



Bildanalyse Code verstehen

Status	In Progress
Project	Praxis: Sichtung, Verstehen & Integration des bisherigen Programmcodes
Tags	

Verwendende Modelle

Zweck	Modelle	Unterzweck	Output - Werte
Objekterkennung	Detectron2		Superkategorie
Menschliche Attribute:	Deep Expectation Modell (DEX)	Altersklassifikation	Altergruppenintervalle: [0, 8] if 0 < age <= 8 else \ [9, 17] if 9 <= age < 18 else \ [18, 30] if 18 <= age < 31 else \ [31, 60] if 31 <= age < 61 else \ [61, 100] if 61 <= age < 100 else \ [100, 200]
	Deep Expectation Modell (DEX)	Geschlecht	[1] = man / [0] = woman
	Deepface	Ethnie	„black“, „asian“, „white“, „latino hispanic“, „middle eastern“, „indian“
	Facial Expression Recognition (FER)	Stimmung	„anger“, „disgust“, „fear“, „happy“, „neutral“, „sad“, „surprise“

Bildkomposition: Bildanteil und Quadranten berechnen

1. Bounding Box von einem Objekt finden
2. Unterteilung der Frames in neun Quadranten
3. Mittelpunkt der Bounding Box finden und in Quadranten finden

Aufgetretene Problem:

Ursprünglich wurden für diese Modelle die Gesamtframes als Eingabe verwendet. Bei einer manuellen Überprüfung der Frames wurde jedoch festgestellt, dass die Modelle gelegentlich menschliche Attribute fälschlicherweise für andere Objekte vorhersagten, anstatt sie nur für Menschen zu erkennen. Zusätzlich kam es vor, dass bei zwei Personen im selben Frame, selbst wenn eine Person im Hintergrund stand, das Modell häufig

Lösung: zugeschnittene Bild

Um diese Fehler zu vermeiden und die Präzision der Modelle zu steigern, wurde entschieden, ausschließlich ein zugeschnittenes Bild der erkannten Person aus dem Gesamtframe zu verwenden, um die Vorhersagen zu optimieren. Diese Vorgehensweise führte nicht nur zu einer Verringerung der Berechnungszeit (Effizienz), sondern steigerte auch die Vorhersagegenauigkeit (Effektivität). Wenn in

die im Vordergrund stehende Person vorhersagte, manchmal sogar zweimal.

einem beschnittenen Bild mehr als ein Gesicht erschien, wurde mit der größeren Bounding Box der beiden oder mehreren Objekte (in diesem Fall Personen) gearbeitet.

Allegemeine Vorgehen von Bildanalyse (Mainskript)

1. Bibliotheken importieren

2. Vortrainierte Modelle installieren

3. Model functions

3. Frame extraction (3 frames pro Sekunde)

3.2 detectron2

3.2.1 Finde die Quadranten

3.2.2 function for

1. finding the quadrant_number
2. Cropp the image based on the bounding box
3. Function for visualising outputs
4. def detectron2_analysis(im) → This will return a list of predictions with object properties and additional attributes (like age, gender, etc.) if the object is a human.

1. DEX: Deep EXpectation of apparent age from a single image (Age and Gender)

- `detect_faces`
 - **Purpose:** Detects faces in an image using a Haar cascade classifier.
 - **Output:** List of face bounding boxes, with the largest face first.
- `dex_visualisation`
 - **Process:** Draws a rectangle around a detected face and adds text for age and gender.
 - **Output:** Saves the annotated image with a unique filename.
- `dex_analysis`
 - **Process:** Detects the largest face, or all faces if modified, crops each face, performs predictions, and optionally visualizes.
 - **Output:** Returns age, age group, and gender for each face as a list of dictionaries.

2. FER: Facial expression recognition

- a. `fer_visualisation` : Saves the annotated image with a unique filename indicating emotion analysis for each face
- b. `fer_analysis` : Returns the dominant emotion for the largest detected face, or `"_"` if no face is found

3. DeepFace

- a. `deepface_visualiser` : Saves an image with bounding box and ethnicity annotation.

- b. `deepface_analysis`: Returns the dominant ethnicity for the largest face or `" "` if no face is found.

3.3 Summary

1. `generate_summary`: The function updates the Excel file with summarized data for both object classes and human attributes, organized in separate sheets.
 - **Process:**
 1. **Read Data:** Loads the "Predictions" sheet from the specified Excel file.
 2. **Filter by Size:** Filters objects based on a threshold (`schwellenwert_proportion_detectron_2`) for minimum object size.
 3. **Class Summary (for each class) :**
 - Aggregates counts, average confidence, average object proportion, and frame numbers for each class.
 - Computes a `frame_ratio` for each class (frames containing each object type relative to the total frame count).
 - Filters out classes with fewer frames than `schwellenwert_frame_nr_detectron_2`.
 4. **Human Attribute Summary:**
 - For each attribute (e.g., gender, ethnicity), filters data and applies thresholds to focus on relevant entries.
 - Computes counts, average object proportions, and frame ratios for each attribute.
 - Filters based on minimum frames, `schwellenwert_frame_nr_human_attributes`, to retain significant attributes.
 - `generate_summary_gender_ethnicity` Function: Updates the Excel file with a new sheet ("Summary_Gender_Ethnicity") containing detailed tables for gender and ethnicity distributions across frames. Creating detailed summaries based on object size, frame counts, and attribute distributions.
 - **Process:**
 1. **Data Loading and Filtering:** Reads the "Predictions" sheet and filters entries by size and frame count thresholds (`schwellenwert_proportion_human_attributes` and `schwellenwert_frame_nr_human_attributes`).
 2. **Gender Analysis:**
 - Groups data by unique frames, calculates gender counts, object proportions, quadrant positions, and frame ratios.
 - Compiles these metrics into `Gender_sheet`, organizing data by cases (e.g., number of men and women in each frame).
 3. **Ethnicity Analysis:**
 - Similar to gender analysis, but with detailed summaries for each detected ethnicity (e.g., Asian, Black).

- Calculates average proportions, quadrants, and frame ratios for each ethnicity and stores in `Ethnicity_sheet`.
4. **Saving Data:** Writes the original data and summaries (for both gender and ethnicity) to new sheets in the Excel file.

4. Main

- Extracting Frames
- Creating the output files
- Creating the summaries

Erstellten Summaries (Excel Blättern) :

- Excel-Blättern mit den unterschiedlichen Ausgaben erstellt:
 - Excel-Blatt „Predictions“:** detected object in one frame
 - `gender_prediction` : Predicted gender.
 - `ethnicity_prediction` : Predicted ethnicity.
 - `object_propotion` : Object's proportion in the frame.
 - `quadrant_number` : Quadrant location of the detected object.
 - Excel-Blatt „Summary_Objects“:** Aggregates detected objects across frames
 - Excel-Blatt „Summary_Gender_Ethnicity Sheet“:**
 - Contains two sub-sheets for gender and ethnicity, summarizing cases and averages.

Gender Sub-Sheet:

- Columns:
 - `case` : Unique identifier combining gender counts.
 - `count` : Occurrence count of the case.
 - `frame_ratio` : Frame ratio for the case.
 - `avg_object_propotion_women` and `avg_object_propotion_men` : Average proportions for women and men.
 - `avg_quadrant_number_women` and `avg_quadrant_number_men` : Average quadrants for women and men.
 - `quadrant_numbers_women` and `quadrant_numbers_men` : List of quadrants where women and men appear.

Ethnicity Sub-Sheet:

- Columns (similar structure for each ethnicity):
 - `case` : Unique identifier for ethnicity cases.

- `count` : Occurrence count.
- `frame_ratio` : Ratio of frames with the ethnicity.
- For each ethnicity (e.g., `asian` , `black` , etc.):
 - `avg_object_propotion_*` : Average object proportion.
 - `avg_quadrant_number_*` : Average quadrant position.
 - `quadrant_numbers_*` : List of quadrants for the ethnicity.

! Es wurde festgestellt (Stand im Projektbericht), dass Objekte und menschliche Attribute fehlerhaft klassifiziert wurden, wenn sie weniger als eine bestimmte Anzahl Frames oder Sekunden in einem Werbespot erscheinen. Daher wurden Schwellenwerte festgelegt: Objekte werden nur bei mindestens 6 Frames und menschliche Attribute bei mindestens 9 Frames berücksichtigt, wobei menschliche Attribute eine Bounding Box von mindestens 10% des Frames einnehmen müssen, während bei Objekten der Schwellenwert unverändert bleibt, da kleine Objekte oft korrekt erkannt werden.

Videoanalyse Ordner Aufbau

1. **models: Erforderliche Modelle für die Bildanalyse**

a. Coco supercategorie

- i. super Kategorie für den Datensatz namens Coco (Super Kategorie: Überkategorie für die Klassen)
- ii. Detectron2 (für die Objekterkennung ausgewählte Modell) wurde mit dem COCO-Dataset trainiert und kann daher die im Datensatz vorhandenen Objekte erkennen. In der Analyse repräsentieren die verschiedenen Klassen die erkannten Objekte

b. Dex

- i. Deep Expectation Modell (DEX) Modell

c. emotion_model

- i. Facial Expression Recognition (FER) Modell

2. **main_Script: Umfassender Python-Code für die Bildanalyse.**

3. **Heatmaps_Bildkomposition: Python-Code zur Erstellung von Heatmaps.**

- a. Code für Heatmaps, die die Geschlechterrepräsentativität in einem Werbung darstellen → durchschnittlichen Bildanteil und Quadranten

4. **Manuelle Evaluation: Evaluation der Modelle. (nicht relevant für unsere Projekt, da die Modelle bereits implementiert sind)**

- a. Excel Tabelle für Vergleich verschiedene Modelle (Accuracy etc.) um zu bestimmen welche Modelle implementiert werden sollte