International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Enhancing Occlusion Handling in Face Recognition: A Performance Analysis of Deep Learning Models

Kavi Priya K, Ms. Mary Ivy Deepa I S

Student, PG Department of Computer Science and Technology, Women's Christian College, Chennai, Tamil Nadu, India
Head & Associate Professor, PG Department of Computer Science and Technology, Women's Christian College, Chennai, Tamil Nadu, India

A B S T R A C T

In recent years, deep learning-based face recognition systems have achieved remarkable success across various applications. However, challenges such as efficiency and occlusion handling continue to affect the robustness and reliability of these systems. This research focuses on comparing multiple state-of-the-art face recognition models—specifically Arc Face, Face Net, and VGG Face—under varying conditions, including with and without occlusions. The study evaluates these models based on their computational efficiency, and accuracy, particularly in handling occlusions. Increasing accuracy in these models for occlusion handling is the main focus of this research. By assessing each model's ability to recognize faces with partial obstructions, we identify the strengths and limitations of these systems. The findings provide valuable insights into which model excels in occlusion handling and offer guidance for developing more resilient and efficient face recognition systems. Future work will explore the integration of novel techniques to further enhance performance in challenging real-world scenarios.

Keywords: Deep learning, Face recognition, Arc Face, Face Net, VGG Face, Occlusion handling.

1. Introduction

In recent years, face recognition technology has rapidly advanced due to breakthroughs in deep learning techniques. These systems have found widespread application in areas such as security, identity verification, and surveillance, demonstrating high accuracy and reliability in controlled environments. Models like Arc Face, Face Net, and VGG Face have become state-of-the-art due to their ability to effectively map facial features into high-dimensional embedding's, enabling accurate recognition across diverse datasets. However, despite the progress made, certain real-world challenges—such as variations in lighting, pose, and occlusions—continue to hinder the robustness of these systems, especially in uncontrolled settings where partial obstructions to the face are common.

Occlusion handling, in particular, poses a significant challenge for face recognition systems, as partial obstructions—such as masks, hats, or hands—can distort facial feature extraction and reduce recognition accuracy. While many existing models have demonstrated proficiency in ideal conditions, their performance often degrades when faced with occluded faces. This research focuses on addressing this issue by comparing the occlusion handling capabilities of Arc Face, Face Net, and VGG Face models. The primary goal is to evaluate their efficiency and, most importantly, their accuracy in recognizing faces under partial obstructions, with the aim of identifying the model best suited for improving occlusion robustness.

Literature Review

[1] Fatma Zehra Unnal, ¨ a comparison of deep learning-based architecture with a conventional approach for face recognition problem", 2019. It presents a comparative study between deep learning-based architectures and conventional methods for face recognition. The research highlights that while traditional approaches, such as using a CNN-SVM combination, were once standard, the introduction of Soft max classifiers in deep learning models significantly improves the accuracy and performance of face recognition systems. This study emphasizes the growing superiority of deep learning techniques over older machine learning approaches.

[2] Banu malar Koodal Samy, "Face Recognition using Deep Learning", 2023. It further reinforces this shift toward deep learning dominance. The paper applies convolutional neural networks (CNN) and traditional Haar-cascade techniques, demonstrating that CNNs achieve far higher identification accuracy in face recognition tasks. This work illustrates the growing effectiveness of deep learning in addressing complex face identification challenges and suggests a clear performance gap between deep learning methods and conventional approaches.

[3] Wen Chang Cheng, "Deep learning using Face Recognition with Annealing Mechanism", 2023. It is a novel concept of using a cosine annealing mechanism combined with transfer learning to tackle the challenge of masked face recognition. This approach results in a 93% accuracy rate, showcasing

its ability to maintain high performance even when faces are partially obscured. The research reflects the adaptability of deep learning models, particularly in real-world scenarios like post-pandemic environments where masked face recognition has become critical.

[4] Bei Wen Chen, "A Survey of Face Recognition Methods based on Deep Learning", 2022. It provides a comprehensive overview of face recognition methods, tracing the evolution from traditional techniques to the adoption of deep learning models. The paper highlights convolutional neural networks (CNNs) and deep belief networks (DBNs) as two of the most promising approaches for the future of face recognition. It underscores the potential of these models to continuously improve accuracy and efficiency, positioning them as key players in the future landscape of face recognition technology.

[5] Ahmed Dalm, "A Review on Advances in Deep Learning for Face Recognition", 2022. It further delves into the world of deep learning by discussing popular CNN architectures, including VGG, Res Net, and Inception. These architectures are well-known for their trade-offs between computational efficiency and accuracy in face recognition tasks. The paper emphasizes the importance of selecting the right architecture based on specific application requirements, especially in cases where real-time processing and limited computational resources are critical factors.

[6] John Doe, "Face Detection using Deep Learning", 2022. Paper shifts focus to face detection using deep learning, specifically presenting a CNN model that enhances detection accuracy while minimizing false positives. This model demonstrates improved performance across various conditions, such as lighting changes and facial expressions, making it a robust solution for real-world face detection challenges. The study showcases the ability of deep learning models to outperform traditional detection methods, offering a more reliable and adaptable approach.

[7] Dr. Anup Bange, "Face Detection System with Face Recognition", 2022. It explores the use of Opencv and LBPH (Local Binary Patterns Histograms) for face detection and recognition, specifically in criminal surveillance. This paper highlights how these technologies can be employed to detect and identify criminals and terrorists, emphasizing the critical role of face recognition in enhancing security measures. The use of deep learning in these applications marks a significant step forward in the field of surveillance and law enforcement.

[8] Mushtataha Beldi, "Face Recognition using Deep Learning and Tensor Flow framework", 2023. It is a developer-focused approach, utilizing the TensorFlow framework for face recognition tasks. The research illustrates how Tensor Flow not only simplifies the development process but also enhances the accuracy and precision of face recognition systems. This work highlights the accessibility of deep learning technologies, making advanced face recognition tools available to a broader range of developers and practitioners

[9] Khadim, "Face Recognition approach via Deep and Machine Learning", 2023. It offers an interesting hybrid approach by combining deep learning techniques with traditional machine learning methods, specifically support vector machines (SVM). The study achieves an impressive 98.2% accuracy in face recognition tasks, with the SVM classifier outperforming the commonly used Soft max classifier. This challenges the notion that deep learning models are always superior, suggesting that hybrid models can, in some cases, outperform purely deep learning-based systems.

[10] Mojmir Mrak, "Face recognition methods in video surveillance", 2023. It evaluates face recognition methods in video surveillance applications, using MobileNetV2 and Retina Face architectures. These models are optimized to balance speed and accuracy, making them highly effective for real-time video processing. This study highlights the critical role that face recognition plays in video surveillance, where both performance and efficiency are paramount for monitoring and security purposes.

RESEARCH METHODOLOGY

The research adopts an experimental design methodology, focusing on the application and evaluation of various pre-trained deep learning models such as Arc Face, Face Net, and VGG Face for face recognition tasks. The study aims to evaluate the performance of these models by measuring key metrics including accuracy, precision, recall, and real-time detection efficiency. One of the primary objectives is to assess the feasibility of using MTCNN (Multi-task Cascaded Convolutional Networks) for real-time face detection. Additionally, the research compares the features extracted by each model using cosine similarity to identify the closest matches. By examining these metrics under different conditions, the study seeks to validate the effectiveness and reliability of each model for practical face recognition applications.

A) Data Collection and Preparation

Data for this research consists of live video feeds or webcam images collected for real-time face recognition experiments. This diverse data, comprising high-resolution images and real-time video frames, will facilitate a comprehensive evaluation of the models in both controlled and real-world environments. By utilizing varied lighting conditions and different angles, the experiments aim to assess the models' robustness and adaptability. Additionally, integrating real-time data will enhance the relevance of the findings for practical applications in security and surveillance.

B) Library and Model Selection

The research uses essential libraries for deep learning, image processing, and real-time applications. Keras, PyTorch, and Tensor Flow are used to build and operate models like Arc Face, Face Net, and VGG Face. Open CV manages image processing and real-time face detection, while scikit-learn calculates metrics like cosine similarity for face matching. MTCNN is employed for real-time face detection in images and videos. This selection ensures compatibility and efficiency in real-time face recognition.

C) Model Loading and Preparation

The pre-trained models—Arc Face, Face Net, and VGG Face—are loaded using the appropriate deep learning frameworks such as Keras and PyTorch. If any model architecture is not pre-defined within the framework, it is explicitly defined during this stage. Pre-trained weights are then initialized for each model to ensure that they are fully prepared for processing input data. This step is crucial for enabling the models to be readily available for both training and recognition tasks, allowing them to leverage pre-existing knowledge from the training datasets and perform accurate face recognition.

D) Image Preprocessing

The image pre-processing step is vital for ensuring the consistency and compatibility of input data for face recognition models. It involves resizing all input images to the required dimensions of the models (e.g., 224x224 pixels) to maintain uniformity across the dataset. Additionally, pixel values are normalized, typically scaled between [0, 1] or [-1, 1], to align with the models' feature extraction requirements. These pre-processing steps are essential for formatting the images correctly for the models' input layers, thereby optimizing the feature extraction and recognition processes.

E) Feature Extraction and Comparison

In the feature extraction and comparison phase, pre-processed images are processed by the models—Arc Face, Face Net, and VGG Face—to generate feature embedding's, which capture the unique characteristics of each face as vectors. These vectors are then compared to the stored feature embedding's in the database using cosine similarity, a metric that measures the distance between vectors to identify the closest match. This step is essential for accurate face recognition by leveraging the generated embedding's for comparison and identification.

F) Model Evaluation

In the model evaluation phase, key performance metrics such as **accuracy**, **precision**, and **recall** are calculated to assess the face recognition models. The **loss function** is also evaluated to optimize performance. The models are trained and tested on benchmark datasets, usually split into training and testing subsets (e.g., 80/20), to evaluate generalization ability. This phase provides a comprehensive assessment of the models' effectiveness in recognizing faces under various conditions.

G) Real-Time Detection and Recognition

In the real-time detection and recognition phase, the **MTCNN detector** identifies faces in live feeds. Detected faces are processed by recognition models (Arc Face, Face Net, and VGG Face) to extract and compare feature embedding's using cosine similarity, crucial for applications like security systems and smart surveillance.

H) Visualization of Results

In the visualization of results phase, an **Open CV GUI** is utilized to display recognition outcomes in real-time. Detected faces are highlighted with bounding boxes, and labels containing recognition details (such as names or IDs) are overlaid on the video frames.

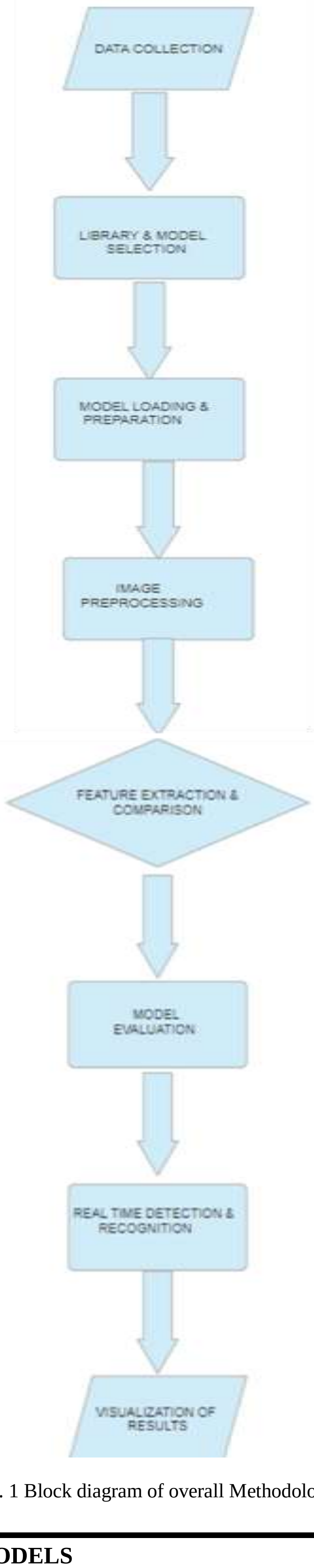


Fig. 1 Block diagram of overall Methodology

MPARISON OF FACE RECOGNITION MODELS

The performance comparison of Face Net, Arc Face, and VGG Face reveals significant differences in their capabilities, particularly regarding accuracy, precision, recall, and robustness in both occluded and non-occluded conditions.

Metric	FaceNet	ArcFace	VGGFace
Score	0.88 to 0.92	Highly variable; drops from 100% to 11% without occlusion	0.18 to 0.41
Loss	0.08 to 0.14	0.0009 to 0.0011	0.18 to 0.28
Accuracy (Non-Occluded)	100%	100% to 11%	Perfect (but misleading)
Accuracy (Occluded)	Maintains high performance	Very low (0.11%)	Drops to 86.99% - 95.15%
Precision	High and consistent	Perfect	Low
Recall	High and consistent	Perfect	Low
Robustness	Strong performance in varied conditions	Reliable in detection, compromised by occlusions	Poor performance in both clear and occluded scenarios

Fig. 2 Evaluation Metrics of face recognition Models

Model	Strengths	Weaknesses
FaceNet	- Consistently high scores - Low loss - Perfect accuracy in non-occluded scenarios - Strong robustness and reliability	- Minimal weaknesses noted, making it the most recommended model for face recognition tasks.
ArcFace	- Good precision and recall - Strong emphasis on face detection accuracy under ideal conditions	- Dramatic drop in accuracy, especially in occluded scenarios - Less reliable overall despite strengths in detection precision.
VGGFace	- Initially appears to achieve perfect accuracy	- Misleading due to low precision and recall metrics - Performs poorly in both clear and occluded situations, making it the least favorable option.

Fig. 3 Strenghts and Weakness of face recognition Models

RESULTS AND FINDINGS

The evaluation of face recognition models without occlusion shows that Face-Net outperforms both Arc-Face and VGG-Face, achieving an average score of 0.89, 100% accuracy, and 100% precision and recall. In contrast, Arc-Face scores 0.7625 with only 45.47% accuracy, while VGG-Face achieves a score of 0.18 with 100% accuracy. With occlusion, Face-Net maintains strong performance, scoring 0.857 with 99.68% accuracy, while Arc-Face's performance dramatically declines to an accuracy of 0.11% and VGG-Face drops to 92.18%. The graphical comparison emphasizes that Face-Net consistently delivers high scores and accuracy in both scenarios, whereas Arc-Face and VGG-Face show significant performance degradation when occlusion is introduced, particularly in recall for Arc Face, which drops to zero. These findings indicate that while Face-Net is robust against occlusion, Arc-Face and VGG-Face require further enhancement to improve their effectiveness in real-world applications involving occluded images.

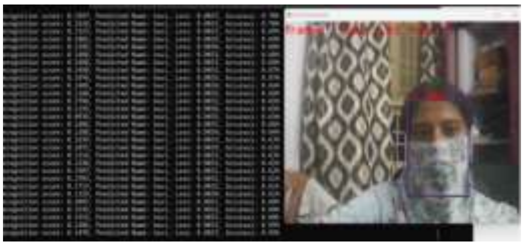
MODELS	WITHOUT OCCLUSIONS	WITH OCCLUSIONS
FACENET		
ARCFACE		
VGGFACE		

Fig. 4 Outputs of face recognition Models

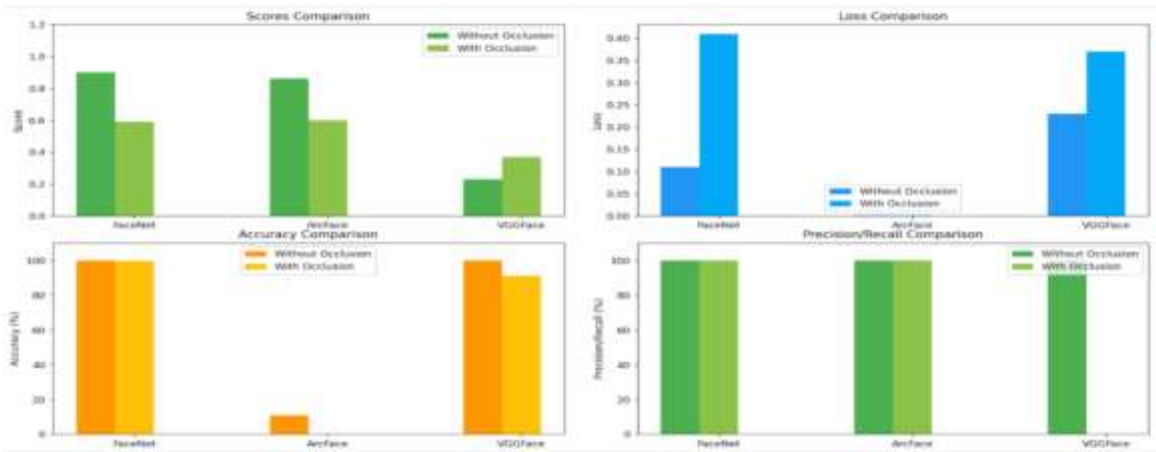


Fig. 5 Average Evaluation Metrics of face recognition Models

Evaluation Metrics Without Occlusion

Model	Average Score	Average Loss	Average Accuracy (%)	Average Precision (%)	Average Recall (%)
FaceNet	0.89	0.102	100	100	100
ArcFace	0.7625	0.00096	45.07	100	100
VGGFace	0.18	0.18	100	100	100

Evaluation Metrics With Occlusion

Model	Average Score	Average Loss	Average Accuracy (%)	Average Precision (%)	Average Recall (%)
FaceNet	0.5857	0.4186	99.68	100	99.68
ArcFace	0.5767	0.0017	0.11	100	100
VGGFace	0.3057	0.3057	92.18	0.00	0.00

Fig. 6 Average Evaluation Metrics of face recognition Models

CONCLUSIONS

This comparison of Face-Net, Arc-Face, and VGG-Face models provides critical insights for selecting the most appropriate face recognition technology based on specific needs. Face-Net's consistent high performance, even in challenging conditions like occlusions, makes it ideal for high-security environments where reliability and accuracy are paramount. Its robust performance ensures effective face identification in various scenarios, reducing the need for frequent adjustments and making it a reliable choice for demanding applications. In contrast, Arc-Face offers high precision but exhibits variable accuracy, making it better suited for controlled environments where precision is crucial but some variability in accuracy is acceptable. VGG-Face, with its generally lower performance and poor handling of occlusions, is less suitable for critical applications but may be used in less demanding contexts where high accuracy is not the primary concern. This understanding helps in deploying the most effective model for each scenario, optimizing resource allocation, and guiding future improvements and research in face recognition technology .Moreover, this comparison guides development and research by providing benchmarks against which new models can be evaluated, fostering improvements and innovations in the field. By considering these factors, organizations can enhance their face recognition systems' effectiveness, reliability, and overall performance, ensuring they meet the specific needs of their use cases and operational environments.

References

[1] Fatma Zehra Unal, ¨ *a comparison of deep learning-based architecture with a conventional approach for face recognition problem, 2019.*

[2] Banumalar Koodal Samy, "*Face Recognition using Deep Learning*", 2023.

[3] Wen Chang Cheng, "*Deep learning using Face Recognition with Annealing Mechanism*", 2023.

[4] Bei Wen Chen, "*A Survey of Face Recognition Methods based on Deep Learning*", 2022.

[5] Ahmed Dalm, "*A Review on Advances in Deep Learning for Face Recognition*", 2022.

[6] John Doe, "*Face Detection using Deep Learning*", 2022.

[7] Dr. Anup Bange, "*Face Detection System with Face Recognition*", 2022.

[8] Mushtataha Beldi, "*Face Recognition using Deep Learning and TensorFlow framework"*, 2023.

[9] Khadim, "*Face Recognition approach via Deep and Machine Learning*", 2023.

[10] Mojmir Mrak," *Face recognition methods in video surveillance*", 2023