Université Côte d'Azur

Master Informatics

Face Recognition Application

AI Models & Applications

Team Members: Camara Ibrahima
Diallo Mamadou Ougailou
Taha Rania

Instructor: Enrico Formenti

Academic Year: 2025–2026

Abstract

Face recognition is a key domain of computer vision, widely applied in areas such as security systems, biometric authentication, and identity verification. The objective of this project, carried out within the *AI Models & Applications* course, is to design and implement a complete face recognition application for celebrity identification. The proposed system is based on a structured pipeline that includes face detection, deep feature extraction using pre-trained neural networks, and supervised classification. Several models and approaches are evaluated and compared in order to assess their performance, robustness, and practical usability.

Contents

1 Introduction

Face recognition is a major research area in computer vision and artificial intelligence. It aims at identifying or verifying the identity of a person from digital images or video streams. Over the past decade, the emergence of deep learning techniques, particularly convolutional neural networks (CNNs), has led to significant improvements in the accuracy and robustness of face recognition systems.

These advances have enabled the deployment of face recognition technologies in a wide range of real-world applications, including biometric authentication, access control, surveillance, and human-computer interaction. Modern systems typically rely on deep feature representations, also known as facial embeddings, which capture discriminative information while remaining robust to variations such as pose, illumination, and facial expressions.

The objective of this project, conducted within the *AI Models & Applications* course, is to design and implement a complete face recognition application focused on celebrity identification. The proposed system is based on a structured processing pipeline that integrates multiple stages: person and face detection, deep feature extraction using pre-trained models, and supervised classification for identity prediction.

Beyond the implementation of the recognition pipeline, this project also emphasizes experimental evaluation and practical usability. Several models and approaches are compared in order to analyze their respective strengths and limitations in terms of accuracy and computational efficiency. In addition, a graphical user interface is developed to facilitate interaction with the system, enabling both single-image and batch processing modes.

2 Dataset and Data Preparation

2.1 Dataset Description

The dataset used in this project is the *Pins Face Recognition* dataset, publicly available on the Kaggle platform. This dataset is composed of facial images of well-known celebrities and contains a total of 105 distinct identities. Each class corresponds to a single celebrity and includes a variable number of images, reflecting real-world conditions such as data imbalance and diversity in image quality.

The images present significant variations in pose, illumination, facial expression, background, and resolution. Such variability makes the dataset particularly suitable for evaluating the robustness and generalization capabilities of face recognition systems in realistic scenarios.

2.2 Data Collection and Organization

The dataset was retrieved directly from Kaggle and organized into a structured directory hierarchy, where each subfolder corresponds to a specific identity label. This organization allows for straightforward loading of the data using standard deep learning frameworks and facilitates the management of class labels during training and evaluation.

To ensure reproducibility and proper project management, raw datasets were excluded from the version control system using appropriate configuration files, while all processing scripts and trained models were tracked.

2.3 Preprocessing

Several preprocessing steps were applied to prepare the images for face recognition models. First, all images were converted to the RGB color space to ensure consistency across inputs. Then, images were resized to a fixed resolution of 224×224 pixels, which corresponds to the input size required by the pre-trained convolutional neural networks used for feature extraction.

Face detection techniques were applied to localize and crop facial regions when necessary, thereby reducing background noise and improving the quality of the extracted facial features. Finally, images were normalized according to the requirements of the selected deep learning models, ensuring numerical stability during inference.

2.4 Dataset Splitting

The dataset was split into training, validation, and test sets following a 70% / 20% / 10% ratio, respectively. A fixed random seed was used during the splitting process in order to guarantee reproducibility of the experimental results. This separation ensures that model evaluation is performed on previously unseen data, providing a reliable estimate of generalization performance.

3 Methodology

This section describes the methodology adopted to design and implement the face recognition application. The proposed approach follows a modular and structured pipeline, allowing each component of the system to be developed, evaluated, and improved independently. The methodology was designed to address the main challenges of face recognition, including robust face localization, discriminative feature extraction, and reliable identity classification.

3.1 Overall Pipeline

The face recognition system is based on a sequential pipeline composed of three main stages: face detection, feature extraction, and identity prediction. Given an input image, the system first detects the presence of persons and localizes facial regions. These detected faces are then processed by a deep neural network to extract compact and discriminative feature representations, commonly referred to as facial embeddings. Finally, a supervised classification model is applied to these embeddings in order to predict the corresponding identity.

Figure 1 illustrates the overall architecture of the proposed face recognition system, highlighting the sequential processing stages from input image to identity prediction.

This modular design ensures flexibility and facilitates experimentation with different models at each stage of the pipeline.

3.2 Project Architecture

The project follows a modular and well-structured architecture, designed to separate data preparation, core processing, visualization, and reporting. This organization facilitates code readability, maintenance, and reproducibility.

The `pre` directory contains all scripts related to data preparation, preprocessing, and post-processing tasks, including face extraction, statistics computation, and result visualization. These scripts are used to prepare the data and analyze intermediate results without interfering with the core application logic.

The `src` directory includes the main implementation of the face recognition system. It contains the core pipeline responsible for orchestrating detection, feature extraction, and classification, as well as the graphical user interface used for interaction and visualization.

All graphical resources, such as figures, diagrams, and generated plots, are stored in the `img` directory. This includes the pipeline overview diagram and the experimental result visualizations used in the report.

Finally, the `rep` directory contains the report sources written in LaTeX, ensuring full reproducibility of the final document and compliance with submission requirements.

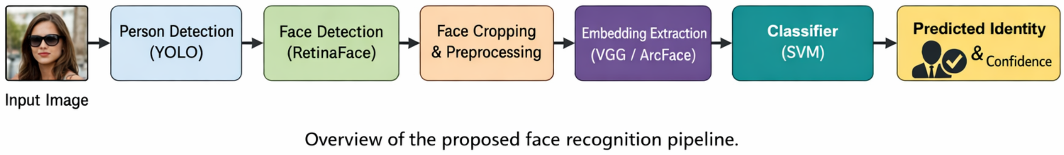

Figure 1: Overview of the proposed face recognition pipeline

3.3 Face Detection

Face detection plays a critical role in the overall performance of the recognition system, as inaccurate localization may significantly degrade the quality of the extracted features. To address this issue, two complementary detection approaches were explored.

First, a person detection stage is applied to locate human figures within the image. This step allows the system to focus on relevant regions and reduce the influence of background noise. Subsequently, a specialized face detection model is used to precisely localize facial regions within the detected person bounding boxes.

This two-step detection strategy improves robustness in complex scenes and ensures accurate face cropping, even in the presence of variations in pose, scale, and illumination.

3.4 Feature Extraction

Once facial regions are detected and cropped, deep feature extraction is performed using pre-trained convolutional neural networks. These networks transform facial images into fixed-length embedding vectors that capture high-level and discriminative facial information.

Multiple deep models were considered in order to analyze their impact on recognition performance. Pre-trained architectures were selected due to their strong generalization capabilities and their ability to leverage large-scale training on external datasets. Using pre-trained models also reduces computational cost and training time while maintaining high recognition accuracy.

3.5 Identity Classification

The extracted facial embeddings are used as input to a supervised classification model responsible for identity prediction. This approach decouples feature learning from classification, allowing the classifier to operate on compact and meaningful representations.

A classical machine learning classifier is employed to learn decision boundaries between different identities based on the embedding space. This choice provides interpretability, efficient training, and good performance, particularly when combined with high-quality deep facial embeddings.

3.6 Evaluation Strategy

To objectively assess the effectiveness of the proposed methodology, the system is evaluated on a held-out test set containing unseen images. Quantitative evaluation is performed using standard classification metrics, enabling a fair comparison between different model configurations. This evaluation strategy ensures that the reported results accurately reflect the generalization capabilities of the face recognition system.

4 Application and User Interface

To ensure practical usability of the proposed face recognition system, a graphical user interface was developed on top of the underlying recognition pipeline. The main objective of this interface is to allow users to interact with the system in an intuitive manner, without requiring prior knowledge of the implemented models or algorithms.

The application was designed with a focus on responsiveness, transparency of the processing steps, and real-time feedback. It provides clear visual indicators for each stage of the pipeline and supports both qualitative inspection and quantitative evaluation of the system.

4.1 Application Overview

The interface integrates the complete face recognition workflow, from image selection to final identity prediction. All processing stages, including person detection, face extraction, feature embedding computation, and classification, are executed automatically once an input is provided.

The modular architecture of the application allows each component of the pipeline to operate independently. This design choice improves robustness, facilitates debugging, and enables future extensions or model replacements with minimal changes to the overall system.

4.2 Processing Modes

The application supports two complementary processing modes.

The first mode focuses on single-image processing. In this mode, the user selects an image through the interface and receives immediate visual feedback, including detected regions, predicted identity, confidence level, and processing time. This mode is particularly useful for demonstrations and qualitative analysis of individual predictions.

The second mode enables batch processing of large image collections. Images are processed sequentially in an asynchronous manner, ensuring that the user interface remains responsive throughout execution. This mode allows efficient large-scale evaluation of the system and supports the generation of aggregated statistics, such as processing time, throughput, and confidence distribution across the dataset.

4.3 Pipeline Visualization

To improve interpretability and user understanding, the interface provides a visual representation of the face recognition pipeline. Each processing stage is associated with a dedicated status indicator that reflects its current state, such as waiting, running, successful completion, or error.

Detected persons and extracted faces are highlighted directly on the input images using bounding boxes, while predicted identities are displayed alongside confidence indicators. This visualization facilitates qualitative assessment of the system's behavior and helps identify potential sources of error.

4.4 Monitoring and User Experience

In addition to prediction results, the application provides real-time monitoring information during execution. This includes progress indicators, elapsed time, and summary statistics related to the recognition outcomes. Informative messages and explicit error notifications are displayed when necessary, contributing to a smooth and transparent user experience.

Overall, the graphical interface transforms the face recognition pipeline into a complete and user-oriented application, bridging the gap between algorithmic development and practical deployment.

5 Results and Evaluation

This section presents the experimental results obtained with the proposed face recognition system and analyzes its performance under different configurations. The evaluation focuses on the accuracy of identity prediction, as well as on the practical behavior of the system when applied to real-world image collections.

5.1 Evaluation Protocol

To ensure a fair and objective evaluation, experiments were conducted on a held-out test set composed of images that were not used during training or validation. The evaluation protocol follows a closed-set identification scenario, where each test image is assumed to belong to one of the known identities present in the dataset.

All models were evaluated under identical conditions, using the same data splits and preprocessing pipeline. This protocol allows a direct comparison between different feature extraction models and classification strategies.

5.2 Evaluation Metrics

System performance was primarily assessed using classification accuracy, defined as the ratio of correctly identified images over the total number of test samples. In addition to accuracy, qualitative indicators such as confidence scores and error cases were analyzed in order to better understand the behavior of the system.

For batch processing experiments, aggregated statistics were computed, including the number of correctly identified images, unknown predictions, and processing failures. These metrics provide a global view of system reliability and robustness.

5.3 Quantitative Results

The quantitative evaluation focuses on the impact of data quality and model configuration on face recognition performance. Rather than reporting isolated metrics, the analysis emphasizes comparative trends across pipeline stages and experimental settings, providing insight into the robustness and reliability of the proposed system.

Figure 2 shows the impact of reference database quality on recognition performance. Cleaning the celebrity reference images leads to a clear improvement in accuracy, increasing from 56% to 61.1%. This result highlights the importance of high-quality facial representations for embedding-based face recognition systems.

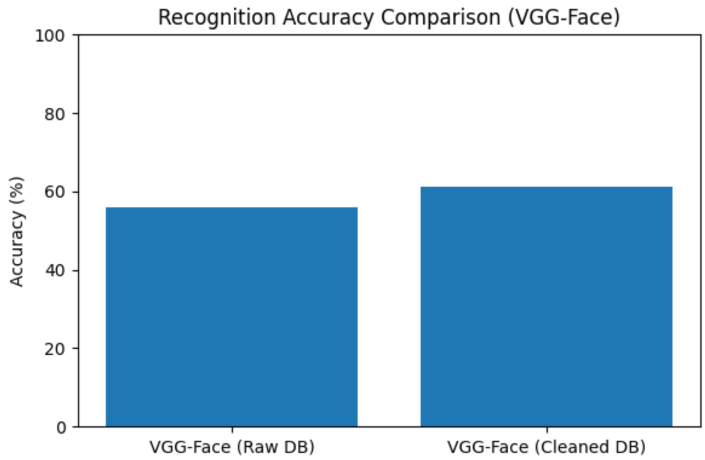

Figure 2: Recognition accuracy comparison before and after reference database cleaning using the VGG-Face model

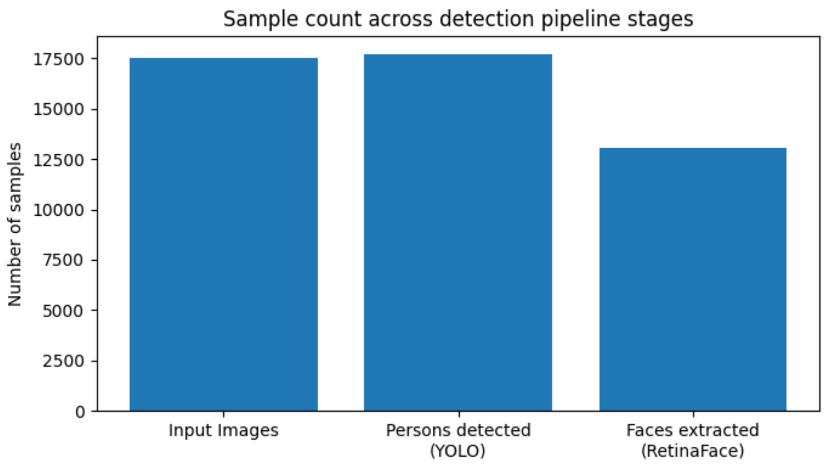

Figure 3: Number of samples retained at each stage of the detection pipeline

Figure 3 illustrates the evolution of the number of samples throughout the detection pipeline.

Pipeline Stage	Method	Samples / Images	Remarks
Input dataset	Raw images	17,513	High variability and noise
Person detection	YOLO	17,707	Background individuals detected
Face detection	RetinaFace (thr = 0.5)	13,043	Quality-based filtering
Reference database	Celebrity_db	105	One high-quality image per identity
Face recognition	VGG-Face + SVM	61.1% accuracy	Cleaned reference data

Table 1: Summary of detection and recognition results across the pipeline

Table 1 provides a global overview of the face recognition pipeline, highlighting the progressive filtering of samples and the impact of data quality on recognition performance.

The summarized results emphasize the cumulative effect of design choices across the pipeline. While early detection stages aim to preserve recall, later filtering steps prioritize face quality, which directly impacts recognition accuracy. This confirms that system performance is not determined by a single component, but rather by the interaction between detection, data preparation, and feature representation.

5.4 Detection and Recognition Statistics

The evaluation of the detection pipeline shows that the YOLO-based person detection stage successfully identifies a large majority of relevant subjects, with 17,707 detected persons across 17,513 images. This result confirms the robustness of the detection stage in complex scenes containing multiple individuals.

Following face detection with RetinaFace, a total of 13,043 valid face images were extracted. The reduction in the number of samples highlights the impact of strict face localization and quality filtering, which is necessary to ensure reliable feature extraction in later stages.

For the recognition stage, a dedicated reference database (*celebrity_db*) was created, containing one high-quality image per celebrity. A total of 105 reference faces were extracted using RetinaFace without lowering the detection threshold. This curated dataset serves as a clean and stable basis for identity recognition using deep embeddings.

5.5 Qualitative Analysis

Beyond numerical results, qualitative analysis was performed by inspecting prediction outputs and visual annotations produced by the application. Correctly classified examples demonstrate the system's ability to generalize across different facial appearances, while misclassified cases often reveal limitations related to occlusions, low resolution, or ambiguous facial features.

The visualization of detected faces, bounding boxes, and confidence indicators proved particularly useful for diagnosing errors and understanding the decision process of the recognition pipeline.

5.6 Discussion of Results

Overall, the results confirm the effectiveness of the proposed methodology for celebrity face recognition. While higher-performance models yield better accuracy, they also introduce additional computational cost. This trade-off highlights the importance of balancing accuracy and efficiency depending on the intended application scenario.

The batch processing experiments further demonstrate the practical applicability of the system, enabling large-scale evaluation and consistent performance monitoring.

6 Discussion

The experimental results obtained in this project provide valuable insights into the behavior and limitations of face recognition systems based on deep feature embeddings. The comparison between different configurations highlights the strong influence of the feature extraction stage on overall recognition performance.

High-capacity embedding models demonstrate superior discriminative power and improved robustness to variations in pose, illumination, and facial expression. However, these gains come at the cost of increased computational complexity, which may impact real-time performance in resource-constrained environments. This observation underlines the importance of selecting models according to the specific requirements of the target application.

Error analysis reveals that most misclassifications occur in challenging scenarios, such as low-resolution images, partial occlusions, or visually similar identities. These limitations are largely inherent to the data quality and highlight the dependency of face recognition systems on well-curated datasets and accurate face detection.

From a practical perspective, the integration of visualization tools and aggregated statistics within the application proved to be particularly useful. These features facilitate model diagnosis, system monitoring, and qualitative assessment, making the application not only a recognition system but also an effective experimental platform.

6.1 Challenges and Limitations

Despite the encouraging results obtained, several challenges and limitations were identified throughout the development of the face recognition system.

A first limitation is related to data quality. The original dataset contains a significant amount of noise, including blurred images, occlusions, and background individuals. Although the use of strict face detection thresholds improves the quality of extracted faces, it also leads to a reduction in the number of usable samples, which may impact recall.

Another limitation concerns the dependency of recognition performance on the reference database. Experimental results clearly show that the accuracy of embedding-based recognition systems is highly sensitive to the quality of reference images. Poorly aligned or low-quality reference faces can significantly degrade performance, even when using strong deep models.

In addition, the current system operates in a closed-set recognition scenario, where all test identities are assumed to be known in advance. This assumption limits applicability in real-world settings, where unknown identities may frequently appear.

Finally, computational cost represents a practical constraint. Deep feature extraction, particularly with high-capacity models, introduces non-negligible processing time, which may limit scalability or real-time deployment without further optimization.

7 Conclusion and Perspectives

In this project, a complete face recognition application was designed and implemented within the framework of the *AI Models & Applications* course. The proposed system integrates all key components of a modern face recognition pipeline, including face detection, deep feature extraction, supervised classification, and a user-friendly graphical interface.

The experimental evaluation demonstrates that the chosen methodology is effective for celebrity identification and highlights the advantages of deep embedding-based approaches. The application successfully supports both single-image analysis and large-scale batch processing, enabling comprehensive evaluation and practical usability.

Several perspectives for future work can be considered. These include extending the system to open-set recognition scenarios, fine-tuning deep models on domain-specific data, and exploring alternative classification strategies. Additionally, deploying the application as a web-based or real-time system could further enhance its applicability.

Overall, this project demonstrates that effective face recognition in unconstrained environments relies as much on data quality and pipeline design as on the choice of deep models. The results underline the importance of careful preprocessing and evaluation when deploying recognition systems in real-world scenarios.

8 Team Members and Contributions

This project was carried out collaboratively by the members of Team 3. Each member was responsible for a specific part of the system in order to ensure an efficient and structured development process.

- **Diallo Mamadou Ougailou** was responsible for the design and implementation of the detection pipeline, including person detection and face localization. This contribution also involved qualitative evaluation of detection results and the analysis of associated limitations.

- **Camara Ibrahima** focused on the face recognition component of the system. This included feature extraction using deep learning models, training and evaluation of the classification module, and implementation of the inference pipeline for identity prediction.
- **Taha Rania** was in charge of system evaluation and result analysis. This role included the computation of evaluation metrics, visualization of results, and the rédaction of the final report, as well as verification of compliance with the project requirements.