
| Name | Nicholas E. Blanton | | | | |
| Date | 01/23/2025 | Telephone | 860.266.1779 | Registration No. | 78,975 |

☐ There is more than one inventor and I am one of the inventors who are jointly identified as the applicant. The required additional certification form(s) signed by the other joint inventor(s) are included with this form.

PTO/AIA/22 (01-25)
Approved for use through 11/30/2027. OMB 0651-0031
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

# PETITION FOR EXTENSION OF TIME UNDER 37 CFR 1.136(a) (NOT for Provisional Applications)

Docket Number (Optional): P1409US01

Application Number: 17/609,911

Filed: 11/09/2021

For: METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

Art Unit: 2674

Examiner: MICHAEL S OSINSKI

This is a request under the provisions of 37 CFR 1.136(a) to extend the period for filing a reply in the above-identified application.

The requested extension and fee are as follows (check time period desired and enter the appropriate fee below):

| | | Fee | Small Entity Fee | Micro Entity Fee | |
|---|---|---|---|---|---|
| ☐ | One month (37 CFR 1.17(a)(1)) | \$235 | \$94 | \$47 | \$ |
| ☐ | Two months (37 CFR 1.17(a)(2)) | \$690 | \$276 | \$138 | \$ |
| ☑ | Three months (37 CFR 1.17(a)(3)) | \$1,590 | \$636 | \$318 | \$ 318 |
| ☐ | Four months (37 CFR 1.17(a)(4)) | \$2,495 | \$998 | \$499 | \$ |
| ☐ | Five months (37 CFR 1.17(a)(5)) | \$3,395 | \$1,358 | \$679 | \$ |

- ☑ Applicant asserts small entity status. See 37 CFR 1.27.
- ☐ Applicant certifies micro entity status. See 37 CFR 1.29.
  Form PTO/SB/15A or B or equivalent must either be enclosed or have been submitted previously.
- ☐ A check in the amount of the fee is enclosed.
- ☐ Payment by credit card. Form PTO-2038 is attached.
- ☐ The Director has already been authorized to charge fees in this application to a Deposit Account.
- ☐ The Director is hereby authorized to charge any fees which may be required, or credit any overpayment, to Deposit Account Number ____________.
- ☑ Payment made via USPTO patent electronic filing system.

**WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.**

I am the

- ☐ applicant.
- ☑ attorney or agent of record. Registration number 78,975.
- ☐ attorney or agent acting under 37 CFR 1.34. Registration number ____________.

/Nicholas Blanton/
Signature

January 23, 2025
Date

Nicholas E. Blanton, Esq.
Typed or printed name

860.266.1779
Telephone Number

**NOTE:** This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4 for signature requirements and certifications. Submit multiple forms if more than one signature is required, see below*.

☐ * Total of ____________ forms are submitted.

UNITED STATES
PATENT AND TRADEMARK OFFICE

P.O. Box 1450
Alexandria, VA 22313 - 1450
www.uspto.gov

# ELECTRONIC ACKNOWLEDGEMENT RECEIPT

| APPLICATION # | RECEIPT DATE / TIME | ATTORNEY DOCKET # |
|---|---|---|
| **17/609,911** | **01/23/2025 08:14:56 PM Z ET** | **P1409US01** |

## Title of Invention

METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

## Application Information

| | | | |
|---|---|---|---|
| APPLICATION TYPE | Utility - U.S. National Stage under 35 USC 371 | PATENT # | - |
| CONFIRMATION # | 1651 | FILED BY | Nicholas Blanton |
| PATENT CENTER # | 68889065 | FILING DATE | 11/09/2021 |
| CUSTOMER # | 36601 | FIRST NAMED INVENTOR | Lucinda Lewis |
| INTL. APPLICATION # | - | INTL. FILING DATE | - |
| CORRESPONDENCE ADDRESS | - | AUTHORIZED BY | - |

## Documents

# TOTAL DOCUMENTS: 3

| DOCUMENT | PAGES | DESCRIPTION | SIZE (KB) |
|---|---|---|---|
| aia0022_P1409US01.pdf | 1 | Extension of Time | 142 KB |
| P1409US01_Non-final_Response_FF.pdf | 17 | Applicant Arguments/Remarks Made in an Amendment | 1569 KB |
| aia0015A_P1409US01.pdf | 1 | Certification of Micro Entity (Gross Income Basis) | 86 KB |

## Digest

| DOCUMENT | MESSAGE DIGEST(SHA-512) |
|---|---|

| | |
|---|---|
| aia0022_P1409US01.pdf | 00E258A3BE0C053F214A47D92CD14E227E0F45D4A8CB2DEC C796F9D246012DB3A041128B72844BC4EBF3F913DDDBA8EC 28BF4B72A07E16798F81A322A1215E87 |
| P1409US01_Non-final_Response_FF.pdf | 97506C266600478D3ECC8501EACA2CABBA5F9FBCDE6C2382 3E9702A746E92327A9F395FD73AE6437B0FEB0C42BBA45ECB 52A4D87EB108357DE40241C6E0A54E7 |
| aia0015A_P1409US01.pdf | F698D5C5F012956DE232806552A4A3C786A027E8C5FCA8E6A B479958A9B1DADCAE92D673A6785CB2EFD0B8DA8EB14C42 C93957FDB2DA20FC729159D9C101E48C |

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

**New Applications Under 35 U.S.C. 111**

If a new application is being filed and the application includes the necessary components for filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application

**National Stage of an International Application under 35 U.S.C. 371**

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

**New International Application Filed with the USPTO as a Receiving Office**

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | |
|---|---|
| Inventors: Lucinda LEWIS | Attorney Docket No: P1409US01 |
| Application No: 17/609,911 | Confirmation No: 1651 |
| Filed: November 9, 2021 | Art Unit: 2664 |
| Title: METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | Examiner: Osinski, Michael S. |

January 23, 2025

Commissioner for Patents
P.O. Box 1450
Alexandria VA 22313-1450

**RESPONSE TO NON-FINAL OFFICE ACTION**

Dear Sir:

In response to the non-final Office Action mailed on July 23, 2024, please amend the above-identified application, without prejudice, as follows:

**Amendments to the Specification** – beginning on page 2.

**Amendments to the Claims** – beginning on page 3.

**Remarks/Arguments** – beginning on page 8.

***Appendix I*** – begins after the Remarks.

## AMENDMENTS TO THE SPECIFICATION

Applicant respectfully submits that no new matter has been added as a result of the following amendments. Please add the follow subsection heading and paragraph, after the Title of the originally-filed application, as follows:

### CROSS REFERENCE TO RELATED APPLICATIONS

**[0001]** This patent application claims priority to, and the benefit of, PCT/US2020/032149 filed on May 8, 2020, entitled "Methods, Systems And Computer Program Products for Media Processing And Display," which in turn claims priority to, and the benefit of, U.S. Provisional Pat. App. No. 62/845,546 filed on May 9, 2019, entitled "Neural Network for Identifying Vehicles and/or Providing Artificial Intelligence"; each prior application is incorporated by reference herein in its entirety.

## AMENDMENTS TO THE CLAIMS

Applicant respectfully submits the listing of claims below that will replace all prior versions, and listings, of claims in the application. Applicant submits that no new matter has been added as a result of the following amendments.

### *Listing of Claims*

1. (Currently amended) A method adapted for processing multimodal input data based upon multimodal and/or multiclass training data, the method comprising:

training, by a processor, a convolutional neural network (CNN) using the multimodal and/or multiclass training data, which includes authenticated data and a taxonomy, to classify at least one data object in a media asset, authenticate the media asset received by the CNN, and verify the data object received by the CNN;

training, by a processor, a recurrent neural network (RNN) using the multimodal and/or multiclass training data;

receiving, by the processor ~~a processing device~~, a query comprising input data;

classifying, by the trained CNN, the input data with respect to the authenticated data and elements of one or more classes of the taxonomy;

generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of one or more classes of the taxonomy comprising a closest match to the input data; and

displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

2. (Previously presented) The method of claim 1, wherein the authenticated data comprises copyright registered works of authorship, metadata and text.

3. (Currently amended) The method of claim 2, wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations and/or writings.

4. (Previously presented) The method of claim 3, wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information.

5. (Previously presented) The method of claim 1, wherein the authenticated data comprises data from a copyright registered database.

6. (Currently amended) The method of claim 1, wherein the one or more classes ~~elements~~ of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture.

7. (Previously presented) The method of claim 1, wherein the input data comprises one or more of image data, video data, intake data or geographical location data.

8. (Previously presented) The method of claim 1, wherein classifying comprises mapping input data to authenticated data using the taxonomy.

9. (Previously presented) The method of claim 1, wherein the result comprises one or more of an image, a video, text, or sound.

10. (Previously presented) The method of claim 1, wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information.

11. (Previously presented) The method of claim 1, wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy.

12. (Previously presented) The method of claim 11, wherein the probability is determined by a cross-entropy function.

13. (Previously presented) The method of claim 1, wherein the result comprises augmented reality content, wherein displaying the result comprises:

displaying the result in an augmented reality apparatus, comprising:

passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides;

imaging the light at the display device; and

displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data.

14. (Previously presented) The method of claim 13, wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time.

15. (Currently amended) The method of claim 1, further comprising:

a deep learning module including neural memory, and/or

a transformer, alone or in combination with the RNN to process attention models ~~further comprising training a recurrent neural network (RNN) using authenticated data and taxonomy~~.

16. (Currently amended) The method of claim 1 [[15]], wherein the input data comprises unstructured data, the method further comprising:

processing, by the ~~trained~~ RNN, the unstructured data to yield structured data; and

classifying, by the ~~trained~~ CNN, the structured data.

17. (Previously presented) The method of claim 1, wherein the input data comprises user uploaded data, the method further comprising authenticating the user uploaded data using Siamese Neural Networks and adding the authenticated user uploaded data to the authenticated data.

18. (Currently amended) A system adapted for processing multimodal input data based upon multimodal and/or multiclass training data, the system comprising:

a memory;

a processor, coupled to the memory, the processor configured to:

train a convolutional neural network (CNN) using the multimodal and/or multiclass training data including authenticated data and a taxonomy to classify at least one data object in a media asset, authenticate the media asset received by the CNN, and verify the data object received by the CNN;

train a recurrent neural network (RNN) using the multimodal and/or multiclass training data;

receive, by the processor ~~a processing device~~, a query comprising input data;

classify, by the trained CNN, the input data with respect to the authenticated data and elements of one or more classes of the taxonomy;

generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of one or more classes of the taxonomy comprising a closest match to the input data; and

display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

19. (Previously presented) The system of claim 18, wherein the authenticated data comprises copyright registered works of authorship, metadata and text.

20. (Currently amended) A computer-readable non-transitory storage medium comprising executable instructions that, when executed by a computing device, cause the computing device to perform operations, adapted for processing multimodal input data based upon multimodal and/or multiclass training data, comprising:

training a convolutional neural network (CNN) using multimodal and/or multiclass training data, which includes authenticated data and a taxonomy, to classify at least one data object in a media asset, authenticate the media asset received by the CNN, and verify the data object received by the CNN;

training a recurrent neural network (RNN) using the multimodal and/or multiclass training data;

receiving~~, by a processing device,~~ a query comprising input data;

classifying, by the trained CNN and/or the RNN, the input data with respect to the authenticated data and elements of one or more classes of the taxonomy;

generating a result, by the trained CNN and/or the RNN, wherein the result comprises authenticated data and elements of one or more classes of the taxonomy comprising a closest match to the input data; and

displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

**REMARKS/ARGUMENTS**

Applicant responds to the non-final Office Action dated July 23, 2024. Please reconsider the application in view of the following remarks.

***Status and Disposition of Claims***

Applicant amends independent claims 1, 18, and 20 to recite "processing multimodal input data based upon multimodal and/or multiclass training data" to clarify certain elements and limitations, and to overcome the art of record. Support for this amendment may be found, for example, in paragraph [0056] of Applicant's originally-filed PCT specification. See also paragraphs [0019]-[0021]. To further support this amendment, Applicant amends dependent claim 6 (claim 6 being dependent from claim 1) to clarify that the "one or more classes" are "of the taxonomy". In this context, "multiclass" refers to, for example, the different categories of data as set forth in claim 6.

Additionally with respect to independent claims 1, 18, and 20, Applicant clarifies the elements "a plurality of media assets", "data object", and "object" to refer to, for example, the exemplary media asset presented in ***Appendix I*** herein. Per ***Appendix I***, an exemplary "media asset" includes numerous data objects located therein, and also including captions and text (including copyright-protected books and/or articles) suitable for natural language processing. An exemplary "data object" may include a specific vehicle and its accompanying textual definitions and/or descriptions, shown within the media asset. And an exemplary "object" may include a specific vehicle shown with the data object.

Dependent claims 15 and 16 have also been amended for clarification and consistency with claim 1. Support for the amendment to claim 15 may be found, for example, in paragraphs [0017], [0041] and [0084]. Concerning other amendments to the independent claims, an "RNN" is consistent with processing multimodal input data based upon multimodal and/or multiclass training data. See e.g., paragraph [0057].

Dependent claims 3 and 6 have been amended for clarification purposes. In particular, claim 6 has been amended to recite "one or more classes" to comport with claim 1 recitation of "multiclass" "Claims 2, 5, 8-9, and 12-20 are presented as originally filed. Therefore, claims 1-20 remain pending in this case.

Applicant requests the Office review and reconsider the application in view of the claim amendments, priority claim, and the following remarks; Applicant respectfully solicits a notice of allowance thereof.

### ***Cross-Reference to Related Applications***

Applicant submits that the amendment to Applicant's specification at paragraph [0001] (submitted with this paper) reaffirms the claim(s) to priority of the as-filed Application Disclosure Statement (ADS) filed on November 9th, 2021. Since the application data sheet is considered part of the application, the specific reference to an earlier filed provisional or nonprovisional application in the application data sheet satisfies the "specific reference" requirement of 35 U.S.C.119(e)(1) or 120, and it also complies with 37 CFR 1.78(a)(3) or (d)(3). Moreover, 37 CFR 1.76(b)(5) codifies this as the official method: "Providing this information in the application data sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and § 1.78." MPEP 601.05(a). Applicant submits that each and every element and limitation in at least the independent claims herewith are fully supported by Applicant's provisional application filed on May 9, 2019.

### ***Prior Art Under AIA 102(b)(1) Identification of Effective Filing Date***

The effective filing date of the Office's prior art reference to Shebl is May 9, 2019 (hereinafter, "*Shebl*" U.S. Pat. Appl. No. 2019/0278994, Appl. No. 17/609,946). A screenshot, provided below, of the file history's Continuity page shows that date:

**Continuity**

Parent data

| Application # | Description | Parent application # | Filing or 371 (c) date | Status | Patent # | AIA |
|---|---|---|---|---|---|---|
| [illegible] | is a National Stage Entry of | [illegible] | 05/09/2019 | Published | -- | - |

Child data

No child data available for selected application

The effective filing date of Applicant's instant application is May 9, 2019, and therefore prior art is deemed from May 8, 2018 under post-AIA 102(b)(1).

**17/609,911 | P1409US01: METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY**

| Application # | Confirmation # | Attorney Docket # | Patent # | Filing or 371 (c) date | Status |
|---|---|---|---|---|---|
| 17/609,911 | 1951 | P1409US01 Edit | - | 11/09/2021 | Non Final Action Mailed 07/23/2024 |

Application Data
Documents & Transactions
Continuity
Patent Term Adjustment
Foreign priority
Fee payment history
Address & Attorney/Agent Information
Supplemental Content
Assignments
Display References

**Continuity**

Parent data

| Application # | Description | Parent application # | Filing or 371 (c) date | Status | Patent # | AIA |
|---|---|---|---|---|---|---|
| 17/609,911 | is a National Stage Entry of | [illegible] | 05/08/2020 | Published | -- | - |
| [illegible] | Claims priority from a provisional application | [illegible] | 05/09/2019 | Expired | -- | - |

Child data

No child data available for selected application

Therefore, Applicant's instant application precedes SHEBL, and therefore the reference cannot constitute prior art as applied against the claims of Applicant's instant application. Applicant further asserts this fact with respect to any other prior art of record.

***Claim Rejections under 35 U.S.C. § 101***

Claims 1-20 stand rejected under 35 U.S.C. § 101 as allegedly being directed to an abstract idea of mental processes which allegedly are concepts performed in the human mind. Applicant respectfully traverses.

Without acquiescing to the merits of the Office's allegation, Applicant amends each of the independent claims 1, 18, and 20 to recite, *inter alia*, "*training, by a processor, a convolutional neural network (CNN) using the multimodal and/or multiclass training data, which includes authenticated data and a taxonomy, to classify at least one data object in a media asset, authenticate the media asset received by the CNN, and verify the data object received by the CNN*". Support for this amendment may be found, for example, in FIG. 3 and/or paragraph [0039] of Applicant's originally-filed specification. In this context, the intended relationship between a "media asset", a "data object", and an "object" are conveyed in *Appendix I* submitted below. An object may be, for example, a vehicle of a particular make, model, and year. A data object may be, for example, one or more written descriptions of the object. And a media asset may refer to the overall image, video, or the like, wherein the image may convey a plurality of data objects.

The claimed functions of training to "classify," "authenticate the media asset received by the CNN", and "verify the data object received by the CNN" do not constitute an abstract idea, because in line with Applicant's intended claimed invention, these functions constitute a technical improvement above and beyond that which customary computing systems are capable of achieving as measured from the time of filing of Applicant's instant application. Applicant asserts that claims 1, 18, and 20, as amended, result in significant *technical improvement(s)* with respect to training, processing, authenticating, verifying, and/or generating a result above that of a human mental process, and therefore constitutes patentable-eligible subject matter, consistent with MPEP §§ 2106.04-2106.06 and Federal Circuit post-*Enfish* precedent. See also: *Enfish, LLC v. Microsoft Corp.*, 822 F.3d 1327 (Fed. Cir. 2016).

Applicant respectfully requests withdrawal of the rejection under 35 U.S.C. § 101 and solicits a notice of allowance thereof.

***Claim Rejections under 35 U.S.C. § 102 (OA, paragraphs 4-17)***

Claims 1-11 and 18-20 stand rejected under 35 U.S.C. § 102 as allegedly being anticipated by U.S. Pat. Pub. No. 2019/0278994 A1 to Bumpas (hereinafter "*Bumpas*"). Applicant respectfully traverses.

Regarding a first concept, claim 1 (and similarly, independent claims 18 and 20 of Applicant's claim set) requires that the method be *adapted for processing multimodal input data based upon multimodal and/or multiclass training data*. This concept is supported throughout Applicant's specification, for example, FIGS 1, 3, 5-7, and/or paragraphs [0015]-[0016], [0026], [0039, [0056]-[0057], [0071], [0073], and [0083]-[0084]. One technical improvement of Applicant's claims invention is that it results in more accurate identification and matching of various input data types, i.e., voice, text, video, and/or image; and also, those input data types for the corresponding training data set, upon which a match is rendered. The invention also reduces the rate of false narratives and/or results, because it is based upon authenticated and verified data.

In contrast, *Bumpas* only discloses image matching characteristics, such as may be found in the Abstract: "a system for image-based vehicle identification includes a database, an image processor, and a vehicle search engine". *Bumpas* contemplates a solution of a problem where "[o]ften a user may not be privy to the make or model of the car they are looking for, making convention search engines frustrating and/or impossible to use." *Bumpas*, paragraph [0004]. This is not the same as Applicant's claim 1, wherein the method is *adapted for processing multimodal input data based upon multimodal training data*.

Furthermore, Applicant amends claim 1 (also claims 18 and 20) to recite "training, by a processor, a recurrent neural network (RNN) using the multimodal training data". Paragraphs [0040]-[0041] of

Applicant's specification support the amendments of the elements and limitations concerning the RNN amendment.

In contrast, *Bumpas* only discloses "a single-pass convolutional neural network (CNN) trained for…image classification." *Bumpas,* paragraph [0025]. This is not the same as Applicant's amended claim 1, comprising, *inter alia*, a CNN and an RNN, *adapted for processing multimodal input data based upon multimodal training data*.

Because *Bumpas* fails to disclose each and every limitation as claimed in Applicant's amended independent claim 1, (Applicant's claims 18 and 20 now recite analogous amendments), Applicant submits that claim 1, and the claims dependent therefrom, are not anticipated by *Bumpas*. Applicant asserts that independent claims 18 and 20 are allowable for similar reasons. Accordingly, Applicant requests that the rejection under 35 U.S.C. § 102(a)(1) be withdrawn.

Applicant submits that claims 1-11 and 18-20 are allowable, as amended. Applicant solicits a Notice of Allowance thereto.

***Claim Rejections under 35 U.S.C. § 103 (OA, paragraphs 18-29)***

Claims 12-13, (OA, paragraphs19-21):

Claims 12 and 13 stand rejected under 35 U.S.C. § 103 as allegedly being unpatentable over *Bumpas* in view of U.S. Pat. Appl. No. 2019/0005670 to DeTone (Hereinafter, "*DeTone*"). Applicant respectfully traverses.

Claims 12 and 13 are dependent from independent claim 1. *DeTone* also does not appear to disclose or make obvious the missing limitations recited in claims 18 and/or 20, namely, the system and the computer-readable non-transitory storage medium *adapted for processing multimodal input data based upon multimodal training data*. Thus, dependent claims 12 and 13 are patentable for the same reasons discussed above with respect to claim 1. Accordingly, Applicant requests that the rejection under 35 U.S.C. § 103 be withdrawn.

Claim 14, (OA, paragraphs 22-23):

Claim 14 stands rejected under 35 U.S.C. § 103 as allegedly being unpatentable over *Bumpas* in view of U.S. Pat. Appl. No. 2019/0005670 to DeTone (Hereinafter, "*DeTone*"), and in further view of Katz (Hereinafter, "*Katz*"). Applicant respectfully traverses.

Claim 14 is dependent from independent claim 1. *Katz* also does not appear to disclose or make in claim 1, namely, the method *adapted for processing multimodal input data based upon multimodal training data*. Thus, dependent claim 14 is patentable for the same reasons discussed above with respect to claim 1. Accordingly, Applicant requests that the rejection under 35 U.S.C. § 103 be withdrawn.

Claim 15, (OA, paragraphs 24-25):

Claim 15 stands rejected under 35 U.S.C. § 103 as allegedly being unpatentable over *Bumpas* in view of U.S. Pat. Appl. No. 2019/0278994 in view of Shebl (Hereinafter, "*Shebl*"). Applicant respectfully traverses.

First, *Shebl* has been shown to post-date the subject matter of the instant independent claims, and therefore cannot be considered as prior art, when applied against this application.

Second, claim 15 is dependent from independent claim 1. *Shebl* also does not appear to disclose or make obvious the missing limitations recited in claim 1, namely, the method *adapted for processing multimodal input data based upon multimodal training data*. Thus, dependent claim 15 is patentable for the same reasons discussed above with respect to claim 1. Accordingly, Applicant requests that the rejection under 35 U.S.C. § 103 be withdrawn.

Claim 16, (OA, paragraphs 26-27):

Claim 16 stands rejected under 35 U.S.C. § 103 as allegedly being unpatentable over *Bumpas* in view of U.S. Pat. Appl. No. 2019/0278994 in view of Shebl (Hereinafter, "*Shebl*"), and in further view of U.S. Pat. Appl. No. 2019/0114489 to Jackson (Hereinafter, "*Jackson*"). Applicant respectfully traverses.

Claim 16 is dependent from independent claim 1. *Shebl* also does not appear to disclose or make obvious the missing limitations recited in claim 1, (as with independent claims 18 and/or 20), namely, the method *adapted for processing multimodal input data based upon multimodal training data*.

Applicant asserts that Jackson is perhaps multi-modal, given the various teachings on sensor data, including camera data, that is combined with the audio data to provide input to the deep learning model (See e.g., Element: 260, FIG. 2). However, Applicant asserts that even if basic multi-modal learning/computing is taught to some extent, *not all multi-modal systems* are suitable for all tasks, and in particular where specialization is required. For example, the multi-modal system of *Jackson* would not be able to classify an automobile from an image. Therefore, the functionality of the system depends on the specific combination of modes employed in the deep learning model and the input to the classification system (which must be at least a subset thereof).

Thus, dependent claim 16 is patentable for on the basis of the aforementioned argument drawing from language in claim 1. Accordingly, Applicant requests that the rejection under 35 U.S.C. § 103 be withdrawn.

Claim 17, (OA, paragraphs 28-29):

Claim 17 stands rejected under 35 U.S.C. § 103 as allegedly being unpatentable over *Bumpas* in view of U.S. Pat. Appl. No. 2019/0005670 to Bergamo (Hereinafter, "*Bergamo*"). Applicant respectfully traverses.

Claim 17 is dependent from independent claim 1. *Bergamo* also does not appear to disclose or make obvious the missing limitations recited in claim 1, namely, the method *adapted for processing* *multimodal input data based upon multimodal training data*. Thus, dependent claim 17 is patentable for the same reasons discussed above with respect to claim 1. Accordingly, Applicant requests that the rejection under 35 U.S.C. § 103 be withdrawn.

Applicant submits that claims 12-17 are allowable, as amended, as are the remaining claims pending herewith. Applicant solicits a Notice of Allowance thereto.

***Conclusion***

In view of the foregoing, it is respectfully submitted that claims 1-20, now pending, are patentably distinct from the references cited and are in condition for allowance. Reconsideration of the application and withdrawal of the rejections of record are respectfully requested.

Applicant submits herewith a three (3) month extension of time fee, as this response to the non-final Office Action dated July 23, 2024 is being filed within the total six (6) month period provided by statute. Reconsideration of the application and withdrawal of the rejections of record are respectfully requested.

In the event that the Examiner wishes to discuss any aspect of this response, please contact the undersigned at the telephone number indicated below.

Respectfully submitted,

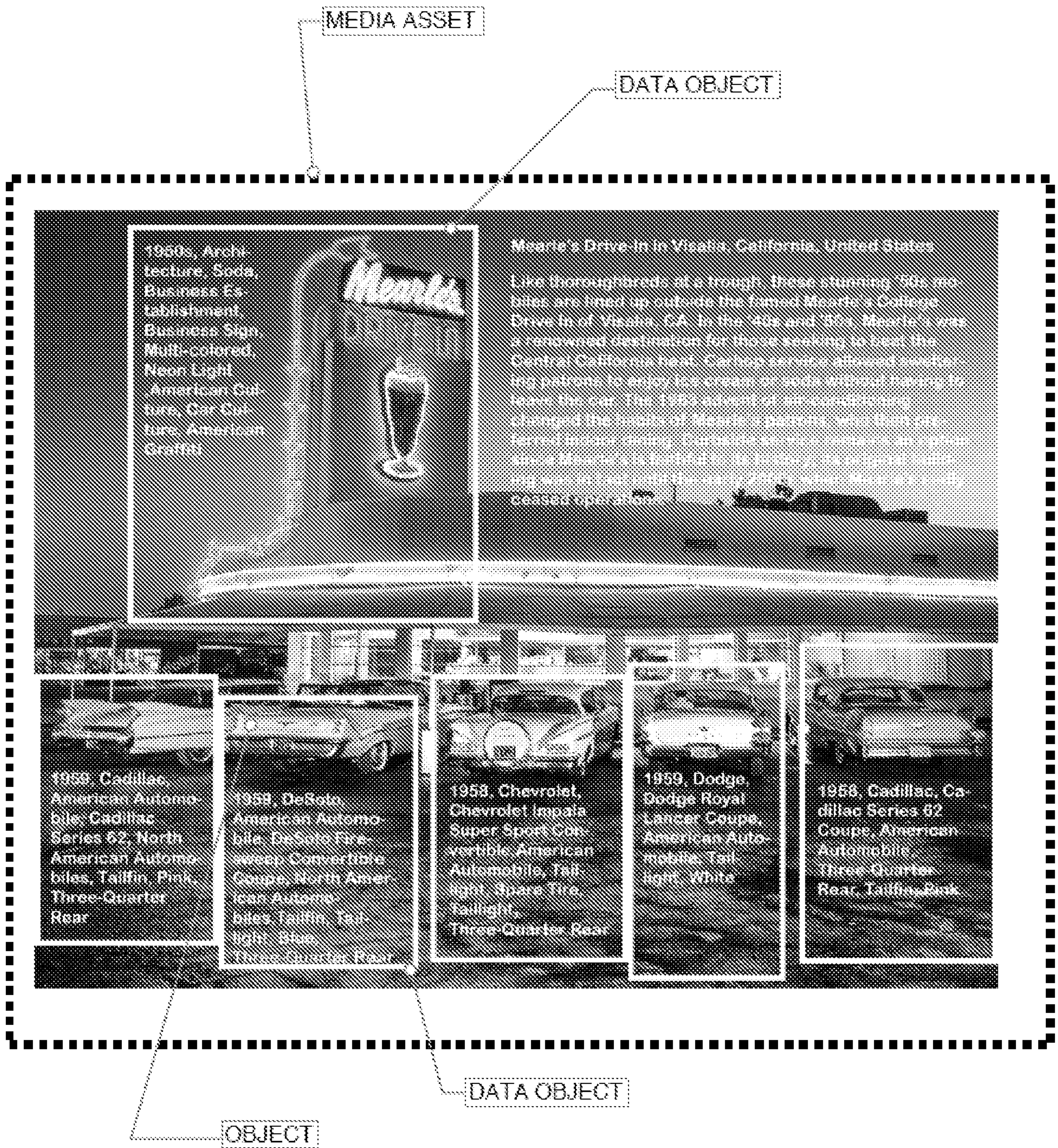
*/Damian Wasserbauer/*
Damian Wasserbauer, Reg. No. 34,749

*/Nicholas Blanton/*
Nicholas Blanton, Reg. No. 78,975

Wasserbauer Law LLC
PO Box 382
Collinsville, CT 06022
(860) 266-1779 | patent@wasserbauerlaw.com

## APPENDIX I

Note: In the exemplary image below, a drive-in restaurant is part of the deep learning aspect of Applicant's invention. Not only is the diner in this context consider as an "object", but the entire image could be considered as an "object", where the vehicles may be considered as providing context to the drive-in restaurant.



PTO/SB/06 (09-11)
Approved for use through 1/31/2014. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

| PATENT APPLICATION FEE DETERMINATION RECORD<br>Substitute for Form PTO-875 | Application or Docket Number<br>17/609,911 | Filing Date<br>11/09/2021 | ☐ To be Mailed |
|---|---|---|---|

ENTITY: ☐ LARGE ☑ SMALL ☐ MICRO

## APPLICATION AS FILED - PART I

| | (Column 1) | (Column 2) | | | |
|---|---|---|---|---|---|
| FOR | NUMBER FILED | NUMBER EXTRA | | RATE ($) | FEE ($) |
| ☐ BASIC FEE (37 CFR 1.16(a), (b), or (c)) | N/A | N/A | | N/A | |
| ☐ SEARCH FEE (37 CFR 1.16(k), (i), or (m)) | N/A | N/A | | N/A | |
| ☐ EXAMINATION FEE (37 CFR 1.16(o), (p), or (q)) | N/A | N/A | | N/A | |
| TOTAL CLAIMS (37 CFR 1.16(i)) | minus 20 = | * | | x $50 = | |
| INDEPENDENT CLAIMS (37 CFR 1.16(h)) | minus 3 = | * | | x $240 = | |
| ☐ APPLICATION SIZE FEE (37 CFR 1.16(s)) | If the specification and drawings exceed 100 sheets of paper, the application size fee due is $310 ($155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s). | | | | |
| ☐ MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j)) | | | | | |
| * If the difference in column 1 is less than zero, enter "0" in column 2. | | | | TOTAL | |

## APPLICATION AS AMENDED - PART II

| | | (Column 1) | | (Column 2) | (Column 3) | | | |
|---|---|---|---|---|---|---|---|---|
| AMENDMENT | 01/23/2025 | CLAIMS REMAINING AFTER AMENDMENT | | HIGHEST NUMBER PREVIOUSLY PAID FOR | PRESENT EXTRA | | RATE ($) | ADDITIONAL FEE ($) |
| | Total (37 CFR 1.16(i)) | * 20 | Minus | ** 20 | = 0 | | x $80 = | 0 |
| | Independent (37 CFR 1.16(h)) | * 3 | Minus | *** 3 | = 0 | | x $240 = | 0 |
| | ☐ Application Size Fee (37 CFR 1.16(s)) | | | | | | | |
| | ☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j)) | | | | | | | |
| | | | | | | | TOTAL ADD'L FEE | 0 |

| | | (Column 1) | | (Column 2) | (Column 3) | | | |
|---|---|---|---|---|---|---|---|---|
| AMENDMENT | | CLAIMS REMAINING AFTER AMENDMENT | | HIGHEST NUMBER PREVIOUSLY PAID FOR | PRESENT EXTRA | | RATE ($) | ADDITIONAL FEE ($) |
| | Total (37 CFR 1.16(i)) | * | Minus | ** | = | | x $0 = | |
| | Independent (37 CFR 1.16(h)) | * | Minus | *** | = | | x $0 = | |
| | ☐ Application Size Fee (37 CFR 1.16(s)) | | | | | | | |
| | ☐ FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j)) | | | | | | | |
| | | | | | | | TOTAL ADD'L FEE | |

| | |
|---|---|
| * If the entry in column 1 is less than the entry in column 2, write "0" in column 3. | LIE |
| ** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20". | /JASON B EADDY/ |
| *** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3". | |

The "Highest Number Previously Paid For" (Total or Independent) is the highest number found in the appropriate box in column 1.

This collection of information is required by 37 CFR 1.16. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 12 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. **SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.**

*If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.*

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 17/609,911 | 11/09/2021 | Lucinda Lewis | P1409US01 | 1651 |

36601 7590 08/06/2024
Wasserbauer Law LLC
PO BOX 382
Collinsville, CT 06022

| EXAMINER | |
|---|---|
| OSINSKI, MICHAEL S | |

| ART UNIT | PAPER NUMBER |
|---|---|
| 2664 | |

| MAIL DATE | DELIVERY MODE |
|---|---|
| 08/06/2024 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

# United States Patent and Trademark Office

*Office of the Commissioner for Patents*

*Dear Applicant,*

***You recently received an Office action from your examiner. Review of your patent application is under way! You can continue to advertise or mark any products covered by your application for your invention as "Patent Pending" for as long as this application is pending. This letter provides information about how to continue through the patent application process and the help available to you on your journey.***

**Where are you in the patent application process?** Your patent application is now in step 3, Application Prosecution (www.uspto.gov/UtilityPatentGuide.) Though the Office action includes a "rejection," that does not mean that you don't have a good idea you can patent. It means your examiner believes your application is not yet ready for allowance.

**Ready to respond?** Responding to the Office action is your opportunity to address the issues raised by your examiner. Responses must be timely, complete, and in writing. In your response, you may want to consider amending your application (www.uspto.gov/patents/maintain/responding-office-actions).

**We provide Pro Se Assistance,** support specifically geared for independent inventors without an agent or attorney. Visit www.uspto.gov/ProSePatents for a consolidated list of resources, including a sample amendment format. Full description of amendment formatting can be found at www.uspto.gov/web/offices/pac/mpep/s714.html. Call us at 866-767-3848, or email proseassistancecenter@uspto.gov.

**Have questions about the Office action?** Contact your examiner! A conversation with your examiner (called an "interview") may help clarify and resolve any issues and possibly identify what you can patent. While an interview does not stop the clock for your official response or substitute for a response, there is no fee for an interview. To schedule a call or video, you can:

1. use the USPTO's Automated Interview Request (AIR) form (www.uspto.gov/patents/uspto-automated-interview-request-air-form);
2. file the Applicant Initiated Interview Request form (PTOL-413A) (www.uspto.gov/sites/default/files/documents/PTOL413A.pdf); or
3. contact your examiner directly via telephone (contact information is provided at the end of your Office action).

**Need help near you?** We have regional offices across the country (www.uspto.gov/locations), as well as Patent and Trademark Resource Centers in libraries in nearly every state (www.uspto.gov/ptrc), with specialists standing by to help you at no cost.

**Become a *registered* electronic filer (registered eFiler).** As a registered eFiler, you can electronically view your submission, track progress of your application, and respond to USPTO correspondence. (www.uspto.gov/NewPatentFilers)

**Contact us**

- General information: USPTO Contact Center — 800-786-9199 • TTY/TDD 800-877-8339 • www.uspto.gov/help
- Patent Center: Patent Electronic Business Center www.uspto.gov/EBC
- Inventors Assistance Center www.uspto.gov/inventorassistance

**Want to learn more?**

- Hear about resources for startups, inventors, and entrepreneurs, as well as our initiatives engaging women, the military community, and others. These programs offer helpful tips and tricks on identifying and managing your IP, finding options for funding, and expanding your network to support your business.
  - www.uspto.gov/startups
  - www.uspto.gov/inventors
  - www.uspto.gov/initiatives
- Access our free services all in one spot. Find all of our contact centers, nationwide services, events, education and training. (www.uspto.gov/FreeServices)
- Connect with us on Instagram, LinkedIn, Twitter, Facebook, and YouTube.
- Hear about the great things we are doing by subscribing at www.uspto.gov/subscribe.

Best,

**Vaishali Udupa**
**Commissioner for Patents**
**United States Patent and Trademark Office**

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
**United States Patent and Trademark Office**
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 17/609,911 | 11/09/2021 | Lucinda Lewis | P1409US00 | 1651 |

36601 7590 07/23/2024

Wasserbauer Law LLC
PO BOX 382
Collinsville, CT 06022

| EXAMINER | |
|---|---|
| OSINSKI, MICHAEL S | |

| ART UNIT | PAPER NUMBER |
|---|---|
| 2664 | |

| MAIL DATE | DELIVERY MODE |
|---|---|
| 07/23/2024 | PAPER |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

| Office Action Summary | Application No. 17/609,911 | Applicant(s) Lewis, Lucinda | |
|---|---|---|---|
| | Examiner MICHAEL S OSINSKI | Art Unit 2664 | AIA (FITF) Status Yes |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTHS FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☑ Responsive to communication(s) filed on 11/9/2021.
☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2a)☐ This action is **FINAL.** 2b) ☑ This action is non-final.

3)☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

4)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims***

5) ☑ Claim(s) 1-20 is/are pending in the application.
5a) Of the above claim(s) _____ is/are withdrawn from consideration.

6) ☐ Claim(s) _____ is/are allowed.

7) ☑ Claim(s) 1-20 is/are rejected.

8) ☐ Claim(s) _____ is/are objected to.

9) ☐ Claim(s) _____ are subject to restriction and/or election requirement

* If any claims have been determined allowable, you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see http://www.uspto.gov/patents/init_events/pph/index.jsp or send an inquiry to **PPHfeedback@uspto.gov.**

**Application Papers**

10)☐ The specification is objected to by the Examiner.

11)☑ The drawing(s) filed on 11/9/2021 is/are: a)☑ accepted or b)☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
**Certified copies:**
a)☐ All b)☐ Some** c)☐ None of the:
1.☐ Certified copies of the priority documents have been received.
2.☐ Certified copies of the priority documents have been received in Application No. _____.
3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

** See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☑ Notice of References Cited (PTO-892)

2) ☑ Information Disclosure Statement(s) (PTO/SB/08a and/or PTO/SB/08b)
Paper No(s)/Mail Date ______.

3) ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date ______.

4) ☐ Other: ______.

### *Notice of Pre-AIA or AIA Status*

The present application, filed on or after March 16, 2013, is being examined under the first inventor to file provisions of the AIA.

## DETAILED ACTION

1. The following Office action is in response to communications filed on 11/9/2021. Claims 1-20 are currently pending within this application.

### *Information Disclosure Statement*

2. The information disclosure statement(s) filed on 11/9/2021 and 7/19/2023 is/are in compliance with the provisions of 37 CFR 1.97, and has/have been considered and a copy/copies is/are enclosed with this Office action.

### *Claim Rejections - 35 USC § 101*

35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

3. Claims 1-20 are rejected under rejected under 35 U.S.C. 101 because the claimed invention is directed to an abstract idea without significantly more.

**<u>Step 1:</u>** Independent claims 1, 18, and 20 recite a method, system, and computer-readable non-transitory storage medium, respectively, which are recognized statutory categories of invention.

**Step 2A, Prong One:** The above mentioned independent claims recite an abstract idea of mental processes which are concepts performed in the human mind (including an observation, evaluation, judgement, and opinion). The claimed functions of "classifying…the input data with respect to the authenticated data and elements of the taxonomy" and "generating a result…wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data" are recited at a high level of generality such that they could be practically performed in the human mind since the limitations, interpreted under their broadest reasonable interpretation and consistent with the specification, cover performance of the limitations in the mind but for generic computer components. These concepts fall into the "mental process" group of abstract ideas, which are observations, evaluations, and/or judgements (classifying input data and generating a result). Other than reciting "a convolutional neural network (CNN)", "a processing device", "a memory", "a processor", "a device", and "a computing device" operated to perform the recited functions, the claims preclude the above mentioned steps from being practically performed in the human mind. *See MPEP 2106.04 and the 2019 PEG.*

**Step 2A, Prong Two:** The abstract idea, as claimed, is not integrated into a practical application. The above mentioned independent claims recite additional structural elements of "a convolutional neural network (CNN)", "a processing device", "a memory", "a processor", "a device", and "a computing device" recited at a high level of generality such that they amount to no more than generic components to implement the abstract idea on a conventional computer. The claims do not point to a specific improvement in computers in their communication role or provides a specific improvement in the way computers operate. As to the additional functional elements of “training a convolutional neural network using authenticated data and a taxonomy”, “receiving, by a processing device, a query comprising input data”, and “displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content”, these are insignificant pre-solution and post-solution data gathering and output activities that do not add more than the insignificant extra-solution activities to the abstract idea (*See MPEP 2016.05(g)*). The claims as a whole, looking at the additional elements individually and in combination, do not integrate the abstract idea into a practical application. See MPEP 2106.04(d).

**Step 2B:** As explained in Step 2A, Prong Two above, the independent claims recite additional elements of “a convolutional neural network (CNN)”, “a processing device”, “a memory”, “a processor”, “a device”, and “a computing device” recited at a high level of generality such that they amount to no more than generic components to implement the abstract idea on a conventional computer. The claim does not point to a specific improvement in computers in their communication role or provides a specific improvement in the way computers operate, As to the additional elements of “training a convolutional neural network using authenticated data and a taxonomy”, “receiving, by a processing device, a query comprising input data”, and “displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content”, these are insignificant pre-solution and post-solution data gathering activities that do not add more than insignificant extra-solution activities to the abstract idea (*See MPEP 2016.05(g)*). The claims as a whole, looking at the additional elements individually and in combination, do not integrate the abstract idea into a practical application and merely append well-understood, routine, and conventional activities known in the industry, specified at a high level of generality, to the judicial exception (*See MPEP 2106.05(d) and Berkheimer Memo*). These limitations generally link the use of the abstract idea to a particular technological environment. Moreover, the additional elements do not reflect an improvement to a technology or technical field, or include the use of a particular machine or particular transformation. The additional elements, taken individually and in combination, do not result in the claim, as a whole, amounting to significantly more than the abstract idea (*See MPEP 2106.04(d) and MPEP 2106.05*).

Therefore, based on the above analysis in conjunction with the 2019 Revised Patent Subject Matter Eligibility Guidance, it is determined that the independent claims are directed towards ineligible subject matter of an abstract idea without significantly more.

Dependent claims 2-17 and 19 are also rejected for being directed towards additional data gathering and data output elements that do not add significantly more to implementing the above identified and discussed abstract idea.

## *Claim Rejections – 35 USC § 102*

4. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
> (a)(1) the claimed invention was patented, described in a printed publication, or in public use, on sale, or otherwise available to the public before the effective filing date of the claimed invention.

5. ***Claims 1-11 and 18-20 are rejected under 35 U.S.C. 102(a)(1) as being anticipated by Bumpas (US PGPub 2019/0278994) [hereafter Bumpas].***

6. As to claim 1, Bumpas discloses a method comprising training a convolutional neural network (CNN) (system 100 shown in Figure 1 including image ingestion module 102, CNN 104, object detection module 108, and image augmentation module 110) using authenticated data (image data/object labels/descriptions) and a taxonomy (categories corresponding to image features/activation maps); receiving, by a processing device, a query (query image 112) comprising input data; classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy (used to derive the trained parameters of the CNN stored in model database 106); generating a result (augmented image 112' and/or image shown in Figures 11 and 14), by the trained CNN, wherein the result comprises authenticated data (image data or a bounding box) and elements of the taxonomy (image features/object labels) comprising a closest match to the input data; and displaying the result on a device (user device as shown in Figure 6), wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality

content (Paragraphs 0027-0028, 0031-0040, 0046-0050, 0062-0064, 0066, 0068-0070, 0073-0074, 0077, 0079-0082).

7. As to claim 2, Bumpas discloses the authenticated data comprises copyright registered works of authorship, metadata and text (Paragraphs 0025, 0032-0033, 0035-0036, 0077-0078).

8. As to claim 3, Bumpas discloses the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings (Paragraphs 0025, 0032-0033, 0035-0036, 0077-0078).

9. As to claim 4, Bumpas discloses the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information (Paragraphs 0025, 0032-0033, 0035-0036, 0077-0078).

10. As to claim 5, Bumpas discloses the authenticated data comprises data from a copyright registered database (database 707) (Paragraphs 0068-0070, 0073).

11. As to claim 6, Bumpas discloses the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors,

vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture (Paragraphs 0032, 0035-0036, 0074, 0077, 0079).

12. As to claim 7, Bumpas discloses the input data comprises one or more of image data, video data, intake data or geographical location data (Paragraphs 0028, 0030, 0034-0035, 0040, 0046-0047, 0070, 0073-0074).

13. As to claim 8, Bumpas discloses classifying comprises mapping input data to authenticated data using the taxonomy (Paragraphs 0031-0033, 0035-0036, 0069-0070, 0073-0074).

14. As to claim 9, Bumpas discloses the result comprises one or more of an image, a video, text, or sound (Paragraphs 0035-0036, 0073, 0079-0082).

15. As to claim 10, Bumpas discloses the result yields one or more of vehicle information, vehicle artifact information or geographical information (Paragraphs 0035-0036, 0073, 0079-0082).

16. As to claim 11, Bumpas discloses generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy (Paragraphs 0042-0044, 0049, 0067, 0069, 0073-0074).

17. As to claims 18-20, the Bumpas reference discloses all claimed subject matter as explained above with respect to the comments of claims 1-2.

## *Claim Rejections – 35 USC § 103*

18. The following is a quotation of 35 U.S.C. 103 which forms the basis for all obviousness rejections set forth in this Office action:

> A patent for a claimed invention may not be obtained, notwithstanding that the claimed invention is not identically disclosed as set forth in section 102 of this title, if the differences between the claimed invention and the prior art are such that the claimed invention as a whole would have been obvious before the effective filing date of the claimed invention to a person having ordinary skill in the art to which the claimed invention pertains. Patentability shall not be negated by the manner in which the invention was made.

19. ***Claims 12-13 are rejected under 35 U.S.C 103 as being unpatentable by Bumpas (US PGPub 2019/0278994) [hereafter Bumpas] in view of DeTone (US PGPub 2019/0005670) [hereafter DeTone].***

20. As to claim 12, it is noted that Bumpas fails to particularly disclose the probability is determined by a cross-entropy function.

On the other hand, DeTone discloses the probability is determined by a cross-entropy function (Paragraphs 0032, 0035).

It would have been obvious to one having ordinary skill in the art before the effective filing date of the invention to include determining a probability through a cross-entropy function because the cited prior art are directed towards augmented reality devices that utilize neural networks for image identification operations and because the claimed limitations are fully disclosed within the cited prior art references and would yield predictable results of tuning the parameters of the neural network using a well-known and established loss function that produces high levels of validation accuracy.

21. As to claim 13, Bumpas discloses the result comprises augmented reality content, wherein displaying the result comprises: displaying the result in an augmented reality apparatus, comprising displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data (Paragraphs 0025, 0035-0036, 0062, 0073, 0080-0082).

It is however noted that Bumpas fails to particularly disclose passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides, and imaging the light at the display device.

On the other hand, DeTone discloses passing light into an eye of a wearer of an augmented reality display device (as shown in Figure 10), said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides, and imaging the light at the display device (Paragraphs 0061-0065).

22. ***Claim 14 is rejected under 35 U.S.C 103 as being unpatentable by Bumpas (US PGPub 2019/0278994) [hereafter Bumpas] and DeTone (US PGPub 2019/0005670) [hereafter DeTone], as applied to claim 13, and in further view of Katz (US PGPub 2018/0084310) [hereafter Katz].***

23. As to claim 14, it is noted that the combination of the Bumpas and DeTone references fails to particularly disclose displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time.

On the other hand, Katz discloses displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time (Paragraphs 0018, 0025, 0072, 0084-0085, 0090-0091).

It would have been obvious to one having ordinary skill in the art before the effective filing date of the invention to include displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time as taught by Katz with the operational method and system of Bumpas and DeTone because the cited prior art are directed towards augmented reality devices that utilize neural networks for image identification operations and because the claimed limitations are fully disclosed within the cited prior art references and would yield predictable results of providing additional verified information to a user of the device.

24. ***Claim 15 is rejected under 35 U.S.C 103 as being unpatentable by Bumpas (US PGPub 2019/0278994) [hereafter Bumpas] in view of Shebl (US PGPub 2022/0207079) [hereafter Shebl].***

25. As to claim 15, it is noted that Bumpas fails to particularly disclose training a recurrent neural network (RNN) using authenticated data and a taxonomy.

On the other hand, Shebl discloses training a recurrent neural network (RNN) using authenticated data and a taxonomy (Paragraphs 0056, 0062, 0089, 0097-0099, 0121, 0137-0139, 0154).

It would have been obvious to one having ordinary skill in the art before the effective filing date of the invention to include training a recurrent neural network (RNN) using authenticated data and a taxonomy as taught by Shebl with the method and operational system of Bumpas because the cited prior art are directed towards augmented reality devices that utilize neural networks for image identification operations and because the claimed limitations are fully disclosed within the cited prior art references and would yield predictable results of enabling semantic text processing to be performed in order to enhance the operational results of the neural network system that identifies and retrieves vehicle information based on user input.

26. ***Claim 16 is rejected under 35 U.S.C 103 as being unpatentable by Bumpas (US PGPub 2019/0278994) [hereafter Bumpas] and Shebl (US PGPub 2022/0207079) [hereafter Shebl], as applied to claim 15, and in further view of Jackson (US PGPub 2019/0114489) [hereafter Jackson].***

27. As to claim 16, it is noted that the combination of the Bumpas and Shebl references fails to particularly disclose the input data comprises unstructured data, the method further comprising: processing, by the trained RNN, the unstructured data to yield structured data; and classifying, by the trained CNN, the structured data.

On the other hand, Jackson discloses the input data comprises unstructured data (sound/audible data received by microphones), the method further comprising: processing, by the trained RNN, the unstructured data to yield structured data; and classifying, by the trained CNN, the structured data (Paragraphs 0035-0037, 0039, 0046-0047).

It would have been obvious to one having ordinary skill in the art before the effective filing date of the invention to include input data comprises unstructured data, the method further comprising: processing, by the trained RNN, the unstructured data to yield structured data; and classifying, by the trained CNN, the structured data as taught by Jackson with the method and operational system of Bumpas and Shelb because the cited prior art are directed towards augmented reality devices that utilize neural networks for image identification operations and because the claimed limitations are fully disclosed within the cited prior art references and would yield predictable results of enabling the content of spoken sentences/words to be transformed into data that can be used by the CNN to classify and generate augmented data.

28. ***Claim 17 is rejected under 35 U.S.C 103 as being unpatentable by Bumpas (US PGPub 2019/0278994) [hereafter Bumpas] in view of Bergamo (US Patent 11853961) [hereafter Bergamo].***

29. As to claim 17, Bumpas discloses the input data comprises user uploaded data, the method further comprising authenticating the user uploaded data and adding the authenticated user uploaded data to the authenticated data (Paragraphs 0028, 0036, 0070, 0073, 0077).

It is however noted that Bumpas fails to particularly disclose using Siamese Neural Networks.

On the other hand, Bergamo discloses using Siamese Neural Networks (Col. 3, 46-64, Col. 4, 9-16, 29-38, Col. 11, 29-48, 54-67, Col. 18, 65-67).

It would have been obvious to one having ordinary skill in the art before the effective filing date of the invention to include using Siamese Neural Networks with the method and operational system of Bumpas because the cited prior art are directed towards augmented reality devices that utilize neural networks for image identification operations and because the claimed limitations are fully disclosed within the cited prior art references and would yield predictable results of enabling the neural networks to be dynamically instantiated during use while reducing memory and processor requirements through the sharing-style network architecture.

## *Conclusion*

30. Any inquiry concerning this communication or earlier communications from the examiner should be directed to MICHAEL S OSINSKI whose telephone number is (571) 270-3949. The examiner can normally be reached on Monday - Friday, 10:00am - 6:00pm. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Nay Maung can be reached on 571-272-7882. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

MO
/MICHAEL S OSINSKI/
Primary Examiner, Art Unit 2664

7/18/2024

| *Notice of References Cited* | Application/Control No. 17/609,911 | Applicant(s)/Patent Under Reexamination Lewis, Lucinda | |
|---|---|---|---|
| | Examiner MICHAEL S OSINSKI | Art Unit 2664 | Page 1 of 4 |

**U.S. PATENT DOCUMENTS**

| * | | Document Number Country Code-Number-Kind Code | Date YYYY-MM-DD | Name | CPC Classification | US Classification |
|---|---|---|---|---|---|---|
| * | A | US-20200005075-A1 | 2020-01-02 | Porter; Bryce Zachary | G06Q30/0278 | 1/1 |
| * | B | US-11783384-B2 | 2023-10-10 | Porter; Bryce Zachary | G06V20/10 | 382/190 |
| * | C | US-20230162488-A1 | 2023-05-25 | Boulanger; Pierre | G06V10/82 | 382/157 |
| * | D | US-11417082-B2 | 2022-08-16 | Zheng; Rui | G06Q30/0643 | 1/1 |
| * | E | US-20200133961-A1 | 2020-04-30 | Wroblewski; Ronald J. | G06N3/08 | 1/1 |
| * | F | US-20210133510-A1 | 2021-05-06 | Boulanger; Pierre | G06F18/241 | 1/1 |
| * | G | US-11568178-B2 | 2023-01-31 | Boulanger; Pierre | G06N3/08 | 1/1 |
| * | H | US-20210012145-A1 | 2021-01-14 | CHAUDHARI; Yogen | G06V30/19167 | 1/1 |
| * | I | US-11544510-B2 | 2023-01-03 | Chaudhari; Yogen | G06V30/413 | 1/1 |
| * | J | US-11790632-B2 | 2023-10-17 | Xu; Juan | G06T7/11 | 382/225 |
| * | K | US-10417499-B2 | 2019-09-17 | Katz; Jeffrey Benjamin | G06V20/46 | 1/1 |
| * | L | US-20190278994-A1 | 2019-09-12 | Bumpas; Derek | G06N3/045 | 1/1 |
| * | M | US-20190138826-A1 | 2019-05-09 | Ghafarianzadeh; Mahsa | G06V20/10 | 1/1 |

**FOREIGN PATENT DOCUMENTS**

| * | | Document Number Country Code-Number-Kind Code | Date YYYY-MM-DD | Country | Name | CPC Classification |
|---|---|---|---|---|---|---|
| | N | | | | | |
| | O | | | | | |
| | P | | | | | |
| | Q | | | | | |
| | R | | | | | |
| | S | | | | | |
| | T | | | | | |

**NON-PATENT DOCUMENTS**

| * | | Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages) |
|---|---|---|
| | U | |
| | V | |
| | W | |
| | X | |

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in YYYY-MM-DD format are publication dates. Classifications may be US or foreign.

| Notice of References Cited | Application/Control No. 17/609,911 | Applicant(s)/Patent Under Reexamination Lewis, Lucinda | |
|---|---|---|---|
| | Examiner MICHAEL S OSINSKI | Art Unit 2664 |  |

**U.S. PATENT DOCUMENTS**

| * | | Document Number Country Code-Number-Kind Code | Date YYYY-MM-DD | Name | CPC Classification | US Classification |
|---|---|---|---|---|---|---|
| * | A | US-20170270625-A1 | 2017-09-21 | Kereth; Lior Ben | G06Q50/184 | 1/1 |
| * | B | US-20170200092-A1 | 2017-07-13 | KISILEV; PAVEL | G06N3/08 | 1/1 |
| * | C | US-20170046613-A1 | 2017-02-16 | Paluri; Balamanohar | G06N3/045 | 1/1 |
| * | D | US-20170039456-A1 | 2017-02-09 | Saberian; Mohammad | G06N3/045 | 1/1 |
| * | E | US-12039648-B2 | 2024-07-16 | Charlton; Ebony James | G06V20/36 | 1/1 |
| * | F | US-12001948-B1 | 2024-06-04 | Teig; Steven L. | G06N20/00 | 1/1 |
| * | G | US-11978106-B2 | 2024-05-07 | Jain; Rahul | G06F18/2413 | 1/1 |
| * | H | US-11853961-B1 | 2023-12-26 | Bergamo; Alessandro | G06V10/454 | 1/1 |
| * | I | US-11741368-B2 | 2023-08-29 | Redford; John | G06V10/751 | 1/1 |
| * | J | US-20230222792-A1 | 2023-07-13 | Badr; Ibrahim | H04N23/633 | 348/169 |
| * | K | US-11475665-B1 | 2022-10-18 | Lin; Qiuying | G06V10/758 | 1/1 |
| * | L | US-11410033-B2 | 2022-08-09 | Neves; Lucas | G06V30/194 | 1/1 |
| * | M | US-11409997-B2 | 2022-08-09 | Yoon; Esther | G06N3/04 | 1/1 |

**FOREIGN PATENT DOCUMENTS**

| * | | Document Number Country Code-Number-Kind Code | Date YYYY-MM-DD | Country | Name | CPC Classification |
|---|---|---|---|---|---|---|
| | N | | | | | |
| | O | | | | | |
| | P | | | | | |
| | Q | | | | | |
| | R | | | | | |
| | S | | | | | |
| | T | | | | | |

**NON-PATENT DOCUMENTS**

| * | | Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages) |
|---|---|---|
| | U | |
| | V | |
| | W | |
| | X | |

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in YYYY-MM-DD format are publication dates. Classifications may be US or foreign.

# *Notice of References Cited*

| | | | |
|---|---|---|---|
| *Notice of References Cited* | Application/Control No. 17/609,911 | Applicant(s)/Patent Under Reexamination Lewis, Lucinda | |
| | Examiner MICHAEL S OSINSKI | Art Unit 2664 | Page 3 of 4 |

**U.S. PATENT DOCUMENTS**

| * | | Document Number Country Code-Number-Kind Code | Date YYYY-MM-DD | Name | CPC Classification | US Classification |
|---|---|---|---|---|---|---|
| * | A | US-20220207079-A1 | 2022-06-30 | SHEBL; Hesham | G06V10/761 | 1/1 |
| * | B | US-11270168-B1 | 2022-03-08 | Otten; Hans | G06F16/55 | 1/1 |
| * | C | US-20220019853-A1 | 2022-01-20 | Hamedi; Jehan | G06N3/08 | 1/1 |
| * | D | US-20220019849-A1 | 2022-01-20 | Kim; Sohyeong | G06N3/08 | 1/1 |
| * | E | US-11227182-B2 | 2022-01-18 | Li; Ruiyu | G06N3/08 | 1/1 |
| * | F | US-20220012533-A1 | 2022-01-13 | Yue; Jun | G06N3/045 | 1/1 |
| * | G | US-20220012526-A1 | 2022-01-13 | PAN; Huadong | G06F16/53 | 1/1 |
| * | H | US-20220012297-A1 | 2022-01-13 | Basu; Suddha Kalyan | G06N3/045 | 1/1 |
| * | I | US-20210382937-A1 | 2021-12-09 | JIE; Zequn | G06F16/55 | 1/1 |
| * | J | US-20210383153-A1 | 2021-12-09 | ZHANG; Zhen | G06V10/82 | 1/1 |
| * | K | US-20210142097-A1 | 2021-05-13 | ZHENG; Rui | G06V10/761 | 1/1 |
| * | L | US-20200356592-A1 | 2020-11-12 | YADA; Ravi Theja | G06F16/538 | 1/1 |
| * | M | US-20190311210-A1 | 2019-10-10 | Chatterjee; Anirban | G06N3/045 | 1/1 |

**FOREIGN PATENT DOCUMENTS**

| * | | Document Number Country Code-Number-Kind Code | Date YYYY-MM-DD | Country | Name | CPC Classification |
|---|---|---|---|---|---|---|
| | N | | | | | |
| | O | | | | | |
| | P | | | | | |
| | Q | | | | | |
| | R | | | | | |
| | S | | | | | |
| | T | | | | | |

**NON-PATENT DOCUMENTS**

| * | | Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages) |
|---|---|---|
| | U | |
| | V | |
| | W | |
| | X | |

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in YYYY-MM-DD format are publication dates. Classifications may be US or foreign.

# *Notice of References Cited*

| | | | |
|---|---|---|---|
| | Application/Control No. 17/609,911 | Applicant(s)/Patent Under Reexamination Lewis, Lucinda | |
| | Examiner MICHAEL S OSINSKI | Art Unit 2664 | Page 4 of 4 |

**U.S. PATENT DOCUMENTS**

| * | | Document Number Country Code-Number-Kind Code | Date YYYY-MM-DD | Name | CPC Classification | US Classification |
|---|---|---|---|---|---|---|
| * | A | US-20200342646-A1 | 2020-10-29 | Wang; Zhaowen | G06N3/045 | 1/1 |
| * | B | US-20190114489-A1 | 2019-04-18 | Jackson; Astrid Glende | G06V10/82 | 1/1 |
| | C | | | | | |
| | D | | | | | |
| | E | | | | | |
| | F | | | | | |
| | G | | | | | |
| | H | | | | | |
| | I | | | | | |
| | J | | | | | |
| | K | | | | | |
| | L | | | | | |
| | M | | | | | |

**FOREIGN PATENT DOCUMENTS**

| * | | Document Number Country Code-Number-Kind Code | Date YYYY-MM-DD | Country | Name | CPC Classification |
|---|---|---|---|---|---|---|
| | N | | | | | |
| | O | | | | | |
| | P | | | | | |
| | Q | | | | | |
| | R | | | | | |
| | S | | | | | |
| | T | | | | | |

**NON-PATENT DOCUMENTS**

| * | | Include as applicable: Author, Title Date, Publisher, Edition or Volume, Pertinent Pages) |
|---|---|---|
| | U | |
| | V | |
| | W | |
| | X | |

*A copy of this reference is not being furnished with this Office action. (See MPEP § 707.05(a).)
Dates in YYYY-MM-DD format are publication dates. Classifications may be US or foreign.

| *Search Notes* | Application/Control No. | Applicant(s)/Patent Under Reexamination |
|---|---|---|
| | 17/609,911 | Lewis, Lucinda |
| | **Examiner** | **Art Unit** |
| | MICHAEL S OSINSKI | 2664 |

| CPC - Searched* | | |
|---|---|---|
| **Symbol** | **Date** | **Examiner** |
| G06N3/08 G06V10/454 G06V10/764 G06V10/82 G06V20/20 G06V20/56 G06N3/042 G06N3/044 G06N3/048 G06F16/433 G06F16/434 G06F16/438 G06F16/90332 G06F18/24143 G06N3/045 G06Q30/018 G06Q30/0241 G06N5/02 G06V2201/08 H04L63/02 | 07/17/2024 | MO |

| CPC Combination Sets - Searched* | | |
|---|---|---|
| **Symbol** | **Date** | **Examiner** |
| | | |

| US Classification - Searched* | | | |
|---|---|---|---|
| **Class** | **Subclass** | **Date** | **Examiner** |
| | | | |

* See search history printout included with this form or the SEARCH NOTES box below to determine the scope of the search.

| Search Notes | | |
|---|---|---|
| **Search Notes** | **Date** | **Examiner** |
| PE2E Search | 07/16/2024 | MO |
| Google Patents Search | 07/16/2024 | MO |
| Consider references cited in IDS | 07/17/2024 | MO |
| Assignee Search | 07/18/2024 | MO |
| Inventor Search | 07/18/2024 | MO |

/MICHAEL S OSINSKI/
Primary Examiner, Art Unit 2664

| **Search Notes** | **Application/Control No.**<br>17/609,911 | **Applicant(s)/Patent Under Reexamination**<br>Lewis, Lucinda |
| --- | --- | --- |
| | **Examiner**<br>MICHAEL S OSINSKI | **Art Unit**<br>2664 |

| **Interference Search** | | | |
| --- | --- | --- | --- |
| **US Class/CPC Symbol** | **US Subclass/CPC Group** | **Date** | **Examiner** |
| | | | |

| | |
| --- | --- |
| /MICHAEL S OSINSKI/<br>Primary Examiner, Art Unit 2664 | |

| *Index of Claims* | Application/Control No.<br>17/609,911 | Applicant(s)/Patent Under Reexamination<br>Lewis, Lucinda |
|---|---|---|
| | Examiner<br>MICHAEL S OSINSKI | Art Unit<br>2664 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| ✓ | Rejected | - | Cancelled | N | Non-Elected | A | Appeal |
| = | Allowed | ÷ | Restricted | I | Interference | O | Objected |

**CLAIMS**

☐ Claims renumbered in the same order as presented by applicant ☐ CPA ☐ T.D. ☐ R.1.47

| CLAIM | | DATE | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Final | Original | 07/18/2024 | | | | | | | | |
| | 1 | ✓ | | | | | | | | |
| | 2 | ✓ | | | | | | | | |
| | 3 | ✓ | | | | | | | | |
| | 4 | ✓ | | | | | | | | |
| | 5 | ✓ | | | | | | | | |
| | 6 | ✓ | | | | | | | | |
| | 7 | ✓ | | | | | | | | |
| | 8 | ✓ | | | | | | | | |
| | 9 | ✓ | | | | | | | | |
| | 10 | ✓ | | | | | | | | |
| | 11 | ✓ | | | | | | | | |
| | 12 | ✓ | | | | | | | | |
| | 13 | ✓ | | | | | | | | |
| | 14 | ✓ | | | | | | | | |
| | 15 | ✓ | | | | | | | | |
| | 16 | ✓ | | | | | | | | |
| | 17 | ✓ | | | | | | | | |
| | 18 | ✓ | | | | | | | | |
| | 19 | ✓ | | | | | | | | |
| | 20 | ✓ | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |
| | | | | | | | | | | |

## Patent Citations (11)

| Publication number | Priority date | Publication date | Assignee | Title |
|---|---|---|---|---|
| Family To Family Citations | | | | |
| EP2745236A1 * | 2011-08-18 | 2014-06-25 | Layar B.V. | Computer-vision based augmented reality system |
| US20170083196A1 * | 2015-09-23 | 2017-03-23 | Google Inc. | Computer-Aided Navigation of Digital Graphic Novels |
| US20170270625A1 * | 2016-03-21 | 2017-09-21 | Facebook, Inc. | Systems and methods for identifying matching content |
| EP3465537A2 * | 2016-05-25 | 2019-04-10 | Metail Limited | Method and system for predicting garment attributes using deep learning |
| US10417499B2 * | 2016-09-21 | 2019-09-17 | GumGum, Inc. | Machine learning models for identifying sports teams depicted in image or video data |
| CN107358596B * | 2017-04-11 | 2020-09-18 | 阿里巴巴集团控股有限公司 | Vehicle loss assessment method and device based on image, electronic equipment and system |
| JP6758250B2 * | 2017-05-22 | 2020-09-23 | 日本電信電話株式会社 | Local feature expression learning device and method |
| US11417082B2 * | 2017-06-16 | 2022-08-16 | Markable, Inc. | Image processing system |
| JP7250709B2 * | 2017-06-28 | 2023-04-03 | マジック リープ, インコーポレイテッド | Method and system for simultaneous localization and mapping using convolutional image transformation |
| US10762424B2 * | 2017-09-11 | 2020-09-01 | Sas Institute Inc. | Methods and systems for reinforcement learning |
| KR102060176B1 * | 2017-09-12 | 2019-12-27 | 네이버 주식회사 | Deep learning method deep learning system for categorizing documents |

* Cited by examiner, † Cited by third party

## Cited By (11)

| Publication number | Priority date | Publication date | Assignee | Title |
|---|---|---|---|---|
| US20220237414A1 * | 2021-01-26 | 2022-07-28 | Nvidia Corporation | Confidence generation using a neural network |
| US20230136902A1 * | 2020-08-24 | 2023-05-04 | Unlikely Artificial Intelligence Limited | Computer implemented method for the automated analysis or use of data |
| US11977854B2 | 2021-08-24 | 2024-05-07 | Unlikely Artificial Intelligence Limited | Computer implemented methods for the automated analysis or use of data, including use of a large language model |
| US11989527B2 | 2021-08-24 | 2024-05-21 | Unlikely Artificial Intelligence Limited | Computer implemented methods for the automated analysis or use of data, including use of a large language model |
| US11989507B2 | 2021-08-24 | 2024-05-21 | Unlikely Artificial Intelligence Limited | Computer implemented methods for the automated analysis or use of data, including use of a large language model |
| Family To Family Citations | | | | |
| CN112528003B * | 2020-12-24 | 2022-10-04 | 北京理工大学 | Multi-item selection question-answering method based on semantic sorting and knowledge correction |
| KR102342055B1 * | 2021-05-28 | 2021-12-27 | 주식회사 애자일소다 | Apparatus and method for processing natural language using structured and unstructured data |
| FR3133475A1 * | 2022-03-08 | 2023-09-15 | Psa Automobiles Sa | Method for analyzing images in a motor vehicle by generating text, associated device and vehicle |
| WO2024084522A1 * | 2022-10-22 | 2024-04-25 | Prashant Huddar | "voice assisted device for object updating and query processing" |
| CN116502640B * | 2023-06-29 | 2023-12-12 | 深圳须弥云图空间科技有限公司 | Text characterization model training method and device based on context |
| CN117274421B * | 2023-11-06 | 2024-04-02 | 北京中数文化科技有限公司 | Interactive scene photo making method based on AI intelligent terminal |

* Cited by examiner, † Cited by third party, ‡ Family to family citation

## Similar Documents

| Publication | Publication Date | Title |
|---|---|---|
| US20220398827A1 | 2022-12-15 | Methods, systems and computer program products for media processing and display |
| Li et al. | 2018 | Humanlike driving: Empirical decision-making system for autonomous vehicles |
| Bhatti et al. | 2021 | Facial expression recognition of instructor using deep features and extreme learning machine |
| US20240046074A1 | 2024-02-08 | Methods, systems and computer program products for media processing and display |
| CN113795851A | 2021-12-14 | Large-scale generation neural network model with reasoning for representation learning using antagonistic training |
| US20190279011A1 | 2019-09-12 | Data anonymization using neural networks |
| Le et al. | 2020 | An overview of deep learning in industry |
| CN117033609B | 2024-02-02 | Text visual question-answering method, device, computer equipment and storage medium |
| CN111797936B | 2023-08-08 | Image emotion classification method and device based on saliency detection and multi-level feature fusion |
| Maybury | 2012 | Multimedia information extraction: Advances in video, audio, and imagery analysis for search, data mining, surveillance and authoring |
| Le et al. | 2017 | DeepSafeDrive: A grammar-aware driver parsing approach to Driver Behavioral Situational Awareness (DB-SAW) |
| Kaur et al. | 2024 | A systematic review of object detection from images using deep learning |
| Ma et al. | 2022 | Vit-dd: Multi-task vision transformer for semi-supervised driver distraction detection |
| Paolicelli et al. | 2022 | Learning semantics for visual place recognition through multi-scale attention |
| Le et al. | 2022 | Tracked-vehicle retrieval by natural language descriptions with domain adaptive knowledge |
| Zhu et al. | 2022 | Image-based storytelling using deep learning |
| Michalski et al. | 2018 | Convolutional neural networks implementations for computer vision |
| Nguyen-Ho et al | 2022 | Text query based traffic video event retrieval with global-local fusion embedding |
| Hou et al. | 2022 | Early warning system for drivers' phone usage with deep learning network |
| CN116977701A | 2023-10-31 | Video classification model training method, video classification method and device |
| Kang et al. | 2021 | ETLi: Efficiently annotated traffic LiDAR dataset using incremental and suggestive annotation |
| Zhao | 2022 | The application of graphic language in animation visual guidance system under intelligent environment |
| Chen | 2022 | A hidden Markov optimization model for processing and recognition of English speech feature signals |
| Kousalya et al. | 2023 | Group Emotion Detection using Convolutional Neural Network |
| Qin et al. | 2021 | Lane recognition system for machine vision |

## Priority And Related Applications

### Priority Applications (1)

| Application | Priority date | Filing date | Title |
|---|---|---|---|
| US17/609,911 | 2019-05-09 | 2020-05-08 | Methods, systems and computer program products for media processing and display |

### Applications Claiming Priority (3)

| Application | Filing date | Title |
|---|---|---|
| US201962845546P | 2019-05-09 | |
| PCT/US2020/032149 | 2020-05-08 | Methods, systems and computer program products for media processing and display |
| US17/609,911 | 2020-05-08 | Methods, systems and computer program products for media processing and display |

## Legal Events

| Date | Code | Title | Description |
| --- | --- | --- | --- |
| 2022-04-19 | AS | Assignment | **Owner name**: AUTOMOBILIA II, LLC, FLORIDA<br>**Free format text**: ASSIGNMENT OF ASSIGNORS INTEREST;ASSIGNOR:LEWIS, LUCINDA;REEL/FRAME:059634/0366<br>**Effective date**: 20220418 |
| 2022-10-03 | STPP | Information on status: patent application and granting procedure in general | **Free format text**: DOCKETED NEW CASE - READY FOR EXAMINATION |

# Bibliographic Data

Application No: 17/609,911

Foreign Priority claimed: ○ Yes ◉ No

35 USC 119 (a-d) conditions met: ☐ Yes ☑ No | ☐ Met After Allowance

Verified and Acknowledged: /MICHAEL S OSINSKI/ Examiner's Signature | Initials

Title: METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

---

| FILING or 371(c) DATE | CLASS | GROUP ART UNIT | ATTORNEY DOCKET NO. |
|---|---|---|---|
| 11/09/2021 | 382 | 2664 | P1409US00 |
| **RULE** | | | |

**APPLICANTS**

Automobilia II, LLC, Coral Gables, FL, UNITED STATES

**INVENTORS**

Lucinda Lewis, Burbank, UNITED STATES

**CONTINUING DATA**

This application is a 371 of PCT/US2020/032149 05/08/2020

**FOREIGN APPLICATIONS**

**IF REQUIRED, FOREIGN LICENSE GRANTED****

07/24/2022

**** MICRO ENTITY ****

**STATE OR COUNTRY**

UNITED STATES

**ADDRESS**

Wasserbauer Law LLC
PO BOX 382
Collinsville, CT 06022
UNITED STATES

**FILING FEE RECEIVED**

$80

Doc code: IDS
Doc description: Information Disclosure Statement (IDS) Filed



# INFORMATION DISCLOSURE STATEMENT BY APPLICANT
**( Not for submission under 37 CFR 1.99)**

| | |
|---|---|
| Application Number | 17609911 |
| Filing Date | 2023-06-14 |
| First Named Inventor | Lucinda LEWIS |
| Art Unit | TBA |
| Examiner Name | TBA |
| Attorney Docket Number | P1409US01 |

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /M.S.O/

## U.S.PATENTS

| Examiner Initial* | Cite No | Patent Number | Kind Code1 | Issue Date | Name of Patentee or Applicant of cited Document | Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear |
|---|---|---|---|---|---|---|
| | 1 | 10140553 | B1 | 2018-11-27 | Vasisht et al. | All relevant. |
| | 2 | 9892133 | B1 | 2018-02-13 | Biessmann et al. | All relevant. |
| | 3 | 10269179 | B2 | 2019-04-23 | Fein et al. | All relevant. |
| | 4 | 10042604 | B2 | 2018-08-07 | Fallon | All relevant. |
| | 5 | 7849491 | B2 | 2010-12-07 | Perlman | All relevant. |
| | 6 | 10192163 | B2 | 2019-01-29 | Wang | All relevant. |

If you wish to add additional U.S. Patent citation information please click the Add button.

## U.S.PATENT APPLICATION PUBLICATIONS

| Examiner Initial* | Cite No | Publication Number | Kind Code1 | Publication Date | Name of Patentee or Applicant of cited Document | Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear |
|---|---|---|---|---|---|---|

| INFORMATION DISCLOSURE STATEMENT BY APPLICANT ( Not for submission under 37 CFR 1.99) | | |
|---|---|---|
| | Application Number | 17609911 |
| | Filing Date | 2023-06-14 |
| | First Named Inventor | Lucinda LEWIS |
| | Art Unit | TBA |
| | Examiner Name | TBA |
| | Attorney Docket Number | P1409US01 |

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /M.S.O/

| | | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 20190050520 | A1 | 2019-02-14 | Alvarez et al. | All relevant. |
| | 2 | 20190114395 | A1 | 2019-04-18 | Lenchner et al. | All relevant. |
| | 3 | 20190102954 | A1 | 2019-04-04 | Bastian et al. | All relevant. |

If you wish to add additional U.S. Published Application citation information please click the Add button.

**FOREIGN PATENT DOCUMENTS**

| Examiner Initial* | Cite No | Foreign Document Number[3] | Country Code[2] i | Kind Code[4] | Publication Date | Name of Patentee or Applicant of cited Document | Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear | T[5] |
|---|---|---|---|---|---|---|---|---|
| | 1 | | | | | | | ☐ |

If you wish to add additional Foreign Patent Document citation information please click the Add button

**NON-PATENT LITERATURE DOCUMENTS**

| Examiner Initials* | Cite No | Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published. | T[5] |
|---|---|---|---|
| | 1 | MADLER, MARK R., "Alternate Futures," San Fernando Valley Business Journal, pp. 9-11, March 18, 2019 . | ☐ |
| | 2 | TAKAHASHI, DEAN, "Gaikai isn't concerned about OnLive's fundamental patent on cloud gaming," VentureBeat, December 15, 2010. | ☐ |
| | 3 | GORDO, ALBERT and LARLUS, DIANE, "Beyond instance-level image retrieval: Leveraging captions to learn a global visual representation for semantic retrieval." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 6589-6598. 2017. | ☐ |

| | | |
|---|---|---|
| **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** ( Not for submission under 37 CFR 1.99) | Application Number | 17609911 |
| | Filing Date | 2023-06-14 |
| | First Named Inventor | Lucinda LEWIS |
| | Art Unit | TBA |
| | Examiner Name | TBA |
| | Attorney Docket Number | P1409US01 |

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /M.S.O/

| | | | |
|---|---|---|---|
| | 4 | WANG, XIAOHUI, "Virtual Reality of 3D Digital Factory Based on Coastal Environment," Journal of Coastal Research, vol. 83, pp. 507-512, Coconut Creek, Florida, USA (2018). | ☐ |

If you wish to add additional non-patent literature document citation information please click the Add button

**EXAMINER SIGNATURE**

| | | | |
|---|---|---|---|
| Examiner Signature | /MICHAEL S OSINSKI/ | Date Considered | 07/17/2024 |

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

| INFORMATION DISCLOSURE STATEMENT BY APPLICANT ( Not for submission under 37 CFR 1.99) | | |
|---|---|---|
| | Application Number | 17609911 |
| | Filing Date | 2023-06-14 |
| | First Named Inventor | Lucinda LEWIS |
| | Art Unit | TBA |
| | Examiner Name | TBA |
| | Attorney Docket Number | P1409US01 |

**CERTIFICATION STATEMENT**

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

**OR**

☒ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☐ See attached certification statement.

☐ The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☒ A certification statement is not submitted herewith.

**SIGNATURE**

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

| Signature | /Damian Wasserbauer/ | Date (YYYY-MM-DD) | 2023-07-18 |
|---|---|---|---|
| Name/Print | Damian Wasserbauer | Registration Number | 34749 |

# PE2E SEARCH - Search History (Prior Art)

| Ref # | Hits | Search Query | DBs | Default Operator | Plurals | British Equivalents | Time Stamp |
|---|---|---|---|---|---|---|---|
| L1 | 50 | (Artificial Intelligence) Similar to: 17/609,911 with 0 CPC Selections and 0 Text Selections<br>Text: (US-20220398827-A1 OR WO-2020227651-A1 OR CA-3139562-A1 OR JP-2022537636-A OR CN-114270412-A OR US-20240046074-A1 OR WO-2022099180-A1 OR US-11847847-B2 OR WO-2022261772-A1 OR US-20230162488-A1 OR US-20200142421-A1 OR DE-102023114645-A1 OR US-20230343068-A1 OR US-20170200092-A1 OR US-20190278994-A1 OR CN-113168502-A OR US-11676182-B2 OR JP-7459425-B2 OR US-20210406996-A1 OR US-11568178-B2 OR US-20210133510-A1 OR US-11978242-B2 OR US-20210216868-A1 OR WO-2022002961-A1 OR US-20170046613-A1 OR US-20220101112-A1 OR US-9805305-B2 OR CN-115605922-A OR CN-112689843-B OR US-11544510-B2 OR US-10963719-B1 OR US-20220189190-A1 OR US-10445597-B2 OR US-20190138826-A1 OR US-10885099-B1 OR CN-108604303-B OR US-20170039456-A1 OR US-20190385023-A1 OR JP-7414367-B2 OR US-20180218429-A1 | (US-PGPUB; USPAT; USOCR; FIT (AU, AP, AT, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, OA, RU, SU, WO, MC, MD, MY, NL, NO, NZ, PH, PL, PT, RO, RS, SE, SG, SI, SK, TH, TN, TR, TW, UA, VN)) | OR | ON | ON | 2024/07/16 04:52 PM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | OR US-10275820-B2 OR US-11809522-B2 OR WO-2022165106-A1 OR US-20230342616-A1 OR DE-102022107480-A1 OR CN-117253069-A OR US-20210073640-A1 OR US-20210012145-A1 OR CN-111204346-A OR WO-2019224164-A1).did. | | | | | |
| L2 | 153 | 2745236A1 "20170083196" "20170270625" 3465537A2 "10417499".pn. "107358596" 6758250B2 "11417082".pn. 7250709B2 "10762424".pn. "102060176" "20230136902" "11977854".pn. "11989527".pn. "11989507".pn. "112528003" "102342055" 3133475A1 "2024084522" "116502640" "117274421" "20220398827" "20240046074" "113795851" "20190279011" "117033609" "111797936" "116977701" | (US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/16 04:58 PM |
| L3 | 224 | L1 L2 | (US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/16 04:58 PM |
| L6 | 7 | 16/458827 | (US-PGPUB; USPAT; USOCR; FIT (AU, AP, AT, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, OA, RU, SU, WO, MC, MD, MY, NL, NO, NZ, PH, PL, PT, RO, RS, SE, SG, SI, SK, | OR | ON | ON | 2024/07/16 05:03 PM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | TH, TN, TR, TW, UA, VN); FPRS; EPO; JPO; DERWENT; IBM_TDB) | | | | |
| L7 | 9 | "20200133961" | (US-PGPUB; USPAT; USOCR; FIT (AU, AP, AT, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, OA, RU, SU, WO, MC, MD, MY, NL, NO, NZ, PH, PL, PT, RO, RS, SE, SG, SI, SK, TH, TN, TR, TW, UA, VN); FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/16 05:13 PM |
| L8 | 6 | "20200406859" | (US-PGPUB; USPAT; USOCR; FIT (AU, AP, AT, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, OA, RU, SU, WO, MC, MD, MY, NL, NO, NZ, PH, PL, PT, RO, RS, SE, SG, SI, SK, TH, TN, TR, TW, UA, VN); FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/16 05:20 PM |
| L9 | 20 | "20190102676" | (US-PGPUB; USPAT; USOCR; FIT (AU, AP, AT, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, OA, RU, SU, WO, MC, MD, MY, NL, NO, NZ, PH, PL, PT, RO, RS, SE, SG, SI, SK, TH, TN, TR, TW, UA, VN); FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/16 05:37 PM |
| L11 | 54292 | ((GAN CNN ANN DNN RNN SVM DBN RBM (support NEAR3 (vector machine)) ((machine | (US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/17 12:15 AM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | convolution convolutional classification classifying deep shallow Siamese Q U dual neural nonlinear non-linear) NEAR3 (network subnetwork sub-network neural nonlinear non-linear model layer tree)) Convnet) WITH (train training trained learn learning learnt learned propagate propagating propagation propagated backpropagate backpropagated backpropagation backpropagating tune tuning tuned optimize optimizing optimization optimized strengthen strengthening strengthened improve improving improvement improved maximize maximizing maximized bolster bolstering bolstered enhance enhancing enhanced enhancement boost boosting boosted revise revising revision revised upgrade upgrading upgraded update updating updated) WITH (image picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene authentic authenticated authenticating authentication copyright copyrighted copyrighting truth story sketch)) AND ((GAN CNN ANN DNN RNN SVM DBN RBM (support NEAR3 (vector machine)) ((machine convolution convolutional classification classifying deep shallow Siamese Q U dual neural | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | nonlinear non-linear) NEAR3 (network subnetwork sub-network neural nonlinear non-linear model layer tree)) Convnet) WITH (train training trained learn learning learnt learned propagate propagating propagation propagated backpropagate backpropagated backpropagation backpropagating tune tuning tuned optimize optimizing optimization optimized strengthen strengthening strengthened improve improving improvement improved maximize maximizing maximized bolster bolstering bolstered enhance enhancing enhanced enhancement boost boosting boosted revise revising revision revised upgrade upgrading upgraded update updating updated) WITH (taxonomy label labeled labeling category categorical hierarchy hierarchical ID identity identification identifying identified pointer context contextual metadata meta-data fringe tag tagging tagged indicator marker pointer header footer extension) WITH (image picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene authentic authenticated authenticating authentication copyright copyrighted copyrighting truth story sketch)) AND ((derive derivation derived deriving acquire | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | acquiring acquirement acquired determine determining determination determined extract extracting extraction extracted infer inferring inference inferred obtain obtaining obtainment obtained generate generating generation generated produce producing production produced process processing processed detect detecting detected detection detector ascertain ascertaining ascertained ascertainment estimate estimating estimation estimated approximate approximating approximation approximated provide providing provided acquisition attain attained attaining attainment receive receiving reception received input inputting inputted supply supplying supplied take taken taking create creating creation created form forming formed retrieve retrieving retrieved retrieval fetch fetching fetched) WITH ((query target reference front frontal first main one reference base main primary principal chief source origin original lead leading premier prime master first one reference base main primary principal chief source origin original lead leading premier prime master start starting started initial initially begin beginning began commence commencement commencing | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | commenced inaugural intro introductory precede precedent preceding prior before past old older oldest known established registered) NEAR3 (image picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene authentic authenticated authenticating authentication copyright copyrighted copyrighting truth story sketch))) | | | | | |
| L12 | 29711 | L11 AND ((GAN CNN ANN DNN RNN SVM DBN RBM (support NEAR3 (vector machine)) ((machine convolution convolutional classification classifying deep shallow Siamese Q U dual neural nonlinear non-linear) NEAR3 (network subnetwork sub-network neural nonlinear non-linear model layer tree)) Convnet) WITH (classify classification classified classifying category classifier categorize categorizing categorized type class grade league tier descriptor description brand branding branded model profile) WITH ((query target reference front frontal first main one reference base main primary principal chief source origin original lead leading premier prime master first one reference base main primary principal chief source origin original lead leading premier prime master | (US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/17 12:15 AM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | start starting started initial initially begin beginning began commence commencement commencing commenced inaugural intro introductory precede precedent preceding prior before past old older oldest known established registered) NEAR3 (image picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene authentic authenticated authenticating authentication copyright copyrighted copyrighting truth story sketch))) | | | | | |
| L13 | 13966 | L12 AND ((choose choice choosing chosen select selecting selection selected elect electing election elected pick picking picked call calling called decide deciding decision decided designate designating designation designated prefer preferring preferred preference preferential want wanted wanting find finding found search searching searched desire desired desiring chose appoint appointed appointment appointing allocate allocation allocated allocating conclude concluding conclusion concluded activate activation activating activated address addressing addressed retrieve retrieving retrieval retrieved output outputting outputted read reading readout receiving | (US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/17 12:16 AM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | received receive reception pull pulling pulled gather gathering gathered extract extraction extracted extracting) WITH (image picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene authentic authenticated authenticating authentication copyright copyrighted copyrighting truth story sketch) WITH (belong belonging belonged correspond corresponding correspondence corresponded correlate correlating correlation correlated match matching matched equal equivalent equaling equaled associate association associated associating connect connection connected connecting liken likening likened link linking linked coincide coincidence coincided coinciding couple coupling coupled join joining joined pairing paired affiliate affiliation affiliated affiliating alliance) WITH ((query target reference front frontal first main one reference base main primary principal chief source origin original lead leading premier prime master first one reference base main primary principal chief source origin original lead leading premier prime master start starting started initial initially begin beginning began commence commencement | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | commencing commenced inaugural intro introductory precede precedent preceding prior before past old older oldest known established registered) NEAR3 (image picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image scene authentic authenticated authenticating authentication copyright copyrighted copyrighting truth story sketch))) | | | | | |
| L14 | 7687 | L13 AND (G06N3/08 G06V10/454 G06V10/764 G06V10/82 G06V20/20 G06V20/56 G06N3/042 G06N3/044 G06N3/048 G06F16/433 G06F16/434 G06F16/438 G06F16/90332 G06F18/24143 G06N3/045 G06Q30/018 G06Q30/0241 G06N5/02 G06V2201/08 H04L63/02).cpc. | (US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/17 12:18 AM |
| L15 | 6213 | L14 AND (((determine determining determination determined decide deciding decision decided discover discovering discovery discovered conclude concluding conclusion concluded ascertain ascertaining ascertained ascertainment check checking checked detect detecting detection detected find finding found obtain obtaining obtainment obtained attain attaining attainment attained | (US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/17 12:23 AM |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | acquire acquiring acquired acquirement measure measuring measurement measured derive derivation derived deriving recognize recognizing recognized recognition sense sensing sensed ID identify identification identified identifying capture capturing captured photographing photographed imaging imaged obtain obtaining obtainment obtained receive receiving reception received search searching searched) WITH (object subject person article body item shoe shirt clothing pants thing region interest ROI region-of-interest product car automobile vehicle cart dolly crane trolley truck SUV van transport train airplane plane aircraft ambulance autocycle auto-cycle bike boat bus cab camper caravan carriage jeep kart limo limousine locomotive motorcycle robot drone UAV railcar tank tractor transportation motorbike home house article object entity structure) WITH (image picture photo photograph video frame movie JPEG MPEG stream live-view through-image preview-image) WITH (type classify classification classified classifying category classifier categorize categorizing categorized type class grade league tier descriptor description brand branding branded model profile make | | | | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | model species genus))) | | | | | |
| L17 | 24 | ("10042604" OR "10140553" OR "10192163" OR "10269179" OR "20190050520" OR "20190102954" OR "20190114395" OR "7849491" OR "9892133" OR "10127250" OR "20120116728" OR "20130066936" OR "20140340423" OR "20150269244" OR "20180084310" OR "20180293552" OR "20190005670" OR "20190102676" OR "6175787" OR "7660437" OR "7817104" OR "8843543" OR "8924415" OR "9818249").pn. | (US-PGPUB; USPAT) | OR | ON | ON | 2024/07/17 04:58 PM |
| L18 | 35 | "20180084310" | (US-PGPUB; USPAT; USOCR; FIT (AU, AP, AT, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, OA, RU, SU, WO, MC, MD, MY, NL, NO, NZ, PH, PL, PT, RO, RS, SE, SG, SI, SK, TH, TN, TR, TW, UA, VN); FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/17 05:01 PM |
| L20 | 2 | 15/842668 | (US-PGPUB; USPAT; USOCR; FIT (AU, AP, AT, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, OA, RU, SU, WO, MC, MD, MY, NL, NO, NZ, PH, PL, PT, RO, RS, SE, SG, SI, SK, TH, TN, TR, TW, UA, VN); FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/17 07:02 PM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| L28 | 11 | L27 AND entropy | (US-PGPUB; USPAT; USOCR; FIT (AP, AT, AU, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, MC, MD, MY, NL, NO, NZ, OA, PH, PL, PT, RO, RS, RU, SE, SG, SI, SK, SU, TH, TN, TR, TW, UA, VN, WO); FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/18 10:38 AM |
| L29 | 1 | 17/609911 AND unstruct$6 | (US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/18 10:42 AM |
| L30 | 309 | L15 AND ((RNN (recurrent NEAR3 network)) WITH (written spoken text letter writing handwriting language voice speak)) | (US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/18 10:43 AM |
| L32 | 4 | "20210012222" AND (CNN convolution$4) | (US-PGPUB; USPAT; USOCR; FIT (AU, AP, AT, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, OA, RU, SU, WO, MC, MD, MY, NL, NO, NZ, PH, PL, PT, RO, RS, SE, SG, SI, SK, TH, TN, TR, TW, UA, VN); FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/18 11:16 AM |
| L33 | 7 | "20210012222" AND (CNN convolution$4 train$4 learn$4 RNN recurrent) | (US-PGPUB; USPAT; USOCR; FIT (AP, AT, AU, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, MC, MD, MY, NL, NO, NZ, OA, PH, PL, PT, RO, RS, RU, SE, SG, SI, SK, SU, TH, TN, TR, TW, UA, VN, WO); FPRS; EPO; | OR | ON | ON | 2024/07/18 11:19 AM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | JPO; DERWENT; IBM_TDB) | | | | |
| L34 | 13 | "20190311210" AND (CNN convolution$4 train$4 learn$4 RNN recurrent) | (US-PGPUB; USPAT; USOCR; FIT (AP, AT, AU, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, MC, MD, MY, NL, NO, NZ, OA, PH, PL, PT, RO, RS, RU, SE, SG, SI, SK, SU, TH, TN, TR, TW, UA, VN, WO); FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/18 11:54 AM |
| L35 | 28 | L17 AND (CNN convolution$4 train$4 learn$4 RNN recurrent) | (US-PGPUB; USPAT; USOCR; FIT (AP, AT, AU, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, MC, MD, MY, NL, NO, NZ, OA, PH, PL, PT, RO, RS, RU, SE, SG, SI, SK, SU, TH, TN, TR, TW, UA, VN, WO); FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/18 12:02 PM |
| L36 | 336 | L15 AND ((RNN (recurrent NEAR3 network)) WITH (input inputting inputted output outputting connect connecting connected connection couple coupled coupling) WITH (CNN (convolution$4 NEAR3 network))) | (US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/18 12:12 PM |
| L38 | 2384 | (G06N3/08 G06V10/454 G06V10/764 G06V10/82 G06V20/20 G06V20/56 G06N3/042 G06N3/044 G06N3/048 G06F16/433 G06F16/434 G06F16/438 G06F16/90332 G06F18/24143 G06N3/045 | (US-PGPUB; USPAT; USOCR; FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/18 12:14 PM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | G06Q30/018 G06Q30/0241 G06N5/02 G06V2201/08 H04L63/02).cpc. AND ((RNN (recur$5 NEAR3 network)) WITH (input inputting inputted output outputting connect connecting connected connection couple coupled coupling) WITH (CNN (convolut$4 NEAR3 network))) | | | | | |
| L39 | 1 | 17/177702 | (US-PGPUB; USPAT; USOCR; FIT (AU, AP, AT, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, OA, RU, SU, WO, MC, MD, MY, NL, NO, NZ, PH, PL, PT, RO, RS, SE, SG, SI, SK, TH, TN, TR, TW, UA, VN); FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/18 12:16 PM |
| L41 | 2 | L40 AND siamese | (US-PGPUB; USPAT; USOCR; FIT (AU, AP, AT, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, OA, RU, SU, WO, MC, MD, MY, NL, NO, NZ, PH, PL, PT, RO, RS, SE, SG, SI, SK, TH, TN, TR, TW, UA, VN); FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/18 12:26 PM |
| L42 | 3 | "20190114489" AND (CNN RNN) | (US-PGPUB; USPAT; USOCR; FIT (AU, AP, AT, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, OA, RU, SU, WO, MC, MD, MY, NL, NO, NZ, PH, PL, PT, RO, | OR | ON | ON | 2024/07/18 03:44 PM |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | RS, SE, SG, SI, SK, TH, TN, TR, TW, UA, VN); FPRS; EPO; JPO; DERWENT; IBM_TDB) | | | | |
| L43 | 21 | ("2017/0178346").urpn. AND (PGPB \| USPT \| USOC).dbnm. | (US-PGPUB; USPAT; USOCR) | OR | ON | ON | 2024/07/18 04:02 PM |
| L45 | 14 | L44 AND (RNN recurrent NEAR3 network) | (US-PGPUB; USPAT; USOCR; FIT (AU, AP, AT, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, OA, RU, SU, WO, MC, MD, MY, NL, NO, NZ, PH, PL, PT, RO, RS, SE, SG, SI, SK, TH, TN, TR, TW, UA, VN); FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/18 04:04 PM |
| L46 | 2 | "20220207079" AND (RNN recurrent NEAR3 network) AND (train$4 learn$4) | (US-PGPUB; USPAT; USOCR; FIT (AU, AP, AT, BE, BG, BR, BY, CA, CH, CN, CS, CU, CZ, DD, DE, DK, EA, EE, EP, ES, FI, FR, GB, HR, HU, ID, IE, IL, IS, IT, JP, KR, LT, LU, LV, MA, OA, RU, SU, WO, MC, MD, MY, NL, NO, NZ, PH, PL, PT, RO, RS, SE, SG, SI, SK, TH, TN, TR, TW, UA, VN); FPRS; EPO; JPO; DERWENT; IBM_TDB) | OR | ON | ON | 2024/07/18 04:07 PM |
| L47 | 2 | ((("LEWIS") near3 ("Lucinda"))).INV. | (US-PGPUB; USPAT; USOCR; EPO; JPO; DERWENT) | OR | ON | ON | 2024/07/18 04:12 PM |
| L48 | 2 | ((("AUTOMOBILIA") near3 ("II") near3 ("LLC"))).AS,AANM. | (US-PGPUB; USPAT) | OR | ON | ON | 2024/07/18 04:12 PM |

## PE2E SEARCH - Search History (Interference)

| |
|---|
| There are no Interference searches to show. |

Doc code: IDS
Doc description: Information Disclosure Statement (IDS) Filed



| **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** ( **Not for submission under 37 CFR 1.99**) | | |
|---|---|---|
| | Application Number | |
| | Filing Date | |
| | First Named Inventor | LEWIS, Lucinda |
| | Art Unit | |
| | Examiner Name | |
| | Attorney Docket Number | 28108-125 |

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /M.S.O/

**U.S.PATENTS** Remove

| Examiner Initial* | Cite No | Patent Number | Kind Code1 | Issue Date | Name of Patentee or Applicant of cited Document | Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear |
|---|---|---|---|---|---|---|
| | 1 | 10127250 | B2 | 2018-11-13 | Dingman et al. | |
| | 2 | 9818249 | B1 | 2017-11-14 | Fraser et al. | |
| | 3 | 8924415 | B2 | 2014-12-30 | Thomas | |
| | 4 | 8843543 | B2 | 2014-09-23 | Damman et al. | |
| | 5 | 7817104 | B2 | 2010-10-19 | Ryu et al. | |
| | 6 | 7660437 | B2 | 2010-02-09 | Breed | |
| | 7 | 6175787 | B1 | 2001-01-16 | Breed | |

If you wish to add additional U.S. Patent citation information please click the Add button. Add

**U.S.PATENT APPLICATION PUBLICATIONS** Remove

| INFORMATION DISCLOSURE STATEMENT BY APPLICANT ( Not for submission under 37 CFR 1.99) | | |
|---|---|---|
| | Application Number | |
| | Filing Date | |
| | First Named Inventor | LEWIS, Lucinda |
| | Art Unit | |
| | Examiner Name | |
| | Attorney Docket Number | 28108-125 |

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /M.S.O/

| Examiner Initial* | Cite No | Publication Number | Kind Code[1] | Publication Date | Name of Patentee or Applicant of cited Document | Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear |
|---|---|---|---|---|---|---|
| | 1 | 20180084310 | A1 | 2018-03-22 | GumGum Inc. | |
| | 2 | 20190005670 | A1 | 2019-01-03 | Magic Leap, Inc. | |
| | 3 | 20190102676 | A1 | 2019-04-04 | SAS Institute Inc. | |
| | 4 | 20180293552 | A1 | 2018-10-11 | Alibaba Group Holding Limited | |
| | 5 | 20150269244 | A1 | 2015-09-24 | Qamar et al. | |
| | 6 | 20140340423 | A1 | 2014-11-20 | Taylor et al. | |
| | 7 | 20130066936 | A1 | 2013-03-14 | Krishnan et al. | |
| | 8 | 20120116728 | A1 | 2012-05-10 | Shear et al. | |

If you wish to add additional U.S. Published Application citation information please click the Add button. Add

FOREIGN PATENT DOCUMENTS Remove

| Examiner Initial* | Cite No | Foreign Document Number[3] | Country Code[2] i | Kind Code[4] | Publication Date | Name of Patentee or Applicant of cited Document | Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear | T[5] |
|---|---|---|---|---|---|---|---|---|

| INFORMATION DISCLOSURE STATEMENT BY APPLICANT ( Not for submission under 37 CFR 1.99) | | |
|---|---|---|
| | Application Number | |
| | Filing Date | |
| | First Named Inventor | LEWIS, Lucinda |
| | Art Unit | |
| | Examiner Name | |
| | Attorney Docket Number | 28108-125 |

ALL REFERENCES CONSIDERED EXCEPT WHERE LINED THROUGH. /M.S.O/

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2018197835 | WO | A1 | 2018-11-01 | BLIPPAR.COM LTD. | | |
| | 2 | 2018065549 | WO | A1 | 2018-04-12 | BLIPPAR.COM LTD. | | |

If you wish to add additional Foreign Patent Document citation information please click the Add button Add

**NON-PATENT LITERATURE DOCUMENTS** Remove

| Examiner Initials* | Cite No | Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published. | T[5] |
|---|---|---|---|
| | 1 | International Search Report for PCT Patent Application No. PCT/US2020/032149 Issued on August 4, 2020 5 pages. | |
| | 2 | White Paper "The Content Authenticity Initiative" Setting the Standard for Digital Content Attribution, Rosenthol, L., Parsons, A., Scouten, E., Athora, J., MacCormack, B., England, P., Levallee, M., Dotan, J., Hanna, S., Farid, H. Gregory, S (Witness) August 2020, 28 pages. | |

If you wish to add additional non-patent literature document citation information please click the Add button Add

**EXAMINER SIGNATURE**

| Examiner Signature | /MICHAEL S OSINSKI/ | Date Considered | 07/17/2024 |
|---|---|---|---|

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

| | | |
|---|---|---|
| **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** ( **Not for submission under 37 CFR 1.99)** | Application Number | |
| | Filing Date | |
| | First Named Inventor | LEWIS, Lucinda |
| | Art Unit | |
| | Examiner Name | |
| | Attorney Docket Number | 28108-125 |

**CERTIFICATION STATEMENT**

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

**OR**

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

See attached certification statement.

The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

X A certification statement is not submitted herewith.

**SIGNATURE**

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

| | | | |
|---|---|---|---|
| Signature | /Mary K. Nicholes/ | Date (YYYY-MM-DD) | 2021-11-09 |
| Name/Print | Mary K. Nicholes | Registration Number | 56,238 |

# Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

July 25, 2023

WASSERBAUER LAW LLP
Attn: DAMIAN WASSERBAUER
220 ALBANY T PKE.
SUITE 112
CANTON, CT 06019

Dear Sir/Madam,

The United States Patent and Trademark Office (USPTO) has reviewed your refund request for Reference Number **17609911** - Refund Request ID **REFND-20230720-31666**. Below is the current status of your refund request assigned to the following processing area: Office of Patent Examination Support Services

| Fee code | Decision | Decision Reason(s) | Refund Amount |
|---|---|---|---|
| 3806 | Approved | No Fee Due | $52.00 |

Generally, approved refunds will be issued in the form that the original payment was provided. You will receive a separate decision letter for any additional fee codes assigned to other processing areas, if applicable. Decisions may be subject to change if an error is identified.

For questions related to this decision, contact the Case Resolution Team at (571) 272-1720 or case.resolution@USPTO.GOV.

Thank you,
Refund Branch

**United States Patent and Trademark Office**

*Office of the Chief Financial Officer*

Document Code:WFEE

User :Rochelle Geter

Sale Adjustment Accounting Date:07/25/2023

| Effective Date | Sale Accounting Date | Sale Item Reference Number |
|---|---|---|
| 07/19/2023 | 07/25/2023 | 17609911 |

| Document Number | Fee Code | Fee Code Description | Amount Paid | Payment Method |
|---|---|---|---|---|
| I20237O044326429 | 3806 | SUBMISSION- INFORMATION DISCLOSURE STMT | $52.00 | |

**United States Patent and Trademark Office**

*Office of the Chief Financial Officer*

Document Code:WFEE

User :Rochelle Geter

Refund Accounting Date:07/25/2023

| Effective Date | Sale Item Reference Number | Refund Total |
|---|---|---|
| 07/19/2023 | 17609911 | $52.00 |

| Document Number | Fee Code | Fee Code Description | Amount Paid | Payment Method | Account Number |
|---|---|---|---|---|---|
| I20237O044326429 | 3806 | SUBMISSION-INFORMATION DISCLOSURE STMT | $52.00 | | *************4007 |

UNITED STATES
PATENT AND TRADEMARK OFFICE

P.O. Box 1450
Alexandria, VA 22313 - 1450
www.uspto.gov

# ELECTRONIC ACKNOWLEDGEMENT RECEIPT

| APPLICATION # | RECEIPT DATE / TIME | ATTORNEY DOCKET # |
|---|---|---|
| **17/609,911** | **07/19/2023 10:46:36 PM ET** | **28108-125** |

## Title of Invention

METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

## Application Information

| | | | |
|---|---|---|---|
| APPLICATION TYPE | Utility - U.S. National Stage under 35 USC 371 | PATENT # | - |
| CONFIRMATION # | 1651 | FILED BY | John Wasserbauer |
| PATENT CENTER # | 62472232 | FILING DATE | 11/09/2021 |
| CUSTOMER # | 36601 | FIRST NAMED INVENTOR | Lucinda Lewis |
| INTL. APPLICATION # | - | INTL. FILING DATE | - |
| CORRESPONDENCE ADDRESS | - | AUTHORIZED BY | Nicholas Blanton |

## Documents

**TOTAL DOCUMENTS: 1**

| DOCUMENT | PAGES | DESCRIPTION | SIZE (KB) |
|---|---|---|---|
| P1409US01_IDS_Refund_Request_Form_7-19-23.pdf | 1 | Refund Request | 713 KB |

## Digest

| DOCUMENT | MESSAGE DIGEST(SHA-512) |
|---|---|
| P1409US01_IDS_Refund_Request_Form_7-19-23.pdf | 487ABE4514C44DE48C02B31277C39E646EAD584C6D1ECBBB0495E3DEDEAB07B3F5AAFC579F5D3BFDB3C114BDB861B9C06116500D1632EDF35D5F6DC317C8B22F |

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

**New Applications Under 35 U.S.C. 111**

If a new application is being filed and the application includes the necessary components for filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application

**National Stage of an International Application under 35 U.S.C. 371**

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

**New International Application Filed with the USPTO as a Receiving Office**

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

# REQUEST FOR REFUND

| **Reference #** *(Patent #, Application #, Trademark Serial #, Registration #, etc.)* | **Title of Invention or Mark Information** | |
|---|---|---|
| 17609911 | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | |
| **Attorney Docket #** *(if applicable)* | **Payment Date** *(mm/dd/yyyy)* | **Refund Request Amount** |
| P1409US01 | 07/19/2023 | $ 52.00 |

**Refund Option** *(Select one)*

- (●) If approved, issue the refund to the account associated with the original payment.
- ( ) If approved, and the original payment was a check, issue the refund to deposit account # ______________.
- ( ) If approved, and the original payment was a check, issue the refund as a U.S. Treasury check.

**Reason for Refund Request** *Refund requests must generally be filed within 2 years of payment date (37 CFR 1.26 and 2.209)*

- [ ] Duplicate Payment
- [x] No Fee Due
- [ ] Office Error
- [ ] Other:
- [ ] Small Entity Later Established*

**(Must be filed within 3 months of payment date, 37 CFR 1.28. There is no refund when micro entity is later established.)*

**Rationale** *(Supporting documentation may be submitted with this form)*

IDS filed before first office action - no fee due, 37 C.F.R. 1.97.

## Requester's Information

| **Company or Firm Name** *(if applicable)* | | |
|---|---|---|
| Wasserbauer Law LLP | | |
| **Address Line 1** | **City** | **State/Region** |
| 220 Albany Tpke. | Canton | CT |
| **Address Line 2** *(if applicable)* | **Country** | **Zip/Postal Code** |
| Ste.112 | USA | 06019 |

| **Email Address** *(You will receive an acknowledgment of receipt only if you provide a valid email address)* | **Phone Number** |
|---|---|
| patent@wasserbauerlaw.com | 202-507-1776 |
| **Requester's Name** | **Registration Number** *(if applicable)* |
| Damian Wasserbauer | 34749 |
| **Signature** | **Date** *(mm/dd/yyyy)* |
| /Damian Wasserbauer/ | 07/19/2023 |

**Submit online:** USPTO's patent electronic filing system (Patent Center or EFS-Web), or
**Fax to:** 571-273-6500, or
**Mail to:** Mail Stop 16, Director of the U.S. Patent and Trademark Office, P.O. Box 1450, Alexandria, VA 22313-1450

| Journal of Coastal Research | SI | 83 | 507–512 | Coconut Creek, Florida | 2018 |
|---|---|---|---|---|---|

# Virtual Reality of 3D Digital Factory Based on Coastal Environment

**Xiaohui Wang**

School of Mechanical Engineering
University of Science and Technology
Beijing 100083, China
wangxh14@ustb.edu.cn

www.cerf-jcr.org

www.JCRonline.org

**ABSTRACT**

Wang, X., 2018. Virtual reality of 3D digital factory based on coastal environment. *In*: Liu, Z.L. and Mi, C. (eds.), *Advances in Sustainable Port and Ocean Engineering. Journal of Coastal Research*, Special Issue No. 83, pp. 507–512. Coconut Creek (Florida), ISSN 0749-0208.

This paper establishes a complete virtual reality based 3D digital factory, to improve the production and training efficiencies and reduce labour intensity. First, the 3D interactive modes in virtual reality scenes are designed, especially those in roam and device operation in detail. Then, the real-time rendering problem of large-scale scenes in digital factories is solved from the aspects of model optimization and dynamic preloading and undoing of models. Finally, the automatic semantic annotation algorithm of 3D models is proposed by using deep learning methods and Internet images, which improves the development efficiency. The results show the feasibility and effectiveness of application of virtual reality technologies in digital factories.

**ADDITIONAL INDEX WORDS:** *Virtual reality, digital factory, interactive mode, model semantics, real-time rendering.*

## INTRODUCTION

Virtual Reality (VR) uses computers to generate a virtual environment, with the help of the natural interaction mode of multi-source information fusion to make users immersed in the virtual environment (Burdea and Coiffet, 1994). Virtual reality technology integrates the computer graphics, image processing, multimedia, sensors, networks and other multidisciplinary achievements. Now virtual reality is widely used in education, health care, art, city planning and many other fields (Biocca and Levy, 1995; Zhang and Jun, 2005). Digital factory is to provide a full control for the manufacturing process by using the information technology (Caggiano, Caiazzo, and Teti, 2015; Canetta, Redaelli, and Flores, 2011). With the development of industry 4.0 and virtual reality technology, the virtual reality based 3D digital factory coastal environment can meet the intelligent and personalized needs, which is the future of the digital factory.

This paper develops a complete virtual reality based 3D digital factory coastal environment by using Unity 3D. Three key issues in this system are discussed. First, how to design interactive modes in virtual reality scenarios is an important factor affecting user experience. This paper takes roam in VR scenes as an example to discuss the interactive modes in detail. Second, real-time rendering of large-scale scenes is one of the important issues to be solved in the 3D digital factory coastal environment, because the number of models in the digital factory scene is very huge. Third, some models in the digital factory contain semantic information and require to be treated differently with their different semantics. For example, the staircase allows users to climb up or down. The manual annotation of model semantics is very time-consuming. How to automatically identify semantic information of models is very helpful to improve the development efficiency.

The rest of the paper is organized as follows: Section 2 gives related work. Section 3 designs the interactive modes in VR Scene by taking the interactive modes of roam and device operations as examples. Section 4 proposes the solution of real-time rendering of large-scale scenes. Section 5 proposes the automatic semantic annotation algorithm of 3D models. Section 6 finally draws the conclusions.

## RELATED WORK

### Virtual Reality

Virtual reality technology contains three characteristics: Immersion, interactivity and imagination (Burdea and Coiffet, 1994). Immersion refers to the real extent of the user in the virtual reality environment. Interactivity is how naturally that users can manipulate objects in a virtual environment and get feedback by means of special devices such as helmet and handles. Creativity refers to the user's ability to create ideas in the virtual world based on acquired multi-source information and their own logical judgment.

The rapid development of virtual reality in recent years attributed to a variety of virtual reality hardware. Virtual reality devices mainly include two categories: Monitors and interactive devices. Monitors contain helmet-based display, such as HTC Vive and Oculus Rift, which are best at present. Google CardBoard, a two-lens smartphone head-mounted display, is the lowest cost. Interactive devices contain handles, data gloves, 3D scanners, force feedback devices and more. In this paper, HTC Vive is adopted, which includes a helmet display, two handles for

DOI: 10.2112/SI83-083.1 received 1 November 2017; accepted in revision 25 February 2018.

manipulating objects in the scene, and an infrared-based spatial tracking system.

The development of virtual reality technology includes 3D modelling, visual real-time dynamic rendering, human-computer interaction, physical simulation and other parts. Most of the 3D modelling uses 3ds Max, Maya, Other parts use 3D game engine, such as Unity 3D and Unreal Engine. We use Unity 3Din this paper.

### 3D Digital Factory Coastal Environment

3D digital factory realizes the integration, simulation, analysis and control of the entire manufacturing plant. It is widely used in petroleum, petrochemical, chemical, electric power and metallurgy areas. There are many enterprises of 3D digital factory coastal environment solutions, such as INOVx (URL: Http://www.inovx.com/). These companies are all focused on the scan recovery and 3D modelling, and have not yet been combined with the virtual reality technology.

The virtual reality based 3D digital factory coastal environment contains the vivid display of the factory and integrates with the background controller. According to the users' operation, it demonstrates the actual state of the plant, which can greatly enhance the user experience. Previous work focuses on the different aspects of the virtual reality based 3D digital factory coastal environment (Żywicki, Zawadzki, and Górski, 2017). Zhou *et al.* (2011) realized the VR based simulation system, which mainly focuses on the connection between the background controller and data simulation. This paper pays more attention on the 3D interaction design in the virtual scenes to improve user experiences. Some work implemented virtual roaming system (Onosato and Iwata, 1993). The number and complexity of the models in these scenes is much smaller and much simpler than those in the real factory. Therefore, this paper explores the real-time rendering of large-scale scenes. Besides, previous work neglected the semantic attributes of the models in the digital factory. But in the real factory, we need to take different actions when roaming and checking different devices based on their semantics. Manual semantic annotation of 3D models in large-scale scenes is time consuming. So, this paper studies the automatic semantic labelling algorithm of 3D model to improve the development efficiency.

## INTERACTIVE MODES IN VR SCENE

Traditional interactive modes with the mouse and keyboard are no longer valid for virtual reality scenarios (Yin and He, 2013). The virtual reality scene simulates the real environment of the real world. Therefore, it is necessary to explore the most natural interactive mode. Users are freed from the mouse and keyboard and return to the most essential interactions such as body movement and gestures.

In this paper, we use HTC Vive, which has two handles, as shown in Figure 1. Based on the keys on this handle, a 3D interactive approach is designed.

### Interactive Modes of Roam

Taking roam as an example, three different roam modes are proposed. First, with the spatial tracking system of HTC Vive, users can walk directly to roam the scene, as shown in Figure 2.

Second, with the touch pad on the handle, 360-degree precision movement in the VR scene is achieved. The speed and direction of the user movement should be set. In our system, the moving speed is proportional to the distance from the touch point to the centre of the touch pad, and the direction of movement is the direction from the centre to the touch point. The formal description is as follows.

*P*: The user's touch point on the touch pad, denoted as $(x, y)$ in the two-dimensional position, and $(x, 0, y)$ in the local coordinate system of the touch pad.

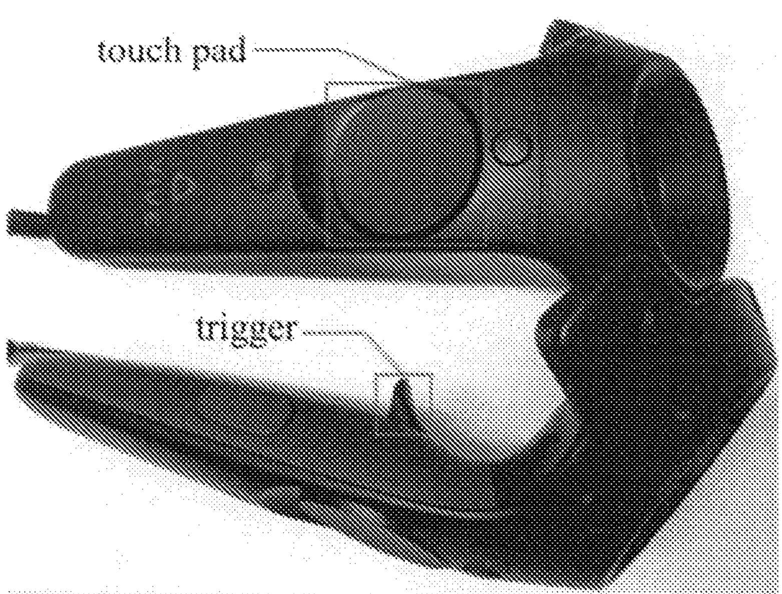



Figure 1. The handles of HTC Vive.

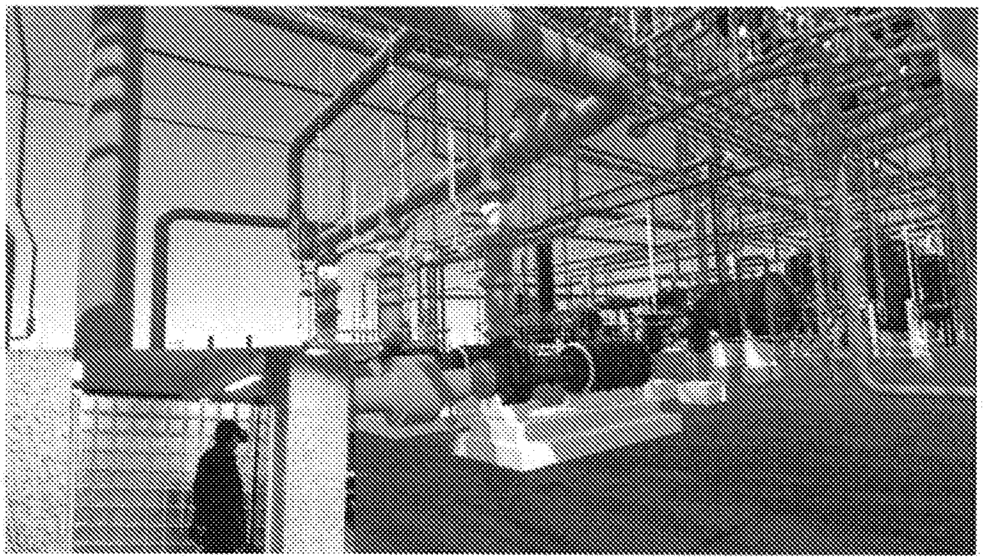


Figure 2. The walking mode of roam in VR scene.

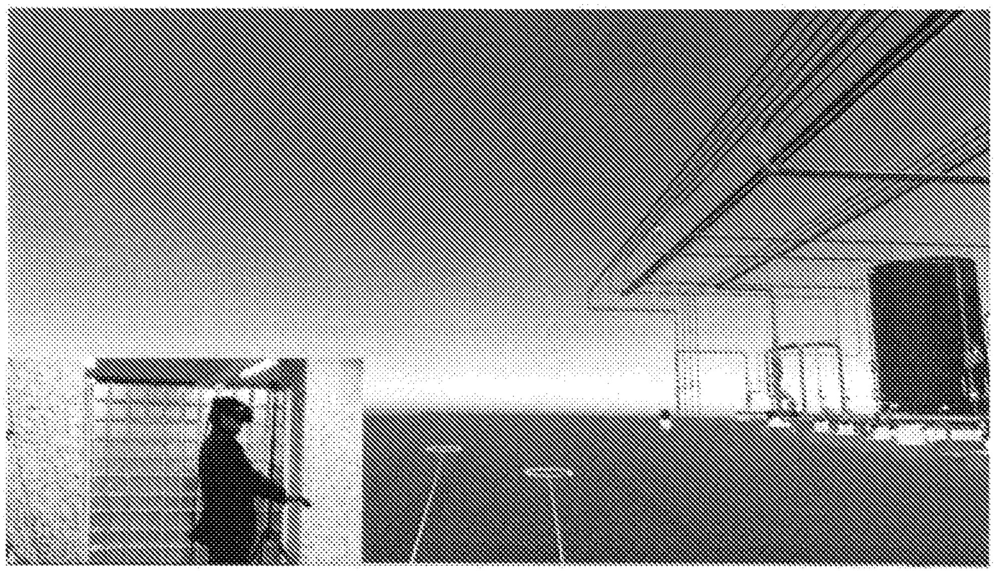


Figure 3. The quick roam in VR scene by the trigger.

$O$: The centre point of the touch pad. Its coordinate is (0,0,0) in the local coordinate system of the touch pad.

$R$: The rotation matrix of the touchpad in the world coordinate.

$P_1$: The coordinate of vector OP in the local coordinate system of the touch pad, denoted as $(x, 0, y)$.

$P_2$: The coordinate of vector OP in the world coordinate system, denoted as $R \cdot P_1$.

$P_3$: The projection of $P_2$ in the horizontal plane in Unity 3D system.

$v$: The velocity.

So, the direction of the user movement is $P_3$ and the speed is $v \cdot P_3$.

Third, the handle points to a position in the scene. At this moment, a ring appears at the position. By activating the trigger, the current position immediately switches to the position of the ring to achieve a quick roam of the scene, as shown in Figure 3.

The key issue here is to obtain the target position. A ray is emitted from the direction of the handle, and the intersection of the ray and the ground is the target point. The algorithm is as follows.

In the local coordinate system of the handle, the direction of the handle is $V_l = (0,0,1)$.

The rotation matrix of the handle in the world coordinate is denoted as $R$, the direction of the handle $V_2 = R \cdot V_1$ in the world coordinate system.

The origin of the handle and $V_2$ determine a ray $V_3$, that is the start point of the ray. The normal of the ground is (0, 1, 1). The height of the ground is the user's height. The above information determines a plane, denoted as $M$. The intersection between $M$ and $V_3$ is the target position.

### Interactive Modes of Device Operations

In the digital factory, there are lots of device operations. For example, you need to check the device parameters in real time when connecting to the background controller of the actual production process. Taking checking the device information as an example, we design and implement the interactive modes of device operation. In traditional 2D digital factory, the interactive mode is to left-click a device and view it in a pop-up dialog box. Due to this familiar mode, we design the similar way in the 3D virtual reality space as follows. The user points to a device by using the handle. And a yellow ray appears in the scene. Then a dialog box is popped in the front of the device and displays the device information toward the user, shown in Figure 4.

The implementation detail is as follows. For each device with information to be shown, a surrounding ellipsoid is established. Find the ellipsoids which the ray goes through. If there are multiple ellipsoids intersecting with the ray, just take the ellipsoid nearest to the user. Then get the device information corresponding to the ellipsoid in the database and display it.

## REAL-TIME RENDERING OF LARGE-SCALE SCENES

In this paper, the real-time rendering of large-scale scenes in the digital factory is comprehensively solved from the aspects of model optimization and dynamic preloading and undoing of models. There are 3D models in many factories before they establish the virtual reality based digital factories. We need to import these models into Unity 3D. VR requires far more real-time rendering than PC, so model optimization is essential for large-scale scenes.

First, remove the fragmented model. When modelling or model recovery based on the scanned point cloud, lots of useless model fragments are generated. This paper adopts the method of surface patch clustering (Belongie, Malik, and Puzicha, 2000; Chavdar, Tim, and Michael, 2015; Gao and Wang, 2017; Gleba, Grulkowski, and Zariczny, 2017; Klaycham, Athisakul, and Chucheepsakul, 2017; Krizhevsky, Sutskever, and Hinton, 2012; Mi *et al.*, 2015; Wang *et al.*, 2012; Yang *et al.*, 2017) and deletes the model fragments in the small categories.

Second, merge models with the same semantics, which can reduce the overhead of API calls and improve rendering efficiency. For large-scale scenes, manual annotation is time consuming, an automatic semantic annotation of 3D models is proposed in this paper, described in detail in the next section.

Finally, the model is preloaded and withdrawn dynamically, especially only the models which can be seen from the user's eyesight are loaded. According to the user behaviour, the user's position and vision range in next seconds are predicted. Then the corresponding models are loaded in advance. And load the rough version of the models farther away. Besides, the models which have not been concerned about in a long time are withdrawn from memory.

## AUTOMATICSEMANTICANNOTATIONOF3DMODELS

With the rapid development of the Internet and the popularity of cameras, data-driven image processing is increasingly attracting the researchers' attention. Chen *et al.* (2009) used Internet images to construct complete images based on the input sketches. Some work used Internet images with semantics for image coloring (Wang *et al.*, 2012). This paper uses Internet images for automatic annotation of 3D models.

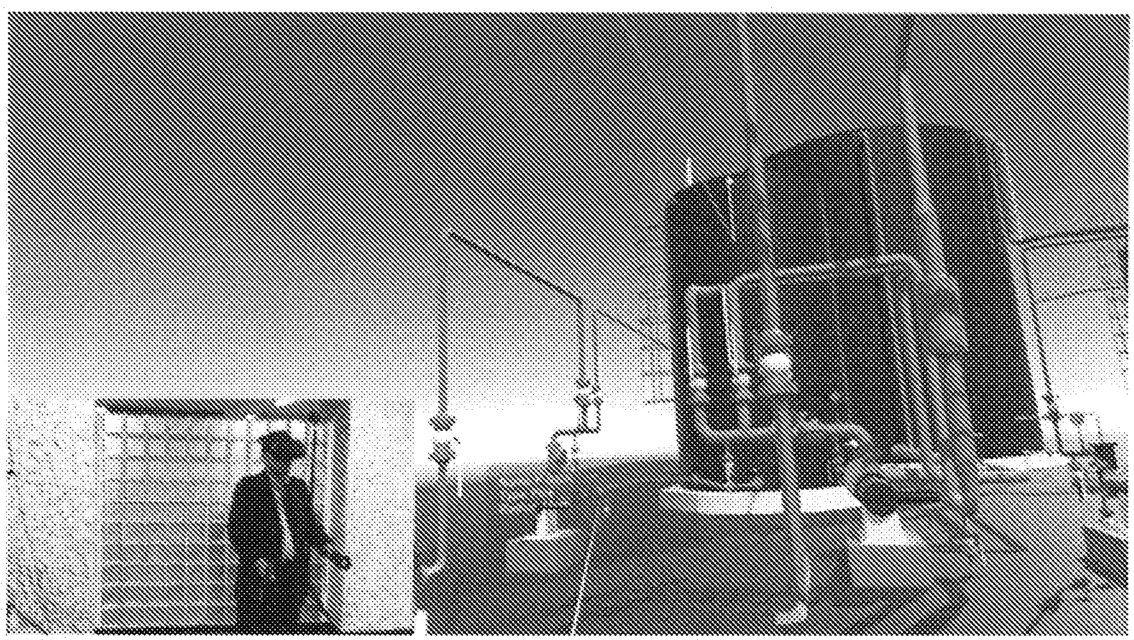


Figure 4. Checking the device information.

In the real digital factory, the staircase is an important model and required to be climbed up and down. In the virtual reality based digital factory, we need to simulate climbing the stairs on the ground. In this section, by taking the stairs as an example, we first introduce how to automatically annotate the semantics of 3D models by using Internet images, then illustrate the implementation of climbing the stairs in the virtual reality.

Taking the stairs as an example, the automatic semantic annotation of 3D models by using Internet images is as follows.

Download the relevant images on the Internet with the keyword "stairs".

For each model in the scene, take photos from different angles.

By using Shape Context (Belongie, Malik, and Puzicha, 2000), calculate the similarity scores between the photos in step (2) and the downloaded images in step (1). By comparing the scores and the predefined threshold, judge whether the model is a staircase or not. When the score is less than the predefined threshold, the model is a staircase.

In addition, deep-neural networks can also be used to obtain high-dimensional features (Krizhevsky, Sutskever, and Hinton, 2012), and improve the accuracy of image matching.

After labelling the stairs, we implement the function of climbing the stairs in the virtual reality. Suppose the staircase is represented

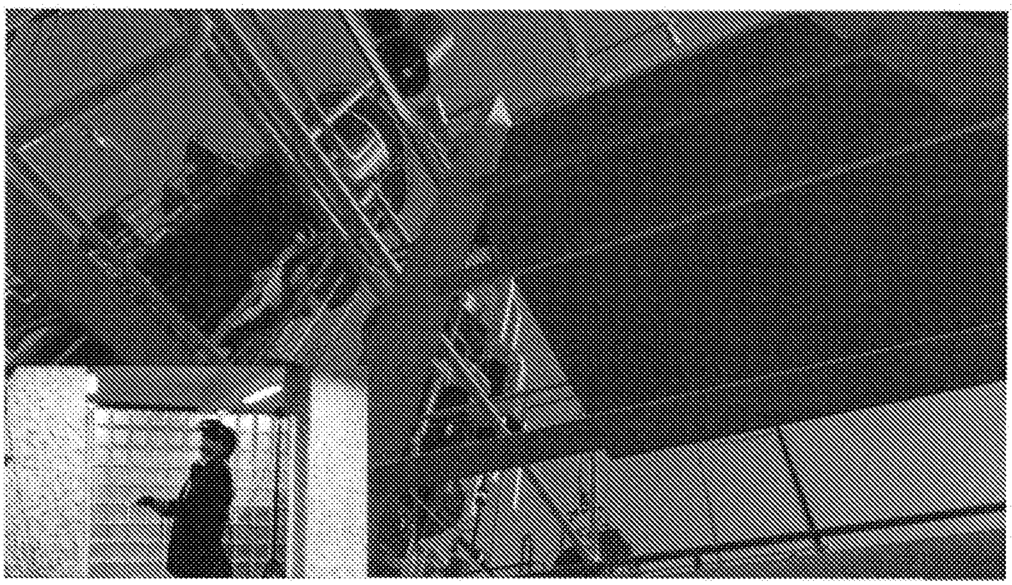


Figure 5. Climbing up and down stairs by the interactive modes of roam.

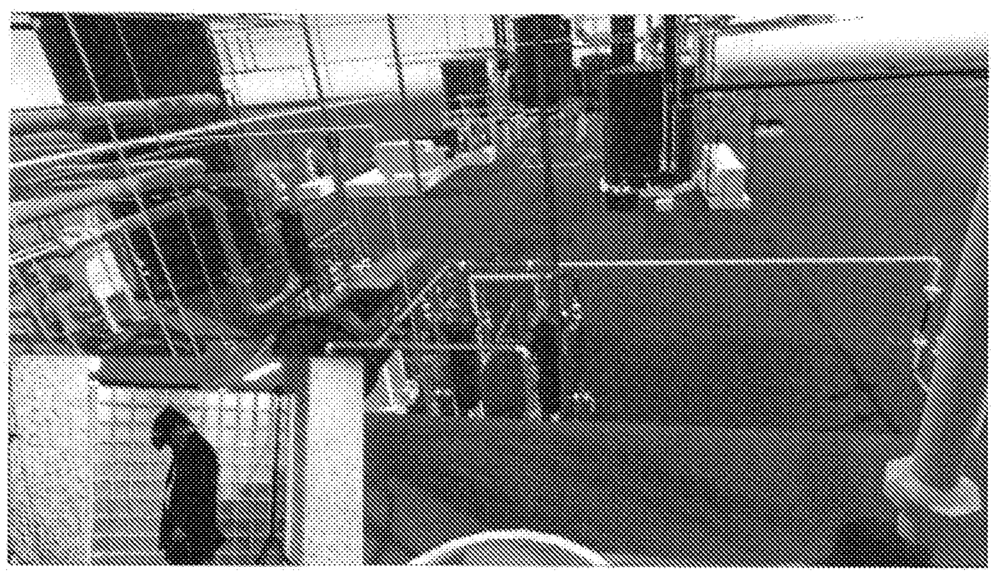


Figure 6. Overlooking the entire factory at the top of the stairs.

as the rectangle $S$ in the three-dimensional space. The algorithm is as follows.

Get the position of the user's head currently in the world coordinate, that is, VR helmet position in the world coordinates.

Get the position of the foot in the world coordinates. Use the height of the VR helmet in the VR local coordinates to get the foot-to-head distance, and then the head position minus this distance to obtain the foot position.

The head and the feet form a line segment $L$. Find whether the rectangle $S$ intersects with the line segment $L$. If yes, obtain the intersection point $P$ between the rectangle $S$ and the line segment $L$ and go to step (4). Otherwise, go to step (5).

If the distance between the lower end of $L$ and $P$ is less than the predefined threshold, it means that the user is already in the stairs and the position of the foot needs to be adjusted to $P$, that is, climbing up the stairs. Then go to step (1).

If $L$ is above $S$, obtain the intersection point $P'$ between the rectangle $S$ and the extension line of segment $L$. Then if the distance between the lower end of $L$ and $P'$ is less than the threshold, then adjust the position of the foot to $P'$, that is, climbing down stairs.

According to the above algorithm, we implement the automatic annotation of stairs and climbing up and down stairs by the interactive modes of roam. The results of climbing up and down stairs by the interactive modes of roam and overlooking the entire factory at the top of the stair are shown in Figure 5 and Figure 6.

## CONCLUSIONS

In this paper, a complete virtual reality based 3D digital factory coastal environment is established. This system vividly displays the environment of the entire factory, and demonstrates the actual state of the plant by connecting to the background controller. Three key issues in this system are discussed. First, the three-dimensional interactive mode in the virtual reality scene is designed by taking roam and device operation as examples. Then, the real-time rendering of large-scale scenes is comprehensively solved from the aspects of model optimization, dynamic preloading and undoing of models. Finally, the automatic semantic annotation of 3D models is realized by means of deep learning algorithm. We implement the system by using Unity 3D and HTC Vive.

The virtual reality based 3D digital factory coastal environment can provide immersive user experience, and display the simulations of various dangerous complex scenes which cannot be presented in a real factory. This system can improve the efficiency of production and training, and further provide effective and practical theories and methods for the application of virtual reality technology to the traditional industries.

## ACKNOWLEDGEMENTS

This work was supported by the National Natural Science Foundation of China (61602033), the Science and Technology Plan Project of Beijing (Z171100001217009), the Social Science Fund of Beijing (16YTC027) and the grant of China Scholarship Council.

## LITERATURE CITED

Belongie, S.; Malik, J., and Puzicha, J., 2000. Shape context: A new descriptor for shape matching and object recognition. *NIPS*, 2, 831-837.

Biocca, F. and Levy, M.R., 2013. *Communication in the Age of Virtual Reality*. Hillsdale, NJ: Lawrence Erlbaum Associates, 21(4), 358-359.

Burdea, G. and Coiffet, P., 2015. Virtual Reality Technology. *Digital Technology and Application*, 95(6), 663-664.

Canetta, L.; Redaelli, C., and Flores M., 2011. Digital factory for human-oriented production systems. *Springer London*, 381 (9876), 1458-1459.

Caggiano, A.; Caiazzo, F., and Teti, R., 2015. Digital factory approach for flexible and efficient manufacturing systems in the aerospace industry. *CIRPe 2015 - Understanding the Life Cycle Implications of Manufacturing*, 37, 122-127.

Chavdar, P.; Tim, K.M., and Michael, J., 2015. Real-time 3D head pose and facial landmark estimation from depth images using triangular surface patch features. *The IEEE Conference on Computer Vision and Pattern Recognition* (CVPR), 4722-4730.

Chen, T.; Cheng, M.; Tan, P.; Shamir, A., and Hu, S., 2009. *Sketch2Photo: Internet Image Montage*. ACM Transactions on Graphics, 28, 124p.

Gao, W. and Wang, F., 2017. New isolated toughness condition for fractional (g,f,n)-critical graphs. *Colloquium Mathematicum*, 147(1), 55-66.

Gleba, R.; Grulkowski, S., and Zariczny, J., 2017. Effect of different structure type traffic on railway line capacity. *Polish Maritime Research*, 24(1), 82-88.

Klaycham, K.; Athisakul, C., and Chucheepsakul, S., 2017. Nonlinear vibration of marine riser with large displacement. *Journal of Marine Science and Technology*, 22(2), 361-375.

Krizhevsky, A.; Sutskever, I., and Hinton, G.E., 2012. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems* 25, NIPS.

Mi, C.; Shen, Y.; Mi, W., and Huang, Y., 2015. Ship identification algorithm based on 3D point cloud for automated ship loader. *In*: Mi, W.; Lee, L.H.; Hirasawa, K., and Li, W. (eds.), *Recent Developments on Port and Ocean Engineering. Journal of Coastal Research*, Special Issue No. 73, pp. 28-34.

Onosato, M. and Iwata, K., 1993. Development of a virtual manufacturing system by integrating product models and factory models. *CIRP Annals* , 42(1), 475-478.

Wang, X.; Jia, J.; Liao, H., and Cai, L., 2012. Affective image colorization. *Journal of Computer Science and Technology*, 27, 1119-1128.

Yang, A.; Han, Y.; Li, S.; Xing, H.; Pan, Y., and Liu, W., 2017. Synthesis and comparison of photocatalytic properties for $Bi_2WO_6$ nanofibers and hierarchical microspheres. *Journal of Alloys and Compounds*, 695, 915-921.

Yin, C. and He, R., 2013. Applications of semiotics in interaction design of VR Gesture. *Packaging Engineering*, 34, 13-15.

Zhang, Y.P. and Jun, H.E., 2005. Applied condition of virtual reality and its application. *Information Technology*, 12, 94-95.

Zhou, Z.; Feng, Y., and Wu, Y., 2011. Virtual reality based process simulation system in refinery. *Computer Engineering and Applications*, 47, 204-208.

Żywicki, K.; Zawadzki, P., and Górski, F., 2017. Virtual reality production training system in the scope of intelligent factory. *International Conference on Intelligent Systems in Production Engineering and Maintenance*, 450-458.

UNITED STATES
PATENT AND TRADEMARK OFFICE

P.O. Box 1450
Alexandria, VA 22313 - 1450
www.uspto.gov

# ELECTRONIC ACKNOWLEDGEMENT RECEIPT

| APPLICATION # | RECEIPT DATE / TIME | ATTORNEY DOCKET # |
|---|---|---|
| **17/609,911** | **07/19/2023 12:47:54 AM ET** | **28108-125** |

## Title of Invention

METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

## Application Information

| | | | |
|---|---|---|---|
| APPLICATION TYPE | Utility - U.S. National Stage under 35 USC 371 | PATENT # | - |
| CONFIRMATION # | 1651 | FILED BY | John Wasserbauer |
| PATENT CENTER # | 62464825 | FILING DATE | 11/09/2021 |
| CUSTOMER # | 36601 | FIRST NAMED INVENTOR | Lucinda Lewis |
| INTL. APPLICATION # | - | INTL. FILING DATE | - |
| CORRESPONDENCE ADDRESS | - | AUTHORIZED BY | Nicholas Blanton |

## Documents

**TOTAL DOCUMENTS: 5**

| DOCUMENT | PAGES | DESCRIPTION | SIZE (KB) |
|---|---|---|---|
| P1409US01_IDS.pdf | 4 | Information Disclosure Statement (IDS) Form (SB08) | 581 KB |

Warning: This is not a USPTO supplied IDS fillable form. Data in the form cannot be automatically loaded to other USPTO systems.

| DOCUMENT | PAGES | DESCRIPTION | SIZE (KB) |
|---|---|---|---|
| Gaikai isn't concerned about OnLive's fundamental patent on cloud gaming VentureBeat.pdf | 6 | Non Patent Literature | 631 KB |
| Gordo_Beyond_Instance- | 10 | Non Patent Literature | 3450 KB |

| | | | |
|---|---|---|---|
| Level_Image_CVPR_2017_paper.pdf | | | |
| ALTERNATE_FUTURES_Despite_cost.PDF | 3 | Non Patent Literature | 1164 KB |
| Virtual_Reality_of_3D_Digital.PDF | 6 | Non Patent Literature | 1415 KB |

## Digest

| DOCUMENT | MESSAGE DIGEST(SHA-512) |
|---|---|
| P1409US01_IDS.pdf | 11538F1220A3B27B361A837D28FE1C2AD9CDDD1AB118B0F957C72CCFFDBCB20D28282859047E5598C41B1C694A97A40BF31A1FD10D3D03A8037769890B72FA7E |
| Gaikai isn't concerned about OnLive's fundamental patent on cloud gaming VentureBeat.pdf | FE4C5AACB9A893259DF5FCE37C8E78D05840495B0B76ABA54BCC8C96C16AAC608E51FB657C75DC94A9D40ED1444310F07374EED8AAD40F2178980521CD9110D9 |
| Gordo_Beyond_Instance-Level_Image_CVPR_2017_paper.pdf | A2CC73B0F9977B79971FA220169FF723CB8744E8C05FD46AED92B928422C60E77630E6579B2C723D0CF4CE2992EACB7CF759A8AE0AF8CF916E5E0A5307043ABB |
| ALTERNATE_FUTURES_Despite_cost.PDF | FF19AE0051CEB9E3B7C1395ABDEC399DE74EFF0767FEC5603B2CFFE7C7D94D46F82D1522B7253D3F99D354FF4F0FCDD3873FDC1393ABA7B7A1759B7B9C28166F |
| Virtual_Reality_of_3D_Digital.PDF | 065C37D386CE1427CE49B4BB7529AB0A1CE37BBE392BB4AE3CFE3E16669787E6BC6E573C38A4BBF88503685C72DFE5CEEBF176C94D3A2FE773C85F86CA395181 |

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

**New Applications Under 35 U.S.C. 111**

If a new application is being filed and the application includes the necessary components for filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application

**National Stage of an International Application under 35 U.S.C. 371**

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

**New International Application Filed with the USPTO as a Receiving Office**

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES
PATENT AND TRADEMARK OFFICE

P.O. Box 1450
Alexandria, VA 22313 - 1450
www.uspto.gov

# ELECTRONIC PAYMENT RECEIPT

| APPLICATION # | RECEIPT DATE / TIME | ATTORNEY DOCKET # |
|---|---|---|
| **17/609,911** | **07/19/2023 12:47:54 AM ET** | **28108-125** |

## Title of Invention

METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

## Application Information

| | | | |
|---|---|---|---|
| APPLICATION TYPE | Utility - U.S. National Stage under 35 USC 371 | PATENT # | - |
| CONFIRMATION # | 1651 | FILED BY | John Wasserbauer |
| PATENT CENTER # | 62464825 | AUTHORIZED BY | Nicholas Blanton |
| CUSTOMER # | 36601 | FILING DATE | 11/09/2021 |
| INTL. APPLICATION # | - | INTL. FILING DATE | - |
| CORRESPONDENCE ADDRESS | - | FIRST NAMED INVENTOR | Lucinda Lewis |

## Payment Information

| PAYMENT METHOD | PAYMENT TRANSACTION ID | PAYMENT AUTHORIZED BY |
|---|---|---|
| **CARD / 4007** | **E20237I149148400** | **John Wasserbauer** |

| FEE CODE | DESCRIPTION | ITEM PRICE($) | QUANTITY | ITEM TOTAL($) |
|---|---|---|---|---|
| 3806 | SUBMISSION OF AN INFORMATION DISCLOSURE STATEMENT | 52.00 | 1 | 52.00 |
| | | | **TOTAL AMOUNT:** | **$52.00** |

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

**New Applications Under 35 U.S.C. 111**

If a new application is being filed and the application includes the necessary components for filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement

Receipt will establish the filing date of the application

**National Stage of an International Application under 35 U.S.C. 371**

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

**New International Application Filed with the USPTO as a Receiving Office**

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

# Beyond instance-level image retrieval: Leveraging captions to learn a global visual representation for semantic retrieval

Albert Gordo and Diane Larlus
Computer Vision group, Xerox Research Center Europe
firstname.name@xrce.xerox.com

## Abstract

*Querying with an example image is a simple and intuitive interface to retrieve information from a visual database. Most of the research in image retrieval has focused on the task of instance-level image retrieval, where the goal is to retrieve images that contain the same object instance as the query image. In this work we move beyond instance-level retrieval and consider the task of semantic image retrieval in complex scenes, where the goal is to retrieve images that share the same semantics as the query image. We show that, despite its subjective nature, the task of semantically ranking visual scenes is consistently implemented across a pool of human annotators. We also show that a similarity based on human-annotated region-level captions is highly correlated with the human ranking and constitutes a good computable surrogate. Following this observation, we learn a visual embedding of the images where the similarity in the visual space is correlated with their semantic similarity surrogate. We further extend our model to learn a joint embedding of visual and textual cues that allows one to query the database using a text modifier in addition to the query image, adapting the results to the modifier. Finally, our model can ground the ranking decisions by showing regions that contributed the most to the similarity between pairs of images, providing a visual explanation of the similarity.*

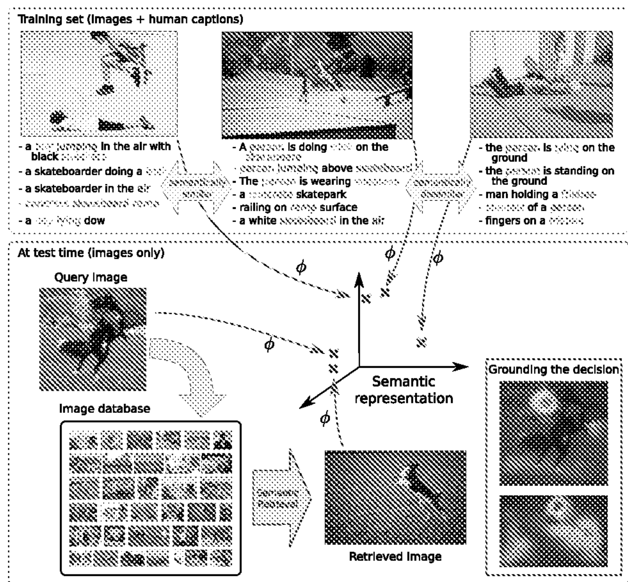



Figure 1. We tackle the **semantic retrieval** task. Leveraging the multiple human captions that are available for images of a training set, we train a semantic-aware representation that improves semantic visual search within a disjoint database of images that do not contain textual annotations. As a by-product, our method highlights regions that contributed the most to the decision.

## 1. Introduction

The task of image retrieval aims at, given a query image, retrieving all images relevant to that query within a potentially very large database of images. This topic has been heavily studied over the years. Initially tackled with bag-of-features representations, large vocabularies, and inverted files [illegible, illegible], and then with feature encodings such as the Fisher vector or the VLAD descriptors [illegible, illegible], the retrieval task has recently benefited from the success of deep learning representations such as convolutional neural networks that were shown to be both effective and computationally efficient for this task [illegible, illegible, illegible]. Among previous retrieval methods, many have focused on retrieving the exact same instance as in the query image, such as a particular landmark [illegible, illegible, illegible] or a particular object [illegible]. Another group of methods have concentrated on retrieving semantically-related images, where "semantically related" is understood as displaying the same object category [illegible, illegible], or sharing a set of tags [illegible, illegible]. This requires to make the strong assumption that all categories or tags are known in advance, which does not hold for complex scenes.

In this paper we are interested in applying the task of *semantic retrieval* to query images that display realistic and complex scenes, where we cannot assume that all the object categories are known in advance, and where the interaction between objects can be very complex. Our first contribution is to *validate that the task of semantic retrieval is well-defined*, particularly in the presence of complex scenes (section 3). Although what different persons understand as a semantically similar scene is subject to interpretation, we show in a user study that there is a high level of consistency between different users.

Following the standard image retrieval paradigm that targets efficient retrieval within databases of potentially millions of images, we aim at learning a global and compact visual representation tailored to the semantic retrieval task that, instead of relying on a predefined list of categories or interactions, implicitly captures information about the scene objects and their interactions. However, directly acquiring enough semantic annotations from humans to train such a model may not be feasible. Our second contribution is to show that *a similarity function based on captions produced by human annotators, which we assume are available at training time, constitutes a good computable surrogate of the true semantic similarity*, and provides enough information to learn a semantic visual representation (section 4).

Our third contribution is a model that leverages the similarity between human-generated captions, *i.e.* privileged information available only at training time, to *learn how to embed images in a semantic space, where the similarity between embedded images is related to their semantic similarity* (section 5.1). Our experiments first show that learning a semantic representation significantly improves over a model pretrained on ImageNet. We also show that it can provide a visual explanation of the semantic similarity by highlighting regions that contributed the most to it.

Our last contribution (section 5.2) is an extension of the previous model that *leverages the image captions explicitly and learns a joint embedding for the visual and textual representations*. We show that this further improves the accuracy of the model, and, more importantly, this allows one to add text modifiers to the query in order to refine the query or to adapt the results towards additional concepts.

## 2. Related work

**Image retrieval.** Image retrieval has been mostly tackled as the problem of instance-level image retrieval [illegible], that focuses on the retrieval of the exact same instance as defined in standard benchmark datasets [illegible]. Moving away from instances, some works have tackled visual search as the retrieval of images that share the same category label [illegible] or a set of tags [illegible]. These works still have a crude understanding of the semantics of a scene. On their synthetic dataset of abstract scenes, Zitnick and Parikh have shown that image retrieval can be greatly improved when detailed semantics is available [illegible]. Explicit modeling of a scene can be done with attributes [illegible], object co-occurrences [illegible], or pairwise relationship between objects [illegible]. As the interaction between objects in a scene can be highly complex, going beyond simple pairwise relations, one extreme interface proposed by Johnson *et al.* [illegible] is to compare explicit scene graph representations instead of visual representations. One shortcoming of their method is that it requires the user to query with a full scene graph, which is a tedious process. We believe that querying with an image is a more intuitive interface.

A number of approaches have cast the task of image captioning as a retrieval problem, first retrieving similar images, and then transferring caption annotations from the retrieved images to the query image [illegible]. Yet these methods use features that are not trained for the task, either simple global features [illegible], features pretrained on ImageNet [illegible] or complex features relying on object detectors, scene classifiers, etc. [illegible]. We believe that the representation should be free of assumptions about the list of objects, attributes, and interactions one might encounter in the scene, and therefore, we learn these representations directly from the training data.

**Joint embeddings of image and text.** Many tasks require to jointly leverage images and natural text, such as zero-shot learning [illegible], language generation [illegible], multimedia retrieval [illegible], image captioning [illegible], and VQA [illegible]. A common solution is to build a joint embedding for textual and visual cues and to compare the modalities directly in that space. The first category of methods for joint embedding is based on CCA [illegible]. Recent methods using CCA include [illegible] and [illegible], a deep extensions of CCA. As an alternative to CCA, previous work has learned the joint embedding with a ranking loss. Among them, WSABIE [illegible] and DeViSE [illegible] learn a linear transformation of visual and textual features with a single-directional ranking loss. Some papers have proposed a bidirectional ranking loss [illegible] possibly with additional constraints [illegible]. Deep methods have also been proposed for this task, based on deep Boltzman machines [illegible], auto-encoders [illegible], LSTMs [illegible], or RNNs [illegible]. These joint image and text embeddings are often used to do cross-modal queries, *i.e.* to retrieve image with textual queries and vice-versa [illegible].

In many of these works learning the joint embedding is, by itself, the end objective. This differs from our work, where the end task is to learn a visual embedding to retrieve images using a query image, and where the joint embedding is used to enrich the visual representation. From that point of view, a connection is also found with the privileged learning framework [illegible]: our improved representation is learned with privileged information in the form of semantic similarity measures provided by the captions that are present at training time. The work of Gomez *et al.* [illegible], in these same proceedings, follows a similar idea, leveraging text from the Wikipedia to learn self-supervised visual embeddings aimed at classification, detection, and retrieval tasks.

## 3. User study

In this section we conduct a user study to acquire annotations related to the semantic similarity between images as perceived by users, and use those to show that the task of semantic retrieval in complex scenes is well-defined and that users tend to agree on their decisions. We also show that visual models pretrained on ImageNet, although better than random, do not reach a high agreement with the users, and that some form of training will be required to achieve good semantic retrieval results using only visual features.

**Dataset.** The computer vision community has made a recent effort in collecting and organizing large-scale datasets allowing for both training and benchmarking of cognitive scene understanding tasks: the MS-COCO dataset [illegible], the VQA dataset [illegible], that adds to MS-COCO a set of question/answer pairs related to the visual content of these images, and, more recently, the Visual Genome dataset [illegible], that is composed of 108k images with a wide range of annotations such as region level-captions, scene graphs, objects, and attributes. This dataset has been designed to evaluate tasks that go beyond image classification and that require to reason about the visual scenes. We adopt the Visual Genome dataset for our experiments, as it is well suited for the task of semantic visual search. We structure it into 80k images for training, 10k for validation, and 10k for test.

**Methodology.** Manually ranking a large set of images according to their semantic relevance to a query image is a very complex, tedious, and time-consuming task. Instead, to ease the task of the annotators, we consider the problem of triplet ranking: given a triplet of images, composed of one query image and two other images, we ask our users to choose the most semantically similar image to the query among the two options. To not bias the annotations towards any interpretation of semantic similarity, we keep the guidelines as open as possible, asking the users to choose, among the two displayed images, the one that "depicted the scene that was most similar to the scene in the query image". The users have the choice to choose one of the two images or to choose that both images were either equally relevant or not relevant to the query.

To construct the triplets we randomly sample query images and then choose two images that are visually similar to the query. This is achieved by extracting image features using ResNet-101 [illegible] pretrained on ImageNet (performing global average pooling after the last convolutional layer) and sampling two images from the 50 nearest neighbors to the query in the visual feature space. The motivation to choose visually similar images is that, in random image triplets, both images will most often be irrelevant to the query. Our study involves 35 annotators (13 women and 22 men), whose annotations spread over 3,000 image triplets. A common set of 50 triplets was answered by 25 users, and most triplets were annotated by at least two annotators. For every triplet we store three values: $o_1$ and $o_2$ encode the number of times the first (resp. second) image was chosen, and $o_3$ the number of times people did not pick any of the two images.

**Inter-user agreement.** We evaluate the agreement between users on this ranking task. We compute a score in a leave-one-user-out fashion, where the decisions of each user are compared against the decisions of all the other users. Given a user and a ranking question, the agreement score $s$ is measured as the proportion of the remaining users that made the same choice as the user, weighted by the proportion of remaining users that made a decision on that triplet, *i.e.*, $s = w\frac{o_i - 1}{o_1 + o_2 - 1}$, with $w = \frac{o_1 + o_2 - 1}{o_1 + o_2 + o_3 - 1}$ and $i \in \{1, 2\}$ is the choice of the user. This score is only computed for triplets where both the user and at least one of the remaining users chose one of the images. The final agreement score for a particular user is the average of the per-triplet agreements. In average, the inter-user agreement score is 89.1, with a standard deviation of 4.6. This shows that people generally agree with each other on the semantic similarity ranking between two images. On the set of 50 images that was annotated by 25 users, we get a similar leave-one-out agreement score of 87.3 $\pm$ 4.5.

**Agreement with visual representations.** We now show that a model pretrained on ImageNet, with no further training, does not achieve a high agreement with the users. We consider an image representation based on the fully-convolutional ResNet-101 architecture [illegible]. Our representation follows the R-MAC [illegible, illegible] architecture, where, after the convolutional layers from [illegible], one performs max-pooling over different grid regions of the image at different scales, normalizes the descriptors of each region independently using PCA with whitening, and finally aggregates and renormalizes the final output to obtain a descriptor of 2048 dimensions. These ResNet R-MAC descriptors can be compared using the dot product.

As in the inter-user agreement case, the agreement between a method and the users is measured as the proportion of users that agree with the ranking decisions produced by the method, weighted by the proportion of users that made a decision on that triplet, averaged through all the triplets with at least one human annotator. Under this setup, our visual baseline, the ResNet with R-MAC, obtains an agreement of 64.0, *cf*. Table 1. This agreement is higher than a random ranking of triplets (50.0 $\pm$ 0.8 over 5 runs), but significantly lower than the inter-user agreement, suggesting that training the visual models is necessary, and that, to that end, semantic annotations will be necessary.

| Method | score |
|---|---|
| Human annotators | 89.1 ± 4.6 |
| Visual baseline: ResNet R-MAC | 64.0 |
| Object annotations | 63.4 |
| Human captions: METEOR | 72.1 |
| Human captions: word2vec + FV | 70.1 |
| Human captions: tf-idf | 76.3 |
| Generated captions: tf-idf | 62.5 |
| Random (x5) | 50.0 ± 0.8 |

Table 1. Top row, inter-human annotation agreement on the image ranking task. Bottom rows: comparison between the semantic ranking provided by human annotators and several visual baselines and methods based on the Visual Genome annotations.

## 4. Proxy measures for semantic similarity

To learn a visual embedding that preserves the semantic similarity between images one would need a large number of annotated image triplets. Unfortunately, requiring human annotators to provide rankings for millions of triplets is not feasible. Instead, we propose to use a surrogate measure. Ideally, this surrogate measure should be efficient to compute and be highly correlated with the ranking given by the human annotators. To this end, we leverage the annotations of the Visual Genome dataset and study which measures yield a high correlation with the human annotators.

Our first representation leverages the objects contained in images. We consider the ground-truth object annotations provided with the Visual Genome dataset [ ], that list all the objects present in one image and, when relevant, their WordNet [ ] synset assignment. We build a histogram representation of each image, counting how many objects of each synset appear in that image, and weight the histograms with a tf-idf mechanism followed by $\ell_2$ normalization. The final representations are compared with the dot product. As seen in Table 1, the agreement of this representation with the users is worse than the visual agreement. This shows that counting objects from a predefined list of categories and neglecting their interactions does not offer a good proxy for semantic similarity, and that more information is needed.

Motivated by this, we consider human captions as a proxy for semantic similarity. Our rationale is that the human annotators will have a bias towards annotating parts of the image that they deem important, and that these annotated parts will be the same that they use to decide if images are semantically similar or not. The Visual Genome dataset contains, on average, 50 region-level captions per image annotated by different users, and this redundancy should further help to capture subtle semantic nuances. Consequently we leverage the provided region-level captions to build several textual representations of the images.

An intuitive way to compare image captions is to use METEOR [ ], a similarity between text sentences typically used in machine translation that has also been used as a standard evaluation measure for image captioning [ ]. To compare two sets of region-level captions $X$ and $Y$ from two images, we perform many-to-many matching with a (non-Mercer [ ]) match kernel of the form

$$K(X,Y) = \frac{1}{|X|+|Y|}\left(\sum_{x\in X}\max_{y\in Y} M(x,y) + \sum_{y\in Y}\max_{x\in X} M(x,y)\right).$$

Note that this requires to evaluate up to thousands of pairs of sentences to compare two images, which may take up to a few seconds for images with more than a hundred captions. Therefore, the scalability of this approach is limited.

To avoid the scalability problem, one option is to merge all the words of all the captions of an image into a single set of words. This sacrifices the structure of the sentences but allows to use other methods based on bags of words. We experiment with two of them. The first one follows [ ] and computes a Fisher vector [ ] (FV) of the word2vec [ ] representations of the captions' words. The semantic similarity between two captioned images is the dot product between the two $\ell_2$-normalized FV representations. The second one is a *tf-idf* weighting of a bag-of-words (BoW) followed by $\ell_2$ normalization, that can also be compared using the dot product. Contrary to the METEOR metric, these two last approaches produce not only a similarity but also a vectorial representation of the text that can potentially be used during training. All learning involved in these representations (vocabulary of 46881 words, idf weights, Gaussian mixture model for the word2vec-based Fisher vector, etc.) is done on our training partition of the Visual Genome dataset.

We compute the agreement score of all these methods by comparing their decision to the users', and report results in Table 1. We observe that the region-level captions provided by human annotators are very good predictors of the semantic similarity between two images, much better than the visual baseline ones. Of these, the tf-idf BoW representation is best, outperforming METEOR and word2vec on this task. Consequently, this is the representation we leverage to train a better visual representation in the next section. As a comparison, we also experimented with automatically-generated captions [ , ] instead of user-generated captions. The score of the automatic captions is significantly lower, highlighting the importance of using human captions for training.

## 5. Learning visual representations

In the previous section we have shown that human generated captions capture the semantic similarity between images. Here we propose to learn a global image representation that preserves this semantic similarity (Section 5.1). We then extend our method to explicitly embed the visual and textual representations jointly (Section 5.2).

### 5.1. Visual embedding

Our underlying visual representation is the ResNet-101 R-MAC network discussed in Section 3. This network is designed for retrieval [illegible] and can be trained in an end-to-end manner [illegible]. Our objective is to learn the optimal weights of the model (the convolutional layers and the projections in the R-MAC pipeline) that preserve the semantic similarity. As a proxy of the true semantic similarity we leverage the tf-idf-based BoW representation over the image captions. Given two images with captions we define their proxy similarity as the dot product between their tf-idf representations.

To train our network we propose to minimize the empirical loss of the visual samples over the training data. If $q$ denotes a query image, $d^+$ a semantically similar image to $q$, and $d^-$ a semantically dissimilar image, we define the empirical loss as as $L = \sum_q \sum_{d^+,d^-} L_v(q, d^+, d^-)$, where

$$L_v(q, d^+, d^-) = \frac{1}{2}\max(0, m - \phi_q^T \phi_+ + \phi_q^T \phi_-), \quad (1)$$

$m$ is the margin and $\phi : \mathcal{I} \rightarrow \mathbb{R}^D$ is the function that embeds the image into a vectorial space, *i.e.* the output of our model. We slightly abuse the notation and denote $\phi(q)$, $\phi(d^+)$, and $\phi(d^-)$, as $\phi_q$, $\phi_+$, and $\phi_-$. We optimize this loss with a three-stream network as in [illegible] with stochastic optimization using ADAM [illegible].

To select the semantically similar $d^+$ and dissimilar $d^-$ images we evaluated two approaches. In the first one we directly sample them such as that $s(q, d^+) > s(q, d^-)$, where $s$ is the semantic similarity between two images, computed as the dot product between their tf-idf representations, as above. However, we observed this sampling strategy not to improve the visual representation. We believe this is because this strategy optimizes the whole ranking at once, and in particular tries to produce a correct ranking for images that are all very relevant, and for images that are all irrelevant, simply based on visual information. This is an extremely challenging task that our model was not able to correctly learn. Instead, for the second approach, we adopt a hard separation strategy. Similar to other retrieval works that evaluate retrieval without strict labels (*e.g.* [illegible]), we consider the $k$ nearest neighbors of each query according to the similarity $s$ as relevant, and the remaining images as irrelevant. This significantly simplifies the problem, as now the goal is to separate relevant images from irrelevant ones given a query, instead of producing a global ranking. Despite the hard thresholding, we observe this approach to learn a much better representation. Note that this thresholding is done only at training time, not at testing time. In our experiments we use $k = 32$, although other values of $k$ led to very similar results. To reduce the impact of this thresholding the loss could also be scaled by a weight involving the semantic similarity, similar to the WARP loss [illegible], although we did not explore this option in this work. Finally, note that the human captions are only needed at training time to select image triplets, and are not used at test time.

### 5.2. A joint visual and textual embedding

In the previous formulation, we only used the textual information (*i.e.* the human captions) as a proxy for the semantic similarity in order to build the triplets of images (query, relevant and irrelevant) used in the loss function. In this section, we propose to leverage the text information in an explicit manner during the training process. This is done by building a joint embedding space for both the visual representation and the textual representation. For this we define two new losses that operate over the text representations associated with the images:

$$L_{t1}(q, d^+, d^-) = \frac{1}{2}\max(0, m - \phi_q^T \theta_+ + \phi_q^T \theta_-), \quad (2)$$

$$L_{t2}(q, d^+, d^-) = \frac{1}{2}\max(0, m - \theta_q^T \phi_+ + \theta_q^T \phi_-). \quad (3)$$

As before, $m$ is the margin, $\phi : \mathcal{I} \rightarrow \mathbb{R}^D$ is the visual embedding of the image, and $\theta : \mathcal{T} \rightarrow \mathbb{R}^D$ is the function that embeds the text associated with the image into a vectorial space of the same dimensionality as the visual features. We define the textual embedding as $\theta(t) = \frac{W^T t}{\|W^T t\|_2}$, where $t$ is the $\ell_2$-normalized tf-idf vector and $W$ is a learned matrix that projects $t$ into a space associated with the visual representation.

The goal of these two textual losses is to explicitly guide the visual representation towards the textual one, which we know is more informative. In particular, the loss in Eq. (2) enforces that text representations can be retrieved using the visual representation as a query, implicitly improving the visual representation, while the loss in Eq. (3) ensures that image representations can be retrieved using the textual representation, which is particularly useful if text information is available at query time. All three losses (the visual and the two textual ones) can be learned simultaneously using a siamese network with six streams – three visual streams and three textual streams. Interestingly, by removing the visual loss (Eq. (1)) and keeping only the joint losses (particularly Eq. (2)), one recovers a formulation similar to popular joint embedding methods such as WSABIE [illegible] or DeViSE [illegible]. In our case, however, retaining the visual loss is crucial as we target a query-by-image retrieval task, and removing the visual loss leads to inferior results. We also note that our visual loss shares some similarities with the structure-preserving loss of [illegible], although they tackle the very different task of cross-modality search (*i.e.* sentence-to-image and image-to-sentence retrieval).

## 6. Experiments

This section validates the representations produced by our proposed semantic embeddings on the semantic re-

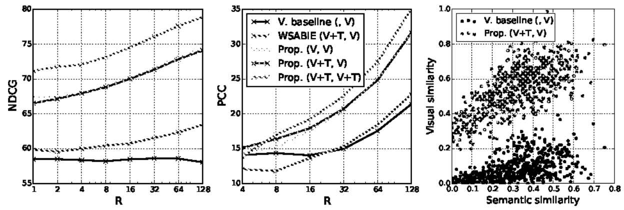



Figure 2. Left and center: NDCG and PCC achieved by the different models as a function of the number of retrieved images R, where the ground truth is determined by the tf-idf similarity. Right: correlation between the ground truth tf-idf similarity and the visual similarity of the baseline and trained models.

| | US | NDCG AUC | PCC AUC |
|---|---|---|---|
| *Text oracle* | | | |
| Caption Tf-idf | 76.3 | 100 | 100 |
| *Query by image* | | | |
| Random (x5) | $50.0 \pm 0.8$ | $10.2 \pm 0.1$ | $-0.2 \pm 0.7$ |
| Visual baseline (, V) | 64.0 | 58.4 | 16.1 |
| WSABIE (V+T, V) | 67.8 | 61.0 | 15.7 |
| Proposed (V, V) | 76.9 | 70.1 | 20.7 |
| Proposed (V+T, V) | 77.2 | 68.8 | 21.1 |
| *Query by image + text* | | | |
| Proposed (V+T, V+T) | 78.6 | 74.4 | 22.5 |

Table 2. Comparison of the proposed methods and baselines evaluated according to User-study (US) agreement score, AUC of the NDCG and PCC curves (*i.e.* NDCG AUC and PCC AUC).

trieval task. We quantitatively evaluate them in two different scenarios. In the first one, we evaluate how well the learned embeddings are able to reproduce the semantic similarity surrogate based on the human captions. In the second scenario, we evaluate our models using the triplet-ranking annotations acquired from users (Section 3), by comparing how well our visual embeddings agree with the human decisions on all these triplets. Then, we propose an experiment that shows which parts of the images led to the matching score. Finally we illustrate how, by leveraging the joint embedding, the results retrieved for a query image can be altered or refined using a text modifier.

## 6.1. Experimental details

**Implementation details.** Our visual model is based on the ResNet-101 architecture [ ] (pretrained on ImageNet) for the convolutional layers followed by the R-MAC pooling, projection, aggregation, and normalization pipeline [ ]. We resize all our images preserving the aspect ratio such as that the largest side is of 576 pixels, and use two scales for the R-MAC pooling. To extract textual features we encode the captions using tf-idf. We stem the words using the Snowball stemmer from NLTK [ ].

Our models are learned with a batch size of 64 triplets (sextuples depending on the setup) using the ADAM optimizer with an initial learning rate of $10^{-5}$, which is reduced to $10^{-6}$ after 8k iterations. To mine triplets for training, we follow a similar approach to [ , ]. We first randomly sample $N = 500$ images. For each of those $N$ samples, we sample 9 relevant images according to the ground truth. This produces a pool of 5000 images, where at least 500 of them have at least 9 relevant images in the pool. Then we extract their features using the current state of the model, and prepare all possible triplets of query image, relevant image, and irrelevant image involving the images in the pool, and where the query is only sampled from the first $N$ images. Finally, the 100 triplets with the largest loss for every query and positive pair are selected as potential candidates to be sampled and used for updating the model. This mining process is repeated after $t = 64$ updates of the model.

**Metrics.** We benchmark our proposed models with two metrics that evaluate how well they correlate with the tf-idf proxy measure, which is the task we optimize for, as well as with the user agreement metric proposed in Section 3. Although the latter corresponds to the exact task that we want to address, the metrics based on the tf-idf similarity provide additional insights about the learning process and allow one to crossvalidate the model parameters. We evaluate our approach using normalized discounted cumulative gain (NDCG) and Pearson's correlation coefficient (PCC). Both measures are designed to evaluate ranking tasks. PCC measures the correlation between ground-truth and predicted ranking scores, while NDCG can be seen as a weighted mean average precision, where every item has a different relevance – in our case, the relevance of one item with respect to the query is the dot product between their tf-idf representations. To evaluate our method in the validation or test splits we choose 1k images from the split, that are used as queries, and use them to rank all the 10k images in the split. The query image is removed from the results. Finally, since we are particularly interested in the top results, we do not report results using the full list of 10k retrieved images. Instead, we report NDCG and PCC after retrieving the top $R$ results, for different values of $R$, and plot the results.

**Methods and baselines.** We evaluate different versions of our embedding. We denote our methods with a tuple of the form ({V, V+T}, {V, V+T}). The first element denotes whether the model was trained using only visual embeddings (V), *cf*. Eq. (1), or joint visual and textual embeddings (V+T), *cf*. Eq. (1)-(3). The second element denotes whether, at test time, one queries only with an image, using its visual embedding (V), or with an image and text, using its joint visual and textual embedding (V+T). In all cases, the database consists only of images represented with visual embeddings, with no textual information.

Our approach is compared to our ResNet-101 R-MAC baseline, pretrained on ImageNet, with no further training, and to a WSABIE-like model, that seeks a joint embedding optimizing the loss in Eq. (2), but does not explicitly optimize the visual retrieval goal of Eq. (1).

## 6.2. Results and discussion

We start by discussing the effect of training in the task of simulating the semantic similarity surrogate function. Figure 2 presents the results using the NDCG@R and PCC@R metrics for different values of R.

Our first observation is that all forms of training improve over the ResNet baseline. Of these, WSABIE is the one that obtains the smallest improvement, as it does not optimize directly the retrieval end goal and only focuses on the joint embedding. All methods that optimize the end goal obtain significantly better accuracies. The second observation is that, when the query consists only of one image, training our model explicitly leveraging the text embeddings – models denoted with (V+T, V) – brings quantitative improvement over (V,V) only on some of the metrics. However, this joint training allows one to query the dataset using both visual and textual information – (V+T, V+T). Using the text to complement the visual information of the query leads to significant improvements.

In Table 2 we evaluate these methods on the human agreement score. For context, we also report the area under the curve (AUC) of the NDCG and PCC curves. As with NDCG and PCC, learning the embeddings brings substantial improvements in the user agreement score. In fact, all of our trained models actually outperform the score of the tf-idf over human captions, which was used as a "teacher" to train our model, following the learning with privileged information terminology. Our model leverages both the visual features as well as the tf-idf similarity during training, and, as such, it is able to exploit the complementary information that they offer. Using text during testing improves agreement with users, and brings considerable improvements in the NDCG and PCC metrics. Additionally, having a joint embedding can be of use even if quantitative results do not improve, for instance for refining the query, see Figure 5.

**Grounding the decisions.** We leverage recent visualization techniques to highlight the regions of a pair of images that contribute the most to their similarity. We follow Grad-CAM [illegible], that displays the aggregated activations of the last convolutional layer weighted with the gradient of the loss for a target class. In our case, instead of using the gradients with respect to a specific class, we use the gradients with respect to the top $k = 5$ dimensions of the final signatures that contributed the most to their similarity. Figure 3 displays pairs of images, where the key regions that most contributed to the similarity are highlighted. Please note how the same image can highlight different regions depending on with which image it has been matched with.

**Qualitative retrieval results.** Figure 4 compares the visual baseline with our trained method (V+T,V), where our method retrieves more semantically meaningful results, such as horses on the beach or newlyweds cutting a wedding cake. Figure 5 shows the effect of text modifiers. The embedding of the query image is combined to the embeddings of textual terms (that can be added or subtracted to the representation) to form a new query with an altered meaning that is able to retrieve different images, and that is only possible thanks to the joint embedding of images and text.

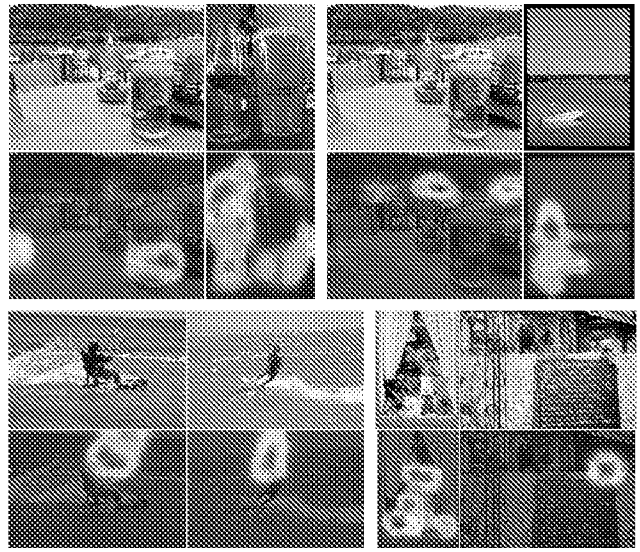


Figure 3. Grounding the decisions. For every pair of images we highlight the parts that contributed the most to their learned visual similarity. Different parts of the same image are highlighted depending on the image it is matched to.

## 7. Conclusions

In this work we focus on the task of semantic image retrieval, where, given a query image, the goal is to retrieve images that depicts similar scenes. To this end we conducted a user study and showed that i) users typically agree on the task of semantically ranking images, and ii) these ranks can be accurately predicted by exploiting human-annotated captions. We leveraged these annotations to learn a visual embedding of the images and showed that this visual embedding predicts very well the human ranking preferences, even better than the human caption proxy we trained with. Our models can also provide visual explanations about why a pair of images is similar. Finally, our joint visual and textual model can leverage text modifiers to refine the meaning of a query image, providing exciting new ways to query image databases.

**Acknowledgments.** We would like to thank Florent Perronnin for fruitful discussions and all of our 35 annotators.


Figure 4. Qualitative results. For every block of images, left: query image. top: top-7 images with the representation pretrained on ImageNet, bottom: top-7 images with our learned representation with the (V+T,V) model.


Figure 5. For a set of query images, we use a text modifier as additional query information (concepts are added or removed) to bias the results. Note that the first query is the last one from Figure 4 refined with additional text.

## References


[1] http://t-satoshi.blogspot.fr/2015/12/image-caption-generation-by-cnn-and-lstm.html. 4

[2] J. Ah-Pine, M. Bressan, S. Clinchant, G. Csurka, Y. Hoppenot, and J. Renders. Crossing textual and visual content in different application scenarios. *Multimedia Tools and Applications*, 42(1):31–56, 2009. 2

[3] J. Ah-Pine, G. Csurka, and S. Clinchant. Unsupervised visual and textual information fusion in cbmir using graph-based methods. *ACM Transactions on Information Systems*, 33(2):9:1–9:31, 2015. 2

[4] Z. Akata, M. Malinowski, M. Fritz, and B. Schiele. Multi-cue zero-shot learning with strong supervision. In *CVPR*, 2016. 2

[5] G. Andrew, R. Arora, J. A. Bilmes, and K. Livescu. Deep canonical correlation analysis. In *ICML*, 2013. 2

[6] S. Antol, A. Agrawal, J. Lu, M. Mitchell, D. Batra, C. L. Zitnick, and D. Parikh. VQA: Visual Question Answering. In *ICCV*, 2015. 2, 3

[7] T. L. Berg, A. C. Berg, and J. Shih. Automatic attribute discovery and characterization from noisy web data. In *ECCV*, 2010. 2

[8] A. Bergamo, L. Torresani, and A. Fitzgibbon. Picodes: Learning a compact code for novel-category recognition. In *NIPS*. 2011. 1, 2

[9] S. Bird, E. Klein, and E. Loper. *Natural Language Processing with Python*. 2009. 6

[10] M. Bucher, S. Herbin, and F. Jurie. Improving semantic embedding consistency by metric learning for zero-shot classiffication. In *ECCV*, 2016. 2

[11] X. Chen, H. Fang, T. Lin, R. Vedantam, S. Gupta, P. Dollár, and C. L. Zitnick. Microsoft COCO captions: Data collection and evaluation server. *CoRR*, abs/1504.00325, 2015. 4

[12] M. J. Choi, J. J. Lim, A. Torralba, and A. S. Willsky. Exploiting hierarchical context on a large database of object categories. In *CVPR*, 2010. 2

[13] M. Denkowski and A. Lavie. Meteor universal: Language specific translation evaluation for any target language. In *EACL Workshop on Statistical Machine Translation*, 2014. 4

[14] C. Desai, D. Ramanan, and C. C. Fowlkes. Discriminative models for multi-class object layout. *IJCV*, 2011. 2

[15] J. Donahue, L. A. Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, K. Saenko, and T. Darrell. Long-term recurrent convolutional networks for visual recognition and description. In *CVPR*, 2015. 2

[16] H. Fang, S. Gupta, F. N. Iandola, R. Srivastava, L. Deng, P. Dollár, J. Gao, X. He, M. Mitchell, J. C. Platt, C. L. Zitnick, and G. Zweig. From captions to visual concepts and back. In *CVPR*, 2015. 2

[17] A. Farhadi, I. Endres, D. Hoiem, and D. Forsyth. Describing objects by their attributes. In *CVPR*, 2009. 2

[18] A. Farhadi, M. Hejrati, M. A. Sadeghi, P. Young, C. Rashtchian, J. Hockenmaier, and D. Forsyth. Every picture tells a story: Generating sentences from images. In *ECCV*, 2010. 2

[19] V. Ferrari and A. Zisserman. Learning visual attributes. In *NIPS*, 2007. 2

[20] A. Frome, G. S. Corrado, J. Shlens, S. Bengio, J. Dean, M. A. Ranzato, and T. Mikolov. DeViSE: A deep visual-semantic embedding model. In *NIPS*, 2013. 2, 5

[21] L. Gomez, Y. Patel, M. Rusiñol, D. Karatzas, and C. V. Jawahar. Self-supervised learning of visual features thorugh embedding images into text topic spaces. In *CVPR*, 2017. 2

[22] Y. Gong, Q. Ke, M. Isard, and S. Lazebnik. A multi-view embedding space for modeling internet images, tags, and their semantics. *IJCV*, 106(2):210–233, 2014. 1, 2

[23] Y. Gong, S. Lazebnik, A. Gordo, and F. Perronnin. Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval. *TPAMI*, 35(12):2916–2929, 2013. 1

[24] Y. Gong, L. Wang, M. Hodosh, J. Hockenmaier, and S. Lazebnik. Improving image-sentence embeddings using large weakly annotated photo collections. In *ECCV*, 2014. 2

[25] A. Gordo, J. Almazán, J. Revaud, and D. Larlus. Deep image retrieval: Learning global representations for image search. In *ECCV*, 2016. 1, 2, 3, 5, 6

[26] D. R. Hardoon, S. Szedmak, and J. Shawe-Taylor. Canonical correlation analysis: An overview with application to learning methods. *Neural computation*, 16(12):2639–2664, 2004. 2

[27] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *CVPR*, 2016. 3, 6

[28] M. Hodosh, P. Young, and J. Hockenmaier. Framing image description as a ranking task: Data, models and evaluation metrics. *JAIR*, 47(1):853–899, May 2013. 2

[29] S. J. Hwang and K. Grauman. Learning the relative importance of objects from tagged images for retrieval and cross-modal search. *IJCV*, 2012. 2

[30] M. Jain, J. van Gemert, T. Mensink, and C. Snoek. Objects2action: Classifying and localizing actions without any video example. In *ICCV*, 2015. 4

[31] H. Jégou and O. Chum. Negative evidences and co-occurences in image retrieval: The benefit of pca and whitening. In *ECCV*. 2012. 1, 2

[32] H. Jégou, M. Douze, and C. Schmid. Hamming embedding and weak geometric consistency for large scale image search. In *ECCV*. 2008. 1, 2

[33] H. Jégou, M. Douze, and C. Schmid. Product quantization for nearest neighbor search. *TPAMI*, 2011. 5

[34] J. Johnson, R. Krishna, M. Stark, L.-J. Li, D. Shamma, M. Bernstein, and L. Fei-Fei. Image retrieval using scene graphs. In *CVPR*, 2015. 2

[35] A. Karpathy and L. Fei-Fei. Deep visual-semantic alignments for generating image descriptions. In *CVPR*, 2014. 2

[36] A. Karpathy, A. Joulin, and L. Fei-Fei. Deep fragment embeddings for bidirectional image sentence mapping. In *NIPS*, 2014. 2

[37] D. P. Kingma and J. L. Ba. Adam: A method for stochastic optimization. In *ICLR*, 2015. 5

[38] R. Kiros, R. Salakhutdinov, and R. S. Zemel. Unifying visual-semantic embeddings with multimodal neural language models. *CoRR*, abs/1411.2539, 2014. 2

[39] B. Klein, G. Lev, G. Sadeh, and L. Wolf. Associating neural word embeddings with deep image representations using fisher vectors. In *CVPR*, 2015. 2

[40] R. Krishna, Y. Zhu, O. Groth, J. Johnson, K. Hata, J. Kravitz, S. Chen, Y. Kalantidis, L.-J. Li, D. A. Shamma, M. Bernstein, and L. Fei-Fei. Visual genome: Connecting language and vision using crowdsourced dense image annotations. *IJCV*, 2016. 3, 4

[41] C. H. Lampert, H. Nickisch, and S. Harmeling. Attribute-based classification for zero-shot visual object categorization. *TPAMI*, 36(3):453–465, 2014. 2

[42] T. Lin, M. Maire, S. J. Belongie, L. D. Bourdev, R. B. Girshick, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick. Microsoft COCO: common objects in context. *CoRR*, abs/1405.0312, 2014. 3

[43] C. Lu, R. Krishna, M. Bernstein, and L. Fei-Fei. Visual relationship detection with language priors. In *ECCV*, 2016. 2

[44] S. Lyu. Mercer kernels for object recognition with local features. In *CVPR*, 2005. 4

[45] M. Malinowski, M. Rohrbach, and M. Fritz. Ask your neurons: A neural-based approach to answering questions about images. In *ICCV*, 2015. 2

[46] J. Mao, W. Xu, Y. Yang, J. Wang, Z. Huang, and A. Yuille. Deep Captioning with Multimodal Recurrent Neural Networks (m-RNN). In *ICLR*, 2015. 2

[47] T. Mensink, E. Gavves, and C. G. M. Snoek. COSTA: Co-Occurrence Statistics for Zero-Shot Classification. In *CVPR*, 2014. 2

[48] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In *NIPS*, 2013. 4

[49] G. A. Miller. Wordnet: A lexical database for english. *Commun. ACM*, 38(11):39–41, 1995. 4

[50] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A. Y. Ng. Multimodal deep learning. In *ICML*, 2011. 2

[51] D. Nister and H. Stewenius. Scalable recognition with a vocabulary tree. In *CVPR*, 2006. 1, 2

[52] V. Ordonez, G. Kulkarni, and T. L. Berg. Im2text: Describing images using 1 million captioned photographs. In *NIPS*, 2011. 2

[53] D. Parikh and K. Grauman. Relative attributes. In *ICCV*, 2011. 2

[54] F. Perronnin and C. Dance. Fisher kernels on visual vocabularies for image categorization. In *CVPR*, 2007. 4

[55] F. Perronnin, Y. Liu, J. Sánchez, and H. Poirier. Large-scale image retrieval with compressed fisher vectors. In *CVPR*, 2010. 1, 2

[56] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Object retrieval with large vocabularies and fast spatial matching. In *CVPR*, 2007. 1, 2

[57] J. Philbin, O. Chum, M. Isard, J. Sivic, and A. Zisserman. Lost in quantization: Improving particular object retrieval in large scale image databases. In *CVPR*, 2008. 1, 2

[58] F. Radenovic, G. Tolias, and O. Chum. CNN image retrieval learns from BoW: Unsupervised fine-tuning with hard examples. In *ECCV*, 2016. 1, 2, 6

[59] M. Ren, R. Kiros, and R. S. Zemel. Exploring models and data for image question answering. In *NIPS*, 2015. 2

[60] R. R. Selvaraju, A. Das, R. Vedantam, M. Cogswell, D. Parikh, and D. Batra. Grad-cam: Why did you say that? visual explanations from deep networks via gradient-based localization. *CoRR*, abs/1610.02391, 2016. 7

[61] J. Sivic and A. Zisserman. Video google: A text retrieval approach to object matching in videos. In *ICCV*, 2003. 1, 2

[62] R. Socher, A. Karpathy, Q. V. Le, C. D. Manning, and A. Y. Ng. Grounded compositional semantics for finding and describing images with sentences. *TACL*, 2:207–218, 2014. 2

[63] N. Srivastava and R. Salakhutdinov. Multimodal learning with deep boltzmann machines. *JMLR*, 15:2949–2980, 2014. 2

[64] G. Tolias, R. Sicre, and H. Jégou. Particular object retrieval with integral max-pooling of cnn activations. In *ICLR*, 2016. 1, 2, 3, 5, 6

[65] L. Torresani, M. Szummer, and A. Fitzgibbon. Learning query-dependent prefilters for scalable image retrieval. In *CVPR*, 2009. 1

[66] V. Vapnik and A. Vashist. A new learning paradigm: Learning using privileged information. In *IJCNN*, 2009. 2

[67] O. Vinyals, A. Toshev, S. Bengio, and D. Erhan. Show and tell: A neural image caption generator. In *CVPR*, 2015. 2, 4

[68] L. Wang, Y. Li, and S. Lazebnik. Learning deep structure-preserving image-text embeddings. In *CVPR*, 2016. 2, 5

[69] J. Weston, S. Bengio, and N. Usunier. WSABIE: Scaling up to large vocabulary image annotation. In *IJCAI*, 2011. 2, 5

[70] C. L. Zitnick and D. Parikh. Bringing semantics into focus using visual abstraction. In *CVPR*, 2013. 2

| **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** ( **Not for submission under 37 CFR 1.99)** | | |
|---|---|---|
| | Application Number | 17609911 |
| | Filing Date | 2023-06-14 |
| | First Named Inventor | Lucinda LEWIS |
| | Art Unit | TBA |
| | Examiner Name | TBA |
| | Attorney Docket Number | P1409US01 |

**U.S.PATENTS**

| Examiner Initial* | Cite No | Patent Number | Kind Code[1] | Issue Date | Name of Patentee or Applicant of cited Document | Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear |
|---|---|---|---|---|---|---|
| | 1 | 10140553 | B1 | 2018-11-27 | Vasisht et al. | All relevant. |
| | 2 | 9892133 | B1 | 2018-02-13 | Biessmann et al. | All relevant. |
| | 3 | 10269179 | B2 | 2019-04-23 | Fein et al. | All relevant. |
| | 4 | 10042604 | B2 | 2018-08-07 | Fallon | All relevant. |
| | 5 | 7849491 | B2 | 2010-12-07 | Perlman | All relevant. |
| | 6 | 10192163 | B2 | 2019-01-29 | Wang | All relevant. |

If you wish to add additional U.S. Patent citation information please click the Add button.

**U.S.PATENT APPLICATION PUBLICATIONS**

| Examiner Initial* | Cite No | Publication Number | Kind Code[1] | Publication Date | Name of Patentee or Applicant of cited Document | Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear |
|---|---|---|---|---|---|---|

# INFORMATION DISCLOSURE STATEMENT BY APPLICANT
( Not for submission under 37 CFR 1.99)

| | |
|---|---|
| Application Number | 17609911 |
| Filing Date | 2023-06-14 |
| First Named Inventor | Lucinda LEWIS |
| Art Unit | TBA |
| Examiner Name | TBA |
| Attorney Docket Number | P1409US01 |

| | | | | | | |
|---|---|---|---|---|---|---|
| | 1 | 20190050520 | A1 | 2019-02-14 | Alvarez et al. | All relevant. |
| | 2 | 20190114395 | A1 | 2019-04-18 | Lenchner et al. | All relevant. |
| | 3 | 20190102954 | A1 | 2019-04-04 | Bastian et al. | All relevant. |

If you wish to add additional U.S. Published Application citation information please click the Add button.

**FOREIGN PATENT DOCUMENTS**

| Examiner Initial* | Cite No | Foreign Document Number3 | Country Code2 i | Kind Code4 | Publication Date | Name of Patentee or Applicant of cited Document | Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear | T5 |
|---|---|---|---|---|---|---|---|---|
| | 1 | | | | | | | ☐ |

If you wish to add additional Foreign Patent Document citation information please click the Add button

**NON-PATENT LITERATURE DOCUMENTS**

| Examiner Initials* | Cite No | Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published. | T5 |
|---|---|---|---|
| | 1 | MADLER, MARK R., "Alternate Futures," San Fernando Valley Business Journal, pp. 9-11, March 18, 2019 . | ☐ |
| | 2 | TAKAHASHI, DEAN, "Gaikai isn't concerned about OnLive's fundamental patent on cloud gaming," VentureBeat, December 15, 2010. | ☐ |
| | 3 | GORDO, ALBERT and LARLUS, DIANE, "Beyond instance-level image retrieval: Leveraging captions to learn a global visual representation for semantic retrieval." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 6589-6598. 2017. | ☐ |

| | | |
|---|---|---|
| **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** ( **Not for submission under 37 CFR 1.99)** | Application Number | 17609911 |
| | Filing Date | 2023-06-14 |
| | First Named Inventor | Lucinda LEWIS |
| | Art Unit | TBA |
| | Examiner Name | TBA |
| | Attorney Docket Number | P1409US01 |

| | | | |
|---|---|---|---|
| | 4 | WANG, XIAOHUI, "Virtual Reality of 3D Digital Factory Based on Coastal Environment," Journal of Coastal Research, vol. 83, pp. 507-512, Coconut Creek, Florida, USA (2018). | ☐ |

If you wish to add additional non-patent literature document citation information please click the Add button

**EXAMINER SIGNATURE**

| | | | |
|---|---|---|---|
| Examiner Signature | | Date Considered | |

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

| | | |
|---|---|---|
| **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** ( **Not for submission under 37 CFR 1.99)** | Application Number | 17609911 |
| | Filing Date | 2023-06-14 |
| | First Named Inventor | Lucinda LEWIS |
| | Art Unit | TBA |
| | Examiner Name | TBA |
| | Attorney Docket Number | P1409US01 |

**CERTIFICATION STATEMENT**

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

☐ That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

**OR**

☒ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

☐ See attached certification statement.

☐ The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☒ A certification statement is not submitted herewith.

**SIGNATURE**

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

| | | | |
|---|---|---|---|
| Signature | /Damian Wasserbauer/ | Date (YYYY-MM-DD) | 2023-07-18 |
| Name/Print | Damian Wasserbauer | Registration Number | 34749 |

VentureBeat
News About Tech, Money and Innovation

# Gaikai isn't concerned about OnLive's fundamental patent on cloud gaming

December 15, 2010 | Dean Takahashi
2 Comments
Share


OnLive announced yesterday that it had received a fundamental patent on cloud gaming, where users can play high-end games on low-end computers by utilizing the computing power of broadband-connected data centers. But one of its chief rivals, Gaikai, said it isn't concerned about the patent because it is operating in a different way.

Both companies can stream game imagery from a data center to a gamer's computer. Most of the heavy-duty computing is done in the servers in the data center and only the compressed images are sent to the user's computer. That's pretty much the definition of cloud gaming, where the user taps the power of the internet "cloud." But David Perry, chief executive of Gaikai in Aliso Viejo, Calif., said in a statement, "We share OnLive's vision that streamed gaming is a key element of the future of the video game industry. We do not expect the general concept of remote gaming to be patentable, as many of us played remote games in the 70's, 80's and 90's. Neither Gaikai nor OnLive were the first to develop technology in this area."

Perry, a well-known game developer, has been working in the industry for a long time. Steve Perlman, chief executive of OnLive, said earlier this week that he filed for his patent back in 2002 and it wasn't granted until last week — almost eight years after the filing. The patent covers technology where video games run on remote servers in data centers and users play via a broadband-connected device such as a TV, PC, Mac or mobile device.

But Perry believes his own company isn't encumbered by OnLive's patent.

"Gaikai has filed a number of patents regarding cloud-based gaming," he said. "We have also been careful to avoid technology where we think that other companies may develop valid patent rights."

He added, “With regard to OnLive's new patent, we are not concerned with making set-top boxes, which is the focus of OnLive's patent, because from the beginning we decided to go frictionless and not require a specific hardware configuration. As a consequence, you are witnessing the evolution of two companies with notably different business models. “

While OnLive is selling games via digital distribution and is offering a subscription service, Gaikai is focused on doing demos and trial version of games which can be embedded on any web site.

“Nearly everything is different in the ways that we approach the consumer marketplace,” Perry said.

**Update:** Jules Urbach, chief executive of Los Angeles-based Otoy, said in an email, “We respect the valid intellectual property rights of others but we'd be surprised to see a valid patent issue today that would preempt the entire field of server-side rendered gaming. This kind of technology has been around and well known since at least the 90's.”

Rob Enderle, an analyst at the Enderle Group, said that playing games on servers does indeed date back to the age of mainframe computers. But he notes that the underlying technology and the experience of server-based games have changed a great deal. Enderle thinks the OnLive patent is defensive in nature. He believes that Gaikai is doing something different for now but eventually both sides will have to work out their intellectual property issues.

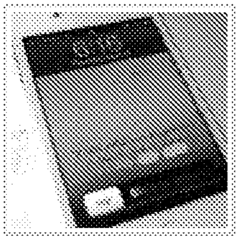
“This does represent the future of gaming and, I expect, before we get there, a lot of attorneys will be vastly wealthier as ownership rights get worked out,” Enderle said.

Next Story: Bing exec explains how social can win the search race
Previous Story: Google launches easy Chrome installer for enterprises — another challenge to Microsoft

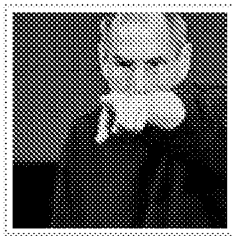


Tags: cloud gaming, patent

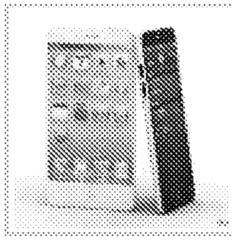
## You might like:



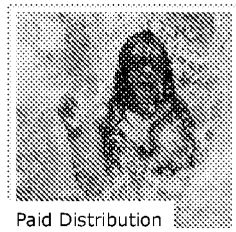
Apple testing out T-Mobile iPhone — oh, and it has an A5 chip
(VentureBeat)



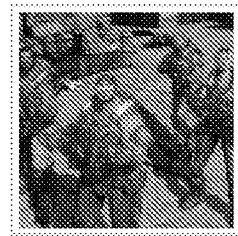
Apple declares war on Adobe, as Lion disables Flash, Photoshop features
(VentureBeat)



New iPhone 5 images based on leaked design specs
(VentureBeat)




Weather Warning: Searing Heat Wave May Hit East Coast in Late August, Scientists Predict
(FoxNews.com)



Hands-on preview of Gears of War 3: expect an epic story with closure
(VentureBeat)

[?]

Like

**Showing 2 comments**

Sort by oldest first

**Perfect_Code**

OnLive = gaming platform

Gaikai = advertising platform

All you need to know.

8 months ago Like Reply

**Robert Shaver**

"he filed for his patent back in 2002 and it wasn't granted until last week — almost eight years after the filing". But I read that the application wasn't actually complete until April of 2003 after which it was rejected five times. So the patent examiners have been burdened with reviewing this thing no less that six times. It was rejected the first five go-rounds. No wonder they're backed up.

But what is this "break-through"? Sounds like they (re)invented the dumb terminal. Is that a novel concept or an old idea with new buzz-words attached?

8 months ago Like Reply

M Subscribe by email S RSS

**Add New Comment** Login

Type your comment here.

**Reactions**

Show more reactions



**MobileBeat**

How Funzio's Crime City shot to the top of the App Store

Motorola's rugged Defy+ is great for clumsy Android fans

Back to school: Amazon releases mobile app for students

**GreenBeat**

Record Hill Wind snags $102M wind power loan guarantee

Google+ represents a big chance for diversification in games

Has cleantech investing lost its glamour? Khosla Ventures loses two partners

**GamesBeat**

How Funzio's Crime City shot to the top of the App Store

Gameforge hits 300M worldwide users for online games

Will Facebook games become faster and richer?

**SocialBeat**

Buddy Media seals $54M funding deal for social media marketing

Will Facebook games become faster and richer?

Blue Mars Lite takes 3D avatars to Google Street View

**Entrepreneur Corner**

VC Roundtable: What makes a great founder?

Entrepreneur Corner: Marketing with Google+ and employee/customer altercations

Infographic: Charting Larry Ellison's Silicon Valley offspring

**MediaBeat**

RGM Group continues to score with ad business aimed at the rich

HBO follows up streaming success with 'MAX Go' app

Netflix now testing awesomely easy interface targeted at kids

**DEMOBeat**

News360 quiets news noise, tailors content to you

DEMO Fall 2011: Opportunities, innovation, inspiration

Zugara licenses Webcam Social Shopper to Banana Flame

**Deals&More**

Google ups its Android patent defense with Motorola purchase

Buddy Media seals $54M funding deal for social media marketing

Google buys Motorola Mobility for $12.5B, says Android will stay open

**DevBeat**

Designers: Create a website, no code required, with Adobe Muse

Will Facebook games become faster and richer?

Hands-on preview of Gears of War 3: an epic story with closure

### In The News

## Companies

GoogleAppleFacebookMicrosoftZyngaTwitterSamsungElectronic ArtsGrouponLinkedInNetflixAT&THTCAmazonNintendo

## People

Steve JobsJohn RiccitielloSteve Perlmanjohn morrisPeter YaredMichael GartenbergSean RyanDavid DrummondYuri MilnerElon MuskBrian CheskyFlorian MuellerWibe Wagemanstim draperMike Morhaime

## Topics

AndroidsmartphonesGoogle PlusiPadsocial gamesiPhonetabletsiOSsecuritysocial networkingpatentsIPOstock marketElectric carelectric vehicle

**VentureBeat Partners**

Powered by WordPress.com VIP

Home

About

Contact

Advertise

Events


Captivated: Customer tries Voyager chair from Positron in Glendale.

# ALTERNATE FUTURES

*Despite costly headsets and motion sickness, Valley companies involved in virtual reality see a market opportunity ahead for enhanced entertainment.*

By **MARK R. MADLER** *Staff Reporter*

Two chairs in a Glendale building can transport a person into the world of virtual reality.

Slip on a headset and a pair of headphones and the red-cushioned chair begins to lean back. Then on the headset viewer comes "How to Train Your Dragon: The Hidden World," and the viewer is introduced to characters Hiccup and Astrid and the dragons, Bonkling, Toothless and Stormfly. The Voyager chair adjusts itself as it tilts back and forward, and rotates side to side, with motion matched to the VR content.

The 4-minute "Dragon" film was made possible by **Walmart Inc.**, which is taking it on a tour in trailers from California to Arkansas that started last month and ends in April. The free film is a marketing tool to get customers into a "How to Train Your Dragon" gift shop.

But for **Jeffrey Travis**, the inventor of the Voyager chair and chief executive of **Positron Inc.**, the Glendale company that developed the device, it is a way to democratize access to the virtual reality format.

An average customer at Walmart does not attend a film festival where VR films are shown, so the trailers

*please see page 10*

*Continued from page 9*

and "Dragon" film are a way to get them to slip on a headset and see it for themselves, Travis said.

"They are able to go up to this trailer from a big film title that they recognize and have this immersive, incredible VR experience they never would have had otherwise," he added. *(An interview with Travis talking about Positron and virtual reality is on page 12.)*

Done in conjunction with DreamWorks Animation, a Glendale-based division of **NBCUniversal** in Universal City, the "Dragon" short was made specifically for use with the Positron chairs.

## Headset issues

Virtual reality companies dot the map of the San Fernando Valley. They run the gamut from hardware developers like Positron to content creators such as **VRWERX**, on the Universal Studios lot, and **Pure Imagination Studios**, in Van Nuys. Elsewhere are location-based entertainment firms such as **Two Bit Circus Corp.**, which has a micro amusement park in downtown Los Angeles that employs virtual reality. The company plans to open another five locations in the next 18 months.

Like digital 3D films from 10 years ago, virtual reality has gone through a cycle of being the next big thing in entertainment to some considering it a big bust.

**Marty Shindler**, of Southern California media consultancy **Shindler Perspective Inc.** was skeptical about why people would want to wear a big headset.

"They don't want to do it," Shindler said. "They haven't done it."

Complaints about the format center on the cost of the headsets, a lack of content and what content there is not being compelling enough to get people to do it more than once. Additionally, there is the fact that some people get eye strain or feel nauseous while watching a VR film.

**Gene Munster**, a virtual reality industry analyst with **Loup Ventures**, in Minneapolis, said the specific problem with VR is all in the hardware; namely, the headsets are expensive, complicated and clumsy.

"Clumsy includes people getting sick," Munster said. "Once the hardware problem is solved, we will see the content piece picking up nicely."

Virtual reality is defined as using computer technologies to create real or imagined worlds seen through a headset. It differs from augmented reality, which is computer-generated imagery or sound placed on top of the real world. Mixed reality is a combination of real and virtual worlds and uses new imagery to interact with the real world.

The history of the format is murky at best. Early examples of a headset were seen in the late 1960s, and for the next 20 years it was limited to training for the military, medical purposes and flight simulation.

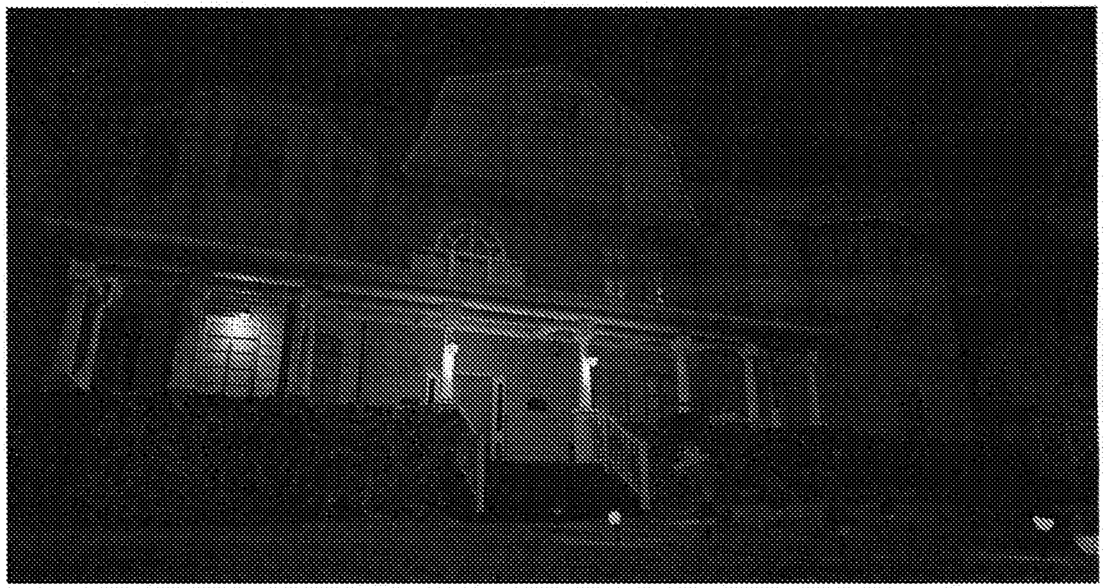
By the early 1990s, **Sega Games Co. Ltd.** had developed a prototype of a headset, Sega VR, but it was never made available to the public. Other companies also worked on headsets.

It wasn't until 2010 and the development of the Oculus Rift prototype that things began to look up for the format. Four years later, **Facebook Inc.** bought Oculus VR for a reported $2 billion.

"History suggests that there will be more platforms to come," Facebook Chief Executive **Mark Zuckerberg** was quoted at the time of the transaction. "Today's acquisition is a long-term bet on the future of computing."

Others began to get into the space. Samsung, **HTC Corp.** and **Sony Corp.** came out with headsets, while **Amazon.com Inc.**, Google, **Apple Inc.** and **Microsoft Corp.** created virtual reality and augmented reality divisions.

**Nancy Bennett**, chief creative officer at Two Bit Circus, said that in the early days of virtual reality, content creators had to build their own cameras and editing systems.



Dark Visions: Scenes from the virtual reality game 'Paranormal Activity: The Lost Soul,' by VRWERX at the Universal Studios lot.

"Now you can get off-the-shelf devices for less than $500," Bennett said.

Bennett was speaking at Digital Entertainment World, an entertainment conference in Marina Del Rey in early February. She appeared on a panel to discuss virtual reality and its market potential.

Another panelist, **Ted Schilowitz**, the futurist for Hollywood studio **Paramount Pictures** who had a similar position with 20th Century Fox, said that when it comes to the "reality" formats, it will be mixed reality that will be the main economic driver while virtual reality will just be the side show.

"It is a better bet to put money on mixed reality than VR," Schilowitz said.

He added that the industry was at least 10 years away from bringing in multi-billion-dollar revenues.

Location-based entertainment centers will be a good starting point. Along with Two Bit Circus, **Dreamscape Immersive**, which counts **AMC Theatres**, **Unibail-Rodamco-Westfield SE** and director **Steven Spielberg** as investors, has a location in Century City showing virtual reality content and will expand to four other cities this year.

## Content creators

On the Universal Studios lot is VRWERX, a virtual reality content creation company headed by **Alex Barder** and **Russell Naftal**.

The company's first offering was a video game, "Paranormal Activity: The Lost Soul." Released in 2017, the game takes players through a 12,000-square-foot haunted house


Poster: 'Paranormal' has a scare engine to alter the action each time the game is played.

where they can pick up objects with virtual hands. The game creates a robust environment in contrast to other early VR games with a single room where the player shoots at zombies or other creatures, Barder said.

Additionally, a proprietary scare engine provides a new experience each time they play.

"For people to replay this, they have to have a different experience each time," Barder said. "We pushed our engineers to create a custom scare engine for us that nobody else was able to do at the time."

The pair are currently in the early stages of creating a VR game based on the "Mission: Impossible" film franchise.

"You'll be doing everything you would expect an agent to do based on the 'Mission: Impossible' universe," Barder said.

*please see page 11*

*Continued from page 10*

"We want people to feel empowered, that they are an actual agent themselves," Naffal added.

Both Barder and Naffal said the virtual reality industry is unlike any other in entertainment. Unlike a film or television show that cannot be changed once it's released, a VR game can be constantly updated. The pair said they are always in contact with players to hear feedback about the "Paranormal Activity" game.

"You don't do what a player says exactly but a lot of times they have great ideas of what they are looking for," Naffal said.

Ultimately what will be a real game changer is the ability of a person wearing a headset to see themselves in a virtual world, Naffal said.

Being able to see yourself in VR is just what **Halsey Minor** has done with his company, **Live Planet**, in San Jose. Minor spent three years building a VR camera system that connects to the cloud to process and store video. The 16-lens camera is unique, he said, in that the person wearing a headset can see themselves in real time. The camera system is currently being offered for $4,950 through the company website.

## Adult entertainment

**Ela Darling** is an adult film actress and also chief marketing officer for **PVR**, a Hong Kong-based company with offices in Los Angeles. PVR has created both a VR camera, the K1 Pro, and a headset called Iris.

The adult industry has long been an early adopter when it comes to new technology and virtual reality is no exception. Darling has been making VR porn for about five years.

The Iris is lightweight, easy to use and, more importantly, discreet because of its unobtrusive look, Darling said.

**Tech Expert: PVR's Ela Darling touts VR.**

"I like to take a headset around to show friends and family and the last thing I want is someone to click over to my VR porn," she added.

The reason adult content was able to harness first videotapes and later DVD and Blu-ray discs was that it had purchasing power in that people were willing to pay for it. With the advent of the internet, an entire generation has grown up expecting adult material for free.

But with virtual reality, Darling foresees a resurgence in a willingness to pay for adult content.

"Customers who would never pay for porn in other mediums are willing to pay for VR because they know that in order for there to be more content, they have to pay for it someway," Darling said. "Otherwise it will not exist."

**VR Bangers**, a Sherman Oaks adult content company, recently updated its Play'a app, which can be used with Oculus Rift and Go, HTC Vive and Samsung Gear headsets or Windows Mixed Reality.

With the Play'a app, a viewer can tailor their virtual reality experience by tilting and scaling the image as well as adjusting the brightness and contrast.

VR Bangers producer **Xander Jones** said that the updated app gives more incentives to the immersion experience of the viewer.

"To us, that is exactly something a professional VR porn producer should think of when trying to provide a premium service," Jones said in a statement.

## Future visions

Shindler, the consultant, is quite skeptical about the market potential for virtual reality.

After all, he pointed out, even IMAX got out of the land-based entertainment VR business starting last year by closing its experience centers, one of which was in Los Angeles.

"After they closed a couple, you knew the rest had to follow," Shindler said.

Virtual reality is following the path of the Gartner Hype Cycle, so-named for information technology research company **Gartner Inc.** in Stamford, Conn., he added.

The cycle starts with an Innovation Trigger and then goes to the Peak of Inflated Expectations where there is a lot of money going toward the technology. Virtual reality experienced that between 2012 and 2016, Shindler said, and then it began to drop off as the tech entered the Trough of Disillusionment. That is the part of the cycle when companies get out of the business as they see it is not going to happen, he added.

Finally, there is the Slope of Enlightenment and the Plateau of Productivity.

"VR may get to the point; it remains to be seen, but not with the headsets that are out there today," Shindler said.

It is outside of the entertainment industry that Shindler thinks virtual reality will make its mark.

"If you are a washing machine repairman and you have a problem, you might be able to dial into a central server and with a VR headset figure out how to do the repair," Shindler said.

Barder and Naffal of VRWERX also believe there are applications for the format outside of video games. Barder recently met a cardiologist who wanted to use VR to look inside the heart and see its condition.

"It's the perfect platform for that kind of thing," Barder said. "And it's way beyond immersive entertainment. VR can save lives."

Munster, the analyst with Loup Ventures, said that while he knows and respects Shindler, he doesn't agree with him about VR. He thinks it will have a more profound impact on society than anyone realizes.

"It is easy to be negative on VR today, but I am still optimistic about what it can do," Munster said.

**Immersive: Live Planet's 16-lens camera.**

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NUMBER | FILING OR 371(C) DATE | FIRST NAMED APPLICANT | ATTY. DOCKET NO./TITLE |
|---|---|---|---|
| 17/609,911 | 11/09/2021 | Lucinda Lewis | 28108-125 |

**CONFIRMATION NO. 1651**

**POA ACCEPTANCE LETTER**

36601
Wasserbauer Law LLC
PO BOX 382
Collinsville, CT 06022

*OC000000058770303*

Date Mailed: 06/28/2023

## NOTICE OF ACCEPTANCE OF POWER OF ATTORNEY

This is in response to the Power of Attorney filed 06/14/2023.

The Power of Attorney in this application is accepted. Correspondence in this application will be mailed to the above address as provided by 37 CFR 1.33.

Questions about the contents of this notice and the requirements it sets forth should be directed to the Office of Data Management, Application Assistance Unit, at **(571) 272-4000** or **(571) 272-4200** or **1-888-786-0101**.

/ytdemisse/

PTO/AIA/80 (07-17)
Approved for use through 03/31/2021. OMB 0651-0035
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no person is required to respond to a collection of information unless it displays a valid OMB control number

# POWER OF ATTORNEY TO PROSECUTE APPLICATIONS BEFORE THE USPTO

I hereby revoke all previous powers of attorney given in the application identified in the attached statement under 37 CFR 3.73(c).

I hereby appoint:

[X] Practitioners associated with Customer Number: 36601

**OR**

[ ] Practitioner(s) named below (if more than ten patent practitioners are to be named, then a customer number must be used):

| Name | Registration Number | Name | Registration Number |
|---|---|---|---|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

As attorney(s) or agent(s) to represent the undersigned before the United States Patent and Trademark Office (USPTO) in connection with any and all patent applications assigned only to the undersigned according to the USPTO assignment records or assignment documents attached to this form in accordance with 37 CFR 3.73(c).

Please change the correspondence address for the application identified in the attached statement under 37 CFR 3.73(c) to:

[X] The address associated with Customer Number: 36601

**OR**

[ ]

| Firm or individual name | | |
|---|---|---|
| Address | | |
| City | State | Zip |
| Country | | |
| Telephone | Email | |

Assignee name and address:

**A copy of this form, together with a statement under 37 CFR 3.73(c) (Form PTO/AIA/96 or equivalent) is required to be filed in each application in which this form is used. The statement under 37 CFR 3.73(c) may be completed by one of the practitioners appointed in this form, and must identify the application in which this Power of Attorney is to be filed.**

**SIGNATURE of Assignee of Record**

The individual whose signature and title is supplied below is authorized to act on behalf of the assignee.

| | |
|---|---|
| Signature /Lucinda Lewis/ | Date 06/13/2023 |
| Name Lucinda LEWIS | Telephone ------------------ |
| Title Inventor / Officer of Automobilia II, LLC | |

UNITED STATES
PATENT AND TRADEMARK OFFICE

P.O. Box 1450
Alexandria, VA 22313 - 1450
www.uspto.gov

# ELECTRONIC ACKNOWLEDGEMENT RECEIPT

| APPLICATION # | RECEIPT DATE / TIME | ATTORNEY DOCKET # |
|---|---|---|
| **17/609,911** | **06/14/2023 02:20:50 PM ET** | **28108-125** |

## Title of Invention

METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

## Application Information

| | | | |
|---|---|---|---|
| APPLICATION TYPE | Utility - U.S. National Stage under 35 USC 371 | PATENT # | - |
| CONFIRMATION # | 1651 | FILED BY | Nicholas Blanton |
| PATENT CENTER # | 62269717 | FILING DATE | 11/09/2021 |
| CUSTOMER # | – | FIRST NAMED INVENTOR | Lucinda Lewis |
| INTL. APPLICATION # | - | INTL. FILING DATE | - |
| CORRESPONDENCE ADDRESS | Lucinda Lewis, -<br>Automobilia II, LLC, 1406 Columbus Boulevard<br>Coral Gables, FLORIDA 33134-2351<br>UNITED STATES | AUTHORIZED BY | - |

## Documents

## TOTAL DOCUMENTS: 1

| DOCUMENT | PAGES | DESCRIPTION | SIZE (KB) |
|---|---|---|---|
| aia0080_signed_F.pdf | 1 | Power of Attorney | 99 KB |

## Digest

| DOCUMENT | MESSAGE DIGEST(SHA-512) |
|---|---|
| aia0080_signed_F.pdf | 4AE87FB5818C3C089099E67D8D0647D6BB7FA9F7BC0B1B6F D9E38390F6207C2FC5F4C89856197716C59695A4834F90C179 |

2F5667C021DB4A7F4E23E72F9E3F24

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

**New Applications Under 35 U.S.C. 111**

If a new application is being filed and the application includes the necessary components for filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application

**National Stage of an International Application under 35 U.S.C. 371**

If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

**New International Application Filed with the USPTO as a Receiving Office**

If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NUMBER | FILING OR 371(C) DATE | FIRST NAMED APPLICANT | ATTY. DOCKET NO./TITLE |
|---|---|---|---|
| 17/609,911 | 11/09/2021 | Lucinda Lewis | 28108-125 |

**CONFIRMATION NO. 1651**

Lucinda Lewis
Automobilia II, LLC
1406 Columbus Boulevard
Coral Gables, FL 33134-2351

**IMPROPER CFR REQUEST**

*OC000000137347021*

Date Mailed: 02/07/2023

## RESPONSE TO REQUEST FOR CORRECTED FILING RECEIPT

***Power of Attorney, Claims, Fees, System Limitations, and Miscellaneous***

In response to your request for a corrected Filing Receipt, the Office is unable to comply with your request because:

- The change of correspondence address submitted on 12/15/2022 has not been entered because it was not properly signed.#

Questions about the contents of this notice and the requirements it sets forth should be directed to the Office of Data Management, Application Assistance Unit, at **(571) 272-4000** or **(571) 272-4200** or **1-888-786-0101**.

/nhassani/

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NUMBER | FILING OR 371(C) DATE | FIRST NAMED APPLICANT | ATTY. DOCKET NO./TITLE |
|---|---|---|---|
| 17/609,911 | 11/09/2021 | Lucinda Lewis | 28108-125 |

**CONFIRMATION NO. 1651**

**PUBLICATION NOTICE**

*OC000000137040098*

Lucinda Lewis
Automobilia II, LLC
1406 Columbus Boulevard
Coral Gables, FL 33134-2351

**Title:**METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

**Publication No.**US-2022-0398827-A1
**Publication Date:**12/15/2022

## NOTICE OF PUBLICATION OF APPLICATION

The above-identified application will be electronically published as a patent application publication pursuant to 37 CFR 1.211, et seq. The patent application publication number and publication date are set forth above.

The publication may be accessed through the USPTO's publically available Searchable Databases via the Internet at www.uspto.gov. The direct link to access the publication is currently http://www.uspto.gov/patft/.

The publication process established by the Office does not provide for mailing a copy of the publication to applicant. A copy of the publication may be obtained from the Office upon payment of the appropriate fee set forth in 37 CFR 1.19(a)(1). Orders for copies of patent application publications are handled by the USPTO's Public Records Division. The Public Records Division can be reached by telephone at (571) 272-3150 or (800) 972-6382, by facsimile at (571) 273-3250, by mail addressed to the United States Patent and Trademark Office, Public Records Division, Alexandria, VA 22313-1450 or via the Internet.

In addition, information on the status of the application, including the mailing date of Office actions and the dates of receipt of correspondence filed in the Office, may also be accessed via the Internet through the Patent Electronic Business Center at www.uspto.gov using the public side of the Patent Application Information and Retrieval (PAIR) system. The direct link to access this status information is currently https://portal.uspto.gov/pair/PublicPair. Prior to publication, such status information is confidential and may only be obtained by applicant using the private side of PAIR.

Further assistance in electronically accessing the publication, or about PAIR, is available by calling the Patent Electronic Business Center at 1-866-217-9197.

Office of Data Managment, Application Assistance Unit (571) 272-4000, or (571) 272-4200, or 1-888-786-0101

RECEIVED
CENTRAL FAX CENTER
DEC 15 2022

PTO/AIA/122 (06-15)
Approved for use through 09/30/2025. OMB 0651-0035
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

# CHANGE OF CORRESPONDENCE ADDRESS
***Application***

Address to:
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

| | |
|---|---|
| Application Number | 17/609,911 |
| Filing Date | 11/09/2021 |
| First Named Inventor | Lucinda Lewis |
| Art Unit | |
| Examiner Name | |
| Attorney Docket Number | |

Please change the Correspondence Address for the above-identified patent application to:

☒ The address associated with Customer Number: 162160

*OR*

☐ Firm or Individual Name Lucinda L Lewis

Address 1825 Ponce De Leon Blvd. #116

City Coral Gables State Florida Zip 33134

Country United States of America

Telephone (818) 633-1176 Email cindy.lewis@me.com

This form cannot be used to change the data associated with a Customer Number. To change the data associated with an existing Customer Number use "Request for Customer Number Data Change" (PTO/SB/124).

I am the:

☒ Applicant

☐ Attorney or agent of record. Registration Number ____________.

☐ Registered practitioner named in the application transmittal papers who acts in a representative capacity under 37 CFR 1.34. See 37 CFR 1.33(a)(1). Registration Number ____________.

Signature /Lucinda Lewis/

Typed or Printed Name Lucinda Lewis

Date 12/15/2022 Telephone (818) 633-1176

NOTE: This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4(d) for signature requirements and certifications. Submit multiple forms if more than one signature is required, see below*.

☐ *Total of ________ forms are submitted.



*If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.*

PTO/AIA/122 (08-12)
Approved for use through 09/30/2025. OMB 0651-0035
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE


# CHANGE OF CORRESPONDENCE ADDRESS
***Application***

Address to:
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

| | |
|---|---|
| Application Number | 17609911 |
| Filing Date | 11/09/2021 |
| First Named Inventor | Lucinda L Lewis |
| Art Unit | |
| Examiner Name | |
| Attorney Docket Number | 28108125 |

Please change the Correspondence Address for the above-identified patent application to:

☒ The address associated with Customer Number: 162160

***OR***

☐ Firm or Individual Name

Address

City | State | Zip

Country

Telephone | Email

This form cannot be used to change the data associated with a Customer Number. To change the data associated with an existing Customer Number use "Request for Customer Number Data Change" (PTO/SB/124).

I am the:

☒ Applicant

☐ Attorney or agent of record. Registration Number ____________________.

☐ Registered practitioner named in the application transmittal papers who acts in a representative capacity under 37 CFR 1.34. See 37 CFR 1.33(a)(1). Registration Number ______________________.

Signature /Lucinda Lewis/

Typed or Printed Name Lucinda L Lewis

Date 12/9/2022 | Telephone 8186331176

NOTE: This form must be signed in accordance with 37 CFR 1.33. See 37 CFR 1.4(d) for signature requirements and certifications. Submit multiple forms if more than one signature is required, see below*.

1 *Total of 1 forms are submitted.

# Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. The United States Patent and Trademark Office (USPTO) collects the information in this record under authority of 35 U.S.C. 2. The USPTO's system of records is used to manage all applicant and owner information including name, citizenship, residence, post office address, and other information with respect to inventors and their legal representatives pertaining to the applicant's/owner's activities in connection with the invention for which a patent is sought or has been granted. The applicable Privacy Act System of Records Notice for the information collected in this form is COMMERCE/PAT-TM-7 Patent Application Files, available in the Federal Register at 78 FR 19243 (March 29, 2013). https://www.govinfo.gov/content/pkg/FR-2013-03-29/pdf/2013-07341.pdf

Routine uses of the information in this record may include disclosure to: 1) law enforcement, in the event that the system of records indicates a violation or potential violation of law; 2) a Federal, state, local, or international agency, in response to its request; 3) a contractor of the USPTO having need for the information in order to perform a contract; 4) the Department of Justice for determination of whether the Freedom of Information Act (FOIA) requires disclosure of the record; 5) a Member of Congress submitting a request involving an individual to whom the record pertains, when the individual has requested the Member's assistance with respect to the subject matter of the record; 6) a court, magistrate, or administrative tribunal, in the course of presenting evidence, including disclosures to opposing counsel in the course of settlement negotiations; 7) the Administrator, General Services Administration (GSA), or their designee, during an inspection of records conducted by GSA under authority of 44 U.S.C. 2904 and 2906, in accordance with the GSA regulations and any other relevant (i.e., GSA or Commerce) directive, where such disclosure shall not be used to make determinations about individuals; 8) another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)); 9) the Office of Personnel Management (OPM) for personnel research purposes; and 9) the Office of Management and Budget (OMB) for legislative coordination and clearance.

If you do not furnish the information requested on this form, the USPTO may not be able to process and/or examine your submission, which may result in termination of proceedings, abandonment of the application, and/or expiration of the patent.

# Additional Uses

Additional USPTO uses of the information in this record may include disclosure to: 1) the International Bureau of the World Intellectual Property Organization, if the record is related to an international application filed under the Patent Cooperation Treaty; 2) the public i) after publication of the application pursuant to 35 U.S.C. 122(b), ii) after issuance of a patent pursuant to 35 U.S.C. 151, iii) if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections, or an issued patent, or iv) without publication of the application or patent under the specific circumstances provided for by 37 CFR 1.14(a)(1)(v)-(vii); and/or 3) the National Archives and Records Administration, for inspection of records.

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 47181086 |
| **Application Number:** | 17609911 |
| **International Application Number:** | |
| **Confirmation Number:** | 1651 |
| **Title of Invention:** | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY |
| **First Named Inventor/Applicant Name:** | Lucinda Lewis |
| **Correspondence Address:** | Lucinda Lewis<br>-<br>Automobilia II, LLC<br>1406 Columbus Boulevard<br>Coral Gables FL 33134-2351<br>US -<br>- |
| **Filer:** | Lucinda Leigh Lewis |
| **Filer Authorized By:** | |
| **Attorney Docket Number:** | 28108-125 |
| **Receipt Date:** | 09-DEC-2022 |
| **Filing Date:** | 09-NOV-2021 |
| **Time Stamp:** | 17:27:14 |
| **Application Type:** | U.S. National Stage under 35 USC 371 |

## Payment information:

| | |
|---|---|
| Submitted with Payment | no |

## File Listing:

| Document Number | Document Description | File Name | File Size(Bytes)/ Message Digest | Multi Part /.zip | Pages (if appl.) |
|---|---|---|---|---|---|
| 1 | Miscellaneous Incoming Letter | Addresschange.pdf | 312091<br>4536fa17793c4d13e4c7ed87e663fc63557b05d7 | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | **Total Files Size (in bytes):** | 312091 | | |

**This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.**

**<u>New Applications Under 35 U.S.C. 111</u>**
**If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.**

**<u>National Stage of an International Application under 35 U.S.C. 371</u>**
**If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.**

**<u>New International Application Filed with the USPTO as a Receiving Office</u>**
**If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.**

# MULTIPLE DEPENDENT CLAIM FEE CALCULATION SHEET

Substitute for Form PTO-1360
(For use with Form PTO/SB/06)

| Application Number | Filing Date |
|---|---|
| 17609911 | |

Applicant(s) Lucinda Lewis

| CLAIMS | AS FILED | | AFTER FIRST AMENDMENT | | AFTER SECOND AMENDMENT | |
|---|---|---|---|---|---|---|
| | Indep | Depend | Indep | Depend | Indep | Depend |
| 1 | 1 | | | | | |
| 2 | | 1 | | | | |
| 3 | | 1 | | | | |
| 4 | | 1 | | | | |
| 5 | | 1 | | | | |
| 6 | | 1 | | | | |
| 7 | | 1 | | | | |
| 8 | | 1 | | | | |
| 9 | | 1 | | | | |
| 10 | | 1 | | | | |
| 11 | | 1 | | | | |
| 12 | | 1 | | | | |
| 13 | | 1 | | | | |
| 14 | | 1 | | | | |
| 15 | | 1 | | | | |
| 16 | | 1 | | | | |
| 17 | | 1 | | | | |
| 18 | 1 | | | | | |
| 19 | | 1 | | | | |
| 20 | 1 | | | | | |
| 21 | | | | | | |
| 22 | | | | | | |
| 23 | | | | | | |
| 24 | | | | | | |
| 25 | | | | | | |
| 26 | | | | | | |
| 27 | | | | | | |
| 28 | | | | | | |
| 29 | | | | | | |
| 30 | | | | | | |
| 31 | | | | | | |
| 32 | | | | | | |
| 33 | | | | | | |
| 34 | | | | | | |
| 35 | | | | | | |
| 36 | | | | | | |
| 37 | | | | | | |
| 38 | | | | | | |
| 39 | | | | | | |
| 40 | | | | | | |
| 41 | | | | | | |
| 42 | | | | | | |
| 43 | | | | | | |
| 44 | | | | | | |
| 45 | | | | | | |
| 46 | | | | | | |
| 47 | | | | | | |
| 48 | | | | | | |
| 49 | | | | | | |
| 50 | | | | | | |
| Total Indep | 3 | | 0 | | 0 | |
| Total Depend | 17 | | 0 | | 0 | |
| Total Claims | 20 | | 0 | | 0 | |

* May be used for additional claims or amendments

| | * | | * | | * | |
|---|---|---|---|---|---|---|
| | Indep | Depend | Indep | Depend | Indep | Depend |
| 51 | | | | | | |
| 52 | | | | | | |
| 53 | | | | | | |
| 54 | | | | | | |
| 55 | | | | | | |
| 56 | | | | | | |
| 57 | | | | | | |
| 58 | | | | | | |
| 59 | | | | | | |
| 60 | | | | | | |
| 61 | | | | | | |
| 62 | | | | | | |
| 63 | | | | | | |
| 64 | | | | | | |
| 65 | | | | | | |
| 66 | | | | | | |
| 67 | | | | | | |
| 68 | | | | | | |
| 69 | | | | | | |
| 70 | | | | | | |
| 71 | | | | | | |
| 72 | | | | | | |
| 73 | | | | | | |
| 74 | | | | | | |
| 75 | | | | | | |
| 76 | | | | | | |
| 77 | | | | | | |
| 78 | | | | | | |
| 79 | | | | | | |
| 80 | | | | | | |
| 81 | | | | | | |
| 82 | | | | | | |
| 83 | | | | | | |
| 84 | | | | | | |
| 85 | | | | | | |
| 86 | | | | | | |
| 87 | | | | | | |
| 88 | | | | | | |
| 89 | | | | | | |
| 90 | | | | | | |
| 91 | | | | | | |
| 92 | | | | | | |
| 93 | | | | | | |
| 94 | | | | | | |
| 95 | | | | | | |
| 96 | | | | | | |
| 97 | | | | | | |
| 98 | | | | | | |
| 99 | | | | | | |
| 100 | | | | | | |

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| U.S. APPLICATION NO. | FIRST NAMED INVENTOR | ATTY. DOCKET NO. |
|---|---|---|
| 17/609,911 | Lucinda Lewis | 28108-125 |

Lucinda Lewis
Automobilia II, LLC
1406 Columbus Boulevard
Coral Gables, FL 33134-2351

| INTERNATIONAL APPLICATION NO. | |
|---|---|
| PCT/US2020/032149 | |
| I.A. FILING DATE | PRIORITY DATE |
| 05/08/2020 | 05/09/2019 |

**CONFIRMATION NO. 1651**
**371 ACCEPTANCE LETTER**

*OC000000135372261*

Date Mailed: 09/08/2022

# NOTICE OF ACCEPTANCE OF APPLICATION UNDER 35 U.S.C 371 AND 37 CFR 1.495

The applicant is hereby advised that the United States Patent and Trademark Office, in its capacity as a Designated / Elected Office (37 CFR 1.495), has ACCEPTED the above identified international application for national patentability examination in the United States Patent and Trademark Office.

The United States Application Number assigned to the application is shown above. A Filing Receipt will be issued for the present application in due course. **THE DATE APPEARING ON THE FILING RECEIPT AS THE "FILING DATE or 371(c) DATE" IS THE DATE ON WHICH THE LAST OF THE 35 U.S.C. 371 (c)(1) and (c)(2) REQUIREMENTS HAS BEEN RECEIVED IN THE OFFICE. THIS DATE IS SHOWN BELOW.** The filing date of the above identified application is the international filing date of the international application (Article 11(3) and 35 U.S.C. 363)

11/09/2021
DATE OF RECEIPT OF 35 U.S.C.
371(c)(1) and (c)(2) REQUIREMENTS

The following items have been received:

- Copy of the International Application filed on 11/09/2021
- Copy of the International Search Report filed on 11/09/2021
- Information Disclosure Statements filed on 11/09/2021
- Inventor's Oath or Declaration filed on 04/19/2022
- Request for Immediate Examination filed on 11/09/2021
- U.S. Basic National Fees filed on 11/09/2021
- Authorize Access to Search Results filed on 11/09/2021
- Priority Documents filed on 11/09/2021
- Authorization to Permit Access filed on 11/09/2021
- Application Data Sheet (37 CFR 1.76) filed on 11/09/2021

Applicant is reminded that any communications to the United States Patent and Trademark Office must be mailed to the address given in the heading and include the U.S. application no. shown above (37 CFR 1.5)

LESLIE R CHASE

Telephone: (703) 756-1026

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NUMBER | FILING or 371(c) DATE | GRP ART UNIT | FIL FEE REC'D | ATTY.DOCKET.NO | TOT CLAIMS | IND CLAIMS |
|---|---|---|---|---|---|---|
| 17/609,911 | 11/09/2021 | | 355 | 28108-125 | 20 | 3 |

**CONFIRMATION NO. 1651**

**FILING RECEIPT**

OC000000135372260

Lucinda Lewis
Automobilia II, LLC
1406 Columbus Boulevard
Coral Gables, FL 33134-2351

Date Mailed: 09/08/2022

Receipt is acknowledged of this non-provisional utility patent application. The application will be taken up for examination in due course. Applicant will be notified as to the results of the examination. Any correspondence concerning the application must include the following identification information: the U.S. APPLICATION NUMBER, FILING DATE, NAME OF FIRST INVENTOR, and TITLE OF INVENTION. Fees transmitted by check or draft are subject to collection.

**Please verify the accuracy of the data presented on this receipt.** If an error is noted on this Filing Receipt, please submit a written request for a corrected Filing Receipt, including a properly marked-up ADS showing the changes with strike-through for deletions and underlining for additions. If you received a "Notice to File Missing Parts" or other Notice requiring a response for this application, please submit any request for correction to this Filing Receipt with your reply to the Notice. When the USPTO processes the reply to the Notice, the USPTO will generate another Filing Receipt incorporating the requested corrections provided that the request is grantable.

**Inventor(s)**

Lucinda Lewis, Burbank;

**Applicant(s)**

Automobilia II, LLC, Coral Gables, FL;

**Power of Attorney:** None

**Domestic Priority data as claimed by applicant**

This application is a 371 of PCT/US2020/032149 05/08/2020
which claims benefit of 62/845,546 05/09/2019

**Foreign Applications** for which priority is claimed (You may be eligible to benefit from the **Patent Prosecution Highway** program at the USPTO. Please see http://www.uspto.gov for more information.) - None.
*Foreign application information must be provided in an Application Data Sheet in order to constitute a claim to foreign priority. See 37 CFR 1.55 and 1.76.*

**Permission to Access Application via Priority Document Exchange:** Yes

**Permission to Access Search Results:** Yes

Applicant may provide or rescind an authorization for access using Form PTO/SB/39 or Form PTO/SB/69 as appropriate.

**If Required, Foreign Filing License Granted:** 07/24/2022

The country code and number of your priority application, to be used for filing abroad under the Paris Convention, is **US 17/609,911**

**Projected Publication Date:** 12/15/2022

**Non-Publication Request:** No

**Early Publication Request:** No
**** MICRO ENTITY ****
**Title**

METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

**Preliminary Class**

**Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications:** No

## PROTECTING YOUR INVENTION OUTSIDE THE UNITED STATES

Since the rights granted by a U.S. patent extend only throughout the territory of the United States and have no effect in a foreign country, an inventor who wishes patent protection in another country must apply for a patent in a specific country or in regional patent offices. Applicants may wish to consider the filing of an international application under the Patent Cooperation Treaty (PCT). An international (PCT) application generally has the same effect as a regular national patent application in each PCT-member country. The PCT process **simplifies** the filing of patent applications on the same invention in member countries, but **does not result** in a grant of "an international patent" and does not eliminate the need of applicants to file additional documents and fees in countries where patent protection is desired.

Almost every country has its own patent law, and a person desiring a patent in a particular country must make an application for patent in that country in accordance with its particular laws. Since the laws of many countries differ in various respects from the patent law of the United States, applicants are advised to seek guidance from specific foreign countries to ensure that patent rights are not lost prematurely.

Applicants also are advised that in the case of inventions made in the United States, the Director of the USPTO must issue a license before applicants can apply for a patent in a foreign country. The filing of a U.S. patent application serves as a request for a foreign filing license. The application's filing receipt contains further information and guidance as to the status of applicant's license for foreign filing.

Applicants may wish to consult the USPTO booklet, "General Information Concerning Patents" (specifically, the section entitled "Treaties and Foreign Patents") for more information on timeframes and deadlines for filing foreign patent applications. The guide is available either by contacting the USPTO Contact Center at 800-786-9199, or it can be viewed on the USPTO website at http://www.uspto.gov/web/offices/pac/doc/general/index.html.

For information on preventing theft of your intellectual property (patents, trademarks and copyrights), you may wish to consult the U.S. Government website, http://www.stopfakes.gov. Part of a Department of Commerce initiative, this website includes self-help "toolkits" giving innovators guidance on how to protect intellectual property in specific countries such as China, Korea and Mexico. For questions regarding patent enforcement issues, applicants may call the U.S. Government hotline at 1-866-999-HALT (1-866-999-4258).

## LICENSE FOR FOREIGN FILING UNDER

## Title 35, United States Code, Section 184

## Title 37, Code of Federal Regulations, 5.11 & 5.15

**GRANTED**

The applicant has been granted a license under 35 U.S.C. 184, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" followed by a date appears on this form. Such licenses are issued in all applications where the conditions for issuance of a license have been met, regardless of whether or not a license may be required as set forth in 37 CFR 5.15. The scope and limitations of this license are set forth in 37 CFR 5.15(a) unless an earlier license has been issued under 37 CFR 5.15(b). The license is subject to revocation upon written notification. The date indicated is the effective date of the license, unless an earlier license of similar scope has been granted under 37 CFR 5.13 or 5.14.

This license is to be retained by the licensee and may be used at any time on or after the effective date thereof unless it is revoked. This license is automatically transferred to any related applications(s) filed under 37 CFR 1.53(d). This license is not retroactive.

The grant of a license does not in any way lessen the responsibility of a licensee for the security of the subject matter as imposed by any Government contract or the provisions of existing laws relating to espionage and the national security or the export of technical data. Licensees should apprise themselves of current regulations especially with respect to certain countries, of other agencies, particularly the Office of Defense Trade Controls, Department of State (with respect to Arms, Munitions and Implements of War (22 CFR 121-128)); the Bureau of Industry and Security, Department of Commerce (15 CFR parts 730-774); the Office of Foreign AssetsControl, Department of Treasury (31 CFR Parts 500+) and the Department of Energy.

**NOT GRANTED**

No license under 35 U.S.C. 184 has been granted at this time, if the phrase "IF REQUIRED, FOREIGN FILING LICENSE GRANTED" DOES NOT appear on this form. Applicant may still petition for a license under 37 CFR 5.12, if a license is desired before the expiration of 6 months from the filing date of the application. If 6 months has lapsed from the filing date of this application and the licensee has not received any indication of a secrecy order under 35 U.S.C. 181, the licensee may foreign file the application pursuant to 37 CFR 5.15(b).

---

# PATENT APPLICATION FEE DETERMINATION RECORD
Substitute for Form PTO-875

Application or Docket Number
17/609,911

## APPLICATION AS FILED - PART I

| FOR | NUMBER FILED (Column 1) | NUMBER EXTRA (Column 2) | SMALL ENTITY RATE($) | SMALL ENTITY FEE($) | OR | OTHER THAN SMALL ENTITY RATE($) | OTHER THAN SMALL ENTITY FEE($) |
|---|---|---|---|---|---|---|---|
| BASIC FEE (37 CFR 1.16(a), (b), or (c)) | N/A | N/A | N/A | | | N/A | 80 |
| SEARCH FEE (37 CFR 1.16(k), (i), or (m)) | N/A | N/A | N/A | | | N/A | 35 |
| EXAMINATION FEE (37 CFR 1.16(o), (p), or (q)) | N/A | N/A | N/A | | | N/A | 200 |
| TOTAL CLAIMS (37 CFR 1.16(i)) | 20 minus 20 = | * | | | OR | x 25 = | 0.00 |
| INDEPENDENT CLAIMS (37 CFR 1.16(h)) | 3 minus 3 = | * | | | | x 120 = | 0.00 |
| APPLICATION SIZE FEE (37 CFR 1.16(s)) | If the specification and drawings exceed 100 sheets of paper, the application size fee due is $310 ($155 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s). | | | | | | 0.00 |
| MULTIPLE DEPENDENT CLAIM PRESENT (37 CFR 1.16(j)) | | | | | | | 0.00 |
| | | | TOTAL | | | TOTAL | 315 |

* If the difference in column 1 is less than zero, enter "0" in column 2.

## APPLICATION AS AMENDED - PART II

| AMENDMENT A | CLAIMS REMAINING AFTER AMENDMENT (Column 1) | | HIGHEST NUMBER PREVIOUSLY PAID FOR (Column 2) | PRESENT EXTRA (Column 3) | SMALL ENTITY RATE($) | SMALL ENTITY ADDITIONAL FEE($) | OR | OTHER THAN SMALL ENTITY RATE($) | OTHER THAN SMALL ENTITY ADDITIONAL FEE($) |
|---|---|---|---|---|---|---|---|---|---|
| Total (37 CFR 1.16(i)) | * | Minus | ** | = | x = | | OR | x = | |
| Independent (37 CFR 1.16(h)) | * | Minus | *** | = | x = | | OR | x = | |
| Application Size Fee (37 CFR 1.16(s)) | | | | | | | | | |
| FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j)) | | | | | | | OR | | |
| | | | | | TOTAL ADD'L FEE | | OR | TOTAL ADD'L FEE | |

| AMENDMENT B | CLAIMS REMAINING AFTER AMENDMENT (Column 1) | | HIGHEST NUMBER PREVIOUSLY PAID FOR (Column 2) | PRESENT EXTRA (Column 3) | RATE($) | ADDITIONAL FEE($) | OR | RATE($) | ADDITIONAL FEE($) |
|---|---|---|---|---|---|---|---|---|---|
| Total (37 CFR 1.16(i)) | * | Minus | ** | = | x = | | OR | x = | |
| Independent (37 CFR 1.16(h)) | * | Minus | *** | = | x = | | OR | x = | |
| Application Size Fee (37 CFR 1.16(s)) | | | | | | | | | |
| FIRST PRESENTATION OF MULTIPLE DEPENDENT CLAIM (37 CFR 1.16(j)) | | | | | | | OR | | |
| | | | | | TOTAL ADD'L FEE | | OR | TOTAL ADD'L FEE | |

* If the entry in column 1 is less than the entry in column 2, write "0" in column 3.
** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 20, enter "20".
*** If the "Highest Number Previously Paid For" IN THIS SPACE is less than 3, enter "3".
The "Highest Number Previously Paid For" (Total or Independent) is the highest found in the appropriate box in column 1.

| | |
|---|---|
| **To:** | patents@lowenstein.com,, |
| **From:** | PAIR_eOfficeAction@uspto.gov |
| **Cc:** | PAIR_eOfficeAction@uspto.gov |
| **Subject:** | Private PAIR Correspondence Notification for Customer Number 131890 |

Jul 05, 2022 03:53:34 AM

Dear PAIR Customer:

LOWENSTEIN SANDLER LLP / BASF
Patent Docket Administrator
One Lowenstein Drive
Roseland, NJ 07068
UNITED STATES

The following USPTO patent application(s) associated with your Customer Number, 131890 , have new outgoing correspondence. This correspondence is now available for viewing in Private PAIR.

The official date of notification of the outgoing correspondence will be indicated on the form PTOL-90 accompanying the correspondence.

Disclaimer:
The list of documents shown below is provided as a courtesy and is not part of the official file wrapper. The content of the images shown in PAIR is the official record.

| Application | Document | Mailroom Date | Attorney Docket No. |
|---|---|---|---|
| 17609911 | PET.DEC.OIPE | 07/05/2022 | 28108-125 |

To view your correspondence online or update your email addresses, please visit us anytime at https://sportal.uspto.gov/secure/myportal/privatepair.

If you have any questions, please email the Electronic Business Center (EBC) at EBC@uspto.gov with 'e-Office Action' on the subject line or call 1-866-217-9197 during the following hours:

Monday - Friday 6:00 a.m. to 12:00 a.m.

Thank you for prompt attention to this notice,

UNITED STATES PATENT AND TRADEMARK OFFICE

PATENT APPLICATION INFORMATION RETRIEVAL SYSTEM

UNITED STATES DEPARTMENT OF COMMERCE
**United States Patent and Trademark Office**
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 17/609,911 | | Lucinda Lewis | 28108-125 | 1651 |

131890 7590 07/05/2022
LOWENSTEIN SANDLER LLP / BASF
Patent Docket Administrator
One Lowenstein Drive
Roseland, NJ 07068

| EXAMINER | |
|---|---|
| | |

| ART UNIT | PAPER NUMBER |
|---|---|
| | |

| NOTIFICATION DATE | DELIVERY MODE |
|---|---|
| 07/05/2022 | ELECTRONIC |

**Please find below and/or attached an Office communication concerning this application or proceeding.**

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patents@lowenstein.com

Commissioner for Patents
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450
www.uspto.gov

LOWENSTEIN SANDLER LLP / BASF Patent Docket Administrator
One Lowenstein Drive
Roseland, NJ 07068

In re Application of Automobilia II, LLC
Application No.: 17/609,911
Filing Date: 09 November 2021
Attorney Docket No.: 28108-125
For: METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

**DECISION ON PETITION UNDER 37 CFR 1.78(c)**

This is a decision on the petition under 37 CFR 1.78(c), filed 11 May 2022, to accept an unintentionally delayed claim under 35 U.S.C. 119(e) for the benefit of priority to one or more prior-filed provisional applications.

Under 37 CFR 1.78(c), a petition to accept an unintentionally delayed claim under 35 U.S.C. 119(e) for the benefit of a prior-filed application must be accompanied by:

(i) the reference required by 35 U.S.C. 119(e) and 37 CFR 1.78(a)(3) to the prior-filed application, unless previously submitted;

(ii) the petition fee set forth in 37 CFR 1.17(m); and

(iii) a statement that the entire delay between the date the claim was due under 37 CF R 1.78(a)(4) and the date the claim was filed was unintentional. The Director may require additional information where there is a question whether the delay was unintentional.

With regard to item (i), a proper reference to the prior-filed application(s) has been included in an application data sheet as required by 37 CFR 1.78(a)(3).

With regard to item (ii), the petition fee set forth in 37 CFR 1.17(m) has been submitted.

With regard to item (iii), the statement of unintentional delay contained in the petition differs slightly from the language contained in 37 CFR 1.78(c)(3) and is hereby construed as a statement that the entire delay between the date the claim was due under 1.78(a)(4) and the date the claim was filed was unintentional. If this interpretation is incorrect, applicant is required to immediately notify the Office. As construed, the provided statement of unintentional delay is acceptable.

For the reasons above, the petition under 37 CFR 1.78(c) is **GRANTED.**

The priority information requested on petition has been entered into the record. The United States Designated/Elected Office (DO/EO/US) will process the application and issue a filing receipt in due course. The Applicant is advised that the inclusion of a prior-filed application on any filing receipt should not be construed as meaning that the application is necessarily entitled to the benefit of the prior-filed application. In order for an application to be entitled to the benefit of the prior-filed application, all other requirements under 35 U.S.C. 119(e) and 37 CFR 1.78 must be met. The examiner will, in due course, determine whether this application is entitled to the benefit of an earlier filing date.

This matter is being referred to the United States Designated/Elected Office (DO/EO/US) for further processing in accordance with this decision.

/bethannne dayoan/

BETHANNE C DAYOAN
PCT Legal Examiner
International Patent Legal Administration
(571)272-3209

PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

| | |
|---|---|
| In Re Patent Application of:<br><br>Lucinda Lewis<br><br>Application No.: 17/609,911<br><br>Filed: November 9, 2021<br><br>For: METHODS, SYSTEMS AND COMPUTERPROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | Examiner: Unassigned<br><br>Art Unit: Unassigned<br><br>Confirmation No.: 1651 |

**CERTIFICATE OF TRANSMISSION**
I hereby certify that this correspondence is being submitted electronically via EFS Web on the date shown below.

/Katherine Johnson/ May 11 2022
***Katherine Johnson*** ***Date***

**FILED VIA EFS WEB**
Mail Stop Missing Parts
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

PETITION TO ACCEPT UNINTENTIONALLY DELAYED CLAIM OF BENEFIT OF A PROVISIONAL APPLICATION UNDER 35 U.S.C. 119(e)

Dear Sir:

Applicant submits herewith a Petition to accept the unintentionally delayed claim of benefit of a provisional application under 35 U.S.C. 119(e), as outlined in 37 CFR §1.78(c). The entire delay between the date the benefit claim was due and the filing of this submission is unintentional.

A corrected Application Data Sheet, which corrects and includes the priority data for the provisional application, and the petition fee due according to 37 CFR §1.17(m) are included as part of this submission.

Also submitted herewith is the Examination Fee and Search Fee due for the above-referenced patent application. The Basic national stage fee and surcharge for late filing of fees have been previously paid. If there are any additional charges due, please charge Deposit Account No. 50-1358.

Respectfully submitted,

LOWENSTEIN SANDLER LLP

Dated: May 11, 2022

/ Benjamin A. Kimes /
Benjamin A. Kimes
Registration No. 50,870

390 Lytton Avenue
Palo Alto, California 94301
(650) 433-5724

| Application Data Sheet 37 CFR 1.76 | Attorney Docket Number | ~~28108-125~~ |
|---|---|---|
| | Application Number | 17/609,911 |
| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | |

The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76.
This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.

## Secrecy Order 37 CFR 5.2:

☐ Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)

## Inventor Information:

**Inventor 1** Remove

**Legal Name**

| Prefix | Given Name | Middle Name | Family Name | Suffix |
|---|---|---|---|---|
| | Lucinda | | Lewis | |

**Residence Information (Select One)** ○ US Residency ◉ Non US Residency ○ Active US Military Service

| City | Burbank | Country of Residence | US |
|---|---|---|---|

**Mailing Address of Inventor:**

| | | | |
|---|---|---|---|
| **Address 1** | 1406 Columbus Boulevard | | |
| **Address 2** | | | |
| **City** | Coral Gables | **State/Province** | FL |
| **Postal Code** | 33134-2351 | **Country** | US |

All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the **Add** button. Add

## Correspondence Information:

**Enter either Customer Number or complete the Correspondence Information section below.**
**For further information see 37 CFR 1.33(a).** ~~131890~~

☒ **An Address is being provided for the correspondence Information of this application.**

| | | | |
|---|---|---|---|
| **Name 1** | Lucinda Lewis | **Name 2** | |
| **Address 1** | Automobilia II, LLC | | |
| **Address 2** | 1406 Columbus Boulevard | | |
| **City** | Coral Gables | **State/Province** | FL |
| **Country** | US | **Postal Code** | 33134-2351 |
| **Phone Number** | (818) 633-1176 | **Fax Number** | |
| **Email Address** | ~~patents@lowenstein.com~~ cindy.lewis@me.com | Add Email | Remove Email |

| Application Data Sheet 37 CFR 1.76 | Attorney Docket Number | ~~28108-125~~ |
|---|---|---|
| | Application Number | 17/609,911 |
| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | |

## Application Information:

| Title of the Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | | |
|---|---|---|---|
| Attorney Docket Number | ~~28108-125~~ | Small Entity Status Claimed | ☒ |
| Application Type | Nonprovisional | | |
| Subject Matter | Utility | | |
| Total Number of Drawing Sheets (if any) | 7 | Suggested Figure for Publication (if any) | |

## Filing By Reference:

Only complete this section when filing an application by reference under 35 U.S.C. 111(c) and 37 CFR 1.57(a). Do not complete this section if application papers including a specification and any drawings are being filed. Any domestic benefit or foreign priority information must be provided in the appropriate section(s) below (i.e., "Domestic Benefit/National Stage Information" and "Foreign Priority Information").

For the purposes of a filing date under 37 CFR 1.53(b), the description and any drawings of the present application are replaced by this reference to the previously filed application, subject to conditions and requirements of 37 CFR 1.57(a).

| Application number of the previously filed application | Filing date (YYYY-MM-DD) | Intellectual Property Authority or Country |
|---|---|---|
| | | |

## Publication Information:

☐ Request Early Publication (Fee required at time of Request 37 CFR 1.219)

☐ **Request Not to Publish.** I hereby request that the attached application not be published under 35 U.S.C. 122(b) and certify that the invention disclosed in the attached application **has not and will not** be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

## Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32).
Either enter Customer Number or complete the Representative Name section below. If both sections are completed the customer Number will be used for the Representative Information during processing.

| Please Select One: | ◉ Customer Number | ○ US Patent Practitioner | ○ Limited Recognition (37 CFR 11.9) |
|---|---|---|---|
| Customer Number | ~~131890~~ | | |

| Application Data Sheet 37 CFR 1.76 | Attorney Docket Number | ~~28108-125~~ |
|---|---|---|
| | Application Number | 17/609,911 |
| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | |

## Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, 365(c), or 386(c) or indicate National Stage entry from a PCT application. Providing benefit claim information in the Application Data Sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78.
When referring to the current application, please leave the "Application Number" field blank.

| Prior Application Status | Pending | | Remove |
|---|---|---|---|
| Application Number | Continuity Type | Prior Application Number | Filing or 371(c) Date (YYYY-MM-DD) |
| | a 371 of international | PCT/US2020/032149 | 2020-05-08 |
| Prior Application Status | Expired | | Remove |
| Application Number PCT/US2020/032149 | Continuity Type | Prior Application Number | Filing or 371(c) Date (YYYY-MM-DD) |
| ~~62845546~~ | Claims benefit of provisional | 62845546 | 2019-05-09 |

Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the **Add** button.

## Foreign Priority Information:

This section allows for the applicant to claim priority to a foreign application. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55. When priority is claimed to a foreign application that is eligible for retrieval under the priority document exchange program (PDX)[i] the information will be used by the Office to automatically attempt retrieval pursuant to 37 CFR 1.55(i)(1) and (2). Under the PDX program, applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period specified in 37 CFR 1.55(g)(1).

Remove

| Application Number | Country[i] | Filing Date (YYYY-MM-DD) | Access Code[i] (if applicable) |
|---|---|---|---|
| | | | |

Additional Foreign Priority Data may be generated within this form by selecting the **Add** button.

## Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications

☐ This application (1) claims priority to or the benefit of an application filed before March 16, 2013 and (2) also contains, or contained at any time, a claim to a claimed invention that has an effective filing date on or after March 16, 2013.
NOTE: By providing this statement under 37 CFR 1.55 or 1.78, this application, with a filing date on or after March 16, 2013, will be examined under the first inventor to file provisions of the AIA.

| Application Data Sheet 37 CFR 1.76 | Attorney Docket Number | ~~28108-125~~ |
|---|---|---|
| | Application Number | 17/609,911 |
| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | |

## Authorization or Opt-Out of Authorization to Permit Access:

When this Application Data Sheet is properly signed and filed with the application, applicant has provided written authority to permit a participating foreign intellectual property (IP) office access to the instant application-as-filed (see paragraph A in subsection 1 below) and the European Patent Office (EPO) access to any search results from the instant application (see paragraph B in subsection 1 below).

Should applicant choose not to provide an authorization identified in subsection 1 below, applicant **must opt-out** of the authorization by checking the corresponding box A or B or both in subsection 2 below.

**NOTE**: This section of the Application Data Sheet is **ONLY** reviewed and processed with the **INITIAL** filing of an application. After the initial filing of an application, an Application Data Sheet cannot be used to provide or rescind authorization for access by a foreign IP office(s). Instead, Form PTO/SB/39 or PTO/SB/69 must be used as appropriate.

**1. Authorization to Permit Access by a Foreign Intellectual Property Office(s)**

**A. Priority Document Exchange (PDX)** - Unless box A in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the State Intellectual Property Office of the People's Republic of China (SIPO), the World Intellectual Property Organization (WIPO), and any other foreign intellectual property office participating with the USPTO in a bilateral or multilateral priority document exchange agreement in which a foreign application claiming priority to the instant patent application is filed, access to: (1) the instant patent application-as-filed and its related bibliographic data, (2) any foreign or domestic application to which priority or benefit is claimed by the instant application and its related bibliographic data, and (3) the date of filing of this Authorization. See 37 CFR 1.14(h)(1).

**B. Search Results from U.S. Application to EPO** - Unless box B in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the EPO access to the bibliographic data and search results from the instant patent application when a European patent application claiming priority to the instant patent application is filed. See 37 CFR 1.14(h)(2).

The applicant is reminded that the EPO's Rule 141(1) EPC (European Patent Convention) requires applicants to submit a copy of search results from the instant application without delay in a European patent application that claims priority to the instant application.

**2. Opt-Out of Authorizations to Permit Access by a Foreign Intellectual Property Office(s)**

☐ A. Applicant **DOES NOT** authorize the USPTO to permit a participating foreign IP office access to the instant application-as-filed. If this box is checked, the USPTO will not be providing a participating foreign IP office with any documents and information identified in subsection 1A above.

☐ B. Applicant **DOES NOT** authorize the USPTO to transmit to the EPO any search results from the instant patent application. If this box is checked, the USPTO will not be providing the EPO with search results from the instant application.

**NOTE:** Once the application has published or is otherwise publicly available, the USPTO may provide access to the application in accordance with 37 CFR 1.14.

| **Application Data Sheet 37 CFR 1.76** | Attorney Docket Number | ~~28108-125~~ |
|---|---|---|
| | Application Number | 17/609,911 |
| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | |

## Applicant Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

**Applicant 1**

If the applicant is the inventor (or the remaining joint inventor or inventors under 37 CFR 1.45), this section should not be completed. The information to be provided in this section is the name and address of the legal representative who is the applicant under 37 CFR 1.43; or the name and address of the assignee, person to whom the inventor is under an obligation to assign the invention, or person who otherwise shows sufficient proprietary interest in the matter who is the applicant under 37 CFR 1.46. If the applicant is an applicant under 37 CFR 1.46 (assignee, person to whom the inventor is obligated to assign, or person who otherwise shows sufficient proprietary interest) together with one or more joint inventors, then the joint inventor or inventors who are also the applicant should be identified in this section.

Clear

- (●) Assignee
- ( ) Legal Representative under 35 U.S.C. 117
- ( ) Joint Inventor
- ( ) Person to whom the inventor is obligated to assign.
- ( ) Person who shows sufficient proprietary interest

If applicant is the legal representative, indicate the authority to file the patent application, the inventor is:

Name of the Deceased or Legally Incapacitated Inventor:

If the Applicant is an Organization check here. ☒

| Organization Name | Automobilia II, LLC | | |
|---|---|---|---|
| **Mailing Address Information For Applicant:** | | | |
| **Address 1** | 1406 Columbus Boulevard | | |
| Address 2 | | | |
| **City** | Coral Gables | **State/Province** | FL |
| **Country** | US | Postal Code | 33134-2351 |
| Phone Number | | Fax Number | |
| Email Address | | | |

Additional Applicant Data may be generated within this form by selecting the Add button.

## Assignee Information including Non-Applicant Assignee Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

| **Application Data Sheet 37 CFR 1.76** | Attorney Docket Number | ~~28108-125~~ |
|---|---|---|
| | Application Number | 17/609,911 |

| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY |
|---|---|

**Assignee 1**

Complete this section if assignee information, including non-applicant assignee information, is desired to be included on the patent application publication. An assignee-applicant identified in the "Applicant Information" section will appear on the patent application publication as an applicant. For an assignee-applicant, complete this section only if identification as an assignee is also desired on the patent application publication.

If the Assignee or Non-Applicant Assignee is an Organization check here. ☐

| Prefix | **Given Name** | Middle Name | **Family Name** | Suffix |
|---|---|---|---|---|
| | | | | |

**Mailing Address Information For Assignee including Non-Applicant Assignee:**

| | | | |
|---|---|---|---|
| **Address 1** | | | |
| Address 2 | | | |
| **City** | | **State/Province** | |
| **Country** | | Postal Code | |
| Phone Number | | Fax Number | |
| Email Address | | | |

Additional Assignee or Non-Applicant Assignee Data may be generated within this form by selecting the Add button.

## Signature:

**NOTE:** This Application Data Sheet must be signed in accordance with 37 CFR 1.33(b). **However, if this Application Data Sheet is submitted with the INITIAL filing of the application and either box A or B is not checked in subsection 2 of the "Authorization or Opt-Out of Authorization to Permit Access" section, then this form must also be signed in accordance with 37 CFR 1.14(c).**

This Application Data Sheet **must** be signed by a patent practitioner if one or more of the applicants is a **juristic entity** (e.g., corporation or association). If the applicant is two or more joint inventors, this form must be signed by a patent practitioner, **all** joint inventors who are the applicant, or one or more joint inventor-applicants who have been given power of attorney (e.g., see USPTO Form PTO/AIA/81) on behalf of **all** joint inventor-applicants.

See 37 CFR 1.4(d) for the manner of making signatures and certifications.

| **Signature** | /Benjamin A. Kimes/ | | | Date (YYYY-MM-DD) | 2022-05-11 |
|---|---|---|---|---|---|
| First Name | Benjamin A. | Last Name | Kimes | Registration Number | 50870 |

Additional Signature may be generated within this form by selecting the Add button.

| **Application Data Sheet 37 CFR 1.76** | Attorney Docket Number | ~~28108-125~~ |
|---|---|---|
| | Application Number | 17/609,911 |
| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | |

# Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent CooperationTreaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

# Electronic Patent Application Fee Transmittal

| | |
|---|---|
| **Application Number:** | 17609911 |
| **Filing Date:** | |
| **Title of Invention:** | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY |
| **First Named Inventor/Applicant Name:** | Lucinda Lewis |
| **Filer:** | Benjamin Kimes/Katherine Johnson |
| **Attorney Docket Number:** | 28108-125 |

Filed as Micro Entity

**Filing Fees for U.S. National Stage under 35 USC 371**

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Basic Filing:** | | | | |
| NATL STAGE SEARCH FEE - U.S. WAS THE ISA | 3641 | 1 | 35 | 35 |
| NATIONAL STAGE EXAM - ALL OTHER CASES | 3633 | 1 | 200 | 200 |
| **Pages:** | | | | |
| **Claims:** | | | | |
| **Miscellaneous-Filing:** | | | | |
| **Petition:** | | | | |
| PET. DELAY SUB OR RESTORE PRIORITY-CLAIM | 3454 | 1 | 525 | 525 |
| **Patent-Appeals-and-Interference:** | | | | |

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Post-Allowance-and-Post-Issuance:** | | | | |
| **Extension-of-Time:** | | | | |
| **Miscellaneous:** | | | | |
| | | **Total in USD ($)** | | **760** |

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 45690980 |
| **Application Number:** | 17609911 |
| **International Application Number:** | |
| **Confirmation Number:** | 1651 |
| **Title of Invention:** | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY |
| **First Named Inventor/Applicant Name:** | Lucinda Lewis |
| **Customer Number:** | 131890 |
| **Filer:** | Benjamin Kimes/Katherine Johnson |
| **Filer Authorized By:** | Benjamin Kimes |
| **Attorney Docket Number:** | 28108-125 |
| **Receipt Date:** | 11-MAY-2022 |
| **Filing Date:** | |
| **Time Stamp:** | 18:49:44 |
| **Application Type:** | U.S. National Stage under 35 USC 371 |

## Payment information:

| | |
|---|---|
| Submitted with Payment | yes |
| Payment Type | DA |
| Payment was successfully received in RAM | $760 |
| RAM confirmation Number | E20225AI50053634 |
| Deposit Account | 501358 |
| Authorized User | Katherine Johnson |

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

37 CFR 1.21 (Miscellaneous fees and charges)

## File Listing:

| Document Number | Document Description | File Name | File Size(Bytes)/ Message Digest | Multi Part /.zip | Pages (if appl.) |
|---|---|---|---|---|---|
| 1 | Petition for review by the PCT legal office | 28108-0125_L0110_PetitionCorrectPriority.pdf | 102410<br>74d6463104c20df568ff2d1584f5479f6294a655 | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 2 | Application Data Sheet | 28108-0125_CorrectedADS.pdf | 1648296<br>cb4ac1c18a9ccb7e856ed7fa9e2b310fdcee7425 | no | 8 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| This is not an USPTO supplied ADS fillable form | | | | | |
| 3 | Fee Worksheet (SB06) | fee-info.pdf | 44714<br>0d179f41a2064ce443fde6fd9494ee6c60cff06e | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | | **Total Files Size (in bytes):** | 1795420 | |

# DECLARATION (37 CFR 1.63) FOR UTILITY OR DESIGN APPLICATION USING AN APPLICATION DATA SHEET (37 CFR 1.76)

| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY |
|---|---|

As the below named inventor, I hereby declare that:

This declaration is directed to:

☐ The attached application, or

☒ United States application or PCT international application number 17/609,911

filed on November 9, 2021.

The above-identified application was made or authorized to be made by me.

I believe that I am the original inventor or an original joint inventor of a claimed invention in the application.

I hereby acknowledge that any willful false statement made in this declaration is punishable under 18 U.S.C. 1001 by fine or imprisonment of not more than five (5) years, or both.

## WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

LEGAL NAME OF INVENTOR

Inventor: Lucinda Lewis Date (Optional) : 4-18-2022

Signature:

Note: An application data sheet (PTO/SB/14 or equivalent), including naming the entire inventive entity, must accompany this form or must have been previously filed. Use an additional PTO/AIA/01 form for each additional inventor.

# Privacy Act Statement

The **Privacy Act of 1974 (P.L. 93-579)** requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant ( *i.e.*, GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

# Electronic Patent Application Fee Transmittal

| | |
|---|---|
| **Application Number:** | |
| **Filing Date:** | |
| **Title of Invention:** | Your fee transmittal sheet cannot be generated at this time. Please complete your submission.<br>After completing your submission, please contact the Electronic Business Center (EBC) to request the fee transmittal sheet for this submission.<br><br>We apologize for the inconvenience.<br><br>Electronic Business Center (EBC)<br>1-866-217-9197 (toll-free)<br>or 571-272-4100<br>6 a.m. to 12 Midnight Eastern Time<br>Monday -Friday |
| **First Named Inventor/Applicant Name:** | |
| **Filer:** | |
| **Attorney Docket Number:** | |

## Filing Fees

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Basic Filing:** | | | | |
| **Pages:** | | | | |
| **Claims:** | | | | |
| **Miscellaneous-Filing:** | | | | |
| **Petition:** | | | | |
| **Patent-Appeals-and-Interference:** | | | | |
| **Post-Allowance-and-Post-Issuance:** | | | | |
| **Extension-of-Time:** | | | | |

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Miscellaneous:** | | | | |
| | | | | |
| | **Total in USD ($)** | | | |

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 45503523 |
| **Application Number:** | 17609911 |
| **International Application Number:** | |
| **Confirmation Number:** | 1651 |
| **Title of Invention:** | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY |
| **First Named Inventor/Applicant Name:** | Lucinda Lewis |
| **Customer Number:** | 131890 |
| **Filer:** | Mary Kristin Nicholes |
| **Filer Authorized By:** | |
| **Attorney Docket Number:** | 28108-125 |
| **Receipt Date:** | 19-APR-2022 |
| **Filing Date:** | |
| **Time Stamp:** | 17:30:22 |
| **Application Type:** | U.S. National Stage under 35 USC 371 |

## Payment information:

| | |
|---|---|
| Submitted with Payment | yes |
| Payment Type | DA |
| Payment was successfully received in RAM | $40 |
| RAM confirmation Number | E20224IH30420277 |
| Deposit Account | 501358 |
| Authorized User | Mary Nicholes |

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

37 CFR 1.17 (Patent application and reexamination processing fees)

37 CFR 1.19 (Document supply fees)

37 CFR 1.20 (Post Issuance fees)

37 CFR 1.21 (Miscellaneous fees and charges)

37 CFR 1.492 (National application filing, search, and examination fees)

## File Listing:

| Document Number | Document Description | File Name | File Size(Bytes)/ Message Digest | Multi Part /.zip | Pages (if appl.) |
|---|---|---|---|---|---|
| 1 | Oath or Declaration filed | 28108_125_Declaration.pdf | 136020<br>5daa5b51dbff0ba0c00c71dc7b7f825f5489456b | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 2 | Fee Worksheet (SB06) | fee-info.pdf | 30061<br>769380a49c92bc8e06d1c1dfb5959d2273810ce5 | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | | **Total Files Size (in bytes):** | 166081 | |

**This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.**

**New Applications Under 35 U.S.C. 111**
**If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.**
**National Stage of an International Application under 35 U.S.C. 371**
**If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.**
**New International Application Filed with the USPTO as a Receiving Office**
**If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.**

## INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 20/32149

A. CLASSIFICATION OF SUBJECT MATTER
IPC - G06N 20/00, G06K 9/46, G06K 9/62, G06N 3/08 (2020.01)

CPC - G06N 20/00, G06K 9/629, G06N 3/0454, G06N 3/08, G06K 9/62

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
See Search History document

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X<br>--<br>Y | US 2018/0084310 A1 (GumGum, Inc.) 22 March 2018 (22.03.2018), entire document, especially abstract and para [0018]-[0019], [0025]-[0027], [0080]-[0081], [0087], [0108]. | 1-11, 18-20<br>------------------------<br>12-17 |
| Y | US 2019/0005670 A1 (Magic Leap, Inc.) 03 January 2019 (03.01.2019), entire document, especially abstract and para [0032], [0035], [0063]. | 12-14, 17 |
| Y | US 2019/0102676 A1 (SAS Institute Inc.) 04 April 2019 (04.04.2019), entire document, especially abstract and para [0048]-[0049], [0148]-[0150]. | 15-16 |
| A | US 2018/0293552 A1 (Alibaba Group Holding Limited) 11 October 2018 (11.10.2018), entire document. | 1-20 |

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"D" document cited by the applicant in the international application
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 01 July 2020 (01.07.2020) | 04 AUG 2020 |

| Name and mailing address of the ISA/US | Authorized officer |
|---|---|
| Mail Stop PCT, Attn: ISA/US, Commissioner for Patents<br>P.O. Box 1450, Alexandria, Virginia 22313-1450<br>Facsimile No. 571-273-8300 | Lee Young<br>Telephone No. PCT Helpdesk: 571-272-4300 |

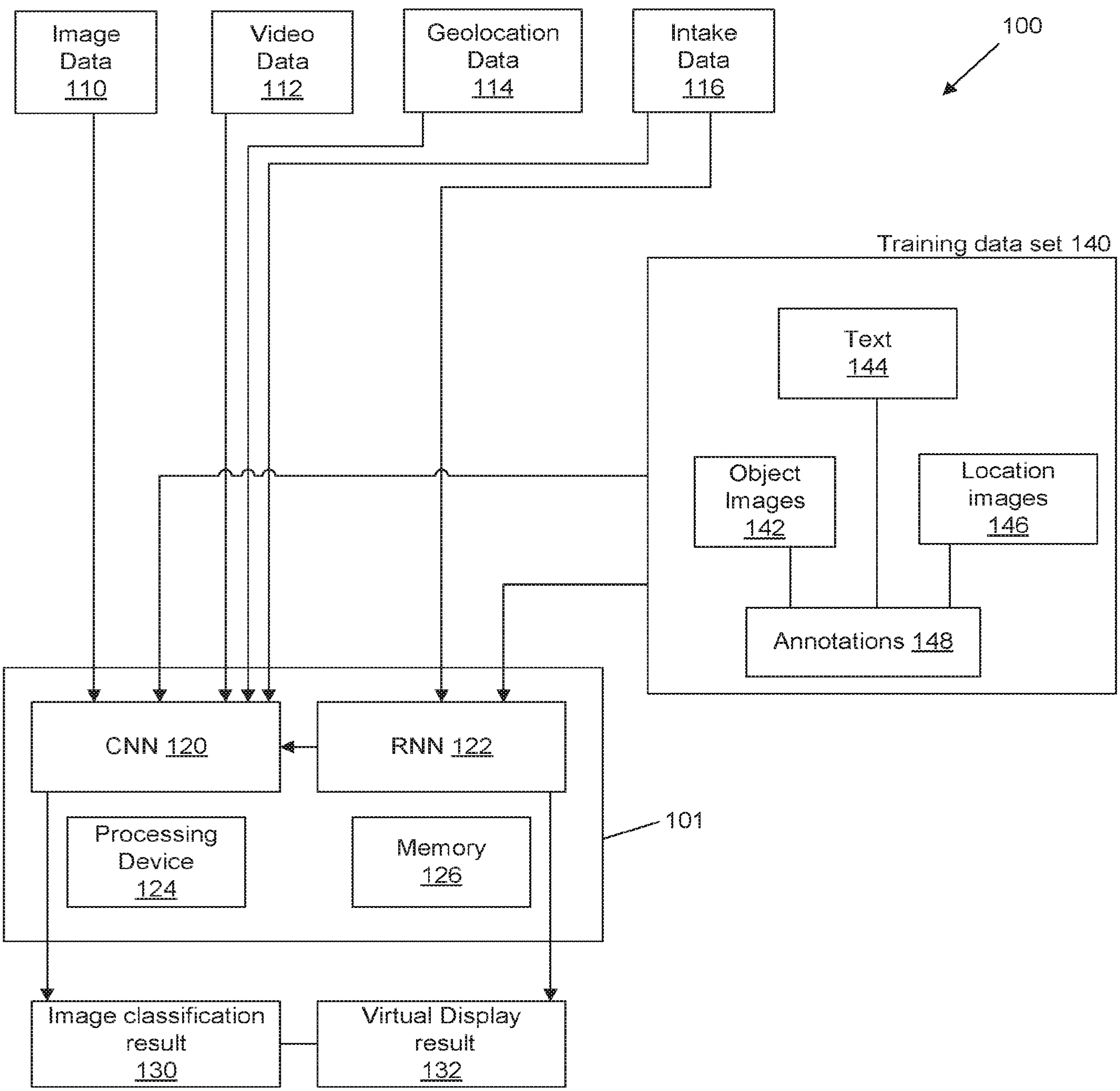

100
Image Data 110
Video Data 112
Geolocation Data 114
Intake Data 116
Training data set 140
Text 144
Object Images 142
Location images 146
Annotations 148
CNN 120
RNN 122
Processing Device 124
Memory 126
101
Image classification result 130
Virtual Display result 132

FIG. 1

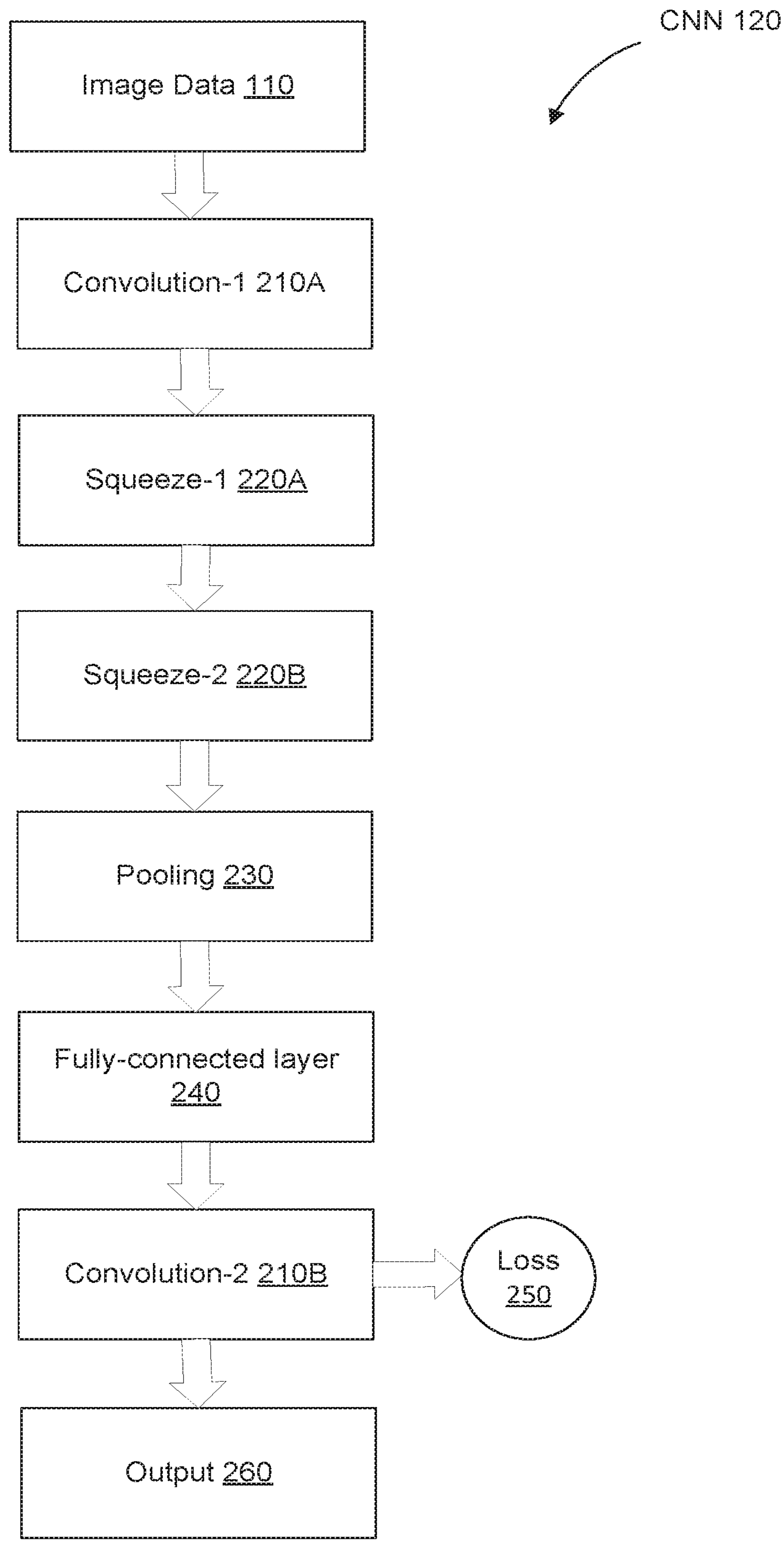

CNN 120
Image Data 110
Convolution-1 210A
Squeeze-1 220A
Squeeze-2 220B
Pooling 230
Fully-connected layer 240
Convolution-2 210B
Loss 250
Output 260

FIG. 2

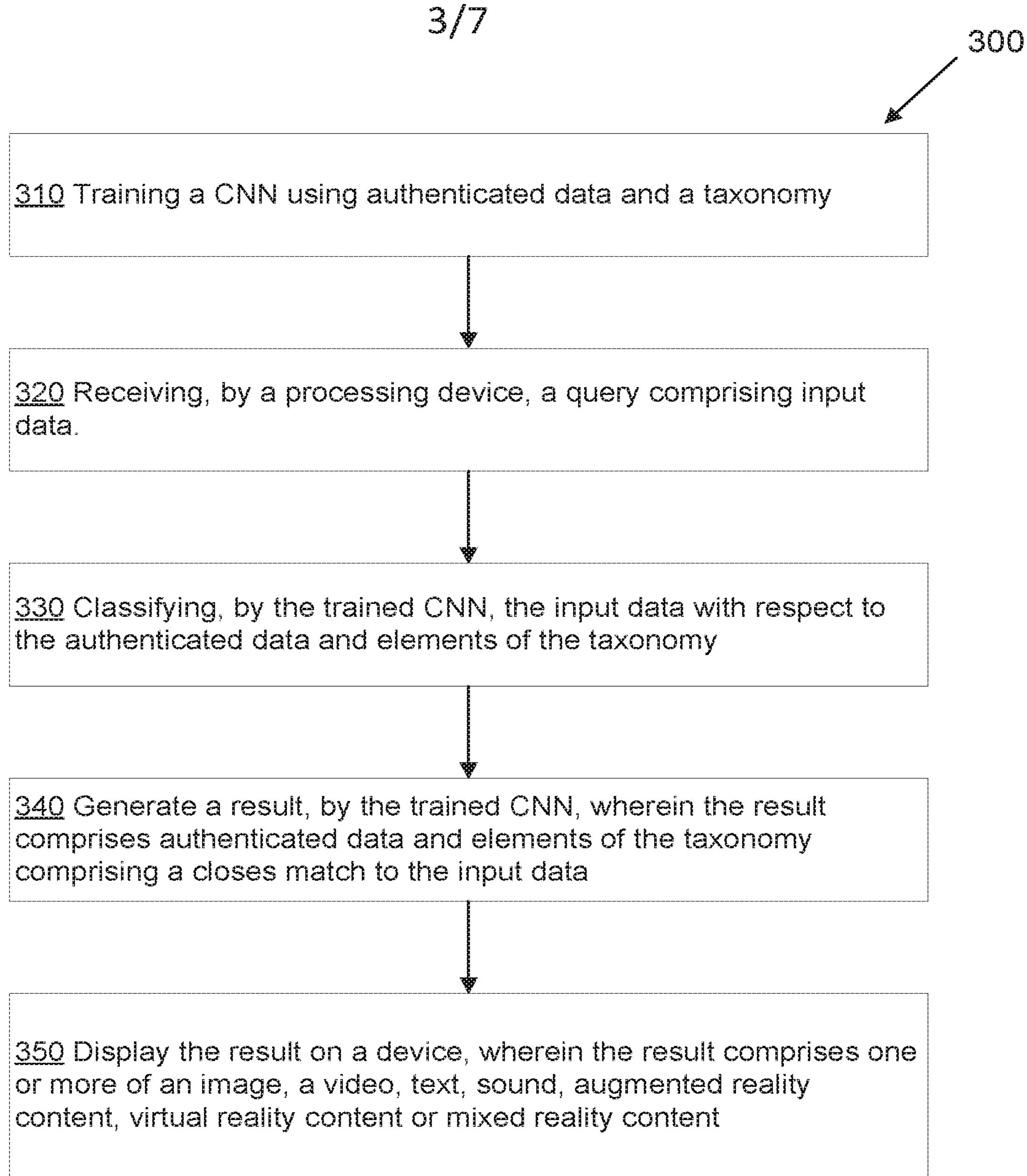

300
310 Training a CNN using authenticated data and a taxonomy
320 Receiving, by a processing device, a query comprising input data.
330 Classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy
340 Generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closes match to the input data
350 Display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content

FIG. 3

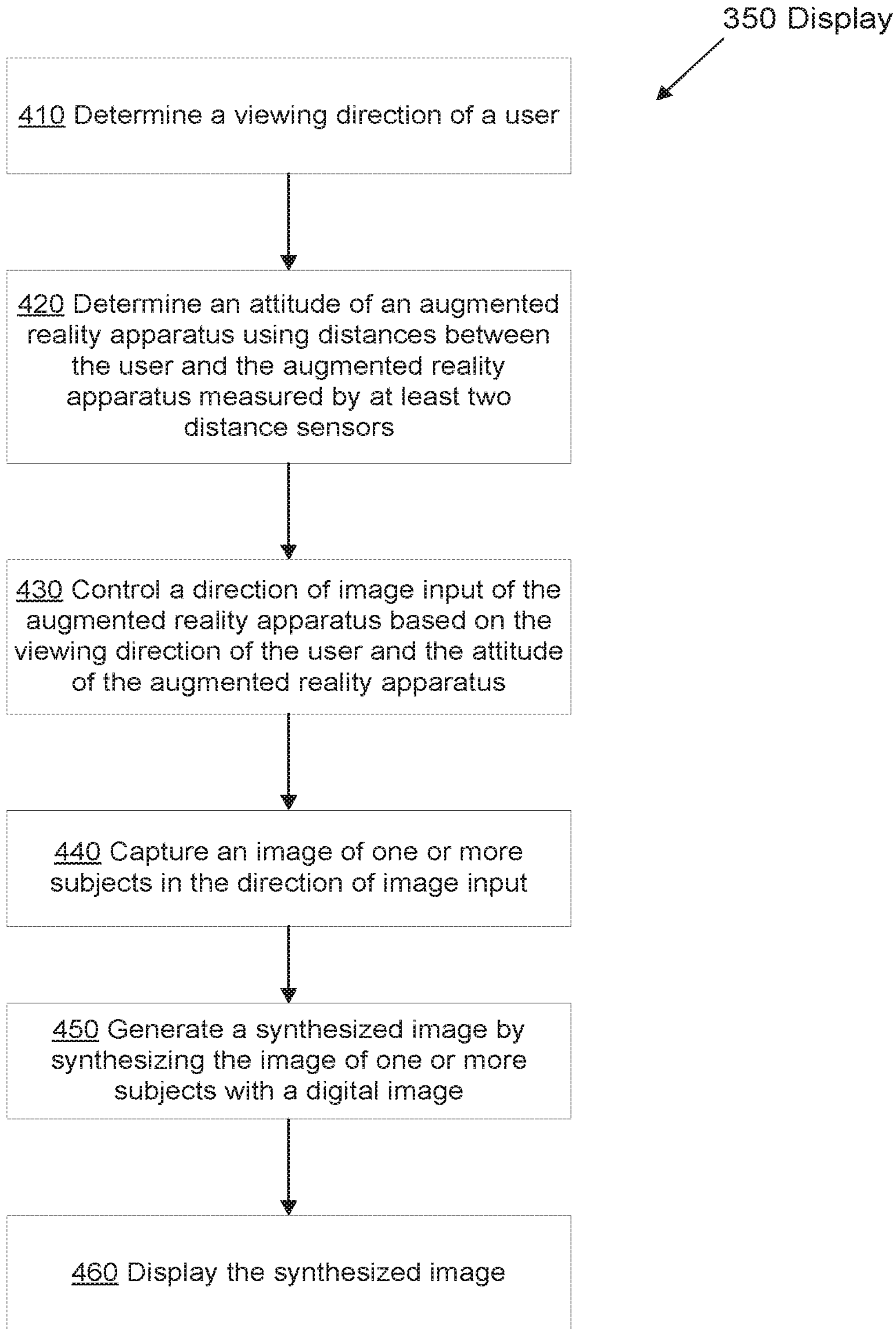

350 Display
410 Determine a viewing direction of a user
420 Determine an attitude of an augmented reality apparatus using distances between the user and the augmented reality apparatus measured by at least two distance sensors
430 Control a direction of image input of the augmented reality apparatus based on the viewing direction of the user and the attitude of the augmented reality apparatus
440 Capture an image of one or more subjects in the direction of image input
450 Generate a synthesized image by synthesizing the image of one or more subjects with a digital image
460 Display the synthesized image

FIG. 4

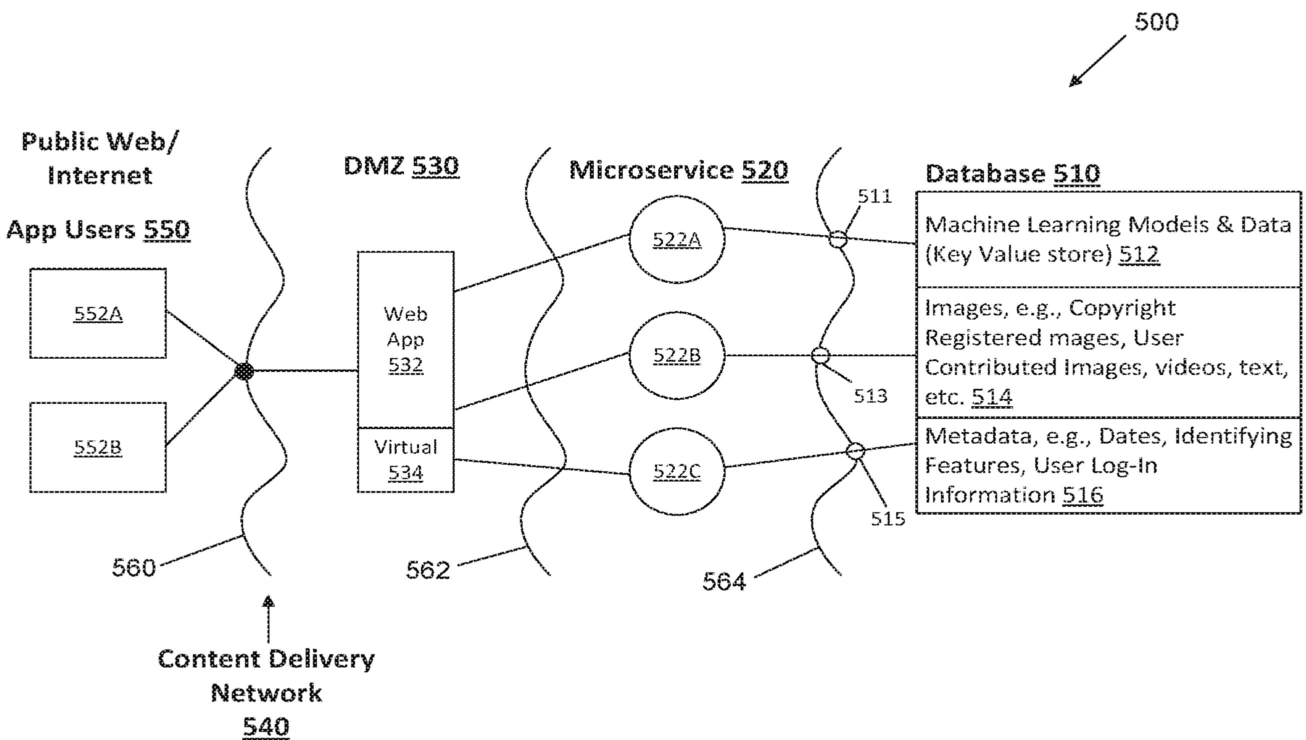

500
Public Web/ Internet
App Users 550
552A
552B
DMZ 530
Web App 532
Virtual 534
Microservice 520
522A
522B
522C
511
513
515
Database 510
Machine Learning Models & Data (Key Value store) 512
Images, e.g., Copyright Registered mages, User Contributed Images, videos, text, etc. 514
Metadata, e.g., Dates, Identifying Features, User Log-In Information 516
560
562
564
Content Delivery Network 540

FIG. 5

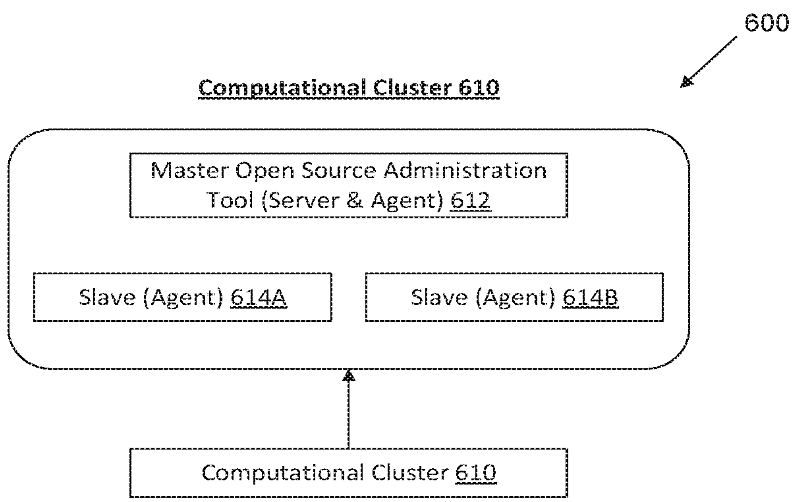

600
Computational Cluster 610
Master Open Source Administration Tool (Server & Agent) 612
Slave (Agent) 614A
Slave (Agent) 614B
Computational Cluster 610

FIG. 6

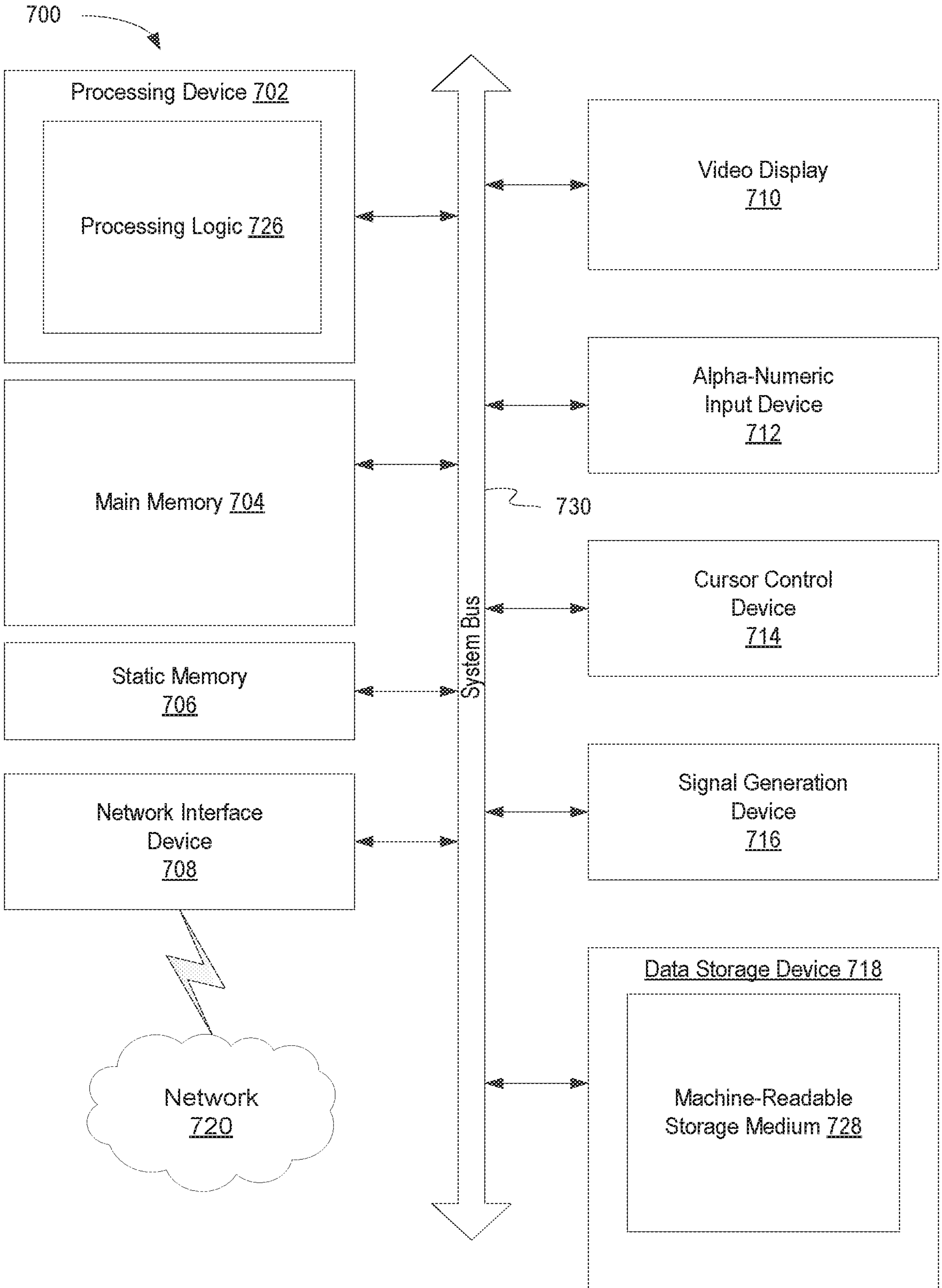

700
Processing Device 702
Processing Logic 726
Main Memory 704
Static Memory
706
Network Interface
Device
708
Network
720
System Bus
730
Video Display
710
Alpha-Numeric
Input Device
712
Cursor Control
Device
714
Signal Generation
Device
716
Data Storage Device 718
Machine-Readable
Storage Medium 728

FIG. 7

# CLAIMS

I/We claim:

1. A method comprising:

training a convolutional neural network (CNN) using authenticated data and a taxonomy;

receiving, by a processing device, a query comprising input data;

classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy;

generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and

displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

2. The method of claim 1, wherein the authenticated data comprises copyright registered works of authorship, metadata and text.

3. The method of claim 2, wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings.

4. The method of claim 3, wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information.

5. The method of claim 1, wherein the authenticated data comprises data from a copyright registered database.

6. The method of claim 1, wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture.

7. The method of claim 1, wherein the input data comprises one or more of image data, video data, intake data or geographical location data.

8. The method of claim 1, wherein classifying comprises mapping input data to authenticated data using the taxonomy.

9. The method of claim 1, wherein the result comprises one or more of an image, a video, text, or sound.

10. The method of claim 1, wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information.

11. The method of claim 1, wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy.

12. The method of claim 11, wherein the probability is determined by a cross-entropy function.

13. The method of claim 1, wherein the result comprises augmented reality content, wherein displaying the result comprises:

displaying the result in an augmented reality apparatus, comprising:

passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides;

imaging the light at the display device; and

displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data.

14. The method of claim 13, wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time.

15. The method of claim 1, further comprising training a recurrent neural network (RNN) using authenticated data and a taxonomy.

16. The method of claim 15, wherein the input data comprises unstructured data, the method further comprising:

processing, by the trained RNN, the unstructured data to yield structured data; and

classifying, by the trained CNN, the structured data.

17. The method of claim 1, wherein the input data comprises user uploaded data, the method further comprising authenticating the user uploaded data using Siamese Neural Networks and adding the authenticated user uploaded data to the authenticated data.

18. A system comprising:

a memory;

a processor, coupled to the memory, the processor configured to:

train a convolutional neural network (CNN) using authenticated data and a taxonomy;

receive, by a processing device, a query comprising input data;

classify, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy;

generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and

display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

19. The system of claim 18, wherein the authenticated data comprises copyright registered works of authorship, metadata and text.

20. A computer-readable non-transitory storage medium comprising executable instructions that, when executed by a computing device, cause the computing device to perform operations comprising:

training a convolutional neural network (CNN) using authenticated data and a taxonomy;

receiving, by a processing device, a query comprising input data;

classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy;

generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and

displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

# METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

## BACKGROUND

**[0001]** Cultural knowledge about one of mankind's most significant inventions, the automobile, is deeply fragmented and not easily searchable. Today, a user may see an object, such as a vehicle, driving down the road and wonder what it is, or an autonomous vehicle driving on a freeway may have a need to identify objects such as vehicles in its vicinity to learn information (e.g., vehicle make, model, year, stopping distance, etc.) and any equipment on board (e.g., communication devices, computer systems, etc.). Such information is currently unauthenticated, fragmented, dispersed, and not readily available in searchable form.

**[0002]** As society becomes increasingly digitalized, benefits have emerged in augmenting human intelligence with artificial intelligence (AI) to answer questions with curated knowledge. The market of automobile information is segmented and incoherent, with bits of information gathered by each segment without any relationship to each other and, thus, the information is not readily searchable. Each market segment has different needs, e.g., a user may have an instant need to identify certain vehicles, auto-related architecture and/or cultural artifacts out of interest whereas an autonomous vehicle may have a need to identify other vehicles in its vicinity and access their capabilities.

**[0003]** In 2018, the automotive advertising segment exceeded $38 trillion, exclusive of advertising for travel and food and automotive repair. The automotive advertising sector is the second largest advertising sector in the overall advertising marketplace. The largest customers are advertisers for new car buyers interested in the heritage of an automotive brand, the collectible car enthusiast market, the automotive parts market, insurance, travel, media archives and libraries with unidentified assets, and consumers with an unidentified photo album showing family vehicles. Additional commercial opportunities reside with government, security, law enforcement, and the entertainment industry.

**[0004]** Existing platforms are unable to sufficiently identify vehicles. These platforms can only make inferences from unauthenticated data. Anyone looking to authenticate a vehicle using these platforms can spend hours and still be uncertain of the exact make, model and year. These platforms have no verifiable source of automotive data and little hope of authoritatively identifying cars. It is difficult and frustrating to credibly authenticate automobiles by searching the Internet or asking vehicle owners; significant knowledge is fading from human memory.

**[0005]** There is a need for methods, systems and computer program products that allow a user to identify and search objects such as vehicles, while simultaneously building provenance and preserving knowledge around such objects and their cultural history, to instantly and properly identify objects by training artificial intelligence tools, to preserve cultural history, to promote historic places influenced by the vehicle's evolution, to establish automotive provenance based on first-hand historical records, to celebrate and educate users about objects (e.g., vehicles), to learn from history and help shape mobility's evolution, to explore human-computer interactions and to address the automotive advertising market. The systems, methods and computer program products described herein can quickly and accurately identify objects such as vehicles using authenticated data and provide an armature for social data to accrue around any object, building provenance around a subject that has lacked the tools to easily authenticate knowledge, e.g., cars and their impact on culture.

## BRIEF DESCRIPTION OF THE DRAWINGS

**[0006]** The present disclosure is illustrated by way of examples, and not by way of limitation, and may be more fully understood with references to the following detailed description when considered in connection with the figures, in which:

**[0007] FIG. 1** schematically illustrates an example system for image processing and data analysis, in accordance with one or more aspects of the present disclosure.

**[0008] FIG. 2** schematically illustrates an example structure of a convolutional neural network (CNN) that may be employed to process data input to an example system for media classification and identification, in accordance with one or more aspects of the present disclosure.

**[0009] FIG. 3** depicts a flow diagram of one illustrative example of a method 300 of data processing and data analysis, in accordance with one or more aspects of the present disclosure.

**[0010] FIG. 4** depicts a flow diagram of one illustrative example of a method 360 of displaying via augmented reality a result from an example system for media classification and identification.

**[0011] FIG. 5** depicts a diagram of a system for implementing the methods and systems described herein.

**[0012] FIG. 6** depicts a diagram of a computational cluster system that may be employed in an example system for media classification and identification, in accordance with one or more aspects of the present disclosure.

**[0013] FIG. 7** depicts a diagram of an illustrative example of a computing device implementing the systems and methods described herein.

## DETAILED DESCRIPTION

**[0014]** Described herein are methods, systems and computer program products for media classification and identification and for displaying results on a display (e.g., a cell phone, a monitor, an augmented reality apparatus, a mixed reality apparatus). While the systems and methods are described with respect to vehicles, vehicle artifacts and geographical locations, the systems and methods are broadly applicable to any objects. In example implementations, the systems and methods may relate to buildings (e.g., architecture), clothing, bridges, tools, highways, mountains, parks, rivers, cities, cars converted to homes and so on. Thus, the systems and methods described herein may be applied to a wide variety of physical objects that may involve various combinations of multiple imaging and/or other image capturing mechanisms.

**[0015]** The present disclosure overcomes the above-noted and other deficiencies by providing systems and methods for image processing and data analysis that may be utilized for identifying, classifying, researching and analyzing objects including, but not limited to vehicles, vehicle parts, vehicle artifacts, cultural artifacts, geographical locations, etc. To identify all of the objects in a photo, alone or in combination with a geographical location and/or a cultural heritage object, and then to associate a narrative with them represents a unique challenge. The converse of this too – the identification of historic places and objects (e.g. Statue of Liberty) alone, or in combination with vehicles— forms a broad descriptive visual narrative that illustrates innovative mapping from natural language processing (NLP) to multi-label image classification and identification.

**[0016]** In example implementations, a repository of photos, videos, keywords and captions of automobiles of proven provenance, with user narratives and comments, can be used to train a unique AI pipeline to map the information to a target space for image classification. For example, given an uploaded user image, the AI models may create the most appropriate summary of the relevant sections of the asset, and perform a multi-labeled classification of the image into the appropriate model of, for example, car manufacturer and year. In the case of an image containing multiple cars and cultural artifacts (e.g., glove compartments, spokes, steering wheels, vehicle lifts, etc.), there may be the additional task of establishing bounding-boxes around each of the recognized objects, and creating summary text appropriate to the image as a whole.

**[0017]** Likewise, the converse problem of taking a vehicle description, i.e., “Show me Prototypes”, and enriching it with AI assisted discovery into a proprietary database of high quality copyrighted images, represents a journey where the feature-vectors comprise the NLP embeddings of the narratives. The target space may be comprised of clusters of automotive

images that share attributes; for example, the query may map to a cluster of experimental cars from a particular decade. This may involve a single machine learning (ML) pipeline where RNN (LSTM/GRU) and BERT-derived attention models interact with CNN-architectures for image classification and Siamese Neural Networks (SNNs) for correct identifications. A collaborative user verification process involving crowd wisdom can be used to improve the accuracy of image-augmentation such that users can point out errors and suggest corrections. Should certain annotations be erroneous and users mark them so, such data will feed into the next round of neural architecture training.

**[0018]** In certain implementations, the systems and methods described herein may perform pixel-level analysis of images (and/or videos) in order to yield images, videos and/or virtual environments (e.g., augmented reality, mixed reality, virtual reality etc.) of vehicles, vehicle artifacts (e.g., images of vehicle tools, feature elements such as goggles, tachometers, wheel spokes, gas cans, etc.) and/or geographical locations. The systems and methods described herein may further determine whether images or videos input to the media processing system contain features that match one or more features of an image, video and/or geographical location of stored in a memory. In implementations, the systems and methods produce a result that comprises the closest matching data (e.g., having the highest probability score based on a cross-entropy function) identified in the training data set and/or database repository. The result may include an image of a vehicle together with text information about the vehicle such as a history, make, model, year, etc. The systems and methods may additionally yield historical information about one or more vehicle and/or geographical location and such information may be displayed in a virtual environment. In certain implementations, the systems and methods as described herein may also be implemented in an autonomous vehicle to capture images and/or video of surrounding vehicles on the road and to produce a result indicating the size, make, model and on-board equipment of the surrounding vehicles. In an example implementation, an autonomous vehicle incorporating the systems and methods described herein can be trained for "platooning." An example of "platooning" is where a vehicle operating in a self-driving or semi-autonomous mode, analyzes other vehicles in its vicinity to determine, for example, which vehicles may be capable of vehicle-to-vehicle (V2V) communication, other equipment on board, the estimated stopping distance of each vehicle and surrounding environmental objects such as children, balls, bicycles, tumbleweeds, etc. The autonomous vehicle may then communicate with the V2V vehicles to maintain a safe speed and distance to those vehicles, that is, the vehicles may move harmoniously together and may stop together at a traffic lights. Any vehicles in the vicinity that are not capable of V2V may be accounted for as an unknown variable. In other

implementations, platooning may involve recognition by the autonomous vehicle of structures that are capable of communicating with the vehicle in a vehicle-to-infrastructure (V2I) configuration. If the infrastructure is equipped with methods and systems as described herein, it may time or adjust the traffic lights to enhance platooning of V2I vehicles taking into consideration the variables of vehicles that are not equipped for V2I communication.

**[0019]** The systems and methods described herein have the benefit of being trained using a proprietary database comprising high-quality, digital copyrighted images of vehicles, vehicle artifacts and/or geographic sites (e.g., historical sites, cultural sites), such that the accuracy of the results produced by the disclosed systems and methods is improved over known methods of researching and analyzing vehicles, vehicle artifacts and/or geographical locations. The database may further include videos, embedded metadata and text. The database may itself be copyrighted. Because the database and data assets (e.g., images, videos, text, etc.) are themselves copyrighted, they form a body of authenticated data on which the neural networks can be trained.

**[0020]** The systems and methods described herein utilize a convolutional neural network (CNN) or a combination of both a CNN and a recurrent neural network (RNN), which form a part of a media processing system. The CNN may process one or more of image data (e.g., containing images, for example, of vehicles, vehicle artifacts, landscapes, etc.), video data (e.g., videos of vehicles, videos of historical sites, etc.), geolocation data (e.g., from a global positioning system) or intake data (e.g., text queries entered via a user interface, voice queries, natural language queries, etc.) to perform classification and with respect to vehicle information, vehicle artifact information, geographical location, etc. and/or to yield the probability of one or more image, video, geolocation and/or intake query matching significant features associated with a vehicle, a vehicle artifact and/or geographical location. The returned images, videos and/or virtual environment may be annotated and/or layered (e.g., overlaid, underlaid) with historical, design, mechanical, etc. information in the form of, for example, text, audio and video.

**[0021]** The RNN processes unstructured data, for example, natural language search queries and/or voice inputs to provide natural language processing (NLP). The unstructured data is transformed into structured data, which is fed to the CNN and processed as described above. In example implementations, the neural network architecture creates a hybridization of natural language processing (RNN) with image classification and identification techniques (CNN) for the purposes of preserving and accumulating data around a key subject area of historical and cultural interest.

**[0022]** A CNN is a computational model based on a multi-staged algorithm that applies a set of pre-defined functional transformations to one or more input (e.g., image pixels) and then utilizes the transformed data to perform, for example, classification, identification, image recognition, pattern recognition, etc. A CNN may be implemented as a feed-forward neural network (FFNN) in which the connectivity pattern between its neurons is inspired by the organization of the animal visual cortex. Individual cortical neurons respond to stimuli in a restricted region of space known as the receptive field. The receptive fields of different neurons partially overlap such that they tile the visual field. The response of an individual neuron to stimuli within its receptive field can be approximated mathematically by a convolution operation. In addition to image processing, the CNN may be used for other input types such as text, audio and video. In implementations, images may be input to a media processing system as described herein and the CNN processes the data. For example, if a user inputs a picture of a Ford Thunderbird automobile, the media processing system may output an image of a Ford Thunderbird together with the make, model, year, history and any known contextual information surrounding the photo and background.

**[0023]** In an illustrative example, a CNN may include multiple layers of various types, including convolution layers, non-linear layers (e.g., implemented by rectified linear units (ReLUs)), pooling layers, and classification (fully-connected) layers. A convolution layer may extract features from the input image by applying one or more learnable pixel-level filters to the input image. In an illustrative example, a pixel-level filter may be represented by a matrix of integer values, which is convolved across the dimensions of the input image to compute dot products between the entries of the filter and the input image at each spatial position, thus producing a feature map that represents the responses of the filter at every spatial position of the input image. The convolution filters are defined at the network training stage based on the training dataset to detect patterns and regions that are indicative of the presence of significant features within the input image.

**[0024]** A non-linear operation may be applied to the feature map produced by the convolution layer. In an illustrative example, the non-linear operation may be represented by a rectified linear unit (ReLU) which replaces with zeros all negative pixel values in the feature map. In various other implementations, the non-linear operation may be represented by a hyperbolic tangent function, a sigmoid function, or by other suitable non-linear function.

**[0025]** A pooling layer may perform subsampling to produce a reduced resolution feature map while retaining the most relevant information. The subsampling may involve averaging and/or determining maximum value of groups of pixels.

**[0026]** In certain implementations, convolution, non-linear, and pooling layers may be applied to the input image multiple times prior to the results being transmitted to a classification (fully-connected) layer. Together these layers extract the useful features from the input image, introduce non-linearity, and reduce image resolution while making the features less sensitive to scaling, distortions, and small transformations of the input image.

**[0027]** The output from the convolutional and pooling layers represent high-level features of the input image. The purpose of the classification layer is to use these features for classifying the input image into various classes. In an illustrative example, the classification layer may be represented by an artificial neural network that comprises multiple neurons. Each neuron receives its input from other neurons or from an external source and produces an output by applying an activation function to the sum of weighted inputs and a trainable bias value. A neural network may include multiple neurons arranged in layers, including the input layer, one or more hidden layers, and the output layer. Neurons from adjacent layers are connected by weighted edges. The term "fully connected" implies that every neuron in the previous layer is connected to every neuron on the next layer.

**[0028]** The edge weights are defined at the network training stage based on the training dataset. In an illustrative example, all of the edge weights are initialized to random values. For every input in the training dataset, the neural network is activated. The observed output of the neural network is compared with the desired output specified by the training data set, and the error is propagated back to the previous layers of the neural network, in which the weights are adjusted accordingly. This process is repeated until the output error is below a predetermined threshold.

**[0029]** The CNN may be implemented in a SNN configuration. A SNN configuration contains two or more identical subnetwork components. In implementations, not only is the architecture of the subnetworks identical, but the weights are shared among them as well. SNNs learn useful data descriptors, which may be used to compare the inputs (e.g., image data, video data, input data, geolocation data, etc.) of the subnetworks. For example, the inputs may be image data with CNNs as subnetworks.

**[0030]** The CNN may be implemented in a Generative Adversarial Network (GAN), which refers to two networks working together. A GAN can include any two networks (e.g., a combination of FFNNs and CNNs), with one tasked to generate content and the other tasked to judge content. The discriminating network receives either training data or generated content from the generative network. The ability of the discriminating network to correctly predict the data source is then used as part of the error for the generating network. This creates a form of competition where the discriminator gets better at distinguishing real data from generated data

and the generator learns to become less predictable to the discriminator. Even quite complex noise-like patterns can become predictable, but generated content similar in features to the input data is harder to learn to distinguish. The dynamics between the two networks need to be balanced; if prediction or generation becomes too good compared to the other, the GAN will not converge as there is intrinsic divergence.

**[0031]** A RNN may be described as a FFNN having connections between passes and through time. A RNN receives not just the current input it is fed, but also what it has perceived previously in time. In a RNN, neurons can be fed information not only from a previous layer, but from a previous pass. A string of text or picture can be fed one pixel or character at a time, so that the time dependent weights can be used for what came before in the sequence, not actually from what happened a specific time (e.g., x seconds) before. The RNN may be implemented as a Long Short-Term Memory (LFTM), which helps preserve the error, back-propagating it through layers and time. A LSTM includes information outside of the normal flow of the RNN in a gated cell. Information can be written to, stored in, or read from a cell, similar to data in a computer's memory. The cell can make decisions about when to allow reads, writes and erasures, and what to store via gates that open and close. These gates are analog, implemented with element-wise multiplication by sigmoids (i.e., all in the range of 0-1).

**[0032]** The RNN may be implemented with Bidirectional Encoder Representations from Transformers (BERT) to perform NLP tasks including *inter alia* question answering and natural language inference. BERT, which uses a Transformer-based language model, is a language representation model that provides accuracy for NLP tasks. A Transformer, in an encoding step, can use learned word embedding to convert words, in one-hot-vector form, into word embedding vectors; for each word-embedding vector, there is one output vector. BERT and its variants and Transformers, alone or in any combination with RNNs, are suitable for NLP tasks according to implementations herein.

**[0033]** Natural Language Processing (NLP) is the ability of a computer program to process and generate human language as spoken and/or written. In implementations, one or more recurrent neural network (RNN) is constructed to perform NLP (e.g., including text classification and text generation). In implementations, a neural network has layers, where each layer includes either input, hidden or output cells in parallel. In general two adjacent layers are fully connected (i.e., every neuron forms one layer to every neuron to another layer). For example, the network can have two input cells and one output cell, which can be used to model logic gates.

**[0034]** Augmented reality, which refers to a combination of a virtual environment and virtual reality, combines real-world images and virtual-world images such as computer graphic

images. Augmented reality is a semi-digital experience. In implementations of augmented reality, an image capturing device (e.g., a camera, a phone, a video recorder, etc.) that receives real images and a display device (e.g., a head mounted display that can display both real images and virtual images) are used together. Using augmented reality, a vehicle can be superimposed over a geographical location, for example, that can be associated with a particular date, time and/or weather condition. Lines can be drawn over an image of a vehicle to identify certain features and or parts, which may or may not be associated with a particular design type, time of history and/or cultural trend.

**[0035]** Virtual reality is a fully digital experience that creates an immersive environment where humans and computers can effectively interact and communicate by enhancing human-computer conversation skills using a variety of input and output techniques. Such techniques include the use of, for example, head-mounted displays, data gloves, or motion capture systems. These techniques receive data regarding variations in the position of a user by monitoring head, hand or other movements (e.g., position, directions, etc.), and transmit the data to a computer, which simulates (e.g., in a 3D coordinate space) the size and depth of an object within the viewing angle of the user.

**[0036]** Mixed reality refers to the merging of the real world with a virtual world to create a new environment where physical and digital objects interact with one another in real-time. In mixed reality, a real image can be captured using an image capturing device (e.g., a camera) and the direction the user faces within the environment is based on the captured real image. The relationship between the user's position and the position of a predetermined object is determined, and data obtained as a result of the calculation is displayed in a virtual space such that the data is laid over the captured real world image. Mixed reality is typically implemented using an image capturing device together with a display device.

**[0037] FIG. 1** schematically illustrates an example system 100 for image processing and data analysis, in accordance with one or more aspects of the present disclosure. As schematically illustrated by **FIG. 1**, the CNN 120 and optionally an RNN 122 together with a processing device 124 and a memory 126 form an media processing system 101. The media processing system 101 may be employed to process image data 110, video data 112, geolocation data 114 and input data 116 to produce an image classification result 130 and/or a virtual display result 132. The image data 110 may include one or more digital images, for example, captured by a camera or scanned, that may be stored in a memory. The video data 112 may include one or more digital videos, for example, captured by an audiovisual recording device or a dubbing device, that may be stored in a memory. The geolocation data 114 may include longitude,

latitude, country, region, historical or cultural place (e.g., Brooklyn Bridge), city, postal/zip code, time zone, way point, cell tower signal, etc. information from, for example, a global positioning system (GPS), entry to a user interface and/or other navigation system. The input data 116 may include structured data such as keyword search query input via a user interface and/or unstructured data input via the user interface. The unstructured data may include written or spoken natural language.

**[0038]** The CNN 120 may be employed to process the image data 110, the video data 112, the geolocation data 114 and the structured input data 116 to produce an image classification result 130 and/or a virtual display result 132, for example, with respect to vehicle information (e.g., a make, model, year, convertible sedan, sports utility vehicle or SUV, prototype, etc. ), vehicle artifacts (e.g., tools, a steering wheel, a lift, spokes, ets.) and/or a geographical location (e.g., a cultural site, a historical site, a landmark, etc.). The RNN 122 may process the unstructured data of the input data 116 (i.e., natural language processing) to produce structured data. The structured data may be fed from the RNN 122 to the CNN 120 for processing as described herein. In an illustrative example, the CNN 120 may correlate the image data 110, video data 112, geolocation data 114, input data 116 and structured data from the RNN 112 with images and data from a database (e.g., a propriety database of high quality copyrighted images and other works) in order to yield the probabilities of, for example, one or more image containing matching significant image features associated with a vehicle and/or a geographical location. The media processing system 101 may also return as part of the image classification result 130 historical, mechanical, cultural and other information with respect to the vehicle and/or geographical location.

**[0039]** The CNN 120 and the RNN 122 may be pre-trained using a comprehensive and precise training data set 140 that comprises *inter alia* non-published data, published data, images, videos, text (e.g., stories, news articles about various (e.g., thousands) vehicles, memoirs, out-of-print books, etc.) and/or geographical locations. For every vehicle or geographical location, the training data set may include multiple pixel-level, optionally annotated, vehicle images 142 and geographical images 146 and also may include related text 144 (e.g., histories, stories, descriptions, books, articles, drawings, sketches, etc.). The training data set 140 according to implementations herein may be a unique, comprehensive and proprietary body comprising copyrighted images, videos and text (e.g., history from the 20$^{th}$ Century surrounding vehicles, books off copyright, personal memoirs about vehicles, stories about racing, metals used, the brass era), built over many decades, that contains approximately 500,000 assets and verifies provenance through its records of timely copyright registrations at the Library of Congress. The

copyrighted works (i.e., authenticated data) identify *inter alia* automobiles and their impact on world culture including the time, place, historical context and significant background architecture captured in the media assets. Initially, the training data set 140 may be comprised in a copyrighted database where all of the assets contained within the training data set are copyrighted and thus authenticated. In further implementations, the training data set 140 may expand to include data input by users and further data assets that are not copyright registered, where such additional assets may be authenticated by other means whether scholarly (e.g., citations and research) or by using SNNs according to embodiments herein. The secondary twin will run a regression against the CNN. While the training data set 140 is comprehensive and precise when used in the systems and methods described herein, it will grow and evolve with more provenance authenticated data.

**[0040]** In certain implementations, training of the CNN 120 may involve activating the CNN 120 for every set of input images in the training dataset. The observed output (e.g., an image produced by the CNN 120) is compared with the desired output (e.g., the expected image) specified by the training data set, the error is calculated, and parameters of the CNN 120 are adjusted. This process is repeated until the output error is below a predetermined threshold.

**[0041]** In certain implementations, training of the RNN 122 may involve activating the RNN 120 for every set of unstructured data inputs in the training dataset. The observed output (e.g., a structured query produced by the CNN 120) is compared with the desired output (e.g., the expected query) specified by the training data set, the error may be calculated, and the parameters of the RNN 122 are adjusted accordingly. In implementations, this process may be repeated until the output error is below a predetermined threshold. In implementations using the RNN and deep learning models, the media processing system 101 may function and draw inferences from cross-related data.

**[0042]** The media processing system 101 may produce a virtual display result 132 in a virtual reality, augmented reality and/or mixed reality environment. In one illustrative example, one or more images, videos, descriptions, audio recordings, etc. may be layered onto an image captured by an image capture device (e.g., a camera, video recorder, etc.) and presented on a display. For example, a user may employ an image capture device (e.g., a cell phone) to capture an image in real time and the media processing system 101 may overlay or underlay images, videos and text onto the captured image as being viewed in an output device (e.g., a head-mounted display).

**[0043]** In one illustrative example, the CNN may be trained to identify automobiles. The media processing system 101 may process data 110, 112, 114, 116 to identify automobiles and preserve vehicle history by allowing a user to query the media processing system 101 to learn facts and

view photos of a specific vehicle. In implementations, the media processing system 101 may provide and provoke the curation of historical information surrounding the returned images (e.g., as a part of the image classification result 130). Multiple query types may be supported including, for example, photo uploads (CNN) and voice inputs (RNN) to query the models.

**[0044]** **FIG. 2** schematically illustrates an example structure of a CNN 120 that may be employed to process image data 110, video data 112, geolocation data 114 and input data 116 in order to produce an image classification result 130 and/or a virtual display result 132, in accordance with one or more aspects of the present disclosure. In certain implementations, acquired images may be pre-processed, e.g., by cropping, which may be performed in order to remove certain irrelevant parts of each frame. In an illustrative example, images having the resolution of 1024 x 1024 pixels may be cropped to remove 100-pixel wide image margins from each side of the rectangular image. In another illustrative example, a car may be outlined and isolated from noisy, non-contributory background elements.

**[0045]** As schematically illustrated by **FIG. 2**, the CNN 120 may include a first convolution layer 210A that receives image data 110 containing one or more images. The first convolution layer 210A is followed by squeeze layers 220A and 220B and a pooling layer 230, which is in turn followed by fully-connected layer 240 and a second convolution layer 210B. The second convolution layer 210B outputs one or more image 260 corresponding to the one or more input image of the image data 110 and may further produce the loss value 250 reflecting the difference between the produced data and the training data set. In certain implementations, the loss value may be determined empirically or set at a pre-defined value (e.g., 0.1).

**[0046]** In certain implementations, the loss value is determined as follows:

$$loss = \sum(x-y)^2\left(\frac{1}{2}+\max(x,y)\right),$$

where $x$ is the pixel value produced by the second convolution layer 210B and y is the value of the corresponding output image pixel.

**[0047]** Each convolution layer 210A, 210B may extract features from a sequence of input images from the input data 110, by applying one or more learnable pixel-level filters to a three-dimensional matrix representing the sequence of input images. The pixel-level filter may be represented by a matrix of integer values, which is convolved across the dimensions of the input image to compute dot products between the entries of the filter and the input image at each spatial position, to produce a feature map representing the responses of the first convolution layer 210A at every spatial position of the input image. In an illustrative example, the first convolution layer 210A may include 10 filters having the dimensions of 2 x 2 x 2. The second

convolution layer 210B may merge all the values produced by previous layers in order produce a matrix representing a plurality of image pixels.

**[0048]** **FIG. 3** depicts a flow diagram of one illustrative example of a method 300 of classifying and identifying input data, in accordance with one or more aspects of the present disclosure. Method 300 and/or each of its individual functions, routines, subroutines, or operations may be performed by one or more processors of the computer system (*e.g.*, system 100 and/or processing device 124 of **FIG. 1**) executing the method. In certain implementations, method 300 may be performed by a single processing thread. Alternatively, method 300 may be performed by two or more processing threads, each thread executing one or more individual functions, routines, subroutines, or operations of the method. In an illustrative example, the processing threads implementing method 300 may be synchronized (*e.g.*, using semaphores, critical sections, and/or other thread synchronization mechanisms). Alternatively, the processing threads implementing method 300 may be executed asynchronously with respect to each other.

**[0049]** At block 310, the processing device performing the method may train a CNN using authenticated data and a taxonomy. In implementations, the authenticated data may include copyright registered works of authorship including, but not limited to, copyrighted images, videos, text, stories, sketches, etc. The authenticated data may be stored in a database and the database itself may be copyright registered. The taxonomy may be used to classify and identify the data assets.

**[0050]** At block 320, the processing device may receive a query comprising input data. The input data can include, but is not limited to, image data, video data, intake data and/or geolocation data according to embodiments herein. In implementations, the intake data may be in the form of a keyword or string of text or may be in the form of unstructured data such as natural language either typed or spoken. The method may further include training an RNN to process the unstructured data of the intake data to form structured data suitable for processing by the CNN. The CNN may then process the structured data.

**[0051]** At block 330, the processing device may classify, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy. During the classification, the CNN may match features of the input data to one or more features of the authenticated data and/or elements of the taxonomy. For example, if the input data comprises an image, the CNN may scan the pixels of the image, identify features and then match the features with the closest matching features in the authenticated data and/or as classified in the taxonomy.

**[0052]** At block 340, the processing device may generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match

to the input data. For example, if five features have probabilities of 80%, 82%, 90%, 95% and 99% of matching five assets of the authenticated data, respectively, then the returned result may include only images with features having a 90% or greater probability of matching the input data.

**[0053]** At block 350, the processing device may display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content and/or mixed reality content. In implementations, the result may be layered with information. For example, a displayed image may be annotated with text, video and/or historical information about an object in the image.

**[0054]** In another illustrative example of a method 300 of classifying and identifying input data, in accordance with one or more aspects of the present disclosure. a processing device performing the method may process a training data set comprising a plurality of input images, in order to determine one or more parameters of a CNN to be employed for processing a plurality of images of one or more vehicle and/or geographical location. In various illustrative examples, the parameters of the CNN may include the convolution filter values and/or the edge weights of the fully-connected layer. In an illustrative example, the plurality of input images comprises one or more vehicle image. The one or more vehicle image may illustrate a vehicle alone or in combination with a geographical location (e.g., a Ford Model T on Route 66).

**[0055]** The processing device performing the method optionally may process a training data set comprising unstructured data in order to determine one or more parameters of a RNN to be employed for processing unstructured data input to the media processing system 101 in the form of natural language queries and voice queries to produce structured data for the CNN. In various illustrative examples, the RNN is trained to perform natural language processing using, for example, unstructured written and/or voice inputs.

**[0056]** The media processing system 101 may receive one or more of: a) image data including at least one input image (e.g., of a vehicle and/or a geographical location), b) video data including at least one input video (e.g., of a vehicle and/or a geographical location), c) input data including at least one of a keyword, a search query and unstructured data (e.g., relating to a vehicle and/or a geographical location), and d) geographical location data including a location of a device. In an illustrative example, the media processing system 101 may receive an image of an automobile alone, or together with a voice request saying "show me the artistic design features of this car."

**[0057]** The processing device performing the method optionally may process, by the RNN of the media processing system 101, any unstructured data of the input data that is received. The RNN outputs structured data that is fed to the CNN for processing. In the foregoing illustrative

example, the RNN may performing natural language processing of the voice request saying "show me the artistic design features of this car."

**[0058]** The processing device performing the method may process by the CNN of the media processing system 101, one or more of: i) the image data 110 to classify at least one input image (e.g., with respect to a vehicle information and/or a geographical location of a vehicle), ii) the video data 112 to classify at least one video (e.g., with respect to a vehicle information and/or a geographical location of a vehicle), iii) the structured input data 116 to classify at least one of a keyword or search query, iv) the structured data from the RNN (330), and v) the geographical location data 114 to produce one or more image, video and/or virtual display, as described in more herein. The probability of the image data, video data, geographical location data, input data and RNN data comprising the significant image features may be determined by a cross-entropy function, the error signal of which is directly proportional to a difference between desired and actual output values. In the foregoing illustrative example, the CNN may process the image of the automobile and the output of the RNN reflecting the voice request saying "show me the artistic design features of this car."

**[0059]** The processing device performing the method may generate a result by the media processing system including at least one of an image (e.g., of a vehicle and/or a geographical location), a video (e.g., of a vehicle and a geographical location), a history (e.g., of a vehicle and/or a geographical location) and/or other textual information. In the foregoing illustrative example, the media processing system 101 may generate an image of the automobile, alone or in combination with text providing the make, model and year of the automobile. The generated image may also be annotated with lines and text that identify artistic features of the automobile.

**[0060]** The processing device performing the method displays the result. The result may be displayed, for example, on a user device such as a cell phone, iPad, monitor or in a virtual device such as a head-mounted display of a virtual reality, augmented reality and/or mixed reality system.

**[0061] FIG. 4** depicts a flow diagram of one illustrative example of a method 360 of displaying a result, in accordance with one or more aspects of the present disclosure. Method 360 and/or each of its individual functions, routines, subroutines, or operations may be performed by one or more processors of the computer system (*e.g.*, system 100 and/or processing device 124 of **FIG. 1**) executing the method. In certain implementations, method 360 may be performed by a single processing thread. Alternatively, method 360 may be performed by two or more processing threads, each thread executing one or more individual functions, routines, subroutines, or operations of the method. In an illustrative example, the processing threads implementing

method 360 may be synchronized (*e.g.*, using semaphores, critical sections, and/or other thread synchronization mechanisms). Alternatively, the processing threads implementing method 300 may be executed asynchronously with respect to each other.

**[0062]** Example method 360 produces an augmented reality display on a display device. At block 410, the processing device performing the method determines a viewing direction of a user wearing and augmented reality apparatus, for example, a head-mounted display. The viewing direction may be determined by angles in relation to a center of the head set (e.g., looking forward) and head position.

**[0063]** At block 420, the processing device performing the method determines an attitude of an augmented reality apparatus using distances between the user and the augmented reality apparatus. The distances may be measured by one or more distance sensors.

**[0064]** At block 430, the processing device performing the method controls a direction of image input of the augmented reality apparatus based on the viewing direction of the user and the attitude of the augmented reality apparatus. In an illustrative example, the augmented reality apparatus may include a driving unit that adjusts the direction of, for example, a digital camera horizontally or vertically so that a subject (e.g., a vehicle) corresponding to a subject image incident upon the digital camera can be chosen even when the augmented reality apparatus is fixed.

**[0065]** At block 440, the processing device performing the method receives an image of one or more subjects (e.g., a vehicle) in the direction of image input. A camera or other device for recording and storing images may be used to capture the image.

**[0066]** At block 450, the processing device performing the method generates a synthesized image by synthesizing the image of the one or more subjects with a digital image. In an illustrative example, the synthesized image will layer images, videos and text with the image of the one or more subjects to produce an augmented reality environment.

**[0067]** At block 460, the processing device performing the method displays the synthesized image. The synthesized image may be displayed on an augmented reality apparatus such as a head-mounted display. In an illustrative example, a user may see the image of a car captured using the user’s cell phone underlaid by a video of geographical locations around the world. The image additionally or alternatively may be annotated with text such as arrows that point out different features of the car. The display may also be accompanied by voice information and/or music, for example, an audio description (e.g., spoken by a human or a bot) of the history of the car.

**[0068] FIG. 5** schematically illustrates an example of the neural network architecture and data pipeline together with a cloud-based, microservices-driven architecture (collectively referred to as "the architecture") 500 for image processing and data analysis, in accordance with one or more aspects of the present disclosure. As schematically illustrated by **FIG. 5**, the architecture 500 includes a memory (not shown) and a database 510 (e.g., MongoDB©, Hbase©) configured for both in-memory and on-disk storage. The database 510 may include one or more trained machine learning models 512 for classifying and identifying images. In implementations, a storage or persistence layer may store images, metadata as a multidimensional cube warehouse, ML models, textual narratives, search-indexes, and software/applications underlying a transactional database. The architecture 500 may further include a plurality of containerized microservices 522A-C. In various example implementations, the runtime logic execution layer may be a collection of docker-container based microservices exposing representational state transfer (REST) application programming interfaces (APIs) (e.g., using Kubernetes18©). Each element 511, 513 and 515 represents a taxonomy feature that the corresponding microservice 522A, B and C is trained to analyze.

**[0069]** System 500 may further include a web application 532 including, for example, the media processing system 101 and system 100 and configured to execute methods 300 and 360. The web application 532 may be stored on a demilitarized zone (DMZ) network 530 to provide security. A virtual memory component 534 comprised of user comments and ratings may also be stored on the DMZ network 530 (i.e., these comments will not be added to the training data set 140 until authenticated). System 500 may further include a content delivery network 540, which is a content distribution network of proxy servers to ensure high availability of content. System 500 may further include a web presentation layer where one or more app users 550 can access the web app 532 and view results on a user device, for example, a cell phone 552A or a laptop 552B. In various example implementations, a presentation layer (e.g., ReactJS©) may be used for rendering the web-presentation layer (e.g., in 19 HTML5/CSS). The architecture may be implemented in a cloud or in a decentralized network (e.g., SOLID via Tim Berners-Lee).

**[0070]** The architecture 500 may further include digital walls 560, 562, 564 providing cybersecurity layers between the components of the architecture 500. Wall 560 may be implemented between the public web and application user devices 550 and the web application 532 and virtual component 534 in the DMZ 530. Wall 562 may be implemented between the DMZ 530 and microservices 520, 522A-C. Wall 564 may be implemented between the microservices 520 and the database 510.

**[0071]** According to various example implementations of the methods and systems described herein, CNN models may be implemented with multi-label classifiers to identify, for example, the make, mode, and year of manufacture of a vehicle. These classifiers may be implemented in, for example, TensorFlow® and Keras® using using ResNet, VGG-19 and/or Inception. In example implementations, these will feed into densely connected layers that predict into a region of an NLP embedding space. This embedding-space may then be used with NLP to identify relevant textual artifacts. According to example implementations, trained SNNs including CNNs may be used for vehicle authentication. The SNNs may use a contrastive loss function to compare a sample to a reference/fingerprint object.

**[0072]** According to various example implementations of the methods and systems described herein, RNN models may be implemented with LSTM, gated recurrent unit (GRU) and attention models. For example, the narratives users contribute, in addition those already curated as authentic history, will feed into NLP models based on RNN (LSTM/GRU) and Attention models like BERT, to assist a user in finding automobiles through descriptions. CNN-recognized objects and their associated meta-tags may play a role in the NLP results to map onto vehicles.

**[0073]** According to various implementations, the media processing system 101 may achieve greater than about 75% accuracy, or greater than about 80% accuracy, or greater than about 85% accuracy, or greater than about 90% accuracy, or greater than about 95% accuracy, or greater than about 99% accuracy when compared against the multi-label classifier. According to example implementations, an accuracy rate of greater than 90% may be achieved for cars that are more popular or common. For rarer cars, an accuracy rate of greater than about 80% may be achieved for vehicles that are less common or have limited production. In various example implementations, vehicle-clusters may be determined from broad descriptions. For example, when a user provides a broad description such as "Show me all Ford Mustangs," the media processing system 101 may provide a greater than 90% accuracy in identifying or recognizing the vehicle described when the text is sufficiently descriptive. For cultural heritage images, in various implementations the media processing system 101 may provide an accuracy of greater than about 80%, or greater than about 85%, or greater than about 90%, or greater than about 95%, or greater than about 99%. For example, the media processing system 101 may return a result with greater than about 80% probability of matching the query.

**[0074]** **FIG. 6** depicts a diagram of a server configuration 600 that may be employed in an example system for image processing and data analysis, in accordance with one or more aspects of the present disclosure. The server configuration 600 may be a computational cluster 610 (e.g., a Hadoop Cluster) having a master open source administration tool server and agent 612 (e.g., an

Ambari server and Ambari agent). The computational cluster 610 may further include a pair of slave agents 614A-B. A Hadoop cluster is a type of computational cluster designed to store and analyze large quantities of unstructured data in a distributed computing environment. Such clusters run Hadoop's open source distributed processing software on low-cost commodity computers. The cluster enables many computers to solve problems requiring massive computation and data.

**[0075] FIG.** 7 illustrates a diagrammatic representation of a machine in the example form of a computer system 700 including a set of instructions executable by systems as described herein to perform any one or more of the methodologies discussed herein. In one implementation, the system may include instructions to enable execution of the processes and corresponding components shown and described in connection with **FIGs. 1-6**.

**[0076]** In alternative implementations, the systems may include a machine connected (e.g., networked) to other machines in a LAN, an intranet, an extranet, or the Internet. The machine may operate in the capacity of a server machine in client-server network environment. The machine may be a personal computer (PC), a neural computer, a set-top box (STB), Personal Digital Assistant (PDA), a cellular telephone, a server, a network router, switch or bridge, or any machine capable of executing a set of instructions (sequential or otherwise) that specify actions to be taken by that machine. Further, while only a single machine is illustrated, the term "machine" shall also be taken to include any collection of machines that individually or jointly execute a set (or multiple sets) of instructions to perform any one or more of the methodologies described herein.

**[0077]** The example computer system 700 can include a processing device (processor) 702, a main memory 704 (e.g., read-only memory (ROM), flash memory, dynamic random access memory (DRAM) such as synchronous DRAM (SDRAM)), a static memory 706 (e.g., flash memory, static random access memory (SRAM)), and a data storage device 718, which communicate with each other via a bus 730.

**[0078]** Processing device 702 represents one or more general-purpose processing devices such as a microprocessor, central processing unit, or the like. More particularly, the processing device 702 may be a complex instruction set computing (CISC) microprocessor, reduced instruction set computing (RISC) microprocessor, very long instruction word (VLIW) microprocessor, or a processor implementing other instruction sets or processors implementing a combination of instruction sets. The processing device 702 may also be one or more special-purpose processing devices such as an application specific integrated circuit (ASIC), a field programmable gate array (FPGA), a digital signal processor (DSP), network processor, or the like. In various implementations of the present disclosure, the processing device 702 is configured to execute

instructions for the devices or systems described herein for performing the operations and processes described herein.

**[0079]** The computer system 700 may further include a network interface device 708. The computer system 700 also may include a video display unit 710 (e.g., a liquid crystal display (LCD) or a cathode ray tube (CRT)), an alphanumeric input device 712 (e.g., a keyboard), a cursor control device 714 (e.g., a mouse), and a signal generation device 716 (e.g., a speaker).

**[0080]** The data storage device 718 may include a computer-readable medium 728 on which is stored one or more sets of instructions of the devices and systems as described herein embodying any one or more of the methodologies or functions described herein. The instructions may also reside, completely or at least partially, within the main memory 704 and/or within processing logic 726 of the processing device 702 during execution thereof by the computer system 700, the main memory 704 and the processing device 702 also constituting computer-readable media.

**[0081]** The instructions may further be transmitted or received over a network 720 via the network interface device 708. While the computer-readable storage medium 728 is shown in an example implementation to be a single medium, the term "computer-readable storage medium" should be taken to include a single medium or multiple media (e.g., a centralized or distributed database, and/or associated caches and servers) that store the one or more sets of instructions. The term "computer-readable storage medium" shall also be taken to include any medium that is capable of storing, encoding or carrying a set of instructions for execution by the machine and that cause the machine to perform any one or more of the methodologies of the present disclosure. The term "computer-readable storage medium" shall accordingly be taken to include, but not be limited to, solid-state memories, optical media, and magnetic media.

**[0082]** In various example implementations described herein, the neural networks are supervised learning models that analyze input data to classify objects, for example, using regression analysis. In an example implementation, a user may upload an image and the media processing system regresses it against many elements to determine the closest match. The supervised learning models are trained using different high-level elements of a taxonomy. The elements are related to categories of the taxonomy wherein the categories are used with ML to train the neural network models. In certain implementations, the elements may include, but are not limited to: actions (e.g., driving), concepts and emotions (e.g., direction), events (e.g., 2007 Tokyo Motor Show), geographic city (e.g., Los Angeles), geographic country (e.g., U.S.A.), geographic places (e.g., LAX Airport), geographic state (e.g., California), geographic location data (e.g., from a GPS), museum collections (e.g., Petersen Automotive Museum), photo environments (e.g., night), photo orientations (e.g., landscape), photo settings (e.g., auto garage), photo techniques (e.g., color),

photo views (e.g., three-quarter front view), signs (e.g., bowling alley), topic subjects (e.g., American culture), vehicle coachbuilder (e.g., Brewster & Co.), vehicle color (e.g., green), vehicle condition (e.g., new), vehicle manufacturer (e.g., including country and continent of origin), vehicle model (e.g., Bentley 6½ Liter Speed Six Tourer), vehicle parts (e.g., 8-track cassette player), vehicle quantity (e.g., one object), vehicle serial number (e.g., Chassis 3559SA), vehicle type (e.g., hydrogen fuel cell) and vehicle year of manufacture (e.g., 1957).

**[0083]** Implementations described herein can preserve and reveal information about vehicles and their impact on society. Using machine learning (ML) algorithms trained upon a proprietary dataset of curated and authenticated photos (image data), videos (video data), input data (e.g., text or voice inputs) and geolocation data, an artificial intelligence (AI) platform (e.g., including one or more convolutional neural network and one ore more recurrent neural network) has been developed that can *inter alia* identify vehicles from 1885 through present day and engages users to annotate images of the vehicles by sharing stories and comments. Machine learning includes, but is not limited to algorithms that find and apply patterns in data. Neural networks can be a form of ML. Implementations described herein provide a kind of time machine chassis capturing alchemical memories of shaped metal propelled through time and space, then identified through a multi-layered neural network. Enabling society to easily access information about vehicles through a searchable media processing system as described herein, augments and preserves human narrative, future transportation solutions and the history of remarkable vehicles.

**[0084]** In implementations, the database upon which the neural network is trained is a transdisciplinary study where abstract concepts (e.g., emotional, verbal, spatial, logical, artistic and social) are represented by semantic keywords expressing different dimensions of intelligence present in the referenced media object. For example, “Art Deco” is a semantic artistic keyword found on numerous vehicles from the 1920s and 1930s due to the visual design-language shown on a particular car or artifact. Using deep learning, the neural networks as described herein can be repeatedly trained for each of these distinct conceptual layers of intelligence found in the media thus resulting in object recognition enhanced through semantic intelligence and linking back to society. Additional databases including engineering information, racing results, car shows, location a car has been exhibited, valuations, etc. can be layered upon the proprietary vehicle database to further enhance information relating to the vehicles. In implementations, users can, for example, page through every Chevrolet Corvette in the library archives and read or listen to entries associated with any vehicle. Similarly, in implementations, a user can experience the development of streamlining or see the condensed vehicle design language of a particular decade of time, e.g. The Fifties. In implementations, a user can hold up a mobile device to an interesting car on the street and learn its story through interaction with the media processing system. For example, the media processing system may be configured to return facts such as “Did you know the V8 engine was invented one hundred years ago?”

**[0085]** Implementations described herein are configured to recognize and identify vehicles input via multiple sensory inputs (e.g., voice, image and text) and record a user’s personal stories about the provenance and significance of vehicles, for example, in telling the story of America. Families can upload shoeboxes of family photos to learn what vehicle a great-grandfather once drove to the Grand Canyon. A user can travel to a historic place site and through the media processing system view hundreds of vehicles and families that previously visited the site over preceding decades. In implementations, family vacation photographs recording an event can be layered upon an existing geographic location to create a virtual environment, for example, using an immersive augmented reality (AR). In implementations, the AR environment can enable a user to see herself or himself, along with his or her ancestors and their vehicles at the same cultural heritage site evoking a ghostly rapture within the time-space continuum. For example, “How many photographs with the family car were taken at the Golden Gate Bridge?”

**[0086]** According to implementations, a proprietary database of images of vehicles contains several labeled images that belong to different categories including, but not limited to “make,” “model” and “year.” “Vehicle” refers to a mechanism for transporting things (e.g., people and goods), including, but not limited to planes, trains, automobiles (e.g., cars, trucks, motorcycles, vans, etc.) and spacecraft. The more images used for each category, the better the model (e.g., the convolutional neural network) can be trained to determine whether an image is, for example, a Porsche image or a Ferrari image. This implementation utilizes supervised machine learning. The model can then be trained using the labeled known images. The images in their extracted forms enter the input side of the model and the labels are in the output side. The purpose is to train the model such that an image with its features coming from the input will match the label in the output. Once the model is trained, it can be used to recognize, classify and/or predict an unknown image. For example, a new image may be recognized, classified or predicted as a 1933 Packard Twelve. As part of the processing by the convolutional neural network, the newly input image also goes through the pixel feature extraction process.

**[0087]** Implementations disclosed herein can address a fundamental problem in the advertising industry: how to authenticate a users’ interest in a particular automotive brand or vehicle model or type. By virtue of a user uploading an unidentified Alfa Romeo to the platform, the user self-identifies the user’s interest in this vehicle brand. Through repeated interactions, the media processing system learns the user’s interests in vehicle information. Such a feedback loop is

valuable to advertisers for better targeting and in turn, can provide intelligence to manufacturers of future vehicles. The proprietary database of vehicles according to implementations, may be authenticated through timely registrations at the Library of Congress Copyright Office, which provides provenance that is preserved in the ML training dataset. In implementations, the training data set 140 may grow including additional data assets, for example, based on data input by users and/or additional assets that may not be copyright registered, but that may be authenticated, for example, by using SNNs as described herein.

**[0088]** Methods, systems and computer program products according to implementations herein relate to an AI platform built by systematically applying natural language processing (NLP) and computer vision, image and video processing to train a convolution and recurrent neural network from a dataset containing high quality, digital images, which may be copyrighted, of automobiles, capable of identifying a particular automobile from about 1885 through present day and into the future.

**[0089]** The essence of being human is to ask questions and AI seeks to provide credible information about a technological evolution: the journey of vehicles (e.g., the automobile), as well as the remaining surrounding artifacts of our vehicle heritage populating culture today. The implementations described herein provide an innovative AI-driven platform for classifying and documenting vehicles. Users can engage in a feedback cycle centered around identified photos, stories, comments, family photos and records. At their core, implementations herein nurture and explore the singular relationship of humans to machines, preserving the bond between the vehicles, design, community, architecture, engineering, history, government and culture to disseminate knowledge about vehicle culture. If a vehicle or important vehicle cultural heritage artifact is unknown, the platform can use the wisdom of crowds to identify and classify the asset. By NLP, the AI agent can begin chat interactions with the user about vehicles and immersive environments shown in the media thus deepening human-computer interaction skills.

**[0090]** Implementations described herein provide for the preservation and accessibility of collated, correlated and curated historical information (e.g., images, video, media, text, unstructured data, geographical location data, etc.) about and concerning vehicles, their use in human society, the environments (e.g., geographical locations) in which they are or have been used (e.g., racing and street), how they are or have been used, jobs they create or have created (e.g., in manufacturing and maintenance, consumer uses, collectors, etc.), technical and design features and special relationships with society.

**[0091]** According to implementations, multi-dimensional inputs query for vehicle attributes and elements from a vehicle dataset trained via ML from a proprietary reference database (e.g., a

neural network) built-upon copyright-registered and authenticated intellectual property about vehicles and their environments from the 1880s through present day and into the future to provide AI in mixed reality applications. Information may be retrieved from implementations described herein using multiple inputs including, but not limited to, audio, text, video, images and global positioning system (GPS) inputs where the input request is referenced against a proprietary-trained vehicle dataset and returns a classification and/or match to the input request in mixed-reality environments. Queries about vehicles can be answered with a probability of correct identification. For example, a user can type into the media processing system: “Auto Union Wanderer W-25” and the system would interpret the words to return an image of the “Auto Union Wanderer W-25.” The probability of the queried vehicle being built by “Auto Union” can be expressed as a percentage, for example, a “95%” probability of the image, video, text, history, etc. returned is an Auto Union and the probability of the image, video, text, history, etc. returned being a model “Wanderer W-25,” for example, “85%.” According to implementations, a short history of a vehicle appears and, using the geolocation services in the media processing system, an identification of where the closest example of this vehicle may be physically available for review relative to the user’s present location.

**[0092]** In implementations, information can be retrieved by uploading to the system (e.g., an app on a cell phone, a website, etc.), via a user interface, a photograph (e.g., a digital image, a scanned image, etc.) of a vehicle the user may encounter in daily life (e.g., on the street, at a car show, etc.). Using image recognition derived from the proprietary database, the media processing system can return, for example, a matching identification and/or classification of a vehicle make, model and year of release (referred to herein as “year”) with probabilities of accuracy based upon the trained dataset rendered through machine learning.

**[0093]** According to implementations described herein, users dictate voice request inputs, queries, text inputs and/or natural language to the media processing system. The input data can include, but is not limited to the make, model and year of a vehicle. For example, a user can speak “show me Ford Thunderbird” into a microphone that inputs data to the media processing system, which returns at least one of an image, a video, a written history, etc. representing the closest match to the “Ford Thunderbird” with additional information provided through mixed reality inputs. The user may refine a query by speaking “show me “red 1957 Ford Thunderbird” and the media processing system would return one or more image having the closest match together with a probability of accuracy. Existing platforms do not have provenance in their training data sets or information repositories (i.e., databases) and so the results returned by such platforms may be inaccurate or incomplete. Such platforms may scrape unauthenticated

information (e.g., not copyright registered, not validated with date, author, metadata, etc.) from publicly available websites such that if a certain number of users say something is, for example, a Thunderbird, then the platform will agree. However, if an asset or image is merely a replica of a Thunderbird, then such information is missed in these existing platforms. In another example, multiple years of Honda Accords may not be able to be recognized with authority by existing platforms.

**[0094]** According to implementations, a user initiates a query by pointing an input device (e.g., a camera, a video recorder, etc.) at a vehicle of interest and the media processing system receives the at least one input image and/or input video and matches it against the ML trained dataset to provide an Augmented Reality (AR) display of information regarding the queried vehicle. Levels of information may be chosen via user setup. For example, a user may only require vehicle make, such as “Ferrari,” or may require vehicle make and model, such as “Ferrari 250 GT,” or may require technical information like engine type, such as “V-8 engines;” the application is configured to return images, videos and/or information that matches V-8 engines from the neural network of information. According to implementations, in each query result, additional educational information about the vehicle is provided depending upon user settings. A brief history of the car can be displayed or overlaid in a mixed reality environment. For example, a user can submit a text or natural language input query, such as “two-tone vehicle interiors,” and matches to the requisition can be displayed on the user device with overlaid text depicting, for example, the make, model, history, design features, etc. of vehicles having two-tone interiors.

**[0095]** The implementations described herein are useful in a variety of fields, for example, where entities may need media assets or user information on vehicles obtainable from an organized, curated, and searchable platform. Example fields include, but are not limited to 1) advertisers: any automobile-related business or ads in which cars appear need authenticated product; 2) automobile manufacturers: marketing need for brand building, loyalty and heritage promotion of manufacturer’s products/services; 3) insurers: verifying vehicles is key in protecting assets and individuals; 4) entertainment: immersive experiences through augmented/virtual reality, skill games; 5) law enforcement: need help in identification of vehicles involved in investigations, possibly from photos taken at/from a crime scene—e.g., by a bystander on his/her cell phone—and fraud detection; 6) vehicle designers: need access to historical examples and perspective for new designs; 7) travel: roadside support, fuel, lodging, food, interesting roads and points of interest along the roadside; 8) classic car market and collectors: buyers, sellers and restorers of vehicles need parts authenticity, provenance

information, special features, and historical context in which the cars appear and are used; and 9) museums and archives: need help with identification, provenance and automotive history found in photo collections.

**[0096]** User interest, expressed by uploads of unidentified photos and by time spent reviewing certain vehicle brand archive sections to the media processing system, self-identifies the user's interest in specific vehicle brands and/or segments that can be sought after targets for advertisers. For example, a user who reads and peruses the Porsche archives is a good target for Porsche brand advertising.

**[0097]** By interacting with the platform, users self-identify interest in a particular automotive brand or vehicle sector thus solving an advertising problem for customers who wish to learn from past automotive designs, verify their illustrated marketing materials and target their communications to potential buyers in the automotive sector of our economy. Users may explore curated information about automobiles and roadside heritage through a virtual library linking other datasets to form a central integrated intelligence platform about automobiles and society. Transportation designers can easily access lessons learned from the last 135 years of automotive design.

**[0098]** Geolocation data, alone or in combination with curated photos of architectural and/or cultural heritage sites, can also be input to the media processing system as described herein. In certain implementations, the application can direct users to roadside services based on personalized user data (e.g., the user's preferred fuel type, fast food and hotel preferences can all be stored) and geolocation data received from a navigation system. For example, suppose a user drives a sports car, which is known from the user's profile stored in a memory accessible by the media processing system. The media processing system may have access to the user's calendar also stored on a memory accessible by the media processing system. The media processing system can receive an input from a navigation program indicating that the user's arrival at a calendar appointment location is estimated for 15 minutes, but there is a great two-lane road the user could drive that would be fun and would still get the user there on time for the calendar appointment. The media processing system would then suggest the alternate route.

**[0099]** Families throughout history have authenticated life in photographs, which may have been taken while on vacations using vehicles. Such image data can be used to virtually augment cultural heritage sites or historic places using historic photos including vehicles. For example, according to various implementations using augmented reality, virtual reality and/or mixed reality, the media processing system can enable users to virtually tour Route 66 throughout history. In implementations, the systems and methods described herein can use augmented,

virtual and/or mixed reality to enhance travel and road trips. For example, a user may drive down Route 66 and, using geolocation data, hold up a device (e.g., a cell phone) and see the present location as it evolved through history within a virtual display device (e.g., a head mounted device).

**[0100]** According to implementations, augmented reality relating to vehicles can be used for cultural heritage tourism to enhance the tourist experience. Linking contextual information found in the backgrounds of family photos, provides the groundwork for creating an authenticated augmented reality system, for example, for America's historic places. For example, implementations described herein are useful for auto clubs such as AAA. A mobile device can be pointed at a vehicle and/or cultural heritage location to capture image data and/or geolocation data and the media processing system can return images of the vehicle and/or cultural heritage location over time at that particular location.

**[0101]** Embodiments of the present disclosure can be described in view of the following clauses. In clause 1, a method comprises: training a convolutional neural network (CNN) using authenticated data and a taxonomy; receiving, by a processing device, a query comprising input data; classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy; generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

**[0102]** In clause 2, the method of clause 1, wherein the authenticated data comprises copyright registered works of authorship, metadata and text. In clause 3, the method of clause 2, wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings. In clause 4, the method of clause 3, wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information. In clause 5, the method of clause 1, wherein the authenticated data comprises data from a copyright registered database. In clause 6, the method of clause 1, wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture. In clause 7, the method of clause 1, wherein the input data comprises one or more of image data,

video data, intake data or geographical location data. In clause 8, the method of clause 1, wherein classifying comprises mapping input data to authenticated data using the taxonomy. In clause 9, the method of clause 1, wherein the result comprises one or more of an image, a video, text, or sound. In clause 10, the method of clause 1, wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information.

**[0103]** In clause 11, the method of clause 1, wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy. In clause 12, the method of clause 11, wherein the probability is determined by a cross-entropy function. In clause 13, the method of clause 1, wherein the result comprises augmented reality content, wherein displaying the result comprises: displaying the result in an augmented reality apparatus, comprising: passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device; and displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data. In clause 14, the method of clause 13, wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time. In clause 15, the method of clause 1, further comprising training a recurrent neural network (RNN) using authenticated data and a taxonomy. In clause 16, the method of clause 15, wherein the input data comprises unstructured data, the method further comprising: processing, by the trained RNN, the unstructured data to yield structured data; and classifying, by the trained CNN, the structured data. In clause 17, the method of clause 1, wherein the input data comprises user uploaded data, the method further comprising authenticating the user uploaded data using SNNs and adding the authenticated user uploaded data to the authenticated data.

**[0104]** In clause 18, a system comprising: a memory; a processor, coupled to the memory, the processor configured to: train a convolutional neural network (CNN) using authenticated data and a taxonomy; receive, by a processing device, a query comprising input data; classify, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy; generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content. In clause 19, the system of clause

18, wherein the authenticated data comprises copyright registered works of authorship, metadata and text. In clause 20, the system of clause 19, wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings. In clause 21, the system of clause 20, wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information. In clause 22, the system of clause 18, wherein the authenticated data comprises data from a copyright registered database. In clause 23, the system of clause 18, wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture. In clause 24, the system of clause 18, wherein the input data comprises one or more of image data, video data, intake data or geographical location data. In clause 25, the system of clause 18, wherein classifying comprises mapping input data to authenticated data using the taxonomy. In clause 26, the system of clause 18, wherein the result comprises one or more of an image, a video, text, or sound. In clause 27, the system of clause 18, wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information. In clause 28, the system of clause 18, wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy. In clause 29, the system of clause 28, wherein the probability is determined by a cross-entropy function. In clause 30, the system of clause 18, wherein the result comprises augmented reality content, wherein displaying the result comprises: displaying the result in an augmented reality apparatus, comprising: passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device; and displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data. In clause 31, the system of clause 30, wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time. In clause 32, the system of clause 18, further configured to train a recurrent neural network (RNN) using authenticated data and a

taxonomy. In clause 33, the method of clause 32, wherein the input data comprises unstructured data, the method further comprising: processing, by the trained RNN, the unstructured data to yield structured data; and classifying, by the trained CNN, the structured data. In clause 34, the method of clause 18, wherein the input data comprises user uploaded data, wherein the system is further configured to authenticate the user uploaded data using SNNs and add the authenticated user uploaded data to the authenticated data.

**[0105]** In clause 35, a computer-readable non-transitory storage medium comprising executable instructions that, when executed by a computing device, cause the computing device to perform operations comprising: training a convolutional neural network (CNN) using authenticated data and a taxonomy; receiving, by a processing device, a query comprising input data; classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy; generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content. In clause 36, the computer-readable non-transitory storage medium of clause 35, wherein the authenticated data comprises copyright registered works of authorship, metadata and text. In clause 37, the computer-readable non-transitory storage medium of clause 36, wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings. In clause 38, the computer-readable non-transitory storage medium of clause 37, wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information. In clause 39, the computer-readable non-transitory storage medium of clause 35, wherein the authenticated data comprises data from a copyright registered database. In clause 40, the computer-readable non-transitory storage medium of clause 35, wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture. In clause 41, the computer-readable non-transitory storage medium of clause 35, wherein the input data comprises one or more of image data, video data, intake data or geographical location data. In clause 42, the computer-readable non-transitory storage medium of clause 35, wherein classifying comprises mapping input data to authenticated data using the

taxonomy. In clause 43, the computer-readable non-transitory storage medium of clause 35, wherein the result comprises one or more of an image, a video, text, or sound. In clause 44, the computer-readable non-transitory storage medium of clause 35, wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information. In clause 45, the computer-readable non-transitory storage medium of clause 35, wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy. In clause 46, the computer-readable non-transitory storage medium of clause 45, wherein the probability is determined by a cross-entropy function. In clause 47, the computer-readable non-transitory storage medium of clause 35, wherein the result comprises augmented reality content, wherein displaying the result comprises: displaying the result in an augmented reality apparatus, comprising: passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device; and displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data. In clause 48, the computer-readable non-transitory storage medium of clause 47, wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time. In clause 49, the computer-readable non-transitory storage medium of clause 35, further comprising training a recurrent neural network (RNN) using authenticated data and a taxonomy. In clause 50, the computer-readable non-transitory storage medium of clause 49, wherein the input data comprises unstructured data, the method further comprising: processing, by the trained RNN, the unstructured data to yield structured data; and classifying, by the trained CNN, the structured data. In clause 51, the computer-readable non-transitory storage medium of clause 35, wherein the input data comprises user uploaded data, the method further comprising authenticating the user uploaded data using SNNs and adding the authenticated user uploaded data to the authenticated data. In clause 52, the medium of clause 51, wherein processing by the CNN at least one of the image data, the video data, the input data and the geographical location data yields one or more image of a vehicle comprising matching features.

**[0106]** The preceding description sets forth numerous specific details such as examples of specific systems, components, methods, and so forth, in order to provide a good understanding of several implementations of the present disclosure. It will be apparent to one skilled in the art, however,

that at least some implementations of the present disclosure may be practiced without these specific details. In other instances, well-known components or methods are not described in detail or are presented in simple block diagram format in order to avoid unnecessarily obscuring the present disclosure. Thus, the specific details set forth are merely presented as examples. Particular implementations may vary from these example details and still be contemplated to be within the scope of the present disclosure. In the above description, numerous details are set forth.

**[0107]** It will be apparent, however, to one of ordinary skill in the art having the benefit of this disclosure, that implementations of the disclosure may be practiced without these specific details. In some instances, well-known structures and devices are shown in block diagram form, rather than in detail, in order to avoid obscuring the description.

**[0108]** Some portions of the detailed description are presented in terms of algorithms and symbolic representations of operations on data bits within a computer memory. These algorithmic descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical, magnetic, or optical signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

**[0109]** It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the above discussion, it is appreciated that throughout the description, discussions utilizing terms such as "rating," "selecting," "comparing," "adjusting," or the like, refer to the actions and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (e.g., electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

**[0110]** Implementations of the disclosure also relate to an apparatus for performing the operations herein. This apparatus may be specially constructed for the required purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, and

magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, or any type of media suitable for storing electronic instructions.

**[0111]** The algorithms and displays presented herein are not inherently related to any particular computer or other apparatus. Various general purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct a more specialized apparatus to perform the required method steps. The required structure for a variety of these systems will appear from the description below. In addition, the present disclosure is not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the disclosure as described herein.

**[0112]** It is to be understood that the above description is intended to be illustrative, and not restrictive. Many other implementations will be apparent to those of skill in the art upon reading and understanding the above description.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) **World Intellectual Property Organization**
International Bureau

(43) **International Publication Date**
**12 November 2020 (12.11.2020)**

WIPO | PCT

(10) **International Publication Number**
**WO 2020/227651 A1**

(51) **International Patent Classification:**
*G06N 20/00* (2019.01) *G06K 9/62* (2006.01)
*G06K 9/46* (2006.01) *G06N 3/08* (2006.01)

(21) **International Application Number:**
PCT/US2020/032149

(22) **International Filing Date:**
08 May 2020 (08.05.2020)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
62/845,546 09 May 2019 (09.05.2019) US

(71) **Applicant: AUTOMOBILIA II, LLC** [US/US]; 301 Clematis Street #3000, West Palm Beach, FL 33401 (US).

(72) **Inventor: LEWIS, Lucinda**; 414 E. Valencia Avenue, #103, Burbank, CA 91505 (US).

(74) **Agent: NICHOLES, Mary, K.** et al.; Lowenstein Sandler LLP, One Lowenstein Drive, Roseland, NJ 07068 (US).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

**Published:**
— *with international search report (Art. 21(3))*

(54) **Title:** METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

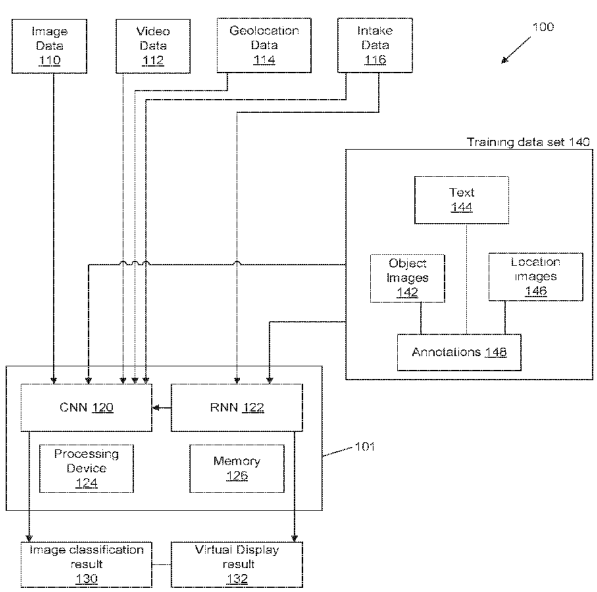



FIG. 1

(57) **Abstract:** The present disclosure relates generally to methods, systems and computer program products for classifying and identifying input data using neural networks and displaying results (e.g., images of vehicles, vehicle artifacts and geographical locations dating from the 1880s to present day and beyond). The results may be displayed on displays or in virtual environments such as on virtual reality, augmented reality and/or mixed-reality devices.

WO 2020/227651 A1

# PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To: MARY K. NICHOLES
LOWENSTEIN SANDLER LLP
ONE LOWENSTIEN DRIVE
ROSELAND, NJ 07068

# PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

| Date of mailing (*day/month/year*) | 04 AUG 2020 |
|---|---|

| Applicant's or agent's file reference | FOR FURTHER ACTION |
|---|---|
| 28108-110 | See paragraph 2 below |

| International application No. | International filing date (*day/month/year*) | Priority date (*day/month/year*) |
|---|---|---|
| PCT/US 20/32149 | 08 May 2020 (08.05.2020) | 09 May 2019 (09.05.2019) |

International Patent Classification (IPC) or both national classification and IPC

IPC - G06N 20/00, G06K 9/46, G06K 9/62, G06N 3/08 (2020.01)

CPC - G06N 20/00, G06K 9/629, G06N 3/0454, G06N 3/08, G06K 9/62

Applicant AUTOMOBILIA II, LLC

1. This opinion contains indications relating to the following items:

- [x] Box No. I Basis of the opinion
- [ ] Box No. II Priority
- [ ] Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- [ ] Box No. IV Lack of unity of invention
- [x] Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step and industrial applicability; citations and explanations supporting such statement
- [ ] Box No. VI Certain documents cited
- [ ] Box No. VII Certain defects in the international application
- [ ] Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

| Name and mailing address of the ISA/US | Date of completion of this opinion | Authorized officer |
|---|---|---|
| Mail Stop PCT, Attn: ISA/US<br>Commissioner for Patents<br>P.O. Box 1450, Alexandria, Virginia 22313-1450<br>Facsimile No. 571-273-8300 | 01 July 2020 (01.07.2020) | Lee Young<br>PCT Help Desk<br>Telephone No. 571-272-4300 |

**WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY**

International application No.
PCT/US 20/32149

**Box No. I Basis of this opinion**

1. With regard to the **language**, this opinion has been established on the basis of:

☒ the international application in the language in which it was filed.

☐ a translation of the international application into ______________ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

2. ☐ This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(b)).

3. ☐ With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of a sequence listing:

   a. ☐ forming part of the international application as filed:

      ☐ in the form of an Annex C/ST.25 text file.

      ☐ on paper or in the form of an image file.

   b. ☐ furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.

   c. ☐ furnished subsequent to the international filing date for the purposes of international search only:

      ☐ in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).

      ☐ on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

4. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

**WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY**

International application No.
PCT/US 20/32149

**Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step and industrial applicability; citations and explanations supporting such statement**

1. Statement

| | | | |
|---|---|---|---|
| Novelty (N) | Claims | 12-17 | YES |
| | Claims | 1-11, 18-20 | NO |
| Inventive step (IS) | Claims | None | YES |
| | Claims | 1-20 | NO |
| Industrial applicability (IA) | Claims | 1-20 | YES |
| | Claims | None | NO |

2. Citations and explanations:

Claims 1-11 and 18-20 lack novelty under PCT Article 33(2) as being anticipated by US 2018/0084310 A1 to GumGum, Inc. (hereinafter 'GumGum').

As to claim 1, GumGum teaches a method comprising: training a convolutional neural network (CNN) using authenticated data and a taxonomy (para [0018]-[0019], [0025]-[0027], [0080] - "A common existing use of video fingerprints is in the field of Digital Rights Management ("DRM") in order for content owners (such as film studios or publishers) to identify when files containing their copyrighted video content are uploaded to video sharing websites or other file sharing networks"; "An image object that is the subject of a classifier that has been created or trained to identify that particular image object may be referred to as a "target image object." For example, a target image object may be a visual representation of a company's logo, and a classifier may be generated specifically to identify at least that logo. In some embodiments, the target image object may generally refer to a class or group of related image objects that may be identified using a particular classifier"; "these neural networks may be trained using video data, but may then be used with either video data or image data as input. In one embodiment, block 612 may be implemented at least in part using a convolutional neural network"); receiving, by a processing device, a query comprising input data (para [0018]-[0019], [0025]-[0027], [0108] - "Video fingerprinting is generally used to compactly represent a video or portion of a video in a manner that allows for efficient searching of other videos to identify either duplicates or slight variations of the same video. For example, computer software may identify, extract and compress characteristics of video data to create a digital file that serves as a "fingerprint" of the specific video data"; "a computing system may analyze a live event broadcast for brand exposures (e.g., instances where a company name, logo or other visual indicator is present in a frame of broadcasted video). For each video sequence that contains a brand exposure, the system may create a digital fingerprint to identify a portion of the video in which the brand exposure occurs. The system may subsequently search video data of potentially hundreds or thousands of subsequent broadcasts across different channels, websites, streaming services, or social media networks to identify occurrences of the digital fingerprint"; "system can query the API with an input parameter such as a keyword, tag, or category, and/or to retrieve media item result sets"); classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy (para [0018]-[0019], [0025]-[0027], [0080]-[0081] - "the computing system at block 612 may use classification models that have each been trained"; "the computing system may access neural networks or other machine learning models that have been trained to identify objects associated with the relevant sponsor(s) for the given team(s) and/or the home venue, then may proceed to sponsor logo identification"); generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data (para [0025]-[0027], [0031], [0087] - "trained image classifier may return a certain probability that a given frame of the video includes a target image object (such as a specific logo), and the system may deem those frames associated with a probability that satisfies a certain threshold likelihood (e.g., 75%) as including the target image object"; "one or more classifiers may provide a separate probability for each of the different companies and/or for each different logo"; "the system may determine confidence levels based on how many pixels the original video portion and subsequent video portion have in common (or pixels that are substantially the same, accounting for slight variations), then based on this confidence level may perform image recognition to verify that a similar video frame is in fact a positive match"; "to supplement and/or confirm the confidence level determinations of the image and video classification models discussed above, the computing system may use text classification models applied to text of a social media post or page in which the image or video is embedded or linked. For example, a textual classifier may provide a confidence score indicating how confident the model is that a particular social media post relates to a specific rights holder and/or sponsor based on the text of the post. This confidence score may be weighted and combined with a corresponding confidence score determined from the image or video itself"); and displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content (para [0036], [0051], [0085] - "the system may generate various spreadsheets, summary views, dashboards, reports, user interfaces and/or other output intended for consumption by a sponsor or other user. The data in one embodiment may include, for each of a number of different URLs, videos, images or social media posts: an indication whether the account is owned and operated by the brand"; "media report may include information for a number of different sponsors' logos, and may be based on the results of multiple passes through the illustrative method").

As to claim 2, GumGum further teaches wherein the authenticated data comprises copyright registered works of authorship, metadata and text (para [0018]-[0019], [0086]-[0087]).

As to claim 3, GumGum further teaches wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings (para [0018]-[0019], [0086]-[0087]).

(See Continuation Box)

**Supplemental Box**

In case the space in any of the preceding boxes is not sufficient.
Continuation of:
Box V.2. Citations and explanations:

As to claim 4, GumGum further teaches wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information (para [0034], [0098]).

As to claim 5, GumGum further teaches wherein the authenticated data comprises data from a copyright registered database (para [0018]-[0019], [0090]).

As to claim 6, GumGum further teaches wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture (para [0025], [0074], [0084], [0098]-[0099]).

As to claim 7, GumGum further teaches wherein the input data comprises one or more of image data, video data, intake data or geographical location data (para [0071]).

As to claim 8, GumGum further teaches wherein classifying comprises mapping input data to authenticated data using the taxonomy (para [0025]-[0027], [0083]-[0087]).

As to claim 9, GumGum further teaches wherein the result comprises one or more of an image, a video, text, or sound (para [0036], [0051], [0085], [0098]-[0099]).

As to claim 10, GumGum further teaches wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information (para [0036], [0051], [0085], [0098]-[0099]).

As to claim 11, GumGum further teaches wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy (para [0024]-[0027], [0031]).

As to claim 18, GumGum teaches a system comprising: a memory; a processor, coupled to the memory (para [0038]-[0042]), the processor configured to: train a convolutional neural network (CNN) using authenticated data and a taxonomy (para [0018]-[0019], [0025]-[0027], [0080]); receive, by a processing device, a query comprising input data (para [0018]-[0019], [0025]-[0027], [0108]); classify, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy (para [0018]-[0019], [0025]-[0027], [0080]-[0081]); generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data (para [0025]-[0027], [0031], [0087]); and display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content (para [0036], [0051], [0085]).

As to claim 19, GumGum further teaches wherein the authenticated data comprises copyright registered works of authorship, metadata and text (para [0018]-[0019], [0086]-[0087]).

As to claim 20, GumGum teaches a computer-readable non-transitory storage medium comprising executable instructions that, when executed by a computing device, cause the computing device to perform operations comprising: training a convolutional neural network (CNN) using authenticated data and a taxonomy (para [0018]-[0019], [0025]-[0027], [0080]); receiving, by a processing device, a query comprising input data (para [0018]-[0019], [0025]-[0027], [0108]); classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy (para [0018]-[0019], [0025]-[0027], [0080]-[0081]); generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data (para [0025]-[0027], [0031], [0087]); and displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content (para [0036], [0051], [0085]).

Claims 12-14 and 17 lack an inventive step under PCT Article 33(3) as being obvious over GumGum in view of US 2019/0005670 A1 to Magic Leap, Inc. (hereinafter 'Magic Leap').

As to claim 12, GumGum further teaches wherein the probability is determined by a function (para [0024]-[0027], [0031], [0083]), but does not explicitly teach cross-entropy. However, Magic Leap does teach cross-entropy (para [0032], [0035] - "The network is trained using a standard cross entropy loss"). It would have been obvious to one of ordinary skill in the art to utilize cross-entropy as taught by Magic Leap in order to provide systems and methods for augmenting video data based on automated identification of one or more objects depicted in the video data as taught by GumGum capable of improved and more efficient network training and probabilistic analysis. Both GumGum and Magic Leap are directed to systems and methods utilizing machine learning in conjunction with training of neural networks and augmented reality application.

(See Next Continuation Box)

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 20/32149

**Supplemental Box**

In case **the space in any of the preceding boxes is not sufficient.**
Continuation of:
Box V.2. Citations and explanations:

As to claim 13, GumGum further teaches wherein the result comprises augmented reality content, wherein displaying the result comprises: displaying the result in an augmented reality apparatus, comprising: and displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data (para [0042], [0098]-[0099]), but does not explicitly teach passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device. However, Magic Leap does teach passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device (para [0063] - "Eyepieces 1002A and 1002B may comprise transparent or semi-transparent waveguides configured to direct light from projectors 1014A and 1014B, respectively. Specifically, processing module 1050 may cause left projector 1014A to output a left projected image 1022A into left eyepiece 1002A, and may cause right projector 1014B to output a right projected image 1022B into right eyepiece 1002B. In some embodiments, each of eyepieces 1002 may each comprise a plurality of waveguides corresponding to different colors and/or different depth planes"). It would have been obvious to one of ordinary skill in the art to utilize passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device as taught by Magic Leap in order to provide systems and methods for augmenting video data based on automated identification of one or more objects depicted in the video data as taught by GumGum capable of improved provision of augmented reality user interaction/interfaces.

As to claim 14, GumGum further teaches wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time (para [0018], [0025], [0072], [0084]-[0085], [0090]-[0091]).

As to claim 17, GumGum further teaches wherein the input data comprises user uploaded data, the method further comprising authenticating the user uploaded data using Neural Networks and adding the authenticated user uploaded data to the authenticated data (para [0018]-[0019], [0030]-[0033]), but does not explicitly teach Siamese. However, Magic Leap does teach Siamese (para [0042] - "a siamese network can be used"). It would have been obvious to one of ordinary skill in the art to utilize Siamese as taught by Magic Leap in order to provide systems and methods for augmenting video data based on automated identification of one or more objects depicted in the video data as taught by GumGum capable of improved similarity determination/fingerprinting and data authentication.

Claims 15-16 lack an inventive step under PCT Article 33(3) as being obvious over GumGum in view of US 2019/0102676 A1 to SAS Institute Inc. (hereinafter 'SAS').

As to claim 15, GumGum further teaches further comprising training using authenticated data and a taxonomy (para [0018]-[0019], [0025]-[0027], [0080]), but does not explicitly teach a recurrent neural network (RNN). However, SAS does teach a recurrent neural network (RNN) (para [0148]-[0150] - "the neural network 1200 can be trained"; "the neural network 1200 is a recurrent neural network"). It would have been obvious to one of ordinary skill in the art to utilize a recurrent neural network (RNN) as taught by SAS in order to provide systems and methods for augmenting video data based on automated identification of one or more objects depicted in the video data as taught by GumGum capable of improved and more efficient network training. Both GumGum and SAS are directed to systems and methods utilizing machine learning in conjunction with training of neural networks.

As to claim 16, GumGum further teaches wherein the input data comprises data, the method further comprising: and classifying, by the trained CNN, the structured data (para [0018]-[0019], [0025]-[0027], [0080]-[0081], [0151]). SAS further teaches unstructured (para [0048] - "network-attached data stores 110 may hold unstructured (e.g., raw) data, such as manufacturing data"), processing, by the trained RNN (para [0148]-[0150]), the unstructured data to yield structured data (para [0049] - "computing environment 114 may be used to analyze the unstructured data in a variety of ways to determine the best way to structure (e.g., hierarchically) that data, such that the structured data is tailored to a type of further analysis that a user wishes to perform on the data").

Claims 1-20 have industrial applicability as defined by PCT Article 33(4) because the subject matter can be made or used in industry.

# SEARCH HISTORY

| | |
|---|---|
| Application Number | PCT/US 20/32149 |
| Search Conducted By | JL |
| Search Approved By | JL |

| | |
|---|---|
| US/IPC Classifications Searched | CPC: G06N 20/00, G06K 9/629, G06N 3/0454, G06N 3/08, G06K 9/62; IPC(8): G06N 20/00, G06K 9/46, G06K 9/62, G06N 3/08 (2020.01); USPC: 706/12, 706/15, 706/20 |
| Date Conducted | 01 July 2020 (01.07.2020) |

| | |
|---|---|
| Documentation Searched | CPC: G06N 20/00, G06K 9/629, G06N 3/0454, G06N 3/08, G06K 9/62; USPC: 706/12, 706/15, 706/20 (keyword limited; terms below) |
| Search Terms Used | training convolutional neural network authenticated taxonomy classify input display, etc. |
| Date Conducted | 01 July 2020 (01.07.2020) |

| | |
|---|---|
| Electronic Database Searched | PatBase |
| Files Searched | Full-text: AU BE BR CA CH CN DE DK EP ES FI FR GB IN JP KR SE TH TW US WO<br>Bibliographic: (Europe) AT BA BE BG CH CS CY CZ DD DK EE ES FI GE GR HR HU IE IS IT LT LU LV MC MD MT NL NO PL PT RO RS SE SI SK SM TR UA YU (Asia) EA GC HK ID IL IN KZ MN MY PH RU SG SU TH TJ TW UZ VN (North America) CA CR CU DO GT HN MX NI PA SV TT (South America) AR BR CL CO EC PE UY (Australasia) AU NZ (Africa) AP DZ EG KE MA MW OA ZA ZM ZW |
| Date Conducted | 01 July 2020 (01.07.2020) |

**Search Logic:**

*** CLASSIFICATION-BASED SEARCH ***

Search 1: (IC=(G06N20/00 or G06K9/46 or G06K9/62 or G06N3/08) or UC=(706/12 or 706/15 or 706/20) or CPC=(G06N20/00 or G06K9/629 or G06N3/0454 or G06N3/08 or

G06K9/62)) (Results 171786)
Search 2: (neural* w2 network*) (Results 133811)
Search 3: (1 and 2) (Results 46792)
Search 4: 3 and ((convol* w3 (neural* or network*)) or CNN) (Results 18392)
Search 5: 4 and (train* w5 (convol* or neural* or network* or CNN)) (Results 10999)
Search 6: 5 and taxon* (Results 159)
Search 7: 6 and (search* or quer*) (Results 117)
Search 8: 7 and classif* (Results 111)
Search 9: 8 and authenticat* (Results 31)
Search 10: 9 and ((match* or similar*) w3 (close* or near* or best or threshold* or scor* or rank*)) (Results 19)
Search 11: 10 and (display* or output* or view* or ((augmented* or virtual* or mixed*) w2 reality*)) (Results 19)

*** FREEFORM SEARCH ***

Search 12: 2 and ((convol* w3 (neural* or network*)) or CNN) (Results 28177)
Search 13: 12 and (train* w5 (convol* or neural* or network* or CNN)) (Results 15234)
Search 14: 13 and taxon* (Results 219)
Search 15: 14 and (search* or quer*) (Results 157)
Search 16: 15 and classif* (Results 149)
Search 17: 16 and authenticat* (Results 41)
Search 18: 17 and ((match* or similar*) w3 (close* or near* or best or threshold* or scor* or rank*)) (Results 21)
Search 19: 18 and (display* or output* or view* or ((augmented* or virtual* or mixed*) w2 reality*)) (Results 21)

| | |
|---|---|
| **Electronic Database Searched** | Google |
| **Files Searched** | Google Scholar (excluding patents) |
| **Date Conducted** | 01 July 2020 (01.07.2020) |

**Search Logic:**

About 16,500 results (0.10 sec) on training convolutional neural network authenticated taxonomy

About 16,400 results (0.17 sec) on training convolutional neural network authenticated taxonomy classify input display

| | |
|---|---|
| **Electronic Database Searched** | PubWEST |
| **Files Searched** | PGPB, USPT, EPAB, JPAB |
| **Date Conducted** | 01 July 2020 (01.07.2020) |

**Search Logic:**

**DATE: Wednesday, July 01, 2020** Purge Queries Printable Copy Create Case

| **Order By** | **Set Name** Side by Side | **Query** | **Hit Count** | **Set Name** Result Set | **Set Name** Grid | **Set Name** Classification |
|---|---|---|---|---|---|---|
| | | *DB=PGPB,USPT,EPAB,JPAB; PLUR=YES; OP=ADJ* | | | | |
| Date | L12 | L11 and copyright$ | 56 | L12 | L12 | L12 |
| Date | L11 | L10 and (vehic$ or car or truck or automobile or airplane or aircraft or geograph$ or geolocat$ or geoposition$ or GPS$) | 388 | L11 | L11 | L11 |
| Date | L10 | L9 and (display$ or output$ or view$ or ((augmented$ or virtual$ or mixed$) near2 reality$)) | 445 | L10 | L10 | L10 |
| Date | L9 | L8 and ((match$ or similar$) near3 (close$ or "nearest" or best or threshold$ or scor$ or rank$)) | 445 | L9 | L9 | L9 |
| Date | L8 | L7 and ((authentic$ or verif$ or validat$ or copyright$) near5 (data$ or element or authorship or image or photo$ or picture or video or audio$ or illustrat$ or writing or content or media)) | 870 | L8 | L8 | L8 |
| Date | L7 | L6 and classif$ | 2711 | L7 | L7 | L7 |
| Date | L6 | L5 and (search$ or quer$) | 2837 | L6 | L6 | L6 |
| Date | L5 | L4 and (taxon$ or category$ or class or classif$) | 4614 | L5 | L5 | L5 |
| Date | L4 | L3 and (train$ near5 (convol$ or neural$ or network$ or CNN)) | 5218 | L4 | L4 | L4 |
| Date | L3 | L2 and ((convol$ near3 (neural$ or network$)) or CNN) | 9152 | L3 | L3 | L3 |
| Date | L2 | L1 and @pd<20190509 | 122378 | L2 | L2 | L2 |
| Date | L1 | (neural$ near2 network$) | 159570 | L1 | L1 | L1 |

END OF SEARCH HISTORY

# PCT

## REQUEST

The undersigned requests that the present international application be processed according to the Patent Cooperation Treaty.

For receiving Office use only

PCT/US20/32149
International Application No.

08 May 2020 (08.05.2020)
International Filing Date

RO/US
Name of receiving Office and "PCT International Application"

Applicant's or agent's file reference *(if desired) (12 characters maximum)* 28108-110

**Box No. I TITLE OF INVENTION**

METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

**Box No. II APPLICANT** ☐ This person is also inventor

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)*

Automobilia II, LLC
301 Clematis Street #3000
West Palm Beach, FL 33401
United States of America

Telephone No.

Facsimile No.

Applicant's registration No. with the Office

**E-mail authorization**: Marking one of the check-boxes below authorizes the receiving Office, the International Searching Authority, the International Bureau and the International Preliminary Examining Authority to use the e-mail address indicated in this Box to send, notifications issued in respect of this international application to that e-mail address if those offices are willing to do so.

☐ as advance copies followed by paper notifications; or ☐ exclusively in electronic form (no paper notifications will be sent).

E-mail address:

State *(that is, country)* of nationality: US

State *(that is, country)* of residence: US

This person is applicant for the purposes of: ☒ all designated States ☐ the States indicated in the Supplemental Box

**Box No. III FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)**

☒ Further applicants and/or (further) inventors are indicated on a continuation sheet.

**Box No. IV AGENT OR COMMON REPRESENTATIVE; OR ADDRESS FOR CORRESPONDENCE**

The person identified below is hereby/has been appointed to act on behalf of the applicant(s) before the competent International Authorities as: ☒ agent ☐ common representative

Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country.)*

NICHOLES, Mary K.
LOWENSTEIN SANDLER LLP
One Lowenstein Drive
Roseland, NJ 07068
United States of America

Telephone No.
973-422-6536

Facsimile No.
973-597-6537

Agent's registration No. with the Office
56,238

**E-mail authorization**: Marking one of the check-boxes below authorizes the receiving Office, the International Searching Authority, the International Bureau and the International Preliminary Examining Authority to use the e-mail address indicated in this Box to send, notifications issued in respect of this international application to that e-mail address if those offices are willing to do so.

☒ as advance copies followed by paper notifications; or ☐ exclusively in electronic form (no paper notifications will be sent).

E-mail address: patents@lowenstein.com

☐ **Address for correspondence:** Mark this check-box where no agent or common representative is/has been appointed and the space above is used instead to indicate a special address to which correspondence should be sent.

**Box No. III FURTHER APPLICANT(S) AND/OR (FURTHER) INVENTOR(S)**

*If none of the following sub-boxes is used, this sheet should not be included in the request.*

| Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)* | This person is: |
|---|---|
| Lucinda Lewis<br>414 E. Valencia Avenue<br>#103<br>Burbank, CA 91505<br>United States of America | ☐ applicant only<br>☐ applicant and inventor<br>☒ inventor only *(If this check-box is marked, do not fill in below.)*<br>Applicant's registration No. with the Office |
| State *(that is, country)* of nationality: | State *(that is, country)* of residence: |
| This person is applicant for the purposes of: ☐ all designated States | ☐ the States indicated in the Supplemental Box |

| Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)* | This person is: |
|---|---|
| | ☐ applicant only<br>☐ applicant and inventor<br>☐ inventor only *(If this check-box is marked, do not fill in below.)*<br>Applicant's registration No. with the Office |
| State *(that is, country)* of nationality: | State *(that is, country)* of residence: |
| This person is applicant for the purposes of: ☐ all designated States | ☐ the States indicated in the Supplemental Box |

| Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)* | This person is: |
|---|---|
| | ☐ applicant only<br>☐ applicant and inventor<br>☐ inventor only *(If this check-box is marked, do not fill in below.)*<br>Applicant's registration No. with the Office |
| State *(that is, country)* of nationality: | State *(that is, country)* of residence: |
| This person is applicant for the purposes of: ☐ all designated States | ☐ the States indicated in the Supplemental Box |

| Name and address: *(Family name followed by given name; for a legal entity, full official designation. The address must include postal code and name of country. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below.)* | This person is: |
|---|---|
| | ☐ applicant only<br>☐ applicant and inventor<br>☐ inventor only *(If this check-box is marked, do not fill in below.)*<br>Applicant's registration No. with the Office |
| State *(that is, country)* of nationality: | State *(that is, country)* of residence: |
| This person is applicant for the purposes of: ☐ all designated States | ☐ the States indicated in the Supplemental Box |

☐ Further applicants and/or (further) inventors are indicated on another continuation sheet.

**Supplemental Box** *If the Supplemental Box is not used, this sheet should not be included in the request.*

1. *If, in any of the Boxes, except Boxes Nos. VIII(i) to (v) for which a special continuation box is provided, **the space is insufficient** to furnish all the information: in such case, write "Continuation of Box No...." (indicate the number of the Box) and furnish the information in the same manner as required according to the captions of the Box in which the space was insufficient, in particular:*
   - (i) ***if more than one person is to be indicated as applicant and/or inventor** and no "continuation sheet" is available: in such case, write "Continuation of Box No. III" and indicate for each additional person the same type of information as required in Box No. III. The country of the address indicated in this Box is the applicant's State (that is, country) of residence if no State of residence is indicated below;*
   - (ii) *if, in Box No. II or in any of the sub-boxes of Box No. III, the indication **"the States indicated in the Supplemental Box"** is checked: in such case, write "Continuation of Box No. II" or "Continuation of Box No. III" or "Continuation of Boxes No. II and No. III" (as the case may be), indicate the name of the applicant(s) involved and, next to (each) such name, the State(s) (and/or, where applicable, ARIPO, Eurasian, European or OAPI patent) for the purposes of which the named person is applicant;*
   - (iii) *if, in Box No. II or in any of the sub-boxes of Box No. III, **the inventor or the inventor/applicant is not inventor for the purposes of all designated States**: in such case, write "Continuation of Box No. II" or "Continuation of Box No. III" or "Continuation of Boxes No. II and No. III" (as the case may be), indicate the name of the inventor(s) and, next to (each) such name, the State(s) (and/or, where applicable, ARIPO, Eurasian, European or OAPI patent) for the purposes of which the named person is inventor;*
   - (iv) *if, in addition to the agent(s) indicated in Box No. IV, there are **further agents**: in such case, write "Continuation of Box No. IV" and indicate for each further agent the same type of information as required in Box No. IV;*
   - (v) *if, in Box No. VI, there are **more than three earlier applications whose priority is claimed**: in such case, write "Continuation of Box No. VI" and indicate for each additional earlier application the same type of information as required in Box No. VI.*
2. *If the applicant intends to make an indication of the wish that the international application be treated, in certain designated States, as an application for a patent of addition, certificate of addition, inventor's certificate of addition or utility certificate of addition: in such case, write the name or two-letter code of each designated State concerned and the indication "**patent of addition,**" "**certificate of addition,**" "**inventor's certificate of addition**" or "**utility certificate of addition,**" the number of the parent application or parent patent or other parent grant and the date of grant of the parent patent or other parent grant or the date of filing of the parent application (Rules 4.11(a)(i) and 49bis.1(a) or (b)).*
3. *If the applicant intends to make an indication of the wish that the international application be treated, in the United States of America, as a continuation or continuation-in-part of an earlier application: in such case, write "United States of America" or "US" and the indication "**continuation**" or "**continuation-in-part**" and the number and the filing date of the parent application (Rules 4.11(a)(ii) and 49bis.1(d)).*

Continuation of Box No. IV

Kesslen, Mark P. (Reg. No. 34,848)
Buckingham, Stephen R. (Reg. 40,538)
Sierchio, Daniel D. (Reg. No. 53,591)
Toma, David (Reg. No. 57,380)
Portnova, Marina (Reg. No. 45,750)
Kimes, Benjamin A. (Reg. No. 50,870)
Paradiso, Robert J. (Reg. No. 41,240)
Grange, Kevin O. (Reg. No. 60,793)
Krueger, Paul M. (Reg. No. 69,944)
Campolongo, Michael J. (Reg. No. 69,034)
Nicholes, Mary K. (Reg. No. 56,238)
Kopelevich, Sofia (Reg. No. 72,909)
Jones, Joseph G. (Reg. No. 62,151)
Lazoff, David (Reg. No. 42,783)
Andreev, Dmitry (Reg. No. 57,428)
Miedema, Garrett (Reg. No. 61,938)
Manali Joglekar (Reg. No. 67,339)
Shekher, Rahul (Reg. No. 69,049)
Greene, Nathan (Reg. No. 56,956)
Shangeeta, Rinat (Reg. No. 74,257)

all of:
Lowenstein Sandler LLP
One Lowenstein Drive
Roseland, New Jersey 07068
United States of America

**Box No. V DESIGNATIONS**

The filing of this request **constitutes under Rule 4.9(a) the designation** of all Contracting States bound by the PCT on the international filing date, for the grant of every kind of protection available and, where applicable, for the grant of both regional and national patents.

However,

☐ DE Germany **is not designated** for any kind of national protection

☐ JP Japan **is not designated** for any kind of national protection

☐ KR Republic of Korea **is not designated** for any kind of national protection

*(The check-boxes above may only be used to exclude (irrevocably) the designations concerned if, at the time of filing or subsequently under Rule 26bis.1, the international application contains in Box No. VI a priority claim to an earlier national application filed in the particular State concerned, in order to avoid the ceasing of the effect, under the national law, of this earlier national application.)*

**Box No. VI PRIORITY CLAIM AND DOCUMENT**

**The priority of the following earlier application(s) is hereby claimed:**

| Filing date of earlier application *(day/month/year)* | Number of earlier application | Where earlier application is: | | |
|---|---|---|---|---|
| | | national application: country or Member of WTO | regional application: regional Office | international application: receiving Office |
| item (1) | | | | |
| item (2) | | | | |
| item (3) | | | | |

☐ Further priority claims are indicated in the Supplemental Box.

**Furnishing the priority document(s):**

☒ The **receiving Office** is requested to prepare and transmit to the International Bureau a certified copy of the earlier application(s) *(only if the earlier application(s) was filed with the receiving Office which, for the purposes of this international application, is the receiving Office)* identified above as:

☒ all items ☐ item (1) ☐ item (2) ☐ item (3) ☐ other, see Supplemental Box

☐ The **International Bureau** is requested to obtain from a digital library a certified copy of the earlier application(s) identified above, using, where applicable, the access code(s) indicated below *(if the earlier application(s) is available to it from a digital library)*:

☐ item (1) access code ______________ ☐ item (2) access code ______________ ☐ item (3) access code ______________ ☐ other, see Supplemental Box

**Restore the right of priority:** the receiving Office is requested to restore the right of priority for the earlier application(s) identified above or in the Supplemental Box as item(s) (______________________________). *(See also the Notes to Box No. VI; further information **must** be provided to support a request to restore the right of priority.)*

**Incorporation by reference:** where an element of the international application referred to in Article 11(1)(iii)(d) or (e) or a part of the description, claims or drawings referred to in Rule 20.5(a) is not otherwise contained in this international application but is completely contained in an earlier application whose priority is claimed on the date on which one or more elements referred to in Article 11(1)(iii) were first received by the receiving Office, that element or part is, subject to confirmation under Rule 20.6, incorporated by reference in this international application for the purposes of Rule 20.6.

**Box No. VII INTERNATIONAL SEARCHING AUTHORITY**

**Choice of International Searching Authority (ISA)** *(if more than one International Searching Authority is competent to carry out the international search, indicate the Authority chosen; the two-letter code may be used)*:

ISA/ US

**Box No. IX** **CHECK LIST for EFS-Web filings** - this sheet is only to be used when filing an international application with RO/US via **EFS-Web**

| This international application **contains** the following: | Number of sheets | This international application is **accompanied by** the following item(s) *(mark the applicable check-boxes below and indicate in right column the number of each item)*: | Number of items |
|---|---|---|---|
| (a) request form PCT/RO/101 (including any declarations and supplemental sheets) : | 5 | 1. ☒ fee calculation sheet : | |
| (b) description (excluding any sequence listing part of the description, see (f), below) : | 33 | 2. ☐ original separate power of attorney : | |
| (c) claims : | 4 | 3. ☐ original general power of attorney : | |
| (d) abstract : | 1 | 4. ☐ copy of general power of attorney; reference number: : | |
| (e) drawings (if any) : | 7 | 5. ☐ priority document(s) identified in Box No. VI as item(s) : | |
| (f) sequence listing part of the description in the form of an **image file** (e.g. PDF) : | | 6. ☐ Translation of international application into *(language)*: : | |
| **Total number of sheets** (including the sequence listing part of the description **if filed as an image file**) : | 50 | 7. ☐ separate indications concerning deposited microorganism or other biological material : | |
| (g) sequence listing part of the description | | 8. ☐ *(only where item (f) is marked in the left column)* copy of the sequence listing in electronic form (Annex C/ST.25 text file) not forming part of the international application but **furnished only for the purposes of international search** under Rule 13*ter* : | |
| ☐ filed in the form of an **Annex C/ST.25 text file** | | 9. ☐ *(only where item (f) (in the left column) and item 8 (above) are marked)* a statement confirming that "the information recorded in electronic form submitted under Rule 13*ter* is identical to the sequence listing as contained in the international application" as filed via EFS-Web: : | |
| ☐ WILL BE filed separately on physical data carrier(s), on the same day and in the form of an **Annex C/ST.25 text file** | | 10. ☐ copy of results of earlier search(es) (Rule 12*bis*.1(a)) : | |
| Indicate type and number of physical data carrier(s) | | 11. ☐ other *(specify)*: : | |

**Figure of the drawings** which should accompany the abstract: FIG. 1

**Language of filing** of the international application: English

**Box No. X** **SIGNATURE OF APPLICANT, AGENT OR COMMON REPRESENTATIVE**

*Next to each signature, indicate the name of the person signing and the capacity in which the person signs (if such capacity is not obvious from reading the request).*

/Mary K. Nicholes/ May 8, 2020

Mary K. Nicholes
U.S. Reg. No. 56,238

For receiving Office use only

| | |
|---|---|
| 1. Date of actual receipt of the purported international application: 08 May 2020 (08.05.2020) | 2. Drawings: ☐ received: ☐ not received: |
| 3. Corrected date of actual receipt due to later but timely received papers or drawings completing the purported international application: | |
| 4. Date of timely receipt of the required corrections under PCT Article 11(2): | |
| 5. International Searching Authority (if two or more are competent): ISA /US — 6. ☐ Transmittal of search copy delayed until search fee is paid | |

For International Bureau use only

Date of receipt of the record copy by the International Bureau:

1070910

# THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE

United States Patent and Trademark Office

*May 16, 2020*

**THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.**

**APPLICATION NUMBER:** ***62/845,546***
**FILING DATE:** ***May 09, 2019***
**RELATED PCT APPLICATION NUMBER:** ***PCT/US20/32149***

**THE COUNTRY CODE AND NUMBER OF YOUR PRIORITY APPLICATION, TO BE USED FOR FILING ABROAD UNDER THE PARIS CONVENTION, IS** ***US62/845,546***

UNITED STATES PATENT AND TRADEMARK OFFICE
DEPARTMENT OF COMMERCE

Certified by

Andrei Iancu

Under Secretary of Commerce
for Intellectual Property
and Director of the United States
Patent and Trademark Office

PTO/SB/16 (10-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE


# PROVISIONAL APPLICATION FOR PATENT COVER SHEET – Page 1 of 2

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

**Priority Mail Express® Label No.** ____________________

| INVENTOR(S) | | |
|---|---|---|
| Given Name (first and middle [if any]) | Family Name or Surname | Residence (City and either State or Foreign Country) |
| Lucinda | Lewis | Los Angeles, CA |
| | | |
| | | |
| | | |
| | | |

Additional inventors are being named on the ____________ separately numbered sheets attached hereto.

**TITLE OF THE INVENTION (500 characters max):**

*Neural Network for Vehicles in Mixed Reality*

Direct all correspondence to: **CORRESPONDENCE ADDRESS**

☐ The address corresponding to Customer Number:

**OR**

☑ Firm or Individual Name Automobilia II, LLC

Address P O Box 49167

| | | |
|---|---|---|
| City Sarasota | State FL | Zip 34230 |
| Country United States | Telephone | Email cindy.lewis@me.com |

**ENCLOSED APPLICATION PARTS (*check all that apply*)**

☐ Application Data Sheet. See 37 CFR 1.76.
☐ CD(s), Number of CDs ____________
☐ Drawing(s) *Number of Sheets* ____________
☐ Other (specify) ____________
☑ Specification (e.g., description of the invention) *Number of Pages* ____________

**Fees Due:** Filing Fee of $280 ($140 for small entity) ($70 for micro entity). If the specification and drawings exceed 100 sheets of paper, an application size fee is also due, which is $400 ($200 for small entity) ($100 for micro entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).

**METHOD OF PAYMENT OF THE FILING FEE AND APPLICATION SIZE FEE FOR THIS PROVISIONAL APPLICATION FOR PATENT**

☐ Applicant asserts small entity status. See 37 CFR 1.27.
☑ Applicant certifies micro entity status. See 37 CFR 1.29.
Applicant must attach form PTO/SB/15A or B or equivalent.
☐ A check or money order made payable to the *Director of the United States Patent and Trademark Office* is enclosed to cover the filing fee and application size fee (if applicable).
☑ Payment by credit card. Form PTO-2038 is attached.
☐ The Director is hereby authorized to charge the filing fee and application size fee (if applicable) or credit any overpayment to Deposit Account Number: ____________.

$70.00
**TOTAL FEE AMOUNT ($)**

***USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT***

PTO/SB/16 (10-18)
Approved for use through 11/30/2020. OMB 0651-0032
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995 no persons are required to respond to a collection of information unless it displays a valid OMB control number

# PROVISIONAL APPLICATION FOR PATENT COVER SHEET – Page 2 of 2

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government. (NOTE: Providing this information on a provisional cover sheet, such as this Provisional Application for Patent Cover Sheet (Form PTO/SB/16), does not satisfy the requirement of 35 U.S.C. 202(c)(6), which requires that the *specification* contain a statement specifying that the invention was made with Government support and that the Government has certain rights in the invention.)

☑ No.

☐ Yes, the invention was made by an agency of the U.S. Government. The U.S. Government agency name is: ______

☐ Yes, the invention was made under a contract with an agency of the U.S. Government.

The contract number is: ______

The U.S. Government agency name is: ______

In accordance with 35 U.S.C. 202(c)(6) and 37 CFR 401.14(f)(4), the specifications of any United States patent applications and any patent issuing thereon covering the invention, including the enclosed provisional application, must state the following:

"This invention was made with government support under [IDENTIFY THE CONTRACT] awarded by [IDENTIFY THE FEDERAL AGENCY]. The government has certain rights in the invention."

## WARNING:

Petitioner/applicant is cautioned to avoid submitting personal information in documents filed in a patent application that may contribute to identity theft. Personal information such as social security numbers, bank account numbers, or credit card numbers (other than a check or credit card authorization form PTO-2038 submitted for payment purposes) is never required by the USPTO to support a petition or an application. If this type of personal information is included in documents submitted to the USPTO, petitioners/applicants should consider redacting such personal information from the documents before submitting them to the USPTO. Petitioner/applicant is advised that the record of a patent application is available to the public after publication of the application (unless a non-publication request in compliance with 37 CFR 1.213(a) is made in the application) or issuance of a patent. Furthermore, the record from an abandoned application may also be available to the public if the application is referenced in a published application or an issued patent (see 37 CFR 1.14). Checks and credit card authorization forms PTO-2038 submitted for payment purposes are not retained in the application file and therefore are not publicly available.

SIGNATURE /Lucinda Lewis/ DATE 05-09-2019

TYPED OR PRINTED NAME Automobilia II, LLC REGISTRATION NO. ______
(*if appropriate*)

TELEPHONE ______ DOCKET NUMBER ______

# Privacy Act Statement

The **Privacy Act of 1974 (P.L. 93-579)** requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (*i.e.*, GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

# Electronic Acknowledgement Receipt

| | |
|---|---|
| **EFS ID:** | 35968168 |
| **Application Number:** | 62845546 |
| **International Application Number:** | |
| **Confirmation Number:** | 1057 |
| **Title of Invention:** | Neural Network for Identifying Vehicles and/or Providing Artificial Intelligence |
| **First Named Inventor/Applicant Name:** | Lucinda Lewis |
| **Customer Number:** | 162160 |
| **Filer:** | Lucinda Lewis |
| **Filer Authorized By:** | |
| **Attorney Docket Number:** | |
| **Receipt Date:** | 09-MAY-2019 |
| **Filing Date:** | |
| **Time Stamp:** | 15:26:44 |
| **Application Type:** | Provisional |

## Payment information:

| | |
|---|---|
| Submitted with Payment | yes |
| Payment Type | CARD |
| Payment was successfully received in RAM | $70 |
| RAM confirmation Number | 051019INTEFSW15300500 |
| Deposit Account | |
| Authorized User | |

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

## File Listing:

| Document Number | Document Description | File Name | File Size(Bytes)/ Message Digest | Multi Part /.zip | Pages (if appl.) |
|---|---|---|---|---|---|
| 1 | Specification | Lewis-Provisional-PatentApp.pdf | 42819<br>a2177916e6e0d7dd0d4aa1256059e2ce187bac2d | no | 3 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 2 | Certification of Micro Entity (Gross Income Basis) | sb0015a.pdf | 104946<br>10b958c5adaff82cd44ffdde89eb4e5d05799bac | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 3 | Provisional Cover Sheet (SB16) | sb0016-1.pdf | 186266<br>730668d05dc6dac78f8aa1dc3657746fd462fc9f | no | 3 |
| **Warnings:** | | | | | |
| This is not a USPTO supplied Provisional Cover Sheet SB16 form. | | | | | |
| **Information:** | | | | | |
| 4 | Fee Worksheet (SB06) | fee-info.pdf | 29202<br>8ff4637801bec31ae564bbfb15ab34d3cc6f1a8d | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | | **Total Files Size (in bytes):** | 363233 | |

This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.

New Applications Under 35 U.S.C. 111
If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.

National Stage of an International Application under 35 U.S.C. 371
If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.

New International Application Filed with the USPTO as a Receiving Office
If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.

Provisional Patent Application of

Lucinda Lewis

For

# TITLE: **Neural Network for Identifying Vehicles and/or Providing Artificial Intelligence**

*System and Method whereby multi-dimensional inputs may query for vehicle attributes and elements from vehicle dataset trained via machine learning from proprietary reference database (neural network) built-upon copyright-registered and authenticated intellectual property about transportation vehicles and their environments from the 1880s through the present day to provide artificial intelligence in mixed reality applications.*

Information may be retrieved from the system by multiple inputs: audio, text, or visual where the request is referenced against the proprietary-trained vehicle dataset and returns artificial intelligence matching the request in mixed-reality environments.

Queries about vehicles are answered with a probability of correct identification. For example, a user could type into the application: "Auto Union Wanderer W-25" and the system would interpret the words to return an image of the "Auto Union Wanderer W-25". The probability of the queried vehicle being built by "Auto Union" will be expressed as a percentage, for example "95%" probability this is an Auto Union and the probability of the vehicle being a model "Wanderer W-25" may be "85%". Additionally a short history of the vehicle would appear, and using the application geolocation services identify where the closest example of this vehicle may be physically available for review relative to the user's present location.

Information may also be retrieved by uploading a photograph of a vehicle the user may encounter in daily life to the system. Using image recognition derived from the proprietary database, the invention would return an identification of Vehicle Make and Model with probabilities of accuracy based upon the trained dataset rendered through Machine Learning (ML).

Users could also dictate voice requests into the application and/or query using natural language the make and model of a requested vehicle. For example, the user may speak, "show me Ford Thunderbird". The System will return an image or images representing the closest match to the "Ford Thunderbird" with additional information provided through mixed reality inputs. User may refine their query by speaking "show me "red 1957 Ford Thunderbird).

User could also initialize the query by pointing an input device (camera) at the vehicle in question, and the application will match the camera display against the ML trained dataset to provide an Augmented Reality (AR) display of information regarding the queried vehicle. Levels of information may be chosen via user setup. For example, user may only require “ Vehicle Make” aka “Ferrari” or Make, Model (Ferrari 250 GT) or may require technical information, such as “V-8 engines” returning images and information found about V-8 engines from the neural network of information.

In each query result, additional educational information about the vehicle may be provided depending upon user settings. A brief history of the car may be displayed or overlaid in mixed reality environment. For example, user may request “two-tone vehicle interiors” and matches to the requisition would be displayed on the device with overlaid text depicting the make and model of the vehicle.

WIPO
WORLD
INTELLECTUAL PROPERTY
ORGANIZATION

# DOCUMENT MADE AVAILABLE UNDER THE PATENT COOPERATION TREATY (PCT)

| | | |
|---|---|---|
| International application number: | | **PCT/US2020/032149** |
| International filing date: | | **08 May 2020 (08.05.2020)** |
| Document type: | | **Certified copy of priority document** |
| Document details: | Country/Office: | **US** |
| | Number: | **62/845,546** |
| | Filing date: | **09 May 2019 (09.05.2019)** |
| Date of receipt at the International Bureau: | | **17 May 2020 (17.05.2020)** |

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a),(b) or (b-*bis*)

34, chemin des Colombettes
1211 Geneva 20, Switzerland
www.wipo.int

# PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To: MARY K. NICHOLES
LOWENSTEIN SANDLER LLP
ONE LOWENSTIEN DRIVE
ROSELAND, NJ 07068

# PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

Date of mailing (*day/month/year*) 04 AUG 2020

| | | |
|---|---|---|
| Applicant's or agent's file reference 28108-110 | | **FOR FURTHER ACTION** See paragraph 2 below |
| International application No. PCT/US 20/32149 | International filing date (*day/month/year*) 08 May 2020 (08.05.2020) | Priority date (*day/month/year*) 09 May 2019 (09.05.2019) |

International Patent Classification (IPC) or both national classification and IPC
IPC - G06N 20/00, G06K 9/46, G06K 9/62, G06N 3/08 (2020.01)
CPC - G06N 20/00, G06K 9/629, G06N 3/0454, G06N 3/08, G06K 9/62

Applicant AUTOMOBILIA II, LLC

1. This opinion contains indications relating to the following items:

- [x] Box No. I Basis of the opinion
- [ ] Box No. II Priority
- [ ] Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- [ ] Box No. IV Lack of unity of invention
- [x] Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step and industrial applicability; citations and explanations supporting such statement
- [ ] Box No. VI Certain documents cited
- [ ] Box No. VII Certain defects in the international application
- [ ] Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

| Name and mailing address of the ISA/US | Date of completion of this opinion | Authorized officer |
|---|---|---|
| Mail Stop PCT, Attn: ISA/US<br>Commissioner for Patents<br>P.O. Box 1450, Alexandria, Virginia 22313-1450<br>Facsimile No. 571-273-8300 | 01 July 2020 (01.07.2020) | Lee Young<br>PCT Help Desk<br>Telephone No. 571-272-4300 |

**WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY**

International application No. PCT/US 20/32149

**Box No. I Basis of this opinion**

1. With regard to the **language**, this opinion has been established on the basis of:
   - [X] the international application in the language in which it was filed.
   - [ ] a translation of the international application into ________________ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

2. [ ] This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(b)).

3. [ ] With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of a sequence listing:
   - a. [ ] forming part of the international application as filed:
     - [ ] in the form of an Annex C/ST.25 text file.
     - [ ] on paper or in the form of an image file.
   - b. [ ] furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
   - c. [ ] furnished subsequent to the international filing date for the purposes of international search only:
     - [ ] in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).
     - [ ] on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

4. [ ] In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

**WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY**

International application No.
PCT/US 20/32149

**Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step and industrial applicability; citations and explanations supporting such statement**

1. Statement

| | | | |
|---|---|---|---|
| Novelty (N) | Claims | 12-17 | YES |
| | Claims | 1-11, 18-20 | NO |
| Inventive step (IS) | Claims | None | YES |
| | Claims | 1-20 | NO |
| Industrial applicability (IA) | Claims | 1-20 | YES |
| | Claims | None | NO |

2. Citations and explanations:

Claims 1-11 and 18-20 lack novelty under PCT Article 33(2) as being anticipated by US 2018/0084310 A1 to GumGum, Inc. (hereinafter 'GumGum').

As to claim 1, GumGum teaches a method comprising: training a convolutional neural network (CNN) using authenticated data and a taxonomy (para [0018]-[0019], [0025]-[0027], [0080] - "A common existing use of video fingerprints is in the field of Digital Rights Management ("DRM") in order for content owners (such as film studios or publishers) to identify when files containing their copyrighted video content are uploaded to video sharing websites or other file sharing networks"; "An image object that is the subject of a classifier that has been created or trained to identify that particular image object may be referred to as a "target image object." For example, a target image object may be a visual representation of a company's logo, and a classifier may be generated specifically to identify at least that logo. In some embodiments, the target image object may generally refer to a class or group of related image objects that may be identified using a particular classifier"; "these neural networks may be trained using video data, but may then be used with either video data or image data as input. In one embodiment, block 612 may be implemented at least in part using a convolutional neural network"); receiving, by a processing device, a query comprising input data (para [0018]-[0019], [0025]-[0027], [0108] - "Video fingerprinting is generally used to compactly represent a video or portion of a video in a manner that allows for efficient searching of other videos to identify either duplicates or slight variations of the same video. For example, computer software may identify, extract and compress characteristics of video data to create a digital file that serves as a "fingerprint" of the specific video data"; "a computing system may analyze a live event broadcast for brand exposures (e.g., instances where a company name, logo or other visual indicator is present in a frame of broadcasted video). For each video sequence that contains a brand exposure, the system may create a digital fingerprint to identify a portion of the video in which the brand exposure occurs. The system may subsequently search video data of potentially hundreds or thousands of subsequent broadcasts across different channels, websites, streaming services, or social media networks to identify occurrences of the digital fingerprint"; "system can query the API with an input parameter such as a keyword, tag, or category, and/or to retrieve media item result sets"); classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy (para [0018]-[0019], [0025]-[0027], [0080]-[0081] - "the computing system at block 612 may use classification models that have each been trained"; "the computing system may access neural networks or other machine learning models that have been trained to identify objects associated with the relevant sponsor(s) for the given team(s) and/or the home venue, then may proceed to sponsor logo identification"); generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data (para [0025]-[0027], [0031], [0087] - "trained image classifier may return a certain probability that a given frame of the video includes a target image object (such as a specific logo), and the system may deem those frames associated with a probability that satisfies a certain threshold likelihood (e.g., 75%) as including the target image object"; "one or more classifiers may provide a separate probability for each of the different companies and/or for each different logo"; "the system may determine confidence levels based on how many pixels the original video portion and subsequent video portion have in common (or pixels that are substantially the same, accounting for slight variations), then based on this confidence level may perform image recognition to verify that a similar video frame is in fact a positive match"; "to supplement and/or confirm the confidence level determinations of the image and video classification models discussed above, the computing system may use text classification models applied to text of a social media post or page in which the image or video is embedded or linked. For example, a textual classifier may provide a confidence score indicating how confident the model is that a particular social media post relates to a specific rights holder and/or sponsor based on the text of the post. This confidence score may be weighted and combined with a corresponding confidence score determined from the image or video itself"); and displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content (para [0036], [0051], [0085] - "the system may generate various spreadsheets, summary views, dashboards, reports, user interfaces and/or other output intended for consumption by a sponsor or other user. The data in one embodiment may include, for each of a number of different URLs, videos, images or social media posts: an indication whether the account is owned and operated by the brand"; "media report may include information for a number of different sponsors' logos, and may be based on the results of multiple passes through the illustrative method").

As to claim 2, GumGum further teaches wherein the authenticated data comprises copyright registered works of authorship, metadata and text (para [0018]-[0019], [0086]-[0087]).

As to claim 3, GumGum further teaches wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings (para [0018]-[0019], [0086]-[0087]).

(See Continuation Box)

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 20/32149

**Supplemental Box**

In case the space in any of the preceding boxes is not sufficient.
Continuation of:
Box V.2. Citations and explanations:

As to claim 4, GumGum further teaches wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information (para [0034], [0098]).

As to claim 5, GumGum further teaches wherein the authenticated data comprises data from a copyright registered database (para [0018]-[0019], [0090]).

As to claim 6, GumGum further teaches wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture (para [0025], [0074], [0084], [0098]-[0099]).

As to claim 7, GumGum further teaches wherein the input data comprises one or more of image data, video data, intake data or geographical location data (para [0071]).

As to claim 8, GumGum further teaches wherein classifying comprises mapping input data to authenticated data using the taxonomy (para [0025]-[0027], [0083]-[0087]).

As to claim 9, GumGum further teaches wherein the result comprises one or more of an image, a video, text, or sound (para [0036], [0051], [0085], [0098]-[0099]).

As to claim 10, GumGum further teaches wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information (para [0036], [0051], [0085], [0098]-[0099]).

As to claim 11, GumGum further teaches wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy (para [0024]-[0027], [0031]).

As to claim 18, GumGum teaches a system comprising: a memory; a processor, coupled to the memory (para [0038]-[0042]), the processor configured to: train a convolutional neural network (CNN) using authenticated data and a taxonomy (para [0018]-[0019], [0025]-[0027], [0080]); receive, by a processing device, a query comprising input data (para [0018]-[0019], [0025]-[0027], [0108]); classify, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy (para [0018]-[0019], [0025]-[0027], [0080]-[0081]); generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data (para [0025]-[0027], [0031], [0087]); and display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content (para [0036], [0051], [0085]).

As to claim 19, GumGum further teaches wherein the authenticated data comprises copyright registered works of authorship, metadata and text (para [0018]-[0019], [0086]-[0087]).

As to claim 20, GumGum teaches a computer-readable non-transitory storage medium comprising executable instructions that, when executed by a computing device, cause the computing device to perform operations comprising: training a convolutional neural network (CNN) using authenticated data and a taxonomy (para [0018]-[0019], [0025]-[0027], [0080]); receiving, by a processing device, a query comprising input data (para [0018]-[0019], [0025]-[0027], [0108]); classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy (para [0018]-[0019], [0025]-[0027], [0080]-[0081]); generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data (para [0025]-[0027], [0031], [0087]); and displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content (para [0036], [0051], [0085]).

Claims 12-14 and 17 lack an inventive step under PCT Article 33(3) as being obvious over GumGum in view of US 2019/0005670 A1 to Magic Leap, Inc. (hereinafter 'Magic Leap').

As to claim 12, GumGum further teaches wherein the probability is determined by a function (para [0024]-[0027], [0031], [0083]), but does not explicitly teach cross-entropy. However, Magic Leap does teach cross-entropy (para [0032], [0035] - "The network is trained using a standard cross entropy loss"). It would have been obvious to one of ordinary skill in the art to utilize cross-entropy as taught by Magic Leap in order to provide systems and methods for augmenting video data based on automated identification of one or more objects depicted in the video data as taught by GumGum capable of improved and more efficient network training and probabilistic analysis. Both GumGum and Magic Leap are directed to systems and methods utilizing machine learning in conjunction with training of neural networks and augmented reality application.

(See Next Continuation Box)

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 20/32149

**Supplemental Box**

In case the space in any of the preceding boxes is not sufficient.
Continuation of:
Box V.2. Citations and explanations:

As to claim 13, GumGum further teaches wherein the result comprises augmented reality content, wherein displaying the result comprises: displaying the result in an augmented reality apparatus, comprising: and displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data (para [0042], [0098]-[0099]), but does not explicitly teach passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device. However, Magic Leap does teach passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device (para [0063] - "Eyepieces 1002A and 1002B may comprise transparent or semi-transparent waveguides configured to direct light from projectors 1014A and 1014B, respectively. Specifically, processing module 1050 may cause left projector 1014A to output a left projected image 1022A into left eyepiece 1002A, and may cause right projector 1014B to output a right projected image 1022B into right eyepiece 1002B. In some embodiments, each of eyepieces 1002 may each comprise a plurality of waveguides corresponding to different colors and/or different depth planes"). It would have been obvious to one of ordinary skill in the art to utilize passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device as taught by Magic Leap in order to provide systems and methods for augmenting video data based on automated identification of one or more objects depicted in the video data as taught by GumGum capable of improved provision of augmented reality user interaction/interfaces.

As to claim 14, GumGum further teaches wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time (para [0018], [0025], [0072], [0084]-[0085], [0090]-[0091]).

As to claim 17, GumGum further teaches wherein the input data comprises user uploaded data, the method further comprising authenticating the user uploaded data using Neural Networks and adding the authenticated user uploaded data to the authenticated data (para [0018]-[0019], [0030]-[0033]), but does not explicitly teach Siamese. However, Magic Leap does teach Siamese (para [0042] - "a siamese network can be used"). It would have been obvious to one of ordinary skill in the art to utilize Siamese as taught by Magic Leap in order to provide systems and methods for augmenting video data based on automated identification of one or more objects depicted in the video data as taught by GumGum capable of improved similarity determination/fingerprinting and data authentication.

Claims 15-16 lack an inventive step under PCT Article 33(3) as being obvious over GumGum in view of US 2019/0102676 A1 to SAS Institute Inc. (hereinafter 'SAS').

As to claim 15, GumGum further teaches further comprising training using authenticated data and a taxonomy (para [0018]-[0019], [0025]-[0027], [0080]), but does not explicitly teach a recurrent neural network (RNN). However, SAS does teach a recurrent neural network (RNN) (para [0148]-[0150] - "the neural network 1200 can be trained"; "the neural network 1200 is a recurrent neural network"). It would have been obvious to one of ordinary skill in the art to utilize a recurrent neural network (RNN) as taught by SAS in order to provide systems and methods for augmenting video data based on automated identification of one or more objects depicted in the video data as taught by GumGum capable of improved and more efficient network training. Both GumGum and SAS are directed to systems and methods utilizing machine learning in conjunction with training of neural networks.

As to claim 16, GumGum further teaches wherein the input data comprises data, the method further comprising: and classifying, by the trained CNN, the structured data (para [0018]-[0019], [0025]-[0027], [0080]-[0081], [0151]). SAS further teaches unstructured (para [0048] - "network-attached data stores 110 may hold unstructured (e.g., raw) data, such as manufacturing data"), processing, by the trained RNN (para [0148]-[0150]), the unstructured data to yield structured data (para [0049] - "computing environment 114 may be used to analyze the unstructured data in a variety of ways to determine the best way to structure (e.g., hierarchically) that data, such that the structured data is tailored to a type of further analysis that a user wishes to perform on the data").

Claims 1-20 have industrial applicability as defined by PCT Article 33(4) because the subject matter can be made or used in industry.

**PATENT COOPERATION TREATY**

# PCT

**INTERNATIONAL PRELIMINARY REPORT ON PATENTABILITY**
(Chapter I of the Patent Cooperation Treaty)

(PCT Rule 44*bis*)

| Applicant's or agent's file reference<br>28108-110 | **FOR FURTHER ACTION** | See item 4 below |
|---|---|---|
| International application No.<br>PCT/US2020/032149 | International filing date *(day/month/year)*<br>08 May 2020 (08.05.2020) | Priority date *(day/month/year)*<br>09 May 2019 (09.05.2019) |
| International Patent Classification (8th edition unless older edition indicated)<br>See relevant information in Form PCT/ISA/237 | | |
| Applicant<br>AUTOMOBILIA II, LLC | | |

1. This international preliminary report on patentability (Chapter I) is issued by the International Bureau on behalf of the International Searching Authority under Rule 44 *bis*.1(a).

2. This REPORT consists of a total of 6 sheets, including this cover sheet.

   In the attached sheets, any reference to the written opinion of the International Searching Authority should be read as a reference to the international preliminary report on patentability (Chapter I) instead.

3. This report contains indications relating to the following items:

| | | |
|---|---|---|
| ☒ | Box No. I | Basis of the report |
| ☐ | Box No. II | Priority |
| ☐ | Box No. III | Non-establishment of opinion with regard to novelty, inventive step and industrial applicability |
| ☐ | Box No. IV | Lack of unity of invention |
| ☒ | Box No. V | Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement |
| ☐ | Box No. VI | Certain documents cited |
| ☐ | Box No. VII | Certain defects in the international application |
| ☐ | Box No. VIII | Certain observations on the international application |

4. The International Bureau will communicate this report to designated Offices in accordance with Rules 44*bis*.3(c) and 93*bis*.1 but not, except where the applicant makes an express request under Article 23(2), before the expiration of 30 months from the priority date (Rule 44*bis* .2).

| | |
|---|---|
| | Date of issuance of this report<br>02 November 2021 (02.11.2021) |
| The International Bureau of WIPO<br>34, chemin des Colombettes<br>1211 Geneva 20, Switzerland | Authorized officer<br>Yukari Nakamura<br>e-mail pct.team7@wipo.int |

Form PCT/IB/373 (revised January 2020)

| **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** ( **Not for submission under 37 CFR 1.99)** | | |
|---|---|---|
| | Application Number | |
| | Filing Date | |
| | First Named Inventor | LEWIS, Lucinda |
| | Art Unit | |
| | Examiner Name | |
| | Attorney Docket Number | 28108-125 |

**U.S.PATENTS** Remove

| Examiner Initial* | Cite No | Patent Number | Kind Code1 | Issue Date | Name of Patentee or Applicant of cited Document | Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear |
|---|---|---|---|---|---|---|
| | 1 | 10127250 | B2 | 2018-11-13 | Dingman et al. | |
| | 2 | 9818249 | B1 | 2017-11-14 | Fraser et al. | |
| | 3 | 8924415 | B2 | 2014-12-30 | Thomas | |
| | 4 | 8843543 | B2 | 2014-09-23 | Damman et al. | |
| | 5 | 7817104 | B2 | 2010-10-19 | Ryu et al. | |
| | 6 | 7660437 | B2 | 2010-02-09 | Breed | |
| | 7 | 6175787 | B1 | 2001-01-16 | Breed | |

If you wish to add additional U.S. Patent citation information please click the Add button. Add

**U.S.PATENT APPLICATION PUBLICATIONS** Remove

| | | |
|---|---|---|
| **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** ( **Not for submission under 37 CFR 1.99)** | Application Number | |
| | Filing Date | |
| | First Named Inventor | LEWIS, Lucinda |
| | Art Unit | |
| | Examiner Name | |
| | Attorney Docket Number | 28108-125 |

| Examiner Initial* | Cite No | Publication Number | Kind Code[1] | Publication Date | Name of Patentee or Applicant of cited Document | Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear |
|---|---|---|---|---|---|---|
| | 1 | 20180084310 | A1 | 2018-03-22 | GumGum Inc. | |
| | 2 | 20190005670 | A1 | 2019-01-03 | Magic Leap, Inc. | |
| | 3 | 20190102676 | A1 | 2019-04-04 | SAS Institute Inc. | |
| | 4 | 20180293552 | A1 | 2018-10-11 | Alibaba Group Holding Limited | |
| | 5 | 20150269244 | A1 | 2015-09-24 | Qamar et al. | |
| | 6 | 20140340423 | A1 | 2014-11-20 | Taylor et al. | |
| | 7 | 20130066936 | A1 | 2013-03-14 | Krishnan et al. | |
| | 8 | 20120116728 | A1 | 2012-05-10 | Shear et al. | |

If you wish to add additional U.S. Published Application citation information please click the Add button. Add

**FOREIGN PATENT DOCUMENTS** Remove

| Examiner Initial* | Cite No | Foreign Document Number[3] | Country Code[2] i | Kind Code[4] | Publication Date | Name of Patentee or Applicant of cited Document | Pages,Columns,Lines where Relevant Passages or Relevant Figures Appear | T[5] |
|---|---|---|---|---|---|---|---|---|

| | |
|---|---|
| **INFORMATION DISCLOSURE STATEMENT BY APPLICANT** ( **Not for submission under 37 CFR 1.99)** | |
| Application Number | |
| Filing Date | |
| First Named Inventor | LEWIS, Lucinda |
| Art Unit | |
| Examiner Name | |
| Attorney Docket Number | 28108-125 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2018197835 | WO | A1 | 2018-11-01 | BLIPPAR.COM LTD. | | |
| | 2 | 2018065549 | WO | A1 | 2018-04-12 | BLIPPAR.COM LTD. | | |

If you wish to add additional Foreign Patent Document citation information please click the Add button Add

**NON-PATENT LITERATURE DOCUMENTS** Remove

| Examiner Initials* | Cite No | Include name of the author (in CAPITAL LETTERS), title of the article (when appropriate), title of the item (book, magazine, journal, serial, symposium, catalog, etc), date, pages(s), volume-issue number(s), publisher, city and/or country where published. | T5 |
|---|---|---|---|
| | 1 | International Search Report for PCT Patent Application No. PCT/US2020/032149 Issued on August 4, 2020 5 pages. | |
| | 2 | White Paper "The Content Authenticity Initiative" Setting the Standard for Digital Content Attribution, Rosenthol, L., Parsons, A., Scouten, E., Athora, J., MacCormack, B., England, P., Levallee, M., Dotan, J., Hanna, S., Farid, H. Gregory, S (Witness) August 2020, 28 pages. | |

If you wish to add additional non-patent literature document citation information please click the Add button Add

**EXAMINER SIGNATURE**

| Examiner Signature | | Date Considered | |
|---|---|---|---|

*EXAMINER: Initial if reference considered, whether or not citation is in conformance with MPEP 609. Draw line through a citation if not in conformance and not considered. Include copy of this form with next communication to applicant.

[1] See Kind Codes of USPTO Patent Documents at www.USPTO.GOV or MPEP 901.04. [2] Enter office that issued the document, by the two-letter code (WIPO Standard ST.3). [3] For Japanese patent documents, the indication of the year of the reign of the Emperor must precede the serial number of the patent document. [4] Kind of document by the appropriate symbols as indicated on the document under WIPO Standard ST.16 if possible. [5] Applicant is to place a check mark here if English language translation is attached.

| INFORMATION DISCLOSURE STATEMENT BY APPLICANT ( Not for submission under 37 CFR 1.99) | | |
|---|---|---|
| | Application Number | |
| | Filing Date | |
| | First Named Inventor | LEWIS, Lucinda |
| | Art Unit | |
| | Examiner Name | |
| | Attorney Docket Number | 28108-125 |

**CERTIFICATION STATEMENT**

Please see 37 CFR 1.97 and 1.98 to make the appropriate selection(s):

That each item of information contained in the information disclosure statement was first cited in any communication from a foreign patent office in a counterpart foreign application not more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(1).

**OR**

☐ That no item of information contained in the information disclosure statement was cited in a communication from a foreign patent office in a counterpart foreign application, and, to the knowledge of the person signing the certification after making reasonable inquiry, no item of information contained in the information disclosure statement was known to any individual designated in 37 CFR 1.56(c) more than three months prior to the filing of the information disclosure statement. See 37 CFR 1.97(e)(2).

See attached certification statement.

The fee set forth in 37 CFR 1.17 (p) has been submitted herewith.

☒ A certification statement is not submitted herewith.

**SIGNATURE**

A signature of the applicant or representative is required in accordance with CFR 1.33, 10.18. Please see CFR 1.4(d) for the form of the signature.

| Signature | /Mary K. Nicholes/ | Date (YYYY-MM-DD) | 2021-11-09 |
|---|---|---|---|
| Name/Print | Mary K. Nicholes | Registration Number | 56,238 |

## Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these record s.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
01 November 2018 (01.11.2018)

W!P O | PCT

(10) International Publication Number
WO 2018/197835 Al

(51) International Patent Classification:
*G06K 9/46* (2006.0 1) *G06K 9/62* (2006.01)

(21) International Application Number:
PCT/GB20 18/050971

(22) International Filing Date:
12 April 2018 (12.04.2018)

(25) Filing Language: English

(26) Publication Langiiage: English

(30) Priority Data:
62/490,1 19 26 April 2017 (26.04.2017) US

(71) Applicant: BLIPPAR.COM LIMITED [GB/GB]; 5th Floor West, 1 London Bridge, London SE1 9BG (GB).

(72) Inventors: XU, Xun; c/o Blippar.Com Limited, 236 W. 30th Street, New York, New York 10001 (US). THILAK, Vimal; c/o Blippar.Com Limited, 236 W. 30th Street, New York, New York 10001 (US). CAO, Meng; c/o Blippar.Com Limited, 236 W. 30th Street, New York, New York 10001 (US). WANG, Xuejun; c/o Blippar.Com Limited, 236 W. 30th Street, New York, New York 10001 (US).

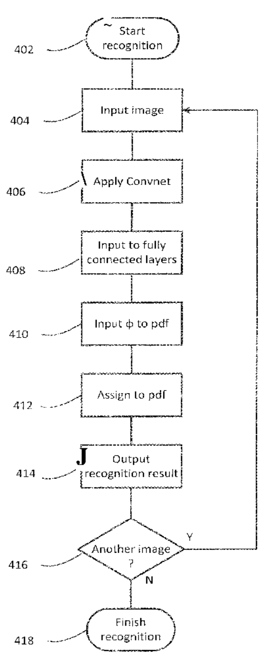
(74) Agent: WP THOMPSON; 138 Fetter Lane, London EC4A 1BT (GB).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

**Published:**
— *with international search report (Art. 21(3))*

**(54) Title:** APPARATUS AND METHOD FOR OPEN-SET OBJECT RECOGNITION




400

Fig. 7

**(57) Abstract:** Data processing apparatus comprising data processing resources comprising a processor and memory, said memory configured to store image data representative of a scene and configured with machine-readable instructions executable by said processor to configure said data processing circuitry to: provide a convolutional neural network trained for closed-set object recognition of a set of objects of an object training class; provide a prediction layer configured to comprise a first likelihood for an object belonging to said training class and a second likelihood for an object not belonging to said training class; receive plural instances of image data each representative of a scene comprising an object belonging to said training class or representative of a scene not comprising an object belonging to said training class and outputting for each of said plural instances of training image data respective output values $\phi$ from the said convolutional neural network indicative of the image content of said training image data, said value $\phi$ comprising a value $\phi$ responsive to the input scene comprising an object belonging to the training class and a value $\bar{\phi}$ for the input scene not comprising an object belonging to the training class; and generate a prediction layer indicative of a probability of an instance of image data being representative of a scene comprising an object belonging to said training class from a distribution of said respective output values $\phi$ and $\bar{\phi}$.

WO 2018/197835 A1

# APPARATUS AND METHOD FOR OPEN-SET OBJECT RECOGNITION

## Field

The present invention relates to an apparatus and method for open-set recognition of images. In particular, but not exclusively, the present invention relates to an apparatus and method for configuring a prediction layer for the open-set recognition of images and an apparatus and method for the open-set recognition of images utilising the prediction layer.

## Background

Deep learning models, specifically deep convolutional neural networks, have been proven to be very effective for visual recognition ([1] Krizhevsky, Sutskever and Hinton 2012; [2] Simonyan and Zisserman 2014; [3] Szegedy et al. 2014; [4] He et al. 2015). However, all these models are designed and trained for closed-set recognition, i.e. the image to be recognized is assumed to comprise an object or feature belonging to a definite set of objects and features determined at the training stage, and the network is configured to pick the correct object or feature out of that set. For the avoidance of doubt, the terms "object" and "feature" may be used interchangeably and the use of one may be considered use of the other unless the context requires otherwise. Objects and features may be referred to as being of a class, e.g. class $k$ or class $l$, to differentiate one type of feature or object from another.

As general background, closed-set recognition is particularly suitable for benchmarking tasks where the objects to be recognized are known to be restricted to a pre-defined set. However, in a real world application it is seldom the case that an object to be recognized in an image is a member of a pre-defined set. In a real world application, a situation may occur in which a recognition apparatus is fed with an arbitrary image, for example an image containing an object or objects not seen at the training stage or not containing any identifiable object at all (e.g. a blurry or cluttered image). For a closed-set recognition the apparatus is configured such that it is required to return a recognition result that is a one of the trained set of objects in the closed-set and if the apparatus is provided with such an arbitrary image the result is likely to be unpredictable, i.e. the apparatus will return a result corresponding to recognition of a one of the closed-set of objects without there being such an object in the image, thus rendering the apparatus unreliable in practice.

A conventional way to mitigate the so-called "open-set recognition problem" ([5] Girshick 2015) is to create a catch-all background class to capture all images that do not belong to the set of recognizable classes. Since the background can simply be anything, there are a prohibitive number of training images, and it is very hard to determine if a sufficient number of background images have been included in the training set.

A related work ([6] Bendale and Boult 2015) also attempts to solve the open-set recognition problem by adapting existing deep networks for closed-set recognition, with a specifically designed *generative* learning approach ([7] Bendale and Boult 2015).

Aspects and embodiments in accordance with the present invention were devised with the foregoing in mind.

## Summary

Viewed from a first aspect there is provided a data processing apparatus comprising data processing resources comprising a processor and memory, said memory configured to store image data representative of a scene and configured with machine-readable instructions executable by said processor to configure said data processing circuitry to:

provide a convolutional neural network trained for closed-set object recognition of a set of objects of an object training class;

provide a prediction layer configured to comprise a first likelihood for an object belonging to said training class and a second likelihood for an object not belonging to said training class;

receive plural instances of image data each representative of a scene comprising an object belonging to said training class or representative of a scene not comprising an object belonging to said training class and outputting for each of said plural instances of training image data respective output values $\phi$ from the said convolutional neural network indicative of the image content of said training image data, said value $\phi$ comprising a value $\phi$ responsive to the input scene comprising an object belonging to the training class and a value $\overline{\phi}$ for the input scene not comprising an object belonging to the training class; and generate a prediction layer indicative of a probability of an instance of image data being representative of a scene comprising an object belonging to said training class from a distribution of said respective output values $\phi$ and $\overline{\phi}$.

Viewed from a second aspect there is provided a method of operating data processing apparatus comprising a processor and memory to train for object recognition, the method comprising:

configuring said processor of said data processing apparatus with a:

convolutional neural network trained for closed-set object recognition of a set of objects of an object training class; and a

prediction layer configured to comprise a first likelihood for an object belonging to said training class and a second likelihood for an object not belonging to said training class;

receiving plural instances of image data to said convolutional neural network, each instance of image data representative of a scene comprising an object belonging to said training class or representative of a scene not comprising an object belonging to said training class;

outputting from said convolutional neural network respective output values $\phi$ from the said convolutional neural network indicative of the image content of said training image data for each of said plural instances of training image data, said value $\phi$ comprising a value $\varphi$ responsive to the input scene comprising an object belonging to the training class and a value $\overline{\phi}$ for the input scene not comprising an object belonging to the training class; and

generating a prediction layer indicative of a probability of an instance of image data being representative of a scene comprising an object belonging to said training class from a distribution of said respective output values $\phi$ and $\overline{\phi}$.

An embodiment may be further configured to provide a convolutional neural network trained for closed-set object recognition of a set of objects of a plurality of object training classes and to provide a prediction layer configured to comprise a first likelihood for an object belonging to each respective training class and a second likelihood for an object not belonging to said each respective training class.

Embodiments in accordance with the first and second aspects provide a prediction layer for the discriminative determination of the probability that an object is present in an image or not present. Thus the object recognition model is trained to recognise when none of the trained objects are present in an input image as well as when trained objects are present, thereby providing a prediction layer suitable for open set recognition.

The data processing circuitry is further configured such that said prediction layer comprises respective probability distribution functions based on output values $\phi$ for objects belonging to a respective training class and not belonging to a respective training class $\overline{\phi}$. Thus, probability distribution functions may be used to discriminate between a likelihood of an object being present in an image and the likelihood of an object not being present in an image.

Suitably, respective probability distribution functions are fitted to a distribution curve. For example, respective probability distribution functions may be fitted to a Gaussian distribution. Other distributions may be utilised at the discretion of the designer of the object recognition model to modify, change or manipulate the recognition profile.

In an embodiment in which object recognition is to be carried out following training, image data for object recognition is received; generation of said prediction layer is inhibited; said output value $\phi$ indicative of the image content of said image data for recognition is output from said convolutional neural network and input to said prediction layer to determine a probability that the scene represented by said image data comprises an object of a one of a trained class; and an indication that said image data comprises an object of a one of a trained class is output for said probability satisfying a threshold criterion.
Thus, a trained embodiment may be used for object recognition if generation of the prediction layer does not occur, ie it is inhibited.

Viewed from a third aspect there is provided a data processing apparatus for object recognition comprising data processing resources comprising a processor and memory, wherein said memory is:

configured to store image data representative of a scene;

configured with a prediction layer indicative of a probability of an instance of image data being representative of a scene comprising an object belonging to a trained class from a distribution of respective output values for input scenes comprising an object belonging to the training class and the input scene not comprising an object belonging to the training class; and

configured with machine-readable instructions executable by said processor to configure said data processing circuitry to:

provide a convolutional neural network trained for closed-set object recognition for the training class for receiving plural instances of image data each representative of a scene;

output from said convolutional neural network an output value $\phi$ indicative of the image content of said image data for recognition;

input said output value $\phi$ to said prediction layer to determine a probability that the scene represented by said image data comprises an object of a one of each trained class; and

output an indication that said image data comprises an object of each trained class for said probability satisfying a threshold criterion.

Viewed from a fourth aspect there is provided a method of operating data processing apparatus for object recognition comprising a processor and memory, the data processing apparatus configured with a prediction layer indicative of a likelihood of an instance of image data being representative of a scene comprising an object belonging to a trained class from a distribution of respective output values for input scenes comprising an object belonging to the training class and the input scene not comprising an object belonging to the training class, the method comprising:

storing image data representative of a scene in said memory;

configuring said processor of data processing apparatus with a convolutional neural network trained for closed-set object recognition for the training class;

receiving to said convolutional neural network plural instances of image data each representative of a scene;

outputting from said closed-set convolutional neural network an output value $\phi$ indicative of the image content of said image data for recognition;

inputting said output value $\phi$ to said prediction layer to determine a probability that the scene represented by said image data comprises an object of each trained class; and

outputting an indication that said image data comprises an object of each trained class for said probability satisfying a threshold criterion.

An embodiment in accordance with the third and fourth aspects does not require a training phase as it is already configured with prediction layers for the trained objects.

Viewed from a fifth aspect there is provided a method of converting a close-set convolutional neural network for closed set recognition to an open-set recognition model, the method comprising removing a soft-max layer from said close-set convolutional neural network, and replacing said soft-max layer with a new prediction layer comprising, for each of the trained objects, two likelihood probability density functions - one for positive (object) class and one for negative (non-object) class, as well as a prior ratio, wherein said two likelihood probability density functions are learned by fitting the feature values extracted from training images. Such an embodiment modifies an existing close-set model to operate as an open-set model. Thus a designer need not create all of the open-set model but utilise an initial part of the close-set model.

Typically, the method further comprises fitting the feature values extracted from training images for each of the trained classes individually.

Viewed from a sixth aspect there is provided a method of predicting the presence of an object in an input image utilising an open-set recognition model derived in accordance with the method set out in one or both of the immediately preceding two paragraphs wherein the posterior probability of each object's existence in the input image is estimated by the prediction layer in accordance with:

$$P_k = \frac{p_k}{p_k + \alpha_k \overline{p_k}},$$

where $\alpha_k = \frac{\overline{\pi_k}}{\pi_k}$ is the prior ratio, and such that if probability $P_k$ meets or exceeds a threshold value the object is considered to exist in the image.

Typically, the method further comprises rejecting an input image for all probabilities from the prediction layer falling below the threshold value.

One or more embodiments recall objects the system was trained to recognize and meanwhile refrain from making unreliable predictions when fed with unfamiliar objects or scenes, resulting in lower false positive rate.

The described system also eliminates, or at least reduces, the need for such a background class.

Utilising a soft-max layer as is conventional in prior art systems converts class-wise features into probabilities through *coupled normalization.* Conversely, in embodiments in accordance with the present invention, each class is processed individually, mapping each feature to a probability score indicating how likely it is that the class is present in the image. Therefore, when the image of an unknown object or scene is presented to the network the probability score of each class is predicted independently If all of the probability scores are sufficiently low the network gracefully rejects the image as an unfamiliar input, instead of making a prediction of one of the trained classes, which it would have to do in closed-set recognition, and which would be an unreliable prediction.

A visual recognition system that is robust in real world scenarios should be able to detect the scenes and objects that fall outside of the trained set and gracefully abstain from making unreliable predictions on those.

In particular, one or more embodiments in accordance with the present invention provide apparatus and or a method that is capable of adapting a pre-trained closed-set recognition model to fulfill such tasks, hence better accommodating real world applications.

The proposed system is capable of adapting a convolutional network which uses soft-max as the classification layer (e.g. a convolutional network such as disclosed in [1] Krizhevsky, Sutskever, and Hinton 2012; [2] Simonyan and Zisserman 2014; [3] Szegedy et al. 2014; [4] He et al. 2015) to accommodate the open-set scenario. One or more embodments reuse the majority of a convolutional network and do not require retraining. Instead, only an add-on learning step is needed operative over the class-wise features that are already available from the training of the existing network.

The system described herein differs from other open-set recognition methodologies, e.g. as disclosed in [6] Bendale, Abhijit, and Terrance Boult "Towards Open Set Deep Networks" *arXiv:1511.06233 [cs],* November 19, 2015. http://arxiv.org/abs/1511.06233 and [7] Bendale, Abhijit, and Terrance Boult "Towards Open World Recognition" *arXiv.1412.5687 [cs],* June 2015, 1893-1902. doi:10.1109/CVPR.2015.7298799, by employing a *discriminative* learning paradigm which effectively exploits information of all training images including the background images. Additionally, one or more embodiments in accordance with the poresent invention may be more flexible than systems using conventional open-set recognition methodologies as the sensitivity of each class can be adjusted individually through the corresponding prior ratio $\alpha_k$ . Finally, the proposed system is conceptually simpler and easier to implement than conventional open-set recognition methodologies.

List of figures

There will now be described one or more embodiments in accordance with the present invention, provided by way of non-limiting example only and with reference to the following drawings:

Figure 1 is a schematic illustration of a filter segment of a feed-forward neural network comprising sparse connectivity and tied weights suitable for receptive field behaviour;

Figure 2 is a schematic illustration of a convolutional neural network for closed-set recognition of an object within an image;

Figure 3 is a schematic illustration of a convolutional neural network in accordance with an embodiment of the present invention;

Figure 4 is a schematic illustration of a system configuration incorporating apparatus comprising a convolutional neural network in accordance with an embodiment of the present invention;

Figure 5 is a graphical representation of an illustrative example of the probability functions for two classes of object for a convolutional neural network in accordance with an embodiment of the present invention;

Figure 6 is a process flow control diagram for the training phase of apparatus 200 in accordance with an embodiment of the present invention; and

Figure 7 is a process flow control diagram for the recognition phase apparatus 200 in accordance with an embodiment of the present invention.

## Description

A fundamental element of a feed-forward convolutional neural network suitable for recognising one or more objects in an image are filter layers which are tied together in order to provide a receptive field behaviour. An example of an arrangement of filter layer for a neural network 50 is schematically illustrated in figure 1. The concept of receptive field behaviour is derived from neurobiology, for example based on the organisation of neurons in the visual cortex of a cat. Arranging a neural network for receptive field behaviour provides for the learning to recognise an object in one location of an image field to be translated or transferred to the same object but located in a different part of the image field without having to have separate weights between respective neurons in different layers as would be required for a fully connected neural network.

In the neural network 50 illustrated in figure 1 there is an input filter layer 52, a hidden layer 54 and an output layer 56. The input filter layer 52 is configured to receive pixel values from a small two-dimensional region of an image undergoing analysis. The output 56 of the neural network 50 is a single value. Neural network 50 acts as a filter on the inputs put into input filter 52. As is now conventional in image processing, input filter layer 52 of neural network 50 is applied repeatedly across an image. That is to say, inputs to input filter layer 52 are the pixel values of respective regions of the image having a size corresponding to the input filter. In the illustrated example, neural network 50 receives as input pixel values from respective groups of five pixels in an image. In other terms, filter layer 52 may be considered to be "moved over" an image so as to have input to the neural network 50 values from respective pixel regions. Applying the neural network 50 filter function across the output of the image at all possible offsets is the convolution that gives convolutional neural networks their name. Filter layer 52 may be moved one pixel at a time across an image or in groups of pixels comprising the width of the filter layer input, i.e. in the described embodiment of five pixels at a time. The particular mode of moving the filter across the image depends on how the model is configured.

Filter weights between each neural connection of input layer 52 and hidden layer 54 are "shared" or "tied" in that they have the same value for the same neural connection vector. That is to say, neural connection vector 53 has the same weight between layers 52 and 54 as illustrated by the long dashed lines, neural connection vector 57 has the same weight between layers 52 and 54 as illustrated by the short dashed lines and neural connection vector 55 has the same weight between layers 52 and 54 as illustrated by the dashed and dotted lines. Weights between respective neural connection vectors 58, 59 and 60 may be different as illustrated by different dashed or dotted outline. Thus, as neural network 50 is moved across the image the same filter is applied at each image region, thereby introducing position invariance into the filter function of neural network 50. Such position invariance may be considered to be an inherent quality of a convolutional neural network when utilised as an image filter. The limited breadth of the input filter 52 also exploits spatially local correlation thereby generating filters which may produce a strong response to a spatially local input pattern. Consequently, a convolutional neural network may provide a position invariant strong response to a local input pattern.

A general outline of a conventional closed-set recognition apparatus 100 will now be described with reference to figure 2 of the drawings. The apparatus 100 receives an image 102 comprising a feature or object 104. Feature 104 belongs to a class of a set of features for which the apparatus was trained as this is a closed-set recognition apparatus. The number of classes may be designated as "C". Such a conventional closed-set recognition apparatus 100 comprises a deep convolutional neural network (CONVNET) for visual recognition consisting of feature extraction layers 106, 108, 110 and 112 followed by a classification layer 114 that outputs a final prediction of the class of object in the input image 102. The final predictions ϕï - ϕC are input to a so-called "soft-max" layer 118 which determines the probability of a class of object being present in image 102 from a log-normalised analysis of the ϕ values output from layer 114.

The input image typically is in 2-dimensional array of pixel values which in a monochrome image may have a single value of "0" or *"1"* indicating either a white or black pixel. For "grey-scale" images each pixel may have a value represented by a byte, i.e. an eight digit binary number, having 256 different values of grey-scale running from zero giving a white pixel to 255 giving a black pixel. Colour images typically have three pixels for any image point, the pixels being red, green and blue and generally having an intensity governed by a byte, i.e. 256 different values of intensity. Consequently, an image is represented by a two-dimensional array of binary numbers.

Each extraction layer filters the input image 102 in a manner that enhances, or at least draws out relative to the rest of the image, characteristics of the feature or object contained in the input image and upon which the apparatus 100 is being or has been trained. Conventionally, convolutional encoding is carried out on image 102 in extraction layer 106 which creates a feature map comprising pixel data output 105 from the filter, e.g. neural network 50 of figure 1, convolved across image 102 and in the specific example illustrated in figure 2 data 105 corresponds to pixel data from a region 103 of image 102. Pixel data 107 in feature map 106 may be reduced by so-called "pooling" or "max-pooling" in pooling layer 108 where representative values 109 for respective groups of pixels such as in a 3 x 3 array is evaluated. Values 111 from the pooling layer undergo convolution in layer 110 resulting in filter output 113. Pixel data 115 convolution layer 110 is then pooled 117 in pooling layer 112 resulting in a final feature map of image 102 represented by a reduced dataset array compared to the number of pixels (image data points) representing the original image 102. Figure 2 illustrates for each extraction layer plural filters 106a, 106 b, 106c ... 106m; 108a, 108b, 108c ... 108n; 110a, 110b, 110c ... HOo; and 112a, 112b, 112c ... 112p each producing a feature map, for convenience referred to herein by the same reference as the corresponding layer filters. Each chain of extraction layer filters results in respective final feature map 112a, 112b, 112c ... 112p.

Final feature maps 112a...112p are input to a fully connected feed forward layer 114 and eventually result in the phi ("φ") layer.

Apparatus 100 is trained for object recognition and typically employs a so-called soft-max layer 118 as a classification layer, such as described in [1] Krizhevsky, Sutskever, and Hinton 2012; [2] Simonyan and Zisserman 2014; [3] Szegedy et al. 2014; [4] He et al. 2015, which is designed for closed-set recognition problems, where the input image is classified into one of the pre-defined set of object classes. Training comprises inputting images 102 including one of the object classes 104 in the closed-set. The soft-max layer 118 performs multi-class classificationby mapping the output values from the previous layer 114 (usually a fully-connected layer as described above) into posterior probabilities that sum up to one. The outputs of the previous layer 114 are called phi ( φ ) and comprise a number of values $\varphi_1$, $\varphi_2$, ..., $\varphi_k$, ..., $\varphi_C$ (where $C$ is the number of trained object classes) which are representative of class-wise features, since each of the values φ corresponds to one trained class and can be considered as the evidence of such a class existing in the input image. Mathematically, the output probability score of the $k$-th class is:

$$Q_k = \exp(\varphi_k)/\sum_{j=1}^{C} \exp(\varphi_j) . \qquad (1)$$

As its name suggests, the soft-max layer compares the class-wise features φ and assigns a high probability to the class with the largest feature, i.e. value of φ, in accordance with equation (1). Such an operation is natural for closed-set recognition, since every input image is known to capture an object belonging to one of the trained classes, and the corresponding class-wise feature is expected to be larger than others. However, under an open-set scenario a soft-max layer approach is problematic. When apparatus 100 receives an image comprising an unknown object or scene which class-wise feature would have the largest result is unpredictable, leading to unreliable probability scores.

Embodiments in accordance with the present invention comprise a recognition system which has replaced the soft-max layer with a different prediction layer that maps the class-wise features to probabilities in a different way from a soft-max layer. That is to say, the neural network is trained with the soft-max layer in place and then the soft-max layer is removed or discarded. Figure 3 schematically illustrates the architecture of an embodiment in accordance with an aspect of the present invention comprising apparatus 200 which may be trained to recognise one or more classes of object for which it has been trained from input images which may comprise one or more objects for which the apparatus 200 has not been trained; classify as not having an object class images which do not comprise one of the trained classes of object; or classify as not having an object class images which do not comprise one of the trained classes of object and input image empty of any object, i.e. an image comprising merely background. Apparatus 200 comprises convolution and pooling pairs 206 and 208 together with fully connected layer 210 (210 corresponds to 114 and 118 in Figure 2) and prediction layer 212. Input images 202 and 203 may comprise respective classes of object which in the described example are respectively a triangle 204 and cross 205. Additionally, images 207 which do not comprise objects for which apparatus 200 has been configured to recognise or are empty of any object 209 may also be input to apparatus 200. Convolution/pooling pairs 206 and 208 have a similar function and indeed may be configured in a similar manner to respective convolutional pooling pairs 106/108 and 110/112. Convolution/pooling pairs 206 and 208 may have the same weight and other configuration parameters as 106/108 and 110/112 to the extent apparatus 200 is configured to recognise an open set environment having the same or similar class objects that apparatus 100 is configured to recognise as part of a closed set.

Figure 4 is a schematic illustration of a server-based embodiment in accordance with the present invention. Server 252 comprises processing resources for implementing apparatus 200 and includes an application program interface 254, a processor 256 and an interface 258. Server 252 also includes memory 262 which stores various program instructions executable by processor 256 for implementing object recognition as well as providing a data store 266. For example, memory 262 may include computer program instructions 264 executable by processor 256 for implementing a convolutional network, computer program instructions 268 executable by processor 256 to implement an application for object recognition and computer program instructions executable by processor 256 for implementing a class feature probability mapping function 270 in accordance with an embodiment of the present invention. The various processing resources and memory are coupled together via data and instruction bus 260. A scanner 272, camera 274 or other image source 276 may be coupled directly to the server to interface 258. Optionally or additionally, scanner 272, camera 274 and image source 276 may be coupled through a computer network, such as the Internet, 278 and into server 252 via API 254.

The class feature probability mapping module 270 comprises processor executable instructions for both a training phase mode and also normal recognition mode. Operation of server 252 will now be described in both the training and a normal recognition mode.

In general terms, for each class $k$ (with reference to figure 3 triangle 204 or cross 205), probability mapping module 270 comprises instructions which generate positive and negative likelihood functions from the training data. The positive likelihood $p_k(\varphi_k | k \in I)$ is the conditional probability density function (p.d.f.) of the $k$-th class-wise feature for and object belonging to class $k$ being present in image $I$. It can be estimated from all training images having a class $k$ object present in them. Similarly, the negative likelihood, $\bar{p}_k(\varphi_k | k \notin I)$ can be estimated from all training images not containing objects belonging to class $k$ (i.e. negative images). Such negative images include both those belonging to other classes, e.g. cross 205 when training for triangle 204, and those not containing any trained classes at all (i.e. background images 209). $p_k$ and $\bar{p}_k$ can take any parametric or nonparametric form. For instance, in the described embodiment referring to one specific implementation $p_k$ and $\bar{p}_k$ are Gaussian functions. The output probability score for an image I including an object belonging to class $k$ is determined by Bayes' theorem:

$$P_k = \frac{p_k}{p_k + \alpha_k \overline{p_k}}, \text{ where } \alpha_k = \frac{\overline{\pi_k}}{\pi_k} \text{ is the prior ratio.} \qquad (2)$$

As is clear, the greater the value of a the lower the probability score for class $k$ and therefore the class would be recognized more conservatively. Additionally, $\alpha_k$ can be optimized as a hyper-parameter.

Figure 5 is a schematic illustration of probability distribution functions generated in respect of two respective class of objects having images input to apparatus 200 comprising objects of respective classes and also objects which do not fall within the trained class of objects, comprise merely background information, or are part of the training of a first object but comprise a second object. For the described embodiment, 310 represents the training result for a first class of object being present in an input image and curve 312 represents the training result from images not comprising the first object. Conversely, 316 is the training result for images comprising a second class of object and curve 314 for images not containing the second class of object. As will be readily apparent to a person of ordinary skill in the art, figure 5 is a schematic representation of the distribution of results and is provided graphically for ease of understanding. Such results will be stored in data store 266 of memory 262 of server 252 illustrated in figure 4. The histogram under curve 310 is representative of the manner in which results are collected, collated and stored in the currently described embodiment. The probability distribution functions illustrated in figure 5 in effect provide the prediction layer 212 referred to with reference to figure 3.

*Training*

Turning now to figure 6, a process flow control diagram 350 schematically illustrating the steps in training apparatus 200 to recognise an object class starts at step 352. Processor executable instructions for implementing an image training and recognition process in accordance with process flow control diagram 350 are stored as application software 268 in memory 262. At step 353 the apparatus is selected to be configured for an image containing an object of the class for which the apparatus is undergoing training ($\phi$ value output) or configured for an image not containing object of the class for which the apparatus is undergoing training ($\bar{\phi}$ value output) and at step 354 the image is input to the apparatus 200. At step 356, application software 268 invokes the convolutional neural network application 264 which applies convolutional neural network processing to the input image. As was described with reference to figure 2 and figure 3 the output of the convolutional and pooling layers is input to fully connected layers 210. Application 268 then invokes the class feature probability mapping software 270 which stores, step 360, the output, $\phi$, from the fully connected layers 210 to a positive ($\phi$ value output) or negative ($\bar{\phi}$ value output) set. At step 362 application 268 determines if the user has input an indication that the training has finished, 362. If no indication that the user has finished training apparatus 200 has been received then process flow control returns to step 354 where the next image is input to apparatus 200 and the process proceeds through steps 356 through to step 362 where the process again determines whether not an instruction to finish training has been received.

During the iteration through steps 353 to 362, a number of images are input into apparatus 200 and a $\phi$ value determined for each image representative of an image containing the particular class of object for which the apparatus is being trained, e.g. in the present example a triangle or a cross. In the described embodiment output value corresponding to an image containing the particular class of object for which the apparatus is being trained is represented by $\phi$. In accordance with one or more embodiments images not containing the particular class of object for which the apparatus is currently being trained, are also input. For example, if the apparatus is undergoing training for a triangle then images not containing a triangle, including an image containing a cross, will be input. For the purposes of clarification and distinction between outputs, the output for an image identified as not containing an object of the class for which the apparatus 200 is undergoing training is represented by $\overline{\phi}$. However, when describing the process in general the symbol $\phi$ is used. All possible values for $\phi$ are divided into "bins" representative of a small range of values for $\phi$ so that a histogram of $\phi$ values may be generated as schematically illustrated in figure 5 for probability distribution function 310. The "bins" are configured in data store 266 and the value for each "bin" is increased, typically by 1, for each $\phi$ value assigned to the bin at step 360. Only the one set of histograms is illustrated for clarity but of course a set of histograms will be produced for each of the curves 312, 314 and 316 illustrated in figure 5.

In response to application 268 determining that the indication training has finished has been received process flow control proceeds to step 364 at which application 268 invokes class feature probability mapping application 270. In the currently described embodiment, class feature probability mapping application 270 fits the histogram collected for the set of training images input for the training phase to produce a curve such as one of curves 310-316 illustrated in figure 5. The line fitted data is then stored for respective classes in data store 266, step 366.

Process flow control then flows to step 368 where training for the class on the training is finished. Training for another class of objects may then take place or the training phase may be halted if all classes of objects have been trained. It should be noted that training for a particular class includes the inputting of images which do not contain the respective class of objects as part of training the apparatus 200 to recognise when an image does not contain a particular class of object, e.g. curve 312 of the probability distribution functions illustrated in figure 5. Curve 312 is representative of images which do not have the class of objects present in them for which curve 310 is derived (i.e. curve 310 is derived from images which do have that class of object present).

In an optional or alternative embodiment, apparatus 200 may already be configured to recognise objects using a "soft-max" layer as known in the prior art and discussed in the introductory description herein. In such an already configured apparatus, probability distributions for outputs of the convolutional neural network, elements 206, 208 and 210 of apparatus 200 illustrated in figure 3 and implemented by module 264 of server 252, for each class of object will have already been derived and stored in data store 266. For example, images comprising a triangle and images comprising a cross will have been input to train apparatus 200 to recognise images comprising a triangle or a cross. Such probability distributions may have the form of curves 310 and 314 illustrated in figure 5, there being no corresponding curves 312 and 316. For such an optional or alternative embodiment the training process is slightly modified. Referring now to figure 6 the modification is such that step 353 is fixed to select training using images which do not include an object of the class undergoing training. For example, training to recognise a triangle will require the input of images not including a triangle to create an appropriate probability distribution function. Referring now to figure 4, if curve 310 is a probability distribution representative of images which contain a triangle then curve 312 will be the result of the further training phase of inputting images not including a triangle. The process as illustrated in figure 6 continues in the same way as described earlier until the training for a particular class of object is completed. Then the training will continue for the next class of object by inputting images not containing that class of object.

In this way, for an illustrative example comprising recognition of just two classes of object probability distribution functions as illustrated in figure 4, i.e. curve 310, 312, 314 and 316, will be generated and the corresponding data stored in memory 266. Thus, the recognition process may be the same for embodiments which utilise existing probability distribution functions as well as embodiments in which the apparatus is trained from a non-trained start point.

*Recognition*

Turning now to figure 7, a process flow control diagram 400 schematically illustrating the steps in apparatus 200 to recognise an object class in an input image starts at step 402. Processor executable instructions for implementing an image training and recognition process in accordance with process flow control diagram 400 are stored as application software 268 in memory 262. At step 404 an image is input to the apparatus 200 and at step 406 application software 268 invokes the convolutional neural network application 264 which applies convolutional neural network processing to the input image. As was described with reference to figure 2 and figure 3 the output of the convolutional and pooling layers is input to fully connected layers 210, step 408. Application 268 then invokes the class feature probability mapping software 270 and inputs the resulting $\phi$ value to the class feature probability mapping software 270, step 410. The class feature probability mapping software 270 feeds the input $\phi$ value into the probability distribution functions stored in data store 266 to compute the posterior probability according to Equation (2), step 412. Using Equation (2) above to determine probability is sufficient since once the probability density function is defined the computation utilising equation (2) is the application of the Bayesian theorem.

At step 414 the recognition result is output, which includes a "no recognised object" output, which occurs when every class has a posterior probability lower than a specified threshold. A typical threshold is 0.5 but the threshold used in any particular implementation is a matter of design choice for the designer of the recognition system. A threshold greater than 0.5 may be used for a system in which a high certainty in the recognition of objects is desired and conversely a lower threshold may be used if a lower certainty is required. At step 416 it is determined whether or not the user has requested that another image is recognised and if so process flow control returns to step 404. Otherwise, process flow control proceeds to step 418 at which the recognition procedure finishes.

The term "empty" is used herein in connection with image data to denote image data representative of a scene not containing a specific class of feature relevant to the context. For example, in a training phase for class $k$ features an empty scene or empty image data would be a scene not containing a $k$ class feature or image data representative of such a scene. An empty scene or empty image data may relate to a scene or data comprising a feature other than the specific class of feature relevant to the context, i.e. in the present discussion a non- $k$ class feature for example a $l$ class feature, as well as a scene or image data comprising no features and which may be described as a "background" image.

Although embodiments have been described utilising convolution and pooling pairs, pooling, max pooling or sub- sampling need not be utilised.

It will be appreciated that any of the optional features of any of the embodiments described herein could also be provided with one or more of any of the other embodiments described herein.

As noted in some of the embodiments above, they may be embodied in hardware. The hardware may be referenced as a hardware element. In general, a hardware element may refer to any hardware structures arranged to perform certain operations. In one embodiment, for example, the hardware elements may include any analogue or digital electrical or electronic elements fabricated on a substrate. The fabrication may be performed using silicon-based integrated circuit (IC) techniques, such as complementary metal oxide semiconductor (CMOS), bipolar, and bipolar CMOS (BiCMOS) techniques, for example. Examples of hardware elements may include processors, microprocessors, circuits, circuit elements (e.g., transistors, resistors, capacitors, inductors, and so forth), integrated circuits, application specific integrated circuits (ASIC), programmable logic devices (PLD), digital signal processors (DSP), field programmable gate array (FPGA), logic gates, registers, semiconductor device, chips, microchips, chip sets, and so forth. The embodiments are not limited in this context. Also noted above, some embodiments may be embodied in software. The software may be referenced as a software element. In general, a software element may refer to any software structures arranged to perform certain operations. In one embodiment, for example, the software elements may include program instructions and/or data adapted for execution by a hardware element, such as a processor. Program instructions may include an organized list of commands comprising words, values or symbols arranged in a predetermined syntax, that when executed, may cause a processor to perform a corresponding set of operations.

The software may be written or coded using a programming language. Examples of programming languages may include C, C++, BASIC, Perl, Matlab, Pascal, Visual BASIC, JAVA, ActiveX, assembly language, machine code, and so forth. The software may be stored using any type of computer-readable media or machine-readable media. Furthermore, the software may be stored on the media as source code or object code. The software may also be stored on the media as compressed and/or encrypted data. Examples of software may include any software components, programs, applications, computer programs, application programs, system programs, machine programs, operating system software, middleware, firmware, software modules, routines, subroutines, functions, methods, procedures, software interfaces, application program interfaces (API), instruction sets, computing code, computer code, code segments, computer code segments, words, values, symbols, or any combination thereof. The embodiments are not limited in this context.

Some embodiments may be implemented, for example, using any computer-readable media, machine-readable media, or article capable of storing software. The media or article may include any suitable type of memory unit, memory device, memory article, memory medium, storage device, storage article, storage medium and/or storage unit, such as any of the examples described with reference to a memory. The media or article may comprise memory, removable or non-removable media, erasable or non-erasable media, writeable or re-writeable media, digital or analog media, hard disk, floppy disk, Compact Disk Read Only Memory (CD-ROM), Compact Disk Recordable (CD-R), Compact Disk Rewriteable (CD-RW), optical disk, magnetic media, magneto-optical media, removable memory cards or disks, various types of Digital Versatile Disk (DVD), subscriber identify module, tape, cassette, electrical signal, radio-frequency signal, optical carrier signal, orthe like. The instructions may include any suitable type of code, such as source code, object code, compiled code, interpreted code, executable code, static code, dynamic code, and the like. The instructions may be implemented using any suitable high-level, low-level, object-oriented, visual, compiled and/or interpreted programming language, such as C, C++, Java, BASIC, Perl, Matlab, Pascal, Visual BASIC, JAVA, ActiveX, assembly language, machine code, and so forth. The embodiments are not limited in this context.

Unless specifically stated otherwise, it may be appreciated that terms such as "processing," and the routine "computing," "calculating," "determining," or the like, refer to the action and/or processes of a computer or computing system, or similar electronic computing device, that manipulates and/or transforms data represented as physical quantities (e.g., electronic) within the computing system's registers and/or memories into other data similarly represented as physical quantities within the computing system's memories, registers or other such information storage, transmission or viewing devices. The embodiments are not limited in this context.

As used herein any reference to "one embodiment" or "an embodiment" means that a particular element, feature, structure, or characteristic described in connection with the embodiment is included in at least one embodiment. The appearances of the phrase "in one embodiment" or the phrase "in an embodiment" in various places in the specification are not necessarily all referring to the same embodiment.

As used herein, the terms "comprises," "comprising," "includes," "including," "has," "having" or any other variation thereof, are intended to cover a non-exclusive inclusion. For example, a process, method, article, or apparatus that comprises a list of elements is not necessarily limited to only those elements but may include other elements not expressly listed or inherent to such process, method, article, or apparatus. Further, unless expressly stated to the contrary, "or" refers to an inclusive or and not to an exclusive or. For example, a condition A or B is satisfied by any one of the following: A is true (or present) and B is false (or not present), A is false (or not present) and B is true (or present), and both A and B are true (or present).

In addition, use of the "a" or "an" are employed to describe elements and components of the invention. This is done merely for convenience and to give a general sense of the invention. This description should be read to include one or at least one and the singular also includes the plural unless it is obvious that it is meant otherwise.

In view of the foregoing description it will be evident to a person skilled in the art that various modifications may be made within the scope of the invention. For example, although an embodiment has been described utilising a normal Gaussian distribution to which the likelihood estimations p and $\bar{p}$ are fitted, other distributions may be utilised, for example multiples of a Gaussian distribution may be utilised as well as other generally bell-shaped distributions such as Cauchy, Student's t-distribution , and logistic distributions or any other suitable distribution. Additionally, although an embodiment in accordance with the present invention has been described utilising "binning" of results and fitting such bins to a Gaussian function the ordinarily skilled person will readily understand that binning results is not necessary and that result could be fitted to a Gaussian function, or other function, in a parametric manner.

Although an embodiment in accordance with the present invention has been described with reference to images 202 and 203 which may be considered to only comprise a single shape 204 and 205 respectively, the images the not be so simplistic and may comprise other features and background features as would normally be the case in images obtained in the real world. In practice, real world images having other features and background features in them will be used in training in order to generate a probability distribution function suited for real world applications. In the described embodiment, $\phi$ values are assigned to respective "bins", each "bin" covering a range of possible $\phi$ values in order to get a meaningful distribution or histogram to which a curve may be fitted. In an optional or additional embodiment, $\phi$ values are not assigned to respective "bins" as they are evaluated. Instead, respective $\phi$ values are stored and then collected into suitable equal range groups. Class feature probability mapping module 270 then determines how many instances of $\phi$ values falls within respective equal range groups to determine a suitable histogram to which a line may be fitted. However, if a very large number of training images are utilised it would not be necessary to assign $\phi$ values to "bins" since a sufficient number of values will be collected for a meaningful distribution to which a Gaussian curve, or other suitable distribution, may be fitted. Optionally, if Gaussian or other parametric form of probability density function is to be used, no binning is needed at all, with any number of training images the system may be configured to calculate the parameters from the $\phi$ values directly.

The specific embodiment described herein has been disclosed with reference to a training phase. However, if an existing deep learning convolutional neural network for a closed-set recognition system is adapted to replace the "soft-max" layer with the probability distribution function layer in accordance with an embodiment of the present invention the deep learning convolutional neural network need not be retrained for respective classes of objects since it is already configured by virtue of its previous training and use for a closed-set recognition system. Instead, the training merely comprises inputting images so that the probability distribution functions may be derived for use in an open-set system.

The described embodiment utilises not just images comprising objects of the classes to be recognised but also images which do not include an object to be recognised or in the class - wise feature training phase including an object of another class to that of the class for which the apparatus 200 is undergoing training.

The scope of the present disclosure includes any novel feature or combination of features disclosed therein either explicitly or implicitly or any generalisation thereof irrespective of whether or not it relates to the claimed invention or mitigate against any or all of the problems addressed by the present invention. The applicant hereby gives notice that new claims may be formulated to such features during prosecution of this application or of any such further application derived therefrom. In particular, with reference to the appended claims, features from dependent claims may be combined with those of the independent claims and features from respective independent claims may be combined in any appropriate manner and not merely in specific combinations enumerated in the claims.

**Works Cited**

[1] Krizhevsky, Alex, Ilya Sutskever, and Geoff Hinton. "ImageNet Classification with Deep Convolutional Neural Networks." In *Advances in Neural information Processing Systems 25,* 1106-14, 2012. http://www.cs.toronto.edu/~hinton/absps/imagenet.pdf.

[2] Simonyan, Karen, and Andrew Zisserman. "Very Deep Convolutional Networks for Large-Scale Image Recognition." *arXiv:1409.1556 [cs],* September 4, 2014. http://arxiv.org/abs/1409.1556.

[3] Szegedy, Christian, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. "Going Deeper with Convolutions." *arXiv.-1409.4842 [cs],* September 16, 2014. http://arxiv.org/abs/1409.4842.

[4]He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep Residual Learning for Image Recognition." *arXiv:1512.03385 [cs],* December 10, 2015. http://arxiv.org/abs/1512.03385.

[5] Girshick, Ross. "Fast R-CNN." *arXiv.1504.08083 [cs],* April 30, 2015. http://arxiv.org/abs/1504.08083.

[6] Bendale, Abhijit, and Terrance Boult. "Towards Open Set Deep Networks." *arXiv:1511.06233 [cs],* November 19, 2015. http://arxiv.org/abs/1511.06233.

[7] Bendale, Abhijit, and Terrance Boult. "Towards Open World Recognition." *arXiv.1412.5687 [cs],* June 2015, 1893-1902. doi:10.1109/CVPR. 2015. 7298799.

Claims:

1. Data processing apparatus comprising data processing resources comprising a processor and memory, said memory configured to store image data representative of a scene and configured with machine-readable instructions executable by said processor to configure said data processing circuitry to:

provide a convolutional neural network trained for closed-set object recognition of a set of objects of an object training class;

provide a prediction layer configured to comprise a first likelihood for an object belonging to said training class and a second likelihood for an object not belonging to said training class;

receive plural instances of image data each representative of a scene comprising an object belonging to said training class or representative of a scene not comprising an object belonging to said training class and outputting for each of said plural instances of training image data respective output values $\phi$ from the said convolutional neural network indicative of the image content of said training image data, said value $\phi$ comprising a value $\phi$ responsive to the input scene comprising an object belonging to the training class and a value $\overline{\phi}$ for the input scene not comprising an object belonging to the training class; and

generate a prediction layer indicative of a probability of an instance of image data being representative of a scene comprising an object belonging to said training class from a distribution of said respective output values $\phi$ and $\overline{\phi}$.

2. Data processing apparatus according to claim 1, further configured provide a convolutional neural network trained for closed-set object recognition of a set of objects of a plurality of object training classes and to provide a prediction layer configured to comprise a first likelihood for an object belonging to each respective training class and a second likelihood for an object not belonging to said each respective training class.

3. Data processing apparatus according to claim 1 or claim *2,* wherein said data processing circuitry is further configured such that said prediction layer comprises respective probability distribution functions based on output values $\phi$ for objects belonging to a respective training class and not belonging to a respective training class
—
<p.

4. Data processing apparatus according to claim 3, wherein said data processing circuitry is further configured such that said respective probability distribution functions are fitted to a distribution curve.

5. Data processing apparatus according to claim 4, wherein said data processing circuitry is further configured such that said respective probability distribution functions are fitted to a Gaussian distribution .

6. Data processing apparatus according to any preceding claim, said machine-readable instructions executable by said processor to further configure said data processing apparatus to:

receive image data for object recognition;

inhibit generation of said prediction layer;

output from said convolutional neural network said output value $\phi$ indicative of the image content of said image data for recognition;

input said output value $\phi$ to said prediction layer to determine a probability that the scene represented by said image data comprises an object of a one of a trained class; and

output an indication that said image data comprises an object of a one of a trained class for said probability satisfying a threshold criterion .

7. Data processing apparatus for object recognition comprising data processing resources comprising a processor and memory, wherein said memory is:

configured to store image data representative of a scene;

configured with a prediction layer indicative of a likelihood of an instance of image data being representative of a scene comprising an object belonging to a trained class from a distribution of respective output values for input scenes comprising an object belonging to the training class and the input scene not comprising an object belonging to the training class; and

configured with machine-readable instructions executable by said processor to configure said data processing circuitry to:

provide a convolutional neural network trained for closed-set object recognition for the training class for receiving plural instances of image data each representative of a scene;

output from said convolutional neural network an output value $\phi$ indicative of the image content of said image data for recognition;

input said output value $\phi$ to said prediction layer to determine a probability that the scene represented by said image data comprises an object of a one of a trained class; and

output an indication that said image data comprises an object of said trained class for said probability satisfying a threshold criterion.

8. Data processing apparatus according to claim 7, further configured provide a convolutional neural network trained for closed-set object recognition of a set of objects of a plurality of object training classes and to provide a prediction layer configured to comprise a first likelihood for an object belonging to each respective training class and a second likelihood for an object not belonging to said each respective training class.

9. A method of operating data processing apparatus comprising a processor and memory to train for object recognition, the method comprising:

configuring said processor of said data processing apparatus with a:

convolutional neural network trained for closed-set object recognition of a set of objects of an object training class; and a prediction layer configured to comprise a first likelihood for an object belonging to said training class and a second likelihood for an object not belonging to said training class;

receiving plural instances of image data to said convolutional neural network, each instance of image data representative of a scene comprising an object belonging to said training class or representative of a scene not comprising an object belonging to said training class;

outputting from said convolutional neural network respective output values $\phi$ indicative of the image content of said training image data for each of said plural instances of training image data, said value $\phi$ comprising a value $\varphi$ responsive to the input scene comprising an object belonging to the training class and a value $\overline{\phi}$ for the input scene not comprising an object belonging to the training class; and

generating a prediction layer indicative of a probability of an instance of image data being representative of a scene comprising an object belonging to said training class from a distribution of said respective output values $\phi$ and $\overline{\phi}$.

10. A method according to claim 9, further comprising providing a convolutional neural network trained for closed-set object recognition of a set of objects of a plurality of object training classes and to provide a prediction layer configured to comprise a first likelihood for an object belonging to each respective training class and a second likelihood for an object not belonging to said each respective training class.

11. A method according to claim 9 or claim 10, wherein said prediction layer comprises respective probability distribution functions based on output values $\phi$ for objects belongingto a respective training class and not belongingto a respective training class $\overline{\phi}$.

12. A method according to claim 11, further comprising fitting said respective probability distribution functions to a distribution curve.

13. A method according to claim 12, wherein said respective probability distribution functions are fitted to a Gaussian distribution.

14. A method according to any of claim 9 to claim 13, further comprising:

receiving image data for object recognition at said convolutional neural network;

inhibiting generation of said prediction layer at said convolutional neural network;

outputting from said convolutional neural network said output value $\phi$ indicative of the image content of said image data for recognition;

inputting said output value $\phi$ to said prediction layer to determine a probability that the scene represented by said image data comprises an object of a one of a trained class; and

outputting an indication that said image data comprises an object of a one of a trained class for said probability satisfying a threshold criterion.

15. A method of operating data processing apparatus for object recognition comprising a processor and memory, the data processing apparatus configured with a prediction layer indicative of a probability of an instance of image data being representative of a scene comprising an object belonging to a trained class from a distribution of respective output values for input scenes comprising an object belonging to the training class and the input scene not comprising an object belonging to the training class, the method comprising:

storing image data representative of a scene in said memory;

configuring said processor of data processing apparatus with a convolutional neural network trained for closed-set object recognition for the training class;

receiving to said convolutional neural network plural instances of image data each representative of a scene;

outputting from said convolutional neural network an output value $\phi$ indicative of the image content of said image data for recognition;

inputting said output value $\phi$ to said prediction layer to determine a probability that the scene represented by said image data comprises an object of said trained class; and outputting an indication that said image data comprises an object of said trained class for said probability satisfying a threshold criterion.

16. A method according to claim 15, wherein said data processing apparatus is configured with a prediction layer configured to comprise a first likelihood for an object belonging to a respective training class of a plurality of training classes and a second likelihood for an object not belonging to said respective training class.

17. A method of converting a close-set convolutional neural network for closed set recognition to an open-set recognition model, the method comprising removing a soft-max layer from said close-set convolutional neural network, and replacing said soft-max layer with a new prediction layer comprising, for each of the trained objects, two likelihood probability density functions - one for positive (object) class and one for negative (non-object) class, as well as a prior ratio, wherein said two likelihood probability density functions are learned by fitting the feature values extracted from training images.

18. A method according to claim 17, further comprising fitting the feature values extracted from training images for each of the trained classes individually.

19. A method of predicting the presence of an object in an input image utilising an open-set recognition model derived in accordance with the method of claim 9 or claim 10, wherein the posterior probability of each object's existence in the input image is estimated by the prediction layer in accordance with :

$$P_k = \frac{p_k}{p_k + \alpha_k \bar{p}_k},$$

where $\alpha_k = \frac{\bar{\pi}_k}{\pi_k}$ is the prior ratio, and such that if probability $P_k$ meets or exceeds a threshold value the object is considered to exist in the image.

20. A method according to claim 19, further comprising rejecting an input image for all probabilities from the prediction layer falling below the threshold value.

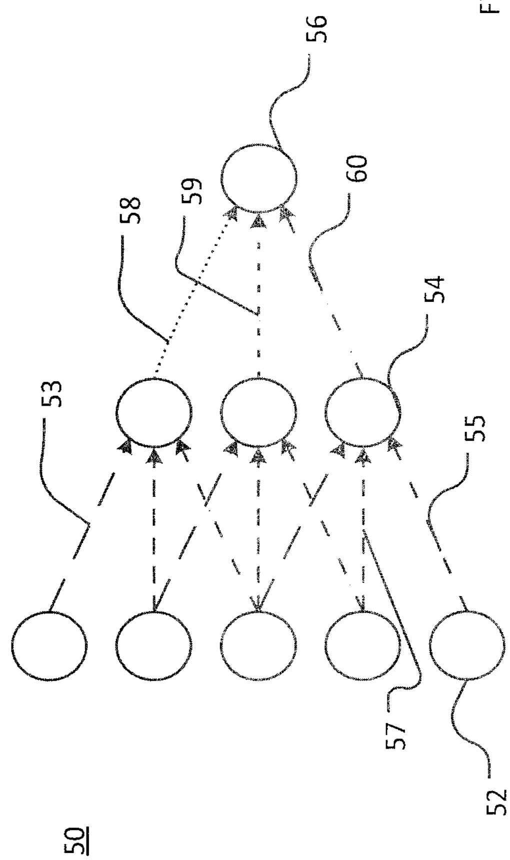

50
52
53
54
55
56
57
58
59
60

Fig. 1

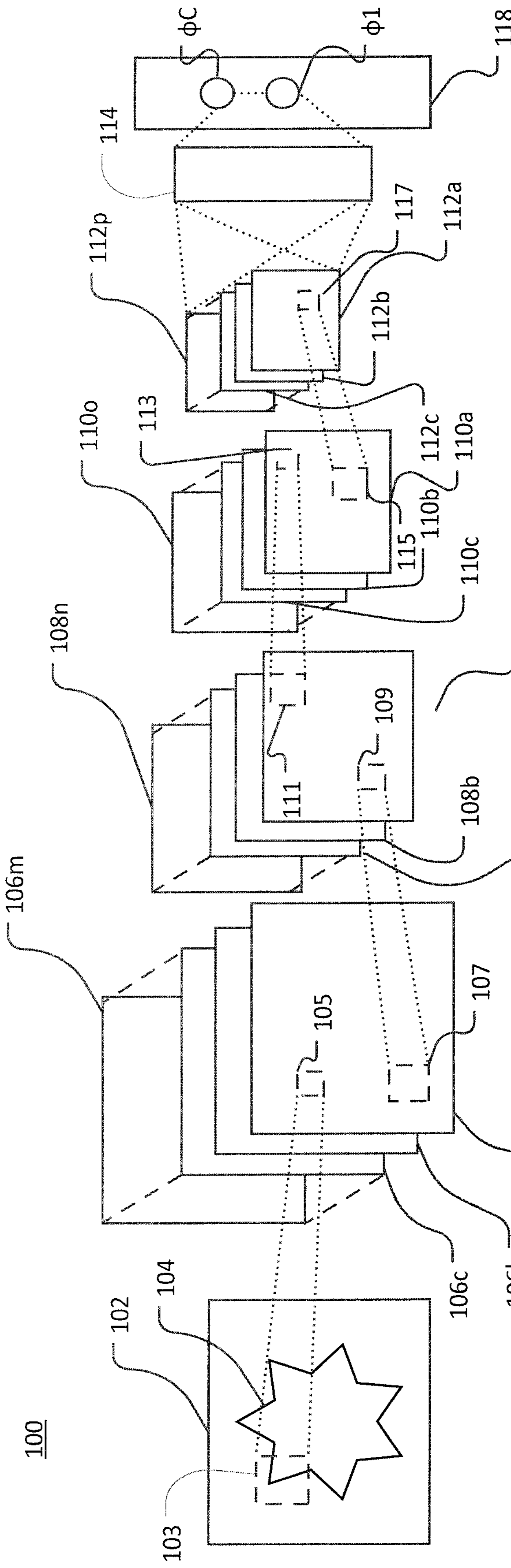

100
102
103
104
105
106a
106b
106c
106m
107
108a
108b
108c
108n
109
110a
110b
110c
110o
111
112a
112b
112c
112p
113
114
115
117
118
ϕC
ϕ1

Fig. 2

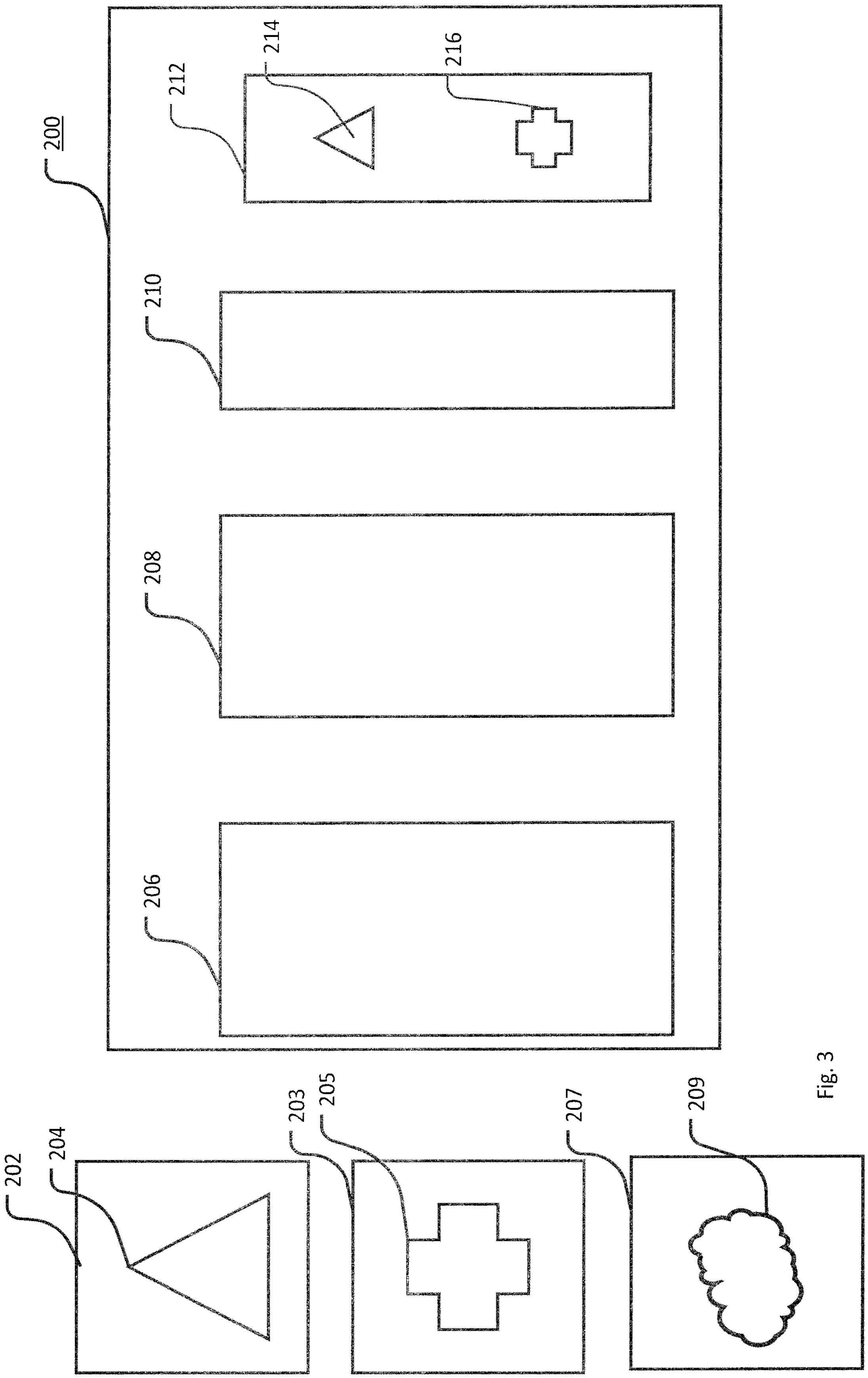



Fig. 3

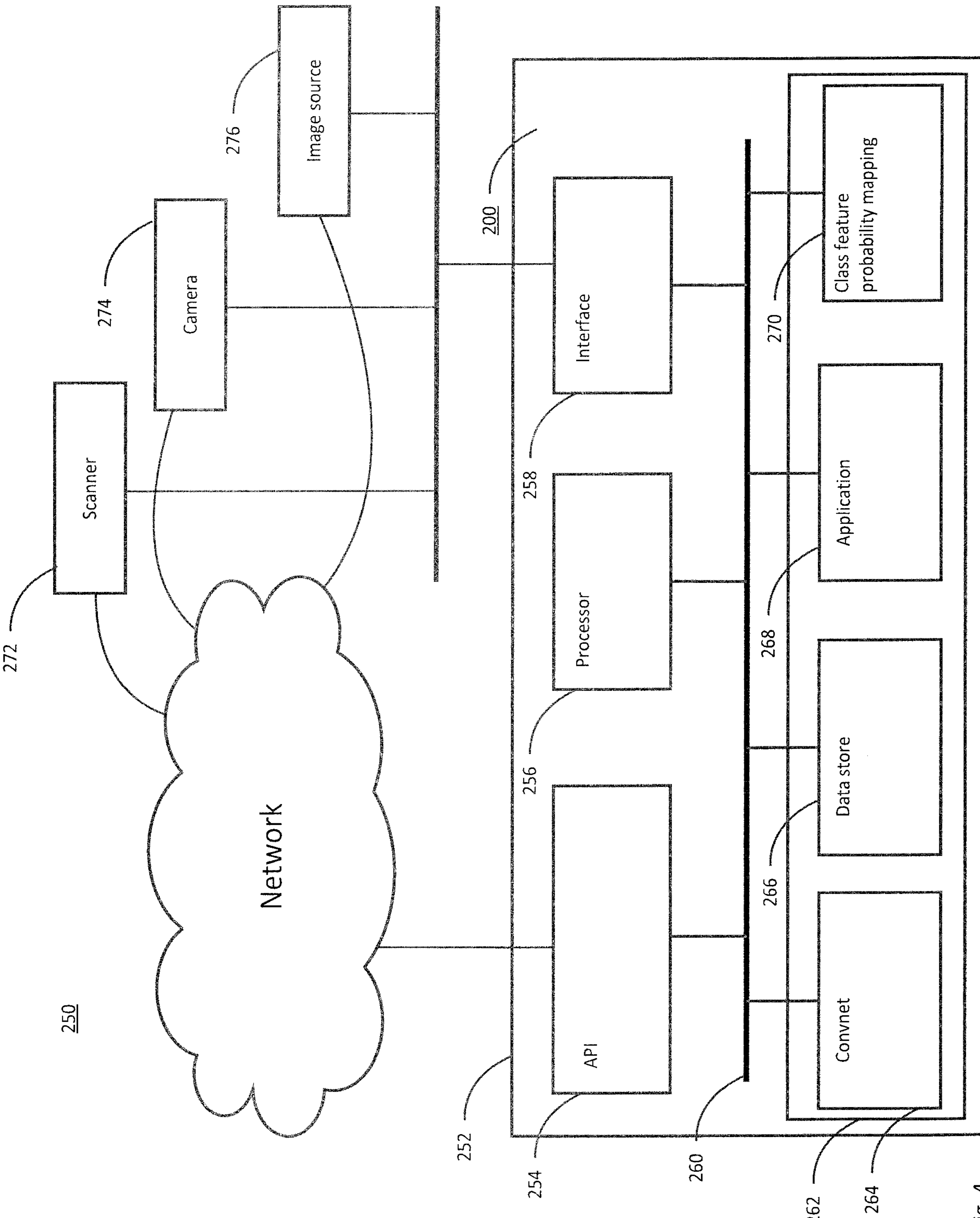

250
Network
272
Scanner
274
Camera
276
Image source
200
252
254
API
256
Processor
258
Interface
260
262
264
Convnet
266
Data store
268
Application
270
Class feature probability mapping

Fig. 4

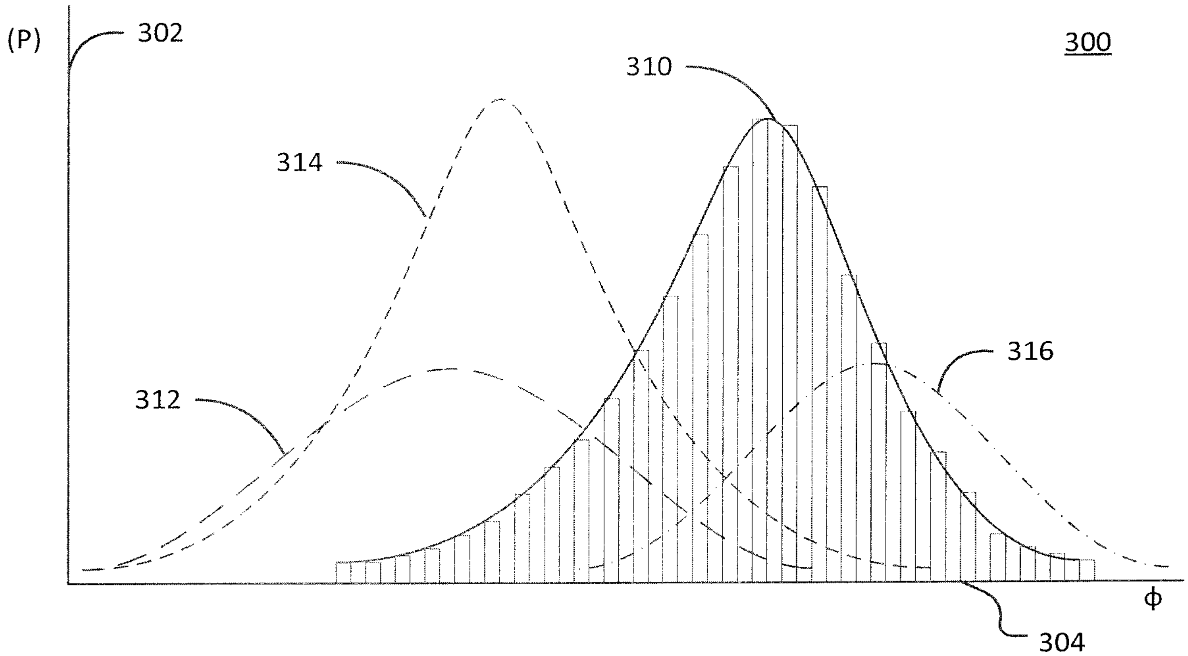



Fig. 5

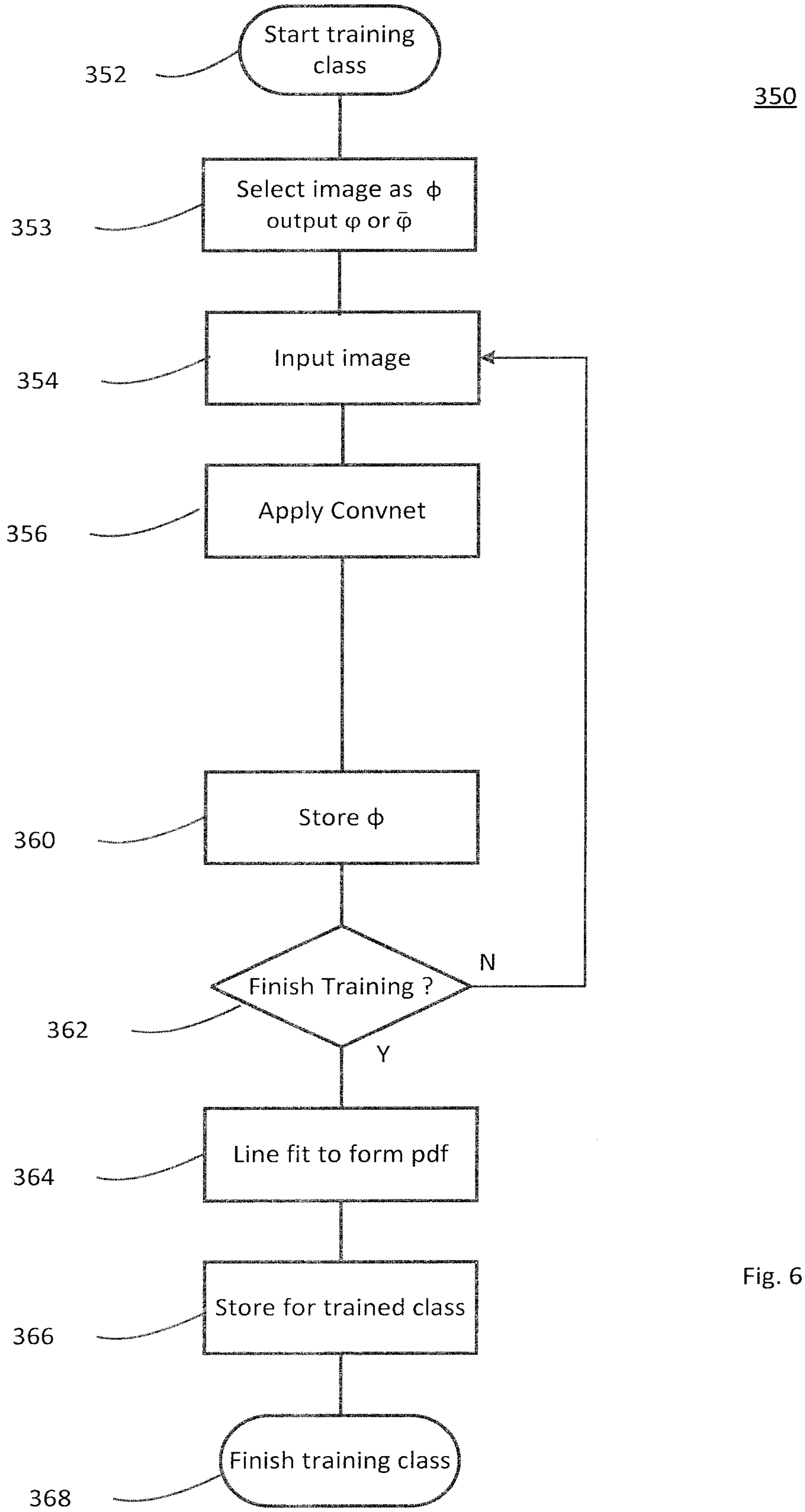
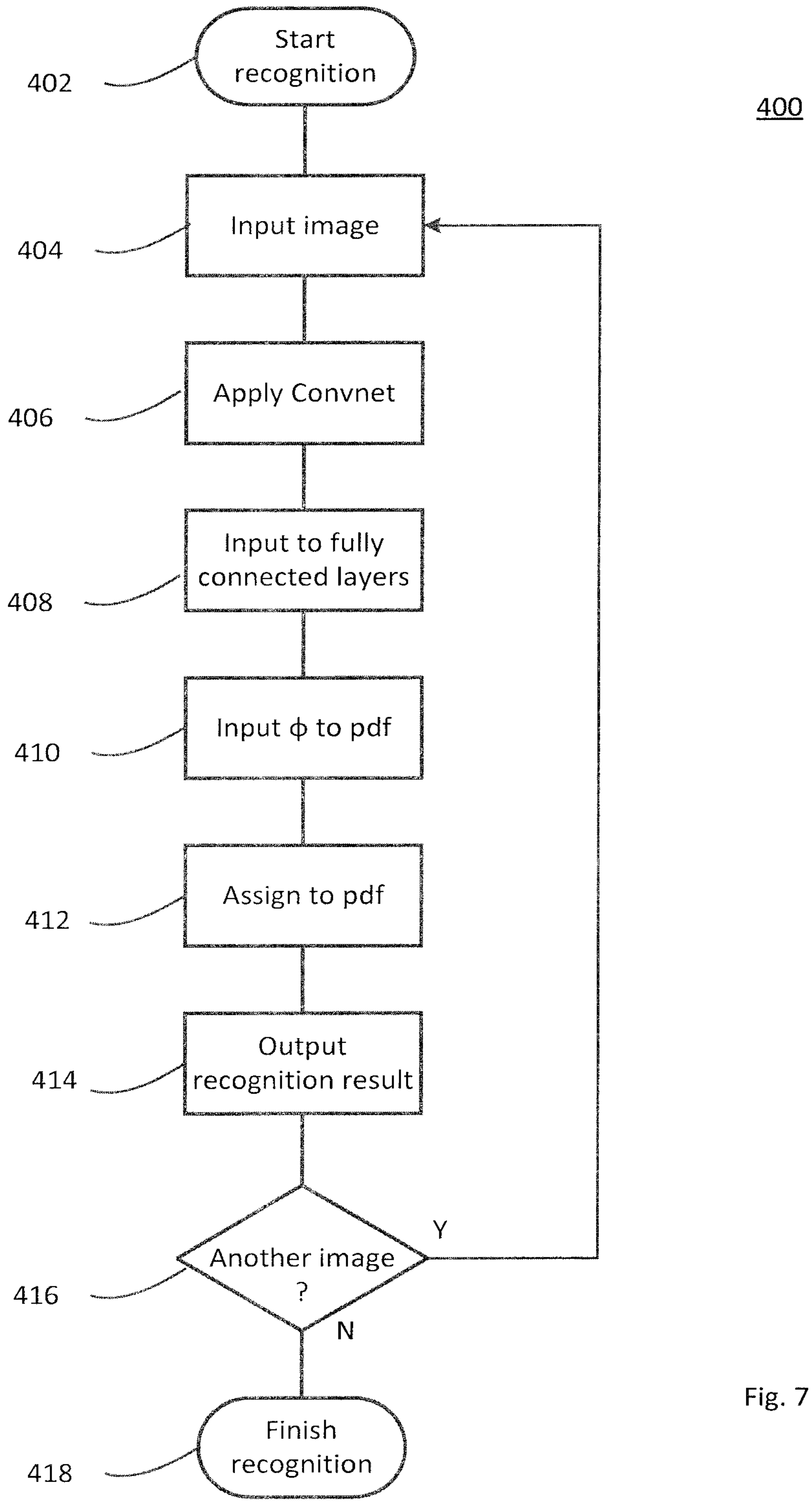

Start training class
352
350
Select image as ϕ output φ or φ̄
353
Input image
354
Apply Convnet
356
Store ϕ
360
Finish Training ?
N
Y
362
Line fit to form pdf
364
Store for trained class
366
Finish training class
368

Fig. 6



400
402
Start recognition
404
Input image
406
Apply Convnet
408
Input to fully connected layers
410
Input ϕ to pdf
412
Assign to pdf
414
Output recognition result
416
Another image ?
Y
N
418
Finish recognition


Fig. 7

# INTERNATIONAL SEARCH REPORT

International application No
PCT/GB2018/050971

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06K9/46 G06K9/62
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | BENDALE; ABHIJIT; TERRANCE BOULT: "Towards Open Set Deep Networks", ARXIV: 1511.06233 [CS], 19 November 2015 (2015-11-19), XP002782024, cited in the application the whole document ----- | 1-20 |

-/-

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 15 June 2018 | 26/06/2018 |

| Name and mailing address of the ISA/ | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2<br>NL - 2280 HV Rijswijk<br>Tel. (+31-70) 340-2040,<br>Fax: (+31-70) 340-3016 | Grigorescu, Cosmin |

2

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | Mostafa Ghazi ET AL: "Open-set Plant Identification Using an Ensemble of Deep Convolutional Neural Networks", CLEF Conference, 1 September 2016 (2016-09-01), pages 1-8, XP055484560, DOI: 10.1186/S13640-018-0261-2 Retrieved from the Internet: URL:https://www.researchgate.net/publication/309210779_0pen-set_Plant_Identification_Using_an_Ensemble_of_Deep_Convolutional_Neural_Networks [retrieved on 2018-06-14] the whole document ----- | 1-20 |
| A | US 2006/093208 AI (LI FAYIN [US] ET AL) 4 May 2006 (2006-05-04) the whole document ----- | 1-20 |

2

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/GB2018/050971

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| US 2006093208 A1 | 04-05-2006 | NONE | |

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 April 2018 (12.04.2018)

W!PO | PCT

(10) International Publication Number
WO 2018/065549 Al

(51) International Patent Classification:
*H04N 21/431* (201 1.01) *H04N 21/6587* (201 1.01)
*H04N 21/234* (201 1.01) *H04N 21/81* (201 1.01)
*H04N 21/4402* (201 1.01) *H04N 21/45* (201 1.01)

(21) International Application Number:
PCT/EP20 17/0754 16

(22) International Filing Date:
05 October 2017 (05.10.2017)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/404,554 05 October 2016 (05.10.2016) US

(71) Applicant: **BLIPPAR.COM LIMITED** [GB/GB]; 5th Floor West, 1 London Bridge, London SE1 9BG (GB).

(72) Inventors: **GRAHAM, Andrew;** c/o Blippar.com Limited, 5th Floor West, 1 London Bridge, London, Greater London SE1 9BG (GB). **QULI, Ariff;** c/o Blippar.com Limited, 5th Floor West, 1 London Bridge, London, Greater London SE1 9BG (GB). **SOLTANI, Dana;** c/o Blippar.com Limited, 5th Floor West, 1 London Bridge, London, Greater London SE1 9BG (GB). **JOSHI, Vikrum;** c/o Blippar.com Limited, 5th Floor West, 1 London Bridge, London, Greater London SE1 9BG (GB). **HILLER, Patrick;** c/o Blippar.com Limited, 5th Floor West, 1 London Bridge, London, Greater London SE 1 9BG (GB). **HARRIS, Michael;** c/o Blippar.com Limited, 5th Floor West, 1 London Bridge, London, Greater London SE1 9BG (GB). **MORALES, Brian;** c/o Blippar.com Limited, 5th Floor West, 1 London Bridge, London, Greater London SE1 9BG (GB). **SCHMALTZ, Ryan;** c/o Blippar.com Limited, 5th Floor West, 1 London Bridge, London, Greater London SE1 9GB (GB).

(54) **Title:** APPARATUS, DEVICE, SYSTEM AND METHOD

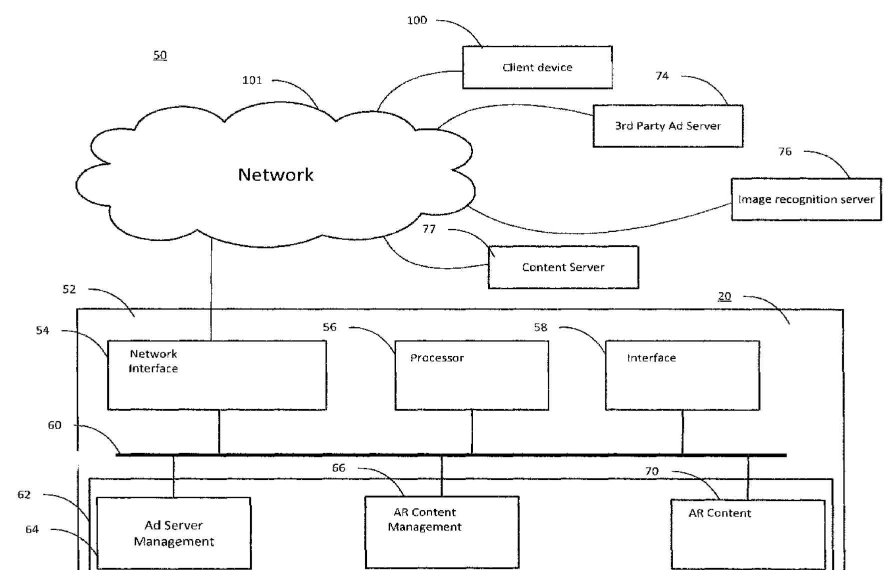



Fig. 1

(57) **Abstract:** A server apparatus (52), comprising a processor resource (56); and a memory comprising program elements operative to configure said processor to control. The server apparatus (52) to receive client metadata from a client device (100); select a content subject for delivery to a client device (100) in dependence on said received client metadata; transmit AR player ( 118a) program elements to be loaded onto a client device ( 100); transmit AR content (70) corresponding to said content subject to a client device ( 100) responsive to determining access permitted to a camera feed (122) of that client device (100); receive an image from a camera feed (122) of a client device (100) and input said image to an image recognition service (76) to determine a presence of a marker image in said image; responsive to determining presence of said marker transmit an AR initiate instruction to AR player program ( 118a) elements loaded onto a client device (100) to initiate rendering (348) of said AR content (70) over said marker in a camera feed (122).

WO 2018/065549 A1

[Continued on nextpage]

**(74) Agent: POTTER, Julian;** 138 Fetter Lane, London EC4A 1BT (GB).

**(81) Designated States** *(unless otherwise indicated, for every kind of national protection available):* AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

**(84) Designated States** *(unless otherwise indicated, for every kind of regional protection available):* ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

**Published:**

— *with international search report (Art. 21(3))*

# Apparatus, device, system and method

## Field

The present invention relates to an apparatus, device, system and method for providing specific data to a client device dependent on sensed input to the device. In particular, but not exclusively, embodiments of the present invention relate to delivering augmented reality content to a client device dependent on an image in a camera feed of the client device.

## Background

Conventionally, before an advertisement (hereinafter also referred to as an "ad", "advert" or "AD") is displayed on a webpage or through an "application" (also termed hereinafter "app") running on ("inside an application") a mobile device, the webpage or mobile device app sends a request to an ad server to serve an ad to it. Serving an ad in this sense means sending an actual picture or video making up the ad, i.e. what the viewer sees as the ad and also termed the "creative", back to the web page or mobile app before the web page fully loads on the device or into the application. In order to decide which creative must be sent to a webpage or mobile app, the webpage or mobile app supplies a number of data points to the ad server. These might include any one or more of the following non-exhaustive list:

-URL of the web page or the name of the mobile app;
-A subset of user cookies;
-Mobile device id (IDFA or Android Advertising ID);
-Size of the container for the ad;
-Duration of the ad (for video ads); and
-Location of the user.

IDFA is a so-called Advertising Identifier which is a unique identity for each device running an iOS operating system. Such information may be referred to as client metadata as it concerns information relating to a client device and its use.

Based on one or more of the data points, the ad server identifies an advertiser from a pool of advertisers who are interested to show an ad to the device corresponding to the data. In an automated programmatic model this is done based on some form of a second price auction. Once the ad server decides which advertiser gets the ad opportunity, a corresponding creative is sent to the web page or mobile app. The web page or mobile app displays the creative to the user thereby effecting service of the ad.

Many web pages or mobile apps accept a HTML5/JavaScript based creative instead of a simple static image or video file. This is the basis of the interactive ads. However, in the standard interactive ads, the HTML5/JavaScript creative is a self-contained unit that can render itself inside the container on the page or inside the mobile app. There is no additional input needed.

Summary

Viewed from a first aspect, the present invention provides a server apparatus, comprising a processor resource and a memory comprising program elements operative to configure said processor to control said server apparatus to:

receive client metadata from a client device;

select a content subject for delivery to a client device in dependence on said received client metadata;

transmit AR player program elements to be loaded onto a client device;

transmit AR content corresponding to said content subject to a client device responsive to determining access permitted to a camera feed of that client device;

receive an image from a camera feed of a client device and input said image to an image recognition service to determine a presence of a marker image in said image;

responsive to determining presence of said marker transmit an AR initiate instruction to AR player program elements loaded onto a client device to initiate rendering of said AR content over said marker in a camera feed.

Viewed from a second aspect the present invention provides a method of operating a server, comprising:

receiving client metadata from a client device;

selecting a content subject for delivery to a client device in dependence on said received client metadata;

transmitting AR player program elements to be loaded onto a client device;

transmitting AR content corresponding to said content subject to a client device responsive to determining access permitted to a camera feed of that client device;

receiving an image from a camera feed of a client device and inputting said image to an image recognition service to determine a presence of a marker image in said image;

transmitting an AR initiate instruction to AR player program elements loaded onto a client device to initiate rendering of said AR content over said marker in a camera feed responsive to determining presence of said marker.

Embodiments in accordance with first and or second aspects of the present invention provides for the delivery and execution of augmented reality (AR) advertisements to enhance the interactive use of a client device without the client device having to be pre-configured for augmented reality advertisement delivery. Typically, the marker is an image of something associated with the content subject and or AR content, the AR content being appropriate for rendering, in particular 3-D-rendering, over the marker in the camera feed.

Typically, said image is transmitted over a communications network to an image recognition server apparatus to implement said image recognition to determine said presence of said marker in said image and receive a message indicative of said marker present in said image. Transmitting the image to an image recognition server avoids the need to implement image recognition functionality on a server functioning as an augmented reality advertisement server. In this way, image recognition may be centralised and in particular a specialist image recognition service may be utilised. Optionally, image recognition functionality may be implemented on server apparatus itself.

Suitably, plural images from a camera feed are received by the server apparatus and determination of the presence of a marker is carried out for each of the more than one image. In this way, a greater certainty as to the identification of a marker or not may be achieved. In such an arrangement, plural images are transmitted over a communications network to an image recognition server apparatus to implement said image recognition to determine said presence of said marker in said plural images and receive a message indicative of said marker present in said plural images.

In a particularly suitable embodiment, the server apparatus may be configured to receive a message from AR player program elements loaded onto a client device indicative of access permitted to a camera feed. Thus, if any reason the server has been unable to determine whether or not access to the camera feed of the client device is permitted, for example by utilising a lookup table or some other reference source, the AR player program itself may determine whether or not access to the camera feed is permitted and relay the result to the server. Thus, it is possible to implement an embodiment in accordance with the present invention on a client device about which no advance information is known.

Conveniently, non-AR content corresponding to said content subject may be transmitted to a client device for display via a display controller on a screen of a client device in the event that access to the camera feed is not permitted. Thus, an ad may be served to a client device even though an AR operation is not possible.

Suitably, the AR content that is sent to the device may be determined by detection of a specific marker from said camera feed of the said device. The said image recognition server might be capable of identifying a very large number of said markers or real-world objects and the AR advertisement may be dynamically adjusted to what is being viewed through said device camera feed.

The server apparatus may be distributed over a communications network which utilises available server resource and also provides an element of redundancy in the event that there is a failure of any one or more physical devices.

Typically, said AR player program elements, image recognition and AR content are hosted on separate servers in order to take advantage of specific configurations and architecture is suitable for respective functions. For example, certain functions may require a large amount of memory or the ability to download large amounts of data and therefore require large bandwidth.

Conveniently, said AR player program elements and AR content may be comprised in a single or multiple files.

Generally the server apparatus is configured to receive data indicative of interaction with said AR content, said interaction through said client device. In this way, appropriate user metrics may be received processing the success or not of a particular advertising campaign for example. Typically, the server apparatus will receive one or more tracking pixels.

Viewed from a third aspect, the present invention provides a client device, comprising:

a processor resource;

a camera; and

a memory comprising program elements operative to configure said processor to control said device to:

transmit an image from a camera feed of said camera to an image recognition service;

receive notification of a presence of a marker image in said image;

receive AR content from a server; and

responsive to said notification of a presence of said marker render said AR content in said camera feed superposed on said image marker.

Viewed from a fourth aspect the present invention provides a method for operating a client device, comprising:

transmitting an image from a camera feed of a camera of a client device to an image recognition service;

receiving notification of a presence of a marker image in said image;

receiving AR content from a server; and

rendering said AR content in said camera feed superposed on said image marker responsive to said notification of a presence of said marker image in said image.

Embodiments in accordance with the third and or fourth aspects of the present invention, provide a client device environment in which augmented reality content such as augmented reality advertisements may be executed without any native or preconfigured software been loaded onto the client device. For example, any application running on a mobile client device may call an augmented reality ad server to implement the third and/or fourth aspect of the present invention.

Suitably, the client device determines if access is permitted to said camera feed and initiates a request to receive said AR content if access is permitted to said camera feed. In this way, AR content is only downloaded if it can be utilised. This saves on the use of bandwidth and also does not incur processing on the client device for something which may not be capable of being utilised.

Generally, the client device transmits device metadata indicative of the AR content handling capability of said device. In this way, the client device may identify its functions, utilities and permissions utilising generally widely recognised data.

Typically, first, second, third and/or fourth aspects of the present invention may be implemented by way of a computer program, such a computer program is typically provided on a carrier medium. Carrier media may be a different type but can include at least one of the following: a radio frequency signal, an optical signal, an electronic signal, a magnetic disc or tape, solid state memory, an optical disc, a magneto-optical disc, a computer disc, digital versatile disc and a blue-ray disc.

List of figures

Figure 1 illustrates a block diagram of an affinity server and data entities in accordance with an embodiment of the invention;
Figure 2 illustrates a block diagram of a communications device and network elements in accordance with an embodiment of the invention; and
Figure 3 is a process control flow diagram for a system in accordance with the present invention.

Description

Viewed in general outline, when an ad server decides to send an augmented reality (AR) creative in accordance with an embodiment of the present invention to a web page or application running on a mobile device, the content that is finally displayed to an ad viewer is not fully determined at the time the decision as to which creative to send is made. The interaction of a user of a device on which the ad is viewed with the ad will depend on an extra sensory input: in the described embodiment - what is in the camera feed from the device camera lens.

First, the AR creative is sent to the user device and is configured to establish if it has access to the camera feed from the device camera. Not every browser on every Operating System (OS) allows access to the camera feed on a device. Also, not every mobile app has access to the camera feed. For example, a weather app probably does not have access to the mobile device camera feed but the photograph app does.

If the AR creative determines it has been loaded into a container on a web page or a mobile app which has access to the device camera feed, it attempts to capture the camera feed. A "container" may be considered to be an application program or subsystem in which the program building block typically known as a component is run, such as an iframe in an HTML document. If explicit user permission is needed to capture the camera feed, the AR creative displays on the display screen of the device a so-called "call to action". The "call to action" could be an animation, video or a static image asking or informing a user to click or tap it to launch the AR ad. When the camera feed is captured, an image from it is sent to an image recognition service (configured as an Application program Interface (API)) every couple of seconds or so. The image recognition service identifies real world objects in an image supplied through the camera feed, for example, a logo, a face, a flower, a building, or a car of a specific make and model. The ad experience is scripted to begin when a specific marker is identified in the camera feed such as a brand logo or a human face. Once the marker is identified in the camera feed, the AR content of the creative is downloaded and rendered over it creating an augmented reality experience. When network bandwidth is or may be limited, AR content might be downloaded prior to recognition of the marker.

This concept can be applied not only to existing mobile devices such as smartphones and tablets but also wearables that get a constant camera feed (e.g. so called smart glasses). The experience could be something like this: a user walks into a grocery store and when they look around through their smart glass, they can see interactive ads placed on the product boxes that are on the shelves. Or when they walk into a shopping mall, they see interactive ads on the storefronts that advertise the latest fashion/clothing they have and invite the user to come in. Or on the user's smartphone, instead of seeing simple interactive ads about a new watch, the user can see how it looks on their wrist when you point their smartphone camera toward their hand.

Figure 1 illustrates a network architecture 50 comprising an ad server 52 which may communicate with image recognition server 76 for recognition of objects in an image provided to it and a client device 100, such as a computing device or mobile device such as a smartphone, via a network 101 . Other 3rd party ad servers, 74, may also be able to communicate with ad server 52 to provide support services or operate independently thereof. Optionally or additionally the network comprises content server 77 which provides augmented reality content and is accessible over a network 101 by advertisement servers 74 and 52, and client device 100. The ad server 52 is a server dedicated to managing the delivery of AR ads to client devices 100. The ad server 52 includes an ad server computing system 20 comprising a server network interface which communicates with clients over HTTP or HTTPS protocol 54; a server processor resource 56; and an interface module 58 for communication with an Ad operations manager or technician operating the ad server 52. The ad server computing system 20 also includes memory 62 for storing program elements and data to implement various modules for storing data and content.

Memory 62 includes ad server management module 64 for managing serving requests for ads, determining which ad to transmit to a client and generally controlling the operation of server 52. The ad server computing system 20 also comprises an AR content management module 66, comprising program elements for configuring processing resource 56 to manage requests and data received by the ad server 52 from a client device 100 and also transmit data and content to a requesting client device 100. AR content management module 66 also includes program instructions operative when executed to provide overall management of the application and call respective modules as and when required, and in particular to deliver the right version of AR player and content when requested by client device 100 or client interface. Additionally, such instructions may also be input and output through the user interface 58. AR content is stored in module 70 of memory store 62. The various modules and resources of ad server computing system 20 communicate over bus 60.

Memory store 62 also comprises general memory, not shown, in which various data and information may be stored. For example, in accordance with an embodiment of the present invention a list of client device identities of client devices that allow access to their camera feed is stored. Inspection of such a list allows early determination of whether or not an augmented reality content element may be utilised by a client device 100.

The servers 52, 74, 76 and 77 are merely illustrative and other servers and network coupled devices may also communicate with the device 100 and server 52.

References to a network and communication made above should be construed broadly to include many types of network and networks of networks, including the public service telephone network where appropriate, cable systems, wireless networks and the Internet.

Figure 2 illustrates a block diagram of communications device 100, according to one embodiment. The communications device 100 is connectable via network connection 113 over a communications network 101 to a number of servers 52, 74 and 76.

The communications device 100 includes a computer memory 103, a processor 104, an RF (Radio Frequency) interface 105, a screen 106, a keyboard 107, a microphone 108 a speaker 109, a Bluetooth interface 110, an infrared interface 111, connectors 112, a network interface 113, a digital camera 122 and a data bus 125. As will be evident to the person of ordinary skill in the art not all these devices are necessary for an implementation of the present invention but such things are typically included in communications devices.

A number of application programs 114 are stored in computer memory 103. Respective contacts databases 115 of the communication application programs 114 store details of the contacts that are associated with the respective communications application programs. For example, a communications application program 114 corresponding to an email mode of communication would include a contact database consisting of email addresses and contact identifiers.

The ad container 123 resides in one or more of the application programs 114. The ad container communicates with the Ad server 74 through a communications interface 120 which is provided by the application program 114 or operating system of the client device 100. Ad container 123 initiates an Ad request to the Ad server 74 for the download of an AR creative unit 118 which includes AR Player 118a. Optionally, the Ad request may be redirected to AR content distribution server 52. AR Player 118a is configured to communicate with camera feed 121 and provide images from camera feed 121 to image recognition server 76. AR player 118 also includes a 3D rendering engine which can render 3D AR content 119 on camera feed 121 when AR player 118a downloads 3D AR content from AR content distribution server 52 to ad container 123. AR creative unit 118 comprises computer program elements, such as instructions, executable by processor resource 104 to configure it to perform the various functions and activities inside ad container 123 such as the capturing camera feed 12, 3D rendering 118b and communicating with Ad server 74, AR content distribution server 52 and image recognition server 74 through a communications interface 120 provided by the client device operation system.

Also in memory 103 is a key handler 116 comprising a software module configured to respond to keyboard 110 generated events. The key handler 116 provides appropriate inputs to the applications 114 and to a control module 117 that provides overall control of the processes leading to initiation of one or more applications 114 and ad containers 123.

Turning now to figure 3 there is illustrated a process flow control diagram for a server-client network arrangement in accordance with an embodiment of the present invention. In the described embodiment, the AR content relates to an interactive advertisement ("ad"). As is usual, a client device 100 initiates a request, step 306, for an advertisement from within an application or web browser in accordance with the way in which the application is being executed or web page is being displayed by the web browser. The request is communicated through network interface 113 and over communication network 101 . Client metadata usually accompanies such a request. On the server side, 52, the request is received through a network interface 54 and the ad server management module 64 is invoked to review the client metadata to determine if the container is able to access the camera feed for a camera located on the device, step 308. In this regard, ad server management module 64 configures processor 56 to inspect the list of device identities known to provide access to the camera feed and stored in memory 62. If the decision, step 310, is "No" then process control flows to step 312 at which point the ad server management module 64 configures processor 56 to access a store of non-augmented reality advertisements and select one in accordance with the client metadata and/or other criteria for transmission through network interface 54 to client device 100. Client device 100 receives the non-augmented reality advertising through network interface 113 and communication interface 120, and control module 117 configures processor 104 to display the non-augmented reality advertisement directly on the display screen 106, step 314.

If the decision, 310, is "Yes" or "Don't know" because the client identity is not in the stored list, process control flows to step 316 where ad server management module 64 is invoked which configures processor 56 to apply an advertisement selection procedure to determine which one of augmented reality (AR) advertisements stored AR content 70 is to be supplied to client device 100. As with the non-augmented reality advertisement, the client metadata, amongst other things, may be utilised in the selection procedure to determine which AR advertisement is selected.

In the described embodiment, The AR advertisement may be created by a scripting language such as Blippar Augmented Reality Developer API which is available through https://blippar.com/en/products/augmented-reality-api. The AR advertisement is optimized for serving as a rich media advertisement unit in accordance with the ad industry guidelines such as IAB HTML5 for Digital Advertising Guidelines. In the described embodiment, the AR advertisement is hosted on server 52, which may be termed an advertisement server in that it handles requests for advertisements. Optionally, the AR advertisement zip file may be retrieved from a third party content distribution server 77. Respective AR advertisements have a public URL which may be used to access it for downloading. Multiple versions of an AR advertisement are created and stored in order to support serving the advertisement to different environments (different browsers, different mobile or wearable devices, different operating systems or different publisher settings which might limit interactivity or access to the camera feed).

At step 318 ad server management module 64 configures processor resource 56 to identify an AR player compatible with the identified client device 100 from the AR content module 70 and transmit computer program elements for that AR player or its public address to client device 100 through the client interface 54 and over network 101 . In the described embodiment, AR player 118a is based on HTML5 and JavaScript and follows IAB HTML5 for Digital Advertising Guidelines.

At step 320 client device 100 receives the AR player program elements and stores them in the device as appropriate. Process control for client device 100 then flows to step 322 at which the AR player 118a determines whether or not access is granted to the camera feed 121 for the camera 122 of client device 100. If the decision at step 324 is that access is denied to camera feed 121 AR player 118 a configures processor resource 104 to initiate a request for a non-augmented reality advertisement through communication interface 120 and network interface 113 to server 52 over network 101 .

If the decision at step 324 is that access to the camera feed is permitted process control flows to step 328 where AR creative 118 configures processor resource 104 to transmit a request through communication interface 120, network interface 113 and over network 101 to server 52 for an AR advertisement. The request for the AR advertisement is received at server 52, step 330 and ad server management module 64 configures processor resource 56 to transmit the AR advertisement selected from the AR content 70 at step 316 through API 54 over network 101 to client device 100.

Within client device 100, at step 332, user permission is requested for access to the camera feed if the operating system or application/container for which the request for augmented reality advertisement has been made requires such user permission. At step 334, client device 100 receives the augmented reality advertisement content and stores it locally on the device 100. Process control for client device 100 then flows to step 336 at which AR player 118a configures processor resource 104 to capture images from camera feed 121 and transmit them through communication interface 120 and network interface 113 across network 101 to an image recognition service such as provided by image recognition server 76.

At step 338, server 76 receives images and is configured to determine the presence of a marker image in the images received from client device 100. If the decision at step 340 is that there is no marker image present, the image recognition service such as image recognition server 76 transmits a message to client device 100 over network 101 to notify AR player 118a of the decision. AR player 118a continues to capture consecutive images from camera feed 121 and transmit them through communication API 120 and network interface 113 across network 101 to image recognition server 76 until a marker image is positively identified in the image. A user has control over their interaction with the AR advertisment and can terminate that interaction at any time. Using and advertisement for a wristwatch as an example; if a user initiates an AR advertisement by tapping on a "call to action" image while pointing to or then points the device camera to the sky or other space empty of a suitable marker, nothing will happen. The AR player will take consecutive images of the sky and send it to the image recognition server but because the image of a wrist is not present in those images, nothing will happen. The user can close the advertisement or point to their wrist to start the render of the wristwatch on their wrist. Some advertisement containers in which advertisements are timed and the hosting application or web page closes them if the advertisement does not finish processing within a certain time threshold. Consequently, if the decision at step 340 is "No", process control flows back to step 336 the further images are sent from the camera feed to the server. However, if the user has sent an instruction to stop the advertisement, step 346 then step 342 provides a "Yes" decision and process control flows to step 354 where the process is stopped.

In one embodiment, there may be a store of marker images corresponding to the AR ads that can be delivered and which is inspected by the image recognition service. Optionally or additionally, the marker image may be something that is recognised as being something the AR player can render AR content over.

If the decision at step 340 is "Yes", then at step 344 the image recognition service, e.g. image recognition server 76 notifies AR player 118a over network 101 of the presence of a marker image in the images provided to the server 76. In response to receiving the notification of the presence of marker images at client device 100, AR player 118a configures processor resource 104 to invoke 3D rendering engine 118b which configures processor resource 104 to render AR advertisement content over the marker present in camera feed 121 . In this way, 3D contents in the AR advertisement can be placed within the image seen through the client device camera 122. For example, if the AR advertisement relates to a wristwatch and the market image is a human wrist, 3D rendering engine 118b may render a 3D image of the wristwatch over the image of the human wrist within the camera feed 121 and show various perspectives and views of the wristwatch as the human wrist is moved.

In the described embodiment, one or more marker images are defined for a particular AR advertisement and acts as a trigger to invoke 3D rendering of the AR content of an AR ad over that marker image in the camera feed. In order to assist a user of a client device in invoking AR content, a suitable message or dialogue box may be displayed instructing the user to view a particular item in order to capture an appropriate marker image in the image feed in order to initiate 3D rendering of the AR content.

At step 350 client device 100 monitors user activity during playing of the AR advertisement reports various tracking metrics up to server 52, which receives the tracking metrics report, step 352, and stores it or forward it to appropriate storage location or entity for further processing.

3D rendering engine 118b may be built based on any suitable 3D rendering program or software. In the described embodiment the 3D rendering engine is based on WebGL or web graphic library or other suitable 3D rendering engine.

The AR player instructs processor resource 104 to fire custom tracking pixels that track specific actions inside an AR advertisement, e.g. a specific button is tapped or user has rotated the phone to see a different view of the AR content. These custom tracking pixels are defined by the developer of an AR content in accordance with the present invention and have to be embedded in the AR advertisement configuration. The AR player 118a also configures processor resource 104 to fire third-party tracking pixels for delivery verification, demographic verification, and viewability. These tracking pixels are added to the AR advertisement configuration.

In the described embodiment, the AR player 118a can render an AR ad inside an HTML iframe or an MRIAD compatible mobile container.

Insofar as embodiments of the invention described above are implementable, at least in part, using a software-controlled programmable processing device such as a general purpose processor or special-purposes processor, digital signal processor, microprocessor, or other processing device, data processing apparatus or computer system it will be appreciated that a computer program for configuring a programmable device, apparatus or system to implement the foregoing described methods, apparatus and system is envisaged as an aspect of the present invention. The computer program may be embodied as any suitable type of code, such as source code, object code, compiled code, interpreted code, executable code, static code, dynamic code, and the like. The instructions may be implemented using any suitable high-level, low-level, object-oriented, visual, compiled and/or interpreted programming language, such as C, C++, Java, BASIC, Perl, Matlab, Pascal, Visual BASIC, JAVA, ActiveX, assembly language, machine code, and so forth. A skilled person would readily understand that term "computer" in its most general sense encompasses programmable devices such as referred to above, and data processing apparatus and computer systems in whatever format they may arise, for example, desktop personal computer, laptop personal computer, tablet, smartphone, smart glass, or other computing device.

Suitably, the computer program is stored on a carrier medium in machine readable form, for example the carrier medium may comprise memory, removable or non-removable media, erasable or non-erasable media, writeable or re-writeable media, digital or analog media, hard disk, floppy disk, Compact Disk Read Only Memory (CD-ROM), Compact Disk Recordable (CD-R), Compact Disk Rewriteable (CD-RW), optical disk, magnetic media, magneto-optical media, removable memory cards or disks, various types of Digital Versatile Disk (DVD) subscriber identity module, tape, cassette solid-state memory. The computer program may be supplied from a remote source embodied in the communications medium such as an electronic signal, radio frequency carrier wave or optical carrier waves. Such carrier media are also envisaged as aspects of the present invention.

The computer program may comprise more than instructions, for example parameters and of data for the computer program and as such a computer program may generally be referred to as comprising computer program elements in order to encompass such parameters, data and other things for the computer program.

As used herein any reference to "one embodiment" or "an embodiment" means that a particular element, feature, structure, or characteristic described in connection with the embodiment is included in at least one embodiment. The appearances of the phrase "in one embodiment" or the phrase "in an embodiment" in various places in the specification are not necessarily all referring to the same embodiment.

As used herein, the terms "comprises," "comprising," "includes," "including," "has," "having" or any other variation thereof, are intended to cover a non-exclusive inclusion. For example, a process, method, article, or apparatus that comprises a list of elements is not necessarily limited to only those elements but may include other elements not expressly listed or inherent to such process, method, article, or apparatus. Further, unless expressly stated to the contrary, "or" refers to an inclusive or and not to an exclusive or. For example, a condition A or B is satisfied by any one of the following: A is true (or present) and B is false (or not present), A is false (or not present) and B is true (or present), and both A and B are true (or present).

In addition, use of the "a" or "an" are employed to describe elements and components of the invention. This is done merely for convenience and to give a general sense of the invention. This description should be read to include one or at least one and the singular also includes the plural unless it is obvious that it is meant otherwise.

In view of the foregoing description it will be evident to a person skilled in the art that various modifications may be made within the scope of the invention. For example, although languages such as HTML5 and JavaScript have been referred to in the specific description such reference is by way of illustrative example only and other comparable languages may be used, including earlier or later versions of those explicitly mentioned if appropriate. Also, although formats such as JPEG and PNG have been referred to in the specific description such reference is by way of illustrative example only and other comparable formats may be used, including earlier or later versions of those explicitly mentioned if appropriate.

In the foregoing various references have been made to webpages, applications and mobile devices. Such terms have their normal meaning within the art but without limitation also include the following. A webpage, or web page, is document that is suitable for distribution over the World Wide Web and for display on web browsers. A web browser is software that displays webpages on a display screen of a computer or mobile device such as a smartphone. The webpage is the displayed aspect of a computer file typically written in a so-called "mark-up" language such as Hyper Text Markup Language (HTML) or similar language. Web pages are available from web servers accessible over a network by network coupled computer or mobile devices. Applications are computer programs generally designed to provide some form of user functionality such as a word processor function or a drawing function. When applied to mobile devices such as smart phones and tablets the term application may refer to such things as a word processor function but also refers to very specific functions for which a specific application may be invoked, for example a calculator function, a connection to a news website (a collection of webpages), a telephone function or text messaging function. A mobile device is any device which does not require a wired connection to operate, either in terms of power or in communications connectivity. Without limitation, a client device may be any device, mobile or not, that can communicate with a server apparatus over a communications channel, for example a network such as the Internet.

The term HTML5 used herein refers to the fifth version of the HTML standard available at https://html.spec.whatwq.org . JavaScript refers to a web browser presentation language the standard for which may be found here: http://www.ecmainternational.org/publications/standards/Ecma-262.htm .

The term "JPEG" used herein refers to a coding and compression standard of the Joint Photographic Experts Group and details may be found here: https://jpeg .org/jpeg/. The term "PNG" used herein refers to the PNG Specification which may be found here: https://www.w3.orq/TR/PNG/ .

Various examples of client metadata have been disclosed in the foregoing but embodiments in accordance with the invention are not limited to such examples or combinations thereof. Other information or data concerning a client device, its use and the applications on it may also be utilised and may also be considered metadata. Additionally, although client devices such as a mobile device or smartphone have been disclosed, client devices may also include wearable devices such as smart glasses, for example Google Glass®. Other display devices may also be configured as client devices or to display information provided by client devices.

In the embodiment described above, image recognition is conducted in an image recognition server, 76, separate from the ad server 52. However, functionality of image recognition server 76 may be included in ad server 52 or distributed over a plurality of servers depending upon the requirements of the network architecture or implementation. Additionally, although AR content has been described as originating on ad server 52 it may be provided by another ad server such as third-party ad server 74 or on a content distribution server configured for delivering AR content.

Although embodiments in accordance with the present invention have been described with reference to process control flow diagrams and with reference to "steps" embodiments in accordance with the present invention are not necessarily limited to the specific order in which the steps occur and may be performed in order other than specifically disclosed herein. Additionally, specific examples of 3D rendering engines and software have been disclosed but embodiments in accordance with the present invention are not limited to rendering engines such as are disclosed any suitable 3D rendering engine.
Specific languages will develop the tools mentioned in the present disclosure not limitations respect of embodiments of the present invention and other programming languages or development tools may be utilised.

The scope of the present disclosure includes any novel feature or combination of features disclosed therein either explicitly or implicitly or any generalisation thereof irrespective of whether or not it relates to the claimed invention or mitigate against any or all of the problems addressed by the present invention. The applicant hereby gives notice that new claims may be formulated to such features during prosecution of this application or of any such further application derived therefrom. In particular, with reference to the appended claims, features from dependent claims may be combined with those of the independent claims and features from respective independent claims may be combined in any appropriate manner and not merely in specific combinations enumerated in the claims.

Claims:

1. Server apparatus, comprising:

 a processor resource; and

 a memory comprising program elements operative to configure said processor to control said server apparatus to:

 receive client metadata from a client device;

 select a content subject for delivery to a client device in dependence on said received client metadata;

 transmit AR player program elements to be loaded onto a client device;

 transmit AR content corresponding to said content subject to a client device responsive to determining access permitted to a camera feed of that client device;

 receive an image from a camera feed of a client device and input said image to an image recognition service to determine a presence of a marker image in said image;

 responsive to determining presence of said marker transmit an AR initiate instruction to AR player program elements loaded onto a client device to initiate .rendering of said AR content over said marker in a camera feed.

2. Server apparatus according claim 1, said program elements further operative to transmit said image over a communications network to an image recognition server apparatus to implement said image recognition to determine said presence of said marker in said image and receive a message indicative of said marker present in said image.

3. Server apparatus according to claim 1, said program elements further operative receive plural images from a camera feed.

4. Server apparatus according claim 3, said program elements further operative to transmit said plural images over a communications network to an image recognition server apparatus to implement said image recognition to determine said presence of

said marker in said plural images and receive a message indicative of said marker present in said plural images.

5. Server apparatus according to any preceding claim, said program elements further operative to receive a message from AR player program elements loaded onto a client device indicative of access permitted to a camera feed.

6. Server apparatus according to any preceding claim, said program elements further operative to determine from said metadata access permitted to a camera feed.

7. Server apparatus according to claim 5 or claim 6, said program elements further operative to transmit non-AR content corresponding to said content subject to a client device for display via a display controller on a screen of a client device.

8. Server apparatus according to any preceding claim, wherein said server apparatus is distributed over a communications network.

9. Server apparatus according to claim 8, wherein said AR player program elements, image recognition and AR content are hosted on separate servers.

10. Server apparatus according to any of claims 1 to 8, wherein said AR player program elements and AR content are comprised in a single file.

11. Server apparatus according to any preceding claim, said program elements further configured to receive data indicative of interaction with said AR content, said interaction through said client device.

12. Server apparatus according to claim 12, said program elements further configured to receive one or more tracking pixel fires.

13. A client device, comprising:
a processor resource;

a camera; and

a memory comprising program elements operative to configure said processor to control said device to:

transmit an image from a camera feed of said camera to an image recognition service;

receive notification of a presence of a marker image in said image;

receive AR content from a server; and

responsive to said notification of a presence of said marker image in said image render said AR content in said camera feed superposed on said image marker.

14. A device according to claim 13, said program elements further operative to determine if access permitted to said camera feed and initiating a request to receive said AR content if access is permitted to said camera feed.

15. A device according to claim 13 or claim 14, said program elements further operative to transmit device metadata indicative of the AR content handling capability of said device.

16. A method of operating a server, comprising:

receiving client metadata from a client device;

selecting a content subject for delivery to a client device in dependence on said received client metadata;

transmitting AR player program elements to be loaded onto a client device;

transmitting AR content corresponding to said content subject to a client device responsive to determining access permitted to a camera feed of that client device;

receiving an image from a camera feed of a client device and inputting said image to an image recognition service to determine a presence of a marker image in said image;

transmitting an AR initiate instruction to AR player program elements loaded onto a client device to initiate rendering of said AR content over said marker in a camera feed responsive to determining presence of said marker.

17. A method according claim 16, further comprising transmitting said image over a communications network to an image recognition server apparatus to implement said image recognition to determine said presence of said marker in said image and receiving a message indicative of said marker present in said image.

18. A method according to claim 16, further comprising receiving plural images from a camera feed.

19. A method according claim 18, further comprising transmitting said plural images over a communications network to an image recognition server apparatus to implement said image recognition to determine said presence of said marker in said plural images and receiving a message indicative of said marker present in said plural images.

20. A method according to any of claim 16 to claim 19, further comprising receiving data indicative of interaction with said AR content, said interaction through said client device.

21. A method according to any of claim 16 to claim 20, further comprising receiving a message from AR player program elements loaded onto a client device indicative of access permitted to a camera feed.

22. A method according to any claim 16 to claim 21, further determining from said metadata access permitted to a camera feed.

23. A method according to claim 21 or claim 22, further comprising transmitting non-AR content corresponding to said content subject to a client device for display via a display controller on a screen of a client device.

24. A method according to any of claim 16 to claim 23, wherein said server apparatus is distributed over a communications network.

25. A method according to claim 24, wherein said AR player program elements, image recognition and AR content are hosted on separate servers.

26. A method according to any of claim 16 to claim 24, wherein said AR player program elements and AR content are comprised in a single file.

27. A method according to any of claim 16 to 26, said program elements further configured to receive one or more tracking pixel fires.

28. A method for operating a client device, comprising:

transmitting an image from a camera feed of a camera of a client device to an image recognition service;

receiving notification of a presence of a marker image in said image;

receiving AR content from a server; and

rendering said AR content in said camera feed superposed on said image marker responsive to said notification of a presence of said marker image in said image.

29. A method according to claim 28, further determining if access permitted to said camera feed and initiating a request to receive said AR content if access is permitted to said camera feed.

30. A method according to claim 28 or claim 29, further comprising transmitting device metadata indicative of the AR content handling capability of said device.

31. A computer program, comprising computer program elements operative in a computer processor to implement a method in accordance with any of claim 16 to claim 30.

32. A computer program comprising computer program elements operative in a computer processor to configure server apparatus in accordance with any of claim 1 to claim 12 or in accordance with any of claim 13 to claim 15.

33. A carrier medium carrying a computer program according to claim 31 or claim 32.

34. A carrier medium according to claim 33 comprising at least one of the following: a radio frequency signal, an optical signal, an electronic signal, a magnetic disc or tape, solid state memory, an optical disc, a magneto-optical disc, a computer disc, digital versatile disc and a blue-ray disc.

35. A computer program product comprising a carrier medium according to claim 33 or claim 34 configured as a computer readable medium.

36. Server apparatus substantially as hereinbefore described and with reference to the accompanying drawings.

37. Client device substantially as hereinbefore described and with reference to the accompanying drawings.

38. A method substantially as hereinbefore described and with reference to the accompanying drawings.

1/3

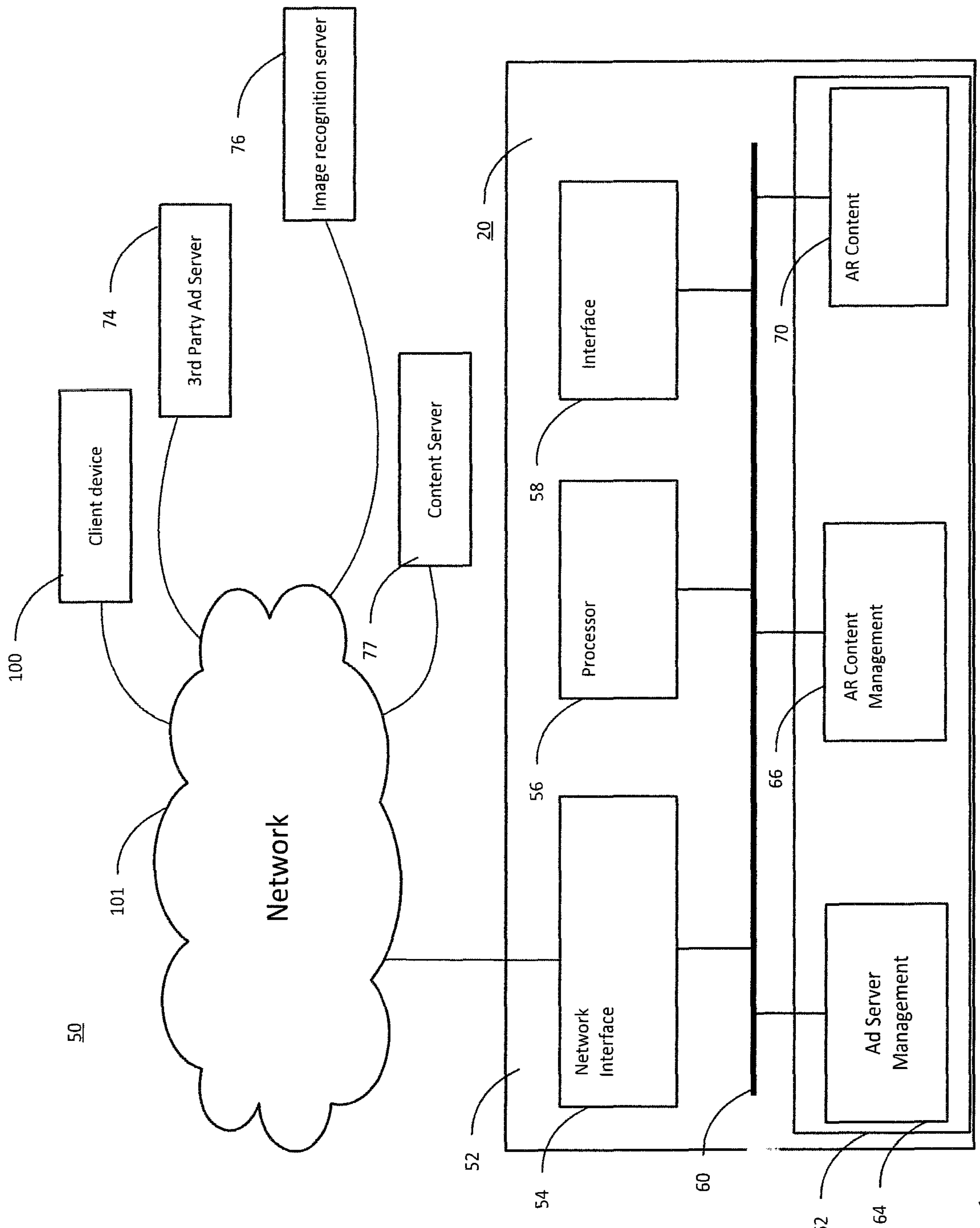

50
Network
101
100
Client device
74
3rd Party Ad Server
76
Image recognition server
77
Content Server
20
52
Network Interface
54
Processor
56
Interface
58
60
Ad Server Management
62
64
AR Content Management
66
AR Content
70

Fig. 1

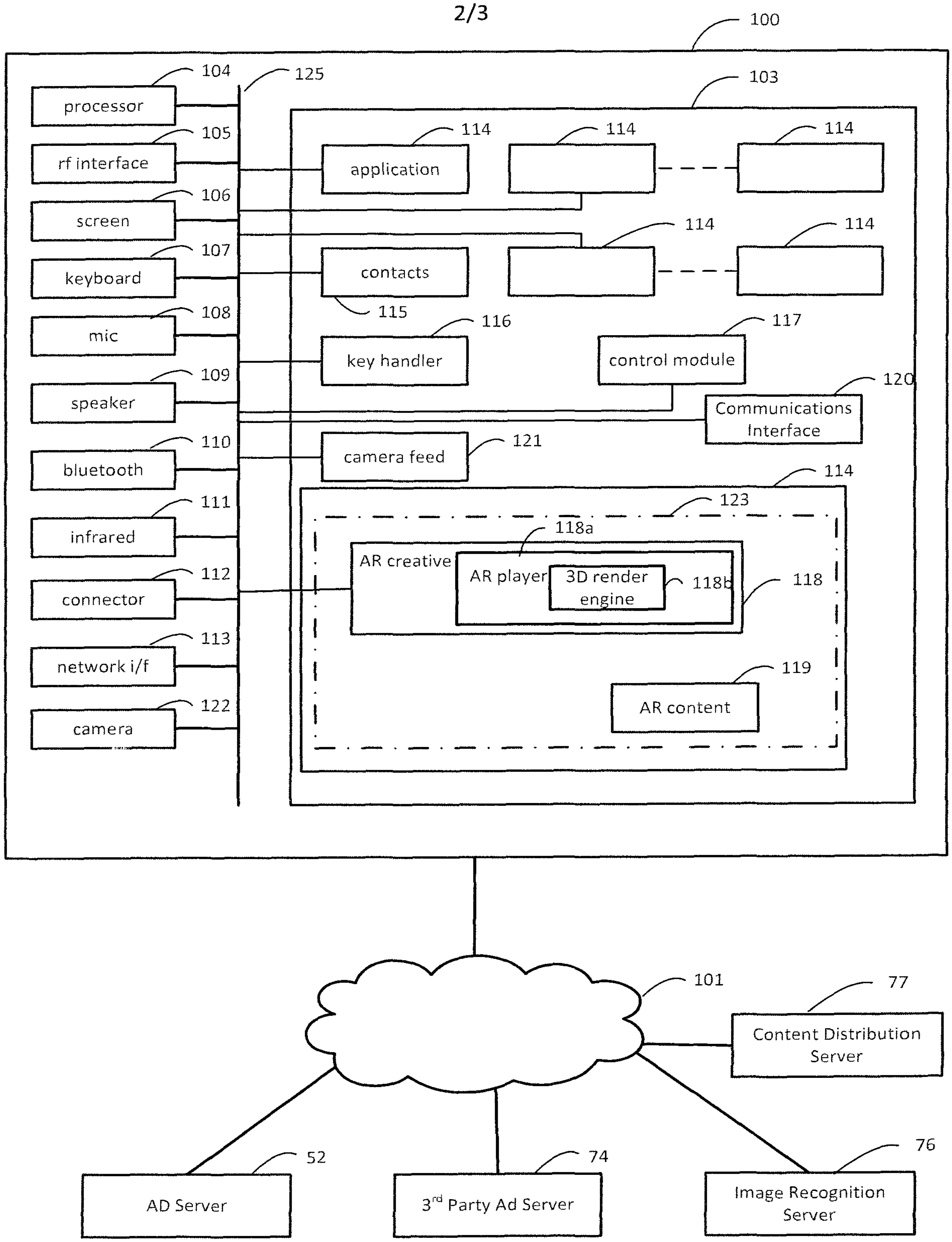
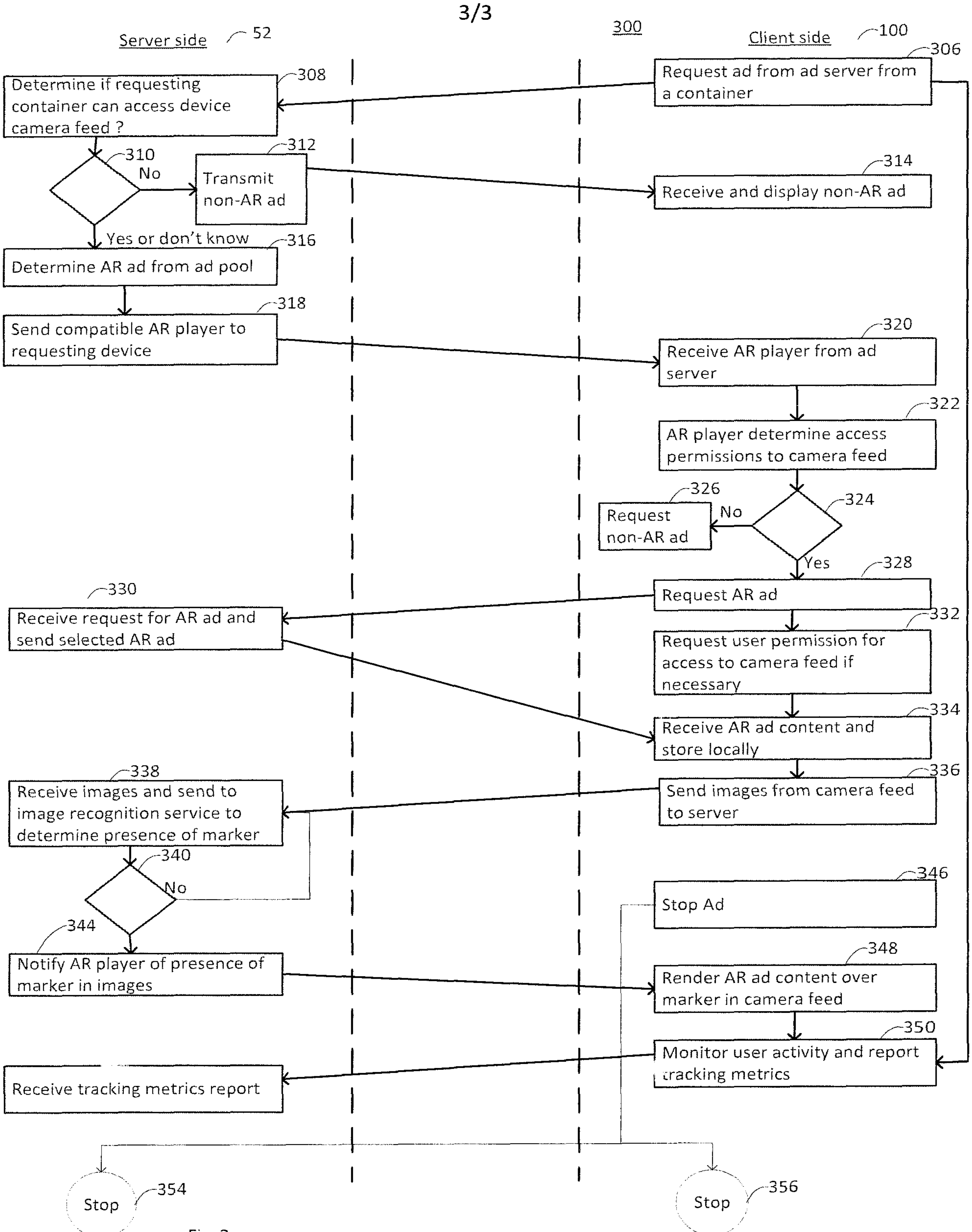

100
104
processor
125
103
105
rf interface
114
application
114
114
106
screen
107
keyboard
114
114
contacts
115
108
mic
116
key handler
117
control module
109
speaker
120
Communications Interface
110
bluetooth
camera feed
121
114
111
infrared
123
118a
AR creative
AR player
3D render engine
118b
118
112
connector
113
network i/f
119
AR content
122
camera
101
77
Content Distribution Server
76
52
74
AD Server
3rd Party Ad Server
Image Recognition Server

Fig. 2



300
Server side
52
Client side
100
Request ad from ad server from a container
306
Determine if requesting container can access device camera feed ?
308
310
No
Transmit non-AR ad
312
Receive and display non-AR ad
314
Yes or don't know
Determine AR ad from ad pool
316
Send compatible AR player to requesting device
318
Receive AR player from ad server
320
AR player determine access permissions to camera feed
322
Request non-AR ad
326
No
324
Yes
Request AR ad
328
Receive request for AR ad and send selected AR ad
330
Request user permission for access to camera feed if necessary
332
Receive AR ad content and store locally
334
Send images from camera feed to server
336
Receive images and send to image recognition service to determine presence of marker
338
340
No
Stop Ad
346
344
Notify AR player of presence of marker in images
Render AR ad content over marker in camera feed
348
Monitor user activity and report tracking metrics
350
Receive tracking metrics report
Stop
354
Stop
356


Fig. 3

# INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2017/075416

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N21/431 H04N21/234 H04N21/4402 H04N21/6587 H04N21/81 H04N21/45
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
H04N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EPO-Internal

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | US 2012/116728 A1 (SHEAR ROBERT MATTHEW [US] ET AL) 10 May 2012 (2012-05-10) figures 2-4 paragraphs [0007], [0033], [0065] - [0081], [0087], [0119] ----- | 1-38 |
| X | US 2014/340423 A1 (TAYLOR DAVID A [US] ET AL) 20 November 2014 (2014-11-20) paragraphs [0005], [0006], [0009], [0043], [0057], [0062], [0072] ----- | 1-38 |
| A | US 2014/171116 A1 (LAMARCA ANTHONY G [US] ET AL) 19 June 2014 (2014-06-19) figure 3 ----- | 1-38 |
| A | US 2014/160161 A1 (BARREIRO PATRICIO [AR] ET AL) 12 June 2014 (2014-06-12) figures 4-9 ----- -/-- | 1-38 |

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 15 December 2017 | 21/12/2017 |

| Name and mailing address of the ISA/ | Authorized officer |
|---|---|
| European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016 | Doyle, Walter |

1

Form PCT/ISA/210 (second sheet) (April 2005)

# INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2017/075416

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
| --- | --- | --- |
| A | US 2012/263154 A1 (BLANCHFLOWER SEAN MARK [GB] ET AL) 18 October 2012 (2012-10-18)<br>figure 8<br>----- | 1-38 |

1

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2017/075416

| Patent document cited in search report | Publication date | Patent family member(s) | Publication date |
|---|---|---|---|
| US 2012116728 A1 | 10-05-2012 | NONE | |
| US 2014340423 A1 | 20-11-2014 | US 2014340423 A1 | 20-11-2014 |
| | | WO 2014145193 A1 | 18-09-2014 |
| US 2014171116 A1 | 19-06-2014 | CN 104798416 A | 22-07-2015 |
| | | CN 107071135 A | 18-08-2017 |
| | | EP 2932771 A1 | 21-10-2015 |
| | | EP 3171648 A1 | 24-05-2017 |
| | | US 2014171116 A1 | 19-06-2014 |
| | | US 2016295359 A1 | 06-10-2016 |
| | | WO 2014092776 A1 | 19-06-2014 |
| US 2014160161 A1 | 12-06-2014 | NONE | |
| US 2012263154 A1 | 18-10-2012 | CN 103781522 A | 07-05-2014 |
| | | EP 2701818 A1 | 05-03-2014 |
| | | US 2012263154 A1 | 18-10-2012 |
| | | US 2013307874 A1 | 21-11-2013 |
| | | US 2015339839 A1 | 26-11-2015 |
| | | US 2016027221 A1 | 28-01-2016 |
| | | WO 2012142332 A1 | 18-10-2012 |

# PATENT COOPERATION TREATY

From the
INTERNATIONAL SEARCHING AUTHORITY

To: MARY K. NICHOLES
LOWENSTEIN SANDLER LLP
ONE LOWENSTIEN DRIVE
ROSELAND, NJ 07068

# PCT

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

(PCT Rule 43*bis*.1)

Date of mailing (*day/month/year*) 04 AUG 2020

| Applicant's or agent's file reference | FOR FURTHER ACTION |
|---|---|
| 28108-110 | See paragraph 2 below |

| International application No. | International filing date (*day/month/year*) | Priority date (*day/month/year*) |
|---|---|---|
| PCT/US 20/32149 | 08 May 2020 (08.05.2020) | 09 May 2019 (09.05.2019) |

International Patent Classification (IPC) or both national classification and IPC

IPC - G06N 20/00, G06K 9/46, G06K 9/62, G06N 3/08 (2020.01)

CPC - G06N 20/00, G06K 9/629, G06N 3/0454, G06N 3/08, G06K 9/62

Applicant AUTOMOBILIA II, LLC

1. This opinion contains indications relating to the following items:

- [x] Box No. I Basis of the opinion
- [ ] Box No. II Priority
- [ ] Box No. III Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
- [ ] Box No. IV Lack of unity of invention
- [x] Box No. V Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step and industrial applicability; citations and explanations supporting such statement
- [ ] Box No. VI Certain documents cited
- [ ] Box No. VII Certain defects in the international application
- [ ] Box No. VIII Certain observations on the international application

2. **FURTHER ACTION**

If a demand for international preliminary examination is made, this opinion will be considered to be a written opinion of the International Preliminary Examining Authority ("IPEA") except that this does not apply where the applicant chooses an Authority other than this one to be the IPEA and the chosen IPEA has notified the International Bureau under Rule 66.1*bis*(b) that written opinions of this International Searching Authority will not be so considered.

If this opinion is, as provided above, considered to be a written opinion of the IPEA, the applicant is invited to submit to the IPEA a written reply together, where appropriate, with amendments, before the expiration of 3 months from the date of mailing of Form PCT/ISA/220 or before the expiration of 22 months from the priority date, whichever expires later.

For further options, see Form PCT/ISA/220.

| Name and mailing address of the ISA/US | Date of completion of this opinion | Authorized officer |
|---|---|---|
| Mail Stop PCT, Attn: ISA/US<br>Commissioner for Patents<br>P.O. Box 1450, Alexandria, Virginia 22313-1450<br>Facsimile No. 571-273-8300 | 01 July 2020 (01.07.2020) | Lee Young<br>PCT Help Desk<br>Telephone No. 571-272-4300 |

**WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY**

International application No.
PCT/US 20/32149

**Box No. I Basis of this opinion**

1. With regard to the **language**, this opinion has been established on the basis of:
   - [x] the international application in the language in which it was filed.
   - [ ] a translation of the international application into ______________ which is the language of a translation furnished for the purposes of international search (Rules 12.3(a) and 23.1(b)).

2. [ ] This opinion has been established taking into account the **rectification of an obvious mistake** authorized by or notified to this Authority under Rule 91 (Rule 43*bis*.1(b)).

3. [ ] With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, this opinion has been established on the basis of a sequence listing:
   a. [ ] forming part of the international application as filed:
      - [ ] in the form of an Annex C/ST.25 text file.
      - [ ] on paper or in the form of an image file.
   
   b. [ ] furnished together with the international application under PCT Rule 13*ter*.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
   
   c. [ ] furnished subsequent to the international filing date for the purposes of international search only:
      - [ ] in the form of an Annex C/ST.25 text file (Rule 13*ter*.1(a)).
      - [ ] on paper or in the form of an image file (Rule 13*ter*.1(b) and Administrative Instructions, Section 713).

4. [ ] In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.

5. Additional comments:

| WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY | International application No. PCT/US 20/32149 |
|---|---|

**Box No. V** **Reasoned statement under Rule 43*bis*.1(a)(i) with regard to novelty, inventive step and industrial applicability; citations and explanations supporting such statement**

1. Statement

| | | | |
|---|---|---|---|
| Novelty (N) | Claims | 12-17 | YES |
| | Claims | 1-11, 18-20 | NO |
| Inventive step (IS) | Claims | None | YES |
| | Claims | 1-20 | NO |
| Industrial applicability (IA) | Claims | 1-20 | YES |
| | Claims | None | NO |

2. Citations and explanations:

Claims 1-11 and 18-20 lack novelty under PCT Article 33(2) as being anticipated by US 2018/0084310 A1 to GumGum, Inc. (hereinafter 'GumGum').

As to claim 1, GumGum teaches a method comprising: training a convolutional neural network (CNN) using authenticated data and a taxonomy (para [0018]-[0019], [0025]-[0027], [0080] - "A common existing use of video fingerprints is in the field of Digital Rights Management ("DRM") in order for content owners (such as film studios or publishers) to identify when files containing their copyrighted video content are uploaded to video sharing websites or other file sharing networks"; "An image object that is the subject of a classifier that has been created or trained to identify that particular image object may be referred to as a "target image object." For example, a target image object may be a visual representation of a company's logo, and a classifier may be generated specifically to identify at least that logo. In some embodiments, the target image object may generally refer to a class or group of related image objects that may be identified using a particular classifier"; "these neural networks may be trained using video data, but may then be used with either video data or image data as input. In one embodiment, block 612 may be implemented at least in part using a convolutional neural network"); receiving, by a processing device, a query comprising input data (para [0018]-[0019], [0025]-[0027], [0108] - "Video fingerprinting is generally used to compactly represent a video or portion of a video in a manner that allows for efficient searching of other videos to identify either duplicates or slight variations of the same video. For example, computer software may identify, extract and compress characteristics of video data to create a digital file that serves as a "fingerprint" of the specific video data"; "a computing system may analyze a live event broadcast for brand exposures (e.g., instances where a company name, logo or other visual indicator is present in a frame of broadcasted video). For each video sequence that contains a brand exposure, the system may create a digital fingerprint to identify a portion of the video in which the brand exposure occurs. The system may subsequently search video data of potentially hundreds or thousands of subsequent broadcasts across different channels, websites, streaming services, or social media networks to identify occurrences of the digital fingerprint"; "system can query the API with an input parameter such as a keyword, tag, or category, and/or to retrieve media item result sets"); classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy (para [0018]-[0019], [0025]-[0027], [0080]-[0081] - "the computing system at block 612 may use classification models that have each been trained"; "the computing system may access neural networks or other machine learning models that have been trained to identify objects associated with the relevant sponsor(s) for the given team(s) and/or the home venue, then may proceed to sponsor logo identification"); generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data (para [0025]-[0027], [0031], [0087] - "trained image classifier may return a certain probability that a given frame of the video includes a target image object (such as a specific logo), and the system may deem those frames associated with a probability that satisfies a certain threshold likelihood (e.g., 75%) as including the target image object"; "one or more classifiers may provide a separate probability for each of the different companies and/or for each different logo"; "the system may determine confidence levels based on how many pixels the original video portion and subsequent video portion have in common (or pixels that are substantially the same, accounting for slight variations), then based on this confidence level may perform image recognition to verify that a similar video frame is in fact a positive match"; "to supplement and/or confirm the confidence level determinations of the image and video classification models discussed above, the computing system may use text classification models applied to text of a social media post or page in which the image or video is embedded or linked. For example, a textual classifier may provide a confidence score indicating how confident the model is that a particular social media post relates to a specific rights holder and/or sponsor based on the text of the post. This confidence score may be weighted and combined with a corresponding confidence score determined from the image or video itself"); and displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content (para [0036], [0051], [0085] - "the system may generate various spreadsheets, summary views, dashboards, reports, user interfaces and/or other output intended for consumption by a sponsor or other user. The data in one embodiment may include, for each of a number of different URLs, videos, images or social media posts: an indication whether the account is owned and operated by the brand"; "media report may include information for a number of different sponsors' logos, and may be based on the results of multiple passes through the illustrative method").

As to claim 2, GumGum further teaches wherein the authenticated data comprises copyright registered works of authorship, metadata and text (para [0018]-[0019], [0086]-[0087]).

As to claim 3, GumGum further teaches wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings (para [0018]-[0019], [0086]-[0087]).

(See Continuation Box)

WRITTEN OPINION OF THE
INTERNATIONAL SEARCHING AUTHORITY

International application No.
PCT/US 20/32149

**Supplemental Box**

In case **the space in any of the preceding boxes is not sufficient.**
Continuation of:
Box V.2. Citations and explanations:

As to claim 4, GumGum further teaches wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information (para [0034], [0098]).

As to claim 5, GumGum further teaches wherein the authenticated data comprises data from a copyright registered database (para [0018]-[0019], [0090]).

As to claim 6, GumGum further teaches wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture (para [0025], [0074], [0084], [0098]-[0099]).

As to claim 7, GumGum further teaches wherein the input data comprises one or more of image data, video data, intake data or geographical location data (para [0071]).

As to claim 8, GumGum further teaches wherein classifying comprises mapping input data to authenticated data using the taxonomy (para [0025]-[0027], [0083]-[0087]).

As to claim 9, GumGum further teaches wherein the result comprises one or more of an image, a video, text, or sound (para [0036], [0051], [0085], [0098]-[0099]).

As to claim 10, GumGum further teaches wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information (para [0036], [0051], [0085], [0098]-[0099]).

As to claim 11, GumGum further teaches wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy (para [0024]-[0027], [0031]).

As to claim 18, GumGum teaches a system comprising: a memory; a processor, coupled to the memory (para [0038]-[0042]), the processor configured to: train a convolutional neural network (CNN) using authenticated data and a taxonomy (para [0018]-[0019], [0025]-[0027], [0080]); receive, by a processing device, a query comprising input data (para [0018]-[0019], [0025]-[0027], [0108]); classify, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy (para [0018]-[0019], [0025]-[0027], [0080]-[0081]); generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data (para [0025]-[0027], [0031], [0087]); and display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content (para [0036], [0051], [0085]).

As to claim 19, GumGum further teaches wherein the authenticated data comprises copyright registered works of authorship, metadata and text (para [0018]-[0019], [0086]-[0087]).

As to claim 20, GumGum teaches a computer-readable non-transitory storage medium comprising executable instructions that, when executed by a computing device, cause the computing device to perform operations comprising: training a convolutional neural network (CNN) using authenticated data and a taxonomy (para [0018]-[0019], [0025]-[0027], [0080]); receiving, by a processing device, a query comprising input data (para [0018]-[0019], [0025]-[0027], [0108]); classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy (para [0018]-[0019], [0025]-[0027], [0080]-[0081]); generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data (para [0025]-[0027], [0031], [0087]); and displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content (para [0036], [0051], [0085]).

Claims 12-14 and 17 lack an inventive step under PCT Article 33(3) as being obvious over GumGum in view of US 2019/0005670 A1 to Magic Leap, Inc. (hereinafter 'Magic Leap').

As to claim 12, GumGum further teaches wherein the probability is determined by a function (para [0024]-[0027], [0031], [0083]), but does not explicitly teach cross-entropy. However, Magic Leap does teach cross-entropy (para [0032], [0035] - "The network is trained using a standard cross entropy loss"). It would have been obvious to one of ordinary skill in the art to utilize cross-entropy as taught by Magic Leap in order to provide systems and methods for augmenting video data based on automated identification of one or more objects depicted in the video data as taught by GumGum capable of improved and more efficient network training and probabilistic analysis. Both GumGum and Magic Leap are directed to systems and methods utilizing machine learning in conjunction with training of neural networks and augmented reality application.

(See Next Continuation Box)

**WRITTEN OPINION OF THE INTERNATIONAL SEARCHING AUTHORITY**

International application No.
PCT/US 20/32149

**Supplemental Box**

In case **the space in any of the preceding boxes is not sufficient.**
Continuation of:
Box V.2. Citations and explanations:

As to claim 13, GumGum further teaches wherein the result comprises augmented reality content, wherein displaying the result comprises: displaying the result in an augmented reality apparatus, comprising: and displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data (para [0042], [0098]-[0099]), but does not explicitly teach passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device. However, Magic Leap does teach passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device (para [0063] - "Eyepieces 1002A and 1002B may comprise transparent or semi-transparent waveguides configured to direct light from projectors 1014A and 1014B, respectively. Specifically, processing module 1050 may cause left projector 1014A to output a left projected image 1022A into left eyepiece 1002A, and may cause right projector 1014B to output a right projected image 1022B into right eyepiece 1002B. In some embodiments, each of eyepieces 1002 may each comprise a plurality of waveguides corresponding to different colors and/or different depth planes"). It would have been obvious to one of ordinary skill in the art to utilize passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device as taught by Magic Leap in order to provide systems and methods for augmenting video data based on automated identification of one or more objects depicted in the video data as taught by GumGum capable of improved provision of augmented reality user interaction/interfaces.

As to claim 14, GumGum further teaches wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time (para [0018], [0025], [0072], [0084]-[0085], [0090]-[0091]).

As to claim 17, GumGum further teaches wherein the input data comprises user uploaded data, the method further comprising authenticating the user uploaded data using Neural Networks and adding the authenticated user uploaded data to the authenticated data (para [0018]-[0019], [0030]-[0033]), but does not explicitly teach Siamese. However, Magic Leap does teach Siamese (para [0042] - "a siamese network can be used"). It would have been obvious to one of ordinary skill in the art to utilize Siamese as taught by Magic Leap in order to provide systems and methods for augmenting video data based on automated identification of one or more objects depicted in the video data as taught by GumGum capable of improved similarity determination/fingerprinting and data authentication.

Claims 15-16 lack an inventive step under PCT Article 33(3) as being obvious over GumGum in view of US 2019/0102676 A1 to SAS Institute Inc. (hereinafter 'SAS').

As to claim 15, GumGum further teaches further comprising training using authenticated data and a taxonomy (para [0018]-[0019], [0025]-[0027], [0080]), but does not explicitly teach a recurrent neural network (RNN). However, SAS does teach a recurrent neural network (RNN) (para [0148]-[0150] - "the neural network 1200 can be trained"; "the neural network 1200 is a recurrent neural network"). It would have been obvious to one of ordinary skill in the art to utilize a recurrent neural network (RNN) as taught by SAS in order to provide systems and methods for augmenting video data based on automated identification of one or more objects depicted in the video data as taught by GumGum capable of improved and more efficient network training. Both GumGum and SAS are directed to systems and methods utilizing machine learning in conjunction with training of neural networks.

As to claim 16, GumGum further teaches wherein the input data comprises data, the method further comprising: and classifying, by the trained CNN, the structured data (para [0018]-[0019], [0025]-[0027], [0080]-[0081], [0151]). SAS further teaches unstructured (para [0048] - "network-attached data stores 110 may hold unstructured (e.g., raw) data, such as manufacturing data"), processing, by the trained RNN (para [0148]-[0150]), the unstructured data to yield structured data (para [0049] - "computing environment 114 may be used to analyze the unstructured data in a variety of ways to determine the best way to structure (e.g., hierarchically) that data, such that the structured data is tailored to a type of further analysis that a user wishes to perform on the data").

Claims 1-20 have industrial applicability as defined by PCT Article 33(4) because the subject matter can be made or used in industry.

August 2020

# The Content Authenticity Initiative

## Setting the Standard for Digital Content Attribution

# Authors


- Leonard Rosenthol (Adobe)
- Andy Parsons (Adobe)
- Eric Scouten (Adobe)
- Jatin Aythora (The British Broadcasting Corporation)
- Bruce MacCormack (CBC/Radio-Canada)
- Paul England (Microsoft Corporation)
- Marc Levallee (The New York Times Company)
- Jonathan Dotan (Stanford Center for Blockchain Research)
- Sherif Hanna (Truepic)
- Hany Farid (University of California, Berkeley)
- Sam Gregory (WITNESS)


*Version 1.0*

# Table of Contents

# 1 Introduction

## 1.1 Overview

With the increasing velocity of digital content and the democratization of powerful creation and editing techniques, robust content attribution is critical to ensure transparency, understanding, and ultimately, trust.

We are witnessing extraordinary challenges to trust in media. As social platforms amplify the reach and influence of certain content via ever more complex and opaque algorithms, mis-attributed and mis-contextualized content spreads quickly. Whether inadvertent *misinformation* or deliberate deception via *disinformation*, collectively *inauthentic content* is on the rise.

Currently, creators who wish to include metadata about their work (for example authorship) cannot do so in a secure, tamper-evident and standard way across platforms. Without this attribution information, publishers and consumers lack critical context for determining the authenticity of media. This is especially true for users of creative tools that enable augmenting reality with AI or even authoring fully synthetic content who need to be empowered to use their tools responsibly.

Ultimately, the solution to the problem of inauthentic content and the erosion of trust it causes will rely on efforts in three distinct areas:

First is *detection* of deliberately deceptive media. Through a combination of algorithmic identification and human-centered verification of intentionally misleading content the amount of inauthentic content can be reduced. However, as techniques for creating misleading content become more sophisticated and accessible we foresee an escalating arms race impeding progress on this front. As malicious purveyors of content become faster and better, detection techniques will struggle to keep pace.

Second, *education* is essential. Well-intentioned creators and consumers will need to understand the danger of disinformation and the use of techniques to eradicate it. They must also understand ways to use sophisticated creative tools responsibly. These are skills that must be learned and passed on through media literacy campaigns and formal education. We must all understand why and when to trust what we see, hear and read. And we must be equipped with the tools and knowledge to do so.

Finally, we must consider *content attribution*, which is the focus of this paper. Often referred to as provenance, attribution empowers content creators and editors, regardless of their geographic location or degree of access to technology, to disclose information about who created or changed an asset, what was changed and how it was changed. While detection can help address the problem of trust in media reactively by identifying content suspected to be deceptive, attribution proactively adds a layer of transparency so consumers can be informed in their decisions. Content with attribution exposes indicators of authenticity so that consumers can have awareness of who has altered content and what exactly has been changed. This ability to provide content attribution for creators, publishers and consumers is essential to engender trust online.

At the same time, it is critically important that those same content creators be able to protect their privacy when necessary. Any solution attempting to restore trust must be globally viable across technology contexts and minimize opportunities to cause unintended harms or risks. It must also have freedom of creative expression in media production at its core.

We seek to address the issue of content authenticity at scale. To accomplish this, we propose an open, extensible approach for content attribution and have begun working toward establishing standards with broad, cross-industry collaboration.

## 1.2 Background

At Adobe MAX 2019, the Content Authenticity Initiative (CAI) was announced by Adobe in partnership with The New York Times Company and Twitter. Since that time, this group has collaborated with a wide set of representatives from commercial organizations (software tools, publishers, social media), human rights organizations and academic research to produce this paper and the approach it describes.

## 1.3 Our Mission

The initial mission of the CAI is to develop the industry standard for content attribution. By augmenting subjective judgments about authenticity with objective facts about how a piece of content came to be, the CAI aims to help content consumers make more informed decisions about what to trust.

Today, most attribution information is embedded in the metadata of assets via long-established standards such as EXIF and XMP. However, most assets appear on the Web without this information intact. Content moderators, fact-checkers and end-users alike are left to reconstruct context through imperfect and inefficient methods. We will provide a layer of robust, tamper-evident attribution and history data built upon XMP, Schema.org and other metadata standards that goes far beyond common uses today. This attribution information will be bound to the assets it describes, which will in turn reduce friction for creators sharing the attribution data and enable intuitive experiences for consumers who use the information to help them decide what to trust.

We balance simplicity in use of the system with security against tampering and strong links to identity. Identity can be that of an individual, where prudent, or that of the trusted cryptographic signing entity.

There is currently no universal approach for storing attribution data appropriate for all use cases. Depending on the systems involved, this information may be large enough to make it impractical to embed in a file containing digital content (hereafter called an asset). Conversely, some creators may have privacy concerns such that no data associated with an asset nor the asset itself can be stored on servers in the cloud. A cloud-based system may provide durability, whereas a file-based workflow optimizes for disconnected workflows and the preservation of anonymity. Therefore, the CAI imagines data storage to comprise a continuum of options ranging from file-based to cloud-based, with hybrid approaches in between. Flexibility for applications in implementing persistence and flexibility for end users to choose where their data is stored is essential for widespread adoption.

Increasing trust in media requires the ongoing engagement of diverse communities. The CAI does not prescribe a unified single platform for authenticity, but instead presents a set of standards that can be used to create and reveal attribution and history for images, documents, time-based media (video, audio) and streaming content. Although the initial implementations will focus on imagery, our intent is to specify a largely uniform method for enabling attribution from various points of view through which diverse stakeholders can build decentralized knowledge graphs about the trustworthiness of media.

# 2 Guiding Principles

To provide clarity on goals, methodology and purpose, CAI efforts including system design and user experience design are governed by a set of guiding principles. These have informed the work done to date and will continue to be employed to ensure consistency in future work streams.

The terms *must*, *must not*, *should*, *should not*, *required*, and *may* are used in accordance with RFC 2119.

## 2.1 Overarching Goals

CAI specifications **should** provide a mechanism for the producers and custodians of any given content to assert, in a verifiable manner, any information they wish to disclose about the creation of that content and any actions taken since the asset's creation. We refer to such information collectively as provenance.

CAI specifications **should not** provide value judgments about whether a given set of provenance data is "good" or "bad," merely whether the data can be verified as associated with the underlying asset, correctly formed, and free from tampering.

## 2.2 Privacy

CAI specifications **must** respect the common privacy concerns of creators, publishers and consumers of content.

## 2.3 Global Audience / Accessibility

CAI specifications **must** take into consideration the needs of interested users throughout the world.

## 2.4 Interoperability

CAI specifications **should** result in an ecosystem of tools for various types of target users which inter-operate successfully to create, maintain and display provenance information about assets.

## 2.5 Fit with Existing Workflows

CAI specifications **must** fit into the existing workflows of each of the target users, typically through incremental additions to existing tools.

## 2.6 Performance

CAI specifications **should** avoid unreasonable performance characteristics for implementors.

## 2.7 Simplicity and Cost Burden

CAI specifications **should** avoid unreasonable technical complexity and cost burden for implementors.

## 2.8 Extensibility

CAI specifications **should** provide extensibility to allow for extension and evolution of authenticity data.

## 2.9 Misuse

CAI specifications **must** be reviewed with a critical eye toward potential abuse and misuse of the framework. In addition, CAI specifications **must** be reviewed for the ability to be abused and cause unintended harm, threats to human rights, or disproportionate risks to vulnerable groups globally.

# 3 Expected Users

While not intended to limit consideration of other interested parties, CAI users can be broadly understood as:

## 3.1 Content creators

A content creator is someone who wishes to assert information about content they've produced in a way that can be trusted. Common examples include:

- Creative professionals
- Knowledge workers
- Journalists and news media organizations
  - Including both professional and citizen journalists
  - Including those operating in high-risk environments
- Human-rights defenders
- Amateur producers of news media content

## 3.2 Content publishers

Content publishers wish to have better information on which to make decisions about what content to trust. They also wish to credit the proper creator. Common examples include:

- News media organizations
- Social media platforms
- Content distribution networks

## 3.3 Content consumers

Content consumers, those users who interact with assets, want to access authentic content and understand the process by which the content they consume was created. Common examples include:

- Legal professionals including lawyers, investigators, and law enforcement

- Fact-checkers
- Consumers of news media and social media content

### 3.4 Implementors

Implementors build software or hardware tools to create, persist, exchange, or consume CAI provenance data in a way that is interoperable with other CAI-enabled systems.

# 4 Workflows

Though not intended to be exhaustive, the following workflows demonstrate the essential characteristics of the system. While these workflows make use of the same fundamental features, there is distinct value derived for different personas.

### 4.1 Photojournalism

The news media industry is facing a growing sense of distrust in digital media and a changing distribution system as the number of people who receive news from social and non-traditional media sources grows.

The following CAI-enabled workflow is meant to provide trust and transparency for photojournalists, editors and content consumers.

1. A photojournalist uses a CAI-enabled capture device during a newsworthy event they are covering. The photojournalist will set the capture application to the preferred settings of the outlet (i.e. authorship, geolocation, time, file storage preference) and register their identity via the device. The photojournalist then captures images from the newsworthy event with CAI capture attribution details included.
2. The photojournalist then moves their files from a capture application into a photo editing application (e.g. Lightroom, Photo Mechanic, Capture One, etc.). The photojournalist will ensure that the editing application has CAI functionality enabled with the proper settings and manually add metadata about subjects and context as well as complete some light editing.
3. The photojournalist then sends their assets and captions to the appropriate photo editor of their publication. The photo editor opens the assets in a digital imaging tool (e.g. Photoshop), verifies the incoming CAI provenance data, and checks that the data meets editorial standards. The editor then makes edits in accordance with their posted

photo editing guidelines. The photo editor ensures they are utilizing CAI-enabled applications with the appropriate settings throughout their work to ensure that their editing actions are captured and documented.

4. The photo editor works with others to finalize the article to be posted on the website of the news organization. The asset is moved into the content management system of the news organization, which has a CAI implementation so that journalistic context can be displayed and carried through to the website. The article is published.

5. The social media manager then posts links to the article on various social media platforms. While social platforms may alter the asset by, for example, compressing and cropping, the CAI metadata survives these alterations. In fact, these modifications would be added to the CAI data captured in the preprocessing pipeline of the social media platform. The resulting post is CAI-enabled and provides consumers the ability to learn more details about the asset (e.g. Who took it? For what publication? When did they take it?).

6. As other social platform users continue sharing the asset (thus disconnecting it from its original affiliation with the media outlet), CAI data will travel with the asset and any user who sees the asset posted by any other user, will be able to investigate the source and original context of the asset.

7. As required, various analysts from fact-checking organizations will verify the CAI data present in the asset, correlate it to any associated article, and then add their own labels and clarifications to it. Together this collection of CAI data creates a rich, verifiable context that amplifies confidence in the authenticity of the asset.

## 4.2 Creative Professional

Creatives rely on attribution details to receive recognition and compensation for their completed works. The scale of digital media today has made it relatively common to disassociate the creator and context of an asset from the asset itself. Additionally, creative professionals often make edits to assets and combine them into new works for creative purposes, without any intention to deceive. In many cases the creative will have obtained permission to use ingredient assets and that permission may bring obligations of proper attribution that need to be included in the final work. Although this example depicts a visual creative flow, the scenario applies to the creation of audio, video and document assets as well.

The following CAI-enabled workflow is meant to empower creatives to continue using innovative editing techniques without removing or losing authorship attribution.

1. A creative professional opens an authoring application (e.g. Photoshop) to create a new asset. They may start with existing assets that they would like to composite, or from a new empty document. Similar to the photojournalist use case, the creative will authenticate via the software and select the appropriate CAI settings before beginning work, to ensure that CAI data capture is enabled.
2. The creative will configure their settings to ensure no detailed edit history or work-in-progress thumbnails are captured. This is because they are not attempting to represent a news image where there would be a higher level of concern for transformational edits to the asset. The creative is also less likely to want to share their detailed edit history as it may include trade techniques.
3. Given #2, the creative may use AI-assisted techniques to transform the asset into something that does not represent a discrete event that occurred in the real world. The creative can use these tools within their authoring application and have the CAI-enabled process capture only "before" and "after" renditions to share the impact of the change without revealing their trade secrets to consumers.
4. After completing their asset, the creative will save to their preferred file storage system and post to their own distribution points, such as a marketplace or a stock asset site. The creative chooses distribution points that support the CAI to ensure their attribution information remains intact. These distributors have strong incentives to inspect and preserve attribution.
5. A consumer may then view the asset on social media and decide to engage with it. The consumer will be able to click on a CAI icon to be presented with a preview of key attribution information including thumbnails, author, date and a link to follow for more information.
6. When a consumer clicks through to see more information, they will be directed a CAI-enabled website. This website will provide access to the full set of the asset's CAI data, which have been made public in accordance with the creative Professional's preferences.

## 4.3 Human Rights Activist

Across the world, people are capturing digital assets as proof of human rights abuses. Many of these activists are not professional content producers. They may live in areas with high levels of surveillance or lacking high connectivity or with lower digital media literacies. Due to the difficulty of sharing assets documenting abuses without exposing themselves to potential harm, human rights activists may use the CAI standards to share their work widely without compromising their identity. There are considerations specific to this workflow for ensuring activists can protect themselves and the reputation of news outlets who publish their assets.

The following CAI-enabled workflow is meant to provide human rights activists as well as non-governmental organizations (NGOs) and media institutions with an option to capture secure and provable details of an asset without unnecessary exposure of privacy details.

1. A human rights activist will select a capture application. Given the non-professional nature of many human rights activists, some will use tools provided by NGOs (e.g. ProofMode) while the majority will use native capture apps on their phone and/or preferred applications. In the path where the user is actively thinking about tools to use beforehand, they will select their CAI preferences within the tool before capture.

2. The human rights activist then captures something of importance within their community on their smartphone. The user may choose to operate online or offline depending on connectivity and privacy concerns. If operating in an offline workflow, CAI data is written to the file itself. While some users may not act beyond this step for significant periods of time, users who do post are likely to share on messaging and social applications (e.g. WhatsApp or Twitter).

3. If connected to an organization, a user may send the asset to the organization in a similar workflow to photojournalists. If the user is not connected to an organization, it is likely that the asset will be later discovered by an NGO or media outlet on messaging and social platforms. These institutions will be able to read the associated CAI data and confirm key attribution aspects before verifying that the asset is accurate.

4. The NGOs and media outlets will then enter their own verification process. In order to verify the assets, they are likely looking at information such as author, time, location, etc. If the asset has preserved CAI data, it will significantly simplify this process. The institution will then likely try to contact the original source as part of verification and may or may not receive the details they desire to feel comfortable sharing on their platform. (If required, this would also be where an institution may decide to use redaction capabilities, e.g. blur out faces before distributing widely.)

5. Assuming the verification process is completed, in a media context, the workflow is then in the hands of a photo editor who will complete a similar processes as the one in "Photojournalism" to post the asset.

6. A key point of differentiation for the asset with CAI information is that it empowers the source, rather than the institutions publishing the asset, to be the trusted entity. As consumers view the content on social media platforms and/or the CAI website, the attribution details that they can see will start with the original capture and mitigate uncertainty for end users who may or may not trust the publishing institution.

### 4.4 Non-ideal Workflows

The scenarios presented above assume wide adoption of CAI standards. We would be remiss not to acknowledge that in the early phases of adoption, many steps in many workflows will not be CAI-enabled. For example, in the photojournalist case the newsroom may not be able to enforce CAI compliant capture due to software/hardware availability and legacy systems. Here, the lack of end-to-end CAI compliance could be addressed by having the newsroom itself vouch for the legitimacy of assets and add time and location as post-hoc CAI assertions. This can be accomplished if the newsroom creates claims through compliant editing software after capture. It is essential here that the organization's identity be captured in the signing process since trust is shifting from a chain of provenance with multiple identities to a single vouching entity.

These are just some of the many factors to explore to enable partial CAI workflows. Although those details are outside the scope of this paper they must be carefully considered in the design of a pragmatic CAI system.

# 5 System Overview

The proposed system is based on a simple structure for storing and accessing cryptographically verifiable metadata created by an entity we refer to as an actor. An actor can be a human or non-human (hardware or software) that is participating in the CAI ecosystem. For example: a camera (capture device), image editing software, or the person using such tools.

This metadata comprises information regarding asset creation, authorship, edit actions, capture device details, software used and many other subjects. There are standardized types common to most use cases as well as custom types, supported through extensibility. They are represented as claims and assertions as described in "Claims and Assertions". In short, assertions represent distinct pieces of information and claims wrap them into verifiable units. These pieces form the provenance of a given asset.

The CAI embraces existing standards. A core philosophy is to enable rapid, wide adoption by creating only the minimum required novel technology and relying on prior, proven techniques wherever possible. This includes standards for encoding, hashing, signing, compression and metadata.

## 5.1 Claims and Assertions

### 5.1.1 Overview

As shown in "Workflows" above, each of the actors that create or process an asset will produce one or more **assertions** about what they did, when they did it, and (if possible) on behalf of whom. An assertion is typically a JSON-based data structure which represents a declaration made by an actor about an asset at a specific time. Some of these actors will be human and add human-generated information (e.g. copyright) while others are machines (software/hardware) providing information they generated (e.g. camera type or device time). Each type of assertion is either defined in the CAI specification, defined by other metadata standards such as XMP or schema.org or can be custom data for a particular actor or workflow.

Assertions are cryptographically hashed and their hashes are gathered together into a **claim**. A claim is a digitally signed data structure that represents a set of assertions along with one or more cryptographic hashes on the data of an asset. The signature ensures the integrity of the claim and makes the system tamper-evident. A claim can be either directly or indirectly embedded into an asset as it moves through the life of the asset.

Each time the asset reaches a specific key point in its lifecycle, such as initial creation, completion of some editing operations, publication to social media, etc. a new set of assertions and a claim are created. Each new claim refers to the previous claim, thus creating a chain of provenance for the asset (see "Figure 1: Creating a Claim").

Because there are various workflows, some of which are more or less cloud-averse, assertions and claims are designed to live either embedded inside an asset or in the cloud or a combination of the two.

## 5.1.2 Figure 1: Creating a Claim

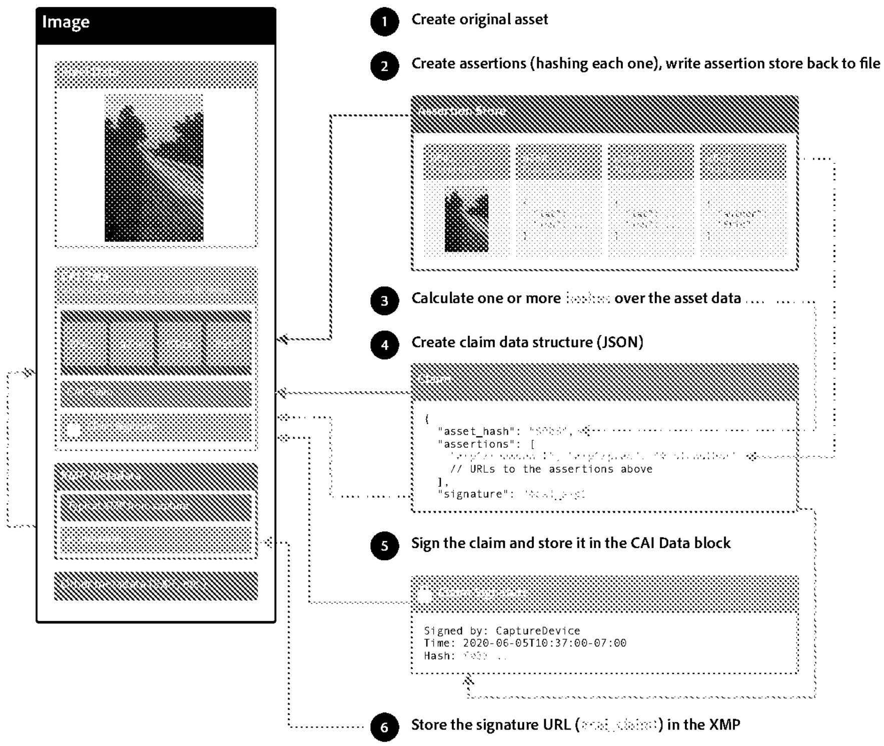



*Diagram of claims and assertions embedded in an image file*

## 5.1.3 Establishing Trust

One key component in establishing trust in the CAI system comes from the entities whose certificates are used for signing the claim. To ensure that only assets signed by trusted actors can be considered properly attributed, it is necessary to create a list of trusted certificates or their certification authorities (CAs).

Similar to the EU Trust List, the Adobe Approved Trust List, and similar lists used by web browsers and operating systems, the members of the CAI will establish their own Trust List of certificates that can be used to sign claims. Details on the governance of the Trust List is outside the scope of this paper.

In many cases, the holder of the certificate will not the individual who created (or edited) an asset, but instead will be the entity responsible for the hardware or software that they used. The signing certificate belongs to the actor (e.g. Truepic Camera, Adobe Photoshop, BBC, etc.) that performed the actions *on behalf of* someone else. This model allows CAI to provide anonymity (and/or pseudonymity) where desired. For scenarios where the certificate holder is able to reliably establish the identity of the individual, and the individual wishes their identity associated with an asset, an identity assertion is used.

### 5.1.4 Identity

One of the assertion types that can be present in a claim is Identity. This digital identity (also sometimes referred to as a *Subject* or an *Entity*) is present when an individual (or organization) is making a clear statement about their association with this claim.

Digital identity fundamentally requires digital identifiers — strings or tokens that are unique within a given scope (globally or locally within a specific domain, community, directory, application, etc.). In order to support a variety of use cases, including those where identity might be anonymous or pseudonymous, it is important that various schemes for the identifiers are available for use. Fortunately, most common identity formats such as Decentralized Identifiers-DID, WebIDs, OpenID, ORCiD and others are all based on URIs. This enables an identity assertion to be expressed in the standard format described in RFC 3986.

#### 5.1.4.1 Decentralized Identifiers

Decentralized Identifiers are particularly well-suited to capturing identity in the CAI attribution system because they adapt well to different authentication scenarios. Although the name implies that a DID is to be used in a decentralized environment (commonly in conjunction with a blockchain) rather than a centralized one, there is nothing in that specification that prevents it. In fact, the specification itself not only calls that out, but points out this flexibility as a benefit of DID.

#### 5.1.4.2 Non-URL formatted identities

Other standards used for identification that are not represented as URLs must be encoded as a URL in order to store it in an identity assertion. For example, ISNI.org recommends that the ISNI - ISO 27729 identifier be added to the end of a standard URL reference to their site like this: http://www.isni.org/isni/0000000114559647.

### 5.1.5 ClaimReview

CAI claims can be augmented with human-generated review assertions so that fact checking professionals can provide additional evidence of authenticity. We can leverage existing standard schemas for this.

One popular schema used for this purpose is ClaimReview which can be used in conjunction with MediaReview to add fact-checking reviews to images and other assets. By having a standard assertion type that contains a ClaimReview instance, these fact checks can now be embedded into the asset itself, enabling additional checks related to the context where the asset appears.

By embracing ClaimReview and other standards like it, the CAI will support fact checking with rich metadata to optimize verification workflows. This will help keep the fact checking ecosystem decentralized and diverse.

### 5.1.6 Redaction of Assertions

In many workflows, there is a need for assertions to be removed by subsequent processes, either because publishing the assertion would be problematic (e.g. the identity of the person who captured a video) or the assertion is no longer valid (e.g. an earlier thumbnail showing something that has since been cropped out). The CAI allows for the redaction of these assertions in a verifiable way that is also part of the provenance of the asset.

In the process of redacting an assertion, a record that something was removed is added to the claim. Because each assertion's reference includes the assertion type, it is clear what type of information (eg. thumbnail, location, etc.) was removed. This enables both humans and machines to apply rules to determine if the removal is acceptable.

NOTE: Assertion redacted only applies to assertions that are part of the CAI data. It does not have anything to do with removal of other metadata (XMP, EXIF, etc.).

## 5.2 Use of XMP

XMP, as defined in ISO 16684-1, is the standard for embedded metadata in numerous asset types. The CAI leverages its standard rules to ensure that the CAI claim can be reliably retrieved using existing technology including a variety of open source tooling.

## 5.3 Composed Documents

It is very common that a content author will integrate other assets (e.g. stock art/photos) into the work that they are creating. In the world of dynamic media, such as video and audio, this is the normal mode of operation where various clips are combined to form the final production. These scenarios produce what are called *Composed Documents* as described in the "pantry and ingredient" model of XMP. (See the Partner Guide to XMP for Dynamic Media.)

To fully understand the complete history and attribution of an asset that has been created as part of a composed document, it is necessary to include or reference each ingredient along with any claims and assertions made on them. The CAI provides for this with a specific type of assertion that references each ingredient's claims and assertions, whether they are embedded into the new composed document or stored in the cloud.

## 5.4 Digital Signatures

In order to ensure the integrity of the CAI information that is embedded into an asset, including the asset's data itself, digital signature technology is employed. That same technology also provides an authentication mechanism connected to the signer, that provides a way to establish trust in the actors involved in the CAI-enabled workflow.

Because the data being signed is, in signature terms, "arbitrary message content," a standard Cryptographic Message Syntax (CMS) signature is used to sign the asset. While a simple CMS signature is sufficient, CAI recommends using a CMS Advanced Electronic Signature (CAdES) instead. Use of a CAdES-compliant signature will ensure that the CAI data complies with the European eIDAS legislation.

It is also recommended that the signer should use a trusted timestamp authority to generate a trusted time-mark or time-stamp token. This is for proving that the signature itself actually existed at a certain date and time and can be incorporated into the CAdES content-time-stamp attribute to create a CAdES-T compliant signature.

## 5.5 Illustrative Flow

A brief, high-level sequence of events for a typical cycle of claim creation and verification is useful to convey how the system operates. This is illustrative, not exhaustive and could apply to any of the "Workflows" above. All data capture is optional; when we refer to identity or location, for example, it is important to note that whether to record such data is a decision of the user. In addition, some of the information in this example is shown embedded in the asset while other information is in the cloud. The choices here are purely for example since any assertion or claim can be stored either in the asset or in the cloud.

1. A capture device captures an image and concurrently creates a set of assertions about the image such as capture location, equipment details, identity of the user, and perhaps a representative thumbnail. These assertions are embedded in the asset.
2. A claim, referring to the set of assertions and including a hash of the image, is created and cryptographically signed by a trusted signing authority on behalf of the user. The claim is then embedded into the image and a reference to it stored in the asset's metadata (see "Figure 1: Creating a Claim").

3. The image, with its attached claim, is imported into an editing tool. The prior claim is verified (see "Figure 2: Validating Claims"), such that upon successful verification its assertions and claims can be carried forward as the image's history is accumulated.

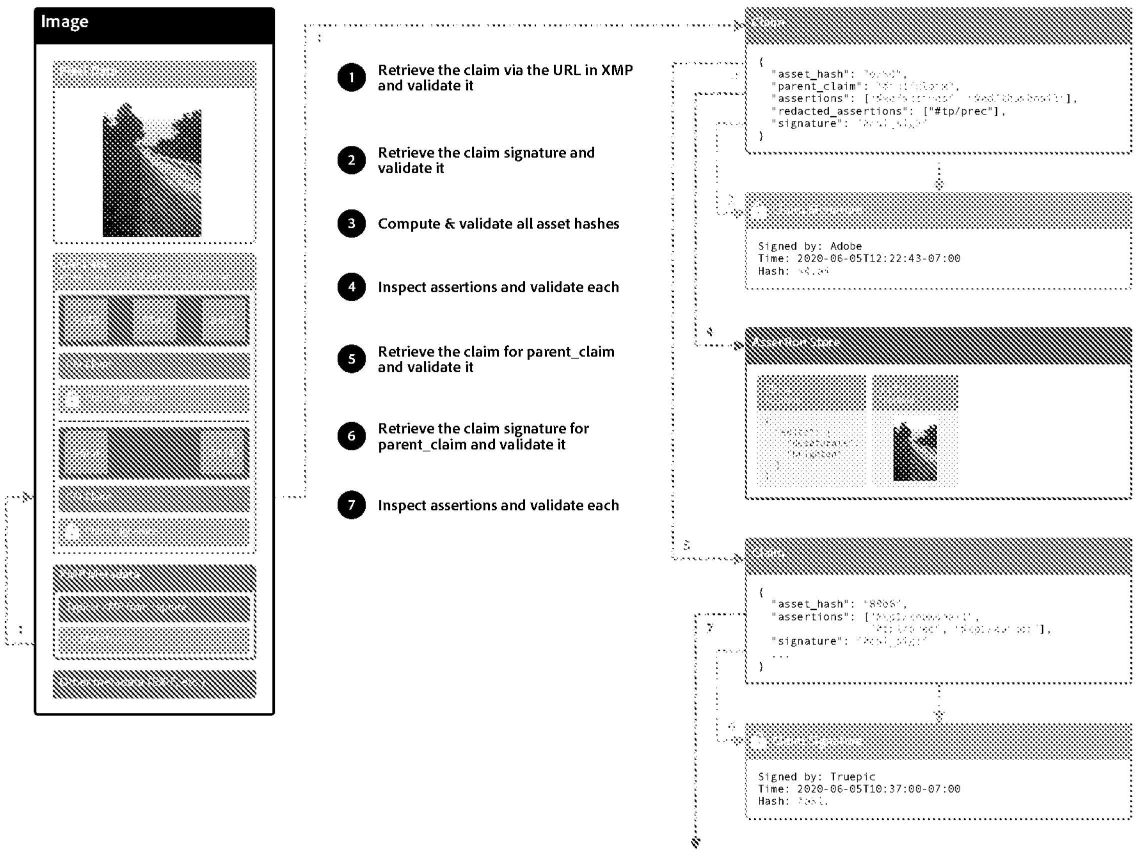
4. As edits are made, additional assertions are captured and stored in a cloud service. Upon image export, these are gathered together as part of a second claim along with the hash of the updated image. A URL to the claim is stored in the asset's metadata. This new claim (which is also stored in the cloud) refers to the prior claim, therefore ensuring that all assertions from both claims are accessible via the exported image without any link to the prior version of the image.

5. The edited image, with its CAI metadata attached, is shared on a platform.

6. A content consumer encounters the image on the Web via a CAI-enabled viewing experience and sees a visual indication that CAI claims are present.

7. Finally, the content consumer interacts with the visual indicator to learn about the image's history. The most recent claim is retrieved and verified along with the entire chain of claims and their component assertions (see "Figure 2: Validating Claims"). The assertions are displayed to the viewer in a clear, time-ordered user experience depicting the image's claim history from inception to display.

## 5.6 Figure 2: Validating Claims




*Validation of Claims and Assertions*

# 6 User Experience

There are two types of user experience (UX) each CAI implementor may need to carefully consider, depending on the goal of the implementation. First, that of creators using the implementor's tools, and second that of consumers viewing content on the implementor's platform.

Both types of experiences should optimize for clarity and provide guidance for users who may have questions about what they are seeing. Above all, experiences should default to the use of consistent terminology and iconography in order to achieve this so that over time expectations of what the CAI provides and its realized value are precisely aligned.

For content creators, it is important to ensure that data sharing via CAI claims is well understood by the user and that publishing CAI claims is not an automatic process. While flows may benefit from a simplified UX, requiring the user to make intentional decisions about what kind of CAI data to record will prevent unintended claim capture. It is strongly suggested that creator tools support previewing claim content before it is signed and attached.

Content consumers will be best served by a different set of UX principles to help them make decisions about what to trust through the presentation of claim evidence. For that reason, it is important that if a visual indication is associated with a valid asset, it should indicate only that CAI data is present and verified, and should *not* appear to indicate whether the content is authentic. In cases where the asset and its CAI data do not match, this should be clearly indicated as well.

For the purposes of this discussion, platforms on which content consumers view content can be native desktop or mobile applications, web sites and web browser integrations.

Since CAI data can contain a depth of detail that is not always relevant for all viewers, we envisage a UX based on the idea of progressive disclosure. This means users are presented a small amount of critical data up front, then empowered to reveal more detail by interacting with the user interface. Which information is most critical for users will have to be carefully evaluated for each situation, but in general it is recommended that the user see when an asset was modified, how it was modified and by whom as top-level information. This model also helps to provide a solution suitable for diverse levels of digital literacy.

There will often be multiple claims for an asset. Implementors should think carefully about how to display them so viewers are provided the simplest and fastest possible path to decisions about what to trust. In some cases a straightforward list of assertions in reverse chronological order might be appropriate. However, a clear visual classification of assertions based on the type of actor that created them (hardware device, software program, human fact checker, etc.) could be a very powerful way to help users decide which assertions matter most in a given context.

Not all information in CAI claims is the same. Each has different dependencies and potential vulnerabilities. For instance, on-device camera information (such as lens used) is different from creation date or location information which is dependent on an external signal (e.g. clock or GPS). An optimal UX for viewers will indicate this, through progressive disclosure, without overly complicating the experience.

# 7 Security, Trust & Privacy Considerations

## 7.1 Educating the User

The presence of the CAI information in an asset means that there is evidence about the asset that the user can use to make their own determination of trust. For a user to fully understand what CAI does (and does not) achieve, it is important that they understand how CAI works, what the disclosed information means, and more. This information needs to be useful but not overwhelming. These considerations apply to both the people who are creating and editing content, who will need to understand why they need to enable it for their assets *but* also how consumers will understand the information to consider it worthwhile.

One key aspect of the user experience presented to users (and perhaps the whole CAI ecosystem) that should be called out specifically is the implication of "trust" in the system. It is extremely important that a consumer of assets not interpret more trust in the presence of valid CAI information than it truly means — specifically that an asset with valid CAI information does not imply anything about the trustworthiness of the content of the asset.

This education is key in preventing social engineering attacks.

## 7.2 Unexpected Disclosure

In cases where adding identity (or any other form of identifying attribution) to an asset at the time of capture could lead to increased risk to content producers or others, it is important that assertions and claims can be added at some later time. The CAI assertion and claim model allows for any type of assertion to be added at any time. In addition, the redaction capability supports the inverse case where *too much* information was added and some needs to be removed.

Of course, once an asset has been released "into the wild" the information contained inside cannot be modified there, as control has been relinquished. When assertions are stored in the cloud, it may be possible to have those assertions removed or made inaccessible by the hosting provider, but it may not be possible to track down and remove all copies of that data cached by third parties.

## 7.3 Certificate Trust

A number of potential concerns arise from attacks against the certificates used to sign the claims of an asset. To address these concerns, the CAI ecosystem is built around its own CAI Trust List for "Establishing Trust". This provides CAI with a well-defined model for allowing only certain approved signers and managing the lifecycle of those certificates (e.g. revocation, expiration, etc.).

Another design goal of the CAI is flexibility in certificates used for signing claims. In most cases, signing certificates tied directly to an individual should not be used, but instead organization certificates (e.g. Adobe, Twitter, etc.) or even one time private keys will be used to sign *on behalf of* the actual user performing the actions on the asset. This is recommended due to the inherent complexity required of users to manage their own keys. Using organization certificates also has the important advantage that the domain of the certificate can be matched to URLs in the claim data, thus preventing link manipulation attacks.

Implementors should take appropriate precautions to ensure that their signing keys cannot be stolen or compromised. Use of a cloud service, hardware enclave and use of a One Time Private Key are some possible approaches that will be taken.

## 7.4 Distributed Ledger Technology

The use of a Trust List is the proposed model for early CAI implementations, but it is not the only possible approach to underwriting trust.

Distributed Ledger Technology (DLT) offers a consensus model for replicated, shared and synchronized data.

While other approaches to secure content attribution have investigated the use of such technology, it was felt that mandating a single ledger for all authored content may not be globally scalable for the CAI and may be at odds with the spirit of decentralization. Implementors in the CAI ecosystem might opt to use a DLT to federate their storage of assertions and claims to achieve an additional level of integrity and transparency.

Having a distributed ledger as a secondary model for storing information about issued claims (e.g. their hashes) could serve as proof that data stored in a given provider's cloud has not been modified or tampered with (either intentionally or unintentionally).

## 7.5 Intentional Misattribution

The CAI attribution model does not prevent a malicious user from stripping all of the CAI data (claims and assertions) from an asset and then adding new claims representing themselves as the originator. Similarly, the "analog hole" or "rebroadcast attack," which are common terms for subverting provenance systems by capturing an image of a photograph or computer screen, are not addressed directly by the model.

However, there are some solutions that can be implemented in concert with the CAI model to achieve resilience against intentional misuse.

- An actor could use watermarking technology to durably embed information (either perceptibly or imperceptibly) about the asset's current claim. The watermark could be subsequently used to recover provenance data.
- A camera device or software could utilize depth mapping to capture scene information (as CAI assertions) which would indicate whether a photograph depicts a 3D scene or a rebroadcast photo of a photo.

- It is possible for systems using "similarity detection" to help users find additional information about an asset. For example, users could be shown whether the asset was published at some prior date or the asset could be contextualized by surfacing substantially similar assets.
- Trusted timestamps can be used to cast doubt on assets with deliberately altered histories. For example, when identical assets with different claim histories are encountered, the earlier CAI claims are likely to be trustworthy while the later ones may represent an attempt to alter history.
- Identity, actions and other assertions that include information about specific domains or organizations could be compared against the domain associated with the signing certificate. This would, for example, prevent the use of a signing certificate from "badsoftware.com" claiming that the user was using "Adobe Photoshop" to edit the image.

While not part of the core CAI infrastructure, such solutions easily integrate with CAI and provide great utility. These techniques support the application of judgment and reason — they are not technological guarantees.

# 8 Future Work Streams

Current and future CAI collaborators will focus on several work streams with the goal of incorporating diverse points of view while stewarding the ideas expressed here toward a unified, pragmatic, adoptable standard.

## 8.1 Establishment of Working Groups

To move from high-level system concepts to detailed specifications, several Working Groups have been created. Working Groups are open to any interested organization or individual. When appropriate, a Working Group will produce one or more specifications for peer review and publication.

## 8.2 Extension of Design for Additional Formats

Many details have yet to be proposed for ensuring the system can embrace time-based media like audio, video, and streaming formats. While early work has been done to ensure the design does not preclude these, it is important for the collaborators to focus on them next. This work would be done in conjunction with experts in the field — not only about the formats but also about common workflows involving them.

### 8.3 Prototype Exploration

With the goal of understanding the practical considerations for implementing CAI-compliant devices and applications, we expect that many prototypes will be created and tested in parallel with Working Group efforts. We define success as wide adoption of carefully vetted specifications. To achieve this, exploratory proofs of CAI concepts will be built by CAI collaborators and learnings shared with the community.

# 9 Conclusion

The collaborators on this paper have explored the challenges of inauthentic media through problem definition, system design and use case research. The results of the exploration are expressed in the design of the CAI provenance system. To achieve widespread adoption we have based the design on existing standards and established techniques, and acknowledge that the system will need to include simple and intuitive user experiences.

However, even an optimally designed system cannot ultimately succeed in a vacuum. We now begin the important work of deeper, more expansive collaboration with leaders in technology, media, academia, advocacy and other disciplines.

With this first step towards an industry standard for digital content attribution, we look optimistically to a future with more trust and transparency in media.

# 10 References


- JSON
- eXtensible Metadata Platform (XMP)
- Partner Guide to XMP for Dynamic Media
- Decentralized Identifiers-DID
- WebIDs
- OpenID
- Cryptographic Message Syntax (CMS)
- European eIDAS legislation

# 11 Contributors

This paper would not have been possible without the substantial contributions of these individuals.

- Will Allen (Adobe)
- Pia Blumenthal (Adobe)
- John Collomosse (Adobe)
- Oliver Goldman (Adobe)
- Andrew Kaback (Adobe)
- Gavin Peacock (Adobe)
- Charlie Halford (The British Broadcasting Corporation)
- Scott Lowenstein (The New York Times Company)
- Thomas Zeng (Truepic)
- Fabiana Meira Pires De Azevedo (Twitter)
- Corin Faife (WITNESS)

# Electronic Patent Application Fee Transmittal

| | |
|---|---|
| **Application Number:** | |
| **Filing Date:** | |
| **Title of Invention:** | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY |
| **First Named Inventor/Applicant Name:** | Lucinda Lewis |
| **Filer:** | Mary Kristin Nicholes/Susan Schuchard |
| **Attorney Docket Number:** | 28108-125 |

Filed as Micro Entity

**Filing Fees for** **U.S. National Stage under 35 USC 371**

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Basic Filing:** | | | | |
| NATIONAL STAGE FEE | 3631 | 1 | 80 | 80 |
| **Pages:** | | | | |
| **Claims:** | | | | |
| **Miscellaneous-Filing:** | | | | |
| **Petition:** | | | | |
| **Patent-Appeals-and-Interference:** | | | | |
| **Post-Allowance-and-Post-Issuance:** | | | | |

| Description | Fee Code | Quantity | Amount | Sub-Total in USD($) |
|---|---|---|---|---|
| **Extension-of-Time:** | | | | |
| **Miscellaneous:** | | | | |
| | **Total in USD ($)** | | | **80** |

# Electronic Acknowledgement Receipt

| | |
|---|---|
| EFS ID: | 44236337 |
| Application Number: | 17609911 |
| International Application Number: | PCT/US2020/032149 |
| Confirmation Number: | 1651 |
| Title of Invention: | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY |
| First Named Inventor/Applicant Name: | Lucinda Lewis |
| Customer Number: | 28221 |
| Filer: | Mary Kristin Nicholes |
| Filer Authorized By: | |
| Attorney Docket Number: | 28108-125 |
| Receipt Date: | 09-NOV-2021 |
| Filing Date: | |
| Time Stamp: | 13:20:34 |
| Application Type: | U.S. National Stage under 35 USC 371 |

## Payment information:

| | |
|---|---|
| Submitted with Payment | yes |
| Payment Type | DA |
| Payment was successfully received in RAM | $80 |
| RAM confirmation Number | E2021A9D20593383 |
| Deposit Account | 501358 |
| Authorized User | Mary Nicholes |

The Director of the USPTO is hereby authorized to charge indicated fees and credit any overpayment as follows:

37 CFR 1.492 (National application filing, search, and examination fees)

37 CFR 1.21 (Miscellaneous fees and charges)

37 CFR 1.20 (Post Issuance fees)

37 CFR 1.19 (Document supply fees)

37 CFR 1.17 (Patent application and reexamination processing fees)

## File Listing:

| Document Number | Document Description | File Name | File Size(Bytes)/ Message Digest | Multi Part /.zip | Pages (if appl.) |
|---|---|---|---|---|---|
| 1 | Transmittal of New Application | 28108_125_371Transmittal_9NOVEMBER2021.pdf | 315784<br>234ca1d3907993bb889e1b517ba5856a1de73d56 | no | 4 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 2 | Certification of Micro Entity (Gross Income Basis) | 28108_125_National_Stage_micro_entity_form_9NOV2021.pdf | 120328<br>8befe79382bbd69fad27de01a8237c65cfb1090f | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 3 | Application Data Sheet | 28108_125_ADS_9NOVEMBER2021.pdf | 1256229<br>811cf0b72d5a0ffe7a71496c5244436c69424336 | no | 8 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 4 | | 28108_125_WO_APPLICATION_9NOV2021.pdf | 2584526<br>81cb18a170c487e4955f73dcfef24284d88c8c0d | yes | 45 |

**Multipart Description/PDF files in .zip description**

| Document Description | Start | End |
|---|---|---|
| Abstract | 1 | 1 |
| Specification | 2 | 34 |
| Claims | 35 | 38 |
| Drawings-only black and white line drawings | 39 | 45 |

**Warnings:**

**Information:**

| | | | | | |
|---|---|---|---|---|---|
| 5 | | 28108_125_APPLICATION_9NOV2021.pdf | 348277<br>b3d630ee9d9b9af494d5729ebf956e7e041e587d | yes | 38 |

**Multipart Description/PDF files in .zip description**

| Document Description | Start | End |
|---|---|---|
| Specification | 1 | 33 |
| Claims | 34 | 37 |
| Abstract | 38 | 38 |

**Warnings:**

**Information:**

| | | | | | |
|---|---|---|---|---|---|
| 6 | Transmittal Letter | 28108_125_IDS_Transmittal_9NOVEMBER2021.pdf | 93691<br>1d60b4a0a81229ba9ce7aac90a5f8a174cdf9017 | no | 2 |

**Warnings:**

**Information:**

| | | | | | |
|---|---|---|---|---|---|
| 7 | Information Disclosure Statement (IDS) Form (SB08) | 28108_125_IDS_9NOVEMBER2021.pdf | 613046<br>22dafd7cc91c14dbd92e9bd642cc47b1d0b185b6 | no | 5 |

**Warnings:**

**Information:**

| | | | | | |
|---|---|---|---|---|---|
| 8 | Foreign Reference | 28108_125_Foreign_Ref_2018197835.pdf | 528429<br>e8487070af582d8d876b6d7e129d4969fa659401 | no | 41 |

**Warnings:**

**Information:**

| | | | | | |
|---|---|---|---|---|---|
| 9 | Foreign Reference | 28108_125_Foreign_Ref_2018065549.pdf | 304583<br>21598065a7607f0d18bb1da9995660117971673b | no | 33 |

**Warnings:**

**Information:**

| | | | | | |
|---|---|---|---|---|---|
| 10 | Non Patent Literature | 28108_125_ISR_9NOV2021.pdf | 351207<br>1807e375e3f979c98ba57407ebce6417e874b050 | no | 5 |

**Warnings:**

| | | | | | |
|---|---|---|---|---|---|
| **Information:** | | | | | |
| 11 | Non Patent Literature | White_Paper_NPL.pdf | 4219516<br>47821da58487c3337739bd5cf31be128935cfc68 | no | 28 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| 12 | Fee Worksheet (SB06) | fee-info.pdf | 37711<br>df00a7585ec7dc10d8e0e4c868599dd21c1d86d1 | no | 2 |
| **Warnings:** | | | | | |
| **Information:** | | | | | |
| | | | **Total Files Size (in bytes):** | 10773327 | |

**This Acknowledgement Receipt evidences receipt on the noted date by the USPTO of the indicated documents, characterized by the applicant, and including page counts, where applicable. It serves as evidence of receipt similar to a Post Card, as described in MPEP 503.**

**New Applications Under 35 U.S.C. 111**

**If a new application is being filed and the application includes the necessary components for a filing date (see 37 CFR 1.53(b)-(d) and MPEP 506), a Filing Receipt (37 CFR 1.54) will be issued in due course and the date shown on this Acknowledgement Receipt will establish the filing date of the application.**

**National Stage of an International Application under 35 U.S.C. 371**

**If a timely submission to enter the national stage of an international application is compliant with the conditions of 35 U.S.C. 371 and other applicable requirements a Form PCT/DO/EO/903 indicating acceptance of the application as a national stage submission under 35 U.S.C. 371 will be issued in addition to the Filing Receipt, in due course.**

**New International Application Filed with the USPTO as a Receiving Office**

**If a new international application is being filed and the international application includes the necessary components for an international filing date (see PCT Article 11 and MPEP 1810), a Notification of the International Application Number and of the International Filing Date (Form PCT/RO/105) will be issued in due course, subject to prescriptions concerning national security, and the date shown on this Acknowledgement Receipt will establish the international filing date of the application.**

PTO-1390 (10-20)
Approved for use through 10/31/2022. OMB 0651-0021
U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE
Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

| **TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A SUBMISSION UNDER 35 U.S.C. 371** | | Attorney Docket No. 28108-125<br>U.S. Application No. (if known, see 37 CFR 1.5) |
|---|---|---|
| International Application No.<br>PCT/US2020/032149 | International Filing Date<br>2020-05-08 | Priority Date Claimed<br>2019-05-09 |

Title of Invention
METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

First Named Inventor
LEWIS, Lucinda

**Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information.**

1. ☒ This is an express request to begin national examination procedures (35 U.S.C. 371(f)). NOTE: The express request under 35 U.S.C. 371(f) will not be effective unless the requirements under 35 U.S.C. 371(c)(1), (2), and (4) for payment of the basic national fee, copy of the International Application and English translation thereof (if required), and the oath or declaration of the inventor(s) have been received.
2. ☒ A copy of the International Application (35 U.S.C. 371(c)(2)) is attached hereto (not required if the International Application was previously communicated by the International Bureau or was filed in the United States Receiving Office (RO/US)).
3. An English language translation of the International Application (35 U.S.C. 371(c)(2))
   a. ☐ is attached hereto.
   b. ☐ has been previously submitted under 35 U.S.C. 154(d)(4).
4. An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4))
   a. ☐ is attached.
   b. ☒ was previously filed in the international phase under PCT Rule 4.17(iv).

**Items 5 to 8 below concern amendments made in the international phase.**

PCT Article 19 and 34 amendments

5. ☐ Amendments to the claims under PCT Article 19 are attached (not required if communicated by the International Bureau) (35 U.S.C. 371(c)(3)).
6. ☐ English translation of the PCT Article 19 amendment is attached (35 U.S.C. 371(c)(3)).
7. ☐ English translation of annexes (Article 19 and/or 34 amendments only) of the International Preliminary Examination Report is attached (35 U.S.C. 371(c)(5)).

Cancellation of amendments made in the international phase

8a. ☐ Do not enter the amendment made in the international phase under PCT Article 19.

8b. ☐ Do not enter the amendment made in the international phase under PCT Article 34.

NOTE: A proper amendment made in English under Article 19 or 34 will be entered in the U.S. national phase application absent a clear instruction from applicant not to enter the amendment(s).

**The following items 9 to 17 concern a document(s) or information included.**

9. ☒ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
10. ☐ A preliminary amendment.
11. ☒ An Application Data Sheet under 37 CFR 1.76.
12. ☐ A substitute specification. NOTE: A substitute specification cannot include claims. See 37 CFR 1.125(b).
13. ☐ A power of attorney and/or change of address letter.
14. ☐ A computer-readable form of the sequence listing in accordance with PCT Rule 13*ter*.3 and 37 CFR 1.821-1.825 (not required if sequence listing in text format was indicated on the PCT Request as part of the International Application and the sequence listing was published as part of the international application).
15. ☐ Assignment papers *(cover sheet and document(s))*. Name of Assignee: ______________________________
16. ☐ 37 CFR 3.73(c) Statement *(when there is an Assignee)*.

| U.S. APPLN. No. (if known – see 37 CFR 1.5) | INTERNATIONAL APPLICATION No. | ATTORNEY DOCKET No. |
|---|---|---|
| | PCT/US2020/032149 | 28108-125 |

17. ☐ Other items or information:

| The following fees have been submitted. | CALCULATIONS |
|---|---|
| 18. ■ Basic national fee (37 CFR 1.492(a)) …… **$320** | $ 80.00 |
| 19. ■ Examination fee (37 CFR 1.492(c))<br>• If the written opinion prepared by ISA/US or the international preliminary examination report prepared by IPEA/US indicates all claims satisfy provisions of PCT Article 33(1)-(4)…… **$0**<br>• All other situations …… **$800** | $ 0 |
| 20. ☐ Search fee (37 CFR 1.492(b))<br>• If the written opinion prepared by ISA/US or the international preliminary examination report prepared by IPEA/US indicates all claims satisfy provisions of PCT Article 33(1)-(4)…… **$0**<br>• Search fee (37 CFR 1.445(a)(2)) has been paid on the international application to the USPTO as an International Searching Authority …… **$140**<br>• International Search Report prepared by an ISA other than the US and provided to the Office or previously communicated to the US by the IB…… **$540**<br>• All other situations …… **$700** | $ 0 |
| **TOTAL OF 18, 19, and 20 =** | $ 80.00 |
| ☐ Additional fee for specification and drawings filed in paper over 100 sheets (excluding sequence listing in compliance with 37 CFR 1.821(c) or (e) in an electronic medium or computer program listing in an electronic medium) (37 CFR 1.492(j)).<br><br>Fee for each additional 50 sheets of paper or fraction thereof …… **$420** | $ |

| Total Sheets | Extra Sheets | Number of each additional 50 or fraction thereof (round **up** to a whole number) | RATE |
|---|---|---|---|
| - 100 = | / 50 = | | x **$420** |

| | |
|---|---|
| Surcharge for furnishing any of the search fee, examination fee, or the oath or declaration after the date of commencement of the national stage (37 CFR 1.492(h))…… **$160** | $ |

| CLAIMS | NUMBER FILED | NUMBER EXTRA | RATE | |
|---|---|---|---|---|
| Total claims | - 20 = | | x **$100** | $ |
| Independent claims | - 3 = | | x **$480** | $ |
| MULTIPLE DEPENDENT CLAIM(S) (if applicable) | | | **+ $860** | $ |

| | |
|---|---|
| Fee for submission of Sequence Listing text file of 300 MB to 800 MB (37 CFR 1.21(o)(1))…… **$1,060** | $ |
| Fee for submission of Sequence Listing text file of more than 800 MB (37 CFR 1.21(o)(2))……**$10,500** | $ |
| Processing fee for furnishing the English translation later than 30 months from the earliest claimed priority date (37 CFR 1.492(i))……**$140** + | $ |
| **TOTAL OF ABOVE CALCULATIONS =** | $ |
| ■ **Applicant asserts small entity status.** See 37 CFR 1.27. Fees above are reduced by ½. | |
| ■ **Applicant certifies micro entity status.** See 37 CFR 1.29. Fees above are reduced by ¾. Applicant must attach form PTO/SB/15A or B or equivalent. | |
| **TOTAL NATIONAL FEE =** | $ 80.00 |
| Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31)…… **$50.00** per property + | $ |
| **TOTAL FEES ENCLOSED =** | $ 80.00 |

a. ☐ A check in the amount of $____________ to cover the above fees is enclosed.

b. ■ Please charge my Deposit Account No. 501358 in the amount of $80.00 to cover the above fees.

c. ■ The Director is hereby authorized to charge additional fees which may be required, or credit any overpayment, to Deposit Account No. 501358 as follows:

i. ■ any required fee.

ii. ☐ any required fee except for excess claims fees required under 37 CFR 1.492(d) and (e) and multiple dependent claim fee required under 37 CFR 1.492(f).

d. ☐ Fees are to be charged to a credit card. **WARNING:** Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038. The PTO-2038 should only be mailed or faxed to the USPTO. However, when paying the basic national fee, the PTO-2038 may NOT be faxed to the USPTO.

**ADVISORY:** If filing by EFS-Web, do **NOT** attach the PTO-2038 form as a PDF along with your EFS-Web submission. Please be advised that this is **not** recommended and by doing so your **credit card information may be displayed via PAIR**. To protect your information, it is recommended to pay fees online by using the electronic payment method.

**NOTE: Where an appropriate time limit under 37 CFR 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the International Application to pending status.**

**Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications**

☐ This application (1) claims priority to or the benefit of an application filed before March 16, 2013, and (2) also contains, or contained at any time, a claim to a claimed invention that has an effective filing date on or after March 16, 2013.

**NOTE 1:** By providing this statement under 37 CFR 1.55 or 1.78, **this application, with a filing date on or after March 16, 2013, will be examined under the first inventor to file provisions of the AIA**.

**NOTE 2:** A U.S. national stage application may not claim priority to the international application of which it is the national phase. The filing date of a U.S. national stage application is the international filing date. See 35 U.S.C. 363.

**Correspondence Address**

■ The address associated with Customer Number: 131890 **OR** ☐ Correspondence address below

| | | | | | |
|---|---|---|---|---|---|
| Name | | | | | |
| Address | | | | | |
| City | | State | | Zip Code | |
| Country | | | | Telephone | |
| Email | | | | | |

| | | | |
|---|---|---|---|
| Signature | /Mary K. Nicholes/ | Date | November 9, 2021 |
| Name (Print/Type) | Mary K. Nicholes | Registration No. (Attorney/Agent) | 56,238 |

# Privacy Act Statement

The **Privacy Act of 1974 (P.L. 93-579)** requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (*i.e.*, GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Doc Code: MES.GIB
Document Description: Certification of Micro Entity Status (Gross Income Basis)

PTO/SB/15A (07-14)

# CERTIFICATION OF MICRO ENTITY STATUS
## (GROSS INCOME BASIS)

| Application Number or Control Number (if applicable): | Patent Number (if applicable): |
|---|---|
| First Named Inventor: LEWIS, Lucinda | Title of Invention: METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY |

The applicant hereby certifies the following—

(1) **SMALL ENTITY REQUIREMENT –** The applicant qualifies as a small entity as defined in 37 CFR 1.27.

(2) **APPLICATION FILING LIMIT –** Neither the applicant nor the inventor nor a joint inventor has been named as the inventor or a joint inventor on more than four previously filed U.S. patent applications, excluding provisional applications and international applications under the Patent Cooperation Treaty (PCT) for which the basic national fee under 37 CFR 1.492(a) was not paid, and also excluding patent applications for which the applicant has assigned all ownership rights, or is obligated to assign all ownership rights, as a result of the applicant's previous employment.

(3) **GROSS INCOME LIMIT ON APPLICANTS AND INVENTORS –** Neither the applicant nor the inventor nor a joint inventor, in the calendar year preceding the calendar year in which the applicable fee is being paid, had a gross income, as defined in section 61(a) of the Internal Revenue Code of 1986 (26 U.S.C. 61(a)), exceeding the "Maximum Qualifying Gross Income" reported on the USPTO Web site at http://www.uspto.gov/patents/law/micro_entity.jsp which is equal to three times the median household income for that preceding calendar year, as most recently reported by the Bureau of the Census.

(4) **GROSS INCOME LIMIT ON PARTIES WITH AN "OWNERSHIP INTEREST" –** Neither the applicant nor the inventor nor a joint inventor has assigned, granted, or conveyed, nor is under an obligation by contract or law to assign, grant, or convey, a license or other ownership interest in the application concerned to an entity that, in the calendar year preceding the calendar year in which the applicable fee is being paid, had a gross income, as defined in section 61(a) of the Internal Revenue Code of 1986, exceeding the "Maximum Qualifying Gross Income" reported on the USPTO Web site at http://www.uspto.gov/patents/law/micro_entity.jsp which is equal to three times the median household income for that preceding calendar year, as most recently reported by the Bureau of the Census.

**SIGNATURE by an authorized party set forth in 37 CFR 1.33(b)**

| | | | | | |
|---|---|---|---|---|---|
| Signature | /Mary K. Nicholes/ | | | | |
| Name | Mary K. Nicholes | | | | |
| Date | 11/09/2021 | Telephone | 973-422-6536 | Registration No. | 56,238 |

☐ There is more than one inventor and I am one of the inventors who are jointly identified as the applicant. The required additional certification form(s) signed by the other joint inventor(s) are included with this form.

## Privacy Act Statement

The **Privacy Act of 1974 (P.L. 93-579)** requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (*i.e.*, GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

| Application Data Sheet 37 CFR 1.76 | Attorney Docket Number | 28108-125 |
|---|---|---|
| | Application Number | |
| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | |

The application data sheet is part of the provisional or nonprovisional application for which it is being submitted. The following form contains the bibliographic data arranged in a format specified by the United States Patent and Trademark Office as outlined in 37 CFR 1.76.
This document may be completed electronically and submitted to the Office in electronic format using the Electronic Filing System (EFS) or the document may be printed and included in a paper filed application.

## Secrecy Order 37 CFR 5.2:

☐ Portions or all of the application associated with this Application Data Sheet may fall under a Secrecy Order pursuant to 37 CFR 5.2 (Paper filers only. Applications that fall under Secrecy Order may not be filed electronically.)

## Inventor Information:

**Inventor** 1 — Remove

**Legal Name**

| Prefix | Given Name | Middle Name | Family Name | Suffix |
|---|---|---|---|---|
| | Lucinda | | Lewis | |

**Residence Information (Select One)** US Residency ● Non US Residency Active US Military Service

| City | Burbank | Country of Residence i | US |
|---|---|---|---|

**Mailing Address of Inventor:**

| | | | |
|---|---|---|---|
| **Address 1** | 1406 Columbus Boulevard | | |
| **Address 2** | | | |
| **City** | Coral Gables | **State/Province** | FL |
| **Postal Code** | 33134-2351 | **Country** i | US |

All Inventors Must Be Listed - Additional Inventor Information blocks may be generated within this form by selecting the **Add** button. Add

## Correspondence Information:

**Enter either Customer Number or complete the Correspondence Information section below.**
**For further information see 37 CFR 1.33(a).**

☐ **An Address is being provided for the correspondence Information of this application.**

| | | | |
|---|---|---|---|
| **Customer Number** | 131890 | | |
| **Email Address** | patents@lowenstein.com | Add Email | Remove Email |

## Application Information:

| | | | |
|---|---|---|---|
| **Title of the Invention** | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | | |
| **Attorney Docket Number** | 28108-125 | **Small Entity Status Claimed** | ☐ |
| **Application Type** | Nonprovisional | | |
| **Subject Matter** | Utility | | |
| **Total Number of Drawing Sheets (if any)** | 7 | **Suggested Figure for Publication (if any)** | |

| Application Data Sheet 37 CFR 1.76 | Attorney Docket Number | 28108-125 |
|---|---|---|
| | Application Number | |

| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY |
|---|---|

## Filing By Reference:

Only complete this section when filing an application by reference under 35 U.S.C. 111(c) and 37 CFR 1.57(a). Do not complete this section if application papers including a specification and any drawings are being filed. Any domestic benefit or foreign priority information must be provided in the appropriate section(s) below (i.e., "Domestic Benefit/National Stage Information" and "Foreign Priority Information").

For the purposes of a filing date under 37 CFR 1.53(b), the description and any drawings of the present application are replaced by this reference to the previously filed application, subject to conditions and requirements of 37 CFR 1.57(a).

| Application number of the previously filed application | Filing date (YYYY-MM-DD) | Intellectual Property Authority or Country |
|---|---|---|
| | | |

## Publication Information:

☐ Request Early Publication (Fee required at time of Request 37 CFR 1.219)

☐ **Request Not to Publish.** I hereby request that the attached application not be published under 35 U.S.C. 122(b) and certify that the invention disclosed in the attached application **has not and will not** be the subject of an application filed in another country, or under a multilateral international agreement, that requires publication at eighteen months after filing.

## Representative Information:

Representative information should be provided for all practitioners having a power of attorney in the application. Providing this information in the Application Data Sheet does not constitute a power of attorney in the application (see 37 CFR 1.32).
Either enter Customer Number or complete the Representative Name section below. If both sections are completed the customer Number will be used for the Representative Information during processing.

| Please Select One: | ● Customer Number | US Patent Practitioner | ◯ Limited Recognition (37 CFR 11.9) |
|---|---|---|---|
| Customer Number | 131890 | | |

## Domestic Benefit/National Stage Information:

This section allows for the applicant to either claim benefit under 35 U.S.C. 119(e), 120, 121, 365(c), or 386(c) or indicate National Stage entry from a PCT application. Providing benefit claim information in the Application Data Sheet constitutes the specific reference required by 35 U.S.C. 119(e) or 120, and 37 CFR 1.78.
When referring to the current application, please leave the "Application Number" field blank.

| Prior Application Status | Pending | | Remove |
|---|---|---|---|
| Application Number | Continuity Type | Prior Application Number | Filing or 371(c) Date (YYYY-MM-DD) |
| | a 371 of international | PCT/US2020/032149 | 2020-05-08 |

| Application Data Sheet 37 CFR 1.76 | Attorney Docket Number | 28108-125 |
|---|---|---|
| | Application Number | |

| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY |
|---|---|

| Prior Application Status | Expired | | Remove |
|---|---|---|---|
| Application Number | Continuity Type | Prior Application Number | Filing or 371(c) Date (YYYY-MM-DD) |
| 62845546 | Claims benefit of provisional | | 2019-05-09 |
| Additional Domestic Benefit/National Stage Data may be generated within this form by selecting the **Add** button. | | | Add |

## Foreign Priority Information:

This section allows for the applicant to claim priority to a foreign application. Providing this information in the application data sheet constitutes the claim for priority as required by 35 U.S.C. 119(b) and 37 CFR 1.55. When priority is claimed to a foreign application that is eligible for retrieval under the priority document exchange program (PDX)[i] the information will be used by the Office to automatically attempt retrieval pursuant to 37 CFR 1.55(i)(1) and (2). Under the PDX program, applicant bears the ultimate responsibility for ensuring that a copy of the foreign application is received by the Office from the participating foreign intellectual property office, or a certified copy of the foreign priority application is filed, within the time period specified in 37 CFR 1.55(g)(1).

| Application Number | Country[i] | Filing Date (YYYY-MM-DD) | Remove<br>Access Code[i] (if applicable) |
|---|---|---|---|
| | | | |
| Additional Foreign Priority Data may be generated within this form by selecting the **Add** button. | | | Add |

## Statement under 37 CFR 1.55 or 1.78 for AIA (First Inventor to File) Transition Applications

☐ This application (1) claims priority to or the benefit of an application filed before March 16, 2013 and (2) also contains, or contained at any time, a claim to a claimed invention that has an effective filing date on or after March 16, 2013.
NOTE: By providing this statement under 37 CFR 1.55 or 1.78, this application, with a filing date on or after March 16, 2013, will be examined under the first inventor to file provisions of the AIA.

| Application Data Sheet 37 CFR 1.76 | Attorney Docket Number | 28108-125 |
|---|---|---|
| | Application Number | |
| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | |

## Authorization or Opt-Out of Authorization to Permit Access:

When this Application Data Sheet is properly signed and filed with the application, applicant has provided written authority to permit a participating foreign intellectual property (IP) office access to the instant application-as-filed (see paragraph A in subsection 1 below) and the European Patent Office (EPO) access to any search results from the instant application (see paragraph B in subsection 1 below).

Should applicant choose not to provide an authorization identified in subsection 1 below, applicant **must opt-out** of the authorization by checking the corresponding box A or B or both in subsection 2 below.

**NOTE**: This section of the Application Data Sheet is **ONLY** reviewed and processed with the **INITIAL** filing of an application. After the initial filing of an application, an Application Data Sheet cannot be used to provide or rescind authorization for access by a foreign IP office(s). Instead, Form PTO/SB/39 or PTO/SB/69 must be used as appropriate.

**1. Authorization to Permit Access by a Foreign Intellectual Property Office(s)**

**A. Priority Document Exchange (PDX)** - Unless box A in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the European Patent Office (EPO), the Japan Patent Office (JPO), the Korean Intellectual Property Office (KIPO), the State Intellectual Property Office of the People's Republic of China (SIPO), the World Intellectual Property Organization (WIPO), and any other foreign intellectual property office participating with the USPTO in a bilateral or multilateral priority document exchange agreement in which a foreign application claiming priority to the instant patent application is filed, access to: (1) the instant patent application-as-filed and its related bibliographic data, (2) any foreign or domestic application to which priority or benefit is claimed by the instant application and its related bibliographic data, and (3) the date of filing of this Authorization. See 37 CFR 1.14(h)(1).

**B. Search Results from U.S. Application to EPO** - Unless box B in subsection 2 (opt-out of authorization) is checked, the undersigned hereby **grants the USPTO authority** to provide the EPO access to the bibliographic data and search results from the instant patent application when a European patent application claiming priority to the instant patent application is filed. See 37 CFR 1.14(h)(2).

The applicant is reminded that the EPO's Rule 141(1) EPC (European Patent Convention) requires applicants to submit a copy of search results from the instant application without delay in a European patent application that claims priority to the instant application.

**2. Opt-Out of Authorizations to Permit Access by a Foreign Intellectual Property Office(s)**

☐ A. Applicant **DOES NOT** authorize the USPTO to permit a participating foreign IP office access to the instant application-as-filed. If this box is checked, the USPTO will not be providing a participating foreign IP office with any documents and information identified in subsection 1A above.

☐ B. Applicant **DOES NOT** authorize the USPTO to transmit to the EPO any search results from the instant patent application. If this box is checked, the USPTO will not be providing the EPO with search results from the instant application.

**NOTE:** Once the application has published or is otherwise publicly available, the USPTO may provide access to the application in accordance with 37 CFR 1.14.

| Application Data Sheet 37 CFR 1.76 | Attorney Docket Number | 28108-125 |
|---|---|---|
| | Application Number | |
| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | |

## Applicant Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

**Applicant** 1 Remove

If the applicant is the inventor (or the remaining joint inventor or inventors under 37 CFR 1.45), this section should not be completed. The information to be provided in this section is the name and address of the legal representative who is the applicant under 37 CFR 1.43; or the name and address of the assignee, person to whom the inventor is under an obligation to assign the invention, or person who otherwise shows sufficient proprietary interest in the matter who is the applicant under 37 CFR 1.46. If the applicant is an applicant under 37 CFR 1.46 (assignee, person to whom the inventor is obligated to assign, or person who otherwise shows sufficient proprietary interest) together with one or more joint inventors, then the joint inventor or inventors who are also the applicant should be identified in this section. Clear

| | | |
|---|---|---|
| ● Assignee | Legal Representative under 35 U.S.C. 117 | Joint Inventor |
| Person to whom the inventor is obligated to assign. | Person who shows sufficient proprietary interest | |

If applicant is the legal representative, indicate the authority to file the patent application, the inventor is:

| | | | |
|---|---|---|---|
| Name of the Deceased or Legally Incapacitated Inventor: | | | |
| If the Applicant is an Organization check here. | ☒ | | |
| Organization Name | Automobilia II, LLC | | |
| **Mailing Address Information For Applicant:** | | | |
| **Address 1** | 1406 Columbus Boulevard | | |
| Address 2 | | | |
| **City** | Coral Gables | **State/Province** | FL |
| **Country** | US | Postal Code | 33134-2351 |
| Phone Number | | Fax Number | |
| Email Address | | | |

Additional Applicant Data may be generated within this form by selecting the Add button. Add

## Assignee Information including Non-Applicant Assignee Information:

Providing assignment information in this section does not substitute for compliance with any requirement of part 3 of Title 37 of CFR to have an assignment recorded by the Office.

| **Application Data Sheet 37 CFR 1.76** | Attorney Docket Number | 28108-125 |
|---|---|---|
| | Application Number | |

| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY |
|---|---|

**Assignee 1**

Complete this section if assignee information, including non-applicant assignee information, is desired to be included on the patent application publication. An assignee-applicant identified in the "Applicant Information" section will appear on the patent application publication as an applicant. For an assignee-applicant, complete this section only if identification as an assignee is also desired on the patent application publication.

Remove

If the Assignee or Non-Applicant Assignee is an Organization check here. ☐

| Prefix | **Given Name** | Middle Name | **Family Name** | Suffix |
|---|---|---|---|---|
| | | | | |

**Mailing Address Information For Assignee including Non-Applicant Assignee:**

| | | | |
|---|---|---|---|
| **Address 1** | | | |
| Address 2 | | | |
| **City** | | **State/Province** | |
| **Country** i | | Postal Code | |
| Phone Number | | Fax Number | |
| Email Address | | | |

Additional Assignee or Non-Applicant Assignee Data may be generated within this form by selecting the Add button. Add

## Signature:

Remove

**NOTE:** This Application Data Sheet must be signed in accordance with 37 CFR 1.33(b). **However, if this Application Data Sheet is submitted with the INITIAL filing of the application and either box A or B is not checked in subsection 2 of the "Authorization or Opt-Out of Authorization to Permit Access" section, then this form must also be signed in accordance with 37 CFR 1.14(c).**

This Application Data Sheet **must** be signed by a patent practitioner if one or more of the applicants is a **juristic entity** (e.g., corporation or association). If the applicant is two or more joint inventors, this form must be signed by a patent practitioner, **all** joint inventors who are the applicant, or one or more joint inventor-applicants who have been given power of attorney (e.g., see USPTO Form PTO/AIA/81) on behalf of **all** joint inventor-applicants.

See 37 CFR 1.4(d) for the manner of making signatures and certifications.

| **Signature** | /Mary K. Nicholes/ | | | Date (YYYY-MM-DD) | 2021-11-09 |
|---|---|---|---|---|---|
| First Name | Mary | Last Name | Nicholes | Registration Number | 56,238 |

Additional Signature may be generated within this form by selecting the Add button. Add

| **Application Data Sheet 37 CFR 1.76** | Attorney Docket Number | 28108-125 |
|---|---|---|
| | Application Number | |

| Title of Invention | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY |
|---|---|

# Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1 The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C. 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether the Freedom of Information Act requires disclosure of these records.

2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.

3 A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.

4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).

5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent CooperationTreaty.

6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).

7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.

8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspections or an issued patent.

9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

## BACKGROUND

**[0001]** Cultural knowledge about one of mankind's most significant inventions, the automobile, is deeply fragmented and not easily searchable. Today, a user may see an object, such as a vehicle, driving down the road and wonder what it is, or an autonomous vehicle driving on a freeway may have a need to identify objects such as vehicles in its vicinity to learn information (e.g., vehicle make, model, year, stopping distance, etc.) and any equipment on board (e.g., communication devices, computer systems, etc.). Such information is currently unauthenticated, fragmented, dispersed, and not readily available in searchable form.

**[0002]** As society becomes increasingly digitalized, benefits have emerged in augmenting human intelligence with artificial intelligence (AI) to answer questions with curated knowledge. The market of automobile information is segmented and incoherent, with bits of information gathered by each segment without any relationship to each other and, thus, the information is not readily searchable. Each market segment has different needs, e.g., a user may have an instant need to identify certain vehicles, auto-related architecture and/or cultural artifacts out of interest whereas an autonomous vehicle may have a need to identify other vehicles in its vicinity and access their capabilities.

**[0003]** In 2018, the automotive advertising segment exceeded $38 trillion, exclusive of advertising for travel and food and automotive repair. The automotive advertising sector is the second largest advertising sector in the overall advertising marketplace. The largest customers are advertisers for new car buyers interested in the heritage of an automotive brand, the collectible car enthusiast market, the automotive parts market, insurance, travel, media archives and libraries with unidentified assets, and consumers with an unidentified photo album showing family vehicles. Additional commercial opportunities reside with government, security, law enforcement, and the entertainment industry.

**[0004]** Existing platforms are unable to sufficiently identify vehicles. These platforms can only make inferences from unauthenticated data. Anyone looking to authenticate a vehicle using these platforms can spend hours and still be uncertain of the exact make, model and year. These platforms have no verifiable source of automotive data and little hope of authoritatively identifying cars. It is difficult and frustrating to credibly authenticate automobiles by searching the Internet or asking vehicle owners; significant knowledge is fading from human memory.

**[0005]** There is a need for methods, systems and computer program products that allow a user to identify and search objects such as vehicles, while simultaneously building provenance and preserving knowledge around such objects and their cultural history, to instantly and properly identify objects by training artificial intelligence tools, to preserve cultural history, to promote historic places influenced by the vehicle's evolution, to establish automotive provenance based on first-hand historical records, to celebrate and educate users about objects (e.g., vehicles), to learn from history and help shape mobility's evolution, to explore human-computer interactions and to address the automotive advertising market. The systems, methods and computer program products described herein can quickly and accurately identify objects such as vehicles using authenticated data and provide an armature for social data to accrue around any object, building provenance around a subject that has lacked the tools to easily authenticate knowledge, e.g., cars and their impact on culture.

**BRIEF DESCRIPTION OF THE DRAWINGS**

**[0006]** The present disclosure is illustrated by way of examples, and not by way of limitation, and may be more fully understood with references to the following detailed description when considered in connection with the figures, in which:

**[0007] FIG. 1** schematically illustrates an example system for image processing and data analysis, in accordance with one or more aspects of the present disclosure.

**[0008] FIG. 2** schematically illustrates an example structure of a convolutional neural network (CNN) that may be employed to process data input to an example system for media classification and identification, in accordance with one or more aspects of the present disclosure.

**[0009] FIG. 3** depicts a flow diagram of one illustrative example of a method 300 of data processing and data analysis, in accordance with one or more aspects of the present disclosure.

**[0010] FIG. 4** depicts a flow diagram of one illustrative example of a method 360 of displaying via augmented reality a result from an example system for media classification and identification.

**[0011] FIG. 5** depicts a diagram of a system for implementing the methods and systems described herein.

**[0012] FIG. 6** depicts a diagram of a computational cluster system that may be employed in an example system for media classification and identification, in accordance with one or more aspects of the present disclosure.

**[0013] FIG. 7** depicts a diagram of an illustrative example of a computing device implementing the systems and methods described herein.

## DETAILED DESCRIPTION

**[0014]** Described herein are methods, systems and computer program products for media classification and identification and for displaying results on a display (e.g., a cell phone, a monitor, an augmented reality apparatus, a mixed reality apparatus). While the systems and methods are described with respect to vehicles, vehicle artifacts and geographical locations, the systems and methods are broadly applicable to any objects. In example implementations, the systems and methods may relate to buildings (e.g., architecture), clothing, bridges, tools, highways, mountains, parks, rivers, cities, cars converted to homes and so on. Thus, the systems and methods described herein may be applied to a wide variety of physical objects that may involve various combinations of multiple imaging and/or other image capturing mechanisms.

**[0015]** The present disclosure overcomes the above-noted and other deficiencies by providing systems and methods for image processing and data analysis that may be utilized for identifying, classifying, researching and analyzing objects including, but not limited to vehicles, vehicle parts, vehicle artifacts, cultural artifacts, geographical locations, etc. To identify all of the objects in a photo, alone or in combination with a geographical location and/or a cultural heritage object, and then to associate a narrative with them represents a unique challenge. The converse of this too – the identification of historic places and objects (e.g. Statue of Liberty) alone, or in combination with vehicles— forms a broad descriptive visual narrative that illustrates innovative mapping from natural language processing (NLP) to multi-label image classification and identification.

**[0016]** In example implementations, a repository of photos, videos, keywords and captions of automobiles of proven provenance, with user narratives and comments, can be used to train a unique AI pipeline to map the information to a target space for image classification. For example, given an uploaded user image, the AI models may create the most appropriate summary of the relevant sections of the asset, and perform a multi-labeled classification of the image into the appropriate model of, for example, car manufacturer and year. In the case of an image containing multiple cars and cultural artifacts (e.g., glove compartments, spokes, steering wheels, vehicle lifts, etc.), there may be the additional task of establishing bounding-boxes around each of the recognized objects, and creating summary text appropriate to the image as a whole.

**[0017]** Likewise, the converse problem of taking a vehicle description, i.e., “Show me Prototypes”, and enriching it with AI assisted discovery into a proprietary database of high quality copyrighted images, represents a journey where the feature-vectors comprise the NLP embeddings of the narratives. The target space may be comprised of clusters of automotive

images that share attributes; for example, the query may map to a cluster of experimental cars from a particular decade. This may involve a single machine learning (ML) pipeline where RNN (LSTM/GRU) and BERT-derived attention models interact with CNN-architectures for image classification and Siamese Neural Networks (SNNs) for correct identifications. A collaborative user verification process involving crowd wisdom can be used to improve the accuracy of image-augmentation such that users can point out errors and suggest corrections. Should certain annotations be erroneous and users mark them so, such data will feed into the next round of neural architecture training.

**[0018]** In certain implementations, the systems and methods described herein may perform pixel-level analysis of images (and/or videos) in order to yield images, videos and/or virtual environments (e.g., augmented reality, mixed reality, virtual reality etc.) of vehicles, vehicle artifacts (e.g., images of vehicle tools, feature elements such as goggles, tachometers, wheel spokes, gas cans, etc.) and/or geographical locations. The systems and methods described herein may further determine whether images or videos input to the media processing system contain features that match one or more features of an image, video and/or geographical location of stored in a memory. In implementations, the systems and methods produce a result that comprises the closest matching data (e.g., having the highest probability score based on a cross-entropy function) identified in the training data set and/or database repository. The result may include an image of a vehicle together with text information about the vehicle such as a history, make, model, year, etc. The systems and methods may additionally yield historical information about one or more vehicle and/or geographical location and such information may be displayed in a virtual environment. In certain implementations, the systems and methods as described herein may also be implemented in an autonomous vehicle to capture images and/or video of surrounding vehicles on the road and to produce a result indicating the size, make, model and on-board equipment of the surrounding vehicles. In an example implementation, an autonomous vehicle incorporating the systems and methods described herein can be trained for "platooning." An example of "platooning" is where a vehicle operating in a self-driving or semi-autonomous mode, analyzes other vehicles in its vicinity to determine, for example, which vehicles may be capable of vehicle-to-vehicle (V2V) communication, other equipment on board, the estimated stopping distance of each vehicle and surrounding environmental objects such as children, balls, bicycles, tumbleweeds, etc. The autonomous vehicle may then communicate with the V2V vehicles to maintain a safe speed and distance to those vehicles, that is, the vehicles may move harmoniously together and may stop together at a traffic lights. Any vehicles in the vicinity that are not capable of V2V may be accounted for as an unknown variable. In other

implementations, platooning may involve recognition by the autonomous vehicle of structures that are capable of communicating with the vehicle in a vehicle-to-infrastructure (V2I) configuration. If the infrastructure is equipped with methods and systems as described herein, it may time or adjust the traffic lights to enhance platooning of V2I vehicles taking into consideration the variables of vehicles that are not equipped for V2I communication.

**[0019]** The systems and methods described herein have the benefit of being trained using a proprietary database comprising high-quality, digital copyrighted images of vehicles, vehicle artifacts and/or geographic sites (e.g., historical sites, cultural sites), such that the accuracy of the results produced by the disclosed systems and methods is improved over known methods of researching and analyzing vehicles, vehicle artifacts and/or geographical locations. The database may further include videos, embedded metadata and text. The database may itself be copyrighted. Because the database and data assets (e.g., images, videos, text, etc.) are themselves copyrighted, they form a body of authenticated data on which the neural networks can be trained.

**[0020]** The systems and methods described herein utilize a convolutional neural network (CNN) or a combination of both a CNN and a recurrent neural network (RNN), which form a part of a media processing system. The CNN may process one or more of image data (e.g., containing images, for example, of vehicles, vehicle artifacts, landscapes, etc.), video data (e.g., videos of vehicles, videos of historical sites, etc.), geolocation data (e.g., from a global positioning system) or intake data (e.g., text queries entered via a user interface, voice queries, natural language queries, etc.) to perform classification and with respect to vehicle information, vehicle artifact information, geographical location, etc. and/or to yield the probability of one or more image, video, geolocation and/or intake query matching significant features associated with a vehicle, a vehicle artifact and/or geographical location. The returned images, videos and/or virtual environment may be annotated and/or layered (e.g., overlaid, underlaid) with historical, design, mechanical, etc. information in the form of, for example, text, audio and video.

**[0021]** The RNN processes unstructured data, for example, natural language search queries and/or voice inputs to provide natural language processing (NLP). The unstructured data is transformed into structured data, which is fed to the CNN and processed as described above. In example implementations, the neural network architecture creates a hybridization of natural language processing (RNN) with image classification and identification techniques (CNN) for the purposes of preserving and accumulating data around a key subject area of historical and cultural interest.

**[0022]** A CNN is a computational model based on a multi-staged algorithm that applies a set of pre-defined functional transformations to one or more input (e.g., image pixels) and then utilizes the transformed data to perform, for example, classification, identification, image recognition, pattern recognition, etc. A CNN may be implemented as a feed-forward neural network (FFNN) in which the connectivity pattern between its neurons is inspired by the organization of the animal visual cortex. Individual cortical neurons respond to stimuli in a restricted region of space known as the receptive field. The receptive fields of different neurons partially overlap such that they tile the visual field. The response of an individual neuron to stimuli within its receptive field can be approximated mathematically by a convolution operation. In addition to image processing, the CNN may be used for other input types such as text, audio and video. In implementations, images may be input to a media processing system as described herein and the CNN processes the data. For example, if a user inputs a picture of a Ford Thunderbird automobile, the media processing system may output an image of a Ford Thunderbird together with the make, model, year, history and any known contextual information surrounding the photo and background.

**[0023]** In an illustrative example, a CNN may include multiple layers of various types, including convolution layers, non-linear layers (e.g., implemented by rectified linear units (ReLUs)), pooling layers, and classification (fully-connected) layers. A convolution layer may extract features from the input image by applying one or more learnable pixel-level filters to the input image. In an illustrative example, a pixel-level filter may be represented by a matrix of integer values, which is convolved across the dimensions of the input image to compute dot products between the entries of the filter and the input image at each spatial position, thus producing a feature map that represents the responses of the filter at every spatial position of the input image. The convolution filters are defined at the network training stage based on the training dataset to detect patterns and regions that are indicative of the presence of significant features within the input image.

**[0024]** A non-linear operation may be applied to the feature map produced by the convolution layer. In an illustrative example, the non-linear operation may be represented by a rectified linear unit (ReLU) which replaces with zeros all negative pixel values in the feature map. In various other implementations, the non-linear operation may be represented by a hyperbolic tangent function, a sigmoid function, or by other suitable non-linear function.

**[0025]** A pooling layer may perform subsampling to produce a reduced resolution feature map while retaining the most relevant information. The subsampling may involve averaging and/or determining maximum value of groups of pixels.

**[0026]** In certain implementations, convolution, non-linear, and pooling layers may be applied to the input image multiple times prior to the results being transmitted to a classification (fully-connected) layer. Together these layers extract the useful features from the input image, introduce non-linearity, and reduce image resolution while making the features less sensitive to scaling, distortions, and small transformations of the input image.

**[0027]** The output from the convolutional and pooling layers represent high-level features of the input image. The purpose of the classification layer is to use these features for classifying the input image into various classes. In an illustrative example, the classification layer may be represented by an artificial neural network that comprises multiple neurons. Each neuron receives its input from other neurons or from an external source and produces an output by applying an activation function to the sum of weighted inputs and a trainable bias value. A neural network may include multiple neurons arranged in layers, including the input layer, one or more hidden layers, and the output layer. Neurons from adjacent layers are connected by weighted edges. The term "fully connected" implies that every neuron in the previous layer is connected to every neuron on the next layer.

**[0028]** The edge weights are defined at the network training stage based on the training dataset. In an illustrative example, all of the edge weights are initialized to random values. For every input in the training dataset, the neural network is activated. The observed output of the neural network is compared with the desired output specified by the training data set, and the error is propagated back to the previous layers of the neural network, in which the weights are adjusted accordingly. This process is repeated until the output error is below a predetermined threshold.

**[0029]** The CNN may be implemented in a SNN configuration. A SNN configuration contains two or more identical subnetwork components. In implementations, not only is the architecture of the subnetworks identical, but the weights are shared among them as well. SNNs learn useful data descriptors, which may be used to compare the inputs (e.g., image data, video data, input data, geolocation data, etc.) of the subnetworks. For example, the inputs may be image data with CNNs as subnetworks.

**[0030]** The CNN may be implemented in a Generative Adversarial Network (GAN), which refers to two networks working together. A GAN can include any two networks (e.g., a combination of FFNNs and CNNs), with one tasked to generate content and the other tasked to judge content. The discriminating network receives either training data or generated content from the generative network. The ability of the discriminating network to correctly predict the data source is then used as part of the error for the generating network. This creates a form of competition where the discriminator gets better at distinguishing real data from generated data

and the generator learns to become less predictable to the discriminator. Even quite complex noise-like patterns can become predictable, but generated content similar in features to the input data is harder to learn to distinguish. The dynamics between the two networks need to be balanced; if prediction or generation becomes too good compared to the other, the GAN will not converge as there is intrinsic divergence.

**[0031]** A RNN may be described as a FFNN having connections between passes and through time. A RNN receives not just the current input it is fed, but also what it has perceived previously in time. In a RNN, neurons can be fed information not only from a previous layer, but from a previous pass. A string of text or picture can be fed one pixel or character at a time, so that the time dependent weights can be used for what came before in the sequence, not actually from what happened a specific time (e.g., x seconds) before. The RNN may be implemented as a Long Short-Term Memory (LFTM), which helps preserve the error, back-propagating it through layers and time. A LSTM includes information outside of the normal flow of the RNN in a gated cell. Information can be written to, stored in, or read from a cell, similar to data in a computer's memory. The cell can make decisions about when to allow reads, writes and erasures, and what to store via gates that open and close. These gates are analog, implemented with element-wise multiplication by sigmoids (i.e., all in the range of 0-1).

**[0032]** The RNN may be implemented with Bidirectional Encoder Representations from Transformers (BERT) to perform NLP tasks including *inter alia* question answering and natural language inference. BERT, which uses a Transformer-based language model, is a language representation model that provides accuracy for NLP tasks. A Transformer, in an encoding step, can use learned word embedding to convert words, in one-hot-vector form, into word embedding vectors; for each word-embedding vector, there is one output vector. BERT and its variants and Transformers, alone or in any combination with RNNs, are suitable for NLP tasks according to implementations herein.

**[0033]** Natural Language Processing (NLP) is the ability of a computer program to process and generate human language as spoken and/or written. In implementations, one or more recurrent neural network (RNN) is constructed to perform NLP (e.g., including text classification and text generation). In implementations, a neural network has layers, where each layer includes either input, hidden or output cells in parallel. In general two adjacent layers are fully connected (i.e., every neuron forms one layer to every neuron to another layer). For example, the network can have two input cells and one output cell, which can be used to model logic gates.

**[0034]** Augmented reality, which refers to a combination of a virtual environment and virtual reality, combines real-world images and virtual-world images such as computer graphic

images. Augmented reality is a semi-digital experience. In implementations of augmented reality, an image capturing device (e.g., a camera, a phone, a video recorder, etc.) that receives real images and a display device (e.g., a head mounted display that can display both real images and virtual images) are used together. Using augmented reality, a vehicle can be superimposed over a geographical location, for example, that can be associated with a particular date, time and/or weather condition. Lines can be drawn over an image of a vehicle to identify certain features and or parts, which may or may not be associated with a particular design type, time of history and/or cultural trend.

**[0035]** Virtual reality is a fully digital experience that creates an immersive environment where humans and computers can effectively interact and communicate by enhancing human-computer conversation skills using a variety of input and output techniques. Such techniques include the use of, for example, head-mounted displays, data gloves, or motion capture systems. These techniques receive data regarding variations in the position of a user by monitoring head, hand or other movements (e.g., position, directions, etc.), and transmit the data to a computer, which simulates (e.g., in a 3D coordinate space) the size and depth of an object within the viewing angle of the user.

**[0036]** Mixed reality refers to the merging of the real world with a virtual world to create a new environment where physical and digital objects interact with one another in real-time. In mixed reality, a real image can be captured using an image capturing device (e.g., a camera) and the direction the user faces within the environment is based on the captured real image. The relationship between the user's position and the position of a predetermined object is determined, and data obtained as a result of the calculation is displayed in a virtual space such that the data is laid over the captured real world image. Mixed reality is typically implemented using an image capturing device together with a display device.

**[0037] FIG. 1** schematically illustrates an example system 100 for image processing and data analysis, in accordance with one or more aspects of the present disclosure. As schematically illustrated by **FIG. 1**, the CNN 120 and optionally an RNN 122 together with a processing device 124 and a memory 126 form an media processing system 101. The media processing system 101 may be employed to process image data 110, video data 112, geolocation data 114 and input data 116 to produce an image classification result 130 and/or a virtual display result 132. The image data 110 may include one or more digital images, for example, captured by a camera or scanned, that may be stored in a memory. The video data 112 may include one or more digital videos, for example, captured by an audiovisual recording device or a dubbing device, that may be stored in a memory. The geolocation data 114 may include longitude,

latitude, country, region, historical or cultural place (e.g., Brooklyn Bridge), city, postal/zip code, time zone, way point, cell tower signal, etc. information from, for example, a global positioning system (GPS), entry to a user interface and/or other navigation system. The input data 116 may include structured data such as keyword search query input via a user interface and/or unstructured data input via the user interface. The unstructured data may include written or spoken natural language.

**[0038]** The CNN 120 may be employed to process the image data 110, the video data 112, the geolocation data 114 and the structured input data 116 to produce an image classification result 130 and/or a virtual display result 132, for example, with respect to vehicle information (e.g., a make, model, year, convertible sedan, sports utility vehicle or SUV, prototype, etc. ), vehicle artifacts (e.g., tools, a steering wheel, a lift, spokes, ets.) and/or a geographical location (e.g., a cultural site, a historical site, a landmark, etc.). The RNN 122 may process the unstructured data of the input data 116 (i.e., natural language processing) to produce structured data. The structured data may be fed from the RNN 122 to the CNN 120 for processing as described herein. In an illustrative example, the CNN 120 may correlate the image data 110, video data 112, geolocation data 114, input data 116 and structured data from the RNN 112 with images and data from a database (e.g., a propriety database of high quality copyrighted images and other works) in order to yield the probabilities of, for example, one or more image containing matching significant image features associated with a vehicle and/or a geographical location. The media processing system 101 may also return as part of the image classification result 130 historical, mechanical, cultural and other information with respect to the vehicle and/or geographical location.

**[0039]** The CNN 120 and the RNN 122 may be pre-trained using a comprehensive and precise training data set 140 that comprises *inter alia* non-published data, published data, images, videos, text (e.g., stories, news articles about various (e.g., thousands) vehicles, memoirs, out-of-print books, etc.) and/or geographical locations. For every vehicle or geographical location, the training data set may include multiple pixel-level, optionally annotated, vehicle images 142 and geographical images 146 and also may include related text 144 (e.g., histories, stories, descriptions, books, articles, drawings, sketches, etc.). The training data set 140 according to implementations herein may be a unique, comprehensive and proprietary body comprising copyrighted images, videos and text (e.g., history from the 20$^{th}$ Century surrounding vehicles, books off copyright, personal memoirs about vehicles, stories about racing, metals used, the brass era), built over many decades, that contains approximately 500,000 assets and verifies provenance through its records of timely copyright registrations at the Library of Congress. The

copyrighted works (i.e., authenticated data) identify *inter alia* automobiles and their impact on world culture including the time, place, historical context and significant background architecture captured in the media assets. Initially, the training data set 140 may be comprised in a copyrighted database where all of the assets contained within the training data set are copyrighted and thus authenticated. In further implementations, the training data set 140 may expand to include data input by users and further data assets that are not copyright registered, where such additional assets may be authenticated by other means whether scholarly (e.g., citations and research) or by using SNNs according to embodiments herein. The secondary twin will run a regression against the CNN. While the training data set 140 is comprehensive and precise when used in the systems and methods described herein, it will grow and evolve with more provenance authenticated data.

**[0040]** In certain implementations, training of the CNN 120 may involve activating the CNN 120 for every set of input images in the training dataset. The observed output (e.g., an image produced by the CNN 120) is compared with the desired output (e.g., the expected image) specified by the training data set, the error is calculated, and parameters of the CNN 120 are adjusted. This process is repeated until the output error is below a predetermined threshold.

**[0041]** In certain implementations, training of the RNN 122 may involve activating the RNN 120 for every set of unstructured data inputs in the training dataset. The observed output (e.g., a structured query produced by the CNN 120) is compared with the desired output (e.g., the expected query) specified by the training data set, the error may be calculated, and the parameters of the RNN 122 are adjusted accordingly. In implementations, this process may be repeated until the output error is below a predetermined threshold. In implementations using the RNN and deep learning models, the media processing system 101 may function and draw inferences from cross-related data.

**[0042]** The media processing system 101 may produce a virtual display result 132 in a virtual reality, augmented reality and/or mixed reality environment. In one illustrative example, one or more images, videos, descriptions, audio recordings, etc. may be layered onto an image captured by an image capture device (e.g., a camera, video recorder, etc.) and presented on a display. For example, a user may employ an image capture device (e.g., a cell phone) to capture an image in real time and the media processing system 101 may overlay or underlay images, videos and text onto the captured image as being viewed in an output device (e.g., a head-mounted display).

**[0043]** In one illustrative example, the CNN may be trained to identify automobiles. The media processing system 101 may process data 110, 112, 114, 116 to identify automobiles and preserve vehicle history by allowing a user to query the media processing system 101 to learn facts and

view photos of a specific vehicle. In implementations, the media processing system 101 may provide and provoke the curation of historical information surrounding the returned images (e.g., as a part of the image classification result 130). Multiple query types may be supported including, for example, photo uploads (CNN) and voice inputs (RNN) to query the models.

**[0044]** **FIG. 2** schematically illustrates an example structure of a CNN 120 that may be employed to process image data 110, video data 112, geolocation data 114 and input data 116 in order to produce an image classification result 130 and/or a virtual display result 132, in accordance with one or more aspects of the present disclosure. In certain implementations, acquired images may be pre-processed, e.g., by cropping, which may be performed in order to remove certain irrelevant parts of each frame. In an illustrative example, images having the resolution of 1024 x 1024 pixels may be cropped to remove 100-pixel wide image margins from each side of the rectangular image. In another illustrative example, a car may be outlined and isolated from noisy, non-contributory background elements.

**[0045]** As schematically illustrated by **FIG. 2**, the CNN 120 may include a first convolution layer 210A that receives image data 110 containing one or more images. The first convolution layer 210A is followed by squeeze layers 220A and 220B and a pooling layer 230, which is in turn followed by fully-connected layer 240 and a second convolution layer 210B. The second convolution layer 210B outputs one or more image 260 corresponding to the one or more input image of the image data 110 and may further produce the loss value 250 reflecting the difference between the produced data and the training data set. In certain implementations, the loss value may be determined empirically or set at a pre-defined value (e.g., 0.1).

**[0046]** In certain implementations, the loss value is determined as follows:

$$loss = \sum(x - y)^2(\frac{1}{2} + \max(x, y)),$$

where $x$ is the pixel value produced by the second convolution layer 210B and y is the value of the corresponding output image pixel.

**[0047]** Each convolution layer 210A, 210B may extract features from a sequence of input images from the input data 110, by applying one or more learnable pixel-level filters to a three-dimensional matrix representing the sequence of input images. The pixel-level filter may be represented by a matrix of integer values, which is convolved across the dimensions of the input image to compute dot products between the entries of the filter and the input image at each spatial position, to produce a feature map representing the responses of the first convolution layer 210A at every spatial position of the input image. In an illustrative example, the first convolution layer 210A may include 10 filters having the dimensions of 2 x 2 x 2. The second

convolution layer 210B may merge all the values produced by previous layers in order produce a matrix representing a plurality of image pixels.

**[0048] FIG. 3** depicts a flow diagram of one illustrative example of a method 300 of classifying and identifying input data, in accordance with one or more aspects of the present disclosure. Method 300 and/or each of its individual functions, routines, subroutines, or operations may be performed by one or more processors of the computer system (*e.g.*, system 100 and/or processing device 124 of **FIG. 1**) executing the method. In certain implementations, method 300 may be performed by a single processing thread. Alternatively, method 300 may be performed by two or more processing threads, each thread executing one or more individual functions, routines, subroutines, or operations of the method. In an illustrative example, the processing threads implementing method 300 may be synchronized (*e.g.*, using semaphores, critical sections, and/or other thread synchronization mechanisms). Alternatively, the processing threads implementing method 300 may be executed asynchronously with respect to each other.

**[0049]** At block 310, the processing device performing the method may train a CNN using authenticated data and a taxonomy. In implementations, the authenticated data may include copyright registered works of authorship including, but not limited to, copyrighted images, videos, text, stories, sketches, etc. The authenticated data may be stored in a database and the database itself may be copyright registered. The taxonomy may be used to classify and identify the data assets.

**[0050]** At block 320, the processing device may receive a query comprising input data. The input data can include, but is not limited to, image data, video data, intake data and/or geolocation data according to embodiments herein. In implementations, the intake data may be in the form of a keyword or string of text or may be in the form of unstructured data such as natural language either typed or spoken. The method may further include training an RNN to process the unstructured data of the intake data to form structured data suitable for processing by the CNN. The CNN may then process the structured data.

**[0051]** At block 330, the processing device may classify, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy. During the classification, the CNN may match features of the input data to one or more features of the authenticated data and/or elements of the taxonomy. For example, if the input data comprises an image, the CNN may scan the pixels of the image, identify features and then match the features with the closest matching features in the authenticated data and/or as classified in the taxonomy.

**[0052]** At block 340, the processing device may generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match

to the input data. For example, if five features have probabilities of 80%, 82%, 90%, 95% and 99% of matching five assets of the authenticated data, respectively, then the returned result may include only images with features having a 90% or greater probability of matching the input data.

**[0053]** At block 350, the processing device may display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content and/or mixed reality content. In implementations, the result may be layered with information. For example, a displayed image may be annotated with text, video and/or historical information about an object in the image.

**[0054]** In another illustrative example of a method 300 of classifying and identifying input data, in accordance with one or more aspects of the present disclosure. a processing device performing the method may process a training data set comprising a plurality of input images, in order to determine one or more parameters of a CNN to be employed for processing a plurality of images of one or more vehicle and/or geographical location. In various illustrative examples, the parameters of the CNN may include the convolution filter values and/or the edge weights of the fully-connected layer. In an illustrative example, the plurality of input images comprises one or more vehicle image. The one or more vehicle image may illustrate a vehicle alone or in combination with a geographical location (e.g., a Ford Model T on Route 66).

**[0055]** The processing device performing the method optionally may process a training data set comprising unstructured data in order to determine one or more parameters of a RNN to be employed for processing unstructured data input to the media processing system 101 in the form of natural language queries and voice queries to produce structured data for the CNN. In various illustrative examples, the RNN is trained to perform natural language processing using, for example, unstructured written and/or voice inputs.

**[0056]** The media processing system 101 may receive one or more of: a) image data including at least one input image (e.g., of a vehicle and/or a geographical location), b) video data including at least one input video (e.g., of a vehicle and/or a geographical location), c) input data including at least one of a keyword, a search query and unstructured data (e.g., relating to a vehicle and/or a geographical location), and d) geographical location data including a location of a device. In an illustrative example, the media processing system 101 may receive an image of an automobile alone, or together with a voice request saying “show me the artistic design features of this car.”

**[0057]** The processing device performing the method optionally may process, by the RNN of the media processing system 101, any unstructured data of the input data that is received. The RNN outputs structured data that is fed to the CNN for processing. In the foregoing illustrative

example, the RNN may performing natural language processing of the voice request saying "show me the artistic design features of this car."

**[0058]** The processing device performing the method may process by the CNN of the media processing system 101, one or more of: i) the image data 110 to classify at least one input image (e.g., with respect to a vehicle information and/or a geographical location of a vehicle), ii) the video data 112 to classify at least one video (e.g., with respect to a vehicle information and/or a geographical location of a vehicle), iii) the structured input data 116 to classify at least one of a keyword or search query, iv) the structured data from the RNN (330), and v) the geographical location data 114 to produce one or more image, video and/or virtual display, as described in more herein. The probability of the image data, video data, geographical location data, input data and RNN data comprising the significant image features may be determined by a cross-entropy function, the error signal of which is directly proportional to a difference between desired and actual output values. In the foregoing illustrative example, the CNN may process the image of the automobile and the output of the RNN reflecting the voice request saying "show me the artistic design features of this car."

**[0059]** The processing device performing the method may generate a result by the media processing system including at least one of an image (e.g., of a vehicle and/or a geographical location), a video (e.g., of a vehicle and a geographical location), a history (e.g., of a vehicle and/or a geographical location) and/or other textual information. In the foregoing illustrative example, the media processing system 101 may generate an image of the automobile, alone or in combination with text providing the make, model and year of the automobile. The generated image may also be annotated with lines and text that identify artistic features of the automobile.

**[0060]** The processing device performing the method displays the result. The result may be displayed, for example, on a user device such as a cell phone, iPad, monitor or in a virtual device such as a head-mounted display of a virtual reality, augmented reality and/or mixed reality system.

**[0061] FIG. 4** depicts a flow diagram of one illustrative example of a method 360 of displaying a result, in accordance with one or more aspects of the present disclosure. Method 360 and/or each of its individual functions, routines, subroutines, or operations may be performed by one or more processors of the computer system (*e.g.*, system 100 and/or processing device 124 of **FIG. 1**) executing the method. In certain implementations, method 360 may be performed by a single processing thread. Alternatively, method 360 may be performed by two or more processing threads, each thread executing one or more individual functions, routines, subroutines, or operations of the method. In an illustrative example, the processing threads implementing

method 360 may be synchronized (*e.g.*, using semaphores, critical sections, and/or other thread synchronization mechanisms). Alternatively, the processing threads implementing method 300 may be executed asynchronously with respect to each other.

**[0062]** Example method 360 produces an augmented reality display on a display device. At block 410, the processing device performing the method determines a viewing direction of a user wearing and augmented reality apparatus, for example, a head-mounted display. The viewing direction may be determined by angles in relation to a center of the head set (e.g., looking forward) and head position.

**[0063]** At block 420, the processing device performing the method determines an attitude of an augmented reality apparatus using distances between the user and the augmented reality apparatus. The distances may be measured by one or more distance sensors.

**[0064]** At block 430, the processing device performing the method controls a direction of image input of the augmented reality apparatus based on the viewing direction of the user and the attitude of the augmented reality apparatus. In an illustrative example, the augmented reality apparatus may include a driving unit that adjusts the direction of, for example, a digital camera horizontally or vertically so that a subject (e.g., a vehicle) corresponding to a subject image incident upon the digital camera can be chosen even when the augmented reality apparatus is fixed.

**[0065]** At block 440, the processing device performing the method receives an image of one or more subjects (e.g., a vehicle) in the direction of image input. A camera or other device for recording and storing images may be used to capture the image.

**[0066]** At block 450, the processing device performing the method generates a synthesized image by synthesizing the image of the one or more subjects with a digital image. In an illustrative example, the synthesized image will layer images, videos and text with the image of the one or more subjects to produce an augmented reality environment.

**[0067]** At block 460, the processing device performing the method displays the synthesized image. The synthesized image may be displayed on an augmented reality apparatus such as a head-mounted display. In an illustrative example, a user may see the image of a car captured using the user's cell phone underlaid by a video of geographical locations around the world. The image additionally or alternatively may be annotated with text such as arrows that point out different features of the car. The display may also be accompanied by voice information and/or music, for example, an audio description (e.g., spoken by a human or a bot) of the history of the car.

**[0068] FIG. 5** schematically illustrates an example of the neural network architecture and data pipeline together with a cloud-based, microservices-driven architecture (collectively referred to as “the architecture”) 500 for image processing and data analysis, in accordance with one or more aspects of the present disclosure. As schematically illustrated by **FIG. 5**, the architecture 500 includes a memory (not shown) and a database 510 (e.g., MongoDB©, Hbase©) configured for both in-memory and on-disk storage. The database 510 may include one or more trained machine learning models 512 for classifying and identifying images. In implementations, a storage or persistence layer may store images, metadata as a multidimensional cube warehouse, ML models, textual narratives, search-indexes, and software/applications underlying a transactional database. The architecture 500 may further include a plurality of containerized microservices 522A-C. In various example implementations, the runtime logic execution layer may be a collection of docker-container based microservices exposing representational state transfer (REST) application programming interfaces (APIs) (e.g., using Kubernetes18©). Each element 511, 513 and 515 represents a taxonomy feature that the corresponding microservice 522A, B and C is trained to analyze.

**[0069]** System 500 may further include a web application 532 including, for example, the media processing system 101 and system 100 and configured to execute methods 300 and 360. The web application 532 may be stored on a demilitarized zone (DMZ) network 530 to provide security. A virtual memory component 534 comprised of user comments and ratings may also be stored on the DMZ network 530 (i.e., these comments will not be added to the training data set 140 until authenticated). System 500 may further include a content delivery network 540, which is a content distribution network of proxy servers to ensure high availability of content. System 500 may further include a web presentation layer where one or more app users 550 can access the web app 532 and view results on a user device, for example, a cell phone 552A or a laptop 552B. In various example implementations, a presentation layer (e.g., ReactJS©) may be used for rendering the web-presentation layer (e.g., in 19 HTML5/CSS). The architecture may be implemented in a cloud or in a decentralized network (e.g., SOLID via Tim Berners-Lee).

**[0070]** The architecture 500 may further include digital walls 560, 562, 564 providing cybersecurity layers between the components of the architecture 500. Wall 560 may be implemented between the public web and application user devices 550 and the web application 532 and virtual component 534 in the DMZ 530. Wall 562 may be implemented between the DMZ 530 and microservices 520, 522A-C. Wall 564 may be implemented between the microservices 520 and the database 510.

**[0071]** According to various example implementations of the methods and systems described herein, CNN models may be implemented with multi-label classifiers to identify, for example, the make, mode, and year of manufacture of a vehicle. These classifiers may be implemented in, for example, TensorFlow® and Keras® using using ResNet, VGG-19 and/or Inception. In example implementations, these will feed into densely connected layers that predict into a region of an NLP embedding space. This embedding-space may then be used with NLP to identify relevant textual artifacts. According to example implementations, trained SNNs including CNNs may be used for vehicle authentication. The SNNs may use a contrastive loss function to compare a sample to a reference/fingerprint object.

**[0072]** According to various example implementations of the methods and systems described herein, RNN models may be implemented with LSTM, gated recurrent unit (GRU) and attention models. For example, the narratives users contribute, in addition those already curated as authentic history, will feed into NLP models based on RNN (LSTM/GRU) and Attention models like BERT, to assist a user in finding automobiles through descriptions. CNN-recognized objects and their associated meta-tags may play a role in the NLP results to map onto vehicles.

**[0073]** According to various implementations, the media processing system 101 may achieve greater than about 75% accuracy, or greater than about 80% accuracy, or greater than about 85% accuracy, or greater than about 90% accuracy, or greater than about 95% accuracy, or greater than about 99% accuracy when compared against the multi-label classifier. According to example implementations, an accuracy rate of greater than 90% may be achieved for cars that are more popular or common. For rarer cars, an accuracy rate of greater than about 80% may be achieved for vehicles that are less common or have limited production. In various example implementations, vehicle-clusters may be determined from broad descriptions. For example, when a user provides a broad description such as "Show me all Ford Mustangs," the media processing system 101 may provide a greater than 90% accuracy in identifying or recognizing the vehicle described when the text is sufficiently descriptive. For cultural heritage images, in various implementations the media processing system 101 may provide an accuracy of greater than about 80%, or greater than about 85%, or greater than about 90%, or greater than about 95%, or greater than about 99%. For example, the media processing system 101 may return a result with greater than about 80% probability of matching the query.

**[0074]** **FIG. 6** depicts a diagram of a server configuration 600 that may be employed in an example system for image processing and data analysis, in accordance with one or more aspects of the present disclosure. The server configuration 600 may be a computational cluster 610 (e.g., a Hadoop Cluster) having a master open source administration tool server and agent 612 (e.g., an

Ambari server and Ambari agent). The computational cluster 610 may further include a pair of slave agents 614A-B. A Hadoop cluster is a type of computational cluster designed to store and analyze large quantities of unstructured data in a distributed computing environment. Such clusters run Hadoop's open source distributed processing software on low-cost commodity computers. The cluster enables many computers to solve problems requiring massive computation and data.

**[0075]** **FIG. 7** illustrates a diagrammatic representation of a machine in the example form of a computer system 700 including a set of instructions executable by systems as described herein to perform any one or more of the methodologies discussed herein. In one implementation, the system may include instructions to enable execution of the processes and corresponding components shown and described in connection with **FIGs. 1-6**.

**[0076]** In alternative implementations, the systems may include a machine connected (e.g., networked) to other machines in a LAN, an intranet, an extranet, or the Internet. The machine may operate in the capacity of a server machine in client-server network environment. The machine may be a personal computer (PC), a neural computer, a set-top box (STB), Personal Digital Assistant (PDA), a cellular telephone, a server, a network router, switch or bridge, or any machine capable of executing a set of instructions (sequential or otherwise) that specify actions to be taken by that machine. Further, while only a single machine is illustrated, the term "machine" shall also be taken to include any collection of machines that individually or jointly execute a set (or multiple sets) of instructions to perform any one or more of the methodologies described herein.

**[0077]** The example computer system 700 can include a processing device (processor) 702, a main memory 704 (e.g., read-only memory (ROM), flash memory, dynamic random access memory (DRAM) such as synchronous DRAM (SDRAM)), a static memory 706 (e.g., flash memory, static random access memory (SRAM)), and a data storage device 718, which communicate with each other via a bus 730.

**[0078]** Processing device 702 represents one or more general-purpose processing devices such as a microprocessor, central processing unit, or the like. More particularly, the processing device 702 may be a complex instruction set computing (CISC) microprocessor, reduced instruction set computing (RISC) microprocessor, very long instruction word (VLIW) microprocessor, or a processor implementing other instruction sets or processors implementing a combination of instruction sets. The processing device 702 may also be one or more special-purpose processing devices such as an application specific integrated circuit (ASIC), a field programmable gate array (FPGA), a digital signal processor (DSP), network processor, or the like. In various implementations of the present disclosure, the processing device 702 is configured to execute

instructions for the devices or systems described herein for performing the operations and processes described herein.

**[0079]** The computer system 700 may further include a network interface device 708. The computer system 700 also may include a video display unit 710 (e.g., a liquid crystal display (LCD) or a cathode ray tube (CRT)), an alphanumeric input device 712 (e.g., a keyboard), a cursor control device 714 (e.g., a mouse), and a signal generation device 716 (e.g., a speaker).

**[0080]** The data storage device 718 may include a computer-readable medium 728 on which is stored one or more sets of instructions of the devices and systems as described herein embodying any one or more of the methodologies or functions described herein. The instructions may also reside, completely or at least partially, within the main memory 704 and/or within processing logic 726 of the processing device 702 during execution thereof by the computer system 700, the main memory 704 and the processing device 702 also constituting computer-readable media.

**[0081]** The instructions may further be transmitted or received over a network 720 via the network interface device 708. While the computer-readable storage medium 728 is shown in an example implementation to be a single medium, the term "computer-readable storage medium" should be taken to include a single medium or multiple media (e.g., a centralized or distributed database, and/or associated caches and servers) that store the one or more sets of instructions. The term "computer-readable storage medium" shall also be taken to include any medium that is capable of storing, encoding or carrying a set of instructions for execution by the machine and that cause the machine to perform any one or more of the methodologies of the present disclosure. The term "computer-readable storage medium" shall accordingly be taken to include, but not be limited to, solid-state memories, optical media, and magnetic media.

**[0082]** In various example implementations described herein, the neural networks are supervised learning models that analyze input data to classify objects, for example, using regression analysis. In an example implementation, a user may upload an image and the media processing system regresses it against many elements to determine the closest match. The supervised learning models are trained using different high-level elements of a taxonomy. The elements are related to categories of the taxonomy wherein the categories are used with ML to train the neural network models. In certain implementations, the elements may include, but are not limited to: actions (e.g., driving), concepts and emotions (e.g., direction), events (e.g., 2007 Tokyo Motor Show), geographic city (e.g., Los Angeles), geographic country (e.g., U.S.A.), geographic places (e.g., LAX Airport), geographic state (e.g., California), geographic location data (e.g., from a GPS), museum collections (e.g., Petersen Automotive Museum), photo environments (e.g., night), photo orientations (e.g., landscape), photo settings (e.g., auto garage), photo techniques (e.g., color),

photo views (e.g., three-quarter front view), signs (e.g., bowling alley), topic subjects (e.g., American culture), vehicle coachbuilder (e.g., Brewster & Co.), vehicle color (e.g., green), vehicle condition (e.g., new), vehicle manufacturer (e.g., including country and continent of origin), vehicle model (e.g., Bentley 6½ Liter Speed Six Tourer), vehicle parts (e.g., 8-track cassette player), vehicle quantity (e.g., one object), vehicle serial number (e.g., Chassis 3559SA), vehicle type (e.g., hydrogen fuel cell) and vehicle year of manufacture (e.g., 1957).

**[0083]** Implementations described herein can preserve and reveal information about vehicles and their impact on society. Using machine learning (ML) algorithms trained upon a proprietary dataset of curated and authenticated photos (image data), videos (video data), input data (e.g., text or voice inputs) and geolocation data, an artificial intelligence (AI) platform (e.g., including one or more convolutional neural network and one ore more recurrent neural network) has been developed that can *inter alia* identify vehicles from 1885 through present day and engages users to annotate images of the vehicles by sharing stories and comments. Machine learning includes, but is not limited to algorithms that find and apply patterns in data. Neural networks can be a form of ML. Implementations described herein provide a kind of time machine chassis capturing alchemical memories of shaped metal propelled through time and space, then identified through a multi-layered neural network. Enabling society to easily access information about vehicles through a searchable media processing system as described herein, augments and preserves human narrative, future transportation solutions and the history of remarkable vehicles.

**[0084]** In implementations, the database upon which the neural network is trained is a transdisciplinary study where abstract concepts (e.g., emotional, verbal, spatial, logical, artistic and social) are represented by semantic keywords expressing different dimensions of intelligence present in the referenced media object. For example, "Art Deco" is a semantic artistic keyword found on numerous vehicles from the 1920s and 1930s due to the visual design-language shown on a particular car or artifact. Using deep learning, the neural networks as described herein can be repeatedly trained for each of these distinct conceptual layers of intelligence found in the media thus resulting in object recognition enhanced through semantic intelligence and linking back to society. Additional databases including engineering information, racing results, car shows, location a car has been exhibited, valuations, etc. can be layered upon the proprietary vehicle database to further enhance information relating to the vehicles. In implementations, users can, for example, page through every Chevrolet Corvette in the library archives and read or listen to entries associated with any vehicle. Similarly, in implementations, a user can experience the development of streamlining or see the condensed vehicle design language of a particular decade of time, e.g. The Fifties. In implementations, a user can hold up a mobile device to an interesting car on the street and learn its story through interaction with the media processing system. For example, the media processing system may be configured to return facts such as "Did you know the V8 engine was invented one hundred years ago?"

**[0085]** Implementations described herein are configured to recognize and identify vehicles input via multiple sensory inputs (e.g., voice, image and text) and record a user's personal stories about the provenance and significance of vehicles, for example, in telling the story of America. Families can upload shoeboxes of family photos to learn what vehicle a great-grandfather once drove to the Grand Canyon. A user can travel to a historic place site and through the media processing system view hundreds of vehicles and families that previously visited the site over preceding decades. In implementations, family vacation photographs recording an event can be layered upon an existing geographic location to create a virtual environment, for example, using an immersive augmented reality (AR). In implementations, the AR environment can enable a user to see herself or himself, along with his or her ancestors and their vehicles at the same cultural heritage site evoking a ghostly rapture within the time-space continuum. For example, "How many photographs with the family car were taken at the Golden Gate Bridge?"

**[0086]** According to implementations, a proprietary database of images of vehicles contains several labeled images that belong to different categories including, but not limited to "make," "model" and "year." "Vehicle" refers to a mechanism for transporting things (e.g., people and goods), including, but not limited to planes, trains, automobiles (e.g., cars, trucks, motorcycles, vans, etc.) and spacecraft. The more images used for each category, the better the model (e.g., the convolutional neural network) can be trained to determine whether an image is, for example, a Porsche image or a Ferrari image. This implementation utilizes supervised machine learning. The model can then be trained using the labeled known images. The images in their extracted forms enter the input side of the model and the labels are in the output side. The purpose is to train the model such that an image with its features coming from the input will match the label in the output. Once the model is trained, it can be used to recognize, classify and/or predict an unknown image. For example, a new image may be recognized, classified or predicted as a 1933 Packard Twelve. As part of the processing by the convolutional neural network, the newly input image also goes through the pixel feature extraction process.

**[0087]** Implementations disclosed herein can address a fundamental problem in the advertising industry: how to authenticate a users' interest in a particular automotive brand or vehicle model or type. By virtue of a user uploading an unidentified Alfa Romeo to the platform, the user self-identifies the user's interest in this vehicle brand. Through repeated interactions, the media processing system learns the user's interests in vehicle information. Such a feedback loop is

valuable to advertisers for better targeting and in turn, can provide intelligence to manufacturers of future vehicles. The proprietary database of vehicles according to implementations, may be authenticated through timely registrations at the Library of Congress Copyright Office, which provides provenance that is preserved in the ML training dataset. In implementations, the training data set 140 may grow including additional data assets, for example, based on data input by users and/or additional assets that may not be copyright registered, but that may be authenticated, for example, by using SNNs as described herein.

**[0088]** Methods, systems and computer program products according to implementations herein relate to an AI platform built by systematically applying natural language processing (NLP) and computer vision, image and video processing to train a convolution and recurrent neural network from a dataset containing high quality, digital images, which may be copyrighted, of automobiles, capable of identifying a particular automobile from about 1885 through present day and into the future.

**[0089]** The essence of being human is to ask questions and AI seeks to provide credible information about a technological evolution: the journey of vehicles (e.g., the automobile), as well as the remaining surrounding artifacts of our vehicle heritage populating culture today. The implementations described herein provide an innovative AI-driven platform for classifying and documenting vehicles. Users can engage in a feedback cycle centered around identified photos, stories, comments, family photos and records. At their core, implementations herein nurture and explore the singular relationship of humans to machines, preserving the bond between the vehicles, design, community, architecture, engineering, history, government and culture to disseminate knowledge about vehicle culture. If a vehicle or important vehicle cultural heritage artifact is unknown, the platform can use the wisdom of crowds to identify and classify the asset. By NLP, the AI agent can begin chat interactions with the user about vehicles and immersive environments shown in the media thus deepening human-computer interaction skills.

**[0090]** Implementations described herein provide for the preservation and accessibility of collated, correlated and curated historical information (e.g., images, video, media, text, unstructured data, geographical location data, etc.) about and concerning vehicles, their use in human society, the environments (e.g., geographical locations) in which they are or have been used (e.g., racing and street), how they are or have been used, jobs they create or have created (e.g., in manufacturing and maintenance, consumer uses, collectors, etc.), technical and design features and special relationships with society.

**[0091]** According to implementations, multi-dimensional inputs query for vehicle attributes and elements from a vehicle dataset trained via ML from a proprietary reference database (e.g., a

neural network) built-upon copyright-registered and authenticated intellectual property about vehicles and their environments from the 1880s through present day and into the future to provide AI in mixed reality applications. Information may be retrieved from implementations described herein using multiple inputs including, but not limited to, audio, text, video, images and global positioning system (GPS) inputs where the input request is referenced against a proprietary-trained vehicle dataset and returns a classification and/or match to the input request in mixed-reality environments. Queries about vehicles can be answered with a probability of correct identification. For example, a user can type into the media processing system: “Auto Union Wanderer W-25” and the system would interpret the words to return an image of the “Auto Union Wanderer W-25.” The probability of the queried vehicle being built by “Auto Union” can be expressed as a percentage, for example, a “95%” probability of the image, video, text, history, etc. returned is an Auto Union and the probability of the image, video, text, history, etc. returned being a model “Wanderer W-25,” for example, “85%.” According to implementations, a short history of a vehicle appears and, using the geolocation services in the media processing system, an identification of where the closest example of this vehicle may be physically available for review relative to the user’s present location.

**[0092]** In implementations, information can be retrieved by uploading to the system (e.g., an app on a cell phone, a website, etc.), via a user interface, a photograph (e.g., a digital image, a scanned image, etc.) of a vehicle the user may encounter in daily life (e.g., on the street, at a car show, etc.). Using image recognition derived from the proprietary database, the media processing system can return, for example, a matching identification and/or classification of a vehicle make, model and year of release (referred to herein as “year”) with probabilities of accuracy based upon the trained dataset rendered through machine learning.

**[0093]** According to implementations described herein, users dictate voice request inputs, queries, text inputs and/or natural language to the media processing system. The input data can include, but is not limited to the make, model and year of a vehicle. For example, a user can speak “show me Ford Thunderbird” into a microphone that inputs data to the media processing system, which returns at least one of an image, a video, a written history, etc. representing the closest match to the “Ford Thunderbird” with additional information provided through mixed reality inputs. The user may refine a query by speaking “show me “red 1957 Ford Thunderbird” and the media processing system would return one or more image having the closest match together with a probability of accuracy. Existing platforms do not have provenance in their training data sets or information repositories (i.e., databases) and so the results returned by such platforms may be inaccurate or incomplete. Such platforms may scrape unauthenticated

information (e.g., not copyright registered, not validated with date, author, metadata, etc.) from publicly available websites such that if a certain number of users say something is, for example, a Thunderbird, then the platform will agree. However, if an asset or image is merely a replica of a Thunderbird, then such information is missed in these existing platforms. In another example, multiple years of Honda Accords may not be able to be recognized with authority by existing platforms.

**[0094]** According to implementations, a user initiates a query by pointing an input device (e.g., a camera, a video recorder, etc.) at a vehicle of interest and the media processing system receives the at least one input image and/or input video and matches it against the ML trained dataset to provide an Augmented Reality (AR) display of information regarding the queried vehicle. Levels of information may be chosen via user setup. For example, a user may only require vehicle make, such as “Ferrari,” or may require vehicle make and model, such as “Ferrari 250 GT,” or may require technical information like engine type, such as “V-8 engines;” the application is configured to return images, videos and/or information that matches V-8 engines from the neural network of information. According to implementations, in each query result, additional educational information about the vehicle is provided depending upon user settings. A brief history of the car can be displayed or overlaid in a mixed reality environment. For example, a user can submit a text or natural language input query, such as “two-tone vehicle interiors,” and matches to the requisition can be displayed on the user device with overlaid text depicting, for example, the make, model, history, design features, etc. of vehicles having two-tone interiors.

**[0095]** The implementations described herein are useful in a variety of fields, for example, where entities may need media assets or user information on vehicles obtainable from an organized, curated, and searchable platform. Example fields include, but are not limited to 1) advertisers: any automobile-related business or ads in which cars appear need authenticated product; 2) automobile manufacturers: marketing need for brand building, loyalty and heritage promotion of manufacturer’s products/services; 3) insurers: verifying vehicles is key in protecting assets and individuals; 4) entertainment: immersive experiences through augmented/virtual reality, skill games; 5) law enforcement: need help in identification of vehicles involved in investigations, possibly from photos taken at/from a crime scene—e.g., by a bystander on his/her cell phone—and fraud detection; 6) vehicle designers: need access to historical examples and perspective for new designs; 7) travel: roadside support, fuel, lodging, food, interesting roads and points of interest along the roadside; 8) classic car market and collectors: buyers, sellers and restorers of vehicles need parts authenticity, provenance

information, special features, and historical context in which the cars appear and are used; and 9) museums and archives: need help with identification, provenance and automotive history found in photo collections.

**[0096]** User interest, expressed by uploads of unidentified photos and by time spent reviewing certain vehicle brand archive sections to the media processing system, self-identifies the user's interest in specific vehicle brands and/or segments that can be sought after targets for advertisers. For example, a user who reads and peruses the Porsche archives is a good target for Porsche brand advertising.

**[0097]** By interacting with the platform, users self-identify interest in a particular automotive brand or vehicle sector thus solving an advertising problem for customers who wish to learn from past automotive designs, verify their illustrated marketing materials and target their communications to potential buyers in the automotive sector of our economy. Users may explore curated information about automobiles and roadside heritage through a virtual library linking other datasets to form a central integrated intelligence platform about automobiles and society. Transportation designers can easily access lessons learned from the last 135 years of automotive design.

**[0098]** Geolocation data, alone or in combination with curated photos of architectural and/or cultural heritage sites, can also be input to the media processing system as described herein. In certain implementations, the application can direct users to roadside services based on personalized user data (e.g., the user's preferred fuel type, fast food and hotel preferences can all be stored) and geolocation data received from a navigation system. For example, suppose a user drives a sports car, which is known from the user's profile stored in a memory accessible by the media processing system. The media processing system may have access to the user's calendar also stored on a memory accessible by the media processing system. The media processing system can receive an input from a navigation program indicating that the user's arrival at a calendar appointment location is estimated for 15 minutes, but there is a great two-lane road the user could drive that would be fun and would still get the user there on time for the calendar appointment. The media processing system would then suggest the alternate route.

**[0099]** Families throughout history have authenticated life in photographs, which may have been taken while on vacations using vehicles. Such image data can be used to virtually augment cultural heritage sites or historic places using historic photos including vehicles. For example, according to various implementations using augmented reality, virtual reality and/or mixed reality, the media processing system can enable users to virtually tour Route 66 throughout history. In implementations, the systems and methods described herein can use augmented,

virtual and/or mixed reality to enhance travel and road trips. For example, a user may drive down Route 66 and, using geolocation data, hold up a device (e.g., a cell phone) and see the present location as it evolved through history within a virtual display device (e.g., a head mounted device).

**[0100]** According to implementations, augmented reality relating to vehicles can be used for cultural heritage tourism to enhance the tourist experience. Linking contextual information found in the backgrounds of family photos, provides the groundwork for creating an authenticated augmented reality system, for example, for America's historic places. For example, implementations described herein are useful for auto clubs such as AAA. A mobile device can be pointed at a vehicle and/or cultural heritage location to capture image data and/or geolocation data and the media processing system can return images of the vehicle and/or cultural heritage location over time at that particular location.

**[0101]** Embodiments of the present disclosure can be described in view of the following clauses. In clause 1, a method comprises: training a convolutional neural network (CNN) using authenticated data and a taxonomy; receiving, by a processing device, a query comprising input data; classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy; generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

**[0102]** In clause 2, the method of clause 1, wherein the authenticated data comprises copyright registered works of authorship, metadata and text. In clause 3, the method of clause 2, wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings. In clause 4, the method of clause 3, wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information. In clause 5, the method of clause 1, wherein the authenticated data comprises data from a copyright registered database. In clause 6, the method of clause 1, wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture. In clause 7, the method of clause 1, wherein the input data comprises one or more of image data,

video data, intake data or geographical location data. In clause 8, the method of clause 1, wherein classifying comprises mapping input data to authenticated data using the taxonomy. In clause 9, the method of clause 1, wherein the result comprises one or more of an image, a video, text, or sound. In clause 10, the method of clause 1, wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information.

**[0103]** In clause 11, the method of clause 1, wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy. In clause 12, the method of clause 11, wherein the probability is determined by a cross-entropy function. In clause 13, the method of clause 1, wherein the result comprises augmented reality content, wherein displaying the result comprises: displaying the result in an augmented reality apparatus, comprising: passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device; and displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data. In clause 14, the method of clause 13, wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time. In clause 15, the method of clause 1, further comprising training a recurrent neural network (RNN) using authenticated data and a taxonomy. In clause 16, the method of clause 15, wherein the input data comprises unstructured data, the method further comprising: processing, by the trained RNN, the unstructured data to yield structured data; and classifying, by the trained CNN, the structured data. In clause 17, the method of clause 1, wherein the input data comprises user uploaded data, the method further comprising authenticating the user uploaded data using SNNs and adding the authenticated user uploaded data to the authenticated data.

**[0104]** In clause 18, a system comprising: a memory; a processor, coupled to the memory, the processor configured to: train a convolutional neural network (CNN) using authenticated data and a taxonomy; receive, by a processing device, a query comprising input data; classify, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy; generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content. In clause 19, the system of clause

18, wherein the authenticated data comprises copyright registered works of authorship, metadata and text. In clause 20, the system of clause 19, wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings. In clause 21, the system of clause 20, wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information. In clause 22, the system of clause 18, wherein the authenticated data comprises data from a copyright registered database. In clause 23, the system of clause 18, wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture. In clause 24, the system of clause 18, wherein the input data comprises one or more of image data, video data, intake data or geographical location data. In clause 25, the system of clause 18, wherein classifying comprises mapping input data to authenticated data using the taxonomy. In clause 26, the system of clause 18, wherein the result comprises one or more of an image, a video, text, or sound. In clause 27, the system of clause 18, wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information. In clause 28, the system of clause 18, wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy. In clause 29, the system of clause 28, wherein the probability is determined by a cross-entropy function. In clause 30, the system of clause 18, wherein the result comprises augmented reality content, wherein displaying the result comprises: displaying the result in an augmented reality apparatus, comprising: passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device; and displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data. In clause 31, the system of clause 30, wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time. In clause 32, the system of clause 18, further configured to train a recurrent neural network (RNN) using authenticated data and a

taxonomy. In clause 33, the method of clause 32, wherein the input data comprises unstructured data, the method further comprising: processing, by the trained RNN, the unstructured data to yield structured data; and classifying, by the trained CNN, the structured data. In clause 34, the method of clause 18, wherein the input data comprises user uploaded data, wherein the system is further configured to authenticate the user uploaded data using SNNs and add the authenticated user uploaded data to the authenticated data.

**[0105]** In clause 35, a computer-readable non-transitory storage medium comprising executable instructions that, when executed by a computing device, cause the computing device to perform operations comprising: training a convolutional neural network (CNN) using authenticated data and a taxonomy; receiving, by a processing device, a query comprising input data; classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy; generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content. In clause 36, the computer-readable non-transitory storage medium of clause 35, wherein the authenticated data comprises copyright registered works of authorship, metadata and text. In clause 37, the computer-readable non-transitory storage medium of clause 36, wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings. In clause 38, the computer-readable non-transitory storage medium of clause 37, wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information. In clause 39, the computer-readable non-transitory storage medium of clause 35, wherein the authenticated data comprises data from a copyright registered database. In clause 40, the computer-readable non-transitory storage medium of clause 35, wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture. In clause 41, the computer-readable non-transitory storage medium of clause 35, wherein the input data comprises one or more of image data, video data, intake data or geographical location data. In clause 42, the computer-readable non-transitory storage medium of clause 35, wherein classifying comprises mapping input data to authenticated data using the

taxonomy. In clause 43, the computer-readable non-transitory storage medium of clause 35, wherein the result comprises one or more of an image, a video, text, or sound. In clause 44, the computer-readable non-transitory storage medium of clause 35, wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information. In clause 45, the computer-readable non-transitory storage medium of clause 35, wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy. In clause 46, the computer-readable non-transitory storage medium of clause 45, wherein the probability is determined by a cross-entropy function. In clause 47, the computer-readable non-transitory storage medium of clause 35, wherein the result comprises augmented reality content, wherein displaying the result comprises: displaying the result in an augmented reality apparatus, comprising: passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device; and displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data. In clause 48, the computer-readable non-transitory storage medium of clause 47, wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time. In clause 49, the computer-readable non-transitory storage medium of clause 35, further comprising training a recurrent neural network (RNN) using authenticated data and a taxonomy. In clause 50, the computer-readable non-transitory storage medium of clause 49, wherein the input data comprises unstructured data, the method further comprising: processing, by the trained RNN, the unstructured data to yield structured data; and classifying, by the trained CNN, the structured data. In clause 51, the computer-readable non-transitory storage medium of clause 35, wherein the input data comprises user uploaded data, the method further comprising authenticating the user uploaded data using SNNs and adding the authenticated user uploaded data to the authenticated data. In clause 52, the medium of clause 51, wherein processing by the CNN at least one of the image data, the video data, the input data and the geographical location data yields one or more image of a vehicle comprising matching features.

**[0106]** The preceding description sets forth numerous specific details such as examples of specific systems, components, methods, and so forth, in order to provide a good understanding of several implementations of the present disclosure. It will be apparent to one skilled in the art, however,

that at least some implementations of the present disclosure may be practiced without these specific details. In other instances, well-known components or methods are not described in detail or are presented in simple block diagram format in order to avoid unnecessarily obscuring the present disclosure. Thus, the specific details set forth are merely presented as examples. Particular implementations may vary from these example details and still be contemplated to be within the scope of the present disclosure. In the above description, numerous details are set forth.

**[0107]** It will be apparent, however, to one of ordinary skill in the art having the benefit of this disclosure, that implementations of the disclosure may be practiced without these specific details. In some instances, well-known structures and devices are shown in block diagram form, rather than in detail, in order to avoid obscuring the description.

**[0108]** Some portions of the detailed description are presented in terms of algorithms and symbolic representations of operations on data bits within a computer memory. These algorithmic descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical, magnetic, or optical signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

**[0109]** It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the above discussion, it is appreciated that throughout the description, discussions utilizing terms such as “rating,” “selecting,” “comparing,” “adjusting,” or the like, refer to the actions and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (e.g., electronic) quantities within the computer system’s registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

**[0110]** Implementations of the disclosure also relate to an apparatus for performing the operations herein. This apparatus may be specially constructed for the required purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, and

magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, or any type of media suitable for storing electronic instructions.

**[0111]** The algorithms and displays presented herein are not inherently related to any particular computer or other apparatus. Various general purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct a more specialized apparatus to perform the required method steps. The required structure for a variety of these systems will appear from the description below. In addition, the present disclosure is not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the disclosure as described herein.

**[0112]** It is to be understood that the above description is intended to be illustrative, and not restrictive. Many other implementations will be apparent to those of skill in the art upon reading and understanding the above description.

## CLAIMS

I/We claim:

1. A method comprising:

training a convolutional neural network (CNN) using authenticated data and a taxonomy;

receiving, by a processing device, a query comprising input data;

classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy;

generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and

displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

2. The method of claim 1, wherein the authenticated data comprises copyright registered works of authorship, metadata and text.

3. The method of claim 2, wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings.

4. The method of claim 3, wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information.

5. The method of claim 1, wherein the authenticated data comprises data from a copyright registered database.

6. The method of claim 1, wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture.

7. The method of claim 1, wherein the input data comprises one or more of image data, video data, intake data or geographical location data.

8. The method of claim 1, wherein classifying comprises mapping input data to authenticated data using the taxonomy.

9. The method of claim 1, wherein the result comprises one or more of an image, a video, text, or sound.

10. The method of claim 1, wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information.

11. The method of claim 1, wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy.

12. The method of claim 11, wherein the probability is determined by a cross-entropy function.

13. The method of claim 1, wherein the result comprises augmented reality content, wherein displaying the result comprises:

displaying the result in an augmented reality apparatus, comprising:

passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides;

imaging the light at the display device; and

displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data.

14. The method of claim 13, wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time.

15. The method of claim 1, further comprising training a recurrent neural network (RNN) using authenticated data and a taxonomy.

16. The method of claim 15, wherein the input data comprises unstructured data, the method further comprising:

processing, by the trained RNN, the unstructured data to yield structured data; and

classifying, by the trained CNN, the structured data.

17. The method of claim 1, wherein the input data comprises user uploaded data, the method further comprising authenticating the user uploaded data using Siamese Neural Networks and adding the authenticated user uploaded data to the authenticated data.

18. A system comprising:

a memory;

a processor, coupled to the memory, the processor configured to:

train a convolutional neural network (CNN) using authenticated data and a taxonomy;

receive, by a processing device, a query comprising input data;

classify, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy;

generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and

display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

19. The system of claim 18, wherein the authenticated data comprises copyright registered works of authorship, metadata and text.

20. A computer-readable non-transitory storage medium comprising executable instructions that, when executed by a computing device, cause the computing device to perform operations comprising:

training a convolutional neural network (CNN) using authenticated data and a taxonomy;

receiving, by a processing device, a query comprising input data;

classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy;

generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and

displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

## ABSTRACT

The present disclosure relates generally to methods, systems and computer program products for classifying and identifying input data using neural networks and displaying results (e.g., images of vehicles, vehicle artifacts and geographical locations dating from the 1880s to present day and beyond). The results may be displayed on displays or in virtual environments such as on virtual reality, augmented reality and/or mixed-reality devices.

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

| | | | |
|---|---|---|---|
| **Inventor(s):** | LEWIS, Lucinda | **Examiner:** | TBD |
| **Application No.:** | | **Confirmation No.:** | TBD |
| **Filed:** | Herewith | **Art Unit:** | TBD |
| **For:** | METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY | | |

**Mail Stop Amendment**
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

**INFORMATION DISCLOSURE STATEMENT**

Sir/Madam:

In order to comply with the duty of disclosure under 37 C.F.R. § 1.56, transmitted herewith is an Information Disclosure Statement (IDS) in connection with the above-identified patent application. This IDS complies with the requirements under 37 C.F.R. § 1.98 and is being submitted pursuant to 37 C.F.R. § 1.97(b), wherein the IDS is being offered for consideration in a national application before the mailing date of a first Office Action on the merits.

Applicant believes there is no fee required for submission of this IDS. However, if a fee is due, the Commissioner is hereby authorized to charge any deficiency in fees, or credit overpayment of the same, to Deposit Account No. 50-1358.

The Examiner's attention is kindly directed to the documents identified on the enclosed/attached Form PTO/SB/08. Copies of U.S. Patents and/or Patent Application

Publications are not required pursuant to 37 C.F.R. § 1.98 and, therefore, have not been provided in connection with this submission.

It is respectfully requested that these references be fully considered by the U.S. Patent and Trademark Office during the examination of this application and printed on any patent which may issue thereon. Accordingly, Applicant kindly requests that a copy of Forms PTO/SB/08, as considered and initialed by the Examiner, be returned with the next office communication.

An early and favorable action is respectfully requested.

Respectfully submitted,

Lowenstein Sandler LLP
One Lowenstein Drive
Roseland, NJ 07068

Dated: November 9, 2021

/Mary K. Nicholes/
Mary K. Nicholes
Attorney for Applicant
Reg. No. 56,238

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) **World Intellectual Property Organization**
International Bureau

(43) **International Publication Date**
**12 November 2020 (12.11.2020)**

WIPO | PCT

(10) **International Publication Number**
**WO 2020/227651 A1**

(51) **International Patent Classification:**
*G06N 20/00* (2019.01) *G06K 9/62* (2006.01)
*G06K 9/46* (2006.01) *G06N 3/08* (2006.01)

(21) **International Application Number:**
PCT/US2020/032149

(22) **International Filing Date:**
08 May 2020 (08.05.2020)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
62/845,546 09 May 2019 (09.05.2019) US

(71) **Applicant: AUTOMOBILIA II, LLC** [US/US]; 301 Clematis Street #3000, West Palm Beach, FL 33401 (US).

(72) **Inventor: LEWIS, Lucinda**; 414 E. Valencia Avenue, #103, Burbank, CA 91505 (US).

(74) **Agent: NICHOLES, Mary, K.** et al.; Lowenstein Sandler LLP, One Lowenstein Drive, Roseland, NJ 07068 (US).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

**Published:**
— *with international search report (Art. 21(3))*

(54) **Title:** METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

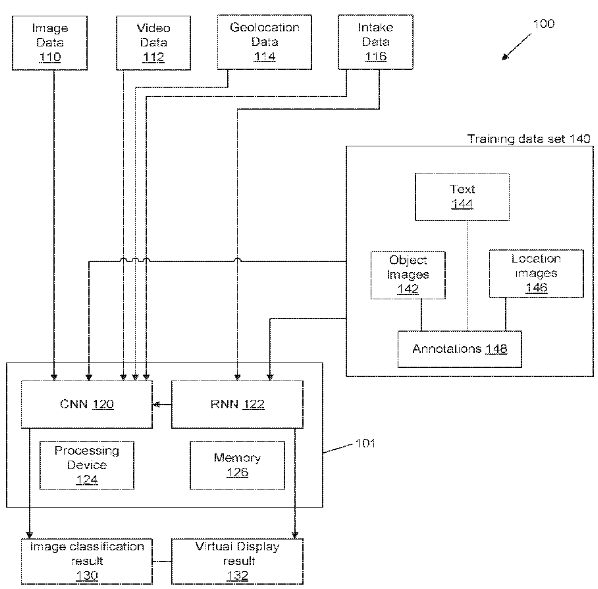



FIG. 1

(57) **Abstract:** The present disclosure relates generally to methods, systems and computer program products for classifying and identifying input data using neural networks and displaying results (e.g., images of vehicles, vehicle artifacts and geographical locations dating from the 1880s to present day and beyond). The results may be displayed on displays or in virtual environments such as on virtual reality, augmented reality and/or mixed-reality devices.

WO 2020/227651 A1

# METHODS, SYSTEMS AND COMPUTER PROGRAM PRODUCTS FOR MEDIA PROCESSING AND DISPLAY

## BACKGROUND

**[0001]** Cultural knowledge about one of mankind's most significant inventions, the automobile, is deeply fragmented and not easily searchable. Today, a user may see an object, such as a vehicle, driving down the road and wonder what it is, or an autonomous vehicle driving on a freeway may have a need to identify objects such as vehicles in its vicinity to learn information (e.g., vehicle make, model, year, stopping distance, etc.) and any equipment on board (e.g., communication devices, computer systems, etc.). Such information is currently unauthenticated, fragmented, dispersed, and not readily available in searchable form.

**[0002]** As society becomes increasingly digitalized, benefits have emerged in augmenting human intelligence with artificial intelligence (AI) to answer questions with curated knowledge. The market of automobile information is segmented and incoherent, with bits of information gathered by each segment without any relationship to each other and, thus, the information is not readily searchable. Each market segment has different needs, e.g., a user may have an instant need to identify certain vehicles, auto-related architecture and/or cultural artifacts out of interest whereas an autonomous vehicle may have a need to identify other vehicles in its vicinity and access their capabilities.

**[0003]** In 2018, the automotive advertising segment exceeded $38 trillion, exclusive of advertising for travel and food and automotive repair. The automotive advertising sector is the second largest advertising sector in the overall advertising marketplace. The largest customers are advertisers for new car buyers interested in the heritage of an automotive brand, the collectible car enthusiast market, the automotive parts market, insurance, travel, media archives and libraries with unidentified assets, and consumers with an unidentified photo album showing family vehicles. Additional commercial opportunities reside with government, security, law enforcement, and the entertainment industry.

**[0004]** Existing platforms are unable to sufficiently identify vehicles. These platforms can only make inferences from unauthenticated data. Anyone looking to authenticate a vehicle using these platforms can spend hours and still be uncertain of the exact make, model and year. These platforms have no verifiable source of automotive data and little hope of authoritatively identifying cars. It is difficult and frustrating to credibly authenticate automobiles by searching the Internet or asking vehicle owners; significant knowledge is fading from human memory.

**[0005]** There is a need for methods, systems and computer program products that allow a user to identify and search objects such as vehicles, while simultaneously building provenance and preserving knowledge around such objects and their cultural history, to instantly and properly identify objects by training artificial intelligence tools, to preserve cultural history, to promote historic places influenced by the vehicle's evolution, to establish automotive provenance based on first-hand historical records, to celebrate and educate users about objects (e.g., vehicles), to learn from history and help shape mobility's evolution, to explore human-computer interactions and to address the automotive advertising market. The systems, methods and computer program products described herein can quickly and accurately identify objects such as vehicles using authenticated data and provide an armature for social data to accrue around any object, building provenance around a subject that has lacked the tools to easily authenticate knowledge, e.g., cars and their impact on culture.

## BRIEF DESCRIPTION OF THE DRAWINGS

**[0006]** The present disclosure is illustrated by way of examples, and not by way of limitation, and may be more fully understood with references to the following detailed description when considered in connection with the figures, in which:

**[0007] FIG. 1** schematically illustrates an example system for image processing and data analysis, in accordance with one or more aspects of the present disclosure.

**[0008] FIG. 2** schematically illustrates an example structure of a convolutional neural network (CNN) that may be employed to process data input to an example system for media classification and identification, in accordance with one or more aspects of the present disclosure.

**[0009] FIG. 3** depicts a flow diagram of one illustrative example of a method 300 of data processing and data analysis, in accordance with one or more aspects of the present disclosure.

**[0010] FIG. 4** depicts a flow diagram of one illustrative example of a method 360 of displaying via augmented reality a result from an example system for media classification and identification.

**[0011] FIG. 5** depicts a diagram of a system for implementing the methods and systems described herein.

**[0012] FIG. 6** depicts a diagram of a computational cluster system that may be employed in an example system for media classification and identification, in accordance with one or more aspects of the present disclosure.

**[0013] FIG.** 7 depicts a diagram of an illustrative example of a computing device implementing the systems and methods described herein.

## DETAILED DESCRIPTION

**[0014]** Described herein are methods, systems and computer program products for media classification and identification and for displaying results on a display (e.g., a cell phone, a monitor, an augmented reality apparatus, a mixed reality apparatus). While the systems and methods are described with respect to vehicles, vehicle artifacts and geographical locations, the systems and methods are broadly applicable to any objects. In example implementations, the systems and methods may relate to buildings (e.g., architecture), clothing, bridges, tools, highways, mountains, parks, rivers, cities, cars converted to homes and so on. Thus, the systems and methods described herein may be applied to a wide variety of physical objects that may involve various combinations of multiple imaging and/or other image capturing mechanisms.

**[0015]** The present disclosure overcomes the above-noted and other deficiencies by providing systems and methods for image processing and data analysis that may be utilized for identifying, classifying, researching and analyzing objects including, but not limited to vehicles, vehicle parts, vehicle artifacts, cultural artifacts, geographical locations, etc. To identify all of the objects in a photo, alone or in combination with a geographical location and/or a cultural heritage object, and then to associate a narrative with them represents a unique challenge. The converse of this too – the identification of historic places and objects (e.g. Statue of Liberty) alone, or in combination with vehicles— forms a broad descriptive visual narrative that illustrates innovative mapping from natural language processing (NLP) to multi-label image classification and identification.

**[0016]** In example implementations, a repository of photos, videos, keywords and captions of automobiles of proven provenance, with user narratives and comments, can be used to train a unique AI pipeline to map the information to a target space for image classification. For example, given an uploaded user image, the AI models may create the most appropriate summary of the relevant sections of the asset, and perform a multi-labeled classification of the image into the appropriate model of, for example, car manufacturer and year. In the case of an image containing multiple cars and cultural artifacts (e.g., glove compartments, spokes, steering wheels, vehicle lifts, etc.), there may be the additional task of establishing bounding-boxes around each of the recognized objects, and creating summary text appropriate to the image as a whole.

**[0017]** Likewise, the converse problem of taking a vehicle description, i.e., "Show me Prototypes", and enriching it with AI assisted discovery into a proprietary database of high quality copyrighted images, represents a journey where the feature-vectors comprise the NLP embeddings of the narratives. The target space may be comprised of clusters of automotive

images that share attributes; for example, the query may map to a cluster of experimental cars from a particular decade. This may involve a single machine learning (ML) pipeline where RNN (LSTM/GRU) and BERT-derived attention models interact with CNN-architectures for image classification and Siamese Neural Networks (SNNs) for correct identifications. A collaborative user verification process involving crowd wisdom can be used to improve the accuracy of image-augmentation such that users can point out errors and suggest corrections. Should certain annotations be erroneous and users mark them so, such data will feed into the next round of neural architecture training.

**[0018]** In certain implementations, the systems and methods described herein may perform pixel-level analysis of images (and/or videos) in order to yield images, videos and/or virtual environments (e.g., augmented reality, mixed reality, virtual reality etc.) of vehicles, vehicle artifacts (e.g., images of vehicle tools, feature elements such as goggles, tachometers, wheel spokes, gas cans, etc.) and/or geographical locations. The systems and methods described herein may further determine whether images or videos input to the media processing system contain features that match one or more features of an image, video and/or geographical location of stored in a memory. In implementations, the systems and methods produce a result that comprises the closest matching data (e.g., having the highest probability score based on a cross-entropy function) identified in the training data set and/or database repository. The result may include an image of a vehicle together with text information about the vehicle such as a history, make, model, year, etc. The systems and methods may additionally yield historical information about one or more vehicle and/or geographical location and such information may be displayed in a virtual environment. In certain implementations, the systems and methods as described herein may also be implemented in an autonomous vehicle to capture images and/or video of surrounding vehicles on the road and to produce a result indicating the size, make, model and on-board equipment of the surrounding vehicles. In an example implementation, an autonomous vehicle incorporating the systems and methods described herein can be trained for “platooning.” An example of “platooning” is where a vehicle operating in a self-driving or semi-autonomous mode, analyzes other vehicles in its vicinity to determine, for example, which vehicles may be capable of vehicle-to-vehicle (V2V) communication, other equipment on board, the estimated stopping distance of each vehicle and surrounding environmental objects such as children, balls, bicycles, tumbleweeds, etc. The autonomous vehicle may then communicate with the V2V vehicles to maintain a safe speed and distance to those vehicles, that is, the vehicles may move harmoniously together and may stop together at a traffic lights. Any vehicles in the vicinity that are not capable of V2V may be accounted for as an unknown variable. In other

implementations, platooning may involve recognition by the autonomous vehicle of structures that are capable of communicating with the vehicle in a vehicle-to-infrastructure (V2I) configuration. If the infrastructure is equipped with methods and systems as described herein, it may time or adjust the traffic lights to enhance platooning of V2I vehicles taking into consideration the variables of vehicles that are not equipped for V2I communication.

**[0019]** The systems and methods described herein have the benefit of being trained using a proprietary database comprising high-quality, digital copyrighted images of vehicles, vehicle artifacts and/or geographic sites (e.g., historical sites, cultural sites), such that the accuracy of the results produced by the disclosed systems and methods is improved over known methods of researching and analyzing vehicles, vehicle artifacts and/or geographical locations. The database may further include videos, embedded metadata and text. The database may itself be copyrighted. Because the database and data assets (e.g., images, videos, text, etc.) are themselves copyrighted, they form a body of authenticated data on which the neural networks can be trained.

**[0020]** The systems and methods described herein utilize a convolutional neural network (CNN) or a combination of both a CNN and a recurrent neural network (RNN), which form a part of a media processing system. The CNN may process one or more of image data (e.g., containing images, for example, of vehicles, vehicle artifacts, landscapes, etc.), video data (e.g., videos of vehicles, videos of historical sites, etc.), geolocation data (e.g., from a global positioning system) or intake data (e.g., text queries entered via a user interface, voice queries, natural language queries, etc.) to perform classification and with respect to vehicle information, vehicle artifact information, geographical location, etc. and/or to yield the probability of one or more image, video, geolocation and/or intake query matching significant features associated with a vehicle, a vehicle artifact and/or geographical location. The returned images, videos and/or virtual environment may be annotated and/or layered (e.g., overlaid, underlaid) with historical, design, mechanical, etc. information in the form of, for example, text, audio and video.

**[0021]** The RNN processes unstructured data, for example, natural language search queries and/or voice inputs to provide natural language processing (NLP). The unstructured data is transformed into structured data, which is fed to the CNN and processed as described above. In example implementations, the neural network architecture creates a hybridization of natural language processing (RNN) with image classification and identification techniques (CNN) for the purposes of preserving and accumulating data around a key subject area of historical and cultural interest.

**[0022]** A CNN is a computational model based on a multi-staged algorithm that applies a set of pre-defined functional transformations to one or more input (e.g., image pixels) and then utilizes the transformed data to perform, for example, classification, identification, image recognition, pattern recognition, etc. A CNN may be implemented as a feed-forward neural network (FFNN) in which the connectivity pattern between its neurons is inspired by the organization of the animal visual cortex. Individual cortical neurons respond to stimuli in a restricted region of space known as the receptive field. The receptive fields of different neurons partially overlap such that they tile the visual field. The response of an individual neuron to stimuli within its receptive field can be approximated mathematically by a convolution operation. In addition to image processing, the CNN may be used for other input types such as text, audio and video. In implementations, images may be input to a media processing system as described herein and the CNN processes the data. For example, if a user inputs a picture of a Ford Thunderbird automobile, the media processing system may output an image of a Ford Thunderbird together with the make, model, year, history and any known contextual information surrounding the photo and background.

**[0023]** In an illustrative example, a CNN may include multiple layers of various types, including convolution layers, non-linear layers (e.g., implemented by rectified linear units (ReLUs)), pooling layers, and classification (fully-connected) layers. A convolution layer may extract features from the input image by applying one or more learnable pixel-level filters to the input image. In an illustrative example, a pixel-level filter may be represented by a matrix of integer values, which is convolved across the dimensions of the input image to compute dot products between the entries of the filter and the input image at each spatial position, thus producing a feature map that represents the responses of the filter at every spatial position of the input image. The convolution filters are defined at the network training stage based on the training dataset to detect patterns and regions that are indicative of the presence of significant features within the input image.

**[0024]** A non-linear operation may be applied to the feature map produced by the convolution layer. In an illustrative example, the non-linear operation may be represented by a rectified linear unit (ReLU) which replaces with zeros all negative pixel values in the feature map. In various other implementations, the non-linear operation may be represented by a hyperbolic tangent function, a sigmoid function, or by other suitable non-linear function.

**[0025]** A pooling layer may perform subsampling to produce a reduced resolution feature map while retaining the most relevant information. The subsampling may involve averaging and/or determining maximum value of groups of pixels.

**[0026]** In certain implementations, convolution, non-linear, and pooling layers may be applied to the input image multiple times prior to the results being transmitted to a classification (fully-connected) layer. Together these layers extract the useful features from the input image, introduce non-linearity, and reduce image resolution while making the features less sensitive to scaling, distortions, and small transformations of the input image.

**[0027]** The output from the convolutional and pooling layers represent high-level features of the input image. The purpose of the classification layer is to use these features for classifying the input image into various classes. In an illustrative example, the classification layer may be represented by an artificial neural network that comprises multiple neurons. Each neuron receives its input from other neurons or from an external source and produces an output by applying an activation function to the sum of weighted inputs and a trainable bias value. A neural network may include multiple neurons arranged in layers, including the input layer, one or more hidden layers, and the output layer. Neurons from adjacent layers are connected by weighted edges. The term "fully connected" implies that every neuron in the previous layer is connected to every neuron on the next layer.

**[0028]** The edge weights are defined at the network training stage based on the training dataset. In an illustrative example, all of the edge weights are initialized to random values. For every input in the training dataset, the neural network is activated. The observed output of the neural network is compared with the desired output specified by the training data set, and the error is propagated back to the previous layers of the neural network, in which the weights are adjusted accordingly. This process is repeated until the output error is below a predetermined threshold.

**[0029]** The CNN may be implemented in a SNN configuration. A SNN configuration contains two or more identical subnetwork components. In implementations, not only is the architecture of the subnetworks identical, but the weights are shared among them as well. SNNs learn useful data descriptors, which may be used to compare the inputs (e.g., image data, video data, input data, geolocation data, etc.) of the subnetworks. For example, the inputs may be image data with CNNs as subnetworks.

**[0030]** The CNN may be implemented in a Generative Adversarial Network (GAN), which refers to two networks working together. A GAN can include any two networks (e.g., a combination of FFNNs and CNNs), with one tasked to generate content and the other tasked to judge content. The discriminating network receives either training data or generated content from the generative network. The ability of the discriminating network to correctly predict the data source is then used as part of the error for the generating network. This creates a form of competition where the discriminator gets better at distinguishing real data from generated data

and the generator learns to become less predictable to the discriminator. Even quite complex noise-like patterns can become predictable, but generated content similar in features to the input data is harder to learn to distinguish. The dynamics between the two networks need to be balanced; if prediction or generation becomes too good compared to the other, the GAN will not converge as there is intrinsic divergence.

**[0031]** A RNN may be described as a FFNN having connections between passes and through time. A RNN receives not just the current input it is fed, but also what it has perceived previously in time. In a RNN, neurons can be fed information not only from a previous layer, but from a previous pass. A string of text or picture can be fed one pixel or character at a time, so that the time dependent weights can be used for what came before in the sequence, not actually from what happened a specific time (e.g., x seconds) before. The RNN may be implemented as a Long Short-Term Memory (LFTM), which helps preserve the error, back-propagating it through layers and time. A LSTM includes information outside of the normal flow of the RNN in a gated cell. Information can be written to, stored in, or read from a cell, similar to data in a computer's memory. The cell can make decisions about when to allow reads, writes and erasures, and what to store via gates that open and close. These gates are analog, implemented with element-wise multiplication by sigmoids (i.e., all in the range of 0-1).

**[0032]** The RNN may be implemented with Bidirectional Encoder Representations from Transformers (BERT) to perform NLP tasks including *inter alia* question answering and natural language inference. BERT, which uses a Transformer-based language model, is a language representation model that provides accuracy for NLP tasks. A Transformer, in an encoding step, can use learned word embedding to convert words, in one-hot-vector form, into word embedding vectors; for each word-embedding vector, there is one output vector. BERT and its variants and Transformers, alone or in any combination with RNNs, are suitable for NLP tasks according to implementations herein.

**[0033]** Natural Language Processing (NLP) is the ability of a computer program to process and generate human language as spoken and/or written. In implementations, one or more recurrent neural network (RNN) is constructed to perform NLP (e.g., including text classification and text generation). In implementations, a neural network has layers, where each layer includes either input, hidden or output cells in parallel. In general two adjacent layers are fully connected (i.e., every neuron forms one layer to every neuron to another layer). For example, the network can have two input cells and one output cell, which can be used to model logic gates.

**[0034]** Augmented reality, which refers to a combination of a virtual environment and virtual reality, combines real-world images and virtual-world images such as computer graphic

images. Augmented reality is a semi-digital experience. In implementations of augmented reality, an image capturing device (e.g., a camera, a phone, a video recorder, etc.) that receives real images and a display device (e.g., a head mounted display that can display both real images and virtual images) are used together. Using augmented reality, a vehicle can be superimposed over a geographical location, for example, that can be associated with a particular date, time and/or weather condition. Lines can be drawn over an image of a vehicle to identify certain features and or parts, which may or may not be associated with a particular design type, time of history and/or cultural trend.

**[0035]** Virtual reality is a fully digital experience that creates an immersive environment where humans and computers can effectively interact and communicate by enhancing human-computer conversation skills using a variety of input and output techniques. Such techniques include the use of, for example, head-mounted displays, data gloves, or motion capture systems. These techniques receive data regarding variations in the position of a user by monitoring head, hand or other movements (e.g., position, directions, etc.), and transmit the data to a computer, which simulates (e.g., in a 3D coordinate space) the size and depth of an object within the viewing angle of the user.

**[0036]** Mixed reality refers to the merging of the real world with a virtual world to create a new environment where physical and digital objects interact with one another in real-time. In mixed reality, a real image can be captured using an image capturing device (e.g., a camera) and the direction the user faces within the environment is based on the captured real image. The relationship between the user's position and the position of a predetermined object is determined, and data obtained as a result of the calculation is displayed in a virtual space such that the data is laid over the captured real world image. Mixed reality is typically implemented using an image capturing device together with a display device.

**[0037] FIG. 1** schematically illustrates an example system 100 for image processing and data analysis, in accordance with one or more aspects of the present disclosure. As schematically illustrated by **FIG. 1**, the CNN 120 and optionally an RNN 122 together with a processing device 124 and a memory 126 form an media processing system 101. The media processing system 101 may be employed to process image data 110, video data 112, geolocation data 114 and input data 116 to produce an image classification result 130 and/or a virtual display result 132. The image data 110 may include one or more digital images, for example, captured by a camera or scanned, that may be stored in a memory. The video data 112 may include one or more digital videos, for example, captured by an audiovisual recording device or a dubbing device, that may be stored in a memory. The geolocation data 114 may include longitude,

latitude, country, region, historical or cultural place (e.g., Brooklyn Bridge), city, postal/zip code, time zone, way point, cell tower signal, etc. information from, for example, a global positioning system (GPS), entry to a user interface and/or other navigation system. The input data 116 may include structured data such as keyword search query input via a user interface and/or unstructured data input via the user interface. The unstructured data may include written or spoken natural language.

**[0038]** The CNN 120 may be employed to process the image data 110, the video data 112, the geolocation data 114 and the structured input data 116 to produce an image classification result 130 and/or a virtual display result 132, for example, with respect to vehicle information (e.g., a make, model, year, convertible sedan, sports utility vehicle or SUV, prototype, etc. ), vehicle artifacts (e.g., tools, a steering wheel, a lift, spokes, ets.) and/or a geographical location (e.g., a cultural site, a historical site, a landmark, etc.). The RNN 122 may process the unstructured data of the input data 116 (i.e., natural language processing) to produce structured data. The structured data may be fed from the RNN 122 to the CNN 120 for processing as described herein. In an illustrative example, the CNN 120 may correlate the image data 110, video data 112, geolocation data 114, input data 116 and structured data from the RNN 112 with images and data from a database (e.g., a propriety database of high quality copyrighted images and other works) in order to yield the probabilities of, for example, one or more image containing matching significant image features associated with a vehicle and/or a geographical location. The media processing system 101 may also return as part of the image classification result 130 historical, mechanical, cultural and other information with respect to the vehicle and/or geographical location.

**[0039]** The CNN 120 and the RNN 122 may be pre-trained using a comprehensive and precise training data set 140 that comprises *inter alia* non-published data, published data, images, videos, text (e.g., stories, news articles about various (e.g., thousands) vehicles, memoirs, out-of-print books, etc.) and/or geographical locations. For every vehicle or geographical location, the training data set may include multiple pixel-level, optionally annotated, vehicle images 142 and geographical images 146 and also may include related text 144 (e.g., histories, stories, descriptions, books, articles, drawings, sketches, etc.). The training data set 140 according to implementations herein may be a unique, comprehensive and proprietary body comprising copyrighted images, videos and text (e.g., history from the 20$^{th}$ Century surrounding vehicles, books off copyright, personal memoirs about vehicles, stories about racing, metals used, the brass era), built over many decades, that contains approximately 500,000 assets and verifies provenance through its records of timely copyright registrations at the Library of Congress. The

copyrighted works (i.e., authenticated data) identify *inter alia* automobiles and their impact on world culture including the time, place, historical context and significant background architecture captured in the media assets. Initially, the training data set 140 may be comprised in a copyrighted database where all of the assets contained within the training data set are copyrighted and thus authenticated. In further implementations, the training data set 140 may expand to include data input by users and further data assets that are not copyright registered, where such additional assets may be authenticated by other means whether scholarly (e.g., citations and research) or by using SNNs according to embodiments herein. The secondary twin will run a regression against the CNN. While the training data set 140 is comprehensive and precise when used in the systems and methods described herein, it will grow and evolve with more provenance authenticated data.

**[0040]** In certain implementations, training of the CNN 120 may involve activating the CNN 120 for every set of input images in the training dataset. The observed output (e.g., an image produced by the CNN 120) is compared with the desired output (e.g., the expected image) specified by the training data set, the error is calculated, and parameters of the CNN 120 are adjusted. This process is repeated until the output error is below a predetermined threshold.

**[0041]** In certain implementations, training of the RNN 122 may involve activating the RNN 120 for every set of unstructured data inputs in the training dataset. The observed output (e.g., a structured query produced by the CNN 120) is compared with the desired output (e.g., the expected query) specified by the training data set, the error may be calculated, and the parameters of the RNN 122 are adjusted accordingly. In implementations, this process may be repeated until the output error is below a predetermined threshold. In implementations using the RNN and deep learning models, the media processing system 101 may function and draw inferences from cross-related data.

**[0042]** The media processing system 101 may produce a virtual display result 132 in a virtual reality, augmented reality and/or mixed reality environment. In one illustrative example, one or more images, videos, descriptions, audio recordings, etc. may be layered onto an image captured by an image capture device (e.g., a camera, video recorder, etc.) and presented on a display. For example, a user may employ an image capture device (e.g., a cell phone) to capture an image in real time and the media processing system 101 may overlay or underlay images, videos and text onto the captured image as being viewed in an output device (e.g., a head-mounted display).

**[0043]** In one illustrative example, the CNN may be trained to identify automobiles. The media processing system 101 may process data 110, 112, 114, 116 to identify automobiles and preserve vehicle history by allowing a user to query the media processing system 101 to learn facts and

view photos of a specific vehicle. In implementations, the media processing system 101 may provide and provoke the curation of historical information surrounding the returned images (e.g., as a part of the image classification result 130). Multiple query types may be supported including, for example, photo uploads (CNN) and voice inputs (RNN) to query the models.

**[0044]** **FIG. 2** schematically illustrates an example structure of a CNN 120 that may be employed to process image data 110, video data 112, geolocation data 114 and input data 116 in order to produce an image classification result 130 and/or a virtual display result 132, in accordance with one or more aspects of the present disclosure. In certain implementations, acquired images may be pre-processed, e.g., by cropping, which may be performed in order to remove certain irrelevant parts of each frame. In an illustrative example, images having the resolution of 1024 x 1024 pixels may be cropped to remove 100-pixel wide image margins from each side of the rectangular image. In another illustrative example, a car may be outlined and isolated from noisy, non-contributory background elements.

**[0045]** As schematically illustrated by **FIG. 2**, the CNN 120 may include a first convolution layer 210A that receives image data 110 containing one or more images. The first convolution layer 210A is followed by squeeze layers 220A and 220B and a pooling layer 230, which is in turn followed by fully-connected layer 240 and a second convolution layer 210B. The second convolution layer 210B outputs one or more image 260 corresponding to the one or more input image of the image data 110 and may further produce the loss value 250 reflecting the difference between the produced data and the training data set. In certain implementations, the loss value may be determined empirically or set at a pre-defined value (e.g., 0.1).

**[0046]** In certain implementations, the loss value is determined as follows:

$$loss = \sum(x-y)^2\left(\frac{1}{2}+\max(x,y)\right),$$

where $x$ is the pixel value produced by the second convolution layer 210B and y is the value of the corresponding output image pixel.

**[0047]** Each convolution layer 210A, 210B may extract features from a sequence of input images from the input data 110, by applying one or more learnable pixel-level filters to a three-dimensional matrix representing the sequence of input images. The pixel-level filter may be represented by a matrix of integer values, which is convolved across the dimensions of the input image to compute dot products between the entries of the filter and the input image at each spatial position, to produce a feature map representing the responses of the first convolution layer 210A at every spatial position of the input image. In an illustrative example, the first convolution layer 210A may include 10 filters having the dimensions of 2 x 2 x 2. The second

convolution layer 210B may merge all the values produced by previous layers in order produce a matrix representing a plurality of image pixels.

**[0048]** **FIG. 3** depicts a flow diagram of one illustrative example of a method 300 of classifying and identifying input data, in accordance with one or more aspects of the present disclosure. Method 300 and/or each of its individual functions, routines, subroutines, or operations may be performed by one or more processors of the computer system (*e.g.*, system 100 and/or processing device 124 of **FIG. 1**) executing the method. In certain implementations, method 300 may be performed by a single processing thread. Alternatively, method 300 may be performed by two or more processing threads, each thread executing one or more individual functions, routines, subroutines, or operations of the method. In an illustrative example, the processing threads implementing method 300 may be synchronized (*e.g.*, using semaphores, critical sections, and/or other thread synchronization mechanisms). Alternatively, the processing threads implementing method 300 may be executed asynchronously with respect to each other.

**[0049]** At block 310, the processing device performing the method may train a CNN using authenticated data and a taxonomy. In implementations, the authenticated data may include copyright registered works of authorship including, but not limited to, copyrighted images, videos, text, stories, sketches, etc. The authenticated data may be stored in a database and the database itself may be copyright registered. The taxonomy may be used to classify and identify the data assets.

**[0050]** At block 320, the processing device may receive a query comprising input data. The input data can include, but is not limited to, image data, video data, intake data and/or geolocation data according to embodiments herein. In implementations, the intake data may be in the form of a keyword or string of text or may be in the form of unstructured data such as natural language either typed or spoken. The method may further include training an RNN to process the unstructured data of the intake data to form structured data suitable for processing by the CNN. The CNN may then process the structured data.

**[0051]** At block 330, the processing device may classify, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy. During the classification, the CNN may match features of the input data to one or more features of the authenticated data and/or elements of the taxonomy. For example, if the input data comprises an image, the CNN may scan the pixels of the image, identify features and then match the features with the closest matching features in the authenticated data and/or as classified in the taxonomy.

**[0052]** At block 340, the processing device may generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match

to the input data. For example, if five features have probabilities of 80%, 82%, 90%, 95% and 99% of matching five assets of the authenticated data, respectively, then the returned result may include only images with features having a 90% or greater probability of matching the input data.

**[0053]** At block 350, the processing device may display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content and/or mixed reality content. In implementations, the result may be layered with information. For example, a displayed image may be annotated with text, video and/or historical information about an object in the image.

**[0054]** In another illustrative example of a method 300 of classifying and identifying input data, in accordance with one or more aspects of the present disclosure. a processing device performing the method may process a training data set comprising a plurality of input images, in order to determine one or more parameters of a CNN to be employed for processing a plurality of images of one or more vehicle and/or geographical location. In various illustrative examples, the parameters of the CNN may include the convolution filter values and/or the edge weights of the fully-connected layer. In an illustrative example, the plurality of input images comprises one or more vehicle image. The one or more vehicle image may illustrate a vehicle alone or in combination with a geographical location (e.g., a Ford Model T on Route 66).

**[0055]** The processing device performing the method optionally may process a training data set comprising unstructured data in order to determine one or more parameters of a RNN to be employed for processing unstructured data input to the media processing system 101 in the form of natural language queries and voice queries to produce structured data for the CNN. In various illustrative examples, the RNN is trained to perform natural language processing using, for example, unstructured written and/or voice inputs.

**[0056]** The media processing system 101 may receive one or more of: a) image data including at least one input image (e.g., of a vehicle and/or a geographical location), b) video data including at least one input video (e.g., of a vehicle and/or a geographical location), c) input data including at least one of a keyword, a search query and unstructured data (e.g., relating to a vehicle and/or a geographical location), and d) geographical location data including a location of a device. In an illustrative example, the media processing system 101 may receive an image of an automobile alone, or together with a voice request saying "show me the artistic design features of this car."

**[0057]** The processing device performing the method optionally may process, by the RNN of the media processing system 101, any unstructured data of the input data that is received. The RNN outputs structured data that is fed to the CNN for processing. In the foregoing illustrative

example, the RNN may performing natural language processing of the voice request saying "show me the artistic design features of this car."

**[0058]** The processing device performing the method may process by the CNN of the media processing system 101, one or more of: i) the image data 110 to classify at least one input image (e.g., with respect to a vehicle information and/or a geographical location of a vehicle), ii) the video data 112 to classify at least one video (e.g., with respect to a vehicle information and/or a geographical location of a vehicle), iii) the structured input data 116 to classify at least one of a keyword or search query, iv) the structured data from the RNN (330), and v) the geographical location data 114 to produce one or more image, video and/or virtual display, as described in more herein. The probability of the image data, video data, geographical location data, input data and RNN data comprising the significant image features may be determined by a cross-entropy function, the error signal of which is directly proportional to a difference between desired and actual output values. In the foregoing illustrative example, the CNN may process the image of the automobile and the output of the RNN reflecting the voice request saying "show me the artistic design features of this car."

**[0059]** The processing device performing the method may generate a result by the media processing system including at least one of an image (e.g., of a vehicle and/or a geographical location), a video (e.g., of a vehicle and a geographical location), a history (e.g., of a vehicle and/or a geographical location) and/or other textual information. In the foregoing illustrative example, the media processing system 101 may generate an image of the automobile, alone or in combination with text providing the make, model and year of the automobile. The generated image may also be annotated with lines and text that identify artistic features of the automobile.

**[0060]** The processing device performing the method displays the result. The result may be displayed, for example, on a user device such as a cell phone, iPad, monitor or in a virtual device such as a head-mounted display of a virtual reality, augmented reality and/or mixed reality system.

**[0061] FIG. 4** depicts a flow diagram of one illustrative example of a method 360 of displaying a result, in accordance with one or more aspects of the present disclosure. Method 360 and/or each of its individual functions, routines, subroutines, or operations may be performed by one or more processors of the computer system (*e.g.*, system 100 and/or processing device 124 of **FIG. 1**) executing the method. In certain implementations, method 360 may be performed by a single processing thread. Alternatively, method 360 may be performed by two or more processing threads, each thread executing one or more individual functions, routines, subroutines, or operations of the method. In an illustrative example, the processing threads implementing

method 360 may be synchronized (*e.g.*, using semaphores, critical sections, and/or other thread synchronization mechanisms). Alternatively, the processing threads implementing method 300 may be executed asynchronously with respect to each other.

**[0062]** Example method 360 produces an augmented reality display on a display device. At block 410, the processing device performing the method determines a viewing direction of a user wearing and augmented reality apparatus, for example, a head-mounted display. The viewing direction may be determined by angles in relation to a center of the head set (e.g., looking forward) and head position.

**[0063]** At block 420, the processing device performing the method determines an attitude of an augmented reality apparatus using distances between the user and the augmented reality apparatus. The distances may be measured by one or more distance sensors.

**[0064]** At block 430, the processing device performing the method controls a direction of image input of the augmented reality apparatus based on the viewing direction of the user and the attitude of the augmented reality apparatus. In an illustrative example, the augmented reality apparatus may include a driving unit that adjusts the direction of, for example, a digital camera horizontally or vertically so that a subject (e.g., a vehicle) corresponding to a subject image incident upon the digital camera can be chosen even when the augmented reality apparatus is fixed.

**[0065]** At block 440, the processing device performing the method receives an image of one or more subjects (e.g., a vehicle) in the direction of image input. A camera or other device for recording and storing images may be used to capture the image.

**[0066]** At block 450, the processing device performing the method generates a synthesized image by synthesizing the image of the one or more subjects with a digital image. In an illustrative example, the synthesized image will layer images, videos and text with the image of the one or more subjects to produce an augmented reality environment.

**[0067]** At block 460, the processing device performing the method displays the synthesized image. The synthesized image may be displayed on an augmented reality apparatus such as a head-mounted display. In an illustrative example, a user may see the image of a car captured using the user's cell phone underlaid by a video of geographical locations around the world. The image additionally or alternatively may be annotated with text such as arrows that point out different features of the car. The display may also be accompanied by voice information and/or music, for example, an audio description (e.g., spoken by a human or a bot) of the history of the car.

**[0068] FIG. 5** schematically illustrates an example of the neural network architecture and data pipeline together with a cloud-based, microservices-driven architecture (collectively referred to as "the architecture") 500 for image processing and data analysis, in accordance with one or more aspects of the present disclosure. As schematically illustrated by **FIG. 5**, the architecture 500 includes a memory (not shown) and a database 510 (e.g., MongoDB©, Hbase©) configured for both in-memory and on-disk storage. The database 510 may include one or more trained machine learning models 512 for classifying and identifying images. In implementations, a storage or persistence layer may store images, metadata as a multidimensional cube warehouse, ML models, textual narratives, search-indexes, and software/applications underlying a transactional database. The architecture 500 may further include a plurality of containerized microservices 522A-C. In various example implementations, the runtime logic execution layer may be a collection of docker-container based microservices exposing representational state transfer (REST) application programming interfaces (APIs) (e.g., using Kubernetes18©). Each element 511, 513 and 515 represents a taxonomy feature that the corresponding microservice 522A, B and C is trained to analyze.

**[0069]** System 500 may further include a web application 532 including, for example, the media processing system 101 and system 100 and configured to execute methods 300 and 360. The web application 532 may be stored on a demilitarized zone (DMZ) network 530 to provide security. A virtual memory component 534 comprised of user comments and ratings may also be stored on the DMZ network 530 (i.e., these comments will not be added to the training data set 140 until authenticated). System 500 may further include a content delivery network 540, which is a content distribution network of proxy servers to ensure high availability of content. System 500 may further include a web presentation layer where one or more app users 550 can access the web app 532 and view results on a user device, for example, a cell phone 552A or a laptop 552B. In various example implementations, a presentation layer (e.g., ReactJS©) may be used for rendering the web-presentation layer (e.g., in 19 HTML5/CSS). The architecture may be implemented in a cloud or in a decentralized network (e.g., SOLID via Tim Berners-Lee).

**[0070]** The architecture 500 may further include digital walls 560, 562, 564 providing cybersecurity layers between the components of the architecture 500. Wall 560 may be implemented between the public web and application user devices 550 and the web application 532 and virtual component 534 in the DMZ 530. Wall 562 may be implemented between the DMZ 530 and microservices 520, 522A-C. Wall 564 may be implemented between the microservices 520 and the database 510.

**[0071]** According to various example implementations of the methods and systems described herein, CNN models may be implemented with multi-label classifiers to identify, for example, the make, mode, and year of manufacture of a vehicle. These classifiers may be implemented in, for example, TensorFlow® and Keras® using using ResNet, VGG-19 and/or Inception. In example implementations, these will feed into densely connected layers that predict into a region of an NLP embedding space. This embedding-space may then be used with NLP to identify relevant textual artifacts. According to example implementations, trained SNNs including CNNs may be used for vehicle authentication. The SNNs may use a contrastive loss function to compare a sample to a reference/fingerprint object.

**[0072]** According to various example implementations of the methods and systems described herein, RNN models may be implemented with LSTM, gated recurrent unit (GRU) and attention models. For example, the narratives users contribute, in addition those already curated as authentic history, will feed into NLP models based on RNN (LSTM/GRU) and Attention models like BERT, to assist a user in finding automobiles through descriptions. CNN-recognized objects and their associated meta-tags may play a role in the NLP results to map onto vehicles.

**[0073]** According to various implementations, the media processing system 101 may achieve greater than about 75% accuracy, or greater than about 80% accuracy, or greater than about 85% accuracy, or greater than about 90% accuracy, or greater than about 95% accuracy, or greater than about 99% accuracy when compared against the multi-label classifier. According to example implementations, an accuracy rate of greater than 90% may be achieved for cars that are more popular or common. For rarer cars, an accuracy rate of greater than about 80% may be achieved for vehicles that are less common or have limited production. In various example implementations, vehicle-clusters may be determined from broad descriptions. For example, when a user provides a broad description such as "Show me all Ford Mustangs," the media processing system 101 may provide a greater than 90% accuracy in identifying or recognizing the vehicle described when the text is sufficiently descriptive. For cultural heritage images, in various implementations the media processing system 101 may provide an accuracy of greater than about 80%, or greater than about 85%, or greater than about 90%, or greater than about 95%, or greater than about 99%. For example, the media processing system 101 may return a result with greater than about 80% probability of matching the query.

**[0074]** **FIG. 6** depicts a diagram of a server configuration 600 that may be employed in an example system for image processing and data analysis, in accordance with one or more aspects of the present disclosure. The server configuration 600 may be a computational cluster 610 (e.g., a Hadoop Cluster) having a master open source administration tool server and agent 612 (e.g., an

Ambari server and Ambari agent). The computational cluster 610 may further include a pair of slave agents 614A-B. A Hadoop cluster is a type of computational cluster designed to store and analyze large quantities of unstructured data in a distributed computing environment. Such clusters run Hadoop's open source distributed processing software on low-cost commodity computers. The cluster enables many computers to solve problems requiring massive computation and data.

**[0075] FIG.** 7 illustrates a diagrammatic representation of a machine in the example form of a computer system 700 including a set of instructions executable by systems as described herein to perform any one or more of the methodologies discussed herein. In one implementation, the system may include instructions to enable execution of the processes and corresponding components shown and described in connection with **FIGs. 1-6**.

**[0076]** In alternative implementations, the systems may include a machine connected (e.g., networked) to other machines in a LAN, an intranet, an extranet, or the Internet. The machine may operate in the capacity of a server machine in client-server network environment. The machine may be a personal computer (PC), a neural computer, a set-top box (STB), Personal Digital Assistant (PDA), a cellular telephone, a server, a network router, switch or bridge, or any machine capable of executing a set of instructions (sequential or otherwise) that specify actions to be taken by that machine. Further, while only a single machine is illustrated, the term "machine" shall also be taken to include any collection of machines that individually or jointly execute a set (or multiple sets) of instructions to perform any one or more of the methodologies described herein.

**[0077]** The example computer system 700 can include a processing device (processor) 702, a main memory 704 (e.g., read-only memory (ROM), flash memory, dynamic random access memory (DRAM) such as synchronous DRAM (SDRAM)), a static memory 706 (e.g., flash memory, static random access memory (SRAM)), and a data storage device 718, which communicate with each other via a bus 730.

**[0078]** Processing device 702 represents one or more general-purpose processing devices such as a microprocessor, central processing unit, or the like. More particularly, the processing device 702 may be a complex instruction set computing (CISC) microprocessor, reduced instruction set computing (RISC) microprocessor, very long instruction word (VLIW) microprocessor, or a processor implementing other instruction sets or processors implementing a combination of instruction sets. The processing device 702 may also be one or more special-purpose processing devices such as an application specific integrated circuit (ASIC), a field programmable gate array (FPGA), a digital signal processor (DSP), network processor, or the like. In various implementations of the present disclosure, the processing device 702 is configured to execute

instructions for the devices or systems described herein for performing the operations and processes described herein.

**[0079]** The computer system 700 may further include a network interface device 708. The computer system 700 also may include a video display unit 710 (e.g., a liquid crystal display (LCD) or a cathode ray tube (CRT)), an alphanumeric input device 712 (e.g., a keyboard), a cursor control device 714 (e.g., a mouse), and a signal generation device 716 (e.g., a speaker).

**[0080]** The data storage device 718 may include a computer-readable medium 728 on which is stored one or more sets of instructions of the devices and systems as described herein embodying any one or more of the methodologies or functions described herein. The instructions may also reside, completely or at least partially, within the main memory 704 and/or within processing logic 726 of the processing device 702 during execution thereof by the computer system 700, the main memory 704 and the processing device 702 also constituting computer-readable media.

**[0081]** The instructions may further be transmitted or received over a network 720 via the network interface device 708. While the computer-readable storage medium 728 is shown in an example implementation to be a single medium, the term "computer-readable storage medium" should be taken to include a single medium or multiple media (e.g., a centralized or distributed database, and/or associated caches and servers) that store the one or more sets of instructions. The term "computer-readable storage medium" shall also be taken to include any medium that is capable of storing, encoding or carrying a set of instructions for execution by the machine and that cause the machine to perform any one or more of the methodologies of the present disclosure. The term "computer-readable storage medium" shall accordingly be taken to include, but not be limited to, solid-state memories, optical media, and magnetic media.

**[0082]** In various example implementations described herein, the neural networks are supervised learning models that analyze input data to classify objects, for example, using regression analysis. In an example implementation, a user may upload an image and the media processing system regresses it against many elements to determine the closest match. The supervised learning models are trained using different high-level elements of a taxonomy. The elements are related to categories of the taxonomy wherein the categories are used with ML to train the neural network models. In certain implementations, the elements may include, but are not limited to: actions (e.g., driving), concepts and emotions (e.g., direction), events (e.g., 2007 Tokyo Motor Show), geographic city (e.g., Los Angeles), geographic country (e.g., U.S.A.), geographic places (e.g., LAX Airport), geographic state (e.g., California), geographic location data (e.g., from a GPS), museum collections (e.g., Petersen Automotive Museum), photo environments (e.g., night), photo orientations (e.g., landscape), photo settings (e.g., auto garage), photo techniques (e.g., color),

photo views (e.g., three-quarter front view), signs (e.g., bowling alley), topic subjects (e.g., American culture), vehicle coachbuilder (e.g., Brewster & Co.), vehicle color (e.g., green), vehicle condition (e.g., new), vehicle manufacturer (e.g., including country and continent of origin), vehicle model (e.g., Bentley 6½ Liter Speed Six Tourer), vehicle parts (e.g., 8-track cassette player), vehicle quantity (e.g., one object), vehicle serial number (e.g., Chassis 3559SA), vehicle type (e.g., hydrogen fuel cell) and vehicle year of manufacture (e.g., 1957).

**[0083]** Implementations described herein can preserve and reveal information about vehicles and their impact on society. Using machine learning (ML) algorithms trained upon a proprietary dataset of curated and authenticated photos (image data), videos (video data), input data (e.g., text or voice inputs) and geolocation data, an artificial intelligence (AI) platform (e.g., including one or more convolutional neural network and one ore more recurrent neural network) has been developed that can *inter alia* identify vehicles from 1885 through present day and engages users to annotate images of the vehicles by sharing stories and comments. Machine learning includes, but is not limited to algorithms that find and apply patterns in data. Neural networks can be a form of ML. Implementations described herein provide a kind of time machine chassis capturing alchemical memories of shaped metal propelled through time and space, then identified through a multi-layered neural network. Enabling society to easily access information about vehicles through a searchable media processing system as described herein, augments and preserves human narrative, future transportation solutions and the history of remarkable vehicles.

**[0084]** In implementations, the database upon which the neural network is trained is a transdisciplinary study where abstract concepts (e.g., emotional, verbal, spatial, logical, artistic and social) are represented by semantic keywords expressing different dimensions of intelligence present in the referenced media object. For example, "Art Deco" is a semantic artistic keyword found on numerous vehicles from the 1920s and 1930s due to the visual design-language shown on a particular car or artifact. Using deep learning, the neural networks as described herein can be repeatedly trained for each of these distinct conceptual layers of intelligence found in the media thus resulting in object recognition enhanced through semantic intelligence and linking back to society. Additional databases including engineering information, racing results, car shows, location a car has been exhibited, valuations, etc. can be layered upon the proprietary vehicle database to further enhance information relating to the vehicles. In implementations, users can, for example, page through every Chevrolet Corvette in the library archives and read or listen to entries associated with any vehicle. Similarly, in implementations, a user can experience the development of streamlining or see the condensed vehicle design language of a particular decade of time, e.g. The Fifties. In implementations, a user can hold up a mobile device to an interesting car on the street and learn its story through interaction with the media processing system. For example, the media processing system may be configured to return facts such as "Did you know the V8 engine was invented one hundred years ago?"

**[0085]** Implementations described herein are configured to recognize and identify vehicles input via multiple sensory inputs (e.g., voice, image and text) and record a user's personal stories about the provenance and significance of vehicles, for example, in telling the story of America. Families can upload shoeboxes of family photos to learn what vehicle a great-grandfather once drove to the Grand Canyon. A user can travel to a historic place site and through the media processing system view hundreds of vehicles and families that previously visited the site over preceding decades. In implementations, family vacation photographs recording an event can be layered upon an existing geographic location to create a virtual environment, for example, using an immersive augmented reality (AR). In implementations, the AR environment can enable a user to see herself or himself, along with his or her ancestors and their vehicles at the same cultural heritage site evoking a ghostly rapture within the time-space continuum. For example, "How many photographs with the family car were taken at the Golden Gate Bridge?"

**[0086]** According to implementations, a proprietary database of images of vehicles contains several labeled images that belong to different categories including, but not limited to "make," "model" and "year." "Vehicle" refers to a mechanism for transporting things (e.g., people and goods), including, but not limited to planes, trains, automobiles (e.g., cars, trucks, motorcycles, vans, etc.) and spacecraft. The more images used for each category, the better the model (e.g., the convolutional neural network) can be trained to determine whether an image is, for example, a Porsche image or a Ferrari image. This implementation utilizes supervised machine learning. The model can then be trained using the labeled known images. The images in their extracted forms enter the input side of the model and the labels are in the output side. The purpose is to train the model such that an image with its features coming from the input will match the label in the output. Once the model is trained, it can be used to recognize, classify and/or predict an unknown image. For example, a new image may be recognized, classified or predicted as a 1933 Packard Twelve. As part of the processing by the convolutional neural network, the newly input image also goes through the pixel feature extraction process.

**[0087]** Implementations disclosed herein can address a fundamental problem in the advertising industry: how to authenticate a users' interest in a particular automotive brand or vehicle model or type. By virtue of a user uploading an unidentified Alfa Romeo to the platform, the user self-identifies the user's interest in this vehicle brand. Through repeated interactions, the media processing system learns the user's interests in vehicle information. Such a feedback loop is

valuable to advertisers for better targeting and in turn, can provide intelligence to manufacturers of future vehicles. The proprietary database of vehicles according to implementations, may be authenticated through timely registrations at the Library of Congress Copyright Office, which provides provenance that is preserved in the ML training dataset. In implementations, the training data set 140 may grow including additional data assets, for example, based on data input by users and/or additional assets that may not be copyright registered, but that may be authenticated, for example, by using SNNs as described herein.

**[0088]** Methods, systems and computer program products according to implementations herein relate to an AI platform built by systematically applying natural language processing (NLP) and computer vision, image and video processing to train a convolution and recurrent neural network from a dataset containing high quality, digital images, which may be copyrighted, of automobiles, capable of identifying a particular automobile from about 1885 through present day and into the future.

**[0089]** The essence of being human is to ask questions and AI seeks to provide credible information about a technological evolution: the journey of vehicles (e.g., the automobile), as well as the remaining surrounding artifacts of our vehicle heritage populating culture today. The implementations described herein provide an innovative AI-driven platform for classifying and documenting vehicles. Users can engage in a feedback cycle centered around identified photos, stories, comments, family photos and records. At their core, implementations herein nurture and explore the singular relationship of humans to machines, preserving the bond between the vehicles, design, community, architecture, engineering, history, government and culture to disseminate knowledge about vehicle culture. If a vehicle or important vehicle cultural heritage artifact is unknown, the platform can use the wisdom of crowds to identify and classify the asset. By NLP, the AI agent can begin chat interactions with the user about vehicles and immersive environments shown in the media thus deepening human-computer interaction skills.

**[0090]** Implementations described herein provide for the preservation and accessibility of collated, correlated and curated historical information (e.g., images, video, media, text, unstructured data, geographical location data, etc.) about and concerning vehicles, their use in human society, the environments (e.g., geographical locations) in which they are or have been used (e.g., racing and street), how they are or have been used, jobs they create or have created (e.g., in manufacturing and maintenance, consumer uses, collectors, etc.), technical and design features and special relationships with society.

**[0091]** According to implementations, multi-dimensional inputs query for vehicle attributes and elements from a vehicle dataset trained via ML from a proprietary reference database (e.g., a

neural network) built-upon copyright-registered and authenticated intellectual property about vehicles and their environments from the 1880s through present day and into the future to provide AI in mixed reality applications. Information may be retrieved from implementations described herein using multiple inputs including, but not limited to, audio, text, video, images and global positioning system (GPS) inputs where the input request is referenced against a proprietary-trained vehicle dataset and returns a classification and/or match to the input request in mixed-reality environments. Queries about vehicles can be answered with a probability of correct identification. For example, a user can type into the media processing system: “Auto Union Wanderer W-25” and the system would interpret the words to return an image of the “Auto Union Wanderer W-25.” The probability of the queried vehicle being built by “Auto Union” can be expressed as a percentage, for example, a “95%” probability of the image, video, text, history, etc. returned is an Auto Union and the probability of the image, video, text, history, etc. returned being a model “Wanderer W-25,” for example, “85%.” According to implementations, a short history of a vehicle appears and, using the geolocation services in the media processing system, an identification of where the closest example of this vehicle may be physically available for review relative to the user’s present location.

**[0092]** In implementations, information can be retrieved by uploading to the system (e.g., an app on a cell phone, a website, etc.), via a user interface, a photograph (e.g., a digital image, a scanned image, etc.) of a vehicle the user may encounter in daily life (e.g., on the street, at a car show, etc.). Using image recognition derived from the proprietary database, the media processing system can return, for example, a matching identification and/or classification of a vehicle make, model and year of release (referred to herein as “year”) with probabilities of accuracy based upon the trained dataset rendered through machine learning.

**[0093]** According to implementations described herein, users dictate voice request inputs, queries, text inputs and/or natural language to the media processing system. The input data can include, but is not limited to the make, model and year of a vehicle. For example, a user can speak “show me Ford Thunderbird” into a microphone that inputs data to the media processing system, which returns at least one of an image, a video, a written history, etc. representing the closest match to the “Ford Thunderbird” with additional information provided through mixed reality inputs. The user may refine a query by speaking “show me “red 1957 Ford Thunderbird” and the media processing system would return one or more image having the closest match together with a probability of accuracy. Existing platforms do not have provenance in their training data sets or information repositories (i.e., databases) and so the results returned by such platforms may be inaccurate or incomplete. Such platforms may scrape unauthenticated

information (e.g., not copyright registered, not validated with date, author, metadata, etc.) from publicly available websites such that if a certain number of users say something is, for example, a Thunderbird, then the platform will agree. However, if an asset or image is merely a replica of a Thunderbird, then such information is missed in these existing platforms. In another example, multiple years of Honda Accords may not be able to be recognized with authority by existing platforms.

**[0094]** According to implementations, a user initiates a query by pointing an input device (e.g., a camera, a video recorder, etc.) at a vehicle of interest and the media processing system receives the at least one input image and/or input video and matches it against the ML trained dataset to provide an Augmented Reality (AR) display of information regarding the queried vehicle. Levels of information may be chosen via user setup. For example, a user may only require vehicle make, such as "Ferrari," or may require vehicle make and model, such as "Ferrari 250 GT," or may require technical information like engine type, such as "V-8 engines;" the application is configured to return images, videos and/or information that matches V-8 engines from the neural network of information. According to implementations, in each query result, additional educational information about the vehicle is provided depending upon user settings. A brief history of the car can be displayed or overlaid in a mixed reality environment. For example, a user can submit a text or natural language input query, such as "two-tone vehicle interiors," and matches to the requisition can be displayed on the user device with overlaid text depicting, for example, the make, model, history, design features, etc. of vehicles having two-tone interiors.

**[0095]** The implementations described herein are useful in a variety of fields, for example, where entities may need media assets or user information on vehicles obtainable from an organized, curated, and searchable platform. Example fields include, but are not limited to 1) advertisers: any automobile-related business or ads in which cars appear need authenticated product; 2) automobile manufacturers: marketing need for brand building, loyalty and heritage promotion of manufacturer's products/services; 3) insurers: verifying vehicles is key in protecting assets and individuals; 4) entertainment: immersive experiences through augmented/virtual reality, skill games; 5) law enforcement: need help in identification of vehicles involved in investigations, possibly from photos taken at/from a crime scene—e.g., by a bystander on his/her cell phone—and fraud detection; 6) vehicle designers: need access to historical examples and perspective for new designs; 7) travel: roadside support, fuel, lodging, food, interesting roads and points of interest along the roadside; 8) classic car market and collectors: buyers, sellers and restorers of vehicles need parts authenticity, provenance

information, special features, and historical context in which the cars appear and are used; and 9) museums and archives: need help with identification, provenance and automotive history found in photo collections.

**[0096]** User interest, expressed by uploads of unidentified photos and by time spent reviewing certain vehicle brand archive sections to the media processing system, self-identifies the user's interest in specific vehicle brands and/or segments that can be sought after targets for advertisers. For example, a user who reads and peruses the Porsche archives is a good target for Porsche brand advertising.

**[0097]** By interacting with the platform, users self-identify interest in a particular automotive brand or vehicle sector thus solving an advertising problem for customers who wish to learn from past automotive designs, verify their illustrated marketing materials and target their communications to potential buyers in the automotive sector of our economy. Users may explore curated information about automobiles and roadside heritage through a virtual library linking other datasets to form a central integrated intelligence platform about automobiles and society. Transportation designers can easily access lessons learned from the last 135 years of automotive design.

**[0098]** Geolocation data, alone or in combination with curated photos of architectural and/or cultural heritage sites, can also be input to the media processing system as described herein. In certain implementations, the application can direct users to roadside services based on personalized user data (e.g., the user's preferred fuel type, fast food and hotel preferences can all be stored) and geolocation data received from a navigation system. For example, suppose a user drives a sports car, which is known from the user's profile stored in a memory accessible by the media processing system. The media processing system may have access to the user's calendar also stored on a memory accessible by the media processing system. The media processing system can receive an input from a navigation program indicating that the user's arrival at a calendar appointment location is estimated for 15 minutes, but there is a great two-lane road the user could drive that would be fun and would still get the user there on time for the calendar appointment. The media processing system would then suggest the alternate route.

**[0099]** Families throughout history have authenticated life in photographs, which may have been taken while on vacations using vehicles. Such image data can be used to virtually augment cultural heritage sites or historic places using historic photos including vehicles. For example, according to various implementations using augmented reality, virtual reality and/or mixed reality, the media processing system can enable users to virtually tour Route 66 throughout history. In implementations, the systems and methods described herein can use augmented,

virtual and/or mixed reality to enhance travel and road trips. For example, a user may drive down Route 66 and, using geolocation data, hold up a device (e.g., a cell phone) and see the present location as it evolved through history within a virtual display device (e.g., a head mounted device).

**[0100]** According to implementations, augmented reality relating to vehicles can be used for cultural heritage tourism to enhance the tourist experience. Linking contextual information found in the backgrounds of family photos, provides the groundwork for creating an authenticated augmented reality system, for example, for America's historic places. For example, implementations described herein are useful for auto clubs such as AAA. A mobile device can be pointed at a vehicle and/or cultural heritage location to capture image data and/or geolocation data and the media processing system can return images of the vehicle and/or cultural heritage location over time at that particular location.

**[0101]** Embodiments of the present disclosure can be described in view of the following clauses. In clause 1, a method comprises: training a convolutional neural network (CNN) using authenticated data and a taxonomy; receiving, by a processing device, a query comprising input data; classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy; generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

**[0102]** In clause 2, the method of clause 1, wherein the authenticated data comprises copyright registered works of authorship, metadata and text. In clause 3, the method of clause 2, wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings. In clause 4, the method of clause 3, wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information. In clause 5, the method of clause 1, wherein the authenticated data comprises data from a copyright registered database. In clause 6, the method of clause 1, wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture. In clause 7, the method of clause 1, wherein the input data comprises one or more of image data,

video data, intake data or geographical location data. In clause 8, the method of clause 1, wherein classifying comprises mapping input data to authenticated data using the taxonomy. In clause 9, the method of clause 1, wherein the result comprises one or more of an image, a video, text, or sound. In clause 10, the method of clause 1, wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information.

**[0103]** In clause 11, the method of clause 1, wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy. In clause 12, the method of clause 11, wherein the probability is determined by a cross-entropy function. In clause 13, the method of clause 1, wherein the result comprises augmented reality content, wherein displaying the result comprises: displaying the result in an augmented reality apparatus, comprising: passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device; and displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data. In clause 14, the method of clause 13, wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time. In clause 15, the method of clause 1, further comprising training a recurrent neural network (RNN) using authenticated data and a taxonomy. In clause 16, the method of clause 15, wherein the input data comprises unstructured data, the method further comprising: processing, by the trained RNN, the unstructured data to yield structured data; and classifying, by the trained CNN, the structured data. In clause 17, the method of clause 1, wherein the input data comprises user uploaded data, the method further comprising authenticating the user uploaded data using SNNs and adding the authenticated user uploaded data to the authenticated data.

**[0104]** In clause 18, a system comprising: a memory; a processor, coupled to the memory, the processor configured to: train a convolutional neural network (CNN) using authenticated data and a taxonomy; receive, by a processing device, a query comprising input data; classify, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy; generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content. In clause 19, the system of clause

18, wherein the authenticated data comprises copyright registered works of authorship, metadata and text. In clause 20, the system of clause 19, wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings. In clause 21, the system of clause 20, wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information. In clause 22, the system of clause 18, wherein the authenticated data comprises data from a copyright registered database. In clause 23, the system of clause 18, wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture. In clause 24, the system of clause 18, wherein the input data comprises one or more of image data, video data, intake data or geographical location data. In clause 25, the system of clause 18, wherein classifying comprises mapping input data to authenticated data using the taxonomy. In clause 26, the system of clause 18, wherein the result comprises one or more of an image, a video, text, or sound. In clause 27, the system of clause 18, wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information. In clause 28, the system of clause 18, wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy. In clause 29, the system of clause 28, wherein the probability is determined by a cross-entropy function. In clause 30, the system of clause 18, wherein the result comprises augmented reality content, wherein displaying the result comprises: displaying the result in an augmented reality apparatus, comprising: passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device; and displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data. In clause 31, the system of clause 30, wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time. In clause 32, the system of clause 18, further configured to train a recurrent neural network (RNN) using authenticated data and a

taxonomy. In clause 33, the method of clause 32, wherein the input data comprises unstructured data, the method further comprising: processing, by the trained RNN, the unstructured data to yield structured data; and classifying, by the trained CNN, the structured data. In clause 34, the method of clause 18, wherein the input data comprises user uploaded data, wherein the system is further configured to authenticate the user uploaded data using SNNs and add the authenticated user uploaded data to the authenticated data.

**[0105]** In clause 35, a computer-readable non-transitory storage medium comprising executable instructions that, when executed by a computing device, cause the computing device to perform operations comprising: training a convolutional neural network (CNN) using authenticated data and a taxonomy; receiving, by a processing device, a query comprising input data; classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy; generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content. In clause 36, the computer-readable non-transitory storage medium of clause 35, wherein the authenticated data comprises copyright registered works of authorship, metadata and text. In clause 37, the computer-readable non-transitory storage medium of clause 36, wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings. In clause 38, the computer-readable non-transitory storage medium of clause 37, wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information. In clause 39, the computer-readable non-transitory storage medium of clause 35, wherein the authenticated data comprises data from a copyright registered database. In clause 40, the computer-readable non-transitory storage medium of clause 35, wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture. In clause 41, the computer-readable non-transitory storage medium of clause 35, wherein the input data comprises one or more of image data, video data, intake data or geographical location data. In clause 42, the computer-readable non-transitory storage medium of clause 35, wherein classifying comprises mapping input data to authenticated data using the

taxonomy. In clause 43, the computer-readable non-transitory storage medium of clause 35, wherein the result comprises one or more of an image, a video, text, or sound. In clause 44, the computer-readable non-transitory storage medium of clause 35, wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information. In clause 45, the computer-readable non-transitory storage medium of clause 35, wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy. In clause 46, the computer-readable non-transitory storage medium of clause 45, wherein the probability is determined by a cross-entropy function. In clause 47, the computer-readable non-transitory storage medium of clause 35, wherein the result comprises augmented reality content, wherein displaying the result comprises: displaying the result in an augmented reality apparatus, comprising: passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides; imaging the light at the display device; and displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data. In clause 48, the computer-readable non-transitory storage medium of clause 47, wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time. In clause 49, the computer-readable non-transitory storage medium of clause 35, further comprising training a recurrent neural network (RNN) using authenticated data and a taxonomy. In clause 50, the computer-readable non-transitory storage medium of clause 49, wherein the input data comprises unstructured data, the method further comprising: processing, by the trained RNN, the unstructured data to yield structured data; and classifying, by the trained CNN, the structured data. In clause 51, the computer-readable non-transitory storage medium of clause 35, wherein the input data comprises user uploaded data, the method further comprising authenticating the user uploaded data using SNNs and adding the authenticated user uploaded data to the authenticated data. In clause 52, the medium of clause 51, wherein processing by the CNN at least one of the image data, the video data, the input data and the geographical location data yields one or more image of a vehicle comprising matching features.

**[0106]** The preceding description sets forth numerous specific details such as examples of specific systems, components, methods, and so forth, in order to provide a good understanding of several implementations of the present disclosure. It will be apparent to one skilled in the art, however,

that at least some implementations of the present disclosure may be practiced without these specific details. In other instances, well-known components or methods are not described in detail or are presented in simple block diagram format in order to avoid unnecessarily obscuring the present disclosure. Thus, the specific details set forth are merely presented as examples. Particular implementations may vary from these example details and still be contemplated to be within the scope of the present disclosure. In the above description, numerous details are set forth.

**[0107]** It will be apparent, however, to one of ordinary skill in the art having the benefit of this disclosure, that implementations of the disclosure may be practiced without these specific details. In some instances, well-known structures and devices are shown in block diagram form, rather than in detail, in order to avoid obscuring the description.

**[0108]** Some portions of the detailed description are presented in terms of algorithms and symbolic representations of operations on data bits within a computer memory. These algorithmic descriptions and representations are the means used by those skilled in the data processing arts to most effectively convey the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities. Usually, though not necessarily, these quantities take the form of electrical, magnetic, or optical signals capable of being stored, transferred, combined, compared, and otherwise manipulated. It has proven convenient at times, principally for reasons of common usage, to refer to these signals as bits, values, elements, symbols, characters, terms, numbers, or the like.

**[0109]** It should be borne in mind, however, that all of these and similar terms are to be associated with the appropriate physical quantities and are merely convenient labels applied to these quantities. Unless specifically stated otherwise as apparent from the above discussion, it is appreciated that throughout the description, discussions utilizing terms such as "rating," "selecting," "comparing," "adjusting," or the like, refer to the actions and processes of a computer system, or similar electronic computing device, that manipulates and transforms data represented as physical (e.g., electronic) quantities within the computer system's registers and memories into other data similarly represented as physical quantities within the computer system memories or registers or other such information storage, transmission or display devices.

**[0110]** Implementations of the disclosure also relate to an apparatus for performing the operations herein. This apparatus may be specially constructed for the required purposes, or it may comprise a general purpose computer selectively activated or reconfigured by a computer program stored in the computer. Such a computer program may be stored in a computer readable storage medium, such as, but not limited to, any type of disk including floppy disks, optical disks, CD-ROMs, and

magnetic-optical disks, read-only memories (ROMs), random access memories (RAMs), EPROMs, EEPROMs, magnetic or optical cards, or any type of media suitable for storing electronic instructions.

**[0111]** The algorithms and displays presented herein are not inherently related to any particular computer or other apparatus. Various general purpose systems may be used with programs in accordance with the teachings herein, or it may prove convenient to construct a more specialized apparatus to perform the required method steps. The required structure for a variety of these systems will appear from the description below. In addition, the present disclosure is not described with reference to any particular programming language. It will be appreciated that a variety of programming languages may be used to implement the teachings of the disclosure as described herein.

**[0112]** It is to be understood that the above description is intended to be illustrative, and not restrictive. Many other implementations will be apparent to those of skill in the art upon reading and understanding the above description.

## CLAIMS

I/We claim:

1. A method comprising:

training a convolutional neural network (CNN) using authenticated data and a taxonomy;

receiving, by a processing device, a query comprising input data;

classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy;

generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and

displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

2. The method of claim 1, wherein the authenticated data comprises copyright registered works of authorship, metadata and text.

3. The method of claim 2, wherein the copyright registered works of authorship comprise one or more of images, video recordings, audio recordings, illustrations or writings.

4. The method of claim 3, wherein the copyright registered works of authorship comprise one or more of vehicle information, geographical information or cultural information.

5. The method of claim 1, wherein the authenticated data comprises data from a copyright registered database.

6. The method of claim 1, wherein the elements of the taxonomy are selected from the group consisting of actions, concepts and emotions, events, geographic cities, geographic countries, geographic places, geographic states, geographic location data, museum collections, photo environments, photo orientations, photo settings, photo techniques, photo views, signs, topic subjects, vehicle coachbuilder, vehicle colors, vehicle conditions, vehicle manufacturers, vehicle models, vehicle parts, vehicle quantities, vehicle serial numbers, vehicle type and vehicle year of manufacture.

7. The method of claim 1, wherein the input data comprises one or more of image data, video data, intake data or geographical location data.

8. The method of claim 1, wherein classifying comprises mapping input data to authenticated data using the taxonomy.

9. The method of claim 1, wherein the result comprises one or more of an image, a video, text, or sound.

10. The method of claim 1, wherein generating the result yields one or more of vehicle information, vehicle artifact information or geographical information.

11. The method of claim 1, wherein generating the result yields a probability of the input data matching at least one feature of the authenticated data or of at least one element of the taxonomy.

12. The method of claim 11, wherein the probability is determined by a cross-entropy function.

13. The method of claim 1, wherein the result comprises augmented reality content, wherein displaying the result comprises:

displaying the result in an augmented reality apparatus, comprising:

passing light into an eye of a wearer of an augmented reality display device, said augmented reality display device comprising a light source and a waveguide stack comprising a plurality of waveguides;

imaging the light at the display device; and

displaying on the display device a vehicle alone or in combination with a geographical location and optionally, on a particular date, that has matching features to at least one of the image data, the video data, the input data and the geographical data.

14. The method of claim 13, wherein displaying on the display device comprises at least one of displaying how the geographical location has changed over time, displaying history of vehicles that have passed through the geographical location over time, displaying weather conditions over a period of time.

15. The method of claim 1, further comprising training a recurrent neural network (RNN) using authenticated data and a taxonomy.

16. The method of claim 15, wherein the input data comprises unstructured data, the method further comprising:

processing, by the trained RNN, the unstructured data to yield structured data; and

classifying, by the trained CNN, the structured data.

17. The method of claim 1, wherein the input data comprises user uploaded data, the method further comprising authenticating the user uploaded data using Siamese Neural Networks and adding the authenticated user uploaded data to the authenticated data.

18. A system comprising:

a memory;

a processor, coupled to the memory, the processor configured to:

train a convolutional neural network (CNN) using authenticated data and a taxonomy;

receive, by a processing device, a query comprising input data;

classify, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy;

generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and

display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

19. The system of claim 18, wherein the authenticated data comprises copyright registered works of authorship, metadata and text.

20. A computer-readable non-transitory storage medium comprising executable instructions that, when executed by a computing device, cause the computing device to perform operations comprising:

training a convolutional neural network (CNN) using authenticated data and a taxonomy;

receiving, by a processing device, a query comprising input data;

classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy;

generating a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closest match to the input data; and

displaying the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content.

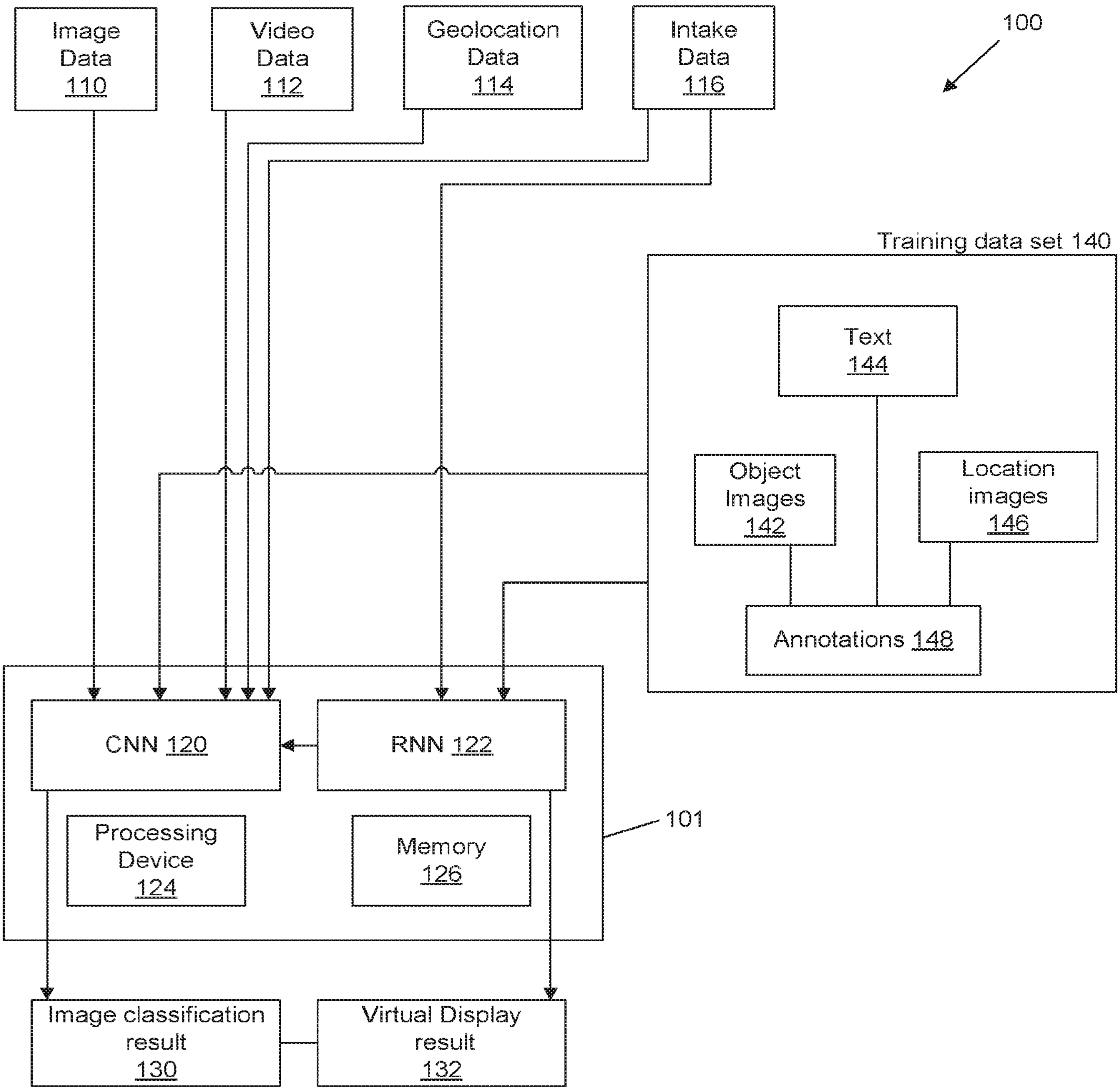


Image Data 110
Video Data 112
Geolocation Data 114
Intake Data 116
100
Training data set 140
Text 144
Object Images 142
Location images 146
Annotations 148
CNN 120
RNN 122
Processing Device 124
Memory 126
101
Image classification result 130
Virtual Display result 132

FIG. 1

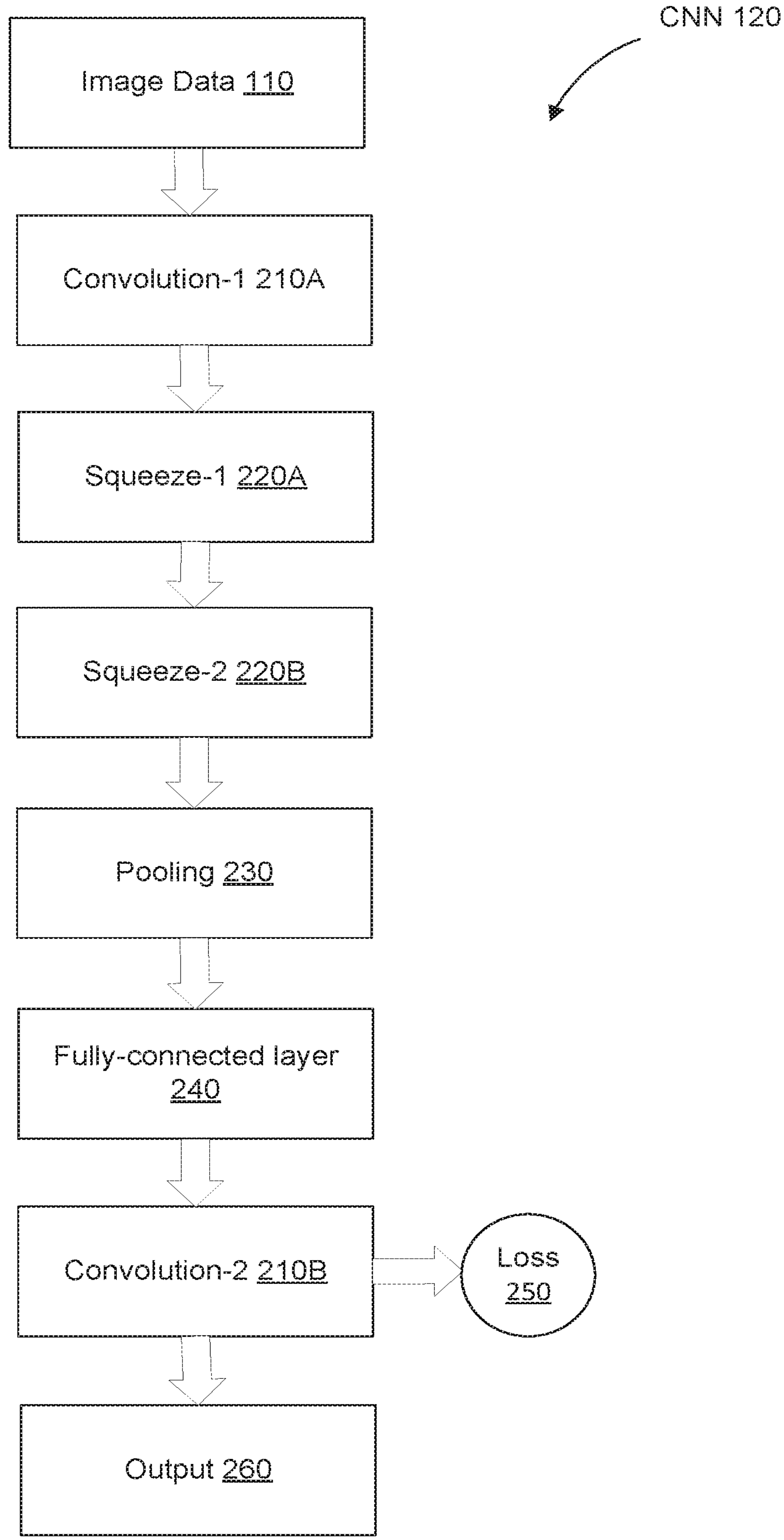

CNN 120
Image Data 110
Convolution-1 210A
Squeeze-1 220A
Squeeze-2 220B
Pooling 230
Fully-connected layer 240
Convolution-2 210B
Loss 250
Output 260

FIG. 2

300

310 Training a CNN using authenticated data and a taxonomy

↓

320 Receiving, by a processing device, a query comprising input data.

↓

330 Classifying, by the trained CNN, the input data with respect to the authenticated data and elements of the taxonomy

↓

340 Generate a result, by the trained CNN, wherein the result comprises authenticated data and elements of the taxonomy comprising a closes match to the input data

↓

350 Display the result on a device, wherein the result comprises one or more of an image, a video, text, sound, augmented reality content, virtual reality content or mixed reality content

FIG. 3

350 Display

410 Determine a viewing direction of a user

420 Determine an attitude of an augmented reality apparatus using distances between the user and the augmented reality apparatus measured by at least two distance sensors

430 Control a direction of image input of the augmented reality apparatus based on the viewing direction of the user and the attitude of the augmented reality apparatus

440 Capture an image of one or more subjects in the direction of image input

450 Generate a synthesized image by synthesizing the image of one or more subjects with a digital image

460 Display the synthesized image

FIG. 4

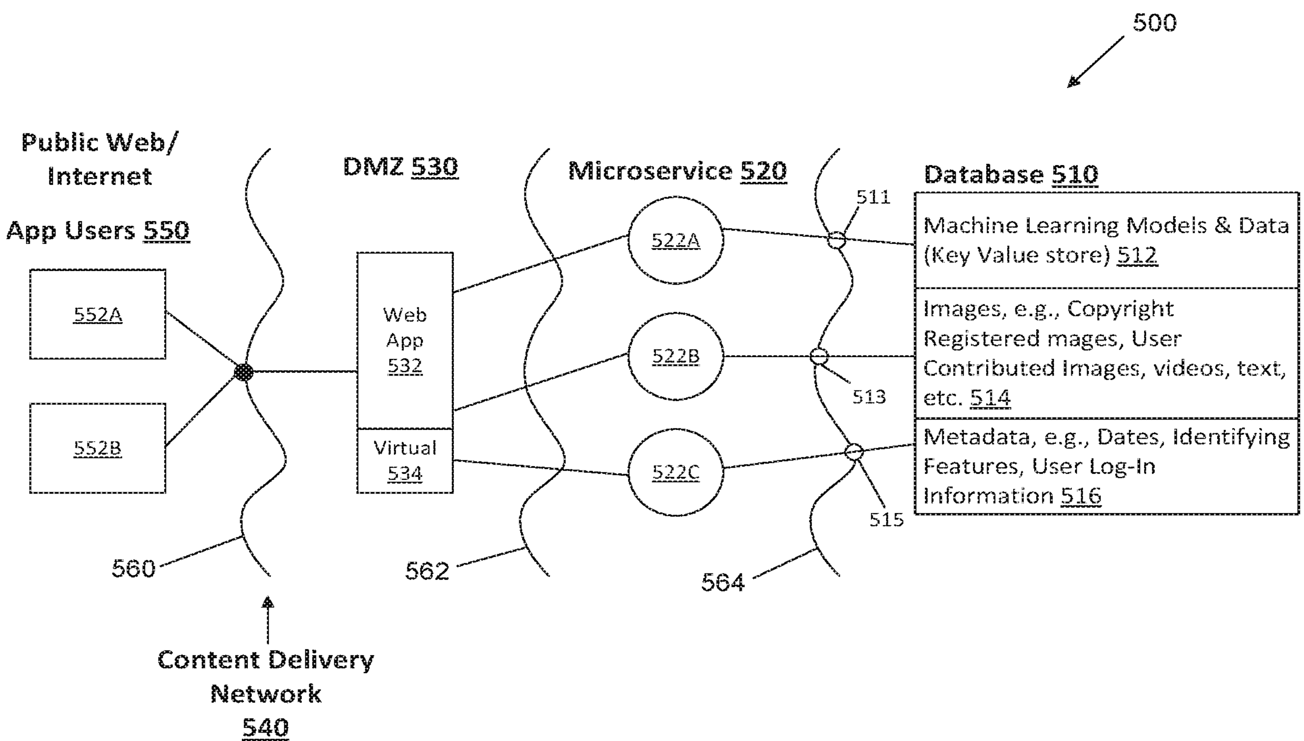

500
Public Web/ Internet
App Users 550
552A
552B
DMZ 530
Web App 532
Virtual 534
Microservice 520
522A
522B
522C
Database 510
511
513
515
Machine Learning Models & Data (Key Value store) 512
Images, e.g., Copyright Registered mages, User Contributed Images, videos, text, etc. 514
Metadata, e.g., Dates, Identifying Features, User Log-In Information 516
560
562
564
Content Delivery Network 540

FIG. 5

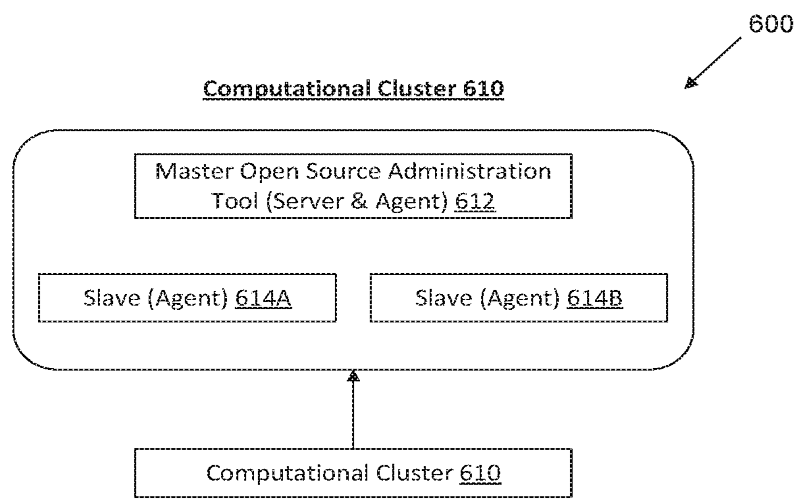

600
Computational Cluster 610
Master Open Source Administration Tool (Server & Agent) 612
Slave (Agent) 614A
Slave (Agent) 614B
Computational Cluster 610

FIG. 6

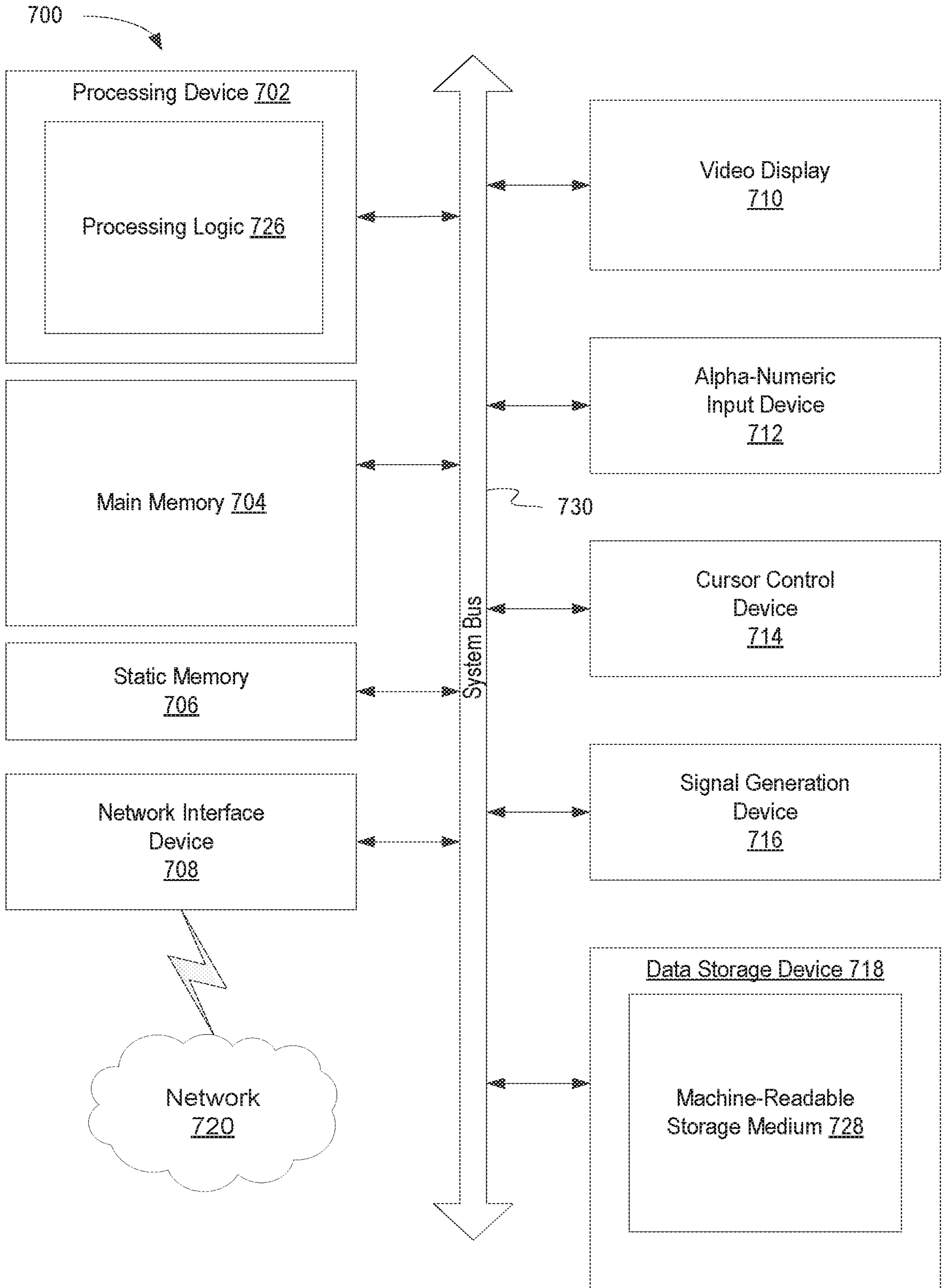

700
Processing Device 702
Processing Logic 726
Main Memory 704
Static Memory 706
Network Interface Device 708
Network 720
System Bus
730
Video Display 710
Alpha-Numeric Input Device 712
Cursor Control Device 714
Signal Generation Device 716
Data Storage Device 718
Machine-Readable Storage Medium 728

FIG. 7