Optimum Science Journal

Journal homepage: https://optimumscience.org

Original Research Article

Optimized Algorithm for Face Recognition using Deepface and Multi-task Cascaded Convolutional Network (MTCNN)

Vaibhav Godase

SKN Sinhgad College of Engineering, Pandharpur, India

ARTICLE INFO

Received 25 March 2025

Accepted 02 May 2025

Available Online 05 May 2025

Keywords:

MTCNN

Deepface

Face Recognition

Optimized Algorithm

ABSTRACT

The combination of MTCNN and Deepface creates a robust and efficient facial recognition and recognition system to determine if a person is authorized. With 96.2% accuracy, MTCNN is known for its face detection and recognition accuracy. In addition to its excellent accuracy, MTCNN is also designed to operate in real-time, making it suitable for applications where instant face detection and recognition are critical. Additionally, MTCNN's offline capability increases its versatility, ensuring consistent performance even without an internet connection. On the other hand, Deepface acts as a light component in this combination. Although its accuracy is slightly lower at 95%, Deepface specializes in analyzing detailed facial features such as eyes, nose, and mouth. With this analysis, Deepface can extract facial features and provide additional information about the facial features of the recognized person as well as not recognized person. Despite its lightweight nature, Deepface's ability to analyze facial features adds depth to the facial recognition process and makes the system more comprehensive. Together, MTCNN and Deepface create a powerful synergy. The high accuracy and real-time capabilities of MTCNN combined with the detailed facial analysis of Deepface results in a system that offers both accuracy and efficiency. In addition, the combination ensures that the system works offline, making it reliable in various settings and conditions. Thus, this integrated approach provides an optimal balance between accuracy, lightweight design, and offline capability, making it well-suited for face recognition tasks where these factors are most important.

To cite this article: Godase, V. (2025). Optimized algorithm for face recognition using Deepface and Multi-task Cascaded Convolutional Network (MTCNN). *Optimum Science Journal*, Issue, pp-pp. http://doi.org/10.5281/zenodo.15341560

1. Introduction

An ordinary door is a common object in our everyday lives, yet it is frequently overlooked despite its vital function in offering protection and privacy. These modest gates, which are made of metal or wood, have patterns that range

from simple to elaborate (Dhope et al., 2024). Doors exist in a variety of types, from folding and revolving to swinging and sliding, and they enhance the beauty and functionality of structures. A modern marvel, the smart door symbolizes a revolutionary way to incorporate technology into our living areas. A smart door, in contrast to its conventional equivalents, has technologies that improve connection, security, and convenience. These doors have sensors, cameras, and smart locks built right in, and you can operate them from a distance using a smartphone or other linked device. Users can monitor and open their doors using keyless access. Smart door systems can easily integrate with home automation systems, allowing consumers to control their whole living space with a touch. The integration of biometric identification (Godase et al., 2024a). Facial recognition and real-time monitoring features enhance security and gives users a greater sense of security. By enabling automated modifications depending on occupancy and weather conditions, smart doors also help to reduce energy use. As technology develops further, smart doors represent the marriage of style and utility, influencing how we use and see our living areas. If we use a high-tech camera it's going to be costly and with cloud computing, it is going to be online, which is also costlier than the device we have created overcomes this problem (Godase et al., 2024b). So, the primary purpose of this research is to develop and evaluate an optimized face recognition system that combines the strengths of Deepface and Multi-task Cascaded Convolutional Network (MTCNN) to achieve high accuracy, real-time performance, and offline capability. The goal is to create a robust and efficient algorithm capable of accurately detecting and identifying human faces for security applications such as smart door access systems. By integrating lightweight and accurate models, the research aims to overcome the limitations of traditional high-cost, cloud-dependent systems, making facial recognition technology more accessible, reliable, and suitable for embedded and offline environments.

1.2. Literature Review

Numerous scientific disciplines show a great deal of interest in the Internet of Things, or IoT (Gupta et al., 2022). Technology has made life both simpler and more complicated. One significant way that smart digital automated systems based on the Internet of Things (IoT) benefit humans is through reducing their burden and incorporating interactive technology into their everyday activities. With our initiative, Smart Doors, we are contributing a small portion to the enormous efforts being made to improve and simplify our lives. This simple solution makes it easier for users to reach doors within a predetermined range. Android software will be able to access the door lock, and data will be transferred via Bluetooth technology. As a result, end users need not be concerned.

The experiment carried out by Nayana and Shashidhar (2019) shows that Today's lives depend heavily on security, which is the primary topic of this work. In this case, an Arduino microcontroller, GSM module, and fingerprint sensor are used. The microcontroller has already recorded fingerprints of authorized individuals. Using a matching algorithm, we can determine the identity of the individual. A One Time Password (OTP) has been issued by GSM to the authorized person's cellphone number if they are authorized. The buzzer will sound like he alerting him that someone is attempting to open the door if the fingerprint does not match the previously recorded fingerprints. Since the individual is not authorized, no OTP will be sent to his cell phone.

In the experiment conducted by Aswini et al. (2021), having a safe home is crucial in the modern world. Even with the use of thick, difficult-to-open metal locks, there are still many reasons to be concerned, such as theft and misplacing the keys. Many innovative technologies have been developed recently to address the shortcomings of conventional door-locking mechanisms. With just a single click, these solutions provide remote door access while simultaneously enhancing home security. One such technology that has greatly eased daily living is the Internet of Things, which offers answers to a wide range of issues. This article discusses a high-security solution for homes that combines OTP-driven technology with an RFID-based door lock system. The one-time password (OTP) that is generated in this device for door access will expire at the specified time. This article discusses the suggested system's working model.

The research conducted by Patil et al. (2018); the purpose of an Android-based smart door locking system is to deter unwanted entry, trees passing, and intrusion. Among the frequent locations for illegal entry, trespassing, and intrusion are banks, business offices, financial institutions, jewelry stores, and government buildings. Typically, the intent behind these kinds of actions is to steal valuables, cash, jewelry, or papers for personal benefit. The Android-based smart door-locking system aims to offer a clever way to get over these obstacles and offers a workable solution. This technique utilizes the idea of a predetermined password. In order to stop an attacker from unlocking something without authorization, it raises the security level. The system allows the user to reset or modify their passwords if they are forgotten. A more secure method of locking and unlocking the machine is provided by this automated password-based lock mechanism.

The research conducted by Djupsjö and Almosawi (2018), this thesis explains the creation of an Internet of Things application based on digitizing a smart door lock so that it may be connected to the Internet and used to identify office workers. The focus of this thesis is security, including a list of common security issues in IOT systems as a whole and an analysis of these issues to create a secure and workable solution from the ground up. For this project, a microcontroller is selected, and a test environment is constructed to explore and create security flaws. Architectural layouts are selected for the Android application as well as the API that is being created. The multi-master database that Azure Active Directory represents is explained in depth, along with its significance in preventing a critical security breech. In the project, a novel method known as Eddy stone is presented to support Bluetooth beacon transmission protocols. The creation of the android application and ensuring that all the established subsystems interact with one another to provide a safe and functioning IoT system flow constitutes the project's final stages.

The experiment carried out by Gupta et al. (2022); the Internet of Things, or IoT, is receiving a lot of interest from several scientific fields. Life is now easier and more complex thanks to technological technology. One important way is that an Internet of Things (IoT)-based smart digital automated system helps people by lowering their workload and integrating interactive technology into daily life. We are adding a tiny bit to the massive efforts being made to enhance and simplify our lives with our initiative, Smart Doors. This is a straightforward project that facilitates users' access to doors within a designated range. The door lock will be accessed by Android software and

Bluetooth technology will be used to transport data. Therefore, end customers don't have to worry about the door lock because they can operate it with their mobile if they're close enough, and they also don't have to worry about carrying around a real key. Both hardware and software development are included in this task. This method aims to eliminate the need for keys to access doors as well as the daily hassle of carrying several keys and having them ready for installation quickly and anywhere. By stopping security breaches, it contributes to the creation of a safe environment both inside and outside.

The experiment conducted by Mehmood et al. (2023); reliable, safe, and intelligent locking systems are desperately needed as security is becoming one of life's necessities. Because they eliminate the need for keys, cards, or insecure communication to avoid transporting, losing, duplicating, or hacking, stand-alone smart security solutions are quite appealing. Here, we present a smart door-locking system (DLS) that uses invisible touch sensors. By pasting the hybrid geometry copper electrodes on cellulose paper, a simple do-it-yourself (DIY) method is used to construct the passive transducer-based touch sensors. Because copper tape and paper are biodegradable and non-toxic materials, this design is an excellent fit for green electronics. The DLS keypad is covered with paper and spray paint to make it invisible for further security. Only the password and the locations of each key on the sensor keypad are necessary to unlock the door. Without any erroneous data, the system is able to identify the precise password pattern with efficiency accurately. Locking systems using invisible touch sensor technology may readily enhance the security of residences, banks, cars, flats, locker rooms, and cabinets.

2. Methodology

To detect and recognize the person's faces the Deepface and MTCNN (Multi-task Cascaded Convolutional Network) models are used (Essien & Ansa, 2023; Peña, 2023). The MTCNN model detects the face along with facial landmark alignment.

2.1. The MTCNN Model

2.1.1. The proposal network (P-Net): Regression vectors of potential windows and their bounding box are obtained using the proposal network. When detecting an object of a preset class—in this case, a face—bounding box regression is a popular technique for box position prediction. Following the acquisition of the bounding box vectors, the overlapping regions are combined through a series of adjustments. All candidate windows are the result of this stage, which lowers the number of candidates following refinement.

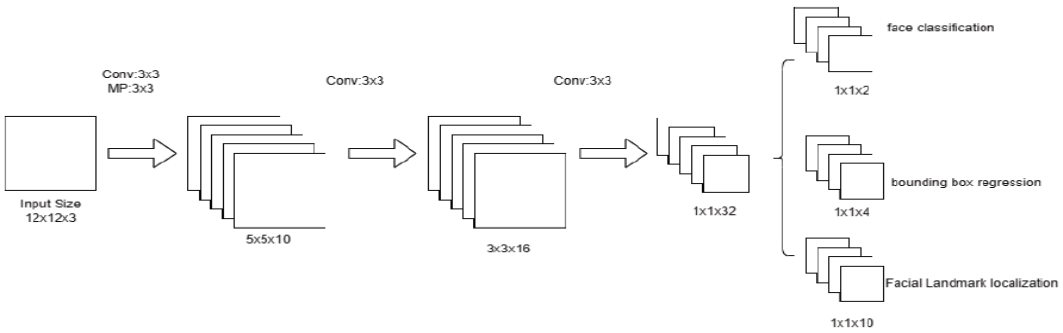

Figure 1. Proposal Network (P Net)

2.1.2. The refine network (R-Net): Refine Network is recommended for all P-Net applicants. Because of the dense layer in the last stage of the network architecture, this network is a CNN rather than an FCN like the one before it. R-Net further minimizes the number of candidates, applies bounding box changes, and merges overlapping candidates using non-maximum damping (NMS). R-Net produces results whether the input is a face. The face's bounding box is a 4-element vector, and the face orientation position is a 10-element vector.

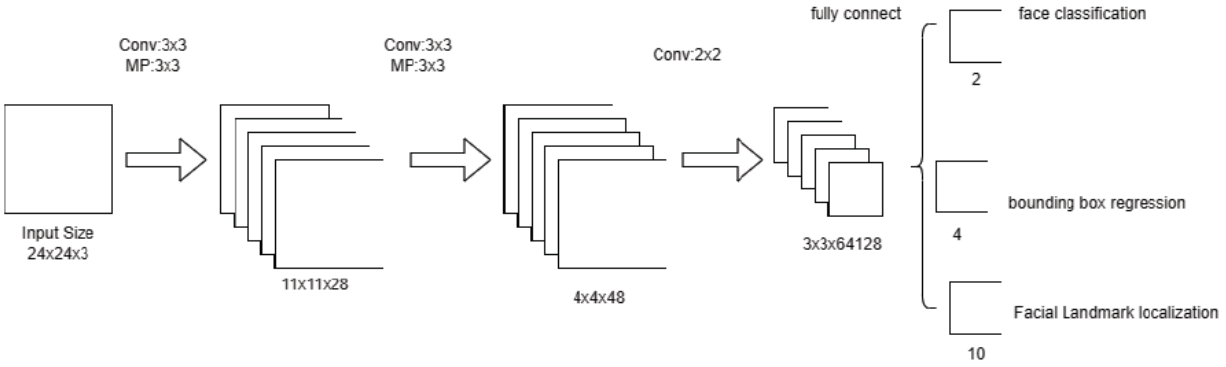

Figure 2. Refine Network

2.1.3. The output network (O-Net): This step is comparable to R-Net, but the output network's goal is to provide a more thorough description of the face and produce five face orientation locations for the mouth, nose, and eyes.

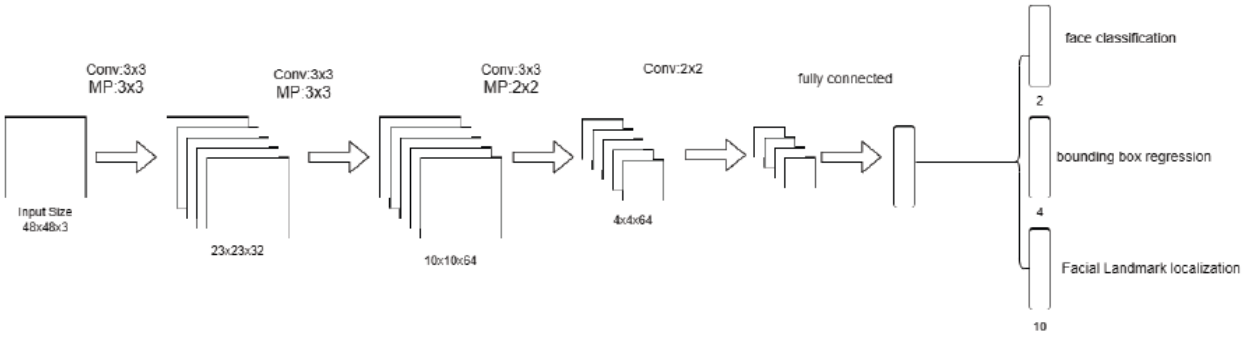

Figure 3. Output Network (O-Net)

Three tasks of MTCNN are performed by the network: face/non-face classification, face orientation, and bounding box regression.

1. Face classification: cross-entropy loss is used to solve this binary classification problem:

$$Li^{det}=-(yi^{det}\log(Pi)+(1-yi^{det})(1-\log(pi))) \quad 1$$

2. Bounding box regression: The learning objective is a regression problem. For each candidate window, the offset between the candidate and the nearest ground truth is calculated. Euclidean loss is employed for this task:

$$Li^{box}=||y_i^{box}-y_i^{box}||^2 \quad 2$$

3. Localizing face landmarks: This task is expressed as a regression problem with Euclidean distance serving as the loss function.

$$L_i^{landmark}=||y_i^{landmark}-y_i^{landmark}|| \quad 3$$

The MTCNN model creates a bounding box when it detects the face, extracts the facial region, and then shows the facial landmarks in the face such as mouth, eyes, and nose. After detecting the person, the model tries to classify whether the person is authorized or not. If the person is authorized the bounding boxes turn green else turn to red and give an alerting sound. The MTCNN model also works offline where it detects the faces from the camera feed directly without requiring an internet connection or online services.

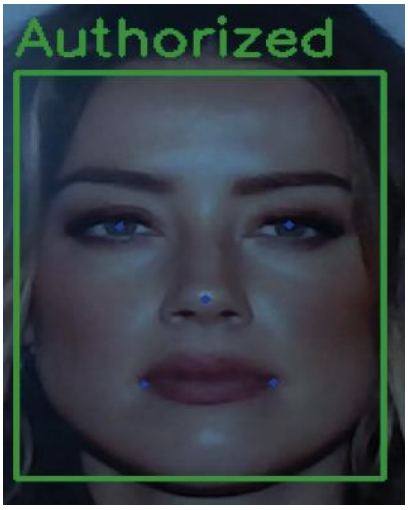

Figure 4. Authorized Image with facial landmarks

2.2. Deepface

The VGG-facial model, a facial recognition-specific variation of the VGG (Visual Geometry Group) architecture, is used by the Deepface method. The 16-layer deep convolutional neural network known as the VGG-16 architecture serves as the foundation for the VGG-Face model.

2.2.1 VGG architecture

There are 16 layers in the VGG-16 architecture: 3 fully connected layers and 13 convolutional layers with 3x3 filters. ReLU activation functions come after each convolutional layer, and max-pooling layers are employed to minimize spatial dimensions. ReLU activation is used in the fully linked layers, while SoftMax is used in the output layer for classification. VGG-16 is usually pre-trained using the ImageNet dataset and includes more than 138 million parameters. A VGG-16 model designed specifically for face recognition is called the VGG-Face model. It can extract characteristics unique to faces since it has been pre-trained on a sizable collection of face photos. The VGG-Face output layer generates a 2622-dimensional feature vector as opposed to class probabilities, in contrast to regular VGG-16. For computing similarity in face recognition tasks, this feature vector is essential. The VGG-Face model functions as a feature extractor in Deepface. A facial picture is sent through a series of layers for processing, and from the final fully linked layer, a high-dimensional feature vector is extracted. Using deep convolutional layers and pre-trained weights, Deepface uses this modified VGG-16 architecture to achieve great accuracy and resilience in face verification and identification.

3. Results and Discussions

MTCNN (Multi-task Cascaded Convolutional Networks) offers an impressive 96.2% accuracy and is particularly known for its accuracy in face detection and recognition. Designed for real-time applications, MTCNN processes

images quickly, making it suitable for applications such as video surveillance and video chat. Its versatility goes beyond just facial recognition; it's also great for aligning and contouring the face, making it a versatile tool for facial care. In addition, MTCNN shows strong performance under different conditions, including different illumination and partial occlusion. Deepface, on the other hand, offers a lightweight option with a slightly lower accuracy at 95 percent. Its design prioritizes computational efficiency, making it less resource-intensive than MTCNN. Deepface is not just facial recognition; it also analyzes facial features such as eyes, ears, and nose. Therefore, it is suitable for applications where the understanding of facial features is crucial. Despite its simplicity, Deepface uses deep learning techniques to extract and compare facial features, enabling reliable face matching.

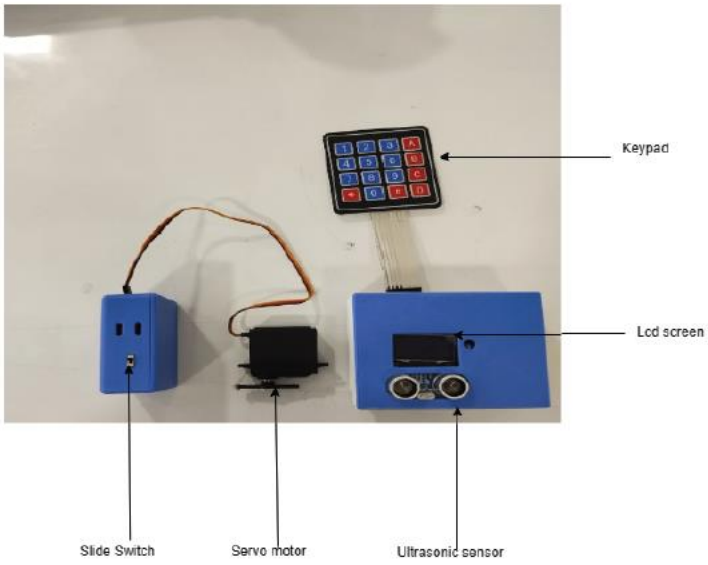

Figure 5. Image of prototype

5. Conclusions

An important development in access control technology is the combination of sophisticated algorithms and useful security features through the integration of an optimized algorithm with a motorized door mechanism. The advanced algorithm of this system effectively detects authorized users, minimizing false positives and negatives. The motor is signaled by the algorithm to open the door when an authorized individual is spotted, allowing for easy entry. On the other hand, an audible buzzer that functions as a real-time alarm and deterrent is activated in response to unauthorized entry attempts. By combining motorized control, auditory feedback for Unauthorized persons, and algorithmic accuracy, this door security solution is both reliable and easy to use, raising the bar for access control systems.

Future enhancements to the proposed face recognition system could focus on integrating edge computing to enable faster, real-time processing without reliance on cloud infrastructure. Incorporating anti-spoofing techniques, such as liveness detection or 3D face modeling, would improve system security against presentation attacks. Expanding and diversifying the training data set to include variations in age, ethnicity, lighting, and occlusion would help

increase the model's robustness. Additionally, the system could be extended to support multimodal biometric authentication by combining facial recognition with voice or gait analysis. Hardware optimization for embedded platforms like Raspberry Pi or Jetson Nano could also support energy-efficient deployment. Lastly, adopting privacy-preserving machine learning techniques, such as federated learning, would enhance user data protection and regulatory compliance.

Acknowledgment

The optimized algorithm used for face recognition is Deepface and MTCNN.

Declaration of Competing Interest and Ethics

The author declares no conflict of interest. This research study complies with research publishing ethics. The scientific and legal responsibility for manuscripts published in OPS Journal belongs to the authors.

References

Aswini, D., Rohindh, R., Manoj Ragavendhara, K.S. & Mridula, C.S. (2021, October). Smart door locking system. *2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA)*, Coimbatore, India, 2021, https://doi.org/10.1109/ICAECA52838.2021.9675590

Dhope, V., Chavan, A., Hadmode, N. & Godase, V. (2024). smart plant monitoring system. *International Journal of Creative Research Thoughts (IJCRT), 12*(5), b844–b849. https://doi.org/10.5281/zenodo.11213525

Djupsjö, K. & Almosawi, M. (2018). *IoT security applied on a smart door lock application*. KTH Royal Institute of Technology. Retrieved from https://www.diva-portal.org/smash/get/diva2:1216681/FULLTEXT01.pdf

Essien, U., & Ansa, G. (2023). A deep learning-based face recognition attendance system. *Global Journal of Engineering and Technology Advances, 17*(1), 9-22. https://doi.org/10.30574/gjeta.2023.17.1.0165

Godase, V., Jadhav, Y., Vishal, K., Metkari, V. & Gangonda, S. (2024b). IOT based greenhouse monitoring and controlling system. *International Journal for Scientific Research and Development, 12*(3), 138-140.

Godase, V., Lawande, A., Mane, K., Davad, K. & Gangonda, S. (2024a). Pipeline survey robot. *International Journal for Scientific Research and Development 12*(3), 141-144.

Gupta, K., Jiwani, N., Uddin Sharif, M.H., Mohammed, M.A. & Afreen, N. (2022, November). Smart door locking system using IoT. *2022 International Conference on Advances in Computing, Communication and Materials (ICACCM)*, Dehradun, India, https://doi.org/10.1109/ICACCM56405.2022.10009534

Mehmood, M. Q., Malik, M. S., Zulfiqar, M. H., Khan, M. A., Zubair, M., & Massoud, Y. (2023). Invisible touch sensors-based smart and disposable door locking system for security applications, *Heliyon, 9*(2), e13586, https://doi.org/10.1016/j.heliyon.2023.e13586

Nayana, R & Shashidhar, R. (2019). Smart door lock system. *International Journal for Modern Trends in Science and Technology*, *5*(2). 36-38.

Patil, A. V., Prakash, S., Akshay, S., Mahadevaswamy, Patgar, C. & Sharath Kumar, A. J. (2018), Android based smart door locking system, *International Journal of Engineering Research & Technology (IJERT) NCESC*, 6(13), 1-4. https://doi.org/10.17577/IJERTCONV6IS13199

Peña, F., Eirea, M., Tanco, M. M., & Andrade, F. (2023, October). Evaluation of face recognition systems for RoboCup@Home service robots. *2023 Latin American Robotics Symposium (LARS), 2023 Brazilian Symposium on Robotics (SBR), and 2023 Workshop on Robotics in Education (WRE)*. Salvador, Brazil, IEEE. https://doi.org/10.1109/LARS/SBR/WRE59448.2023.10333025