JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

IMPROVING MASKED FACE RECOGNITION WITH VGG16: A DEEP LEARNING APPROACH

Girish Katkar[1], Lekha Prajapati[2], Ajay Ramteke[3]
[1]Assistant Professor, [2]Research Scholar, [3]Assistant Professor
[1]Computer Science Department,
[1]Taywade College Mahadula, Koradi, Nagpur, India

***Abstract**:* The recent COVID-19 pandemic has necessitated wearing masks as a protective measure against viral transmission. This has significantly impacted facial recognition technologies, as traditional systems struggle to identify individuals accurately when facial features are partially obscured. The presence of masks has posed substantial challenges for existing facial recognition systems due to the obstruction of key facial landmarks. Consequently, there is an urgent need for new methods that can improve the accuracy and robustness of facial recognition in the context of masked faces. To address the challenges associated with masked face recognition, this study presents a novel method that approach with a modified VGG-16 architecture. The model is fine-tuned to adapt to the masked face recognition task, focusing on the upper face region. The key idea is to leverage transfer learning to extract distinctive features from the upper half of the face, focusing on the forehead and eye regions. This transfer learning approach enables the extraction of the most relevant features from the unobstructed areas. By focusing on the upper half of the face and employing a modified VGG-16 architecture, the system achieves 0.99 recognition accuracy across MFR2 Dataset.
IndexTerms - Masked Face Recognition, Transfer Learning, VGG-16 architecture

I. INTRODUCTION:

The COVID-19 pandemic has led to widespread requirements for wearing masks in public spaces to prevent virus transmission. This change has significantly impacted traditional biometric techniques like fingerprint and facial recognition, as contact-based methods pose health risks. Masked face recognition presents a unique challenge since facial masks obscure key features, making it difficult to identify individuals accurately. Manual inspection in public spaces has become challenging, necessitating the development of automated and contactless recognition systems that can effectively recognize people wearing masks in security-sensitive areas like airports, railway stations, and banks. Traditional facial recognition systems face a considerable decline in accuracy when applied to masked faces because masks cover essential facial features, such as the nose, mouth, and cheeks, which play a significant role in face recognition. Additionally, masks come in various colors, patterns, and shapes, adding further complexity to the task of distinguishing the facial region from the mask. This increased variability, combined with changes in lighting, head poses, and skin tones, poses substantial challenges to existing recognition methods. Consequently, new approaches are needed to maintain high performance under these conditions.

In response to these challenges, deep learning techniques have been leveraged to improve the performance of masked face recognition. Methods such as deep metric learning, combined with models like FaceMaskNet-21 and Inception-v4, have shown promising results in recognizing masked faces in real-time videos and static images. For example, the FaceMaskNet-21 model achieved an accuracy of 88.92%, while the Inception-v4 model reached 82.12%, demonstrating that these advanced models can provide significant improvements over traditional methods. However, occlusions caused by masks still degrade the quality of extracted features, emphasizing the need for further enhancements. To overcome these limitations, researchers have begun employing attention mechanisms that focus on the visible regions of the face, such as the eyes, to extract more discriminative features for masked face recognition. Additionally, collecting large-scale datasets that include various types of masked faces is crucial, but it is a resource-intensive process. Thus, there is an urgent need for low-cost data augmentation techniques to diversify training data effectively. By combining these strategies with optimized model designs, it is possible to enhance the accuracy of masked face recognition, making it more reliable for applications in access control, security checks, and other public settings.

In this research paper, the VGG16 model will be utilized as a core component for masked face recognition due to its proven effectiveness in extracting meaningful features from images. VGG16, with its 16-layer architecture, is known for its ability to capture intricate details through a series of small convolutional filters (3x3), making it well-suited for image recognition tasks, even when dealing with partially occluded faces. By leveraging pre-trained weights on large-scale datasets like ImageNet, VGG16 can be fine-tuned to recognize facial features while focusing on the visible regions of the face.

II. RELEATED WORK

Putthiporn Thanathamathee et al. [1] An Optimized Machine Learning and Deep Learning Framework for Facial and Masked Facial Recognition. This study introduces a novel approach that incorporates grid search for hyperparameter tuning and nested cross-validation during the model verification phase. Unlike earlier studies that relied on single machine learning models without reporting optimal parameter settings, our method enhances model performance through systematic tuning. Results demonstrate that a Support Vector Machine (SVM) model with hyperparameter optimization significantly outperforms other models, achieving an accuracy of 0.99912. The approach is particularly relevant for real-world scenarios, validating the system's ability to recognize masked individuals accurately in practical face recognition applications.

meng zhang et al. [2] Masked Face Recognition with Mask Transfer and Self-Attention Under the COVID-19 Pandemic. This paper presents a novel approach for enhancing the accuracy of face recognition in the presence of masks, addressing the challenges posed by mask-induced occlusions. The proposed method consists of three key components. First, a cost-effective and precise masked face synthesis technique, called mask transfer, is introduced for data augmentation to increase the diversity and robustness of the training dataset. Second, an innovative attention-aware masked face recognition model (AMaskNet) is developed to boost recognition performance. AMaskNet integrates two modules: a feature extractor and a contribution estimator, where the latter is responsible for learning the significance of feature elements and refining feature representation through matrix multiplications. The entire model is optimized using an end-to-end training strategy. Lastly, a mask-aware similarity matching strategy (MS) is incorporated during the inference phase to enhance recognition accuracy.

Chen, C., Kurnosov, I., Ma, G. et al. [3] Masked Face Recognition Using Generative Adversarial Networks by Restoring the Face Closed Part. Researchers introduce a novel approach leveraging image segmentation to remove masks from faces in images. After removing the mask, they reconstruct the facial region that was covered by the mask, allowing us to apply existing face recognition techniques on the complete face. To restore the occluded facial features, they use Generative Adversarial Networks (GANs). his results demonstrate that the proposed method significantly enhances the recognition accuracy of masked faces compared to conventional approaches. Specifically, they compare his approach with the MobileNetV2-based algorithm, showing improved recognition performance.

Rucha Golwalkar and Ninad Mehendale [4] Masked-face recognition using deep metric learning and FaceMaskNet-21. The proposed system leverages deep metric learning along with a custom-built FaceMaskNet-21 deep learning network to facilitate face recognition. The system generates 128-dimensional encodings for accurate identification from various input sources, including static images, live video streams, and video files. It achieves a testing accuracy of 88.92%, with an execution time of under 10 milliseconds, making it capable of real-time masked face recognition. This capability enables the system to be used effectively in monitoring and surveillance scenarios, such as identifying individuals in CCTV footage from malls, banks, ATMs, and other public areas. Its fast-processing speed also makes it suitable for use in schools, colleges, and high-security environments like banks, where it can grant access to authorized individuals without requiring them to remove their masks.

III.RESEARCH METHODOLOGY

In This research paper we proposed Research Methodology with VGG16 Architecture showing in below steps.

1. Dataset Selection: Use datasets containing both masked and unmasked facial images, such as RMFRD (Real-World Masked Face Recognition Dataset), MFR2, or a custom dataset.
2. Data Augmentation: Apply data augmentation techniques to increase variability in the training data. This can include random rotations, flips, scaling, brightness changes, and synthetic mask application to unmasked face images.
3. Image Preprocessing: Resize all images to a fixed size of 224x224 pixels to match the input requirement of VGG16.
4. Transfer Learning with VGG16: Utilize the VGG16 model pre-trained on the ImageNet dataset for transfer learning, which allows the model to leverage the pre-learned features for masked face recognition.
5. Model Customization: Remove the top fully connected layers of the pre-trained VGG16 model and add new layers suitable for face recognition tasks:
    - Flatten Layer: Flatten the output of the last convolutional layer.
    - Dense Layers: Add one or more fully connected layers with a ReLU activation function.
    - Dropout Layer: Include a dropout layer to prevent overfitting during training.
    - Output Layer: Use a final dense layer with softmax activation for multi-class classification or sigmoid activation for binary classification.
6. Training the Model: Divide the dataset into training, validation, and test sets. Ensure that masked and unmasked face images are represented in all sets.

- Loss Function and Optimization:

Use categorical cross-entropy loss for multi-class classification or binary cross-entropy for binary classification tasks. Employ optimizers like Adam or Stochastic Gradient Descent (SGD), with a learning rate scheduler for better convergence.

- Hyperparameter Tuning: Tune hyperparameters such as learning rate, batch size, number of epochs, and dropout rate to optimize model performance.

7. Test and Evaluate: Test model on dataset with VGG16 architecture and generated accuracy to masked face recognition.

IV. EXPERIMENTAL WORK

The experimental work for this research involves developing a masked face recognition system using the VGG16 model architecture, which is fine-tuned on the MFR2 dataset to classify faces into two categories: "Masked" and "Unmasked." The primary goal is to assess the system's ability to generalize and accurately recognize masked faces using transfer learning techniques. Here's a detailed explanation of the experimental steps involved in the research:

Step 1: Model Architecture
VGG16 Pre-trained Model: The VGG16 model, pre-trained on the ImageNet dataset, is used as the base model. This model has demonstrated strong feature extraction capabilities in computer vision tasks, making it a suitable choice for transfer learning. The top layers (fully connected layers) of VGG16 are removed, keeping the convolutional layers that have learned general-purpose features. New custom layers are added: a flattening layer followed by a dense layer with 512 units and ReLU activation, and a dropout layer with a 50% dropout rate to prevent overfitting. The final output layer uses a SoftMax activation function with the number of neurons equal to the number of classes (Masked and Unmasked).
Step 2: Data Preparation
Dataset: The MFR2 dataset, containing 269 images across 53 identities (both celebrities and politicians), is divided into training and testing sets. The images include both masked and unmasked faces, with labels indicating the category.
Data Augmentation: The training images are augmented using transformations like horizontal flipping and zooming to improve the model's ability to generalize. Both training and testing images are normalized by rescaling pixel values to a range of 0 to 1.
Step 3: Training Configuration
Data Generators: The ImageDataGenerator is utilized to preprocess and augment the training and testing data. The images are resized to 224x224 pixels, which is the required input size for the VGG16 model. A batch size of 32 is used for training. The model is trained for binary classification (Masked and Unmasked), which corresponds to the two output classes in the dataset.
Step 4: Model Compilation and Training
The model is compiled using the Adam optimizer with a learning rate of 0. 0001.The loss function used is categorical cross-entropy, suitable for multi-class classification. The training process is carried out for 10 epochs. The training data is used to fit the model, while the testing data serves as the validation set to monitor the model's performance during training.
Step 5: Evaluation: After training, the model is evaluated on the test data to measure its accuracy and loss. The accuracy represents the percentage of correctly classified images, while the loss indicates the model's prediction error, our model is predicted 0.99 accuracy for masked face recognition.
Step 6: Predictions and Metrics
Classification Report and Confusion Matrix:
The model's predictions on the test data are compared with the true labels to generate a classification report, which includes precision, recall, and F1-score for each class. A confusion matrix is also produced to visualize the number of correct and incorrect predictions for each category in below figure 2.

Figure 2. Classification report

	Precision	Recall	F1-score	Support
Masked_MFR2	0.98	0.99	0.98	332
Unmask_MFR2	0.99	0.98	0.99	342
Accuracy			0.99	674
Macro Avg	0.99	0.99	0.99	674
Weighted Avg	0.99	0.99	0.99	674

The training accuracy and loss along with the testing accuracy and loss, are plotted over the epochs to visualize the model's learning process in below figure 3 and figure 4.

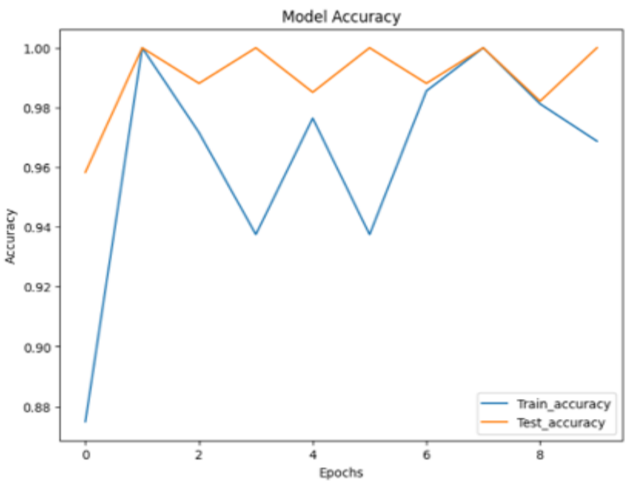

Figure 3. Train and Test accuracy

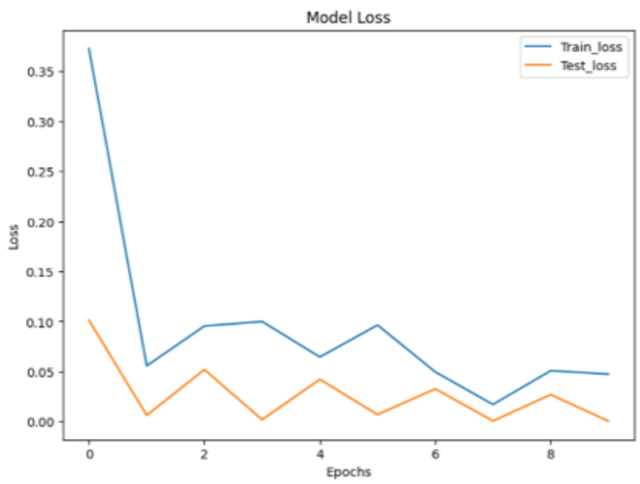

Figure 4. Train and Test

V. RELATED DATASET
The MFR2 dataset is a commonly used dataset in the development of robust masked face recognition systems. It consists of 269 images featuring 53 different identities, including celebrities and politicians. The dataset provides both masked and unmasked face images, which helps in training models to effectively generalize across various individuals, environments, and types of masks.
Key features of the MFR2 dataset:

1. Classes and Labels: The images are categorized into two main classes: Masked and Unmasked. Each image is labelled to indicate whether the person is wearing a mask, making it suitable for binary classification tasks in supervised learning.

2. Diversity: Despite its relatively small size, the dataset's inclusion of diverse face images helps train models that can generalize well across different scenarios. This diversity is critical in improving the recognition performance on real-world masked face recognition tasks.

Researchers often use the MFR2 dataset to fine-tune pre-trained models to better handle the challenges associated with recognizing masked faces. The process involves adapting models to learn from the differences between masked and unmasked faces, making it valuable for evaluating recognition systems under different conditions. For visual reference, Figure 5 in the original document likely illustrates examples of both masked and unmasked images from the MFR2 dataset.

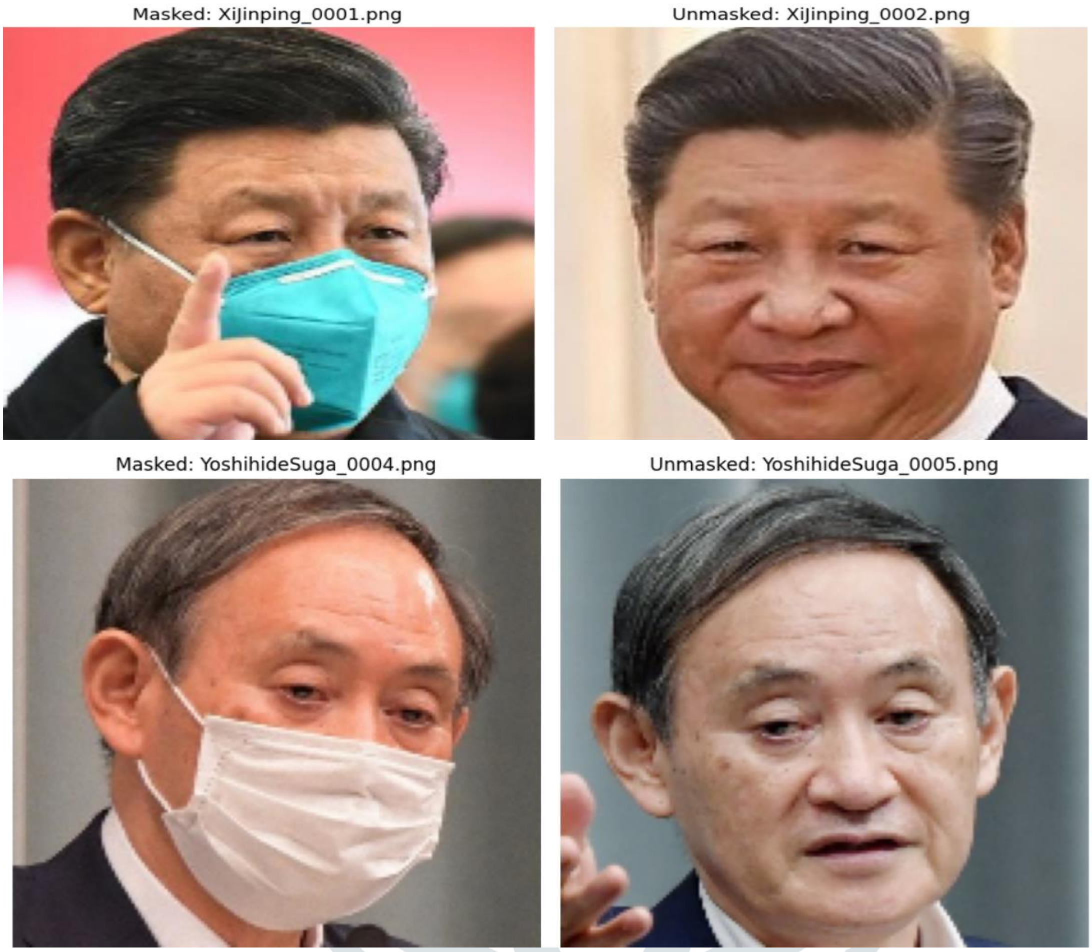

Figure 5. Masked and Unmasked MFR2 dataset images.

VI. CONCLUSION:

The recent COVID-19 pandemic has necessitated wearing masks as a protective measure against viral transmission. This has significantly impacted facial recognition technologies, as traditional systems struggle to identify individuals accurately when facial features are partially obscured. To address the challenges associated with masked face recognition, this study presents a novel method that approach with a modified VGG-16 architecture. The model is fine-tuned to adapt to the masked face recognition task, focusing on the upper face region. The evaluation conducted on the MFR2 dataset shows a remarkable recognition accuracy of 0.99, indicating the model's robustness in handling the variability introduced by different mask-wearing conditions. The results confirm that focusing on unoccluded facial regions while employing transfer learning techniques can substantially enhance the model's ability to recognize masked faces**.**

REFERENCES

[1] Thanathamathee, P., Sawangarreerak, S., Kongkla, P., & Nizam, D. N. M. (2023). An optimized machine learning and deep learning framework for facial and masked facial recognition. *Emerging Science Journal*, *7*(4), 1173-1187.

[2] Zhang, M., Liu, R., Deguchi, D., & Murase, H. (2022). Masked face recognition with mask transfer and self-attention under the COVID-19 pandemic. *IEEE Access*, *10*, 20527-20538.

[3] Chen, C., Kurnosov, I., Ma, G. *et al*. Masked Face Recognition Using Generative Adversarial Networks by Restoring the Face Closed Part. *Opt. Mem. Neural Networks* **32**, 1–13 (2023). https://doi.org/10.3103/S1060992X23010022

[4] Golwalkar, R., & Mehendale, N. (2022). Masked-face recognition using deep metric learning and FaceMaskNet-21. *Applied Intelligence*, *52*(11), 13268-13279.

[5] Muhi, O. A., Farhat, M., & Frikha, M. (2022, May). Transfer learning for robust masked face recognition. In *2022 6th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)* (pp. 1-5). IEEE.

[6] Jignesh Chowdary, G., Punn, N. S., Sonbhadra, S. K., & Agarwal, S. (2020). Face mask detection using transfer learning of inceptionv3. In *Big Data Analytics: 8th International Conference, BDA 2020, Sonepat, India, December 15–18, 2020, Proceedings 8* (pp. 81-90). Springer International Publishing.

[7] Talahua, J. S., Buele, J., Calvopiña, P., & Varela-Aldás, J. (2021). Facial recognition system for people with and without face mask in times of the covid-19 pandemic. *Sustainability*, *13*(12), 6900.

[8] Sikha, O.K., Bharath, B. VGG16-random fourier hybrid model for masked face recognition. *Soft Comput* **26**, 12795–12810 (2022). https://doi.org/10.1007/s00500-022-07289-0

[9] Ibitoye, O. T., & Osaloni, O. O. (2022, October). Masked Faces Classification using Deep Convolutional Neural Network with VGG-16 Architecture. In *2022 3rd International Conference on Electrical Engineering and Informatics (ICon EEI)* (pp. 168-171). IEEE.

[10] Kayali, D., Olawale, P., Kirsal-Ever, Y., & Dimililer, K. (2022, September). The effect of compressor-decompressor networks with different image sizes on mask detection using Convolutional Neural Networks-VGG-16. In *2022 Innovations in Intelligent Systems and Applications Conference (ASYU)* (pp. 1-5). IEEE.