

Face Mask Detection: A Computer Vision Approach for Object Classification

¹ Mr.M.Ganesh Gowd, ² K.Mohan Krishna, ³ U.Harika, ⁴ M.Chandini, ⁵T.Chaitanya Varma

¹Associate Professor, Srinivasa Institute of Engineering and Technology

²³⁴⁵UG Scholar, Srinivasa Institute of Engineering and Technology

doi.org/10.64643/IJIRTV12I10-195099-459

Abstract: In recent times, maintaining public health safety has become extremely important, particularly during global outbreaks such as COVID-19. Wearing face masks plays a vital role in reducing the spread of infectious diseases. However, manually checking whether people follow this rule in crowded areas is difficult and inefficient.

This paper proposes an automated face mask detection system that uses deep learning and computer vision techniques. The model is built using Convolutional Neural Networks (CNN) to classify whether a person is wearing a mask or not. Face detection is performed using pre-trained models, and the system works in real time through live video streams.

The experimental results indicate that the model achieves high accuracy and performs efficiently in real-world conditions. This system can be applied in public places such as hospitals, schools, airports, and shopping centres to ensure safety compliance and reduce human effort.

Keywords: Face Mask Detection, Deep Learning, Convolutional Neural Network (CNN), Computer Vision, Real-Time Monitoring, COVID-19.

I. INTRODUCTION

The COVID-19 pandemic had a major impact on global health, economies, and everyday life. To control the spread of the virus, several preventive measures were introduced, including social distancing, hand hygiene, and wearing face masks. Among these, the use of face masks proved to be one of the most effective and affordable methods to limit airborne transmission [12].

Despite strict regulations, ensuring that people consistently wear masks in crowded areas remains challenging. Manual monitoring by security personnel is not only time-consuming but also unreliable due to human errors and fatigue.

With the advancement of Artificial Intelligence (AI) and Deep Learning (DL), automated systems can now perform complex image analysis tasks.

Computer vision techniques, especially Convolutional Neural Networks (CNNs), are widely used[2][3] for image classification and object detection.

Face mask detection involves identifying human faces [1] in images or videos and determining whether a mask is present. By combining face detection algorithms with deep learning models, an efficient real-time monitoring system can be developed.

The proposed work focuses on designing an automated system that detects face masks using live video streams. This approach reduces manual effort and improves monitoring efficiency, making it suitable for integration with existing surveillance systems.

II. PROBLEM STATEMENT

Maintaining proper mask usage in public areas is essential for public health safety, especially during pandemic situations. Although guidelines have been issued by health organizations [12], enforcing these rules in crowded environments is difficult.

Public places such as railway stations, malls, and hospitals experience heavy crowd movement, making continuous monitoring challenging. Manual supervision requires a large workforce and is often affected by human limitations such as tiredness and lack of attention.

Traditional CCTV systems only record video without providing automatic analysis. As a result, violations may go unnoticed, and there is no provision for real-time alerts. Additionally, variations in lighting, camera angles, and crowd density make detection more complex.

Therefore, there is a need for an automated system that can accurately detect mask usage in real time and reduce dependency on manual monitoring.

III. EXISTING SYSTEM

In the existing system, face mask compliance is primarily monitored through manual observation by security personnel or by reviewing CCTV footage without automated analysis. Surveillance cameras are widely installed in public places such as airports, railway stations, shopping malls, hospitals, and educational institutions; however, these systems function mainly as recording devices rather than intelligent monitoring tools. Security staff are required to continuously observe multiple video streams to identify individuals who are not wearing masks, which is both time consuming and inefficient. This approach heavily depends on human attention and alertness, making it prone to fatigue, oversight, and inconsistent enforcement of safety regulations. Moreover, traditional surveillance systems lack automated alert mechanisms, real-time detection capabilities, and data logging features for tracking compliance. They do not utilize artificial intelligence or deep learning techniques to analyse visual data automatically. As a result, mask violations may go unnoticed, especially in crowded or fast-moving environments. The absence of intelligent processing limits scalability, as increasing the number of cameras also increases the requirement for additional manpower. Therefore, the existing system is not suitable for efficient, large-scale, real time mask compliance monitoring in dynamic public environments.

IV. LITERATURE SURVEY

Recent studies have focused on improving the performance[4] and efficiency of face mask detection systems using deep learning techniques. Several researchers have developed models based on CNN architectures[2][3] that can detect mask usage with high accuracy.

Researchers have also explored lightweight deep learning architectures such as MobileNet[7] and EfficientNet to enable real-time detection on low-power devices. These lightweight models are particularly useful for deployment in embedded

systems, surveillance cameras, and mobile devices where computational resources are limited.

Another area of research involves integrating object detection frameworks such as YOLO[5], SSD (Single Shot Detector) [8], and Faster R-CNN[9] to improve detection speed and accuracy. These models can simultaneously detect multiple faces in a single frame and classify whether each person is wearing a mask or not. Studies have shown that combining face detection algorithms with optimized CNN classifiers significantly improves system performance in real-time environments.

In addition, researchers have investigated the use of transfer learning techniques to improve model accuracy[3] when training data is limited. Transfer learning allows pre-trained deep learning models to adapt to new tasks with minimal training time while maintaining high performance. These approaches have made face mask detection systems more practical and scalable for real-world applications.

V. METHODOLOGY

The proposed system is developed using deep learning and computer vision techniques to identify whether a person is wearing a mask. Initially, a dataset containing images of masked and unmasked faces is collected from publicly available sources[10].

The dataset is pre-processed by resizing images, normalizing pixel values, and applying augmentation techniques such as rotation and flipping to improve model performance. The processed data is then used to train a Convolutional Neural Network (CNN) [2] for classification.

For real-time detection, video input is captured through a webcam or surveillance camera. Face detection is performed using OpenCV[4], and the detected facial regions are passed to the trained CNN model. Based on the prediction, the system displays labels indicating “Mask” or “No Mask” along with bounding boxes.

This process is executed continuously, enabling real-time monitoring with minimal delay.

VI. PROPOSED SYSTEM

The proposed system is a deep learning based Real-Time Face Mask Detection model designed to automatically monitor mask compliance in public

areas. It integrates face detection and mask classification into a single intelligent framework. The system captures live video from a webcam or CCTV camera and processes each frame to identify human faces. Once detected, the facial region is extracted and sent to a trained Convolutional Neural Network (CNN) model for classification [2][3]. The CNN model analyses the facial features and predicts whether the person is wearing a mask or not. Based on the prediction, the system displays a bounding box around the face with a label “Mask” (green colour) or “No Mask” (red colour). This visual indication helps authorities quickly identify violations. The system works continuously with minimal delay, ensuring real-time performance even in moderately crowded environments. Additionally, the proposed system is scalable and can be integrated with existing surveillance infrastructure without requiring major hardware changes. It reduces human effort, increases monitoring accuracy, and provides a cost-effective solution for public safety management. This automated approach enhances compliance monitoring and supports smart surveillance applications in hospitals, airports, educational institutions, and other public spaces.

VII. SYSTEM ARCHITECTURE

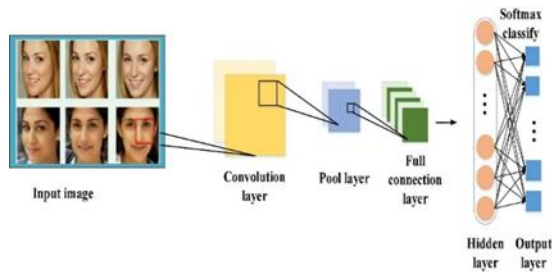


Figure 1: system Architecture

Representation

Figure 1 illustrates the architecture of the proposed system. The process begins with video input from a camera. Each frame undergoes face detection to identify facial regions. The detected faces are pre-processed and passed to the CNN classifier for mask detection. The classification result is displayed with bounding boxes and labels. The integrated system enables real-time detection and compliance monitoring of mask. The above figure illustrates the working architecture of a Convolutional Neural

Network (CNN) used for face mask detection. The process begins with an input image, which is fed into the convolution layer for feature extraction[2]. The CNN automatically learns important visual features from the image, such as edges, textures, and facial patterns. In the Convolution Layer, filters (kernels) are applied to the input image to extract low level features like edges and shapes. These filters slide over the image and generate feature maps that highlight important regions. As the network goes deeper, more complex features such as facial structures and mask patterns are detected. The Pooling Layer follows the convolution layer and reduces the spatial dimensions of the feature maps. This helps in decreasing computational complexity and prevents overfitting. Max pooling is commonly used to retain the most significant features while discarding unnecessary information. After pooling, the extracted features are passed to the Fully Connected Layer. In this stage, the network combines all learned features to make a final prediction. The fully connected layer acts as a traditional neural network and performs classification based on the extracted features. The Hidden Layers process the input features further and apply weights and activation functions to improve learning accuracy. Finally, the Output Layer uses the Softmax activation function[3] to classify the image into categories such as “Mask” and “No Mask.” The Softmax function converts the output into probability values, and the class with the highest probability is selected as the final prediction. Thus, the CNN architecture efficiently extracts features from facial images and accurately determines whether a person is wearing a mask or not.

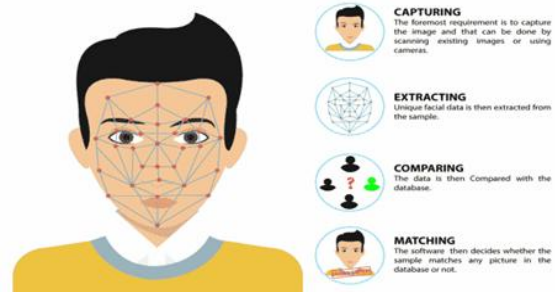


Figure 2: Face Detection Process

The face detection process begins with capturing an image or video frame using a camera, which serves as the input to the system. The captured image is first pre-processed by converting it into grayscale, resizing it, and reducing noise to improve accuracy and speed.

After preprocessing, a face detection algorithm such as Haar Cascade[1], HOG, or a CNN-based model scans the image to identify and locate the face, usually by drawing a bounding box around it. Once the face is detected, important facial features are extracted and converted into numerical data called face encoding. Finally, this encoded data can be compared with stored data for recognition or further analysis, and the system provides the final result based on the matching score.

VIII. RESULTS

The proposed face mask detection system was trained and tested on a labeled dataset[10][11] containing images of individuals with and without masks. The dataset was divided into training and testing sets to evaluate model performance effectively. After training the Convolutional Neural Network (CNN) model, the system achieved high classification accuracy on the test dataset. Performance metrics such as accuracy, precision, recall, and F1-score[3] were used to measure the effectiveness of the model.

The results indicate that the model accurately distinguishes between masked and unmasked faces under different lighting conditions and backgrounds. The confusion matrix analysis shows that the number of misclassifications is minimal, demonstrating the robustness of the trained model. Real-time testing using a webcam further confirmed that the system can detect multiple faces simultaneously and classify them correctly with minimal delay. To evaluate the performance of the proposed system, several experiments were conducted using a labeled dataset containing images of masked and unmasked individuals. The dataset was divided into training and testing subsets to ensure proper model evaluation.

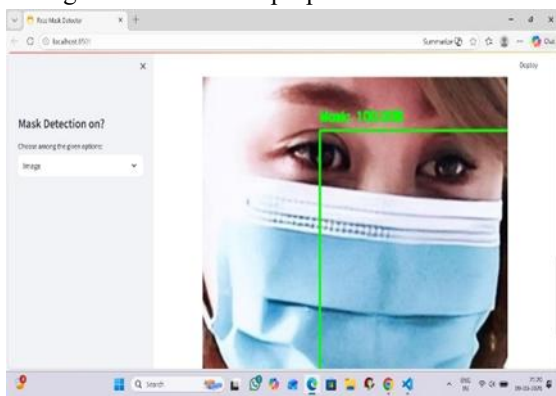


Figure 3: Dataset Process with Mask



Figure 4: Dataset Process Without Mask

Overall, the experimental analysis proves that integrating face detection with CNN based classification provides reliable and efficient real-time mask detection. The system performs well in practical scenarios and can be deployed in public places such as airports, hospitals, educational institutions, and shopping malls to ensure safety compliance. Furthermore, real-time experiments were conducted using a webcam to evaluate the system's performance in live scenarios. The results showed that the system can detect multiple faces simultaneously and classify them correctly with minimal delay. This confirms that the proposed system is suitable for deployment in real-world surveillance environments

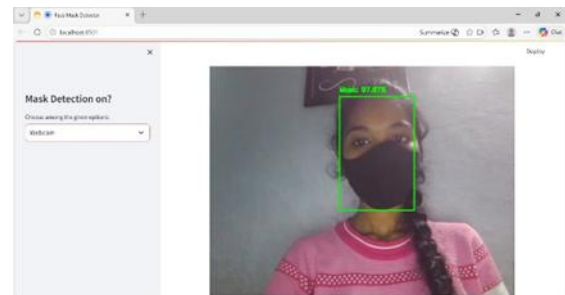


Figure 5: webcam Process with Mask

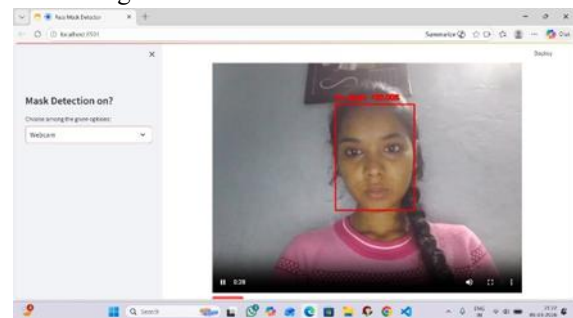


Figure 6: Webcam process Without Mask

VIII. CONCLUSION

This study presents an automated face mask detection system using deep learning techniques [3]. The system

combines face detection with CNN-based classification to identify mask usage accurately. The model was trained using a labelled dataset and demonstrated high performance during testing. Real-time implementation confirmed that the system can detect multiple faces efficiently with minimal delay. The proposed solution reduces manual monitoring efforts and enhances public safety by ensuring compliance with mask regulations. It can be deployed in various public environments such as hospitals, educational institutions, and transportation hubs. Future improvements may include detecting incorrectly worn masks and integrating additional health monitoring features to enhance system capabilities.

IX. FUTURE SCOPE

Although the proposed system performs effectively in detecting face masks, several improvements can be implemented in future research. One possible enhancement is the detection of improperly worn masks, such as masks worn below the nose or chin. Detecting such cases would improve the overall effectiveness of the system.

Another potential improvement is integrating additional health monitoring features such as temperature detection and crowd density monitoring[10]. Combining these features with face mask detection can help create a comprehensive smart surveillance system for public health safety.

Future research may also focus on improving the model's performance in extremely crowded environments and under challenging lighting conditions. Additionally, deploying the system on edge devices and embedded systems can enable large-scale real-time monitoring without requiring powerful computing resources.

X. DISCUSSION

The experimental results obtained from the proposed face mask detection system demonstrate that deep learning techniques can effectively identify whether a person is wearing a mask[10][11] or not in real-time environments. By combining face detection algorithms with a Convolutional Neural Network (CNN) classifier, the system is capable of accurately detecting faces and classifying them into two

categories: "Mask" and "No Mask." The model achieved high accuracy during testing and performed well under different lighting conditions and backgrounds.

One of the major advantages of the proposed system is its ability to process live video streams and detect multiple faces simultaneously. The integration of OpenCV for face detection and the trained CNN model for classification enables the system to operate efficiently with minimal delay. This makes the system suitable for real-time surveillance applications where quick detection and response are required.

During testing, it was observed that the model performs well when faces are clearly visible and properly aligned with the camera. However, certain challenges were also identified. For example, detection accuracy may slightly decrease when faces are partially covered, when lighting conditions are extremely poor, or when individuals are positioned far away from the camera. In crowded environments, overlapping faces may also affect detection performance.

Despite these challenges, the proposed system provides a practical and efficient solution for automated mask compliance monitoring. Compared to traditional manual monitoring systems, the automated approach significantly reduces human effort and improves monitoring accuracy. The system can be easily integrated with existing CCTV infrastructure, allowing organizations to implement intelligent surveillance without requiring major hardware changes.

Overall, the results indicate that deep learning-based face mask detection systems can play an important role in improving public safety and ensuring compliance with health guidelines. With further improvements such as detecting improperly worn masks and optimizing the model for edge devices, the system can become even more effective for large-scale deployment in real-world environments.

REFERENCE

- [1] P. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," IEEE CVPR, 2001.
- [2] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," NIPS, 2012.

- [3] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016.
- [4] A. Rosebrock, “Face Mask Detector with OpenCV, Keras, and TensorFlow,” PyImageSearch, 2020.
- [5] J. Redmon et al., “You Only Look Once: Unified Real-Time Object Detection,” IEEE CVPR, 2016.
- [6] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” IEEE CVPR, 2016.
- [7] A. Howard et al., “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” arXiv, 2017.
- [8] W. Liu et al., “SSD: Single Shot MultiBox Detector,” European Conference on Computer Vision (ECCV), 2016.
- [9] R. Girshick, “Fast R-CNN,” IEEE International Conference on Computer Vision (ICCV), 2015.
- [10] S. Wang et al., “Masked Face Recognition Dataset and Application,” IEEE International Conference on Image Processing, 2020.
- [11] Z. Wang, G. Wang, and B. Huang, “Masked Face Recognition Dataset and Application,” IEEE, 2020.
- [12] World Health Organization, “Advice on the Use of Masks in the Context of COVID-19,” WHO Guidelines, 2020.