

## Analysis of FaceNet and VGG16 for Blind Face Recognition with MTCNN and HaarCascade Detection Methods

Agung Budi Prastyo<sup>1✉</sup>, Arief Setyanto<sup>2</sup>

<sup>1,2</sup> Informatics, PJJ MTI, Universitas Amikom Yogyakarta, Indonesia

✉ Corresponding Author : Agung Budi Prastyo (e-mail: [agungbp@students.amikom.ac.id](mailto:agungbp@students.amikom.ac.id))

### Article Information

#### Article History

Received : February 2, 2025

Revised : February 5, 2025

Published : April 02, 2025

#### Keywords:

*Face Recognition;*

*FaceNet;*

*HaarCascade;*

*MTCNN;*

*VGG16.*

### ABSTRACT

Face recognition is a rapidly growing biometric technology, especially with the application of Convolutional Neural Networks (CNN) such as FaceNet and VGG16. This research aims to evaluate the effectiveness of both CNN models in recognizing the faces of visually impaired people, who face the challenge of limited vision in image retrieval. The research uses two face detection methods, namely MTCNN and HaarCascade, to analyze the effect of face detection on recognition accuracy. The experimental method was conducted by collecting facial data of visually impaired people under various lighting conditions and expressions. The results show that accurate face detection greatly affects the performance of face recognition models, with MTCNN providing better face detection results (93.75% detection accuracy) than HaarCascade (83.75% detection accuracy). Both models, FaceNet and VGG16, show excellent face recognition accuracy (100%) if the face image is correctly detected by MTCNN. Therefore, for face recognition of visually impaired people, it is recommended to use MTCNN as the face detection method, followed by FaceNet or VGG16 for face recognition.

## INTRODUCTION

Face recognition is one of the biometric technologies that has experienced rapid development in recent years, especially with advances in deep learning techniques. In the context of face recognition technology, methods based on Convolutional Neural Networks (CNN) such as FaceNet and VGG16 have proven effective in overcoming major challenges, including lighting variations, facial expressions, and misaligned facial positions (Ma & Long, 2023) (Saragih & To, 2022). CNN methods utilize the ability of neural networks to process facial images to extract unique facial features, which are then used for identification or verification of a person's identity (Sikha & Bharath, 2022).

One of the main challenges in face recognition is face identification of visually impaired people. Visually impaired people have limited visual access to their own faces, which can affect the quality of images captured for face identification. Therefore, it is important to develop face recognition methods that are not only accurate but can also handle the vagueness that often occurs in capturing face images under certain conditions (Shin et al., 2024) (Goel et al., 2021).

In this study, the authors propose the use of FaceNet and VGG16 for face recognition of visually impaired people. FaceNet, which is known for its excellent ability in mapping face images into Euclidean space, enables more efficient face comparison by using triplet loss (Ardiawan & Negarara, 2024). Meanwhile, VGG16, which is one of the most widely used CNN models, has a good ability to identify facial features with high accuracy, even on images with low quality or physical obstructions such as masks (Bewoor et al., 2023) (Firdaus & Munir, 2022).



This is an open access article under the [CC BY](#) license

To detect and verify faces in this face recognition, we use two popular face detection methods: MTCNN and HaarCascade. MTCNN (Multi-task Cascaded Convolutional Networks) is a very effective model in detecting faces with various positions and lighting conditions (Sunarko et al., 2023). While HaarCascade, although simpler, remains a good choice for face detection under more constrained conditions and on devices with limited resources (Sikha & Bharath, 2022).

This research aims to evaluate and analyze the ability of FaceNet and VGG16 to recognize faces of the visually impaired, taking into account the face detection performed by MTCNN and HaarCascade. Through this research, it is hoped that a more efficient and accurate solution can be found for face recognition in more challenging conditions, especially for the visually impaired.

## RESEARCH METHODS

This study aims to analyze the effectiveness of two Convolutional Neural Network (CNN) architectures, namely FaceNet and VGG16, in recognizing faces of visually impaired people. To detect faces in various conditions, we use two popular face detection methods: MTCNN (Multi-task Cascaded Convolutional Networks) and HaarCascade. These methods are applied in two main stages: face detection and face recognition:

### Research Design

This research uses an experimental approach with a face recognition system for blind people. In this research, the first stage is face detection using MTCNN or HaarCascade models. After the face is detected, the second stage is face recognition using two CNN models, FaceNet and VGG16, which are analyzed to evaluate their accuracy and effectiveness in recognizing faces.

### Data and Samples

The data used in this study are facial images taken under varied conditions, including various lighting, facial expressions, and head positions. The dataset used consists of facial images of visually impaired people taken under various conditions to ensure that the model used can work in various real-world scenarios (Asmara et al., 2022).

### Data Collection Technique

The data used in this study was collected by downloading face images from websites that provide open datasets. Datasets were selected to match the relevance to the object of research. Selecting facial data that is diverse in terms of lighting, facial expressions, and head positions that can represent the object of research (Reddy et al., 2023) (Mensah et al., 2024).

### Face Detection Method

Face detection is the first stage in the face recognition process. Two methods used for face detection are:

- **MTCNN:** MTCNN is used to detect faces with robustness to position and lighting variations. This model consists of three stages: Proposal Network (P-Net), Refinement Network (R-Net), and Output Network (O-Net), which enables accurate face detection in images (Asmara et al., 2022).
- **HaarCascade:** HaarCascade is a simpler yet effective face detection method, especially on devices with limited resources. This method is used to detect faces in images with various conditions (Susanto et al., 2024).

### Model Implementation

In this study, two main CNN models were used for face recognition, namely FaceNet and VGG16:

- **FaceNet:** This model uses triplet loss to ensure that similar faces have closer vector distances compared to different faces. FaceNet maps faces to Euclidean space, which allows measuring the distance between facial features for identification (Sukash et al., 2024).

- **VGG16:** This model is used to identify facial features in images with high accuracy. VGG16 has a deep architecture and is widely used for face recognition tasks under various conditions (Chethana et al., 2024).

### Training and Testing Process

After face detection and model application, the detected face images are processed using the CNN model:

- **Model Training:** The detected face dataset is used to train FaceNet and VGG16 models. These models are trained to identify facial features and recognize facial features.
- **Model Testing:** Testing is done using datasets that have never been seen before by the model. This test aims to evaluate the model's ability to recognize the faces of blind people under varying conditions (Pandit et al., 2023).

### Performance Evaluation

Model performance is evaluated using metrics such as accuracy, precision, recall, and F1-score. This evaluation is done to measure how well the model can recognize faces under various conditions, including poor lighting, different facial expressions, misaligned head positions, and especially in the presence of eye feature noise (Sukash et al., 2024).

## RESULTS AND DISCUSSION

This research aims to evaluate the effectiveness of FaceNet and VGG16 in recognizing the faces of blind/visually impaired people, by assessing how both models are affected by the face detection methods used, namely MTCNN and HaarCascade. Based on the experimental results, it was found that face recognition is greatly affected by the quality of face detection, which in turn is affected by factors such as lighting, face position, and facial expressions in the dataset.

### Effect of Face Detection on Face Recognition

The experimental results show that the success of FaceNet and VGG16 in recognizing faces is greatly affected by the success of the face detection process. FaceNet, which functions to generate face embedding and measure face similarity by using triplet loss, shows excellent performance in face recognition, but only when faces are well detected. Once the face fails to be detected, FaceNet cannot function optimally. In this study, MTCNN was shown to provide better face detection results than HaarCascade, especially in images with poor lighting conditions, misaligned head positions, and varied facial expressions. This shows that the face detection process plays an important role in successful face recognition (Balde et al., 2020).

### Face Detection Results using MTCNN and HaarCascade

Table 1 shows that out of a total of 80 tested face data, 13 face data were poorly detected by HaarCascade, which caused the face recognition model to be unable to recognize the faces. In contrast, MTCNN only failed to detect 5 faces out of the total 80 faces, which shows that MTCNN is more effective in handling variations in facial position, lighting, and expression. These results show the importance of face detection quality in determining the success of face recognition, where FaceNet and VGG16 function better when faces are accurately detected (Chethana et al., 2024) (Susanto et al., 2024).

**Table 1.** Face Detection Results using MTCNN and HaarCascade

Face Detection Method	Total Image	Detected	Not Detected	Detection Percentage
MTCNN	80	75	5	93.75%
HaarCascade	80	67	13	83.75%

### Effect of Face Detection Process on the Use of FaceNet and VGG16 Models

FaceNet and VGG16 rely on face detection results to perform recognition. FaceNet, which uses face representation in the form of vector embedding, requires very precise face detection for

the resulting vectors to correctly represent faces. Once the face detection fails, as happened with some data that was not detected by HaarCascade, the face recognition result will also fail. VGG16, despite its strong ability to recognize facial features, is also affected by the quality of face detection provided by the detection system. This shows that although both models have good recognition capabilities, they are highly dependent on the input quality of the face detection process (Mensah et al., 2024) (Chethana et al., 2024).



**Figure 1.** Face Detected Image Data by MTCNN



**Figure 2.** Image Data with No Face Detected by MTCNN



**Figure 3.** Face Detected Image Data by HaarCascade



**Figure 4.** Image Data with No Face Detected by HaarCascade

## Performance Evaluation of Using Models for Face Recognition

After the face detection process, face recognition is performed using KNN. The evaluation shows that FaceNet gives better results than VGG16, but only when faces are successfully detected properly. MTCNN ensures more well-detected faces, which improves face recognition performance. In contrast, HaarCascade resulted in more poorly detected faces, which negatively impacted face recognition using both models (Pandit et al., 2023).

**Table 2.** FaceNet & VGG16 Model Evaluation Results Table

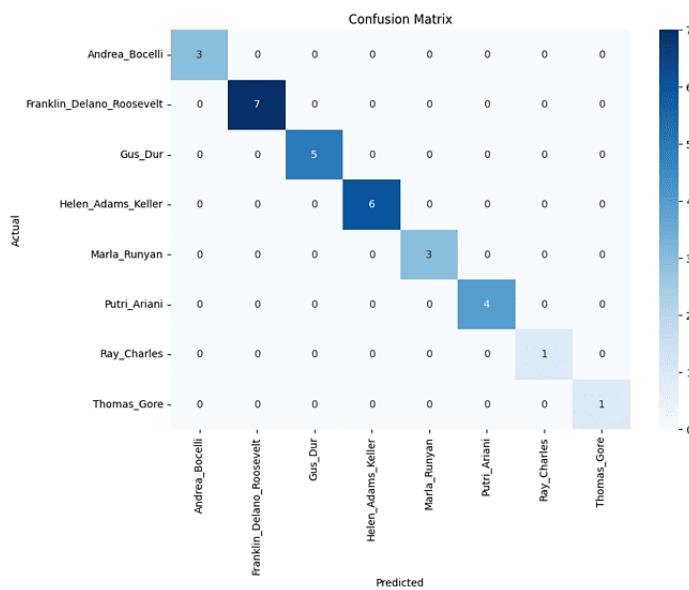
Face Recognition Model	Face Image	Accuracy
FaceNet + MTCNN	75	100%
FaceNet + HaarCascade	67	100%
VGG16 + MTCNN	75	100%
VGG16 + HaarCascade	67	100%

The face recognition process is greatly influenced by the quality of face detection performed previously. The results show that MTCNN performs better than HaarCascade in

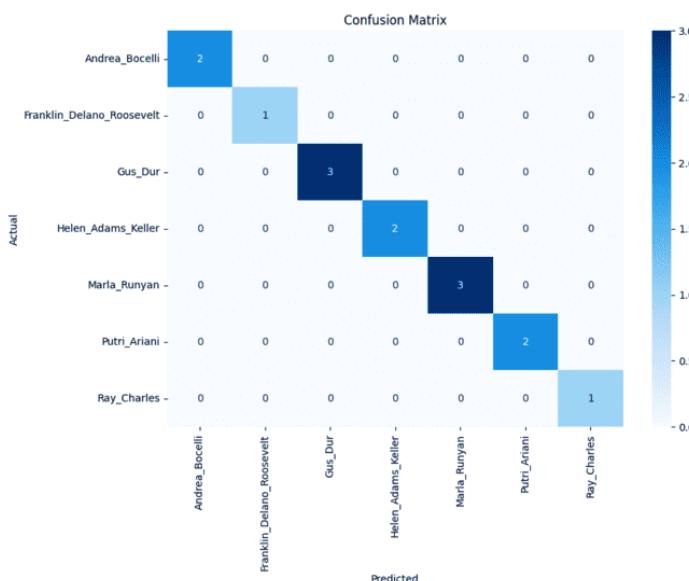
detecting faces, especially when the images have lighting variations, misaligned head positions, or changing facial expressions. In the experiments conducted, HaarCascade was unable to detect 13 faces out of 80 faces tested, while MTCNN only failed to detect 5 faces out of 80 faces. This indicates that an accurate face detection process is crucial for successful face recognition using models such as FaceNet and VGG16. The results of the study are shown in table 3. Accuracy was obtained with a score of 100% for both FaceNet and VGG16 models. Table 3 is reinforced by the results of model evaluation with KNN using google colab shown in Figures 5, 6, 7, and 8.

**Table 3.** Confusion Matrix Table of FaceNet & VGG16 Model Evaluation

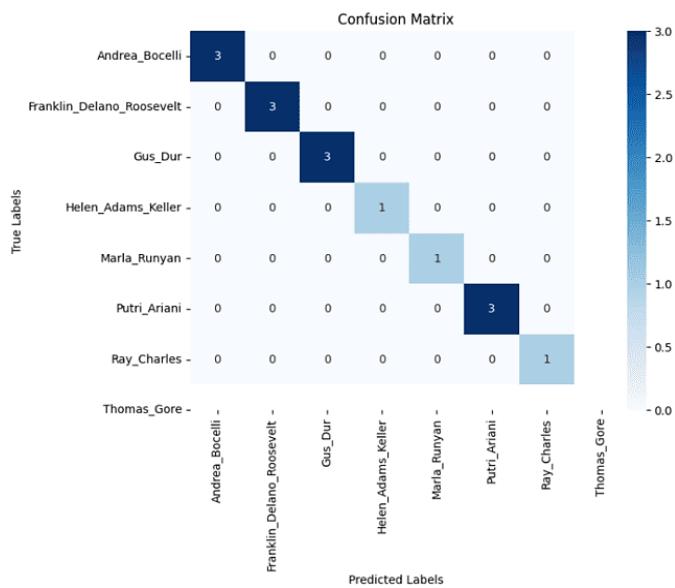
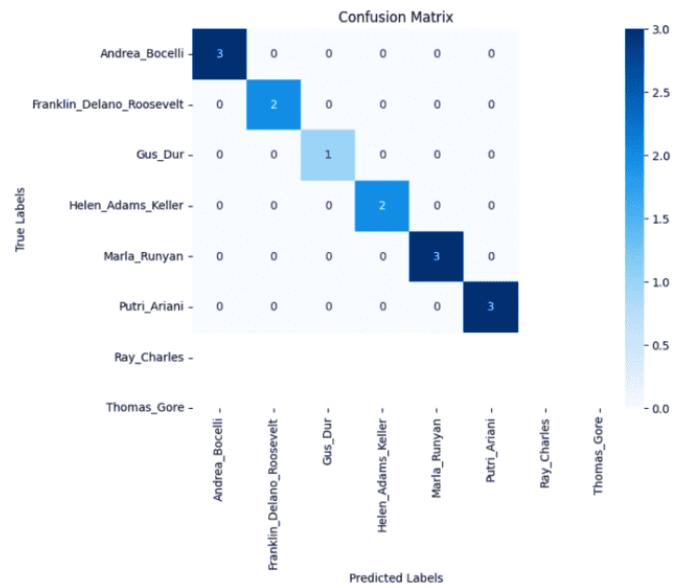
Face Detection Method	Face Recognition Model	Accuracy	Precision	Recall	F1-Score
MTCNN	FaceNet	100%	100%	100%	100%
	VGG16	100%	100%	100%	100%
HaarCascade	FaceNet	100%	100%	100%	100%
	VGG16	100%	100%	100%	100%



**Figure 5.** Confusion Matrix of VGG-MTCNN



**Figure 6.** VGG-HaarCascade Confusion Matrix

**Figure 7.** Confusion Matrix of FaceNet-MTCNN**Figure 8.** Confusion Matrix of FaceNet-HaarCascade

## CONCLUSION

The results show that accurate face detection greatly affects the performance of face recognition models, with MTCNN providing better face detection results (93.75% detection accuracy) than HaarCascade (83.75% detection accuracy). Both models, FaceNet and VGG16, show excellent face recognition accuracy (100%) if the face image is correctly detected by MTCNN.

Based on the results of this experiment, it can be concluded that the face recognition process using FaceNet and VGG16 is greatly affected by the quality of face detection. MTCNN shows an advantage in detecting faces more accurately, even in less than ideal image conditions, compared to HaarCascade. Therefore, face recognition will be more effective if using MTCNN for face detection, which is then followed by using FaceNet or VGG16 for the face recognition process.

Future research could explore more advanced face detection models, such as RetinaFace or BlazeFace, to compare with MTCNN in terms of accuracy and processing time efficiency. In addition, studies can examine the effect of image resolution and quality on face detection performance as well as examine how variations in pose and facial expression affect recognition accuracy. A hybrid approach by combining multiple face detection methods can also be evaluated to improve the reliability of the system.

## ACKNOWLEDGEMENTS

We would like to the Amikom University Yogyakarta community for their knowledge and guidance.

## REFERENCES

- Asmara, R. A., Sayudha, B., Mentari, M., Budiman, R. P. P., Handayani, A. N., Ridwan, M., & Arhandi, P. P. (2022). Face Recognition Using ArcFace and FaceNet in Google Cloud Platform For Attendance System Mobile Application. In *Proceedings of the 2022 Annual Technology, Applied Science and Engineering Conference (ATASEC 2022)* (pp. 134–144). Atlantis Press International BV. [https://doi.org/10.2991/978-94-6463-106-7\\_13](https://doi.org/10.2991/978-94-6463-106-7_13)
- Ardiawan, M. I., & Negarara, G. P. K. (2024). Comparative Analysis of FaceNet, VGGFace, and GhostFaceNets Face Recognition Algorithms For Potential Criminal Suspect Identification. *Journal of Applied Artificial Intelligence*, 5(2), 34–49. <https://doi.org/10.48185/jaai.v5i2.1237>
- Balde, A. A., Jain, A., & Patra, D. (2020, September 28). Improving Facial Recognition of FaceNet in a small dataset using DeepFakes. *4th International Conference on Computer, Communication and Signal Processing, ICCCSP 2020*. <https://doi.org/10.1109/ICCCSP49186.2020.9315253>
- Bewoor, M., Patil, S., Kushwaha, S., Tandon, S., Trivedi, S., & Pawar, A. (2023). Face recognition using open CV and VGG 16 transfer learning. *AIP Conference Proceedings*, 2890(1). <https://doi.org/10.1063/5.0157084>
- Chethana, Nagavi, T. C., Mahesha, Ravi, V., & Gururaj. (2024). FaceNet – A Framework for Age Variation Facial Digital Images. *ICST Transactions on Scalable Information Systems*, 11. <https://doi.org/10.4108/eetsis.5198>
- Firdaus, F., & Munir, R. (2022). Masked Face Recognition using Deep Learning based on Unmasked Area. *ICPC2T 2022 - 2nd International Conference on Power, Control and Computing Technologies, Proceedings*. <https://doi.org/10.1109/ICPC2T53885.2022.9776651>
- Goel, R., Mehmood, I., & Ugail, H. (2021). A study of deep learning-based face recognition models for sibling identification. *Sensors*, 21(15). <https://doi.org/10.3390/s21155068>
- Ma, L., & Long, Z. (2023). A Face Recognition Method Using ResNet34 and RetinaFace. *Academic Journal of Computing & Information Science*, 6(10). <https://doi.org/10.25236/ajcis.2023.061003>
- Mensah, J. A., Appati, J. K., Boateng, E. K. A., Ocran, E., & Asiedu, L. (2024). FaceNet recognition algorithm subject to multiple constraints: Assessment of the performance. *Scientific African*, 23. <https://doi.org/10.1016/j.sciaf.2023.e02007>
- Pandit, A., Nikalje, R., Vishwakarma, N., Vishwasrao, V., & Khose, Prof. T. (2023). Face Authentication using MTCNN and FaceNet. *International Journal for Research in Applied Science and Engineering Technology*, 11(11), 1140–1143. <https://doi.org/10.22214/ijraset.2023.56679>
- Reddy, K. S. S., Ramesh, G., Praveen, J., Surekha, P., & Sharma, A. (2023). A Real-time Automated System for Object Detection and Facial Recognition. *E3S Web of Conferences*, 430. <https://doi.org/10.1051/e3sconf/202343001076>
- Saragih, R. E., & To, Q. H. (2022). A Survey of Face Recognition Based on Convolutional Neural Network. In *Indonesian Journal of Information Systems (IJIS) (Vol. 4, Issue 2)*.

- Shin, D.-K., Kim, T.-H., Seo, D.-J., Kim, H., Cha, J.-G., & Kim, S.-K. (2024). Investigating Digital Illiterate Classification Techniques Based on DeepFace Technology. *IJASEIT*, 14(5), 1496–1503.
- Sikha, O. K., & Bharath, B. (2022). VGG16-random fourier hybrid model for masked face recognition. *Soft Computing*, 26(22), 12795–12810. <https://doi.org/10.1007/s00500-022-07289-0>
- Sukash, Sakthi, Sudharsan, Vignesh, & Tamil. (2024). Real-time Face Recognition Using FaceNet and Haar Cascade for Door Monitoring. *International Journal for Multidisciplinary Research (IJFMR)*, 6(2), 1–7. [www.ijfmr.com](http://www.ijfmr.com)
- Sunarko, H., Hidayat, R., & Hartanto, R. (2023, July 26). Comparative Analysis of Masked and Unmasked for Face Recognition Using VGG Face and MTCNN. *Proceedings of the 5th International Conference on Applied Engineering, ICAE 2022, 5 October 2022, Batam, Indonesia*. <https://doi.org/10.4108/eai.5-10-2022.2327473>
- Susanto, B. M., Surateno, Atmadji, E. S. J., Pramulintang, A. H., Apriliano, G., Wulansari, T., & Gumilang, M. A. (2024). Face recognition using haar cascade classifier and FaceNet (A case study: Student attendance system). *International Journal of Informatics and Communication Technology*, 13(2), 272–284. <https://doi.org/10.11591/ijict.v13i2.pp272-284>