Name: Nguyen Minh Trang

ID: 411021365

Image Processing Final Project

Report Paper

Introduction

In this project, a face authentication system is developed to identify individuals from their facial features. The system, built upon essential image processing techniques and custom algorithms, focuses on enhancing the accuracy of facial recognition despite variations in facial feature positions and image orientation.

Dataset

The dataset utilized in this project was divided into two primary categories: ‘Face_DB’ and ‘Test_DB’. The ‘Face_DB’ folder contains frontal face images of the 10 authorized people. This folder contains two sub-folders: “./Images”, and “./Landmark_data”. The first sub-folder contains images in JPG format. The images are named as “NAME_000.jpg”, where NAME corresponds to the person’s name. The second folder contains CSV files with landmark location data for each image in the “./Images” sub-folder. The CSV files are named as “NAME_000.csv”, where NAME corresponds to the person’s name. The facial landmark location data was annotated following the 300-W 68-landmark indexing scheme shown in the following figure.

The “Test_DB” folder contains two sub-folders: “./Images” and “./Landmark_data”. The files are named in a similar way as for “Face_DB”. In the “./Images” subfolder the images are named as “NAME_00X.jpg”, where NAME corresponds to the person’s name and X is the image number. In the “./Landmark_data” folder the CSV files are named as “NAME_00X.csv”, where NAME is the person’s name and X is the image number corresponding to image, in the “./Images”

subfolder, with the same X value. In total, “Test_DB” folder comprises 6 images per person (90 in total), 6 CSV files per person (90 in total).

Methodology

As the sole developer of this face authentication system, I began by addressing the challenge of comparing facial features in a dataset with significant variations in face angles and feature locations. To effectively measure the similarity between facial features, I opted for the Euclidean distance method. This choice was driven by the method's straightforward nature and its efficacy in representing the physical dissimilarity between multi-dimensional data points, such as facial landmarks.

Upon closely examining the dataset, two key issues became evident: the facial images in 'Face_DB' were all frontal, while those in 'Test_DB' were often taken from various angles, and the feature locations across the dataset were inconsistent.

The inconsistency of feature locations and face orientation across the dataset

Recognizing that applying a distance calculation directly to this raw data might yield inaccurate results, I decided that aligning the faces was a critical first step. For this, I chose the eyes as the reference feature, ensuring they were in the same position in all images. This alignment was vital to mitigate the impact of different face angles and positions, providing a more standardized basis for feature comparison.

Some images after alignment

However, post-alignment, a significant asymmetry in features was noticeable, particularly in images taken from an angle. To address this, I developed a method to make these features symmetric. For each landmark point, I identified its corresponding point on the opposite side of the face, essentially mirroring it across a vertical axis that runs through the center of the face, then calculated the midpoint between each original landmark point and its mirrored counterpart. This midpoint represents a symmetrical position relative to the center of the face. By adjusting the landmarks on each face, I was able to create a more balanced and consistent representation of facial features across all images. It's important to note that this symmetrization process was sensitive to the inherent variations in facial features. The goal was not to create a perfectly symmetrical face but rather to adjust the landmarks to a position that represents a balanced average, reducing the bias introduced by facial angles feature locations in the image.

Another critical realization was the sensitivity of the jawline feature to variations in facial angles in the photographs. Including the jawline directly in the feature vector would likely skew the distance calculations due to these variations.

Same person, different jawline positions due to angle

To resolve this, I devised a strategy to derive a feature vector from within each image. This vector comprised relative distances between certain pairs of facial landmarks, normalized by the width of the left eye. This normalization was crucial in ensuring that the features remained consistent and comparable across different images, regardless of the face's angle or position in the photo.

With these symmetric, aligned, and normalized features, I was then able to apply the Euclidean distance method effectively to compare two images.

The next challenge was determining the optimal threshold for face matching. Rather than arbitrarily choosing a threshold, I undertook a systematic approach. I observed the distances between the test images and those in the database to determine a range of potential thresholds (from 50 to 120). I then created and evaluated models for each threshold within this range,

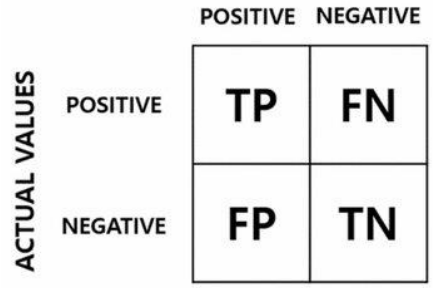
assessing their performance based on accuracy, precision, recall, and F1-score. This thorough evaluation enabled me to select the best threshold that balanced the system's sensitivity and specificity.

Experiments

In this phase, my focus was on exploring the impact of various threshold values on the system's ability to authenticate faces accurately. I conducted experiments using a range of thresholds, from 50 to 120, each corresponding to a different model of the face authentication system. These experiments were crucial for determining the most effective threshold setting that would balance sensitivity and specificity.

1. Methodology of Experiments

- For each threshold value within the specified range, I created a separate instance of the face authentication system.
- Each model was then tested using the 'Test_DB'. This involved attempting to authenticate each face in the test set and recording the system's decision.
- To assess the performance of each model, I calculated key metrics such as accuracy, precision, recall, and F1-score. These metrics provided insights into how well each threshold setting was performing in terms of correctly identifying both authentic and imposter faces.



$$Precision = \frac{TP}{TP + FP} \qquad Recall = \frac{TP}{TP + FN}$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN}$$

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

  - Accuracy is the proportion of total predictions (both authentic and imposter faces) that the system got right.
  - Precision indicates the proportion of positive identifications that were actually correct. In the context of this project's topic, it tells the number of correct face matches out of all the faces the system identified.
  - Recall measures the proportion of actual positives that were correctly identified. For this system, it reflects how many actual known faces were correctly recognized by the system.
  - F1-score is the harmonic mean of precision and recall. It provides a single metric that balances both precision and recall, ensuring that the system is neither too strict (rejecting genuine individuals) nor too lenient (accepting imposters).
- The calculated metric results of all tested threshold are in **'performance.csv'**

2. Observations from Experiments:

- The experiments revealed varying levels of performance across different thresholds. Lower thresholds generally led to higher precision but lower recall, indicating a stricter system that tended to reject more faces. Conversely, higher thresholds showed increased recall but reduced precision, suggesting a more lenient system with a tendency to accept more faces.
- Through careful analysis of these metrics, I decided to use F1-score as the key metric to decide which threshold results in the best performance of a model.

Results

1. Model selection

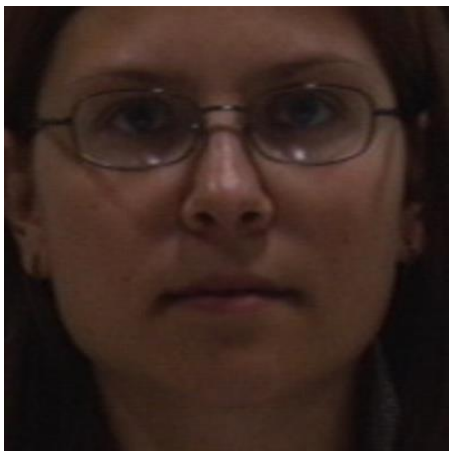
   Following the extensive experiments, I selected the best threshold for the system, which is 73. This threshold was the one that provided the most balanced performance in terms of both recognizing authentic individuals and rejecting unauthorized ones.

2. Implementation of the best model
   I implemented the face authentication system using the identified best threshold. The system was then subjected to a final round of testing with the 'Test_DB' to evaluate its effectiveness.
3. Final results of the best model

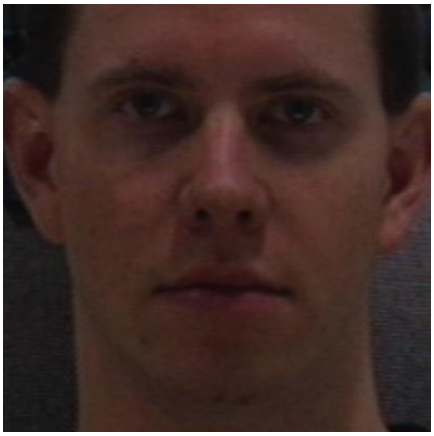
- The calculated metrics for this threshold were: Accuracy = 0.5222, Precision = 0.5769, Recall = 0.5882, F1-score = 0.5825
- The results of each image in the test dataset were then examined, looking at whether each face was authorized ('is_auths' – is authorized), whether it exists in the database ('is_in_DBs' – is authentic), the identity assigned by the system ('authorize_persons' – matched person), and the final verdict of each test ('results' – verdict).
- The results of the whole test set were recorded and saved in **'best_threshold_result.csv'**.
- Results of testing with some images (according to the recorded results from before)
  - Correct authorization:

Test image: Chloe_001.jpg

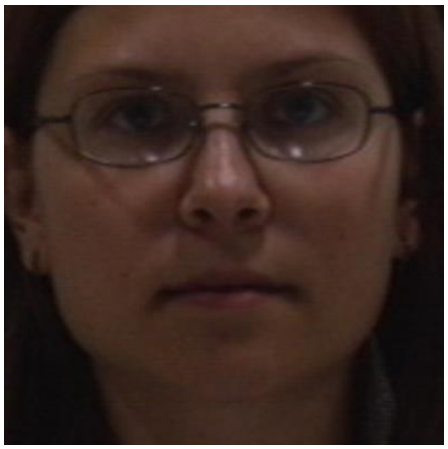
Database image: Chloe_000.jpg



- Incorrect authorization:
  i. Accepting imposters
  Test image: Ethan_002.jpg
  Matched person in database: Sebastian



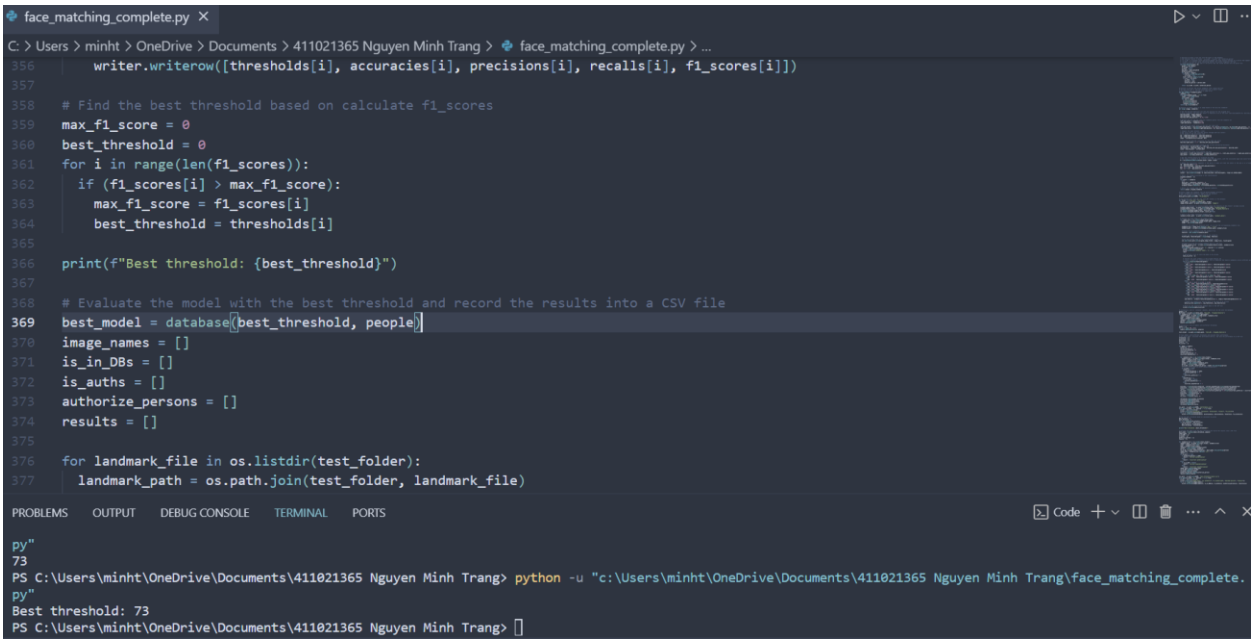
  ii. Wrong identity
  Test image: Zoe_000.jpg
  Matched person in database: Chloe



- Correct unauthorization
  Test image: Lucas_004.jpg

- Incorrect unauthorization
  Test image: Wyatt_004.jpg

- Code running: To show that the code runs properly without any error, I tried to print out the best threshold. As can be seen from the below image, “Best threshold: 73” was printed out as the result of the code.



Discussion

The process of developing and testing this face authentication system has provided a practical understanding of the challenges and considerations involved in facial recognition technology. While the system shows potential, it also highlights areas where further development and refinement are needed.

Through careful experimentation with different threshold values and the implementation of basic image processing techniques, this project has shed light on the complexities of accurately

identifying and authenticating individuals based on facial features. The approach of aligning faces using the eyes as reference points and making the features more symmetric helped address some of the inherent difficulties posed by the variability of facial orientations and expressions in the dataset.

The decision to adjust the facial features for symmetry and to exclude the jawline from direct feature vector calculations was significant. It made the system more resistant to errors caused by variations in the way faces were positioned in the photographs. This step was a key factor in enhancing the system's ability to produce consistent results.

Conclusion

Through the development of this face authentication system, I have gained substantial knowledge and practical skills in the field of facial recognition technology. While the face authentication system I developed is an initial step and might not be groundbreaking, it has significantly contributed to my understanding of the challenges and intricacies involved in such systems.

One of the key insights I gained is the critical role of preprocessing in image analysis. Techniques like aligning faces based on eye landmarks and balancing facial features taught me the importance of consistent and standardized input data for effective analysis.

Additionally, the process of testing various threshold levels and observing their impact on the system's performance has deepened my understanding of empirical evaluation and data-driven decision-making. It highlighted the importance of carefully assessing each parameter's influence on the overall system functionality.

In essence, this project has significantly enhanced my technical expertise in image processing and facial recognition, as well as strengthened my problem-solving and analytical skills. These learnings are invaluable and will undoubtedly aid in my continued exploration and contributions to image processing fields.