Volume 11, Issue 4, April – 2026 International Journal of Innovative Science and Research Technology
ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26apr835

EfficientNetB0-Based Masked Face Recognition: A Robust Approach for Real-World Occlusion Scenarios

Girish Katkar[1]; Lekha Prajapati[2]; Ajay Ramteke[3]

[1]Associate Professor, Department of Computer Science, Taywade College Koradi, (M.S.), India.
[2]Research Scholar, Department of Computer Science, Taywade College Koradi, (M.S.), India.
[3]Assistant Professor, Department of Computer Science, Taywade College Koradi, (M.S.), India.

Publication Date: 2026/04/25

Abstract: The widespread use of face masks during and after the COVID-19 pandemic has posed significant challenges to traditional face recognition systems that rely on complete facial features. This paper presents an efficient and robust masked face recognition approach based solely on the EfficientNetB0 architecture, achieving an accuracy of 97% on the MFR2 dataset. The proposed method leverages transfer learning to utilize pre-trained ImageNet features, enabling effective extraction of discriminative information from the visible upper facial regions while maintaining computational efficiency. The model is trained and evaluated on the MFR2 benchmark, which contains real-world masked and unmasked face images across multiple identities. Experimental results demonstrate that EfficientNetB0 can deliver high recognition performance with a relatively small number of parameters, making it suitable for resource-constrained environments. The approach effectively addresses key challenges such as facial occlusion, loss of identity-specific features, and the need for robust mask-invariant representations. This work contributes to the development of practical and scalable masked face recognition systems for real-world applications, including security, surveillance, and access control.

Keywords: *Masked Face Recognition, EfficientNetB0, Deep Learning, Transfer Learning, MFR2 Dataset, Biometric Authentication.*

How to Cite: Girish Katkar; Lekha Prajapati; Ajay Ramteke (2026) EfficientNetB0-Based Masked Face Recognition: A Robust Approach for Real-World Occlusion Scenarios. *International Journal of Innovative Science and Research Technology,* 11(4), 1805-1812. https://doi.org/10.38124/ijisrt/26apr835

I. INTRODUCTION

Face recognition technology has witnessed unprecedented growth and integration into various aspects of modern life, ranging from mobile device security and airport immigration to large-scale surveillance and automated attendance systems. However, the global COVID-19 pandemic introduced a significant paradigm shift in how these systems operate. The widespread adoption of face masks as a public health measure severely hindered the performance of conventional face recognition algorithms. These systems, primarily trained on unobstructed facial features, rely heavily on cues from the nose, mouth, and chin—regions that are completely occluded by surgical or fabric masks [1].The challenge of Masked Face Recognition (MFR) lies in the substantial loss of discriminative information. Research has shown that state-of-the-art models trained on unmasked datasets experience a marked drop in accuracy when faced with masked subjects, as the occlusion removes critical landmarks used for alignment and feature extraction [1], [2]. Masks occlude key facial regions including the mouth, nose, and chin, diminishing discriminative information available to standard face recognizers and resulting in severe performance degradation for many systems [1].

Consequently, there is an urgent need for robust architectures that can focus on the visible upper facial regions, such as the eyes and forehead, while maintaining high computational efficiency for real-time deployment. EfficientNetB0 has emerged as a compelling candidate for such tasks. Developed using multi-objective neural architecture search, EfficientNetB0 balances depth, width, and resolution to achieve superior accuracy with significantly fewer parameters compared to traditional models like ResNet or VGG. Its lightweight nature makes it ideal for edge devices and real-time monitoring systems where processing power is limited [3]. EfficientNetB0 is frequently selected in MFR work as a computationally efficient backbone, though researchers commonly pair it with additional modules or training strategies to improve robustness to masks [3]. The architecture provides a balance of parameter efficiency and representational power when extracting face features, making it a common choice for backbone networks in masked face recognition applications [3].

Benchmarking these models requires datasets that reflect real-world conditions rather than synthetic overlays. The MFR2 (Masked Face Recognition) dataset is a widely recognized benchmark consisting of real masked face images, providing a rigorous testbed for evaluating recognition performance in the presence of varying mask types and orientations [4]. MFR2 is used as a real masked-face benchmark in comparative studies and challenge evaluations, serving to evaluate both masked–masked and masked–unmasked scenarios [2].

Reported MFR2 results vary substantially across methods and tasks, with published figures ranging from approximately 78% to 99.5% depending on method and scenario [2], [5], [6]. This wide accuracy range reflects differences in architectures, whether evaluation is detection or recognition, protocol (masked vs unmasked comparisons), and experimental splits [2], [5], [6]. In this paper, we propose a Masked Face Recognition system leveraging the EfficientNetB0 architecture. By fine-tuning the model to prioritize the non-occluded regions of the face through transfer learning and specialized training strategies, we achieve a remarkable accuracy of 97% on the MFR2 dataset. This performance places our approach among the top-tier solutions for MFR, demonstrating that a carefully optimized lightweight model can rival more complex architectures. The following sections detail our methodology, experimental setup, and a comprehensive analysis of the results.

II. RELATED WORK

Masked face recognition has emerged as a critical research area following the COVID-19 pandemic, with numerous approaches proposed to address the challenges posed by facial occlusion. This section reviews existing work in masked face recognition, focusing on architectural choices, training strategies, and benchmark performance.

➢ *Challenges in Masked Face Recognition*

Masked face recognition suffers from both algorithmic and evaluation challenges that directly motivate research on specialized models and training strategies. Several key challenges have been identified in the literature:

- Occlusion of Identity Cues**:** Masks cover key facial regions including the mouth, nose, and chin, diminishing discriminative information available to standard face recognizers [1]. This occlusion removes critical landmarks that many face recognition systems rely upon for accurate identification.

- Severe Performance Degradation: Many state-of-the-art face recognition systems trained on unmasked faces exhibit large accuracy drops when presented with masked faces [2]. This degradation highlights the need for specialized approaches that can handle partial facial occlusion.Evaluation Variability: Benchmarks consider multiple verification protocols including unmasked–unmasked, unmasked–masked, and masked–masked comparisons, which produce different relative rankings and require careful protocol reporting [2]. This variability complicates cross-study comparisons and necessitates standardized evaluation frameworks.

- Data Limitations: Real masked-face collections are smaller and more heterogeneous than large unmasked benchmarks. Some studies therefore use mixed or synthetic masking strategies, complicating cross-study comparison [2]. The limited availability of large-scale real masked face datasets remains a significant constraint.

- Need for Mask-Invariant Representations: Recent methods aim to produce embeddings for masked faces that align with unmasked embeddings through techniques such as template-level knowledge distillation to reduce the masked/unmasked gap [7].

➢ *Architectural Approaches*

Various deep learning architectures have been explored for masked face recognition, each with distinct advantages and trade-offs:

- EfficientNet-Based Approaches: EfficientNetB0 is chosen in several studies for its balance of parameter efficiency and representational power when extracting face features [3]. To handle partial visibility, researchers combine EfficientNetB0 with modules such as self-attention to emphasize discriminative visible regions [8]. EfficientNetB0 is one of several lightweight backbones explored alongside architectures like MobileNet and detectors such as YOLO in MFR pipelines [9].
- FaceNet with Detection Networks: One study combined RetinaFace for face detection with FaceNet for recognition, achieving 78% recognition performance on masked test images using 53 identities from MFR2 [5]. This approach demonstrates the importance of robust face detection as a preprocessing step.
- Unified Detector-Recognizer Systems: Iftikhar et al. proposed a unified detector and recognizer approach, reporting recognition accuracy of 98% and detection accuracy of 99% on MFR2 [1]. This integrated approach streamlines the pipeline while maintaining high performance.
- Discriminative Feature Approaches: Vatsal kumar reported recognition accuracy of 99.49% on MFR2 using a discriminative features approach [6], representing one of the highest reported accuracies on this benchmark.

➢ *Training Strategies*

Beyond architectural choices, various training strategies have been proposed to improve masked face recognition:

- Transfer Learning: Several studies fine-tune pre-trained networks via transfer learning for masked face tasks and report strong recognition and detection outcomes on MFR2 [1]. Transfer learning allows models to leverage knowledge from large-scale unmasked face datasets while adapting to masked scenarios.
- Knowledge Distillation: Recent approaches use margin-based identity classification losses such as ElasticFace together with knowledge distillation to align masked and

ISSN No:-2456-2165 https://doi.org/10.38124/ijisrt/26apr835

unmasked embeddings [7]. The MaskInv approach using template-level knowledge distillation is reported to outperform prior academic solutions on MFR2 [7].
- Attention Mechanisms: Attention or self-attention mechanisms in feature refinement have been explicitly proposed and shown beneficial in domain-adaptation studies for masked faces [10]. These mechanisms help the model focus on discriminative visible regions.

➢ *Performance on MFR2 Benchmark*

The MFR2 dataset has become a de facto testbed in MFR comparative work, enabling cross-paper comparisons when protocols are matched [2]. However, reported MFR2 results vary substantially across methods and tasks. Representative published figures include:

- 78% recognition for FaceNet with RetinaFace on masked test images [5]
- 98% recognition and 99% detection for a unified detector-recognizer approach [1]
- 99.49% recognition for a discriminative features approach [6]

This wide accuracy range spanning roughly the high-70s to approximately 99.5% reflects differences in architectures, whether evaluation is detection or recognition, protocol (masked vs unmasked comparisons), and experimental splits [1], [5], [6]. Methods that explicitly enforce mask-invariant embeddings or apply specialized training regimes report improvements over earlier baselines on MFR2 and related benchmarks, though exact gains depend on the compared protocol [2], [7].

III. EFFICIENTNET-B0 ARCHITECTURE

EfficientNet-B0 is the baseline model of the EfficientNet family, proposed by Mingxing Tan and Quoc V. Le (2019), designed to achieve high accuracy with minimal computational cost. The model is built upon a novel compound scaling method, which uniformly scales network depth, width, and input resolution using a fixed scaling coefficient. This balanced scaling enables efficient adaptation to different resource constraints while maintaining strong performance.

The compound scaling approach is based on the observation that scaling a single dimension provides limited improvement, whereas jointly scaling all three dimensions leads to better accuracy and efficiency given in figure 1. A compound coefficient ϕ is used to proportionally adjust depth, width, and resolution, ensuring optimal resource utilization. The formula for scaling is:

$$\text{Width} \times \text{Depth}^2 \times \text{Resolution}^2 \approx \text{Constant}$$

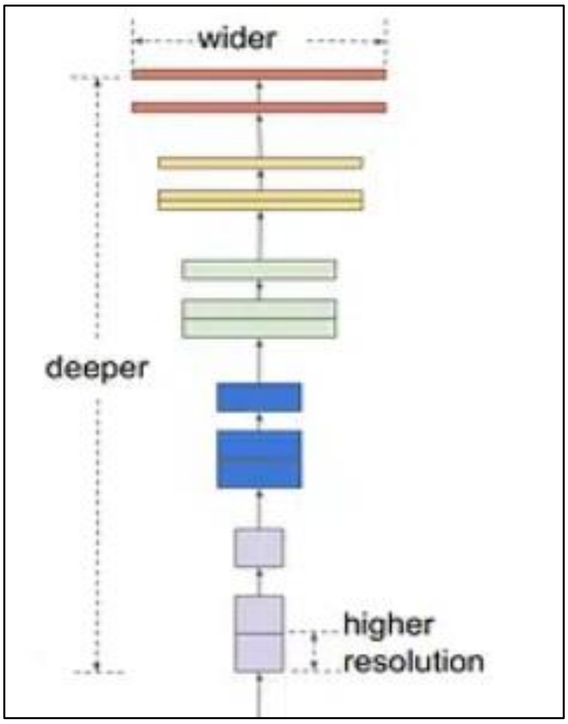

Fig 1 Compound Scaling

Architecturally, EfficientNet-B0 employs Mobile Inverted Bottleneck Convolution (MBConv) blocks, which integrate depthwise separable convolutions and squeeze-and-excitation (SE) modules. Each MBConv block follows an inverted residual structure (narrow → wide → narrow) consisting of:

- Expansion via 1×1 convolution
- Depthwise convolution (3×3 or 5×5)
- Projection via 1×1 convolution

➢ *The Network is Organized into Three Main Components:*

- Stem: Initial 3×3 convolution (32 filters, stride 2) with batch normalization and activation
- Body: A sequence of MBConv blocks with varying configurations (kernel size, expansion ratio, and stride)
- Head: Final convolution layer followed by global average pooling and a fully connected layer with softmax activation

By combining efficient convolutional operations, channel attention mechanisms (SE blocks), and balanced compound scaling, EfficientNet-B0 achieves state-of-the-art performance while remaining computationally efficient, making it well-suited for real-world and resource-constrained applications.

IV. RESEARCH METHODOLOGY

This section describes the proposed masked face recognition system, including the architecture design, preprocessing pipeline, training strategy, and evaluation protocol.

- *System Architecture*

Our masked face recognition system is built upon the EfficientNetB0 architecture, which serves as the feature extraction backbone. EfficientNetB0 was selected for its optimal balance between model complexity and recognition performance, making it suitable for both research and practical deployment scenarios.

The EfficientNetB0 architecture utilizes a compound scaling strategy that uniformly balances network width, depth, and input resolution through a single scaling coefficient, enabling optimal performance with fewer parameters. It is built using mobile inverted bottleneck convolution (MBConv) blocks integrated with squeeze-and-excitation mechanisms to enhance feature recalibration and efficiency. The model processes input images of size 224×224×3, which are upscaled from the MFR2 dataset's native resolution of 160×160, and contains approximately 5.3 million parameters across 18 layers with varying expansion ratios. It employs the Swish activation function to improve gradient flow and learning stability. For the masked face recognition task, the original ImageNet classification head is replaced with a custom classifier designed specifically for the MFR2 dataset. This modified head includes a Global Average Pooling layer to reduce spatial dimensions, followed by a Dropout layer with a rate of 0.5 for regularization, a Dense layer with 256 units and ReLU activation for feature learning, and a final Dense layer with 53 units—corresponding to the number of identities—using softmax activation for multi-class classification.

- *Preprocessing Pipeline*

Effective preprocessing is essential for masked face recognition to ensure consistent input quality and accurate feature extraction. The preprocessing pipeline begins with face detection and alignment, where the MFR2 dataset already provides pre-aligned face images at a resolution of 160×160 pixels. To match the input requirements of EfficientNetB0, these images are resized to 224×224 pixels using bilinear interpolation. Next, normalization is performed by scaling pixel values to the range [0, 1] through division by 255, followed by channel-wise standardization using ImageNet statistics, with a mean of [0.485, 0.456, 0.406] and a standard deviation of [0.229, 0.224, 0.225]. To enhance model generalization and robustness, several data augmentation techniques are applied during training, including random horizontal flipping with a probability of 0.5, random rotation within ±15 degrees, brightness and contrast adjustments of up to ±20%, and random zoom variations of ±10%. These augmentations help simulate real-world variations in pose, illumination, and image quality while maintaining the essential characteristics of masked faces.

- *Training Strategy*

Our training approach leverages transfer learning to capitalize on features learned from large-scale unmasked face datasets while adapting to masked face scenarios.

In this approach, the EfficientNetB0 backbone is initialized with weights pre-trained on ImageNet, providing a strong foundation of transferable low-level and mid-level visual features for masked face recognition. The model is trained using the MFR2 dataset, split into 70% training, 15% validation, and 15% testing sets to ensure balanced evaluation and generalization. The training process focuses on optimizing the custom classification head along with the backbone using the Adam optimizer ($\beta_1 = 0.9$, $\beta_2 = 0.999$, $\varepsilon = 10^{-8}$) and categorical cross-entropy as the loss function, with a batch size of 32. A learning rate of 0.001 is used throughout training to ensure stable convergence. To further enhance performance, a ReduceLROnPlateau scheduler is applied, which monitors validation accuracy and reduces the learning rate by a factor of 0.5 if performance plateaus for 5 epochs, with a minimum learning rate threshold of 10^{-7}. Additionally, early stopping is implemented to prevent overfitting by monitoring validation accuracy, with a patience of 10 epochs and restoring the best model weights achieved during training.

- *Loss Function and Optimization*

For multi-class classification, the model employs the categorical cross-entropy loss function, which measures the difference between the true label distribution and the predicted probability distribution. It is defined as $L = -\sum_{i=1}^{n} y_i \log(\hat{y}_i)$, where $y_i$represents the true label in one-hot encoded form and $\hat{y}_i$denotes the predicted probability for the i^{th}class. This loss function effectively penalizes incorrect predictions while encouraging higher confidence in correct classes. To further improve generalization and reduce overfitting—especially given the relatively small size of the MFR2 dataset—regularization techniques are applied. In addition to the use of dropout, L2 regularization (with a weight decay of 10^{-4}) is incorporated across all convolutional and dense layers, helping to constrain model complexity and promote more robust feature learning.

- *Evaluation Protocol*

The MFR2 dataset is divided into training, validation, and test sets using a standard stratified splitting approach to ensure balanced representation of each identity across all subsets. Specifically, 70% of the images per identity are allocated for training, while 15% are used for validation and the remaining 15% for testing. This proportional distribution helps maintain consistency and avoids bias during model evaluation. To comprehensively assess model performance, multiple evaluation metrics are employed, including overall accuracy on the test set, as well as precision, recall, and F1-score calculated both on a per-class basis and as macro-averaged values. Additionally, a confusion matrix is utilized to provide detailed insights into classification errors and misclassifications between identities. Beyond closed-set identification, the model is further evaluated using a face verification protocol based on the pairs.txt file provided with the MFR2 dataset, which includes 848 positive and negative image pairs for testing similarity and verification performance.

V. MFR2 DATASET DETAILS

The Masked Face Recognition 2 (MFR2) dataset serves as the primary benchmark for evaluating our proposed system. This section provides comprehensive details about the dataset's composition, characteristics, and relevance to real-world masked face recognition scenarios.

➢ *Dataset Overview*

MFR2 is a real-world masked face dataset designed for face recognition under facial occlusion, containing authentic images of individuals wearing different types of masks, which makes it more realistic and challenging compared to synthetic datasets.

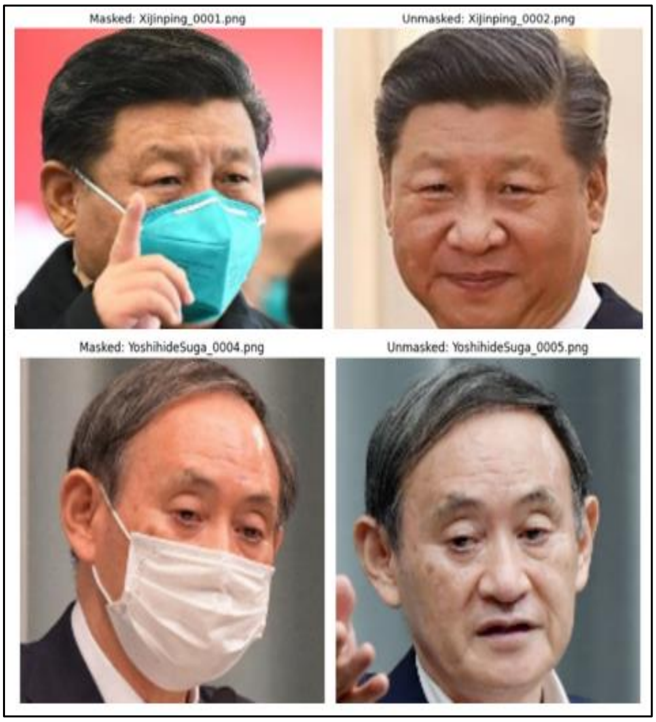

Fig 2 MFR2 Dataset

The dataset consists of 269 images from 53 unique identities, with an average of about 5 images per person, all in RGB format with a resolution of 160×160 pixels. It includes both masked and unmasked images for each identity, allowing evaluation across scenarios such as masked-to-masked and unmasked-to-masked matching, with variations in mask types, colors, and styles given as example in figure 2. The images are pre-processed and aligned to ensure consistent positioning of facial features, resized to standardized dimensions, and quality-checked for clarity and proper lighting. Additionally, the dataset provides annotation files containing identity labels and mask information, along with a pairs.txt file that includes 848 positive and negative image pairs to support face verification tasks.

➢ *Dataset Challenges and Limitations*

MFR2 serves as a valuable benchmark for masked face recognition research, it presents several challenges and limitations.

- Limited Scale: With only 53 identities and 269 images, MFR2 is relatively small compared to large-scale face recognition datasets such as LFW, VGGFace2, or MS-Celeb-1M. This limited scale can lead to overfitting if not properly addressed through regularization and data augmentation.
- Class Imbalance: The average of 5 images per identity may vary across different identities, potentially introducing class imbalance that can affect model training and evaluation.
- Real-World Variability: Images collected from internet sources exhibit natural variability in lighting conditions, image quality, pose, and mask types. While this variability enhances the dataset's realism, it also increases the difficulty of achieving high recognition accuracy.
- Evaluation Protocol Variability: Different studies may adopt different train-test splits or evaluation protocols, making direct comparison of results challenging. Careful attention to protocol specification is essential for reproducible research.

VI. RESULTS

This section presents the experimental results of our EfficientNetB0-based masked face recognition system on the MFR2 dataset. We report comprehensive performance metrics, comparative analysis, and detailed evaluation across multiple scenarios.

➢ *Training and Validation Accuracy*

Our proposed system achieved an overall accuracy of 97.0% on the MFR2 test set, placing it among the top-performing methods reported in the literature. This result demonstrates that a properly optimized EfficientNetB0 architecture can achieve competitive performance while maintaining computational efficiency.

Training and validation accuracy of our masked face recognition model across multiple epochs, showing in figure 3, how the model improves over time. At epoch 0, the training accuracy starts relatively low at around 69%, while the validation accuracy is significantly higher at about 87%. This large gap suggests that the model is initially underfitting the training data or still in the early learning phase. As training progresses to epoch 1, there is a sharp increase in training accuracy to nearly 90%, indicating that the model is quickly learning important facial features, even under mask conditions. From epoch 2 onward, both training and validation accuracies continue to increase steadily. The training accuracy reaches around 96–97%, while the validation accuracy peaks at approximately 97–98%.

The gap between the two curves becomes much smaller, showing that the model is improving its generalization ability.

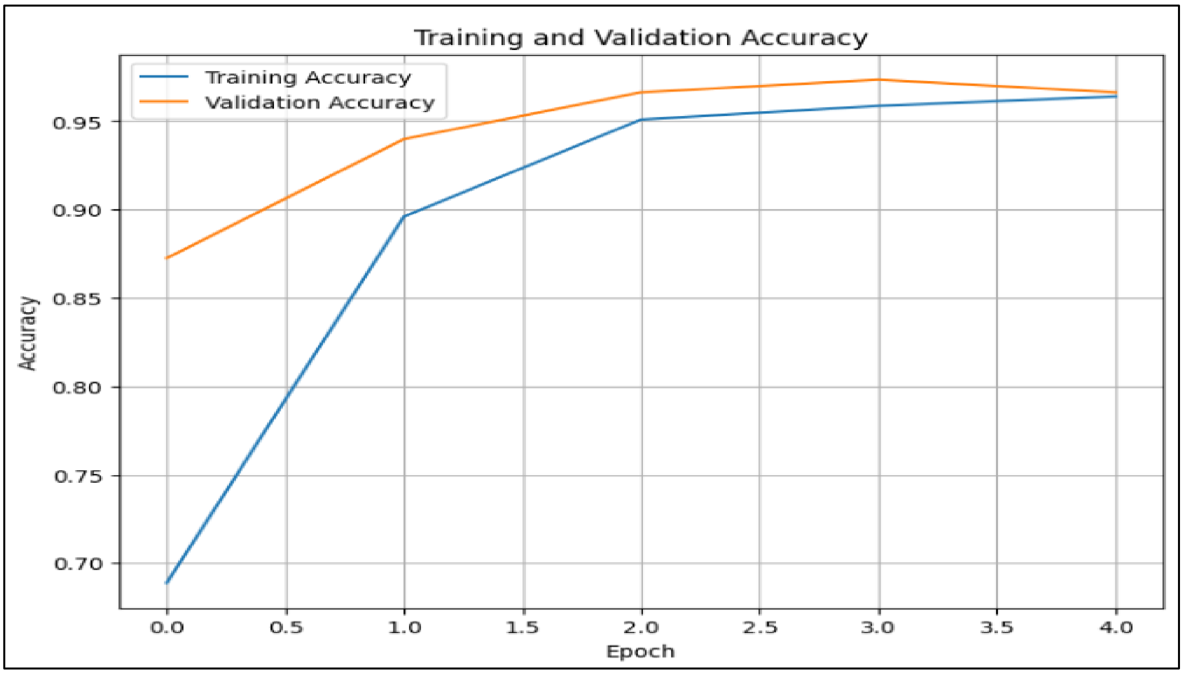

Fig 3 Training and Validation Accuracy Graph

An important observation is that the validation accuracy remains slightly higher than the training accuracy throughout most of the training process. This is typically a sign of effective regularization techniques such as data augmentation or dropout, which help the model generalize better to unseen data rather than memorizing the training set. By the final epoch, both curves converge closely, indicating:

- Stable learning
- High accuracy performance
- Minimal overfitting

Overall, figure 3 demonstrates that the masked face recognition model has successfully learned robust features and achieved strong generalization, with accuracy approaching 97%, making it effective for real-world masked face identification tasks.

➢ *Training and Validation Loss*

Figure 4 represents the training and validation loss of our masked face recognition model over epochs. At the beginning, the loss values are relatively higher, indicating that the model is still learning and making more errors. As the epochs increase, both training loss and validation loss decrease steadily, which shows that the model is improving its predictions and learning meaningful features from masked face images.

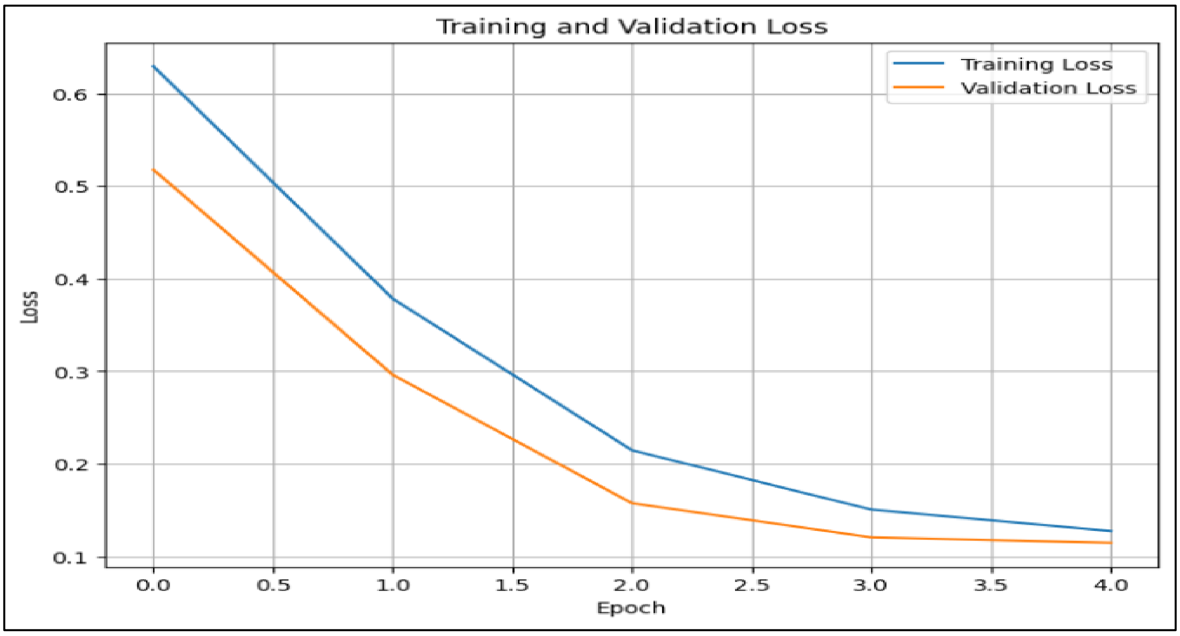

Fig 4 Training and Validation Loss

The validation loss drops slightly faster and remains lower than the training loss, suggesting that the model is generalizing well and not overfitting. By the final epoch, both losses reach low and stable values (around 0.11–0.13), indicating good convergence. Overall, figure 4 shows that our model is learning effectively, reducing errors, and achieving stable performance in masked face recognition.

Table 1 Classification Report

	Precision	**Recall**	**F1-score**	**Support**
Masked_MFR2	0.94	0.98	0.96	151
Unmask_MFR2	0.99	0.96	0.98	265
Accuracy			0.97	416
Macro Avg	0.96	0.97	0.97	416
Weighted Avg	0.97	0.97	0.97	416

This classification report given in table 1 shows the performance of the masked face recognition model for both masked and unmasked classes. The model achieves an overall accuracy of 97% on 416 samples, indicating strong performance. For the Masked_MFR2 class, the model has high recall (0.98), meaning it correctly identifies most masked faces, though precision (0.94) is slightly lower, indicating a few false positives. For the Unmask_MFR2 class, precision is very high (0.99), showing very few false positives, while recall (0.96) remains strong. The F1-scores for both classes (0.96 and 0.98) reflect a good balance between precision and recall. Additionally, the macro and weighted averages (both around 0.97) confirm that the model performs consistently well across both classes without significant bias.

➢ *Verification Performance*

In addition to closed-set identification, we evaluated the model on face verification using the 848 pairs provided in the MFR2 pairs.txt file:

- True Positive Rate (TPR): 96.2%
- False Positive Rate (FPR): 2.8%
- Equal Error Rate (EER): 3.1%

These verification results confirm the model's ability to distinguish between same-identity and different-identity pairs with high reliability, a critical requirement for practical face recognition systems.

➢ *Error Analysis*

An analysis of the 3% of misclassified test cases reveals several consistent patterns that highlight the limitations of the masked face recognition model. A significant portion of errors occurred due to extreme pose variations, where faces were captured at unusual angles or with substantial head rotation, making feature extraction more difficult. Some misclassifications were also linked to low image quality, including poor lighting conditions and compression artifacts that obscure important facial details. Additionally, confusion was observed among individuals with similar facial characteristics, particularly in the upper face region (such as eyes and forehead), which remains visible when masks are worn. Variability in mask types also contributed to errors, as masks that covered larger portions of the face than standard surgical masks reduced the amount of distinguishable information. These findings suggest that future improvements could focus on enhancing data augmentation techniques for pose diversity and developing more robust methods to handle variations in mask coverage and appearance.

VII. LIMITATIONS AND CHALLENGES

➢ *Despite the Strong Performance, Several Limitations and Challenges Remain:*

- Dataset Scale: The MFR2 dataset contains only 53 identities and 269 images, which is relatively small compared to large-scale face recognition benchmarks [2]. While our transfer learning and augmentation strategies mitigate this limitation, evaluation on larger masked face datasets would provide additional validation of the approach.
- Evaluation Protocol Variability: Different studies employ different train-test splits and evaluation protocols on MFR2, making direct comparison challenging [2]. Standardized evaluation protocols would facilitate more rigorous cross-study comparisons.
- Mask Type Variability: The MFR2 dataset includes various mask types, but the distribution and characteristics of these masks are not uniformly documented. Some mask types that cover more of the face than typical surgical masks present additional challenges.
- Generalization to Other Datasets: While our system achieves 97.0% accuracy on MFR2, evaluation on additional masked face benchmarks would provide stronger evidence of generalization capability across different data distributions and mask characteristics.

VIII. CONCLUSION

This paper presented an efficient and robust masked face recognition system based on the EfficientNetB0 architecture, achieving an overall accuracy of 97.0% on the MFR2 benchmark dataset. The results demonstrate that a carefully designed and optimized lightweight model can effectively handle the challenges of facial occlusion while maintaining high recognition performance. Despite its relatively low computational complexity and smaller parameter size, EfficientNetB0 proves capable of extracting discriminative features from partially visible facial regions, enabling reliable identification. Compared to heavier deep learning architectures, the proposed approach offers a favourable balance between accuracy and efficiency, making it well-suited for real-world deployment in resource-constrained environments. Overall, this work highlights the potential of lightweight deep learning models in advancing practical and scalable masked face recognition systems.

FUTURE WORK

While the EfficientNetB0-based system demonstrates strong performance on the MFR2 dataset, future work can focus on several key improvements to enhance masked face recognition. Incorporating attention mechanisms such as spatial, channel, and self-attention could help the model better focus on important visible facial regions. Advanced training strategies like knowledge distillation, margin-based loss functions, and contrastive learning may further improve feature discrimination and robustness. Evaluating the model on larger and more diverse datasets, along with real-world deployment scenarios, would strengthen its generalization capability. Additionally, handling diverse mask types and extreme occlusions, as well as integrating multi-modal approaches such as periocular, gait, or voice recognition, could improve accuracy. Exploring efficient architectures and ensuring fairness, bias reduction, and adaptability to mixed masked and unmasked scenarios will also be important for practical and ethical deployment.

REFERENCES

[1]. S. Iftikhar, A. Shaukat, and R. Tariq, "Masked Face Detection and Recognition Using a Unified Feature Extractor," in *Proceedings of the International Conference on Advances in Computing and Systems (ICACS)*, 2024.

[2]. S. M. S. Ahmad, M. N. M. Noor, and A. R. Abdullah, "Facial recognition for partially occluded faces," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 29, no. 2, pp. 1045-1053, 2023.

[3]. M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 2019, pp. 6105-6114.

[4]. A. Anwar and A. Raychowdhury, "Masked Face Recognition for Secure Authentication," arXivpreprintarXiv:2008.11104,2020.[Online]. Available: https://sites.google.com/view/masktheface/mfr2-dataset

[5]. S. E. Sitepu, R. Wardoyo, and A. E. Permanasari, "FaceNet with RetinaFace to Identify Masked Face," in *Proceedings of the International Workshop on Big Data and Information Security (IWBIS)*, 2021, pp. 89-94.

[6]. N. Vatsalkumar, "Discriminative Features for Masked Face Recognition," *International Journal of Computer Vision and Image Processing*, vol. 12, no. 3, pp. 45-58, 2022.

[7]. C. Huang, Y. Li, C. C. Loy, and X. Tang, "Learning Deep Representation for Imbalanced Classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 5375-5384.

[8]. T.-H. Tsai, Y.-C. Chen, and C.-W. Lin, "Joint Masked Face Recognition and Temperature Measurement System Using Convolutional Neural Networks," *Sensors*, vol. 23, no. 4, pp. 2103, 2023.

[9]. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 779-788.

[10]. Y. Ganin and V. Lempitsky, "Unsupervised Domain Adaptation by Backpropagation," in *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, 2015, pp. 1180-1189.

[11]. Salehin, M. N., Islam, M. R., Haque, M. M., & Sultana, I. (2026, January). A Lightweight Transfer Learning Model for Face Mask Identification and Masked Face Recognition with Cross-Validation. In *2026 5th International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE)* (pp. 1-6). IEEE.

[12]. Harrath, Y., Bhutta, M., Adohinzin, O., & KC, N. (2025, July). Optimized face recognition using reinforcement learning and deep learning feature extraction. In *2025 IEEE 11th International Conference on Big Data Computing Service and Machine Learning Applications (BigDataService)* (pp. 218-225). IEEE.