US 20250255289A1

(19) United States

(12) Patent Application Publication

(10) Pub. No.: US 2025/0255289 A1

(43) Pub. Date: Aug. 14, 2025

Chellappan et al.

(54) SMART MOSQUITO TRAP FOR MOSQUITO CLASSIFICATION

(71) Applicant: **University of South Florida**, Tampa, FL (US)

(72) Inventors: **Sriram Chellappan**, Tampa, FL (US); **Ryan Marc Carney**, Temple Terrace, FL (US); **Brandon Wolfram**, Seffner, FL (US); **Ahsan Wahab**, Tampa, FL (US); **Ismail Mustali**, Tampa, FL (US)

(21) Appl. No.: **19/190,061**

(22) Filed: **Apr. 25, 2025**

Related U.S. Application Data

(63) Continuation-in-part of application No. 17/496,563, filed on Oct. 7, 2021.

(60) Provisional application No. 63/198,254, filed on Oct. 7, 2020.

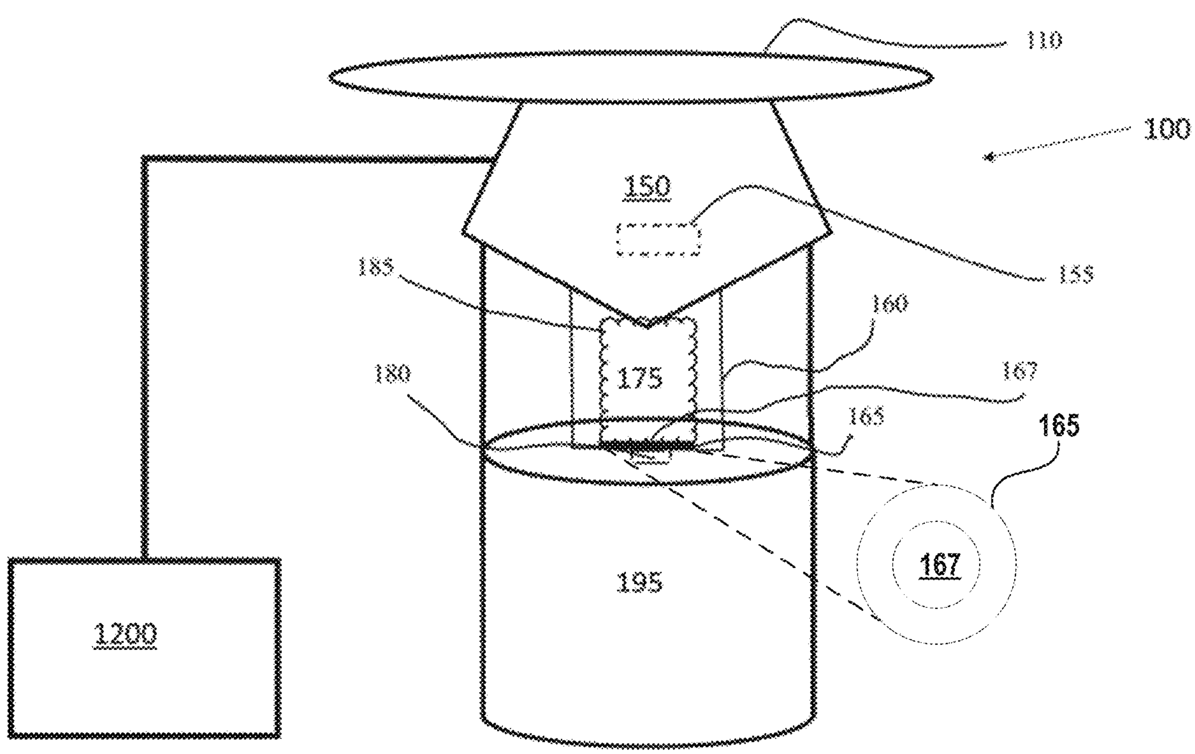
Publication Classification

(51) **Int. Cl.**
A01M 1/02 (2006.01)
A01M 1/06 (2006.01)
A01M 1/14 (2006.01)

(52) **U.S. Cl.**
CPC ***A01M 1/026*** (2013.01); ***A01M 1/023*** (2013.01); ***A01M 1/06*** (2013.01); ***A01M 1/14*** (2013.01); *A01M 2200/012* (2013.01)

(57) **ABSTRACT**

An insect trap includes a combination of one or more components used to classify the insect according to a genus and species. In one embodiment, an apparatus can include an entry aperture configured to lead one or more insects to an imaging chamber within the apparatus. A device may include a door on a side surface of the apparatus, wherein a rear surface of the door is positioned within the imaging chamber in a closed position, and wherein the rear surface of the door comprises a removable/replaceable adhesive material or surface configured to trap one or more insects thereon. A device may include at least one imaging sensor positioned opposite the rear surface of the door configured to capture images of the one or more insects adhering to the adhesive material or surface.



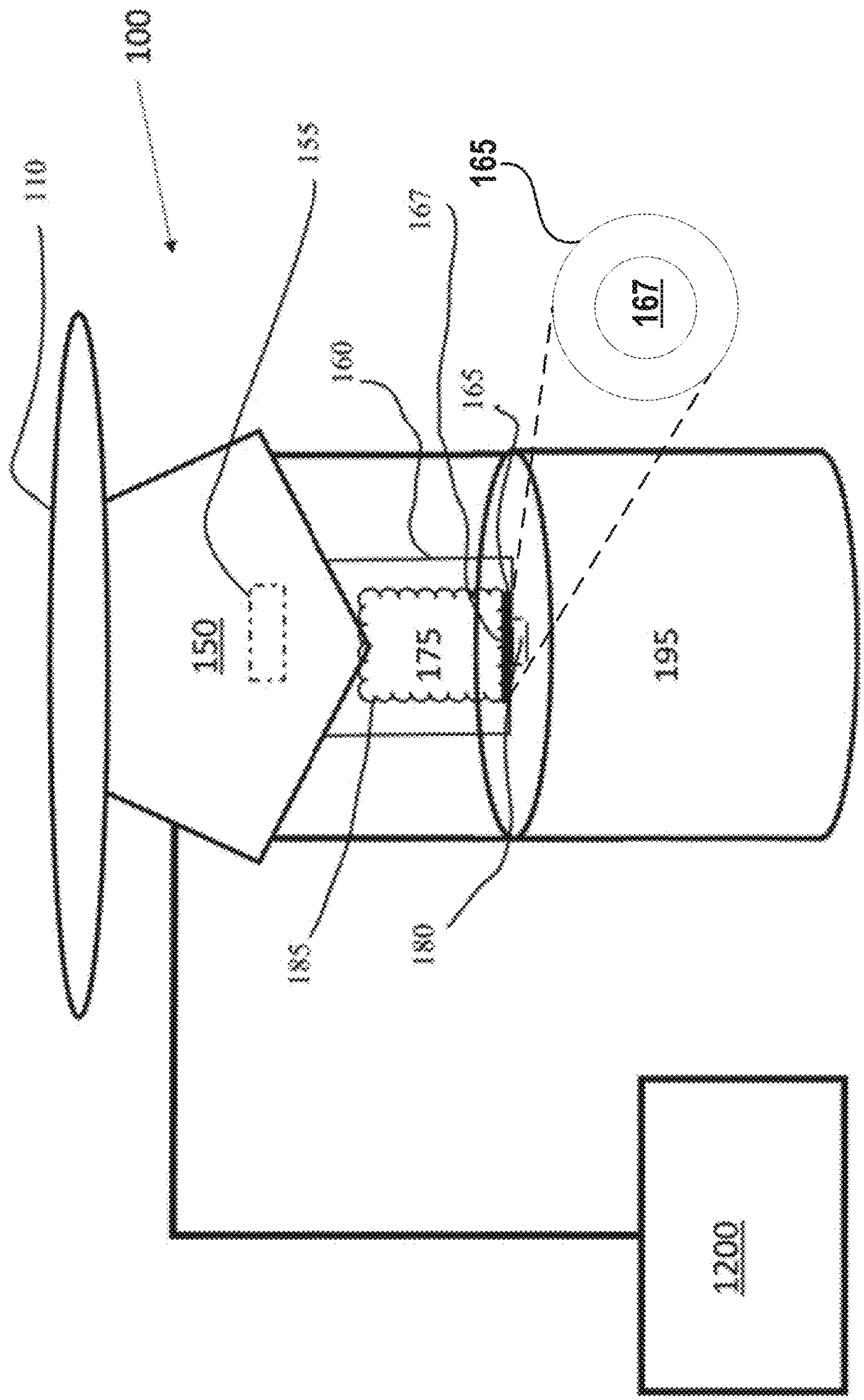


FIG. 1A

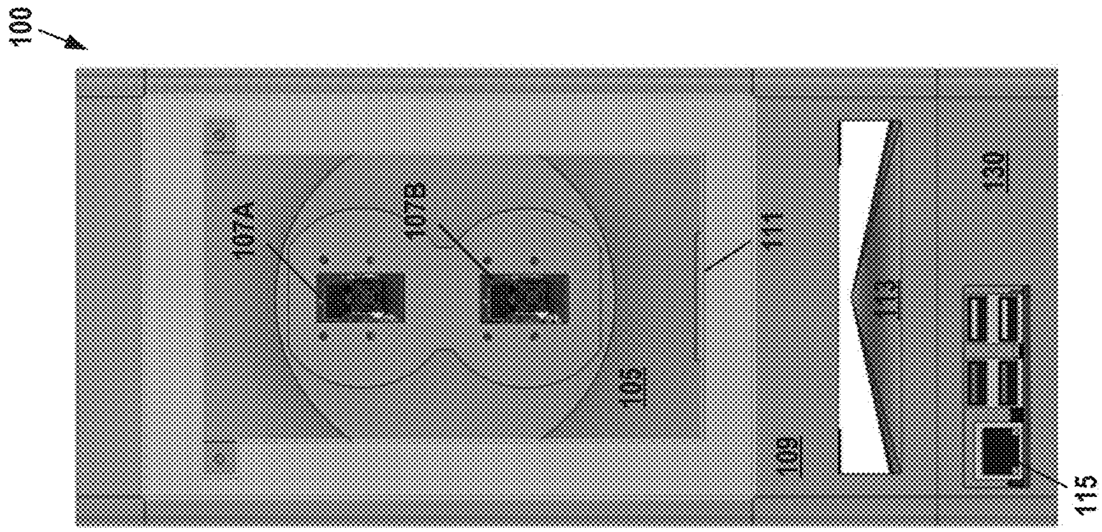

100
107A
107B
111
105
109
113
130
115

FIG. 1C

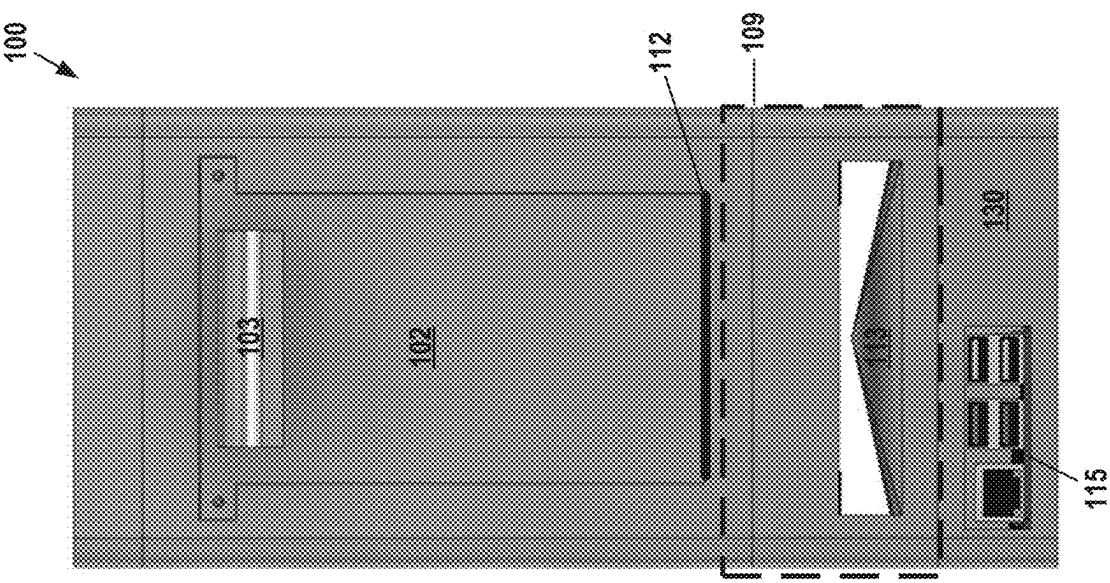

100
112
109
103
102
130
115

FIG. 1B

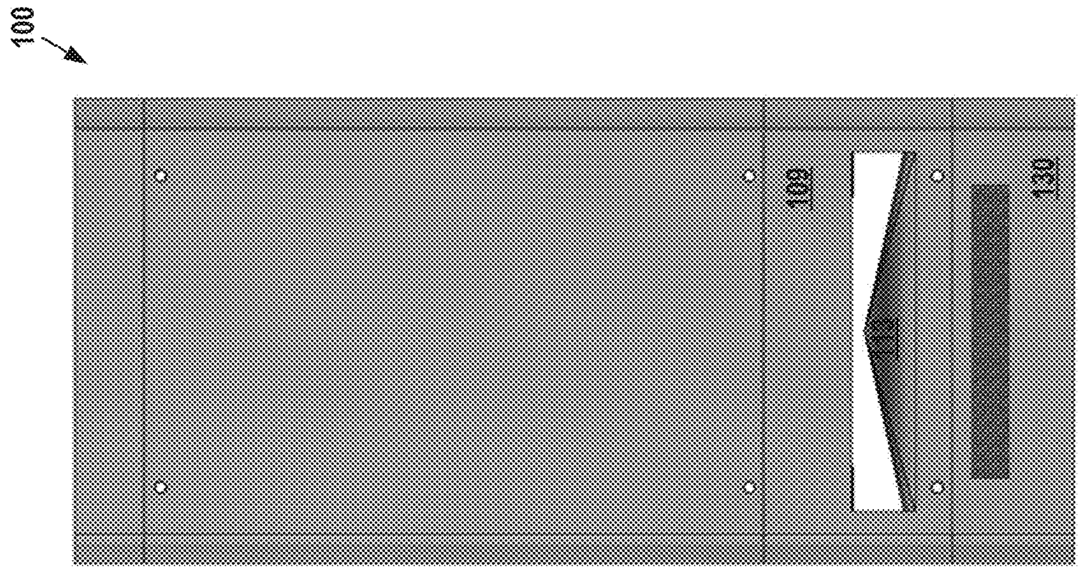


FIG. 1E

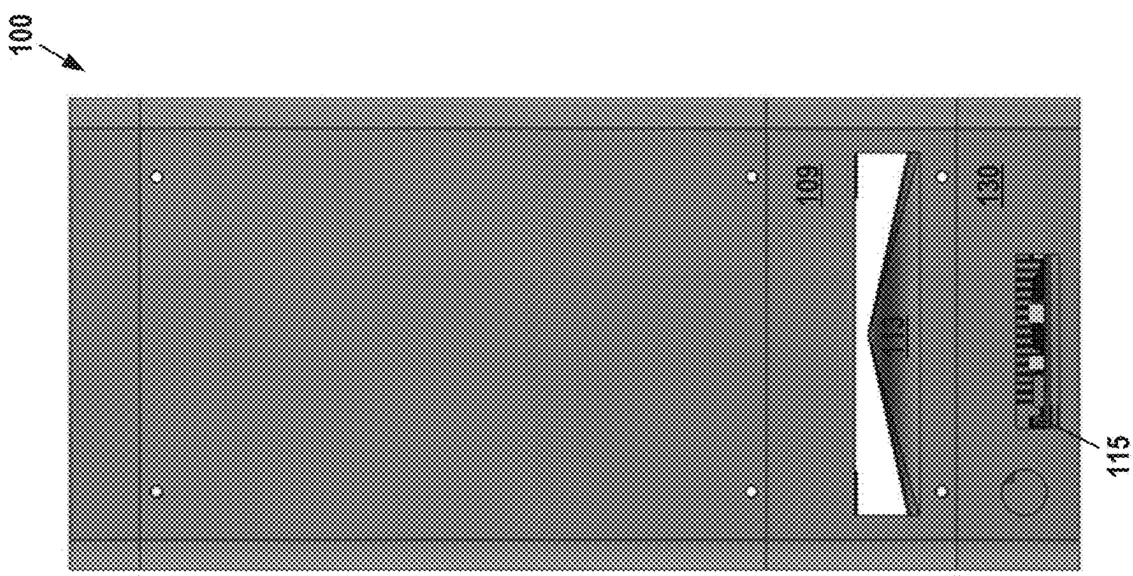


FIG. 1D

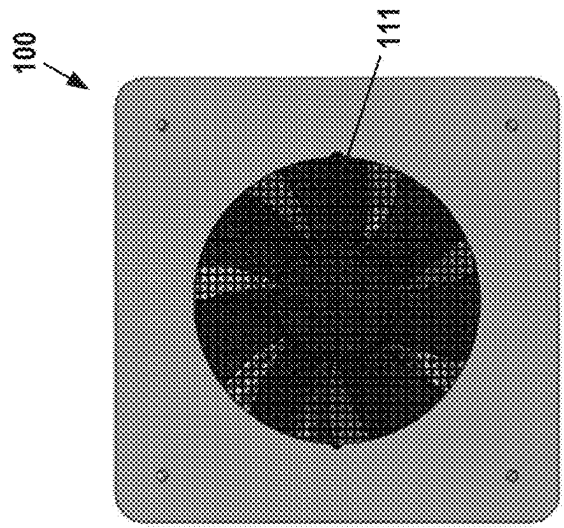


FIG. 1G

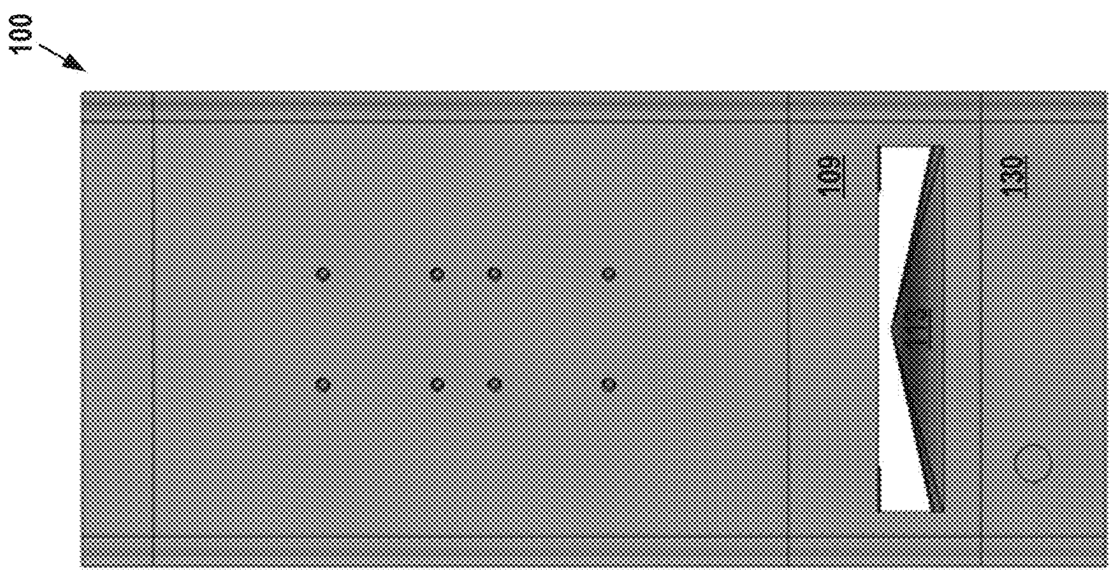


FIG. 1F

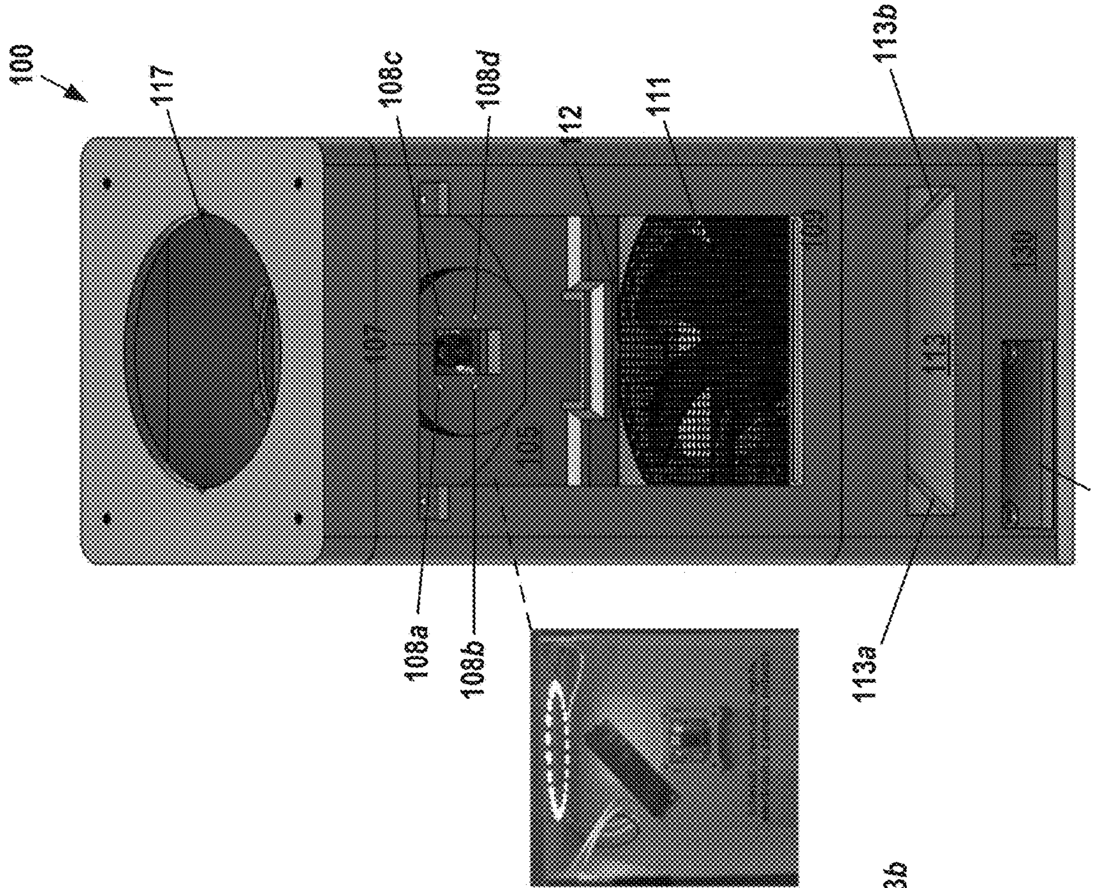

100
117
108c
108d
112
111
113b
107
106
109
113
130
115
113a
108a
108b

FIG. 1I

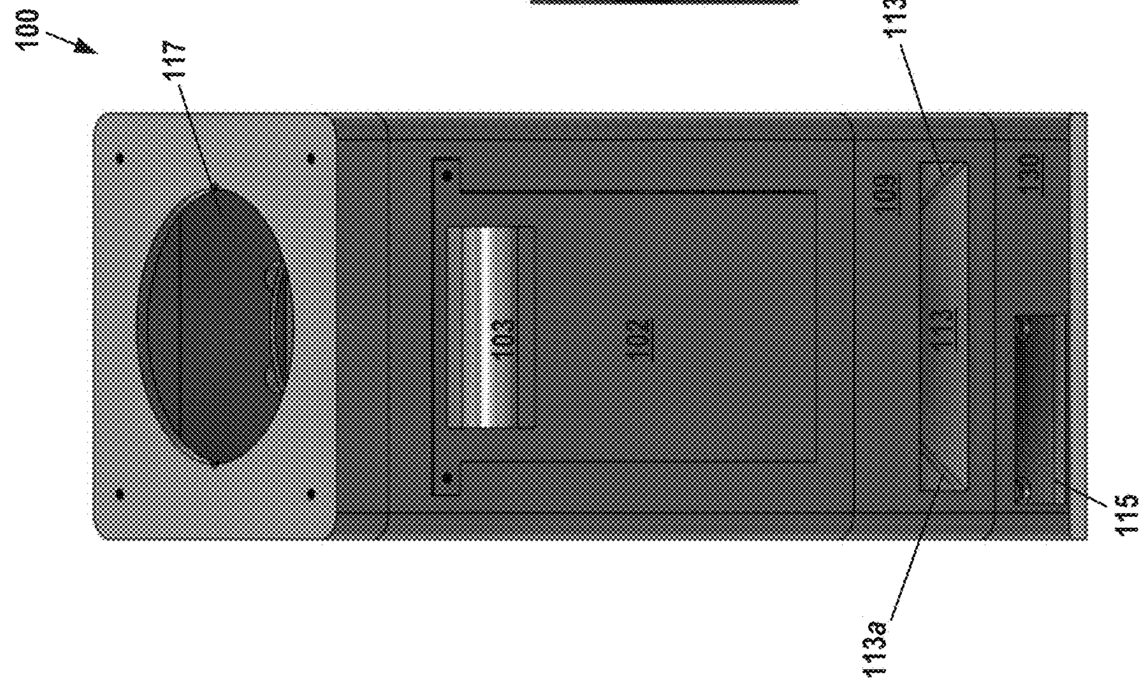

100
117
113b
109
130
103
102
113
115
113a

FIG. 1H

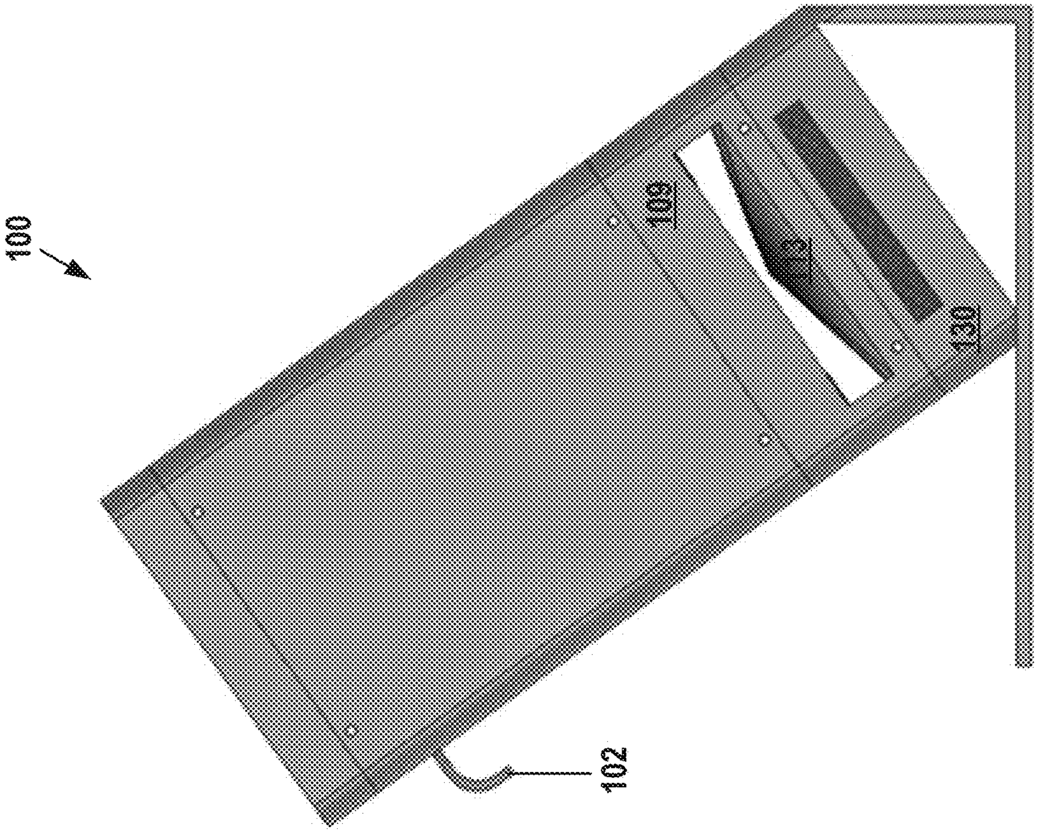


FIG. 1K

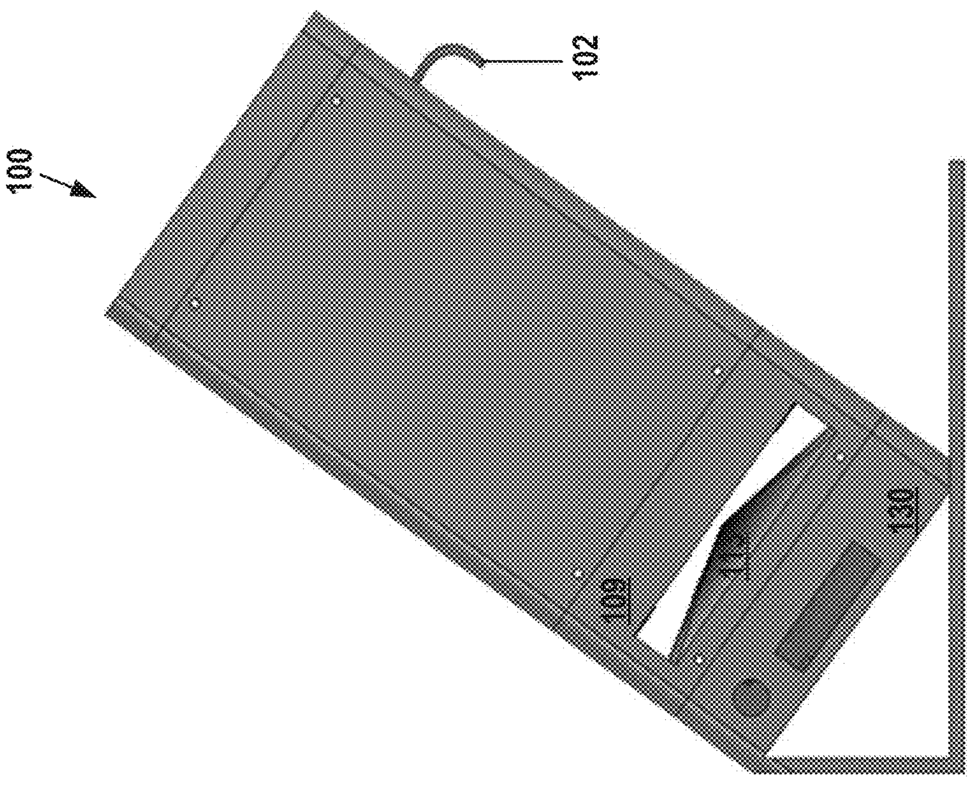


FIG. 1J

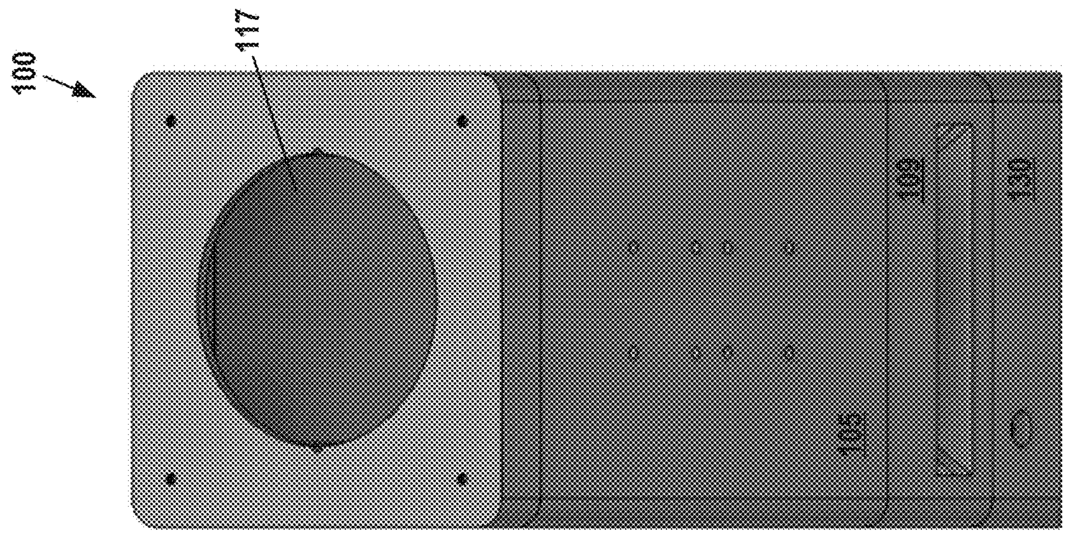


FIG. 1M

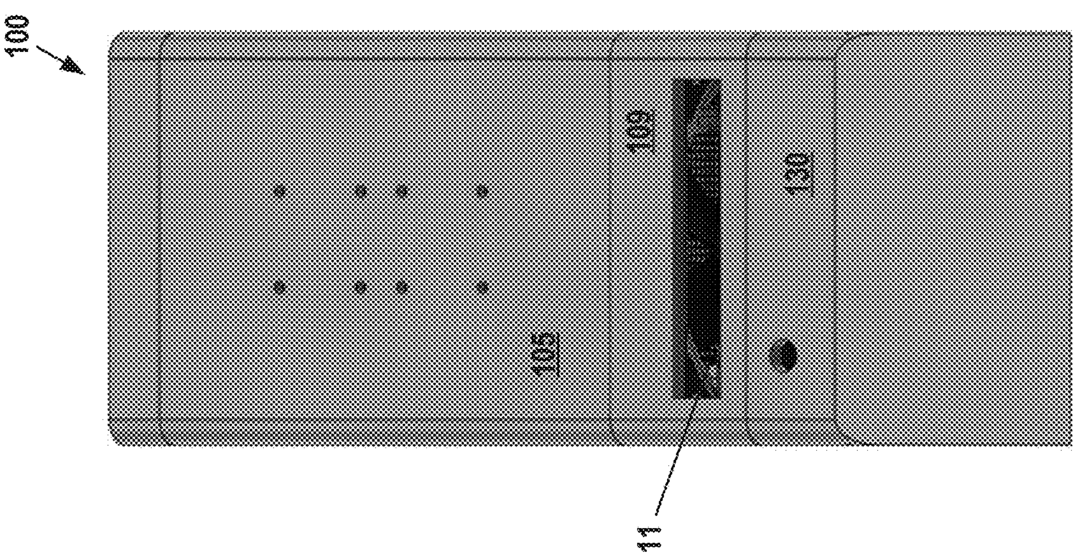


FIG. 1L

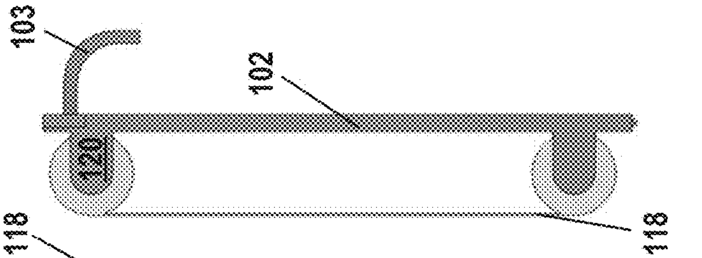


FIG. 1Q

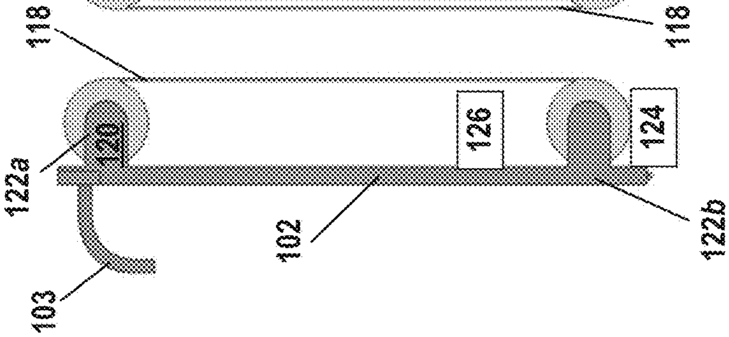


FIG. 1P

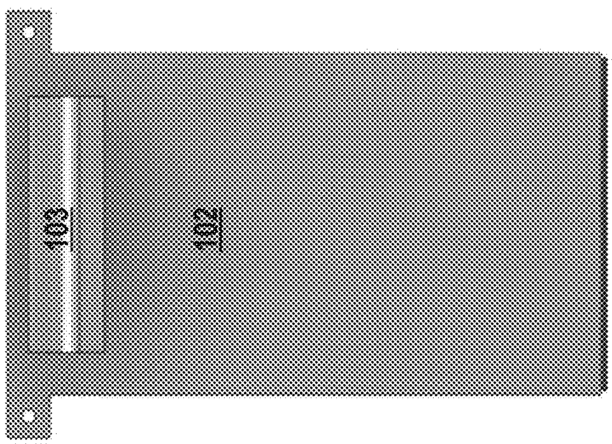


FIG. 1O

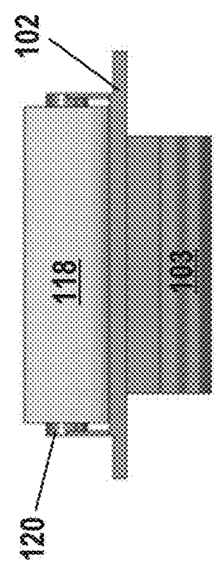


FIG. 1R

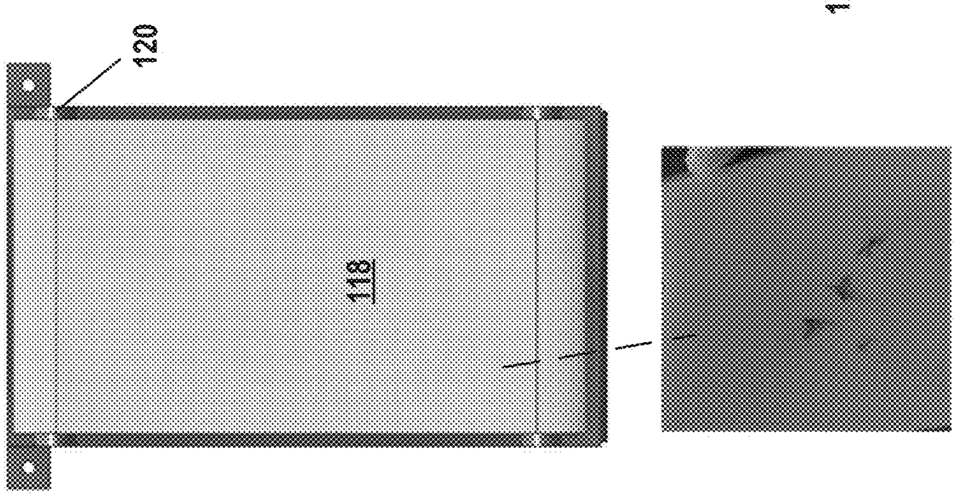


FIG. 1N

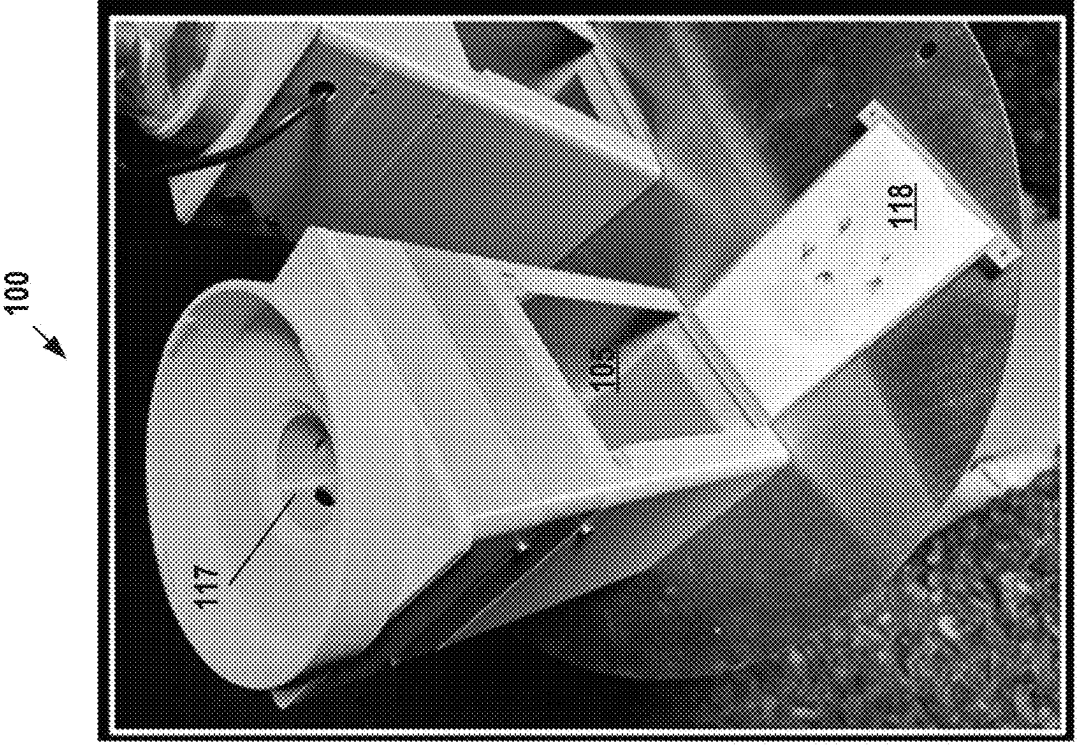


FIG. 1S

FIG. 1T

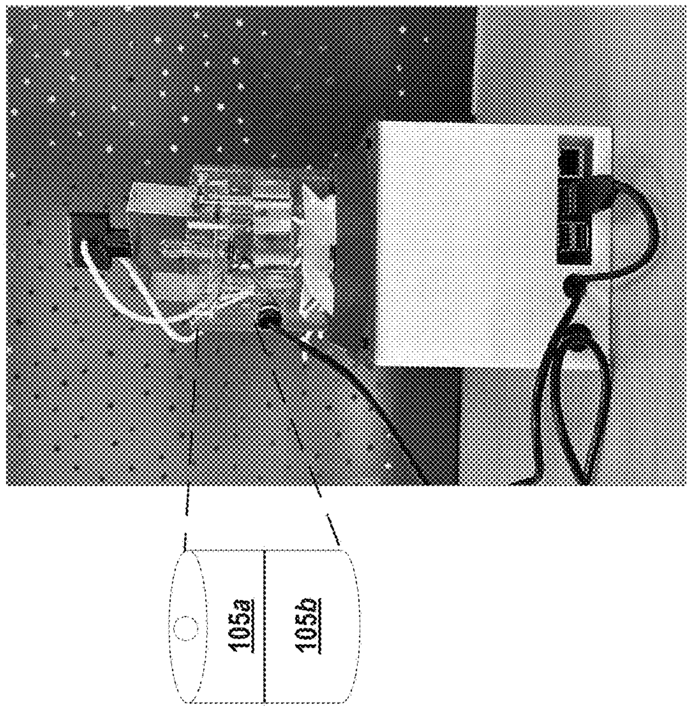


FIG. 1W

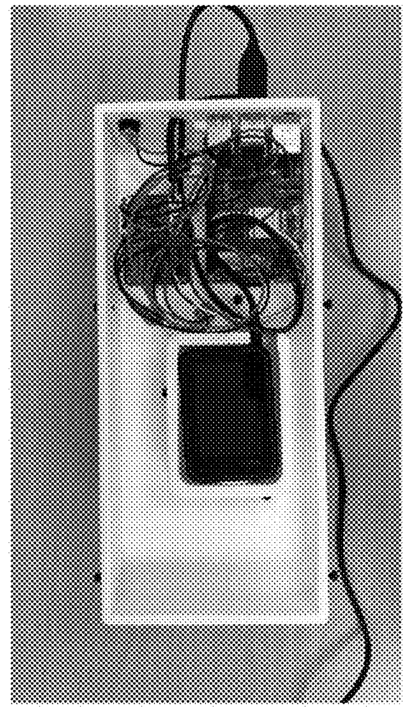


FIG. 1X

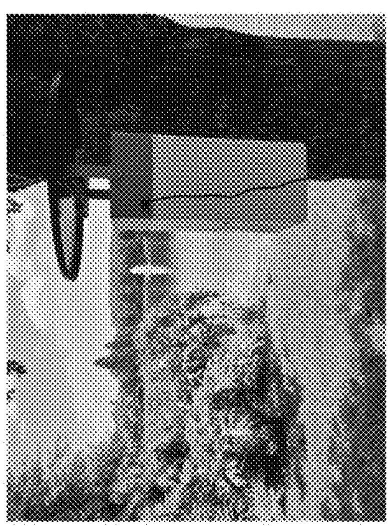


FIG. 1U

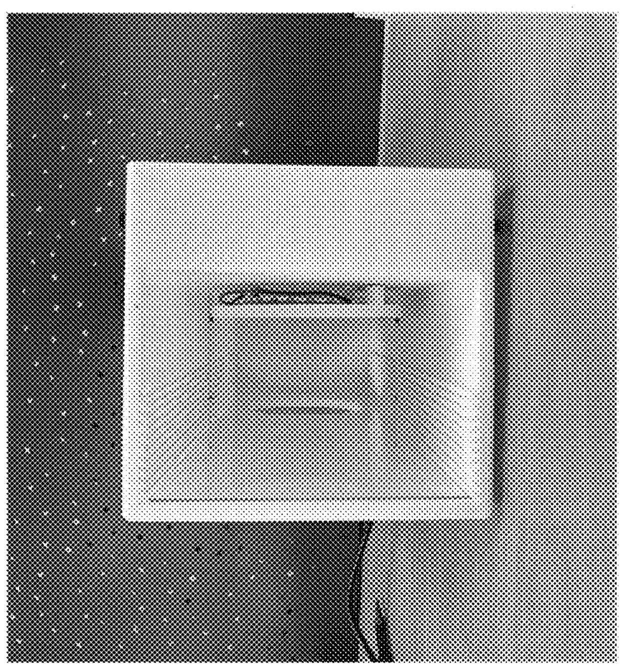


FIG. 1V

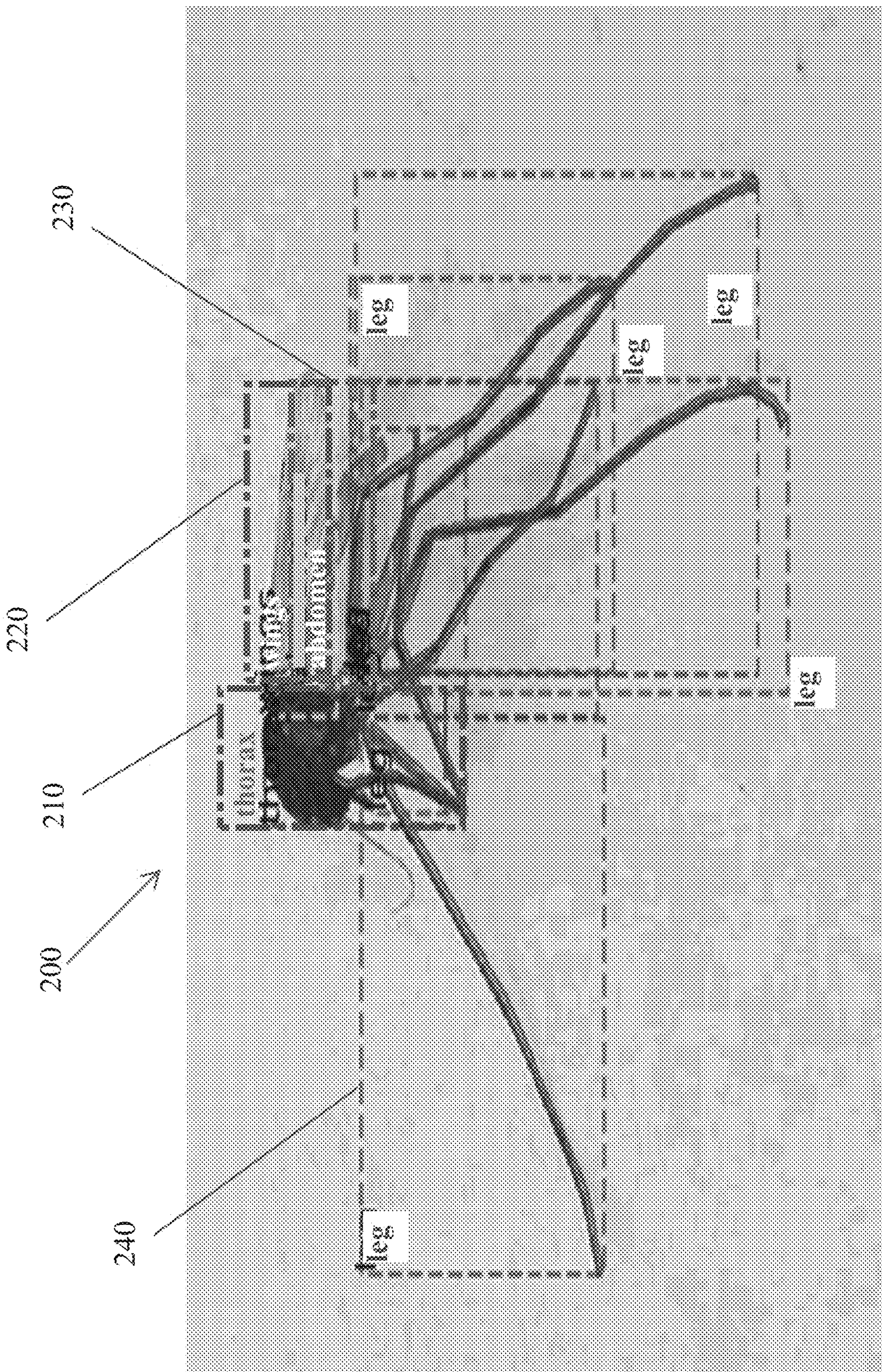


FIG. 2

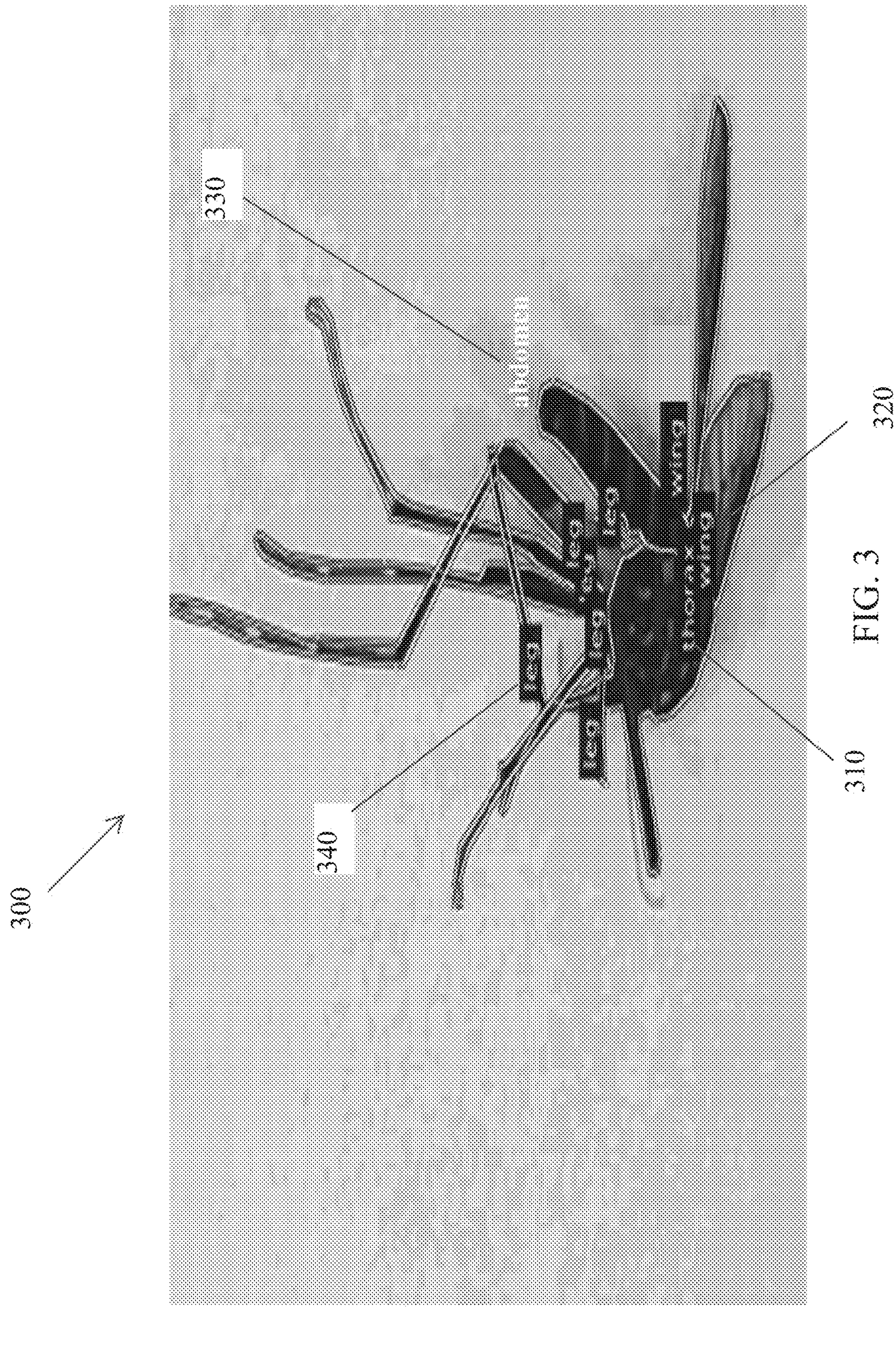


FIG. 3

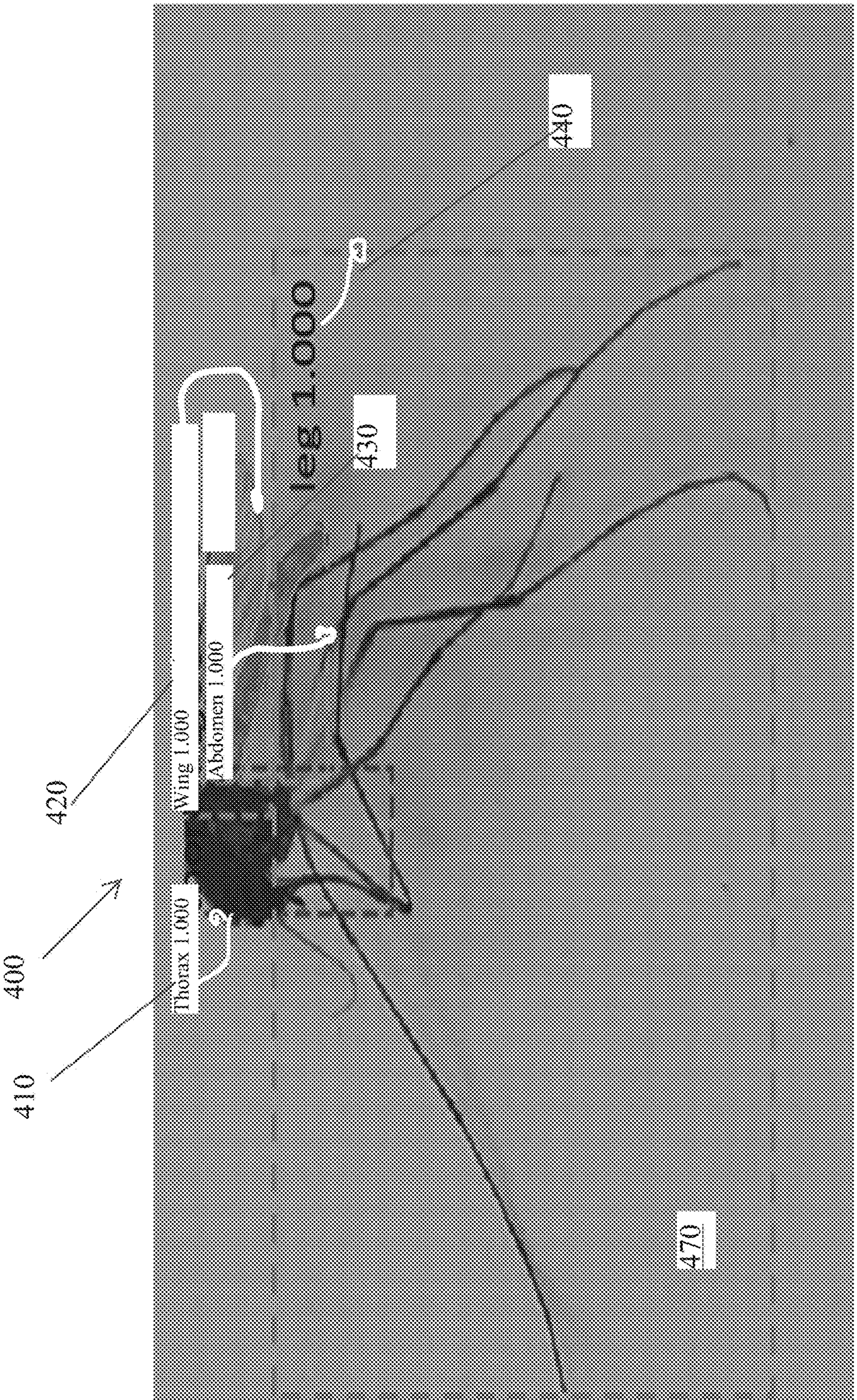

400
410
420
430
440
470
Thorax 1.000
Wing 1.000
Abdomen 1.000
leg 1.000

FIG. 4

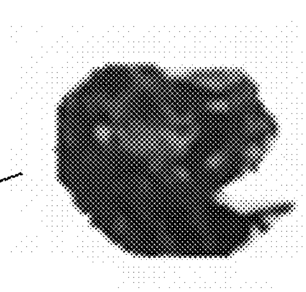


FIG. 5A

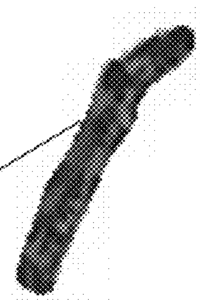


FIG. 5B

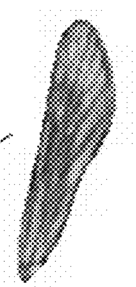


FIG. 5C

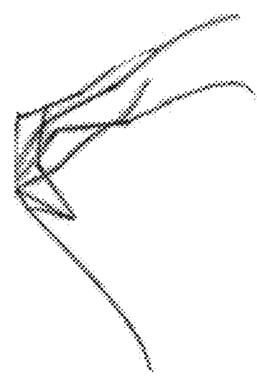


FIG. 5D

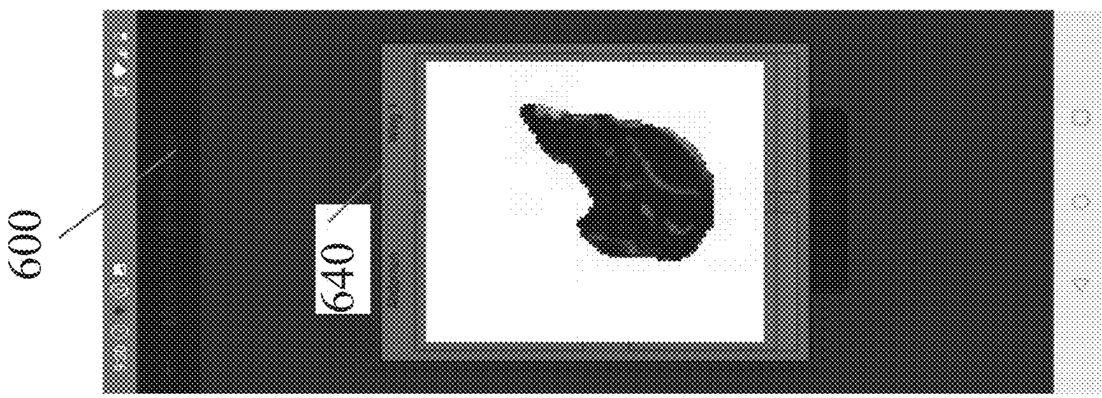


FIG. 6C

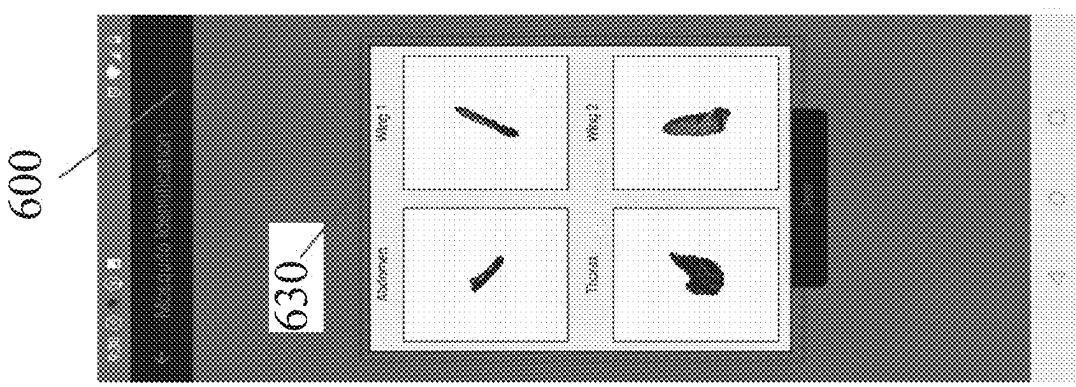


FIG. 6B

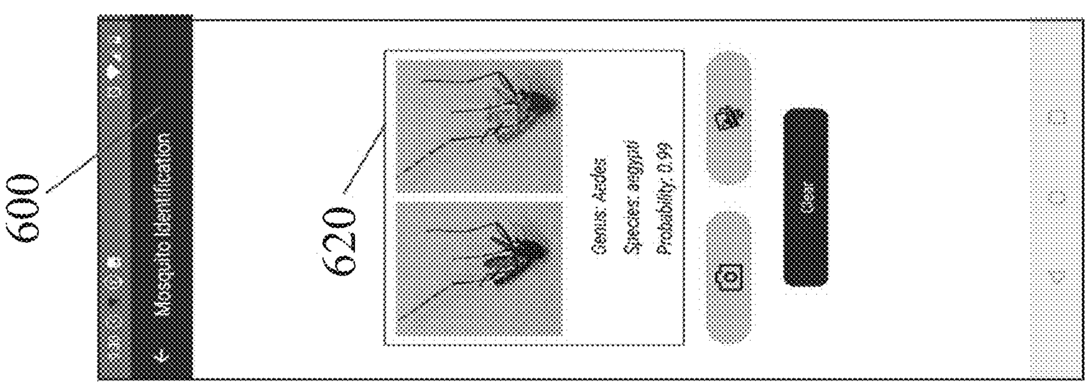


FIG. 6A

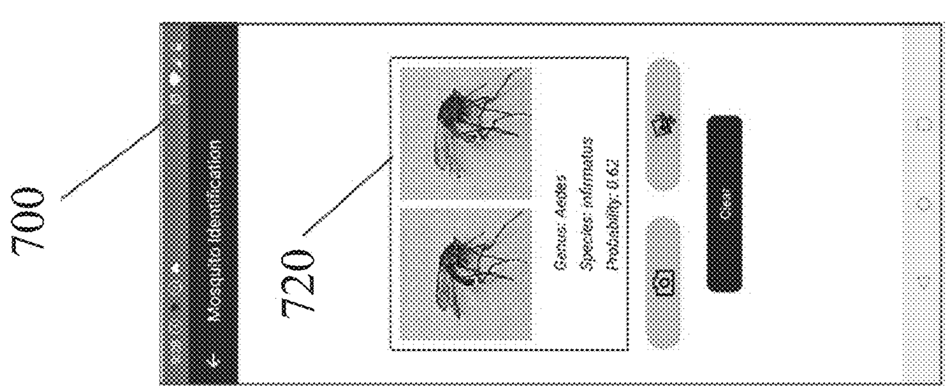


FIG. 7A

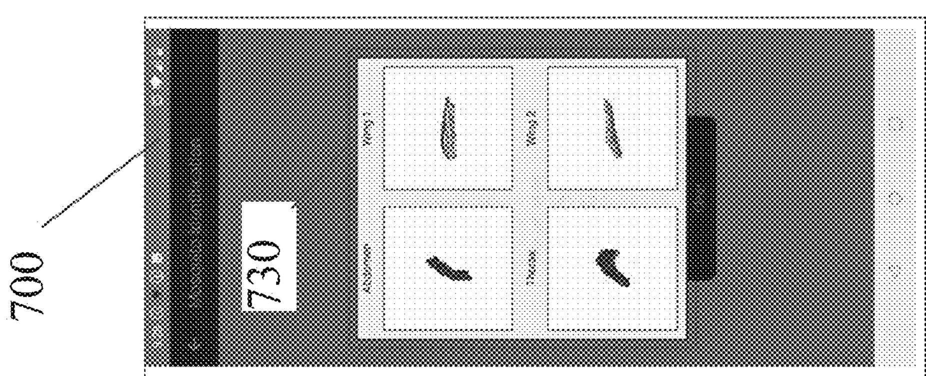


FIG. 7B

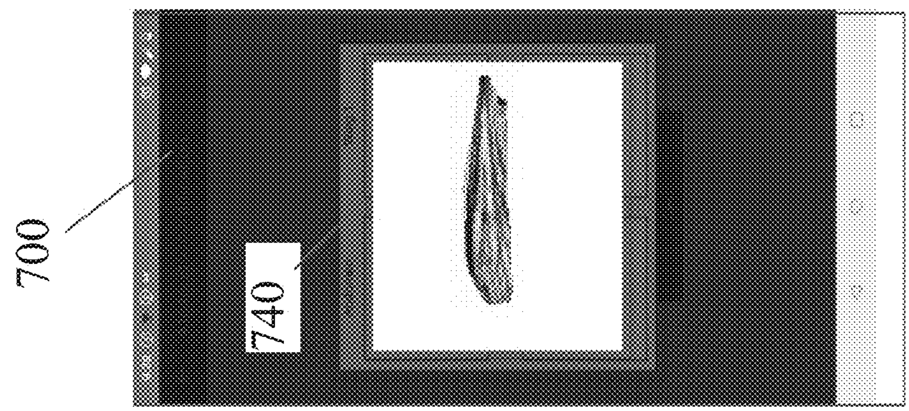


FIG. 7C

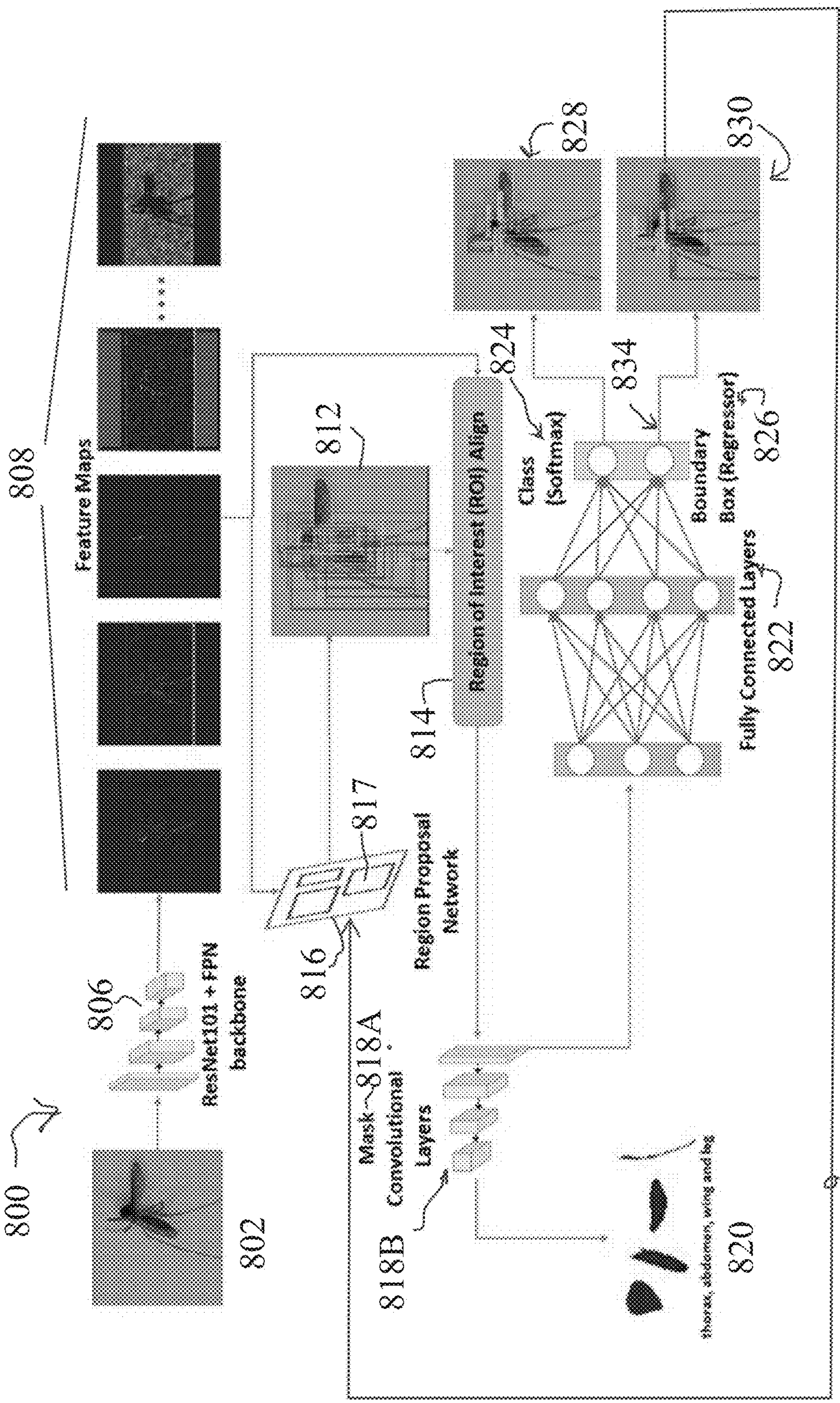

800
802
806
ResNet101 + FPN backbone
808
Feature Maps
812
814
Region of Interest (ROI) Align
816
817
Region Proposal Network
818A
818B
Mask Convolutional Layers
820
822
Fully Connected Layers
824
Class (Softmax)
826
Boundary Box (Regressor)
828
830
834
850

FIG. 8

900A

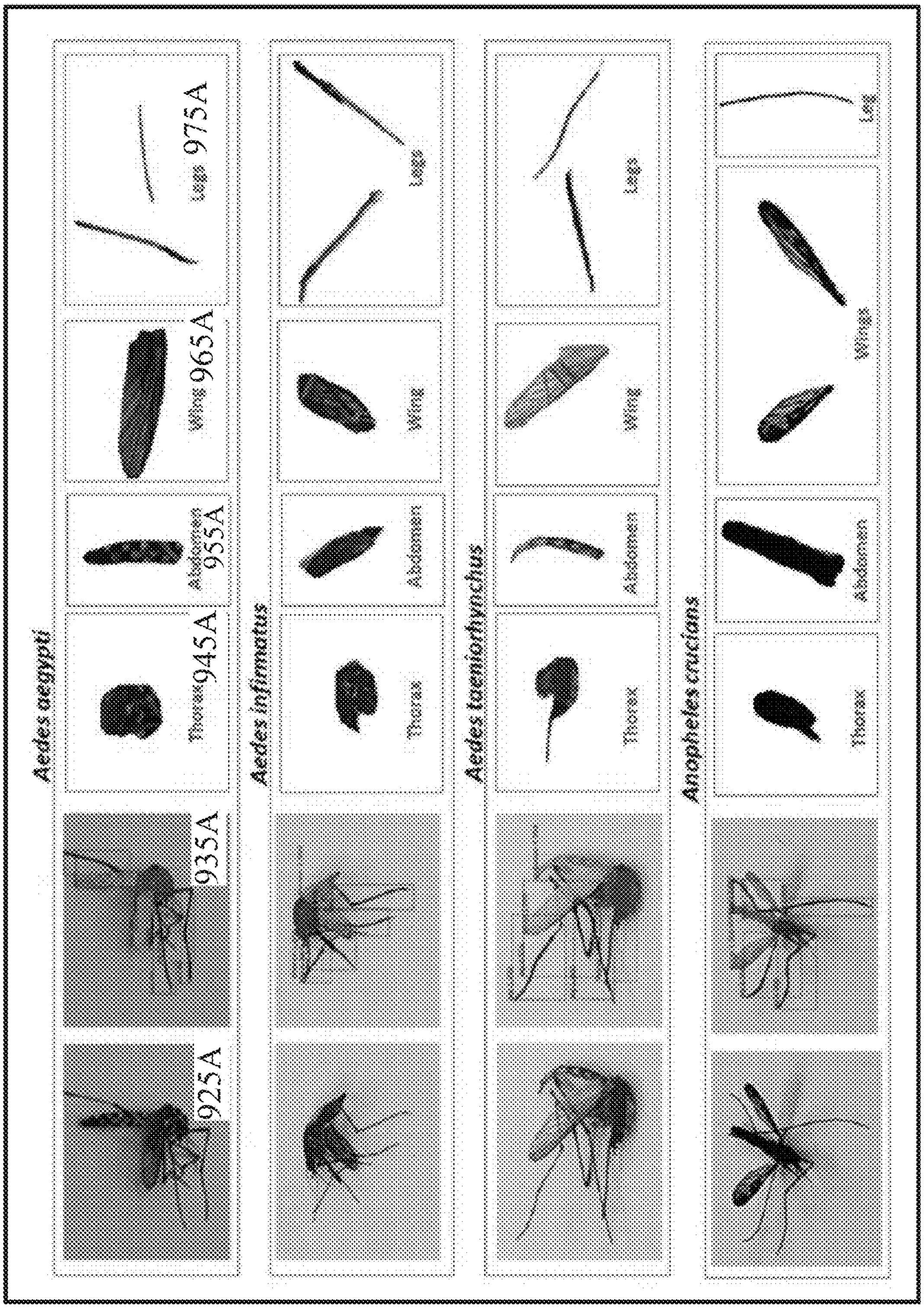

Aedes aegypti
925A
935A
Thorax 945A
Abdomen 955A
Wing 965A
Legs 975A
Aedes infirmatus
Thorax
Abdomen
Wing
Legs
Aedes taeniorhynchus
Thorax
Abdomen
Wing
Legs
Anopheles crucians
Thorax
Abdomen
Wings
Leg

FIG. 9A

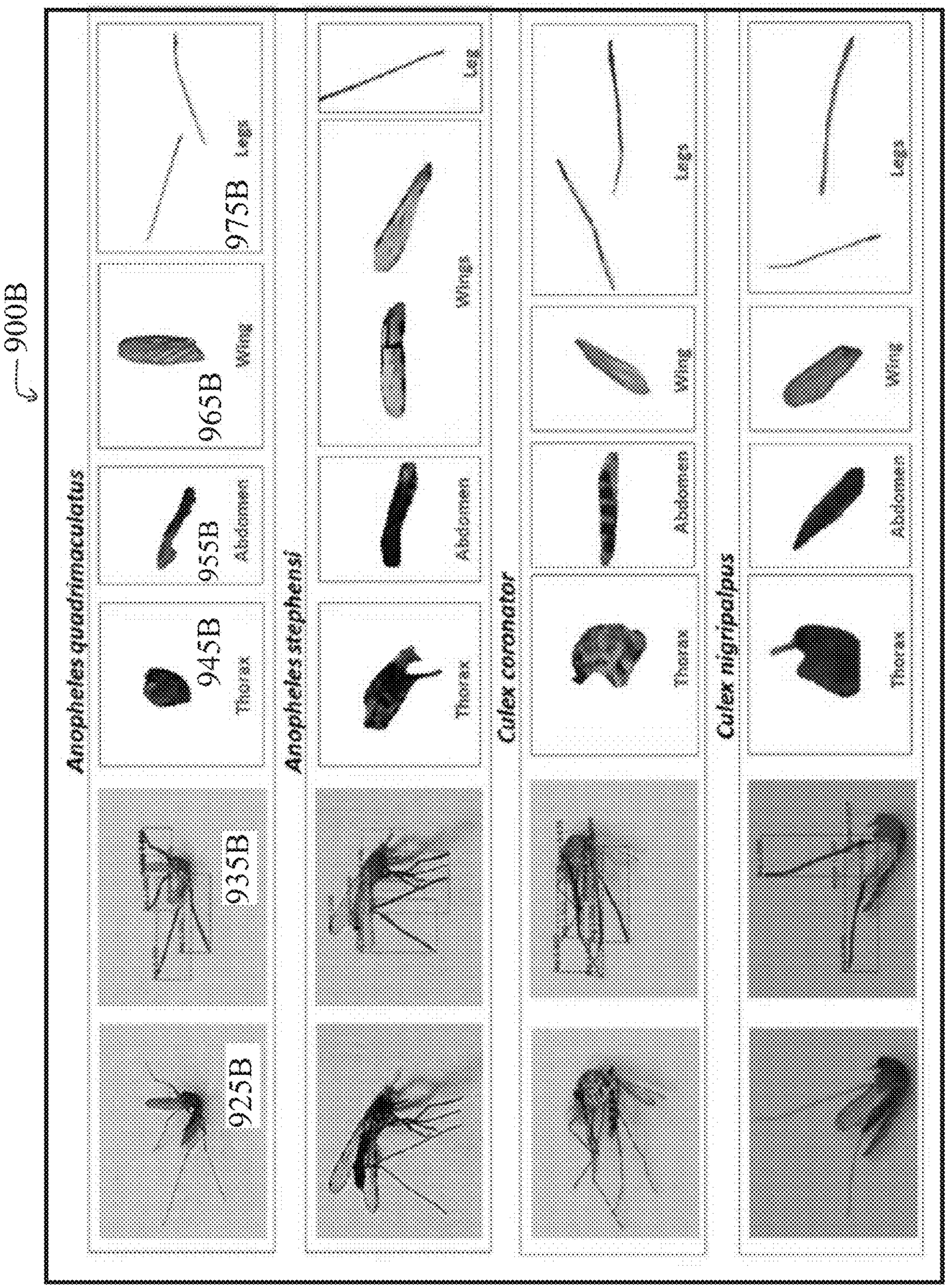

900B
Anopheles quadrimaculatus
925B
935B
945B
Thorax
955B
Abdomen
965B
Wing
975B
Legs
Anopheles stephensi
Thorax
Abdomen
Wings
Leg
Culex coronator
Thorax
Abdomen
Wing
Legs
Culex nigripalpus
Thorax
Abdomen
Wing
Legs

FIG. 9B

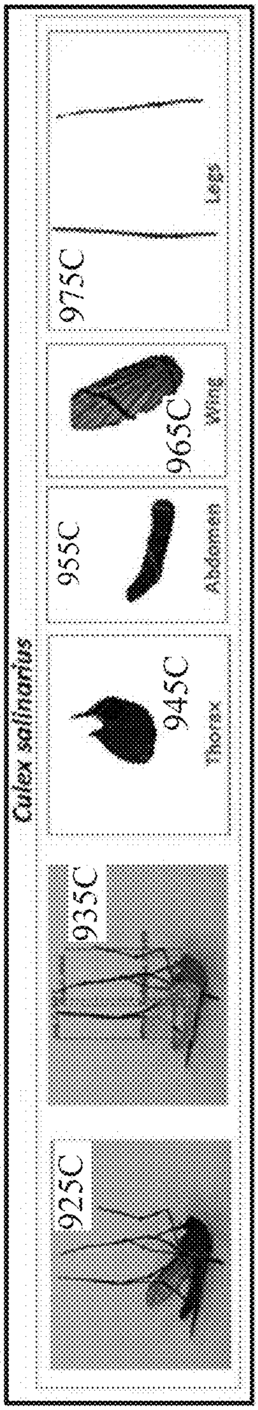

900C
Culex salinarius
925C
935C
945C
Thorax
955C
Abdomen
965C
Wing
975C
Legs

FIG. 9C

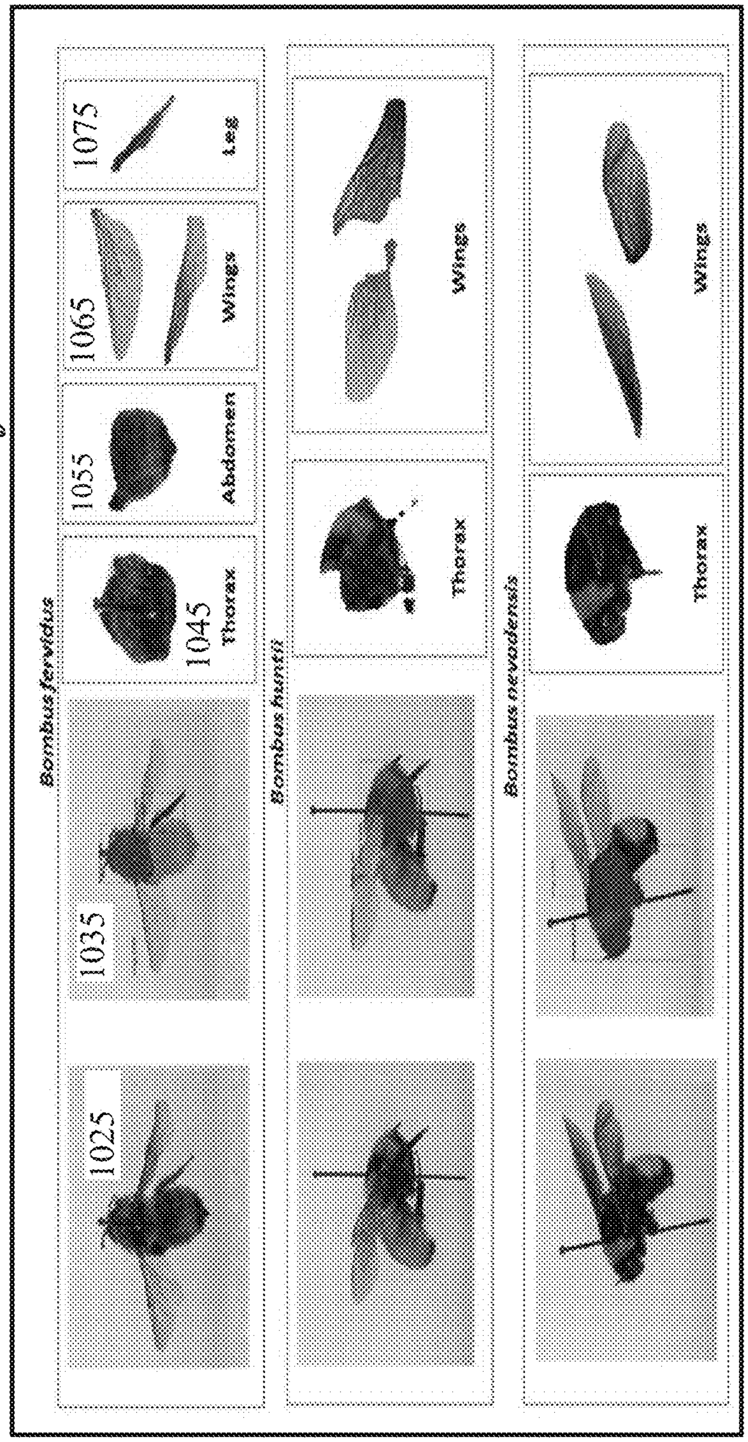

1000
Bombus fervidus
1025
1035
1045
Thorax
1055
Abdomen
1065
Wings
1075
Leg
Bombus huntii
Thorax
Wings
Bombus nevadensis
Thorax
Wings

FIG. 10

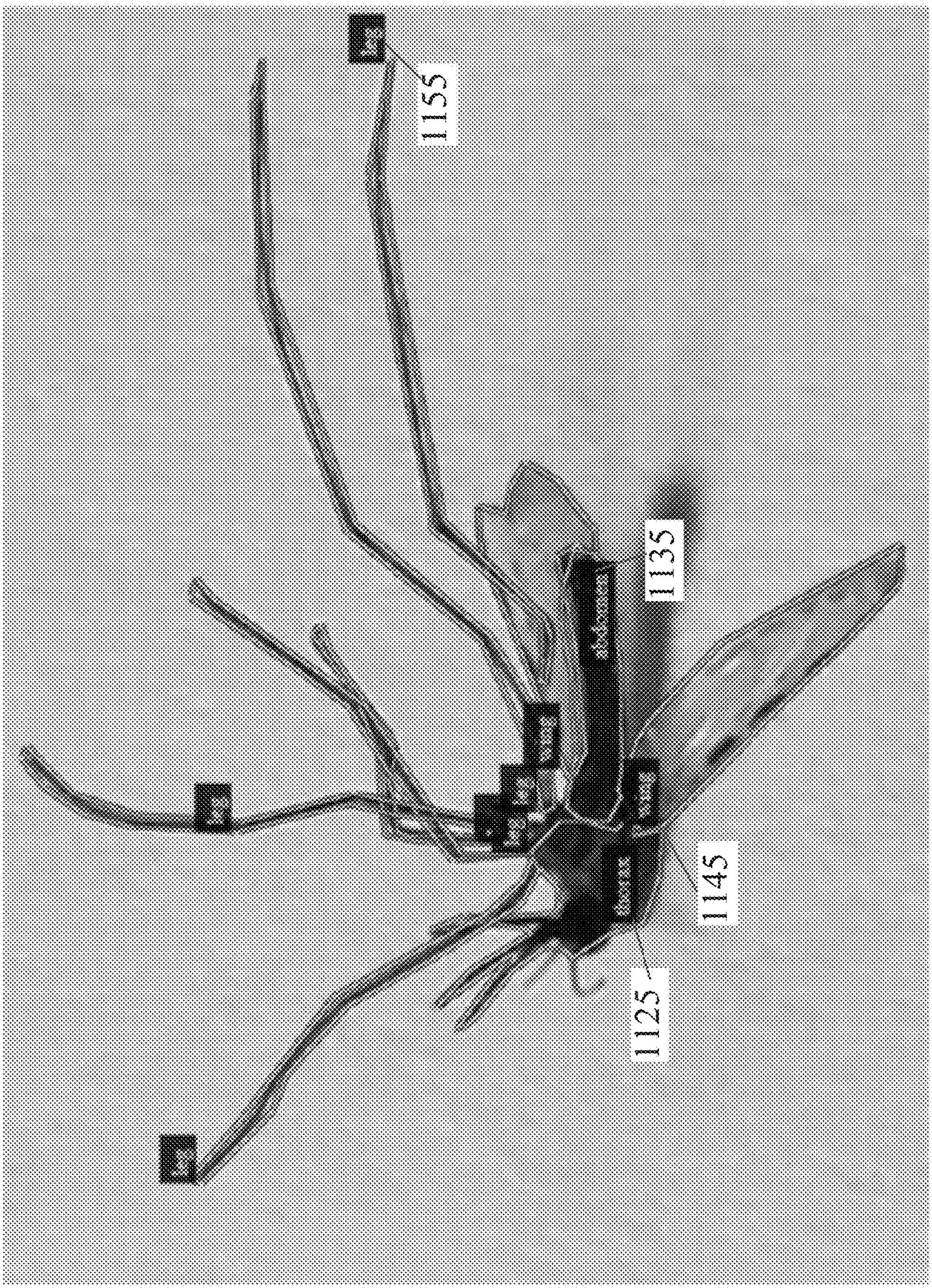


FIG. 11

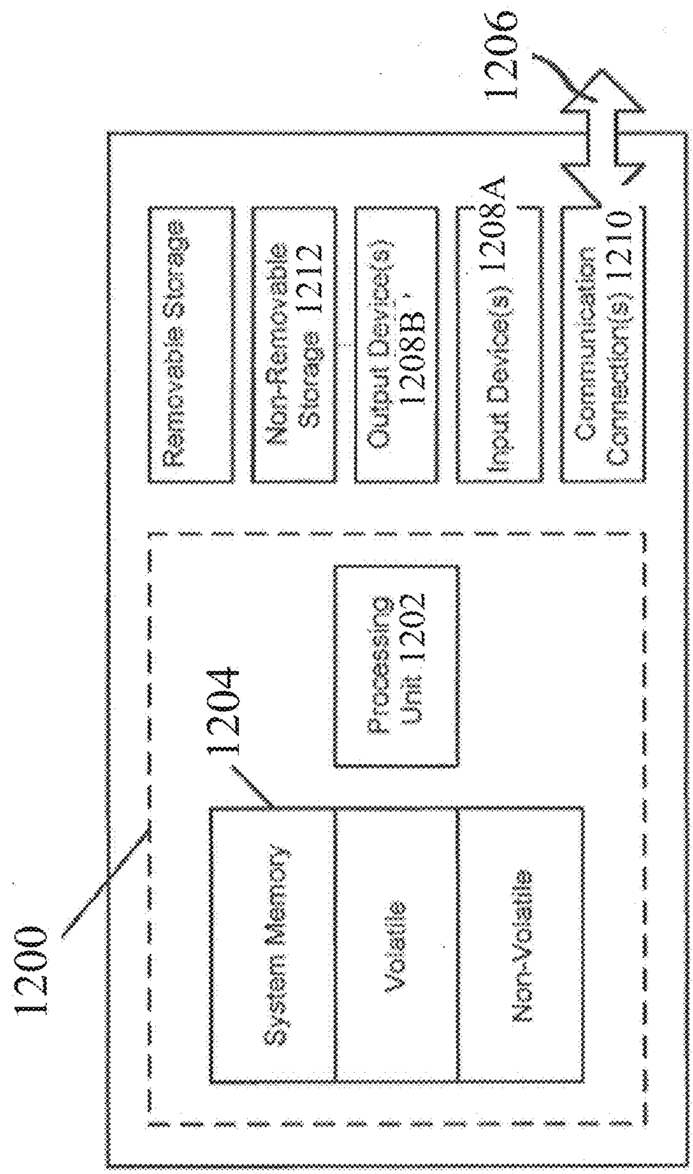

1200
1204
System Memory
Volatile
Non-Volatile
Processing Unit 1202
Removable Storage
Non-Removable Storage 1212
Output Device(s) 1208B
Input Device(s) 1208A
Communication Connection(s) 1210
1206

FIG. 12

SMART MOSQUITO TRAP FOR MOSQUITO CLASSIFICATION

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application is a continuation in part of U.S. Non-Provisional patent application Ser. No. 17/496,563, filed on Oct. 7, 2021, which claims priority to, and the benefit of, U.S. Provisional Patent Application Ser. No. 63/198,254 filed on Oct. 7, 2020, the contents of which are hereby incorporated by reference herein in their entireties.

BACKGROUND

[0002] Taxonomy is the process of classifying organisms in nature. Entomology is the study of insect organisms. Taxonomy in the context of entomology is a relatively obscure discipline in the era of modern sciences. Very few people want their professional careers spent with hours poring through a microscope trying to identify what genus and species an insect is. In the context of mosquitoes, there are close to 4500 different species of mosquitoes, and training to identify all of these mosquitoes is hard if not impossible. In countries like India, Bangladesh and even the US, it is simply not possible to train professionals to identify all mosquitoes that are endemic in these countries (e.g., there are 400 species of mosquitoes endemic to India; and about 150 species in the US). With increasing travel and global connectivity among nations, mosquitoes can invade to newer places, and identifying the "new" mosquitoes becomes impossible by local professionals. Mosquitos and other insects are considered "vectors" because they can carry viruses, bacteria, and strains of diseases and transmit them to humans. The term "vector" is therefore given its broadest meaning in the art of infectious diseases.

[0003] Modern entomology updates have focused on eliminating or minimizing human involvement in classifying genus and species of mosquitoes during disease outbreak. There are close to 4500 different species of mosquitoes in the world spread across **45** or so genera. Out of these, only handfuls of species across three genus types spread the deadliest diseases. These mosquitoes belong to *Aedes* (Zika, Dengue, Chikungunya, Yellow Fever), *Culex* (West Nile Virus, and EEE), and *Anopheles* (Malaria). Within these three genera, the deadliest species are *Aedes aegypti, Aedes albopictus, Culex nigripalpus, Anopheles gambiae* and *Anopheles stephensi*. When a mosquito-borne disease, say Dengue affects a region, then identifying the presence of the particular vectors for Dengue (i.e., *Aedes aegypti* and *Aedes albopictus*) becomes important. This is hard and expensive. For instance in India, there are close to 450 types of mosquitoes spread all over. Accordingly, public health experts lay traps in disease prone areas, and sometimes hundreds of mosquitoes get trapped. Now, however, they can identify which of those is the genus and species they are looking for. Because, once they identify the right mosquitoes, they can then take those mosquitoes to the lab for DNA testing etc. to see if the pathogen (i.e., virus) is there within the trapped mosquito. Naturally, if they find a reasonable large number of those mosquitoes with the virus in them, there is a public health crisis, and corrective action needs to be taken.

[0004] Other efforts have focused on detecting foreign mosquitoes at borders. This is a problem that is attracting a lot of global attention—the need to identify if a mosquito in borders of a nation (land or sea or air or road) is a foreign mosquito. For instance, consider a scenario in which mosquitos, e.g., both a domestic vector and one non-native to the US, are on a vehicle entering the US borders.

[0005] Assuming that borders do have mosquito traps, it is likely that this "new" breed of mosquito could get trapped along with other local mosquitoes. The question here is how public health authorities identify that a "foreign" mosquito is in one such trap. Current entomology classification systems would require going periodically to these traps, collecting and studying subjects through a microscope, and identifying specimens one by one. This is impossibly cumbersome if the goal is to only detect a particular type of "foreign" mosquito.

[0006] Current disease models rely upon proper classification of infection vectors. The entomology classification systems need to be improved for use in specialized and detail intensive instances, such as the hypothetical above. A need exists in the art of entomological classification to include algorithms that are adaptable for use in resolving important, yet hard to pinpoint issues, such as identifying the hypothetical "foreign" mosquito that did indeed get trapped. Updated algorithms are needed to provide researchers with options in entomological classification for specialized situations, such as the hypothetical random occurrence of new insects affecting a local population.

[0007] Continuing with the "foreign" mosquito example, the art of entomological classifications needs improved techniques and models that have been trained with images of the foreign mosquito (provided by international partners) to identify the genus and species directly from initial observations. In the alternative, a need should be met to enable running the foreign mosquito through models trained with other mosquitoes. These techniques would allow researchers to notify public health officials that a new mosquito, that appears to be previously unknown in a given location, has been currently trapped. In either case, there is significant benefit for public health at borders.

[0008] As detailed in this disclosure, to address the above noted inadequacies, digitizing anatomies of mosquito specimens across the globe (with citizen and expert involvement) will help create a massive repository of mosquito anatomy images tagged with genus and species types. This repository could then be used for training personnel, and also for automatic identification using algorithms in this disclosure (when a picture is uploaded). For instance and without limiting this disclosure, the Florida Medical Entomology Lab in Vero Beach trains a very small number of personnel each year (both international and also domestic military personnel) in the detailed art of insect classification. From prior investigations, space is very limited, and many are turned away from these kinds of training programs. With a digital repository in place, the training programs can be globally expanded as well with potentially thousands of images to train interested personnel.

[0009] The need for these kinds of improvements in entomological classification is apparent in at least one example. M any states and counties in India (especially those at borders) have been and are currently willing to pay for such a service. Such a service with appropriate mosquito traps and can be deployed in international airplanes, ships and buses.

[0010] In another expression of the needs in this arena, soldiers going to countries where mosquito-borne diseases are common are routinely trained to help local communities identify mosquitoes and other vectors for disease. A digital repository can train soldiers remotely without having to physically travel to a location in need for these services. Furthermore, soldiers and even personnel from government agencies engaged in traveling and residing overseas might benefit from a trap in the bases and/or homes that can tell them decipher the type of mosquitoes trapped in their vicinity, and how dangerous they are.

[0011] Finally, in another expression of needs in this area, mosquito traps today do not have digital advancements in them. Current traps are mostly based on light attractants or carbon dioxide (CO2) attractants (sometimes augmented with human sweat as an attractant). Current traps with these attractants lure mosquitoes in. Once inside the trap, a fan typically pulls the mosquitoes into a chamber from where they cannot get out. Some traps use chemicals in these chambers to actually kill the mosquitoes. The next day, taxonomists pick up mosquitoes in these chambers for identification of genus and species. Other traps can detect if an insect falling in is a mosquito or not-they use passive infrared sensors ("PIR sensors) to detect the mosquito entering the trap. A PIR sensor may incorporate a pyroelectric sensor that detects changes in levels of infrared radiation. PIR sensors are digitally/electrically compatible with communications over a network to other computers; however, there is no ability to capture images or videos in traps know to date.

[0012] A need currently exists for a smart mosquito trap that uses electronics to gather data about mosquito anatomies, behaviors, movements, and the like in order to classify the mosquito specimens according to at least genus and species.

BRIEF SUMMARY OF THE DISCLOSURE

[0013] An insect trap includes a combination of one or more components used to classify the insect according to a genus and species. The trap includes an imaging device, a digital microphone, and passive infrared sensors at the entrance of the trap to sense wing-beat frequencies and size of the insect (to identify entry of a mosquito). A lamb-skin membrane, filled with an insect attractant such as carbon dioxide mixed with gas air inside, mimics human skin so that the insect can rest on the membrane and even pierce the membrane as if a blood meal is available. An imaging device such as a passive infrared sensor or a camera gathers image data of the insect. The insect may be a mosquito.

[0014] This disclosure presents a system to design state of the art artificial intelligence (AI) techniques, namely techniques to use digital images to classify mosquitos. The digital images may be analyzed based on mask region-based convolutional neural networks to extract anatomical components of mosquitoes from digital images and archiving them permanently based on genus, species and other taxonomies.

[0015] In some implementations, an apparatus is provided. The apparatus can include: an entry aperture configured to lead one or more insects to an imaging chamber within the apparatus; a door on a side surface of the apparatus, wherein a rear surface of the door is positioned within the imaging chamber in a closed position, and wherein the rear surface of the door includes a removable/replaceable adhesive material or surface configured to trap one or more insects thereon; and at least one imaging sensor positioned opposite the rear surface of the door configured to capture images or video of the one or more insects adhering to the adhesive material or surface.

[0016] In some implementations, the at least one imaging sensor includes at least two autofocus cameras.

[0017] In some implementations, the apparatus is in electronic communication with an external camera configured to capture image data in a vicinity of or external to the apparatus.

[0018] In some implementations, the apparatus further includes: at least one light source positioned adjacent to the at least one imaging sensor within the imaging chamber.

[0019] In some implementations, the adhesive material or surface is positioned at approximately 35 degrees relative to an inner side surface of the apparatus.

[0020] In some implementations, the apparatus further includes: a fan chamber positioned below the imaging chamber, wherein the fan chamber includes at least one fan configured to generate negative air pressure by pulling air through the entry aperture and imaging chamber.

[0021] In some implementations, the apparatus further includes: a mesh grill separating the imaging chamber from the fan chamber.

[0022] In some implementations, the apparatus further includes: a pyramidal airflow diverter positioned at a bottom of the fan chamber configured to redirect airflow horizontally outward through open sides.

[0023] In some implementations, the apparatus further includes: an electronics chamber positioned below the fan chamber.

[0024] In some implementations, the electronics chamber includes an image capture and processing module configured to transmit the images to a cloud-based platform for species identification and classification using artificial intelligence or machine learning techniques.

[0025] In some implementations, the door includes: a frame configured to support the adhesive material or surface within the imaging chamber.

[0026] In some implementations, the frame includes: a top roller and a bottom roller mounted within the frame, and a motor operatively connected to and configured to incrementally rotate the bottom roller to expose fresh adhesive material or surface.

[0027] In some implementations, the apparatus further includes: one or more sensors operatively coupled to the frame or door, wherein the one or more sensors are configured to: detect an amount of remaining adhesive material or surface, and generate an alert when replacement of the adhesive material or surface is required.

[0028] In some implementations, the apparatus further includes: an entry aperture configured to lead one or more insects into an entry chamber within the apparatus; a door on a side surface of the apparatus, wherein a rear surface of the door is positioned within the entry chamber in a closed position; and at least one camera assembly positioned opposite the rear surface of the door configured to capture video data of the one or more insects within the entry chamber.

[0029] In some implementations, the apparatus further includes: a reversible fan configured to generate a controlled airflow through the entry chamber, wherein the reversible fan is operable in a first direction to create a negative pressure to draw insects into the entry chamber and house them therein for imaging, and wherein the reversible fan is operable in a second direction to expel insects from the entry chamber into a secondary holding chamber.

[0030] In some implementations, the secondary holding chamber is physically contiguous with the entry chamber and is sealed off from the external environment except through airflow routes controlled by the reversible fan.

[0031] In some implementations, the apparatus further includes: at least one transparent sheet positioned within the entry chamber configured to maintain the one or more insects at a specified focal range relative to the camera assembly.

[0032] In some implementations, the apparatus further includes: at least one lighting element adjacent to the camera assembly configured to illuminate an interior of the entry chamber.

[0033] In some implementations, the at least one lighting element includes one or more circular light emitting diode (LED) strips surrounding or adjacent to a camera assembly mount.

[0034] In some implementations, the apparatus further includes: an electronics chamber including at least one of a microcontroller, printed circuit boards, and a power distribution system, wherein the electronics chamber is positioned above or adjacent to the entry chamber.

[0035] In some implementations, the apparatus further includes at least one of: a mesh grill separating the entry chamber from the reversible fan, and a lid to secure or cover at least a portion of the electronics chamber.

[0036] In some implementations, a system is provided. The system can comprise: at least one apparatus, the at least one apparatus comprising: an entry aperture configured to lead one or more insects to an imaging chamber within the apparatus, a door on a side surface of the apparatus, wherein a rear surface of the door is positioned within the imaging chamber in a closed position, and wherein the rear surface of the door comprises a removable/replaceable adhesive material or surface configured to trap one or more insects thereon, and at least one imaging sensor positioned opposite the rear surface of the door configured to capture images or video of the one or more insects adhering to the adhesive material or surface, and an electronics including at least a processor or microcontroller; and at least one external camera operatively in electronic communication with the at least a processor or microcontroller, wherein the at least one external camera is configured to capture image or video data in a vicinity of or external to the apparatus (e.g., an area adjacent to the apparatus).

BRIEF DESCRIPTION OF THE FIGURES

[0037] The patent application file or the patent issuing therefrom contains at least one drawing executed in color. Copies of this patent or patent application publication with the color drawing(s) will be provided by the Office upon request and payment of the necessary fee. Reference will now be made to the accompanying drawings, which are not necessarily drawn to scale.

[0038] FIG. **1**A is a schematic representation of a smart trap assembly according to embodiments of this disclosure.

[0039] FIG. **1**B, FIG. **1**C, FIG. **1**D, FIG. **1**E, FIG. **1**F, FIG. **1**G, FIG. **1**H, FIG. **1**I, FIG. **1**J, FIG. **1**K, FIG. **1**L, FIG. **1**M, FIG. **1**N, FIG. **1**O, FIG. **1**P, FIG. **1**Q, FIG. **1**R, FIG. **1**S, FIG. **1**T, FIG. **1**U, FIG. **1**V, FIG. **1**W, FIG. **1**X are schematic diagrams showing example traps and associated components in accordance with certain embodiments described herein.

[0040] FIG. **2** is a pixel-wise segmentation of a pest using boxes within a neural network in accordance with this disclosure.

[0041] FIG. **3** is a schematic view of an output of a neural network that segments pixels of respective pest body parts and annotates an image of the pest in accordance with this disclosure.

[0042] FIG. **4** is a schematic view of a masked anatomy result of a convolutional neural network operation on a digital image according to this disclosure.

[0043] FIG. **5**A is a schematic view of a cropped thorax anatomy result of a convolutional neural network operation on a digital image according to this disclosure.

[0044] FIG. **5**B is a schematic view of a cropped abdomen anatomy result of a convolutional neural network operation on a digital image according to this disclosure.

[0045] FIG. **5**C is a schematic view of a cropped wing anatomy result of a convolutional neural network operation on a digital image according to this disclosure.

[0046] FIG. **5**D is a schematic view of cropped legs anatomy result of a convolutional neural network operation on a digital image according to this disclosure.

[0047] FIG. **6**A is a first screen shot of a software application ("app") display using in conjunction with the systems and methods disclosed herein.

[0048] FIG. **6**B is a second screen shot of a software application ("app") display using in conjunction with the systems and methods disclosed herein.

[0049] FIG. **6**C is a third screen shot of a software application ("app") display using in conjunction with the systems and methods disclosed herein.

[0050] FIG. **7**A is a fourth screen shot of a software application ("app") display using in conjunction with the systems and methods disclosed herein.

[0051] FIG. **7**B is a fifth screen shot of a software application ("app") display using in conjunction with the systems and methods disclosed herein.

[0052] FIG. **7**C is a sixth screen shot of a software application ("app") display using in conjunction with the systems and methods disclosed herein.

[0053] FIG. **8** is a schematic diagram of a convolutional neural network ("CNN") used in a computer environment configured to implement the computerized methods of this disclosure.

[0054] FIG. **9**A is a schematic illustration of consolidated segmenting results of respective species of mosquitos with each species represented by an original image, a feature map illustrated with bounding boxes used by a respective convolutional network and individually segmented anatomy parts separated out of the original image for display on a graphical processing unit of a computer used herein.

[0055] FIG. **9**B is a schematic illustration of consolidated segmenting results of respective species of mosquitos with each species represented by an original image, a feature map illustrated with bounding boxes used by a respective convolutional network, and individually segmented anatomy parts separated out of the original image for display on a graphical processing unit of a computer used herein.

[0056] FIG. **9**C is a schematic illustration of consolidated segmenting results of respective species of mosquitos with each species represented by an original image, a feature map illustrated with bounding boxes used by a respective convolutional network, and individually segmented anatomy parts separated out of the original image for display on a graphical processing unit of a computer used herein.

[0057] FIG. **10** is a schematic illustration of a training test for a convolutional neural network providing consolidated segmenting results of respective species of bumble bees with each species represented by an original image, a feature map illustrated with bounding boxes used by a respective convolutional network, and individually segmented anatomy parts separated out of the original image for display on a graphical processing unit of a computer used herein. The training test was used to illustrate that a convolutional neural network used to identify mosquito anatomy is adaptable to other species of pests.

[0058] FIG. **11** is a segmented image of a mosquito in which components of the mosquito anatomy are identified as foreground pixels and filtered out by a convolutional neural network of this disclosure.

[0059] FIG. **12** is a schematic illustration of a computer environment in which neural networks according to this disclosure are processed for anatomy segmentation and pest identification according to embodiments of this disclosure.

DETAILED DESCRIPTION

[0060] This disclosure presents the hardware design of a smart-mosquito trap that will automatically capture images, videos and wing-beat frequencies of mosquitoes that get trapped. Once the digital data is made available artificial intelligence ("AI") algorithms, such as those shown below, will be executed to identify the genus and species of the trapped specimens. This disclosure presents the design details and prototype implementation of a smart mosquito-trap that can use any attractant (CO2, Light, Human Sweat, Animal Odors, Lactic acid, etc.). The trap is embedded with the following components that make it unique a) a digital microphone/passive infrared sensors (PIR) at the entrance of the trap to sense wing-beat frequencies and size of the insect (i.e., to identify entry of a mosquito); b) a lamb-skin membrane filled with CO2 with gas air inside that mimics human skin so that the mosquito can rest on the membrane to pierce as if the mosquito is could actually consume blood; and c) multiple embedded cameras to capture images of the mosquito on the lamb-skin membrane for post processing. The combination of these components in a mosquito-trap is novel and has not been attempted before.

[0061] The process of identifying mosquitoes from trapped specimens is laborious. Mosquitoes get attracted to traps placed at strategic locations. They are then identified one-by-one by expert personnel via visual inspection through a microscope. This process, across the world, takes hours to do so and in inefficient. With the disclosed smart-trap, images, videos and wing-beat frequencies are captured immediately as a mosquito gets attracted to the trap. Then, AI algorithms can automate the process of identifying mosquito specimens. The images and results of classification and anatomies extracted can be relayed to a remote control center instantly. Currently, no such technology exists to be able to record images, videos and audio wing beat frequencies of trapped mosquitoes.

[0062] Terms used to describe the apparatus are given their broadest plain meaning. For example, an imaging device may be any device that gathers and/or transmits information to a computer by discerning physical image features of a subject. Accordingly, an imaging device may be any kind of camera but also passive infrared sensors and other kinds of data collection equipment using light transmission.

[0063] In one non-limiting embodiment, the apparatus for collecting mosquito specimens is a CO2 trap (that can be modified into a light trap, with ability to add more attractants), that emits warm CO2 to attract mosquitoes. Mosquitoes are attracted to the trap, since they are tricked into believing that the CO2 comes from a source that is a potential blood meal. FIG. **1**A is an example schematic. In FIG. **1**A, an apparatus **100** for collecting mosquito specimens has a camera **150** (including but not limited to a raspberry pi (R-Pi) based imaging device), a microphone **155**, and a network connection or direct connection to a computer processor **1200** storing software in computerized memory that is configured to receive image data and audio data to determine the classification of the species in the camera field of view. The apparatus **100** also includes a membrane body **175** that mimics an animal or human skin that the mosquito can try to pierce seeking blood. While the mosquito is perched on the membrane body, the camera **150** or any other imaging device can gather image data within the field of view. The apparatus further includes an enclosed chamber **195** to collect mosquitoes after data capture. The mosquitoes may be alive or dead and can be used for data collection verification by other manual methods. A cover **110** prevents interested mosquitoes from flying away.

[0064] The membrane body **175** is mounted on a brass puck **165** having the following non-limiting features: 1) the membrane has a gas inlet and four holes to allow the membrane to fill with CO2 gas; 2) the center of the brass puck has a recess **167** for the placement of lactic acid, or any other kind of mosquito attractor; 3) the brass puck is heated using a heater **180** (e.g., a cartridge heater) and the temperature monitored by a thermocouple (TC). 4) A raspberry pi processor controls the temperature so that warm CO2 is present to further enhance mosquito attraction. In some embodiments, the skin surface **185** is a lamb-skin prophylactic as lamb-skin is known to serve as an excellent resting surface for mosquitos. The skin surface **185** of the membrane body **175** is attached to the brass puck **165** using a rubber band as the inside diameter of the skin is slightly larger than the outside diameter of the puck. CO2 gas permeates the membrane body **175** from inside the membrane enclosure **160** and wafts to the top of the apparatus **100** to attract the mosquitos.

[0065] The camera **150** is triggered with its built-in microphone **155** that is continuously listening inside the trap. In software stored in computer memory, the audio spectrum is converted to a frequency representation of the audio using a Fast Fourier transform. Next, the software produces a frequency representation of the audio (e.g., augmax function). If the frequency is within the range of the mosquito wing beat frequency, the software triggers the camera. In non-limiting embodiments, this range is currently set to 400 to 1000 Hz. Based on performance, a user can change the value to increase or decrease the overall sensitivity. In some embodiments, Passive Infrared Sensors connected to the camera and the other electronics detect whether or not an insect coming into the trap is a mosquito or not.

[0066] This disclosure presents the design details and prototype implementation of a smart mosquito-trap that can use any attractant (CO2, Light, Human Sweat, Lactic acid, etc.). This disclosure presents a trap that is embedded with a combination of one or more of the following components that make it unique a) a digital microphone/passive infrared sensors at the entrance of the trap to sense wing-beat frequencies and size of the insect (to identify entry of a mosquito); b) a lamb-skin membrane filled with CO2 with gas air inside that mimics human skin so that the mosquito can rest on the membrane to pierce (to trick the mosquito into thinking that a blood meal is available); c) multiple embedded cameras to capture images and videos of the mosquito on the lamb-skin membrane for post processing; d) hardware platforms that can execute artificial intelligence algorithms processing the captured data to classify mosquitoes, into genus and species; e) a depth-sensing camera system to track and video/audio record the mosquitoes in flight; f) a simple funnel to reduce the entry diameter, which will be connected to the instrument chamber; g) An ultraviolet light source to detect presence of fluorescence in the insect; h) multiple chambers in the trap, each of which will be triggered by a smart actuation system, wherein the mosquito (based on how it was classified) will be made to fall for collection in the appropriate chamber; i) options for the trap to be provisioned with yeast and water to attract gravid mosquitoes; j) hardware to export data to a cloud server via WiFi/Cellular connectivity; k) sensors in the trap to collect ambient environmental data including temperature, altitude, humidity, latitude/longitude, time, ambient light sensors; I) hardware to be able to freeze the mosquitoes upon capture into the trap.

[0067] In one embodiment, the apparatus is a mosquito trap having a digital microphone and infrared sensor at the entrance of the trap to sense the wing-beat frequency. Both of these sensors will be used to classify the entry of a mosquito. The active infrared sensor can be used for a tripwire to trigger the microphone and the camera so that both instruments can record and it could be used to confirm the wingbeat frequency reported from the microphone. This can be done by applying a Fast Fourier Transformation to both and getting the frequency graph from both sensors.

[0068] In another embodiment, a lamb-skin membrane filled with CO2 is used to mimic human skin and respiration. The mosquitoes can rest on the membrane and pierce the membrane (to trick the mosquito into thinking that a blood meal is available). The lamb-skin membrane is slightly porous to allow for the CO2 gas to slowly leak out, creating a slow flow of CO2 to help attract the mosquitoes to the trap. The membrane will be below the sensors and the camera to ensure the mosquitoes go through the sensors.

[0069] In another embodiment, embedded cameras are configured to capture images of the mosquitoes on the landing surfaces. This will enable a user to capture images from multiple different angles so that the certain machine learning algorithms, incorporated by example below, can have multiple different views for the classifier. Giving the classifier different views will allow for a better classification from the images as well as better extraction of insect anatomy.

[0070] In another embodiment, using embedded hardware platforms allow for the execution of machine learning algorithms to process the captured data to classify the mosquitoes. To decrease the inference time when we use these platforms the apparatus also will employ the use of purpose-built A SIC processors such as the Google Coral Accelerator module. This will allow us to reduce the latency and reduce the amount of network traffic.

[0071] In another embodiment, a global shutter camera will also be used to capture images of the mosquitoes in flight. Being able to capture images of the mosquito in flight will allow the user to reduce the trap size and possibly avoid having a landing surface all together. A global shutter camera will also provide clearer images with reduced motion blur which will help the machine learning algorithms classify the image better.

[0072] A simple funnel will be employed to reduce the entry diameter from the top of the apparatus. The entryway will be reduced by a factor of three in one non-limiting embodiment. This funnel will also be attached to the top of the sensor chamber where the infrared sensor and microphone will be. The funnel is aimed at reducing the number of mosquitoes that enter the traps to allow for the sensors to sense one mosquito at a time.

[0073] In other non-limiting embodiments, an ultraviolet light source will be employed to detect the presence of fluorescence on the insect. With the release of genetically modified mosquitoes in south Florida, being able to detect genetically modified mosquitoes is very important. These modified mosquitoes have a fluorescence marker on them to identify them as modified. In the right light conditions, provided by the ultraviolet light, the camera will be able to pick up on these markers. The system will be able to count these markers as they come into the trap.

[0074] In some embodiments, the base of the trap will have multiple chambers. These chambers will be triggered based on the classification of the mosquito or other insect that flies into the trap. Based on the classification, a smart actuation system using a servo/motor may allow for the specimen to fall for collection in the appropriate chamber. If a general vector mosquito, collect in Chamber 1. If *Aedes aegypti*, collect in Chamber 2. If *Aedes aegypti*, and it has a fluorescent marker, then, collect in Chamber 3. If other insects, collect in the general chamber.

[0075] The trap will have multiple options for attractants. The user can use yeast or dry ice to generate CO2. The CO2 will be dispensed using a solenoid valve in the trap to control the gas output during only the specified times. Secondly, water can be used to attract gravid mosquitoes that had a blood meal and want to lay eggs. The user can also use solid state attractants that are already available on the market to allow for a more cost-effective attractant solution for the trap.

[0076] The trap will send the classified audio and images to a cloud server to be accessed by the user as they wish. To facilitate this, the trap will have WiFi and a cellular module to allow for the user to connect to the network of their choice. For remote locations, a satellite communications module can be added to allow for low bandwidth communication for counts and other critical information. Also, the trap can use LoRa for radio communications over long distances without the need for satellite communications.

[0077] Additional sensors on the trap will allow for the collection of ambient environmental data. This data will include temperature, air pressure, altitude, humidity, latitude, longitude, time, CO2 concentration and ambient light conditions. The sensors we will use include the Bosch® BME 280 for temperature and humidity, Bosch® BM P390 for air pressure and altitude, a MediaTek MTK 3339-based GPS module for latitude and longitude, Sensirion® SCD-40 for CO2 concentration, and a Rohm® BH1750 for ambient light.

[0078] The trap will be able to freeze the mosquitoes when they are in the chamber. This is done using a peltier module to bring a metal plate below freezing. The peltier module will be assisted with a fan to further reduce the temperature of the plate and allow for the peltier module to keep the plate at a constant temperature. Freezing the mosquitoes will slow the degradation of the specimens and allow for the user to do further research on the captured mosquitoes if they wish.

[0079] As noted above, the apparatus **100** is a smart trap that is useful for catching and gathering data regarding mosquitoes. The rest of this disclosure provides an example of the kinds of image and data analysis that may be used as the above noted artificial intelligence. This disclosure, therefore, incorporates by reference U.S. Pat. No. 10,963,742, entitled Leveraging Smart-Phone Cameras and Image Processing Techniques To Classify Mosquito Genus and Species, as issued on Mar. 30, 2021. The '742 patent discloses a computerized method of identifying an insect specimen, such as the genus and species of a mosquito, and includes gathering a plurality of digital images of the insect specimen positioned within a respective set of image backgrounds. The disclosure continues by extracting image portions from each digital image, wherein the image portions include body pixels of image data corresponding to the insect specimen and excluding image background pixels. The method further includes converting the body pixels into a selected color space data set and identifying textural features of the image portions from the selected color space data set.

[0080] This disclosure also incorporates by reference U.S. Pat. No. 11,048,928, entitled Systems and Methods of Entomology Classification Based on Extracted Anatomies, as issued on Jun. 29, 2021. The '928 patent illustrates a Deep Neural Network Framework to extract anatomical components, such as but not limited to, thorax, wings, abdomen and legs from mosquito images. The technique is based on the notion of Mask R-CNN **800** of FIG. **8**, wherein artificial intelligence iteratively learns feature maps **808** from images **802**, emplaces anchors (shown as bounding boxes **812**, **817** in the Figures but can be any shape) around foreground components, followed by segmenting **820** and classification **824** of pixels corresponding to the anatomical components within anchors. In some embodiments, results of this disclosure show that the techniques are favorable when interpreted in the context of being able to glean descriptive morphological markers for classifying mosquitoes.

[0081] In one embodiment of U.S. Pat. No. 11,048,928, a system for identifying a genus and species of an insect includes an imaging device configured to generate images of the insect. A computer processor is connected to memory storing computer implemented commands in software, and the memory receives the images, wherein the software implements a following computerized method with respective images beginning with a step of applying a first convolutional neural network to the respective images to develop feature maps directed to anatomical pixels in the respective images that correspond to a body part of the insect. A computer then applies anchors to the feature maps, wherein the anchors identify portions of respective layers of image data in the feature maps that contain respective anatomical pixels for respective body parts. Generating a mask allows the system to segment the respective anatomical pixels from the respective layers. The system extracts fully connected layers from the respective layers that have had the first convolutional neural network applied thereto. The fully connected layers are applied to a regressor network and a classification network, wherein generating the mask for segmenting, applying the fully connected layers to a regressor network, and applying the fully connected layers to a classification network are parallel operations conducted by the software.

[0082] Also in the embodiments of U.S. Pat. No. 11,048,928, systems and methods are disclosed for extracting information about anatomical components of a living creature from a digital image. The method includes training a mask-region based convolutional neural network with a set of training images segmented with computerized algorithms that identify ground truth anatomical components to a set degree of accuracy. The training includes classifying respective anatomical components in the training images and comparing the training images to the ground truth images; tightening bounding boxes surrounding the anatomical components in the digital images; and generating a mask for use in extracting information of a second set of images. A general discussion of CNNs and associated terminology can be found in numerous references cited below. For example, Reference **35** (Stewart) explains how filters, made of multiple kernels (weighted matrices) are convolved onto original images to create feature maps of numerous layers and adaptable data density. Stewart explains using the feature maps to pool certain layers with techniques such as max pooling, that separates out those feature maps with maximum values to reduce complexity. Rectified Non-Linear Unit (Re-LU) data sets are added to the feature maps to identify areas that should be accounted for but were missed when the changes from one pixel to the next were below a filtering threshold. In very simplistic terms, the Re-Lu is an activation function operated on the image to produce layers that may be appended to the feature maps as shown in FIG. **8**. Generally, in some non-limiting embodiments, the Re-LU may retain a certain filter's value at a respective output matrix index or insert a zero if that certain index value is negative. The overall concept of a convolutional neural network, therefore, incorporates convolutional layers as feature maps of the original image, pooling layers and Re-LU layers for added detail, as well as fully connected layers that are data rich outputs that are combined. As noted at Ref. 35 (Stewart), the fully connected layers, such as those shown in the non-limiting example of FIG. **8**, aggregate all information into a finally replicated image.

[0083] In some aspects, the present disclosure relates to computerized apparatuses, computer implemented methods, and computerized systems that use digital image analysis to identify species of insect specimens, such as, but not limited to mosquitos. The disclosure presents a system wherein a user (expert or an ordinary citizen) takes a photo of a mosquito or other pests, using a smart-phone, and then the image is immediately sent to a central server along with GPS information data of the smart-phone.

[0084] The server will implement algorithms described in this disclosure to a) identify the genus of the mosquito; b) identify the species of the mosquito; c) separate the body parts of the image into objects of interest like wings, legs, proboscis, abdomen, scutum etc.; d) give feedback on species and genus back to user, along with information as to what diseases the species carry, and more interesting information like flight range etc. Potential uses are in mosquito identification, since it is a painful and cognitively demanding problem now. School districts could also use this software application to teach kids about biology and other areas of science, given that these kinds of scientific analysis skills may eventually be mandatory for schools in many areas. Defense and Homeland Security agencies and other government agencies may see a need for the computerized application described herein.

[0085] One non-limiting value proposition of this disclosure is the ability to bypass humans (that peer through a microscope currently) for classification, and instead use digital cameras and proposed techniques for automated classification of genus and species type. A secondary value proposition is the ability of a system with large scale citizen and expert generated imagery, with tagging, to start digitizing anatomies of mosquitoes across the globe. This database could prove invaluable for training, and global information sharing in the context of mosquito, and especially vector surveillance.

[0086] Although example embodiments of the present disclosure are explained in detail herein, it is to be understood that other embodiments are contemplated. Accordingly, it is not intended that the present disclosure be limited in its scope to the details of construction and arrangement of components set forth in the following description or illustrated in the drawings. The present disclosure is capable of other embodiments and of being practiced or carried out in various ways. For example, the test results and examples all pertain to identification of genus and species of mosquitos from the mosquito traits and features extracted from digital images. The techniques and concepts utilized and claimed in this disclosure, however, are not limited to mosquitos, but can be used with other kinds of identification processes for other animals, humans, plants and the like.

[0087] FIG. **1**A illustrates a system that utilizes an apparatus **100** for pests that captures the pests, takes their picture using a built-in camera **150**, and sends the pictures to the cloud. In the cloud, algorithms that are implemented by computers and servers on various networks are designed to identify the type of pest. Information is fed back to farmers who can then plan accordingly to treat their crops. However, anatomies are not extracted in this product.

[0088] Referring now to FIG. **1**B, FIG. **1**C, FIG. **1**D, FIG. **1**E and FIG. **1**F, views of an example apparatus **100** (the terms mosquito trap device, device, trap, and apparatus are used interchangeably herein) in accordance with certain embodiments described herein are provided. In various implementations, the apparatus **100** is configured to preserve mosquito integrity by avoiding the use of adhesives, toxins, or mechanical impellers in the capture zone, thereby facilitating the collection of undamaged, live specimen footage. In some examples, high-speed video footage can be collected and processed by a machine learning or artificial intelligence model trained to identify mosquito species known to carry waterborne diseases. The apparatus **100** is configured to obtain image and/or video data of trapped mosquitos for identification and classification. This disclosure contemplates that other types of insects can be identified using the apparatus **100** including, for example, but not limited to, flies (e.g., fruit flies, blow flies, tsetse flies), fleas, bees, moths, and the like.

[0089] As shown in FIG. **1**B and FIG. **1**C, the example apparatus **100** comprises a rectangular shaped housing including a vertically oriented door **102** that is accessible via a handle **103**. In some implementations, the housing is printed in white Polylactic acid (PLA) plastic with approximately 15% infill to create a uniform background and enhance contrast in the recorded video footage.

[0090] FIG. **1**B depicts a front view of the apparatus **100** with the door **102** closed, while FIG. **1**C depicts a front view of the apparatus **100** with the door **102** open. FIG. **1**D shows a left side view (e.g., side portion or panel) of the apparatus **100** and FIG. **1**E shows a right side view of the apparatus **100**. FIG. **1**F shows a rear view (e.g., back portion or panel) of the apparatus **100**. FIG. **1**G shows a top view of the apparatus **100**. In some embodiments, the trap housing is at least partially 3D-printed from Polylactic acid (PLA) plastic and dimensioned to form a substantially rectangular enclosure comprising two mated bodies, each having a width of about 120 mm, a combined length of about 288 mm, and a height of about 113.9 mm (excluding any top lid).

[0091] In various implementations, one or more mosquitos or other insects enter the apparatus **100** through a specialized entry chamber/aperture on a surface of the apparatus/trap **100**. In some embodiments, the apparatus **100** includes an entry aperture (e.g., **117** shown on a top surface of the apparatus **100** in FIG. **1**I) positioned on a surface (e.g., top surface, side surface, bottom surface) of the apparatus **100** that is configured for mosquito entry. An internal portion of the apparatus **100** comprises/defines an imaging chamber **105** for obtaining images and/or video data of the one or more mosquitos. A rear surface of the trap door **102** (i.e., oriented toward or facing the inner portion/imaging chamber **105**) can comprise an adhesive material, referred to herein as a sticky pad or surface (e.g., **118** as shown in FIG. **1**N) configured to trap an insect thereon. The sticky pad (e.g., **118**) is removable and replaceable for maintenance and mosquito retrieval.

[0092] As depicted in FIG. **1**C, in some implementations, the apparatus **100** includes one or more imaging sensors **107**A, **107**B (e.g., cameras, camera assemblies, autofocus cameras, or the like) positioned opposite the sticky pad (e.g., **118**) within the imaging chamber **105** to capture images and/or video data of mosquitoes adhering to the pad **118** located within (e.g., or flying within) the imaging chamber **105**. The imaging chamber **105** can comprise a geometry designed to funnel mosquitoes toward the center focal point of the cameras **107**A, **107**B or to a focal point of a single camera, ensuring improved clarity and reducing occurrences of the mosquitoes being captured at the periphery of the field of view. In some embodiments, the imaging chamber **105** includes a light source (e.g., shown in FIG. **1**L) positioned adjacent to the cameras **107**A, **107**B within the imaging chamber **105** to provide illumination ensuring clear image or video capture. In some implementations, the apparatus **100** is in electronic communication with an external camera configured to capture image data in a vicinity of or external to the apparatus. In some embodiments, the external camera can be either static or rotating about an angle that is controlled by a controller of the apparatus (e.g., R-PI) to continuously record the ambient environment, and also the external housing of the apparatus **100**. The external camera can be used for troubleshooting purposes should the apparatus/trap **100** malfunction.

[0093] In some embodiments, as shown in FIG. **1**C, the apparatus **100** includes a fan chamber **109** positioned below the imaging chamber **105**, containing a fan **111** (e.g., reversible fan) configured to generate negative air pressure by pulling air through the top entry hole (e.g., **117** as shown in FIG. **1**I) and imaging chamber **105**. In some implementations, the apparatus **100** includes a mesh grill **112** separating the imaging chamber **105** from the fan chamber **109**, preventing movement or passage of the mosquito from the imaging chamber **105** into the fan chamber **109**. In other words, the mesh grill **112** is dimensioned to allow airflow but prevent mosquitoes from passing through. In some embodiments, the fan **111** can be positioned and/or reoriented (e.g., vertically, blowing air out of the trap sideways) to ensure that mosquitoes do not contact the fan blades during entry of the apparatus **100**. In some embodiments, the fan **111** is arranged to generate a controlled airflow through the entry chamber/imaging chamber **105**. For example, the fan **111** can operate in a first direction to create a negative pressure, drawing mosquitoes into the entry chamber and keeping/trapping them there for imaging. In some embodiments, the fan **111** is powered by a CPU-style motor (e.g., motor **124** shown in FIG. **1**P) that is configured to operate continuously or intermittently so as to maintain negative and positive pressure that retains mosquitoes within the chamber during the imaging process.

[0094] Additionally, in some implementations, the fan **111** is reversibly operable in a second direction to push mosquitoes out of the entry chamber into the secondary holding chamber (e.g., from entry chamber **105***a* to secondary holding chamber **105***b* as shown in FIG. **1**W). In some implementations, the secondary holding chamber **105***b* is positioned adjacent to or downstream of the entry chamber **105***a* configured to temporarily store mosquitoes after recording. In some embodiments, the secondary holding chamber **105***b* is physically contiguous with the entry chamber **105***a* and is sealed off from the external environment except through airflow routes controlled by the reversible fan **111**.

[0095] In some implementations, a camera assembly is mounted within the imaging chamber **105** such that its lens is directed into the mosquito entry chamber (**105***a*). The camera assembly can comprise a GoPro Hero 12 Black camera or equivalent high-speed camera capable of recording at about 240 frames per second and 2.7K resolution. The camera can incorporate a macro lens attachment configured to capture magnified footage of live mosquitoes. The camera mount can be formed by a cutout or aperture in the housing that allows the GoPro Hero 12 Black lens (with macro attachment) to extend into the mosquito entry chamber **105***a*, ensuring close-up capture of mosquitoes within the chamber.

[0096] In some embodiments, at least one transparent acrylic sheet is positioned within the mosquito entry chamber (**105***a*) and configured to maintain mosquitoes within a focal range of about 63 mm from the camera lens while preventing direct contact between the camera (**107**) and the mosquitoes. In some embodiments, the transparent acrylic sheet is positioned to prevent mosquitoes from escaping the focal range, while also minimizing image distortion and preventing contact damage to the mosquitoes. Optionally, at least two lighting elements (e.g., light sources **108***a-d* shown in FIG. **1**I) are arranged adjacent to the camera assembly and configured to illuminate the interior of the mosquito entry chamber (**105***a*) for high-speed video capture. The lighting elements **108***a-d* can be powered and controlled by on-board electronics in the electronics chamber **130**. In some embodiments, as shown in FIG. **11**, the lighting element(s) comprise two circular LED strips arranged around or near the camera mount, providing uniform illumination for high-resolution and high-speed footage of mosquito flight. Additionally, in some embodiments, the apparatus **100** includes a lid or cover secured above the electronics chamber **130**. The lid can facilitate inspection or maintenance of internal components without disturbing the camera assembly or transparent acrylic sheet.

[0097] In some implementations, a microcontroller is located in the electronics chamber **130** above or adjacent to the mosquito entry chamber **105***a*. In various examples, the microcontroller is configured to control video recording intervals of the camera, control activation of the lighting element, upload recorded footage to a cloud storage service, control fan speed and direction, and/or control a motorized door **102**. In some embodiments, the microcontroller is a Raspberry Pi configured with an automated script that periodically records about 15 seconds of video every 30 seconds and subsequently uploads said video to a remote storage platform for species identification using machine learning.

[0098] In some embodiments, the example apparatus **100**/electronics chamber **130** includes a power distribution system comprising at least one custom printed circuit board (PCB) and any necessary wiring pass-throughs to supply the camera **107**, lighting element(s) **108***a-d*, CPU fan **111** with power and data/signals. The electronics chamber **130** can further include ports for: connecting the camera **107** to external wiring, at least one power adapter input, access points for microcontroller, enabling external data retrieval or system updates, HDMI and USB inputs, and/or external wiring for motorized door **102** and CPU fan **111**.

[0099] In some implementations, the apparatus **100** includes a pyramidal airflow diverter **113** positioned at the bottom of the fan chamber **109** configured to redirect airflow horizontally outward through open sides of the diverter **113**. Additionally, as shown, the apparatus **100** includes an electronics chamber **130** positioned below the fan chamber **109**, housing electronics **115** that control the cameras **107**A, **107**B, lighting (e.g., one or more light sources **108***a-d* shown in FIG. **1**I), and the fan **111**. As depicted in FIGS. **1**B-F, the pyramidal airflow diverter **113** comprises sloped surfaces designed to evenly distribute airflow through multiple open sides (e.g., apertures or opening on side surfaces of the apparatus **100**) of the fan chamber **109**. The electronics **115** can include an image capture and processing module configured to transmit captured mosquito images to a cloud-based platform for species identification and classification using artificial intelligence or machine learning techniques.

[0100] Referring now to FIG. **1**H, FIG. **1**I, FIG. **1**J, FIG. **1**K, FIG. **1**L, and FIG. **1**M views of another example apparatus **100** in accordance with certain embodiments described herein are provided. As illustrated in FIGS. **1**H and **1**I, the example apparatus **100** comprises a rectangular shaped housing including a vertically oriented door **102** that is accessible via a handle **103**. FIG. **1**H depicts a front view of the apparatus **100** with the door **102** closed, while FIG. **1**I depicts a front view of the apparatus **100** with the door **102** open. FIG. **1**J and FIG. **1**K show a left side view and right side view of the apparatus **100**, respectively. FIG. **1**L shows a rear view of the apparatus **100**. FIG. **1**M shows a top view of the apparatus **100**.

[0101] As shown in FIG. **1**H and FIG. **1**L, similar to the example shown in FIG. **1**B and FIG. **1**C, the apparatus **100** includes an entry aperture (e.g., **117** shown in FIG. **1**H and FIG. **1**I) positioned on a surface (e.g., top surface, side surface) of the apparatus **100** that is configured for mosquito entry. An internal portion of the apparatus **100** comprises/defines an imaging chamber **105** for obtaining images of one or more mosquitos. A rear surface of the trap door **102** (i.e., oriented toward or facing the inner portion/imaging chamber **105**) can comprise a sticky pad or surface (e.g., **118**) configured to trap an insect thereon. The sticky pad (e.g., **118**) is removable and replaceable for maintenance and mosquito retrieval.

[0102] In the example shown in FIG. **1**L, the imaging chamber **105** is positioned at the top of the apparatus **100**, defining a top portion (e.g., section) of the apparatus **100**. As shown, the imaging chamber **105**/camera **107** is tilted approximately 35 degrees towards an inner surface of the door **102**. Additionally, the apparatus **100** comprises a top entry hole **117** configured for mosquito entry. The door **102** is oriented toward an inner bottom surface (e.g., the floor side) within the imaging chamber **105** and is configured to hold a replaceable sticky pad (e.g., **118** shown in FIG. **1**I) for capturing mosquitoes is positioned at approximately 35 degrees (more or less) relative to an inner side surface of the apparatus **100**. The apparatus **100** can include at least two autofocus cameras (**107**A, **107**B) positioned opposite the sticky pad **118** within the imaging chamber **105** to capture images of mosquitoes adhering to the pad **118**. The imaging chamber **105** includes at least one light source (e.g., **108**) positioned adjacent to the cameras **107**A within the imaging chamber **105** to provide illumination ensuring clear image capture. In the example shown in FIG. **1**I, a plurality of light sources **108***a*, **108***b*, **108***c*, **108***d* are positioned around/surround the camera **107**.

[0103] As illustrated in FIG. **1**L, the apparatus **100** comprises a fan chamber **109** positioned below the imaging chamber **105** that contains a fan **111** configured to generate negative air pressure by pulling air through the top entry hole **117** and imaging chamber **105**. In various embodiments, the fan **111** is positioned to ensure mosquitoes do not contact the fan blades during entry. As further depicted, a mesh grill **112** separates the imaging chamber **105** from the fan chamber **109**, preventing mosquito passage therebetween. The mesh grill **112** is dimensioned to allow airflow but prevent mosquitoes from passing through.

[0104] The apparatus **100** includes a pyramidal airflow diverter **113** positioned at the bottom of the fan chamber **109** configured to redirect airflow horizontally outward through open sides (**113***a*, **113***b*). The pyramidal airflow diverter **113** comprises sloped surfaces designed to evenly distribute airflow through multiple open sides of the fan chamber **109**.

[0105] The apparatus **100** includes an electronics chamber **130** positioned below the fan chamber **109**, housing electronics **115** that control the cameras (**107** etc), lighting (**108** etc), and fan **111**. The electronics chamber **130** can include an image capture and processing module configured to transmit captured mosquito images to a cloud-based platform for species identification and classification using artificial intelligence or machine learning techniques.

[0106] Referring now to FIG. **1**N, FIG. **1**O, FIG. **1**P, FIG. **1**Q, and FIG. **1**R, example views showing a trap door **102** and sticky pad **118** in accordance with certain embodiments of the present disclosure are provided. FIG. **1**N shows an inner surface of the door **102**, FIG. **1**O shows an external surface of the door **102**, FIG. **1**P shows a left profile view of the door **102**, FIG. **1**Q shows a right profile view of the door **102**, and FIG. **1**R show a top view of the door **102**.

[0107] As illustrated in FIGS. **1**N-**1**R, the door **102** includes a frame **120** configured to hold a roll of sticky pad material **118**. The frame **120** can be fixedly or removably attached to at least a portion or surface of the door **102**. When attached, the frame **120** is positioned within an imaging chamber **105** of the apparatus **100**. In some embodiments, the frame **120** is dimensioned to orient the sticky pad material **118** at an angle optimized for mosquito capture and image clarity.

[0108] As illustrated, the frame **120** can include a top roller **122***a* mounted within the frame and configured to hold a roll of sticky pad material **118**. The frame **120** also includes a bottom roller **122***b* mounted within the frame **120** that is configured to collect and store used sticky pad material **118**.

[0109] In some implementations, a motor **124** operatively connected to the bottom roller **122***b* operates to incrementally rotate the bottom roller **122***b*, thereby continuously pulling sticky pad material **118** from the top roller **122***a* and exposing fresh sticky pad surface **118**. The motor **124** and/or associated components can be positioned in the electronics chamber **130** of the apparatus **100**. The frame **120** can include means for securing an initial portion of the sticky pad material **118** onto the bottom roller **122***b*, facilitating the controlled advancement of the sticky pad **118**. In some embodiments, the motor **124** is controlled electronically to advance the sticky pad material **118** after mosquito image capture events. The sticky pad material **118** can be replaced by loading a new roll onto the top roller **122***a* and securing its initial portion to the bottom roller **122***b*. In some implementations, the door **102**/frame **120** includes one or more sensors **126** (e.g., proximity sensors, pressure sensors, or the like) configured to detect the amount of sticky pad material **118** remaining and provide alerts when replacement is required.

[0110] FIG. **1**S, FIG. **1**T, and FIG. **1**U are views of an example apparatus **100** in accordance with various embodiments of the present disclosure. FIG. **1**V, FIG. **1**W, and FIG. **1**X are views of an example apparatus **100** in accordance with various embodiments of the present disclosure.

[0111] FIG. **2** illustrates results of an approach used in one non-limiting example of digital segmentation of an insect **200**, i.e., the mosquito as shown, utilizing convolutional neural networks (CNNs). This procedure is based on the notion of Mask R-CNN described in the article cited as He, Kaiming, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. "Mask r-cnn." In Proceedings of the IEEE international conference on computer vision, pp. 2961-2969. 201, cited as Ref. 32. Another example of CNNs is discussed in an online resource, entitled "Simple Introduction to Convolutional Neural Networks" by Dr. Matthew Stewart, cited as Ref. 35. Both of these articles are incorporated by reference as if each is set forth in its entirety herein. This disclosure leverages such advances in convolutional neural networks (CNNs) and segments pixels containing each anatomical component of interest by adding a branch for predicting an object mask (i.e., pixel-wise segmentation discussed further in regard to FIG. **8**) in parallel with the existing branch for recognizing the bounding box of prior art CNNs. FIG. **2** illustrates one step of a CNN anatomical segmentation that uses regressively trained bounding boxes **210**, **220**, **230**, **240** to isolate and identify corresponding pixels on each of a respective thorax, wings, abdomen, and legs portion of an image. As noted above, the term "bounding boxes" is illustrative for example only, as the outlining used for segmenting an anatomy may take any shape, and boxes or rectangles of FIG. **2** are not limiting of this disclosure.

[0112] In this approach, several challenging steps need to be executed. A first step includes training the model using pretrained convolutional neural networks (CNNs) to generate proposals about certain regions where there might be an object within the image. Without limiting this disclosure, one example embodiment used ResNet101 as a backbone convolutional model. Initialization of the model was done using the pretrained MS COCO dataset weights. The MS Microsoft Common Objects in Context (COCO) data set has been disclosed at Lin, Tsung-Yi, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick, "Microsoft coco: Common objects in context" in European conference on computer vision, pp. 740-755; Springer, Cham, 2014, cited at Ref. 34, which is incorporated by reference in its entirety as if set forth fully herein. MS COCO dataset is large scale object detection dataset. It contains 1.5 million object instances, and 80 object categories.

[0113] ResNet has been generally described at He, K aiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016, which is incorporated by reference in its entirety as if set forth fully herein and cited at Ref. 33. ResNet is characterized in part as having a very deep network and introduces a residual connection to get the input from the previous layer to the next layer. The residual connection helps in solving gradient vanishing problems by detecting the smallest of differences between layers of the convolution. The next step is to design an object detector network that does three tasks: classifying the bounding boxes **210**, **220**, **230**, **240** with respective anatomies, tightening the boxes, and generating a mask **818**A (i.e., pixel-wise segmentation **820**) of each anatomical component. In constructing the architecture of the object detector network, non-limiting examples of this disclosure have used per-pixel sigmoid, and binary cross-entropy loss function (to identify the "k" anatomical components) and rigorously train them.

[0114] In regard to generating training data sets, non-limiting examples utilize tools that create a mask for each anatomical component in a subject dataset. To start the training, the procedure first annotated 571 mosquito images, using Visual Geometry Group (VGG) Image annotator tool which is itself a very tedious job. An example of an annotated image **300** is shown in FIG. **3**. Out of 571 mosquito images that were previously annotated by experts, 404 images were separated out as the training images and 167 images are separated out as the validation images. Next, the methods and systems of this disclosure iterate a model to optimize weights and hyper-parameters from these known-to-be-accurate sets of training data. For example, embodiments herein have optimized hyper-parameters like base feature extractor model, learning rate, momentum, optimizer, steps per epoch, validation steps and number of epochs, which are all parameters to be set in convolutional neural networks (CNNs) used herein. Example, non-limiting details are below in Table 1.

TABLE 1

Parameter	Value
Optimizer	Adam
Momentum	0.9
Learning rate	1e−4 for first 50 epochs, 1e−5 for next 50 and 1e−6 for next 100 epochs
Batch Size	2
Steps per epoch	202
Validation steps	84
Number of epochs	200

[0115] The metrics to measure the accuracy of mask R-CNN algorithm is mA P (Mean Average Precision). It was calculated by taking the mean of all the average precision across all classes over all Intersection over Union (IoU) thresholds, and is 0.833. The metric IoU measures the intersection of ratio of pixels that belong to the ground-truth of the object in the bounding box and the union of the predicted and the ground truth ratio of pixels in the box. In one design, the IoU threshold was set as 0.75. FIG. **3** illustrates how a training image **300** can be annotated to illustrate anatomical parts represented by respective anatomy part pixels **310**, **320**, **330**, **340**. In one non-limiting embodiment, a computer implemented software package also adds the written anatomy name for the part of the body that has been segmented for training purposes as shown in FIG. **3**. With the accurately segmented images for examples, each of the hyperparameters of Table 1 may be optimized to repeat the results of the trained data set.

[0116] FIG. **4** shows the masked anatomy **400** in a result in which bounding boxes **410**, **420**, **430**, **440** are honed in via regressive processes of a convolutional neural network on pixels that correspond to particular parts of the anatomy, and FIGS. **5**A-**5**D illustrate the cropped anatomy results for a specimen's thorax pixels **510**, abdomen pixels **520**, wing pixels **530**, and leg pixels **540**, after the convolutional neural networks have found these respective pixel sets via the procedure outlined in FIG. **8** and background pixels **470** have been extracted. Results are subject to cloud storage over a network.

[0117] Non-limiting embodiments of this disclosure led to a development and design for a smartphone app in Android and iOS that when enables a user either to take an image of a mosquito, or choose one from the local storage of the phone. The app will classify the mosquito and also extract anatomical pixels corresponding to anatomies of interest-thorax, wing and abdomen, and even legs as shown in FIGS. **6** and **7**.

[0118] This disclosure incorporates a framework based on a Mask Region-Based Convolutional Neural Network to automatically detect and separately extract pixels corresponding to anatomical components of mosquitoes, particularly the thorax, wings, abdomen and legs from images. In one non-limiting embodiment, a training dataset consisted of 1500 smartphone images of nine mosquito species trapped in Florida. In the proposed technique, the first step is to detect anatomical components within a mosquito image. Then, as discussed herein, the systems and methods of this disclosure localize and classify the extracted anatomical components, while simultaneously adding a branch in a neural network architecture to segment pixels containing only the anatomical components.

[0119] To evaluate generality, this disclosure tests example architectures on bumblebee images as shown in the schematic illustration **1000** of FIG. **10**, when the architectures have been trained only with mosquito images. The procedures of this disclosure have indicated favorable results.

[0120] Mosquito-borne diseases are still major public health concerns. A cross the world today, surveillance of mosquito vectors is still a manual process. Steps include trap placement, collection of specimens, and identifying each specimen one by one under a microscope to determine the genus and species. Unfortunately, this process is cognitively demanding and takes hours to complete. This is due, in part, because mosquitoes that fall into traps include both vectors for disease as well as many that are not vectors. Recently, AI approaches are being designed to automate classification of mosquitoes. Works like design machine learning models (Refs. 1-4) are based on hand-crafted features from image data that are generated from either smartphones or digital cameras. Two recent papers design deep neural network techniques (that do not need hand-crafted features) to classify mosquitoes from image data generated via smartphones (Refs. 5, 6). Other works process sounds of mosquito flight for classification, based on the notion that wing-beat frequencies are unique across mosquito species (Refs. 7-10).

[0121] In this disclosure, the work demonstrates novel applications for mosquito images when processed using AI techniques. The most descriptive anatomical components of mosquitoes are the thorax, abdomen, wings and legs, and this disclosure presents a technique that extracts just the anatomical pixels corresponding to these specific anatomical components from any mosquito image. The technique is based on M ask Region-based Convolutional Neural Network (Ref. 11).

[0122] This disclosure utilizes procedures of convolutional neural networks (CNNs), including feature maps **808** illustrated in FIG. **8**. In CNN theory, a neural network **800** applies weighted filters to images **802** for respective purposes. In a very simplistic sense, the filters are weighted to extract very precise portions, called features, from the image **802**. In some embodiments, the weights are designed from training scenarios looking for gradient changes denoting edges in certain parts of the image. The respective result of each filter and its weights, as applied to the image, is a feature map **808**. While the feature maps **808** of FIG. **8** are shown as two dimensional schematic representations, this disclosure includes feature maps in which a CNN applies feature filters with layers of kernels that output multi-layer feature maps. In other words, the feature maps **808** may have respective layers of image data. The image data may be processed globally as a set or individually as respective layers that have had a first convolutional neural network **806** applied thereto. A first step of this disclosure includes extracting respective feature maps **808** for anatomical features of interest, from a training dataset of 1500 smartphone images of 200 mosquito specimens spread across nine species trapped in Florida. The first convolutional neural network **806** to extract feature maps **808** is ResNet-101 with a Feature Pyramid Network (Ref. 12) (an architecture that can handle images at multiple scales, and one well suited for our problem).

[0123] As shown in FIG. **8**, the systems and methods herein include applying bounding boxes **817** that are tailored to mark the feature maps **808** for respective features of the image, such as respective anatomical portions of a mosquito's body in the examples of this disclosure. The bounding boxes **812**, **817** are organized pursuant to a region proposal network **816** for each feature of interest. As noted above, the examples of the figures are not limiting, as the bounding boxes may take any shape, including but not limited to rectangular. In one non-limiting example, the bounding boxes **812**, **817** are proposed as shown at **816** based upon computer driven systems learning from the training sets that gradient changes in the pixels of respective convolved image layers may correspond to a certain anatomical feature if the gradient occurs in a certain area of the convolved image layer. The systems and methods utilize regressive processes, loss theories, and feedback from one feature map to the next to make the bounding boxes more and more precise and more tailored to one feature of interest (e.g., a thorax, a leg, an abdomen, a wing for an insect). Suggested separations, or segmenting, of pixels correspond to these features of interest.

[0124] In certain non-limiting embodiments, the process of honing in the bounding boxes **812**, **817** for respective sets of anatomical pixels making up a body part is paired with an alignment process **814** that ensures that the output of the region proposal network **816** still matches outlines set forth in the original feature maps **808**. Once this alignment is complete, and as shown in the non-limiting example of FIG. **8**, the systems and methods disclosed herein are subject to masking operations, or pixel extraction, in a second convolutional neural network **818**B. The second convolutional neural network provides segmented images **820** in which certain examples result in anatomical pixels corresponding to a thorax, abdomen, wing, and leg of an insect. The output of the bounding boxes **812**, **817** applied by the region proposal network **816** is also fed to fully connected neural network layers **822**. It is notable that the second convolutional neural network **818**B utilizes convolutional layers that are filtered so that each "neuron" or matrix index within a data layer subject to a convolution are separately calculated and more sparse. The fully connected layers **822** track each prior layer more closely and are more data rich. The last fully connected layer is transmitted to both a classifier **824** and a boundary box regressor **826**. The fully connected layers **822** are actually tied to each other layer by layer, neuron by neuron as shown by the arrows. The final fully connected layer **834** is the output layer and includes all data for all layers. In separate parallel operations, a boundary box regressor **826** and a classification processor **824** are applied to each layer of the first convolutional neural network **806** and/or the second convolutional neural network **818**B. The bounding box regressor **826** utilizes error function analyses to regressively tighten the bounding boxes **812** more accurately around a respective feature of interest. This kind of feedback loop **850** ensures that the bounding boxes **812**, **817** of the region proposal network **816** provide convolved image layers that are distinct for each feature sought by the feature maps **808**. The classifier **824** provides automated computerized processes to identify and label respective sets **828**, **830** of anatomical pixels identifying each anatomical part of the subject insect from the original image **802**.

[0125] Subsequently, this disclosure sets forth steps to detect and localize anatomical components only (denoted as foreground) in the images in the form of rectangular anchors as illustrated in FIG. **2**. The term anchors is a broader term for the above described "bounding boxes" **812**, **817**. The anchors, therefore, may take any shape and the rectangular bounding boxes **812**, **817** are non-limiting examples of anchors. Once the foreground is detected, the next step is to segment the foreground pixels (e.g., FIG. **4**, Refs. **410**, **420**, **430**, **440**) from the background pixels **470** by adding a branch to mask (i.e., "extract pixels of") each component present in the foreground. This extra branch is shown in FIG. **8** as the second convolutional neural network **818**B that is done in parallel with two other branches **824**, **826** to classify the extracted rectangular anchors and to tighten them to improve accuracy via the feedback loop **850**. FIG. **11** shows how foreground pixels corresponding to a thorax **1125**, an abdomen **1135**, a wing **1145**, and a leg **1155** are extracted and may be reassembled into an image **1100** as shown with literal identifier labels superimposed thereon.

[0126] E valuation of the technique reveals favorable results. As shown in FIG. **4**, one can see that anatomical pixels corresponding to the thorax, wings, abdomen and legs are extracted with high precision (i.e., very low false positives). For legs though, in some non-limiting embodiments false negatives are higher than others, since the number of background pixels overwhelm the number of leg pixels in the image. Nevertheless, one can see that enough descriptive features within the leg of a mosquito are indeed extracted out, since mosquito legs are long, and the descriptive features, such as color bands, do repeat across the leg.

[0127] This disclosure explains that extracting images of mosquito anatomy has an impact towards (a) faster classification of mosquitoes in the wild; (b) new digital-based, larger-scale and low-cost training programs for taxonomists; (c) new and engaging tools to stimulate broader participation in citizen-science efforts and more. Also, to evaluate generality, this disclosure incorporates testing of an architecture trained on mosquito images with images of bumblebees (which are important pollinators).

[0128] Overall, results show excellent accuracy in extracting the wings, and to a certain extent, the thorax, hence demonstrating the generality of the technique for many classes of insects. Training has enabled a Mask Region-Based Convolutional Neural Network (M ask R-CNN) to automatically detect and separately extract anatomical pixels corresponding to anatomical components of mosquitoes-thorax, wings, abdomen and legs from images. For this study, this disclosure illustrates 23 specimens of *Aedes aegypti* and *Aedes infirmatus*, and 22 specimens of *Aedes taeniorhynchus, Anopheles crucians, Anopheles quadrimaculatus, Anopheles stephensi, Culex coronator, Culex nigripalpus* and *Culex salinarius*. After imaging the specimens via multiple smartphones, the dataset was 1600 mosquito images. These were split into 1500 images for training the neural network, and 100 images for validation. Together, this dataset yielded 1600 images of thorax, 1600 images of abdomen, 3109 images of wings and 6223 images of legs. These data were used to train the architecture illustrated in FIG. **8** on an Nvidia graphic processing unit (GPU) cluster of four Geforce GTX TITAN X cards having 3,583 cores and 12 GB memory each. It took 48 hours to train and validate the architecture.

[0129] For testing in this disclosure, the research disclosed herein trapped and imaged (via smartphones) another set of 27 mosquitoes, i.e., three per species. The testing data set consisted of 27 images of thorax and abdomen, 48 images of wings and 105 images of legs. FIG. **9**A, FIG. **9**B, and FIG. **9**C are each schematic illustrations **900**A, **900**B, **900**C, respectively, showing example results. One embodiment presents results of a technique to extract anatomical components **945**A-C, **955**A-C, **965**A-C, **975**A-C of a mosquito in FIGS. **9**A, **9**B, **9**C for one sample image **925**A-C among the nine species in the testing dataset. These figures are representative of all other images tested in FIGS. **9**A, **9**B, **9**C with each species, showing an original image **925**A, **925**B, **925**C, respective output layers **935**A, **935**B, **935**C of a convolutional neural network, and extracted anatomical pixels corresponding to a respective thorax portion **945**A, **945**B, **945**C, abdomen portion **955**A, **955**B, **955**C, wing portion **965**A, **965**B, **965**C, and legs portion **975**A, **975**B, **975**C. The anatomical components are indeed coming out clearly from image data **935**A-C processed with bounding boxes utilized by a convolutional neural network.

[0130] Next, the systems and methods herein quantify performance for the entire dataset using four standard metrics: Precision, Recall, Intersection over Union (IoU) and Mean Average Precision (mA P). Precision is basically the fraction of relevant instances (here, pixels) among those instances (again, pixels) that are retrieved. Recall is the fraction of the relevant instances that were actually retrieved. IoU is a metric that assesses the ratio of areas of the intersection and the union among the predicted pixels and the ground truth. A higher IoU means more overlap between predictions and the ground-truth, and so better classification. To define a final metric, the Mean Average Precision (mAP), this disclosure defines another metric, Average precision (AP), which is the average of all the Precision values for a range of Recall (0 to 100 for our problem) at a certain preset IoU threshold and for a particular class among the four for our problem (i.e., wings, thorax, legs and abdomen). This metric essentially balances both Precision and Recall for a particular value of IoU for one class. Finally, the Mean Average Precision (mA P) is the average of A P values among all our four classes.

[0131] The Precision and Recall values for the validation and testing datasets are presented in Tables 1 and 2 respectively for various values of IoU. As shown, the performance metrics in the validation dataset during training match the metrics during testing (i.e., unseen images) and post training across all IoUs. This is convincing evidence that the architecture is robust and not overfitted.

[0132] Precision for all classes is high, which means that false positives are low. Recall is also high for the thorax, abdomen and wings, indicating low false negatives for these classes. However, Recall for legs class is relatively poor. It turns out that a non-trivial portion of the leg pixels are classified as the background in one non-limiting architecture. While this may seem a bit discouraging, in FIGS. **9**A, **9**B, **9**C, a very good portion of the legs are still identified and extracted correctly by the disclosed architecture (due to the high Precision). As such, the goal of gleaning the morphological markers from all anatomical components is still enabled. Finally, the mean average precision is presented in Table 4 for all classes. The lower numbers in Table 4, are due to poorer performance for classifying legs, as compared to thorax, abdomen and wings.

TABLE 2

Precision and Recall for Different IoU Thresholds on Validation Set						
	IoU ratio = 0.30		IoU ratio = 0.50		IoU ratio = 0.70	
Anatomy	Precision (%)	Recall (%)	Precision (%)	Recall (%)	Precision (%)	Recall (%)
Thorax	94.57	95.15	99.32	89.69	99.09	66.67
Abdomen	95.27	90.96	96.37	85.80	99.17	77.41
Wing	98.17	91.49	98.53	85.50	97.82	76.59
Leg	99.35	37.85	100	25.60	100	21.50

TABLE 3

Precision and Recall for Different IoU Thresholds on Testing Set.						
	IoU ratio = 0.30		IoU ratio = 0.50		IoU ratio = 0.70	
Anatomy	Precision (%)	Recall (%)	Precision (%)	Recall (%)	Precision (%)	Recall (%)
Thorax	96	96	100	87.50	100	52
Abdomen	95.23	95.23	100	85.71	100	61.90
Wing	100	88.36	100	81.81	100	61.36
Leg	95.46	35.76	100	21.40	100	19.25

TABLE 4

mAP scores for masking		
IoU Ratio	Validation Set (%)	Testing Set (%)
0.30	62.50	53.49
0.50	60	52.38
0.70	51	4120

[0133] This disclosure also includes results from a small experiment with bumblebee images of FIG. **10**. The information herein subsequently verified how an AI architecture that was trained only with mosquito images, performs, when tested with images **1025** of bumblebees after the images **1025** have been subject to processing by a neural network as disclosed herein, utilizing bounding box images **1035** to segment the bee anatomy. Bumblebees (Genus: *Bombus*) are important pollinators, and detecting them in nature is vital. FIG. **10** presents example results for one representative image among three species of bumblebees, although the results are only representative of more than 100 bumblebee images we tested. The image source for bumblebees was Smithsonian National Museum of Natural History in Washington, D.C. Images can be found at Ref. 13. As shown in FIG. **10**, one non-limiting technique in accordance with this disclosure is robust in detecting and extracting wing pixels **1065**. While the thorax pixels **1045** are mostly extracted correctly, the ability to extract out the abdomen pixels **1055** and leg pixels **1075** is relatively poor. With these results, a confidence level is quite high for using the disclosed architecture to extract wings of many insects. With appropriate ground-truth data, only minimal tweaks to architecture will be needed to ensure robust extraction of all anatomical components for a wide range of insects.

[0134] This disclosure includes the following discussions on the significance of contributions in this paper.

(a) Faster classification of trapped mosquitoes. A cross the world, where mosquito-borne diseases are problematic, it is standard practice to lay traps, and then come back the next day to pick up specimens, freeze them and bring them to a facility, where expert taxonomists identify each specimen one-by-one under a microscope to classify the genus and species. This process takes hours each day, and is cognitively demanding. During rainy seasons and outbreaks, hundreds of mosquitoes get trapped, and it may take an entire day to process a batch from one trap alone. Based on technologies illustrated herein, mobile cameras can assist in taking high quality pictures of trapped mosquito specimens, and the extracted anatomies can be used for classification by experts by looking at a digital monitor rather than peering through a microscope. This will result in lower cognitive stress for taxonomists and also speed up surveillance efforts. In one non-limiting embodiment, Table 5 presents details on morphological markers that taxonomists look for to identify mosquitoes used in this study and discussed further at Ref. 14.

TABLE 5

Anatomical components and markers aiding mosquito classification Refs. 6-30.				
Species	Thorax	Abdomen	Wing	Leg
Aedes Aegypti	Dark with white lyre shaped pattern and patches of white scales	Dark with narrow white basal bands	Dark	Dark with white basal bands
Aedes infirmatus	Brown with patches of white scales	Dark with basal triangular patches of white scales	Dark	Dark
Aedes Taenio-rhynchus	Dark with patches of white scales	Dark with white basal bands	Dark	Dark with white basal bands
Anopheles Crucians	Gray-Black	Dark	Light and dark scales; dark costa; white wing tip; 3 dark spots on sixth vein	Dark with pale knee spots
Anopheles Quadri-maculatus	Gray-Black	Dark	Light and dark scales; 4 distinct darker spots	Dark with pale knee spots
Anopheles Stephensi	Broad bands of white scales		Four dark spots on costa extending to first vein	Speckling; narrow white band on fifth tarsomere
Culex Coronator	Dark with white scales on the apical and third segments	Sterna with-out dark triangles; mostly pale scaled		Distinct basal and apical bands on hind tarsomeres
Culex Nigri-palpus	Brown copper color; white scales	Dark with lateral white patches	Dark	Dark
Culex Salinarius	Copper; sometimes distinctly red; patches of white scales	Dark with golden basal bands; golden color on seventh segment	Dark	Dark

(b) AI and Cloud Support Education for Training Next-generation Taxonomists. The process of training taxonomists today across the world consists of very few training institutes, which store a few frozen samples of local and non-local mosquitoes. Trainees interested in these programs are not only professional taxonomists, but also hobbyists. The associated costs to store frozen mosquitoes are not trivial (especially in low economy countries), which severely limit entry into these programs, and also make these programs expensive to enroll. With technologies like those of this disclosure, digital support for trainees is enabled. Benefits include, but are not limited to, potential for remote education, reduced operational costs of institutes, reduced costs of enrollment, and opportunities to enroll more trainees. These benefits when enabled in practice will have positive impact to taxonomy, entomology, public health and more.

(c). Digital Preservation of Insect Anatomies under Extinction Threats. Recently, there are concerning reports that insects are disappearing at rapid rates. Digital preservation of their morphologies could itself aid preservation, as more and more citizen-scientists explore nature and share data to identify species under immediate threat. Preservation of insect images may also help educate future scientists across a diverse spectrum.

Image Processing Examples

[0135] Generation of Image Dataset and Preprocessing. In Summer 2019, research included partnering with Hillsborough county mosquito control district in Florida, USA to lay outdoor mosquito traps over multiple days. Each morning after laying traps, methods of this disclosure included collecting all captured mosquitoes, freezing them in a portable container and taking them to the county lab, where taxonomists identified them. This study utilized 23 specimens of *Aedes aegypti* and *Aedes infirmatus*, and 22 specimens of *Aedes taeniorhynchus, Anopheles* crucians, *Anopheles* quadrimaculatus, *Anopheles stephensi, Culex coronator, Culex nigripalpus* and *Culex salinarius*. It is notable to point out that specimens of eight species were trapped in the wild. The *An. stephensi* specimens alone were lab-raised whose ancestors were originally trapped in India.

[0136] Each specimen was then placed on a plain flat surface, and then imaged using a smartphone (among iPhone 8, 8 Plus, and Samsung Galaxy S8, S10) in normal indoor light conditions. To take images, the smartphone was attached to a movable platform 4 to 5 inches above the mosquito specimen, and three photos at different angles were taken. One directly above, and two at 45 degree angles to the specimen opposite from each other. As a result of these procedures, a total of 600 images were generated. Then, 500 of these images were preprocessed to generate the training dataset, and the remaining 100 images were separated out for validation. For preprocessing, the images were scaled down to 1024×1024 pixels for faster training (which did not lower accuracy). The images were augmented by adding Gaussian blur and randomly flipping them from left to right. These methods are standard in image processing, which better account for variances during run-time execution. After this procedure, the training dataset increased to 1500 images. Note here that all mosquitoes used in this study are vectors for disease and illness. Among these, *Aedes aegypti* is particularly dangerous, since it spreads Zika fever, dengue, chikungunya and yellow fever. This mosquito is also globally distributed now.

[0137] Deep Neural Network Framework based on Mask R-CNN. To address the goal of extracting anatomical components from a mosquito image, a straightforward approach is to try a mixture of Gaussian models to remove background from the image. See Refs 1, 15. But this will only remove the background, without being able to extract anatomical components in the foreground separately. There are other recent approaches in the realm also. One technique is U-Net, see Ref. 15, wherein semantic segmentation based on deep neural networks is proposed. However, this technique does not lend itself to instance segmentation (i.e., segmenting and labeling of pixels across multiple classes). Multi-task Network Cascade, see Ref. 16 (MNC), is an instance segmentation technique, but it is prone to information loss, and is not suitable for images as complex as mosquitoes with multiple anatomical components.

[0138] Fully Convolutional Instance-A ware Semantic Segmentation, see Ref. 17 (FCIS), is another instance segmentation technique, but it is prone to systematic errors on overlapping instances and creates spurious edges, which are not desirable. See DeepM ask at Ref. 19, developed by Facebook, extracts masks (i.e., pixels) and then uses Fast R-CNN (Ref. 20) technique to classify the pixels within the mask. This technique though is slow as it does not enable segmentation and classification in parallel. Furthermore, it uses selective search to find out regions of interest, which further adds to delays in training and inference.

[0139] In one aspect, this disclosure leverages M ask R-CNN, see Ref. 11, which is a neural network architecture for extracting masks (i.e. pixels) corresponding to objects of interest within an image which eliminates selective search, and also uses Regional Proposal Network (RPN) of Ref. 20 to learn correct regions of interest. This approach is best suited for quicker training and inference. A part from that, it uses superior alignment techniques for feature maps, which helps prevent information loss. The basic architecture is shown in FIG. **8**. Adapting it for the issues of this disclosure requires a series of steps presented below.

[0140] 1. Annotation for Ground-Truth. First, research herein includes manually annotating training and validation images using the VGG Image Annotator (VIA) tool as set forth in Ref. 21. To do so, this disclosure shows manually (and carefully) emplacing bounding polygons around each anatomical component in our training and validation images. The pixels within the polygons and associated labels (i.e., thorax, abdomen, wing or leg) serve as ground truth. One sample annotated image is shown in FIG. **4**.

[0141] 2. Generate Feature Maps using CNN. Then, the systems and methods disclosed herein use semantically rich features in the training image dataset to recognize the complex anatomical components of the mosquito as shown, for example, in the image **802** of FIG. **8**. To do so, one non-limiting neural network architecture **800** is a combination of the popular Res-Net101 architecture **806** with Feature Pyramid Networks (FPN) as shown in Ref. 12. Very briefly, ResNet-101 (Ref. 23) is a convolutional neural network (CNN) with residual connections, and was specifically designed to remove vanishing gradients at later layers during training. It is relatively simple with 345 layers. Addition of a feature pyramid network to ResNet was attempted in another study, where the motivation was to leverage the naturally pyramidal shape of CNNs, and to also create a subsequent feature pyramid network that combines low resolution semantically strong features with high resolution semantically weak features using a top-down pathway and lateral connections. Ref. 12.

[0142] This resulting architecture is well suited to learn from images at different scales from only minimal input image scales. Ensuring scale-invariant learning is important for this disclosure, since mosquito images can be generated at different scales during run-time, due to diversity in camera hardware and human induced variations. The output of the first convolutional neural network **806** is a set of respective feature maps **808** that isolate anatomical pixels for respective anatomical body parts of the subject insect in the image **802**.

[0143] 3. Emplacing anchors on anatomical components in the image. This step leverages the notion of Regional Proposal Network (RPN) **816**, as set forth in Ref. 20, and results from the previous two steps, to design a simpler CNN that will learn feature maps corresponding to ground-truth tested anatomical components in the training images. One end goal is to emplace anchors (which, in non-limiting examples, are bounding boxes **812**) that enclose the detected anatomical components of interest in the image.

[0144] 4. Classification and pixel-level extraction. Finally, this disclosure aligns the feature maps of the anchors (i.e., region of interest) learned from the above step into fixed sized feature maps. The alignment step **814** provides the fixed sized feature maps as inputs to three branches of the architecture **800** to:

[0145] (a) label the anchors with the anatomical component as illustrated in FIG. **3**;

[0146] (b) extract only the pixels within the anchors that represents an anatomical component as illustrated in FIGS. **5A-5D**; and

[0147] (c) tighten the anchors for improved accuracy as shown at **812** in FIG. **8**.

All three steps are done in parallel.

[0148] 5. Loss functions. For issues considered in this disclosure, one non-limiting scenario recalls that there are three specific sub-problems: labeling the anchors as thorax, abdomen, wings or leg; masking the corresponding anatomical pixels within each anchor; and a regressor to tighten anchors. Embodiments of this disclosure incorporate loss functions used for these three sub-problems. Loss functions are a critical component during training and validation of deep neural networks to improve learning accuracy and avoid overfitting.

[0149] 6. Labeling (or classification) loss. For classifying the anchors, non-limiting embodiments of this disclosure utilize the Categorical Cross Entropy loss function, and it worked well. For a single anchor j, the loss is given by an expression "where p is the model estimated probability for the ground truth class of the anchor."

[0150] 7. Masking loss. M asking is a challenging endeavor in image processing, considering the complexity in a neural network learning to detect only pixels corresponding to anatomical components in an anchor. Non-limiting experiments used in this research used the simple Binary Cross Entropy loss function. With this loss function, good accuracy was shown for pixels corresponding to thorax, wings and abdomen. But, many pixels corresponding to legs were mis-classified as background. This is because of class imbalance highlighted in FIG. **2** wherein we see significantly larger number of background pixels, compared to number of foreground pixels for anchors (colored blue) emplaced around legs. This imbalance leads to poor learning for legs, because the binary class entropy loss function is biased towards the (much more, and easier to classify) background pixels.

[0151] Another investigation utilized another more recently developed loss function called focal loss, discussed at Ref. 23, which lowers the effect of well classified samples on the loss, and rather places more emphasis on the harder samples. This loss function hence prevents more commonly occurring background pixels from overwhelming the not so commonly occurring foreground pixels during learning, hence overcoming class imbalance problems. The focal loss for a pixel i is represented as:

[0152] where p is the model estimated probability for the ground truth class, and gamma—γ—is a tunable parameter, optionally set as 2 in one example model. With these definitions, it is easy to see that when a pixel is mis-classified and p→0, then the modulating factor (1−p) γ tends to 1 and the loss (log(p)) is not affected. However, when a pixel is classified Correctly and when p→1, the loss is down-weighted. In this manner, priority during training is emphasized more on the hard negative classifications, hence yielding superior classification performance in the case of unbalanced datasets. Utilizing the focal loss gave superior classification results for all anatomical components.

[0153] 8. Regressor loss. To tighten the anchors and hence improve masking accuracy, the loss function used in one non-limiting example is based on the summation of Smooth L **1** functions computed across anchor, ground truth and predicted anchors.

[0154] In one example algorithm for Let (x, y) denote the top-left coordinate of a predicted anchor. Let xa and x* denote the same for anchors generated by the RPN, and the manually generated ground-truth. The notations are the same for the y coordinate, width w and height h of an anchor. The procedure may include defining several terms first, following which the loss function L reg used in one non-limiting example architecture is presented.

$$t_x^* = \frac{X*-Xa}{Wa},\ t_y^* = \frac{y*-Ya}{ha},\ t_w^* = \log\left(\frac{W*}{Wa}\right),\ t_h^* = \log\left(\frac{h*}{ha}\right), \tag{3}$$

$$t_x = \frac{X-Xa}{Wa},\ t_y = \frac{y-Ya}{ha},\ t_w = \log\left(\frac{W}{Wa}\right),\ t_h = \log\left(\frac{h*}{ha}\right),$$

$$\text{smooth}_{L1} = 0.5\,x^2,\ \text{if } |x| < 1 \text{ and}$$

$$|x| - 0.5,\ \text{otherwise}$$

$$Lreg(t_i, t_i^*) = \sum i \in x, y, w, h \ \text{smooth}(L1)(t_i^* - t_i).$$

Hyperparameters. For convenience, Table 6 lists values of critical hyperparameters in a finalized architecture.

TABLE 6

Values of Critical Hyperparameters in the Architecture

Hyperparameter	Value
Number of Layers	394
Learning rate	1e−3 for 1-100 epochs 5e−4 for 101-200 epochs 1e−5 for 201-400 epochs 1e−6 for 401-500 epochs
Optimizer	SGD
Momentum	0.9
Weight Decay	0.001
Number of epochs	500

[0155] Accordingly, this disclosure presents a system to design state of the art artificial intelligence (AI) techniques, namely techniques based on M ask Region-based Convolutional Neural Networks to extract anatomical components of mosquitoes from digital images and archive them permanently based on genus, species and other taxonomies. This disclosure indicates that the systems and methods of this disclosure currently have generated close to 30,000 digital images of mosquitoes (taken via smartphones) that are tagged based on genus and species type. Once anatomies of interest are extracted, this disclosure utilizes AI techniques to design a model that can recognize genus and species types of mosquitoes. Should the system be popular among citizens and experts, and users recruit entomologists, there are opportunities to globally scale up the effort to include many more mosquito types and improve our models over time.

[0156] One non-limiting proposition of this disclosure is the ability to bypass humans that peer through a microscope currently for classification, and instead use digital cameras and the proposed technique for automated classification of genus and species type. A secondary value proposition is the ability of the disclosed system, with large scale citizen and expert generated imagery with tagging, to start digitizing anatomies of mosquitoes across the globe. This database could prove invaluable for training, and global information sharing in the context of mosquito, and especially vector surveillance.

Example—Extraction of Anatomies

[0157] The disclosed approach for one example procedure is based on the notion of Mask R-CNN leveraging by which one segment pixels containing each anatomical component of interest by adding a branch for predicting an object mask (i.e., pixel-wise segmentation) in parallel with the existing branch for recognizing the bounding box (see FIG. **4** below).

[0158] In this approach, several critical steps (each of which is challenging) need to be executed. First, one must train the model using pretrained convolutional neural networks to generate proposals about the regions where there might be an object within the image. As discussed above, one non-limiting example of a convolutional neural network used herein is ResNet101, used as a backbone convolutional model. In one non-limiting example, initialization of the model was done using the pretrained MS COCO dataset weights. ResNet [33] has a very deep network and introduces a residual connection to get the input from the previous layer to the next layer. The residual connection helps in solving gradient vanishing problem. MS COCO dataset is large scale object detection dataset. It contains 1.5 million object instances, and 80 object categories. The next step is to design an object detector network that does three tasks: classifying the boxes with respective anatomies, tightening the boxes, and generating a mask (i.e., pixel-wise segmentation) of each anatomical component. In constructing the architecture of the object detector network, this disclosures uses per-pixel sigmoid as an example, along with binary cross-entropy loss function (to identify the k anatomical components) and rigorously train them.

[0159] Note that generating training datasets here are not easy, since a mask is required to be created for each anatomical component in our dataset. To start the training, users first annotated 571 mosquito images, using VGG Image annotator tool which is itself a very tedious job (annotated images are shown in FIG. **3**). Out of 571 images, 404 are the training images and 167 are validation images. Then, the methods iterate the model to optimize weights/hyper-parameters. This disclosure shows that the methods have optimized hyper-parameters like base feature extractor model, learning rate, momentum, optimizer, steps per epoch, validation steps and number of epochs, details are in below Table 7.

TABLE 7

Parameter	Value
Optimizer	Adam
Momentum	0.9
Learning rate	1e−4 for first 50 epochs, 1e−5 for next 50 and 1e−6 for next 100 epochs
Batch Size	2
Steps per epoch	202
Validation steps	84
Number of epochs	200

Result:

[0160] The metrics to measure the accuracy of mask R-CNN algorithm is mA P (Mean Average Precision). It was calculated by taking the mean of all the average precision across all classes over all IoU thresholds, and is 0.833. The metric IoU measures the intersection of ratio of pixels that belong to the ground-truth of the object in the bounding box and the union of the predicted and the ground truth ratio of pixels in the box. In our design, the IoU threshold was set as 0.75.

[0161] FIGS. **6**A, **6**B, **6**C and **7**A, **7**B, **7**C show an example design of a smartphone (e.g., imaging device) **600**, **700** application in Android and iOS that enables a user to either take an image **620**, **720** of a mosquito, or choose one from the local storage of the phone. The app will classify the mosquito and also extract anatomies of interest-thorax, wing and abdomen and legs for display in respective anatomical images **620**, **720**, **630**, **730**, **640**, **740**. All images generated are archived in the cloud.

Example—Classification of Genus and Species Type Based on Extracted Anatomies

[0162] This disclosure explains a design for a neural network based architecture to identify genus and species type of mosquitoes from the whole image body. The network architectures are presented below. The results have achieved close to 80% accuracy in classifying various (currently nine mosquito species), and close to 99% accuracy in identifying a very deadly mosquito—*Aedes aegypti*—that spreads Zika fever, dengue, chikungunya, and yellow fever—and one is prevalent all over the world.

TABLE 8

Species Architecture		
Layer	Size In	Size Out
block17_10_conv (Layer 433 in IRV2)	(None, 17, 17, 384)	(None, 17, 17, 1088)
GlobalAveragePooling2D	(None, 17, 17, 1088)	(1, 1088)
dense_1	(1, 1088)	512
dense_2	512	256
dense_3	256	128
dense_4	128	256
concat_1	(dense_1, dense_7, dense_3, dense_4)	1152
softmax	1152	9

[0163] In another example, users are extending the above architectures to classify based on anatomies also. This is part of on-going work, but the network architectures for each anatomy—thorax, abdomen, wings and leg will be very similar to the above architectures.

[0164] Embodiments of this disclosure include non-limiting combinations of the above described work. Accordingly, in one example, a system **800** for identifying a genus and species of an insect includes an imaging device **600**, **700** configured to generate images of the insect. The imaging device may be a smart phone or other mobile computer devices with camera functions. A computer processor **1202** is connected to memory **1204** storing computer implemented commands in software, and the memory receives the images (i.e., the computer processor, the memory, and the imaging device may be in data communication over a network or a local connection). The software implements a computerized method with respective images, beginning with applying a first convolutional neural network **806** to the respective images **802** to develop feature maps **808** directed to anatomical pixels **510**, **520**, **530**, **540**, e.g., in the respective images that correspond to a body part of the insect. By weighting filters in the first convolutional neural network, the system is set up to identify, within the respective images, the anatomical pixels as foreground pixels and remaining pixels as background pixels to be removed during segmenting operations.

[0165] Next, the system utilizes the processor or other computers to apply anchors **812**, **817** to the feature maps **808**, wherein the anchors identify portions of respective layers of image data in the feature maps that contain respective anatomical pixels for respective body parts. In other words, the feature maps may be multidimensional layers of image data, and the system can operate on individual layers of image data or multiple sets of layers of image data that have resulted from the first convolutional neural network **806**. The anchors may take the form of any polygon that bounds a desired set of anatomical pixels within images, feature maps, or layers of image data. In one non-limiting example, the anchors are bounding boxes that are generated by the computer processor and shown on a graphical display unit as being within or superimposed on the images. The software further includes an alignment function **814** to align layers having anchors **812**, **817** thereon with the original feature maps.

[0166] The system uses these anchors in generating a mask **818**A that segments the respective anatomical pixels from the respective layers of image data. Generating the mask may include applying a second convolutional neural network **818**B to the respective layers, wherein the second convolutional neural network segments the anatomical pixels according to a corresponding respective body part.

[0167] The mask allows for extracting fully connected layers **822** from the respective layers that have had the first convolutional neural network **806** applied thereto, and the system is further configured for applying the fully connected layers to a regressor network **826** and a classification network **824**, wherein generating the mask for segmenting, applying the fully connected layers to a regressor network, and applying the fully connected layers to a classification network are parallel operations conducted by the software. In certain non-limiting embodiments generating the mask includes applying a second convolutional neural network to the respective layers, wherein the second convolutional neural network segments the anatomical pixels according to a corresponding respective body part.

[0168] In some non-limiting embodiments, the parallel operations occur simultaneously.

[0169] The regressor network **826** is a software program implemented by a computer to calculate error values regarding iterative positions for the anchors in the respective layers. The system uses the error values in a feedback loop **850** to tighten the anchors **812**, **817** around anatomical pixels corresponding to a respective body part. The regressor network and associated computer-implemented software calculates error values regarding iterative positions for the anchors in the respective layers and wherein the error values are derived from a binary cross entropy loss function or a focal loss function.

[0170] The anchors **812**, **817** may be bounding boxes, or any other shape, originating from a region proposal network **816** receiving the feature maps **808** as respective layers of image data, and the feedback loop **850** transmits error values from the regressor network **826** to the region proposal network **816** to tighten the boxes onto appropriate pixels corresponding to the respective body parts. The region proposal network is an image processing software implementation that utilizes data from the feature maps to predict probable portions of images and layers of images that contain anatomical pixels corresponding to an insect body part.

[0171] The classification network **824** is a software tool implemented by a computer for generating classification output images **828**, **830** and in some embodiments, these classification output images include updated versions of original images with bounding polygons **812**, **817** therein, labels **1125**, **1135**, **1145**, **1155** for anatomical component names thereon, and even color coding as shown in Table 5 that may aid in genus and species identification. The example embodiments herein shows the system identifying insects such as a mosquito and anatomical component names including wings, legs, thorax, and abdomen corresponding to the respective body parts. In some non-limiting embodiments, the classification network utilizes a per-pixel sigmoid network. In non-limiting uses, the system populates a database storing tested outputs of the classification network, wherein the outputs include image versions with labels of anatomical component names thereon, and wherein the database stores respective genus and species information with corresponding data about respective genera and species.

[0172] The system embodiment may be implemented with at least one computer that performs a computerized method of extracting information about anatomical components of a living creature from an image. The images may include digital images of insects or other animals or even inanimate objects, wherein the digital images include views of respective insects, animals, or inanimate objects from directly above the specimen and from side angles relative to a background holding the respective specimens. By training a mask-region based convolutional neural network with a set of training images, segmented with computerized algorithms, the method begins by identifying ground truth anatomical components to a set degree of accuracy. The training for the convolutional neural networks used in this disclosure generally includes classifying respective anatomical components in the training images and comparing the training images to the ground truth images. By tightening bounding boxes surrounding the anatomical components in the digital images, the method learns how to maximize efficiency and accuracy in ultimately generating a mask for use in extracting information of a second set of images, such as feature maps that have been previously created. For forming the ground truth images, the computerized algorithm may utilize an image annotator tool configured for manual operation. The training iteratively updates hyperparameters that target anatomical pixels in a training data set. This method has a proven track record of tracking, identifying, and archiving genera and species identifying data for a plurality of species of a plurality of genera of insects.

[0173] In example implementations, at least some portions of the activities may be implemented in software provisioned on a networking device. In some embodiments, one or more of these features may be implemented in computer processor **1200**, provided external to these elements, or consolidated in any appropriate manner to achieve the intended functionality. The various network elements may include software (or reciprocating software) that can coordinate image development across domains such as time, amplitude, depths, and various classification measures that detect movement across frames of image data and further detect particular objects in the field of view in order to achieve the operations as outlined herein. In still other embodiments, these elements may include any suitable algorithms, hardware, software, components, modules, interfaces, or objects that facilitate the operations thereof.

[0174] Furthermore, computer systems described and shown herein (and/or their associated structures) may also include suitable interfaces for receiving, transmitting, and/or otherwise communicating data or information in a network environment. Additionally, some of the processors **1202** and memory elements **1204** associated with the various nodes may be removed, or otherwise consolidated such that single processor and a single memory element are responsible for certain activities. In a general sense, the arrangements depicted in the Figures may be more logical in their representations, whereas a physical architecture may include various permutations, combinations, and/or hybrids of these elements. It is imperative to note that countless possible design configurations can be used to achieve the operational objectives outlined here. Accordingly, the associated infrastructure has a myriad of substitute arrangements, design choices, device possibilities, hardware configurations, software implementations, equipment options, etc.

[0175] In some of example embodiments, one or more memory elements (e.g., memory can store data used for the operations described herein. This includes the memory being able to store instructions (e.g., software, logic, code, etc.) in non-transitory media, such that the instructions are executed to carry out the activities described in this Specification. A processor can execute any type of computer readable instructions associated with the data to achieve the operations detailed herein in this Specification. In one example, processors (e.g., processor) could transform an element or an article (e.g., data) from one state or thing to another state or thing. In another example, the activities outlined herein may be implemented with fixed logic or programmable logic (e.g., software/computer instructions executed by a processor) and the elements identified herein could be some type of a programmable processor, programmable digital logic (e.g., a field programmable gate array (FPGA), an erasable programmable read only memory (EPROM), an electrically erasable programmable read only memory (EEPROM)), an ASIC that includes digital logic, software, code, electronic instructions, flash memory, optical disks, CD-ROM s, DVD ROM s, magnetic or optical cards, other types of machine-readable mediums suitable for storing electronic instructions, or any suitable combination thereof.

[0176] These devices may further keep information in any suitable type of non-transitory storage medium **1212** (e.g., random access memory (RAM), read only memory (ROM), field programmable gate array (FPGA), erasable programmable read only memory (EPROM), electrically erasable programmable ROM (EEPROM), etc.), software, hardware, or in any other suitable component, device, element, or object where appropriate and based on particular needs. Any of the memory items discussed herein should be construed as being encompassed within the broad term 'memory element.' Similarly, any of the potential processing elements, modules, and machines described in this Specification should be construed as being encompassed within the broad term 'processor.' See FIG. **12** for a schematic example showing a computing environment for input devices **1208**A, such as imaging devices described above, and output devices **1208**B such as smartphones of FIGS. **6** and **7**. This computer environment is amenable to various network and cloud connections as shown at **1206**.

[0177] It must also be noted that, as used in the specification and the appended claims, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise. By "comprising" or "containing" or "including" is meant that at least the named compound, element, particle, or method step is present in the composition or article or method, but does not exclude the presence of other compounds, materials, particles, method steps, even if the other such compounds, material, particles, method steps have the same function as what is named.

[0178] Ranges may be expressed herein as from "about" or "approximately" one particular value to "about" or "approximately" another particular value. When such a range is expressed, exemplary embodiments include from the one particular value to the other particular value. As used herein, "about" or "approximately" generally can mean within 20 percent, preferably within 10 percent, and more preferably within 5 percent of a given value or range, and can also include the exact value or range. Numerical quantities given herein can be approximate, meaning the term "about" or "approximately" can be inferred if not expressly stated.

[0179] In describing example embodiments, terminology will be resorted to for the sake of clarity. It is intended that each term contemplates its broadest meaning as understood by those skilled in the art and includes all technical equivalents that operate in a similar manner to accomplish a similar purpose. It is also to be understood that the mention of one or more steps of a method does not preclude the presence of additional method steps or intervening method steps between those steps expressly identified. Steps of a method may be performed in a different order than those described herein without departing from the scope of the present disclosure. Similarly, it is also to be understood that the mention of one or more components in a device or system does not preclude the presence of additional components or intervening components between those components expressly identified.

REFERENCES

[0180] 1. Minakshi, M., Bharti, P. & Chellappan, S. Leveraging smart-phone cameras and image processing techniques to classify mosquito species. In *Proceedings of the* 15*th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services,* 77-86 (ACM, 2018).

[0181] 2. De Los Reyes, A. M. M., Reyes, A. C. A., Torres, J. L., Padilla, D. A. & Villaverde, J. Detection of *Aedes aegypti* mosquito by digital image processing techniques and support vector machine. In 2016 *IEEE Region* 10 *Conference* (*TENCON*), 2342-2345 (IEEE, 2016).

[0182] 3. Fuchida, M., Pathmakumar, T., Mohan, R. E., Tan, N. & Nakamura, A. Vision-based perception and classification of mosquitoes using support vector machine. *Appl. Sci.* 7, 51 (2017).

[0183] 4. Favret, C. & Sieracki, J. M. Machine vision automated species identification scaled towards production levels. *Syst. Entomol.* 41, 133-143 (2016).

[0184] 5. Minakshi, M. et al. Automating the surveillance of mosquito vectors from trapped specimens using computer vision techniques. In *ACM COMPASS* (ACM, 2020).

[0185] 6. Park, J., Kim, D. I., Choi, B., Kang, W. & Kwon, H. W. Classification and morphological analysis of vector mosquitoes using deep convolutional neural networks. *Sci. Rep.* 10, 1-12 (2020).

[0186] 7. Chen, Y., Why, A., Batista, G., Mafra-Neto, A. & Keogh, E. Flying insect detection and classification with inexpensive sensors. *J. Vis. Exp.* (*JoVE*) e52111 (2014).

[0187] 8. Mukundarajan, H., Hol, F. J., Castillo, E. A., Newby, C. & Prakash, M. Using mobile phones as acoustic sensors for the surveillance of spatio-temporal mosquito ecology (2016).

[0188] 9. Vasconcelos, D., Nunes, N., Ribeiro, M., Prandi, C. & Rogers, A. Locomobis: a low-cost acoustic-based sensing system to monitor and classify mosquitoes. In 2019 16th *IEEE Annual Consumer Communications & Networking Conference* (*CCNC*), 1-6 (IEEE, 2019).

[0189] 10. Ravi, P., Syam, U. & Kapre, N. Preventive detection of mosquito populations using embedded machine learning on low power iot platforms. In *Proceedings of the 7th Annual Symposium on Computing for Development,* 1-10 (2016).

[0190] 11. He, K., Gkioxari, G., Dollár, P. & Girshick, R. M ask r-cnn. Proceedings of the IEEE International Conference on Computer Vision 2961-2969 (2017).

[0191] 12. Lin, T.-Y. et al. Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2117-2125 (2017).

[0192] 13. Smithsonian homepage. https://www.si.edu/. The importance of learning identification of larvae and adult mosquitoes. https://juniperpublishers.com/jojnhc/pdf/JOJNHC.MS.ID.555636.pdf.

[0193] 14. Stauffer, C. & Grimson, W. E. L. Adaptive background mixture models for real-time tracking. In Computer Vision and Pattern Recognition, 1999. *IEEE Computer Society Conference on*., vol. 2, 246-252 (IEEE, 1999).

[0194] 15. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention,* 234-241 (Springer, 2015).

[0195] 16. Dai, J., He, K. & Sun, J. Instance-aware semantic segmentation via multi-task network cascades. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 3150-3158 (2016).

[0196] 17. Li, Y., Qi, H., Dai, J., Ji, X. & Wei, Y. Fully convolutional instance-aware semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 2359-2367 (2017).

[0197] 18. Pinheiro, P. O., Collobert, R. & Dollar, P. Learning to segment object candidates. In *Advances in Neural Information Processing Systems* 1990-1998 (2015).

[0198] 19. Girshick, R. Fast R-CNN. In Proceedings of the IEEE International Conference on Computer Vision 1440-1448 (2015).

[0199] 20. Ren, S., He, K., Girshick, R. & Sun, J. Faster R-CNN: towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems* 91-99 (2015).

[0200] 21. Dutta, A. & Zisserman, A. The VGG image annotator (via). arX iv preprint arXiv: 1904.10699 (2019).

[0201] 22. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* 770-778 (2016).

[0202] 23. Lin, T.-Y., Goyal, P., Girshick, R., He, K. & Dollár, P. Focal loss for dense object detection. In *Proceedings of the IEEE International Conference on Computer Vision,* 2980-2988 (2017).

[0203] 24. Mask RCNN code. https://github.com/matterport/Mask_RCNN.

[0204] 25. IFAS. Florida medical entomology laboratory. https://fmel.ifas.ufl.edu/mosquito-guide/species-identification-table/species-identification-table-adult/.

[0205] 26. Glyshaw, P. & Wason, E. *Anopheles quadrimaculatus*. https://animaldiversity.org/accounts/*Anopheles_quadrimaculatus*/edu/mosquito-guide/species-identification-table/species-identification-table-adult/.

[0206] 27. Dharmasiri, A. G. et al. First record of *Anopheles stephensi* in Sri Lanka: a potential challenge for prevention of malaria reintroduction. *Malaria J.* 16, 326 (2017).

[0207] 28. IFAS. Florida medical entomology laboratory. https://fmel.ifas.ufl.edu/publication/buzz-words/buzz-words-archive/is-it-*Culex-tarsalis*-or-*Culex-coronator/*

[0208] 29. Floore, T. A., Harrison, B. A. & Eldridge, B. F. The *anopheles* (*anopheles*) *crucians* subgroup in the united states (diptera: Culicidae) (Tech. Rep, Walter Reed Army Inst Of Research Washington Dc Department Of Entomology, 1976)

[0209] 30. Smithsonian Institution. https://collections.si.edu/search/results.htm?fq=tax_kingdom%3A%22animalia%22& fq=online_media_type%3A%22Images%22& fq=datasource%3A%22NMNH++Entomology+Dept.%22&q=NMNH-USNMENT01001576&gfq=CSILP_6.

[0210] 31. Smithsonian https://collections.si.edu/search/results.htm?fq=tax_kingdom%3A%22Animedia_type%3A%22Images%22&fq=data_source%3A%22NMNH+Entomology+Dept.%22&q=NMNH-ENT01006317&gfq=CSILP_6.

[0211] 32. He, K aiming, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. "M ask r-cnn." In Proceedings of the IEEE international conference on computer vision, pp. 2961-2969. 2017.

[0212] 33. He, K aiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. 2016.

[0213] 34. Lin, Tsung-Y i, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. "Microsoft coco: Common objects in context." In European conference on computer vision, pp. 740-755. Springer, Cham, 2014.

[0214] 35. Stewart, Matthew Ph.D. Simple Introduction to Convolutional Neural Networks in *Towards Data Science* located at https://towardsdatascience.com/simple-introduction-to-convolutional-neural-networks-cdf8d3077bac.

What is claimed is:

1. An apparatus comprising:

an entry aperture configured to lead one or more insects to an imaging chamber within the apparatus;

a door on a side surface of the apparatus, wherein a rear surface of the door is positioned within the imaging chamber in a closed position, and wherein the rear surface of the door comprises a removable/replaceable adhesive material or surface configured to trap one or more insects thereon; and

at least one imaging sensor positioned opposite the rear surface of the door configured to capture images of the one or more insects adhering to the adhesive material or surface.

2. The apparatus of claim **1**, wherein the at least one imaging sensor comprises at least two autofocus cameras.

3. The apparatus of claim **1**, wherein the apparatus is in electronic communication with an external camera configured to capture image data in a vicinity of or external to the apparatus.

4. The apparatus of claim **1**, further comprising:

at least one light source positioned adjacent to the at least one imaging sensor within the imaging chamber.

5. The apparatus of claim **1**, wherein the adhesive material or surface is positioned at approximately 35 degrees relative to an inner side surface of the apparatus.

6. The apparatus of claim **1**, further comprising:

a fan chamber positioned below the imaging chamber, wherein the fan chamber comprises at least one fan configured to generate negative air pressure by pulling air through the entry aperture and imaging chamber.

7. The apparatus of claim **6**, further comprising:

a mesh grill separating the imaging chamber from the fan chamber.

8. The apparatus of claim **6**, further comprising at least one of:

a pyramidal airflow diverter positioned at a bottom of the fan chamber configured to redirect airflow horizontally outward through open sides, and

an electronics chamber positioned below the fan chamber.

9. The apparatus of claim **8**, wherein the electronics chamber comprises an image capture and processing module configured to transmit the images to a cloud-based platform for species identification and classification using artificial intelligence or machine learning techniques.

10. The apparatus of claim **1**, wherein the door comprises:

a frame configured to support the adhesive material or surface within the imaging chamber.

11. The apparatus of claim **10**, wherein the frame includes:

a top roller and a bottom roller mounted within the frame, and

a motor operatively connected to and configured to incrementally rotate the bottom roller to expose fresh adhesive material or surface.

12. The apparatus of claim **11**, further comprising one or more sensors operatively coupled to the frame or door, wherein the one or more sensors are configured to:

detect an amount of remaining adhesive material or surface, and

generate an alert when replacement of the adhesive material or surface is required.

13. An apparatus comprising:

an entry aperture configured to lead one or more insects into an entry chamber within the apparatus;

a door on a side surface of the apparatus, wherein a rear surface of the door is positioned within the entry chamber in a closed position; and

at least one camera assembly positioned opposite the rear surface of the door configured to capture video data of the one or more insects within the entry chamber.

14. The apparatus of claim **13**, further comprising:

a reversible fan configured to generate a controlled airflow through the entry chamber, wherein the reversible fan is operable in a first direction to create a negative pressure to draw insects into the entry chamber and house them therein for imaging, and wherein the reversible fan is operable in a second direction to expel insects from the entry chamber into a secondary holding chamber.

15. The apparatus of claim **14**, wherein the secondary holding chamber is physically contiguous with the entry chamber and is sealed off from the external environment except through airflow routes controlled by the reversible fan.

16. The apparatus of claim **15**, further comprising:

at least one transparent sheet positioned within the entry chamber configured to maintain the one or more insects at a specified focal range relative to the camera assembly.

17. The apparatus of claim **15**, further comprising:

at least one lighting element adjacent to the camera assembly configured to illuminate an interior of the entry chamber.

18. The apparatus of claim **17**, wherein the at least one lighting element comprises one or more circular light emitting diode (LED) strips surrounding or adjacent to a camera assembly mount.

19. The apparatus of claim **14**, further comprising at least one of:

an electronics chamber including at least one of a microcontroller, printed circuit boards, and a power distribution system, wherein the electronics chamber is positioned above or adjacent to the entry chamber,

a mesh grill separating the entry chamber from the reversible fan, and

a lid to secure or cover at least a portion of the electronics chamber.

20. A system comprising:

at least one apparatus, the at least one apparatus comprising:

an entry aperture configured to lead one or more insects to an imaging chamber within the apparatus,

a door on a side surface of the apparatus, wherein a rear surface of the door is positioned within the imaging chamber in a closed position, and wherein the rear surface of the door comprises a removable/replaceable adhesive material or surface configured to trap one or more insects thereon, and

at least one imaging sensor positioned opposite the rear surface of the door configured to capture images or video of the one or more insects adhering to the adhesive material or surface, and

an electronics including at least a processor or microcontroller; and at least one external camera operatively in electronic communication with the at least a processor or microcontroller, wherein the at least one external camera is configured to capture image or video data in a vicinity of or external to the apparatus.

* * * * *