KF6007 - Artificial Intelligence and Robotics

Facial Recognition Attendance System Report

W21043564

Contents

[1. Introduction 3](#_Toc167669216)

[2. Literature Review 3](#_Toc167669217)

[3. Novel AI Techniques and Approaches 4](#_Toc167669218)

[4. Product Development Stages 5](#_Toc167669219)

[4.1. Methodology 5](#_Toc167669220)

[4.2. Design 6](#_Toc167669221)

[4.2.1. UI Design 6](#_Toc167669222)

[4.2.2. AI Methods Selection 6](#_Toc167669223)

[4.2.3. Programming Tools and Libraries 6](#_Toc167669224)

[4.2.4. Dataset Preparation 6](#_Toc167669225)

[4.2.5. System Architecture 6](#_Toc167669226)

[4.3. Implementation 8](#_Toc167669227)

[4.3.1. System Interface 8](#_Toc167669228)

[4.3.2. Face Detection and Recognition 8](#_Toc167669229)

[4.3.3. CNN Model Architecture and Training 9](#_Toc167669230)

[4.3.4. Transfer Learning 10](#_Toc167669231)

[4.3.5. Attendance Management 10](#_Toc167669232)

[5. Critical Evaluation 10](#_Toc167669233)

[5.1. Black-box Testing 11](#_Toc167669234)

[9.1. White-box Testing 11](#_Toc167669235)

[9.2. Performance Evaluation 12](#_Toc167669236)

[9.3. Strengths and Limitations 12](#_Toc167669237)

[9.4. Future Work 12](#_Toc167669238)

[References 13](#_Toc167669239)

[**Appendix A: Swimlane Activity Diagrams** 15](#_Toc167669240)

[**Appendix B: Unit Testing** 17](#_Toc167669241)

[**Appendix C: Confusion Matrix Evaluation** 19](#_Toc167669242)

# Introduction

In modern workplaces, accurate and efficient attendance tracking is crucial for operational management. Traditional methods such as RFID check-ins, roll calls, and sign-in sheets are labour-intensive, error-prone, and susceptible to fraudulent practices. To address these challenges, a facial recognition attendance system is proposed that utilizes Haar-cascade model for initial face detection and Convolutional Neural Network (CNN) model for face classification. Transfer learning is incorporated to allow addition of new members.

The following sections of this report provide an overview of the system's development. **Section 2** reviews relevant literature and highlights existing solutions and their limitations. **Section 3** details novel AI techniques and approaches utilized. **Section 4** elaborates on the product’s development stages. **Section 5** offers an evaluation of the system and discusses its effectiveness and areas for improvement.

# Literature Review

Although facial recognition attendance systems offer enhanced efficiency, accuracy, and security in attendance monitoring, due to high setup and maintenance costs, privacy concerns, unstable accuracy, and potential bias, standard attendance practices are commonly adopted.

Generally used in facial recognition, a Haar-cascade classifier is a machine learning-based approach where a cascade function is trained and Haar-like features are utilized to detect objects (Jaiswal, 2022). Abdulhussien and Saud (2022) state that Viola-Jones algorithm (Haar-cascade classifier) offers high detection speeds, low computational costs, and processes images in real-time. However, it struggles in scenarios with occlusion or varying lighting conditions and exhibits limited accuracy compared to deep learning-based methods. It is used in biometrics, surveillance, sign language detection, etc.

Convolutional Neural Networks (CNNs) are deep learning models designed to process structured grid data like images (IBM, 2023). Yeh et al. (2023) performed face classification using CNNs and achieved 70% accuracy, highlighting its high accuracy capabilities across multiple classes. Pooja et al. (2024) utilized CNNs for attendance systems and found that it simplifies attendance management processes and offers high accuracy and efficiency. Although CNNs are capable of learning large amounts of data and outperform traditional machine learning techniques (Alzubaidi et al., 2021), they have high computational requirements and are difficult with small datasets which can lead to overfitting (Younesi et al., 2024).

Moreover, CNNs have the ability to learn hierarchical features from input data and are well-suited for transfer learning which involves leveraging knowledge learned from one domain to improve performance on a related task (Murel & Kavlakoglu, 2024). Salehi et al. (2023) utilized transfer learning in medical imaging and found that it enhances accuracy, reduces resource requirements, and addresses class imbalances. Although it reduces training time and improves generalization, it has an overfitting risk and is difficult to finetune (Vinithavn, 2021; Murel & Kavlakoglu, 2024).

# Novel AI Techniques and Approaches

The proposed system introduces several novel AI techniques and approaches to address the limitations of traditional attendance tracking methods. It employs an integrated approach that synergistically combines the strengths of three powerful techniques: Haar-cascade classifiers for efficient face detection, CNNs for accurate facial recognition, and transfer learning for member additions.

Although Haar-cascade classifiers are well-known for their computational efficiency and real-time processing capabilities that make them ideal for initial face detection, their accuracy can be limited in challenging scenarios (Abdulhussien & Saud, 2022). Hence, the system leverages the superior performance of CNNs for facial recognition tasks. Hong et al. (2021) performed lung disease classification using CNNs and achieved 96.1% accuracy with four-class predictions, further demonstrating CNNs’ state-of-the-art results in computer vision tasks.

A key novel aspect of the proposed system is the incorporation of transfer learning which allows the system to adapt to new members without requiring extensive retraining from scratch. By leveraging this technique, the system can efficiently fine-tune the model with limited data from new individuals and significantly reduce the time and computational resources required for training (Vinithavn, 2021).

Privacy and ethical considerations are vital in the development of facial recognition systems. To address these concerns, the proposed system complies with the data protection laws of UK GDPR by obtaining explicit consent before data collection and processing, only collecting necessary facial data, only using the data for defined purposes, and providing transparency about data usage.

The system explores integrating with existing organizations to streamline data management, enhance security, and improve operational efficiency by consolidating attendance data with other systems.

By combining novel AI techniques, addressing privacy and ethical concerns, and exploring deployment and integration possibilities, the proposed system aims to provide a comprehensive and innovative solution for accurate and efficient attendance tracking in various settings.

# Product Development Stages

## Methodology

The proposed system aims to provide an accurate and efficient solution for attendance tracking in various settings such as workplaces, educational institutions, and organizations. It addresses the limitations of traditional methods which are prone to errors, time-consuming, and susceptible to fraudulent practices.

The system employs a multi-stage approach by leveraging the strengths of different AI techniques. The initial stage involves face detection using a Haar-cascade classifier since they have proven to be computationally efficient and capable of real-time processing (Abdulhussien & Saud, 2022). This stage quickly identifies regions of interest in the input image containing faces.

The CNN model is responsible for facial recognition and classification. CNNs are deep learning models that excel at extracting complex features from raw data, making them well-suited for computer vision tasks (IBM, 2023). The CNN model is trained on a dataset of facial images, allowing it to learn the unique patterns and features associated with each individual.

To enhance the system's adaptability and scalability, transfer learning techniques are employed which enables efficient adaptation of the CNN model to new individuals without requiring extensive retraining from scratch, significantly reducing the computational resources and time required.

The decision-making process for attendance tracking involves the following steps:

1. Face detection using Haar-cascade classifier.
2. Preprocessing and normalization of detected faces.
3. Passing the preprocessed images through the CNN model for facial recognition.
4. Updating the attendance records accordingly.

The system's modular design allows for easy integration with existing organizational systems such as HR, payroll, or access control systems, enabling streamlined data management and improved operational efficiency.

## Design

### UI Design

The UI is designed using the *customtkinter* library which provides a modern interface. The decision to use a graphical user interface (GUI) is made to enhance user experience and accessibility. The GUI allows users to interact with the system without the need for command-line interfaces or complex configurations.

### AI Methods Selection

The system integrates three key AI techniques: Haar-cascade classifier for initial face detection, CNN for facial recognition, and transfer learning for adapting the CNN model to new individuals. This combination is chosen to leverage the strengths of each technique, ensuring both computational efficiency and high accuracy.

### Programming Tools and Libraries

Python is selected as the primary programming language due to its extensive library support and ease of use in developing AI applications. The *OpenCV* library is utilized for image processing and computer vision tasks while *Keras* is employed for building and training the CNN model. *Sqlite3* library was using for database manipulation.

### Dataset Preparation

The dataset is compiled using web-scraped images obtained via the *bing\_image\_downloader* library. A UI feature enables users to add their own facial images for employee addition. After collection, images undergo preprocessing steps including face detection, cropping, resizing, and normalization to ensure consistency and enhance model performance. The dataset is then divided into training, validation, and testing sets for model development and assessment.

### System Architecture

The system follows a modular architecture as depicted in the UML class diagram (**Figure 1**). The core components include a UI, face detector, CNN model for face recognition, preprocessor, display interface for live feed, database to store attendance data, data collector for new members, dataset to store image data, and a web-based image scraper. This modular design ensures scalability and ease of integration with existing organizational systems or databases.

A black background with white text

Description automatically generated

**Figure 1.** UML Class Diagram for Facial Recognition Attendance System.

By carefully considering the design choices for the UI, AI methods, programming tools, dataset preparation, and system architecture, the proposed system aims to strike a balance between accuracy, efficiency, and user-friendliness while maintaining the potential for future expansions and integrations.

## Implementation

### System Interface

The system's UI (**Figure 2**) provides a modern experience displaying an image along with three buttons: "Start," "Add Employee," and "Exit." The "Start" button initiates the facial recognition process, while the "Add Employee" button opens a popup window for adding new employees.

A camera lens with a circular lens

Description automatically generated

**Figure 2.** System’s UI.

### Face Detection and Recognition

The system uses a pre-trained Haar-cascade classifier for real-time face detection from the webcam feed. A square border is shown across the user’s face to indicate successful face detection. The detected face is then extracted, preprocessed, and passed through the trained CNN model for facial recognition. The predicted results (name, prediction confidence, date, time) are displayed to users in real-time after 10 consecutive positive recognitions (**Figure 3**). Additionally, key-binding instructions are displayed for guidance. The classifier constantly listens for key inputs, allowing users to communicate with the display interface effectively. See **Figure A2** for face recognition swimlane diagram.

A hand holding a picture of a person

Description automatically generated

**Figure 3.** System’s Display Interface.

### CNN Model Architecture and Training

The CNN model (**Figure 4**) consists of five convolutional layers with ReLU activation, followed by max-pooling layers for feature extraction. The final layers include a flatten layer, a dense layer with regularization and dropout to prevent overfitting, and an output dense layer with softmax activation for multi-class classification.

A screen shot of a computer program

Description automatically generated

**Figure 4.** CNN Model.

The model is trained on a web-scraped dataset of 175 images per class. The data is preprocessed, split into training (60%), validation (20%), and testing (20%), and normalized. Training uses the Adam optimizer (learning rate 0.0001) for 35 epochs with a batch-size of 80, employing early stopping and learning rate reduction based on validation loss to prevent overfitting.

### Transfer Learning

To adapt the CNN model to new employees without extensive retraining, transfer learning is implemented. When a new employee is added (**Figure 5**), the last layer is replaced with a new dense layer with the updated number of classes. The model is then fine-tuned on the new employee's data. See **Figure A1** for member addition swimlane diagram.

A screenshot of a computer

Description automatically generated

**Figure 5.** Adding new employees.

### Attendance Management

The attendance management system is implemented using an SQLite database (**Figure 6**). When an employee is successfully recognized, their attendance record is updated in the database with the login or logout time. The attendance records can be easily retrieved and integrated with existing HR or payroll systems.

A close up of a login

Description automatically generated

**Figure 6.** Database structure.

# Critical Evaluation

The system was tested for robustness, accuracy, and reliability using black-box and white-box testing.

## Black-box Testing

Black-box testing focused on the system's functionality and user experience, evaluating:

1. **User Interface:** Assessed usability, responsiveness, and aesthetics through tasks like adding employees, attendance tracking, and system exit.
2. **Display Interface:** Tested the camera feed and keyboard input functionality.
3. **Face Detection and Recognition:** Evaluated the system's consistency in detecting and recognizing faces under various conditions.
4. **Attendance Management:** Verified the functionality for updating attendance records for login and logout in the database.

## White-box Testing

White-box testing involved evaluating the CNN model and transfer learning approach. Unit tests were conducted to ensure the correctness of individual components (**Appendix B**).

A screenshot of a diagram

Description automatically generated

**Figure 7.** Test data confusion matrix.

## Performance Evaluation

The CNN model achieved 83.33% accuracy and 0.94 loss, demonstrating its effectiveness in recognizing faces, a precision of 0.84, reflecting its ability to correctly identify positive cases, a recall of 0.83, indicating its capability to capture relevant instances, and an F1-score of 0.83, highlighting the model's overall performance.

The confusion matrix (**Figure 7**) shows high accuracy for most classes, indicated by the strong diagonal. However, notable misclassifications occur, especially between Andy Samberg and Emilia Clarke. This suggests the model is effective but could improve in distinguishing certain classes (see **Appendix C** for details). Enhancing the model or adding more diverse training data could help reduce these misclassifications.

The transfer learning approach for adding employees is suboptimal due to the need for finetuning with each addition. Retraining the entire model is more efficient for automation and requires less maintenance.

## Strengths and Limitations

The system offers efficient and accurate attendance tracking with a modern interface and high model accuracy. It automates the attendance process, provides data analytics, integrates seamlessly with existing systems, and offers a contactless solution. This convenience and the ability to collect valuable data makes it an attractive option for many organizations.

However, it raises privacy concerns and incurs high setup and maintenance costs. Face detection can be affected by varying lighting conditions and is susceptible to 2D image scans. Potential biases in recognition and the need for robust infrastructure present challenges. Transfer learning for adding employees requires manual fine-tuning. Although a lengthy process, automated training from scratch yields better results. Legal and ethical considerations also need to be addressed.

## Future Work

Future work could involve exploring advanced CNN architectures like *ResNet* to improve accuracy and robustness. Implementing continuous learning techniques would enable the system to adapt and learn from new data without frequent retraining. A mobile app could be developed to enable remote attendance. Integrating voice recognition, fingerprint scanning, or infrared sensors could enhance security and reliability. Exploring cloud-based deployment could enable integration with existing organizational systems and facilitate remote access and management.

# References

Abdulhussien, D.M. and Saud, L.J., 2022. AN EVALUATION STUDY OF FACE DETECTION BY VIOLA-JONES ALGORITHM. World Economics and Finance Bulletin, 14, pp.159-164.

Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M. and Farhan, L., 2021. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of big Data, 8, pp.1-74.

Bruce, B.R., Aitken, J.M., Petke, J. (2016). Deep Parameter Optimisation for Face Detection Using the Viola-Jones Algorithm in OpenCV. In: Sarro, F., Deb, K. (eds) Search Based Software Engineering. SSBSE 2016. Lecture Notes in Computer Science(), vol 9962. Springer, Cham. <https://doi.org/10.1007/978-3-319-47106-8_18>

Hong, M., Rim, B., Lee, H., Jang, H., Oh, J. and Choi, S., 2021. Multi-class classification of lung diseases using CNN models. Applied Sciences, 11(19), p.9289.

IBM (2023). *What are Convolutional Neural Networks? | IBM*. [online] www.ibm.com. Available at: <https://www.ibm.com/topics/convolutional-neural-networks>.

Jaiswal A. (2022). *Object Detection Using Haar Cascade: Analytics Vidhya*. [online] Available at: <https://www.analyticsvidhya.com/blog/2022/04/object-detection-using-haar-cascade-opencv/#:~:text=Haar%20cascade%20is%20an%20algorithm>.

Murel J., Kavlakoglu E. (2024). *What is transfer learning? | IBM*. [online] Available at: <https://www.ibm.com/topics/transfer-learning>.

Pooja, Grace, J., Rajalakshmi and Jasna, R. (2024). FACE RECOGNITION ATTENDANCE. [online] 9, p.601. Available at: <https://www.ijnrd.org/papers/IJNRD2404672.pdf> [Accessed 23 May 2024].

Salehi, A.W., Khan, S., Gupta, G., Alabduallah, B.I., Almjally, A., Alsolai, H., Siddiqui, T. and Mellit, A., 2023. A study of CNN and transfer learning in medical imaging: Advantages, challenges, future scope. Sustainability, 15(7), p.5930.

Singh, A. (2019). *Feature Descriptor | Hog Descriptor Tutorial*. [online] Analytics Vidhya. Available at: <https://www.analyticsvidhya.com/blog/2019/09/feature-engineering-images-introduction-hog-feature-descriptor/>.

Thalluri, L.N., Babburu, K., Madam, A.K. et al. Automated face recognition system for smart attendance application using convolutional neural networks. Int J Intell Robot Appl 8, 162–178 (2024). <https://doi.org/10.1007/s41315-023-00310-1>

Vinithavn (2021). *The Power Of Transfer Learning in Deep Learning*. [online] Analytics Vidhya. Available at: <https://medium.com/analytics-vidhya/the-power-of-transfer-learning-in-deep-learning-681f86a62f79>.

Yeh, J.F., Lin, K.M., Chang, C.C. and Wang, T.H., 2023. Expression Recognition of Multiple Faces Using a Convolution Neural Network Combining the Haar Cascade Classifier. Applied Sciences, 13(23), p.12737.

Younesi, A., Ansari, M., Fazli, M., Ejlali, A., Shafique, M. and Henkel, J., 2024. A Comprehensive Survey of Convolutions in Deep Learning: Applications, Challenges, and Future Trends. arXiv preprint arXiv:2402.15490.

# **Appendix A: Swimlane Activity Diagrams**

A black screen with white rectangles

Description automatically generated

**Figure A1.** Swimlane Diagram for Members Addition Task

A screenshot of a black screen

Description automatically generated

**Figure A2.** Swimlane Diagram for Attendance Operations.

# **Appendix B: Unit Testing**

#### UI Functions

For the user interface testing of the system, the functionalities of critical interface elements are validated. Unit testing is performed on the *next* method of the **Employee** class in *app.py* which involves multiple checks such as agreement to terms and conditions, presence check, length validation, and directory creation for new employees. A set of decoupled utility functions are created to facilitate the testing process using the *pytest* library. These functions are:

* **check\_agreement:** Validates whether the user has agreed to the terms and conditions.
* **check\_presence:** Ensures that the employee name is provided.
* **check\_length:** Checks that the employee name meets the length requirements.
* **check\_existing\_employee:** Verifies if the employee name already exists in the directory.
* **create\_employee\_directory:** Attempts to create a new directory for the employee.

By testing these functions individually, it is guaranteed that each piece of the logic works correctly. **Figure B1** shows the successful passed tests for these interface functions.

A computer screen shot of a black screen

Description automatically generated

**Figure B1.** Pass results of UI functions.

#### Database Management

In the attendance management section, the functionalities for handling employee login and logout records are tested. This included creating the database, inserting login times, and updating logout times. The following tests are performed:

* **Database Creation:** Verify that the database and necessary tables are created if they do not already exist.
* **Insert Login:** Ensure that login times are correctly inserted into the database.
* **Insert Logout:** Confirmed that logout times are properly updated in the database for the most recent login entry.

The test cases are designed to cover different scenarios such as valid logins, handling logouts without prior logins, and checking the database integrity after operations. **Figure B2** illustrates the successful passed tests for the attendance management functions.

A screen shot of a computer program

Description automatically generated

**Figure B2.** Pass results of database functions.

By testing the attendance management system, we can ensure that the system updates the database as expected.

# **Appendix C: Confusion Matrix Evaluation**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Class** | **Correctly Classified** | **Misclassified As** | **Total Instances** | **Notable Misclassifications** |
| *Alexandra Daddario* | 34 | 2 (1 as Henry Cavill, 1 as Leonardo DiCaprio) | 36 | Few misclassifications |
| *Andy Samberg* | 24 | 11 (1 as Emilia Clarke, 7 as Henry Cavill, 3 as Leonardo DiCaprio) | 35 | Frequent misclassification as Emilia Clarke and Leonardo DiCaprio |
| *Emilia Clarke* | 31 | 2 (1 as Alexandra Daddario, 1 as Leonardo DiCaprio) | 33 | High accuracy |
| *Henry Cavill* | 30 | 6 (4 as Andy Samberg, 1 as Alexandra Daddario, 1 as Leonardo DiCaprio) | 36 | Some misclassifications, particularly as Andy Samberg |
| *Leonardo DiCaprio* | 26 | 8 (2 as Alexandra Daddario, 2 as Emilia Clarke, 4 as Henry Cavill) | 34 | Some misclassifications, notably as Henry Cavill |

**Table C1.** Confusion matrix critical evaluation.