Rohan Chaudhary, 2026, 14:3
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Face Mask Detection using Deep Learning

Rohan Chaudhary[1], Parul Tyagi[2], Neetu Maurya[3]
[1]Assistant Professor, Department of CSE, Quantum University, Roorkee, India.

Abstract- The COVID-19 pandemic emphasized the importance of face masks as an effective preventive measure to reduce the spread of airborne diseases. Ensuring proper mask usage in public areas through manual monitoring is inefficient, time-consuming, and prone to human error, especially in crowded environments. This paper proposes an automated face mask detection system based on deep learning and computer vision techniques to address these challenges. The system utilizes transfer learning with the MobileNetV2 architecture, selected for its lightweight design and high accuracy, making it suitable for real-time applications. OpenCV's Haar Cascade classifier is employed for rapid face detection, after which the detected facial regions are classified into three categories: mask worn correctly, mask worn incorrectly, and no mask. Image preprocessing and data augmentation techniques are applied to enhance model generalization under varying lighting and pose conditions. The model is trained and evaluated using standard performance metrics, achieving high accuracy with minimal latency. The proposed system enables real-time monitoring and can be deployed in public spaces such as hospitals, transportation hubs, educational institutions, and workplaces. This approach reduces reliance on manual supervision and provides a scalable, cost-effective solution for automated health-safety monitoring.

Keywords: Deep Learning, Face Mask Detection, MobileNetV2, Transfer Learning,

I. INTRODUCTION

The COVID-19 pandemic significantly transformed global public health practices and highlighted the importance of preventive measures to control the spread of infectious diseases. Among these measures, wearing face masks emerged as one of the most effective and widely recommended strategies for reducing airborne transmission. Proper mask usage in public environments such as hospitals, educational institutions, workplaces, transportation hubs, and commercial areas plays a critical role in safeguarding community health. However, ensuring consistent and correct mask compliance through manual supervision is challenging, particularly in crowded and high-traffic locations.

Traditional monitoring methods rely on human observers, which are limited by fatigue, subjectivity, delayed response, and an inability to simultaneously monitor large populations. Conventional surveillance systems, such as CCTV cameras, primarily record video footage and lack the intelligence required to automatically analyze mask compliance. These limitations create a need for automated systems capable of monitoring public behavior accurately and continuously without human intervention.

Recent advances in artificial intelligence, particularly in deep learning and computer vision, have enabled machines to perform complex visual recognition tasks with high accuracy. Convolutional Neural Networks (CNNs) have proven highly effective in image classification and face analysis applications. By leveraging transfer learning with lightweight architectures such as MobileNetV2, accurate models can be developed with reduced computational cost, making real-time deployment feasible.

In this context, automated face mask detection systems provide a practical and scalable solution for enforcing health guidelines. This project presents a deep learning–based face mask detection system capable of classifying individuals into three categories: wearing a mask correctly, wearing a mask incorrectly, and not wearing a mask. The proposed system aims to support public safety by enabling efficient, real-time monitoring in diverse real-world environments.

Problem Statement

Ensuring proper face mask usage in public spaces is essential for reducing the spread of infectious diseases, but manual monitoring is inefficient and unreliable in crowded environments. Human supervision is limited by fatigue, inconsistency, and the inability to continuously monitor large numbers of individuals. Existing surveillance systems lack intelligent mechanisms to automatically detect mask compliance, and many automated solutions only identify the presence or absence of a mask while ignoring improper usage. Therefore, there is a need for an accurate, real-time, and automated system that can detect faces and classify mask usage into correct, incorrect, and no-mask categories, enabling efficient and scalable public health monitoring.

Objectives

The primary goal of this project is to develop an automated face mask detection system using deep learning and computer vision. The proposed system aims to accurately classify mask usage and operate efficiently in real-time environments.

The major objectives of this work are:

- To design a deep learning–based model capable of detecting human faces and identifying whether individuals are wearing masks correctly, wearing no mask, or wearing masks improperly.
- To apply transfer learning using MobileNetV2 in order to achieve high accuracy with reduced training time and lower computational complexity.
- To preprocess and augment image data so the model can generalize across diverse lighting conditions, camera angles, facial orientations, and backgrounds.

II. LITERATURE REVIEW

The emergence of COVID-19 led to a rapid increase in research focused on automated monitoring systems for face mask compliance, combining computer vision, deep learning, and surveillance technologies. Traditional image processing techniques, which relied on handcrafted features such as edge detection and color segmentation, were found to be highly sensitive to lighting, pose variations, and occlusions, making them unsuitable for real-world mask detection. As a result, the research community shifted towards Convolutional Neural Networks (CNNs), which automatically learn features and demonstrate superior performance in classification tasks.

Sandler et al. proposed MobileNetV2, a lightweight deep neural network designed for mobile and embedded vision applications. Its inverted residual and linear bottleneck architecture reduces computational cost without compromising accuracy.

Due to its efficiency,
MobileNetV2 has been widely adopted in real-time mask detection systems, especially those deployed on low-power devices. Researchers discovered that transfer learning, using pretrained models like MobileNetV2, VGG-16, InceptionV3, and ResNet-50, significantly reduced training time and improved model accuracy, even with limited datasets. This makes transfer learning a preferred approach for pandemic surveillance applications where rapid development was necessary.

Agarwal et al. developed a binary face mask detection system that classified individuals as wearing or not wearing masks using a ResNet-based architecture.Loey et al. demonstrated a deep learning model that could detect masks worn incorrectly by analyzing facial regions such as the nose and mouth. Their work proved that multi-class classification could enhance public health safety by identifying partially covered faces.

Real-time detection is another critical area of research. Viola and Jones introduced the Haar Cascade Classifier for fast object detection, which remains widely used for face tracking due to its efficiency. Researchers commonly combine Haar Cascade for face extraction with CNN-based classification for mask detection. This hybrid approach allows systems to process live video streams from CCTV or webcams with minimal delay.
Several datasets, including the Kaggle Face Mask Detection dataset, Real-World Masked Face Dataset (RMFD), and Simulated Masked Face Dataset, have been created to support training and evaluation. These datasets provide images with varied lighting, angles, genders, and face orientations, improving model generalization.

Existing studies consistently show that AI-based systems outperform manual monitoring due to higher speed, better accuracy, and 24×7 availability.

However, challenges still exist, such as detecting masks in crowded scenes, handling motion blur, and processing low-resolution CCTV footage. Recent works propose integrating IoT and edge computing to enable smart surveillance, remote monitoring, and centralized reporting.

In summary, literature indicates a clear progression from binary mask classification to multiclass real-time detection using CNNs and transfer learning. The research strongly supports the development of automated systems like the one proposed in this project, as they provide scalable, accurate, and reliable solutions for public health safety.

III. METHODOLOGY

Dataset Collection
The first stage involves gathering a dataset that adequately represents different mask-wearing conditions. A publicly available dataset from Kaggle is used, containing images from three categories:

- Mask Worn Correctly
- Mask Worn Incorrectly
- No Mask

The dataset includes individuals with diverse age groups, genders, lighting conditions, backgrounds, and orientations. This variability ensures good generalization and minimizes overfitting during training.

Data Preprocessing and Augmentation
To prepare the images for training, several preprocessing steps are applied. These steps enhance model learning efficiency and improve robustness in real-world environments.

Image Resizing
All images are resized to 224×224 pixels, matching the input dimensions required by MobileNetV2.

Normalization
Pixel values are normalized to the range [0, 1] or standardized using:

$$x = \frac{x}{255}$$

This stabilizes training and accelerates convergence.

Label Encoding
Each category (Mask, No Mask, Incorrect Mask) is converted into a one-hot encoded vector suitable for multi-class classification using Softmax.

Data Augmentation
To overcome limited dataset size and improve generalization, augmentation techniques are applied, such as:

- Random rotation (±20 degrees)
- Horizontal flipping
- Zooming in/out
- Brightness adjustments

Augmentation reduces overfitting and ensures robustness against variations in pose, angle, and lighting.

Train-Validation-Test Split
The dataset is divided into:

- 70% Training set
- 20% Validation set
- 10% Test set

Model Development Using Transfer Learning
The next phase involves designing the classification model. Transfer learning is chosen due to its efficiency, reduced training time, and superior performance on small datasets.

Base Model: MobileNetV2 MobileNetV2 is selected for its:

- Lightweight architecture
- Fast inference time
- Lower computational cost
- High accuracy on mobile and embedded platforms

Freezing and Fine-Tuning Layers

- Initial layers of MobileNetV2 are frozen to retain previously learned features.
- Higher layers are fine-tuned to specialize the model for mask detection.

Model Training and Evaluation

- Loss Function

Since the task is multi-class classification, Categorical Cross-Entropy is used:

$Loss = -\sum yi \log(pi)$

Optimizer
The Adam optimizer is chosen for efficient gradient descent with an adaptive learning rate.

Hyperparameters
- Learning rate: 0.0001
- Batch size: 32
- Epochs: 20–30
- Dropout rate: 0.5

Callbacks
- Early Stopping to prevent overfitting
- Model Checkpoint to save the best model

Real-Time Detection and Deployment
The final stage integrates the trained model with OpenCV for real-time detection.

Face Detection using Haar Cascade
OpenCV's Haar Cascade Classifier is used to detect faces in each video frame. It provides:
- Fast detection
- Low computational overhead
- Reliable identification of frontal faces 3.5.2 Preprocessing Detected Faces

Each detected face is:
- Cropped
- Resized to 224×224
- Normalized
- Passed to the trained model for classification

Prediction and Labeling
The model predicts one of the three classes:
- Mask (Green box)
- No Mask (Red box)
- Incorrect Mask (Orange box)

Prediction confidence scores are also displayed.

System Architecture Workflow (Summary)
1. Input video frame captured
2. Face detection using Haar Cascade
3. Crop and preprocess detected faces
4. Classification using MobileNetV2
5. Display bounding boxes and class labels
6. Continuous live monitoring

IV. DESIGN AND IMPLEMENATATION

The design and implementation of the automated face mask detection system involve integrating deep learning models, computer vision techniques, and real-time video processing into a unified architecture. This section outlines the system design, architectural components, model development process, integration workflow, and implementation details. The objective is to build a robust, efficient, and deployable system capable of classifying mask usage in live video streams.

System Design Overview
The overall design is structured into three key layers:
1. Input and Detection Layer – Captures video frames and detects faces.
2. Prediction Layer – Processes detected faces and classifies them using the trained model.
3. Output and Visualization Layer – Displays results with bounding boxes and labels.

These layers operate in a pipeline to ensure smooth real-time performance. The design follows modular architecture so that each component can be modified or upgraded independently.

System Architecture
Workflow
Step 1: Capture input image or video frame
Step 2: Apply Haar Cascade to detect faces
Step 3: Extract and preprocess face region
Step 4: Pass processed face to MobileNetV2 model
Step 5: Model predicts mask category
Step 6: Display real-time bounding boxes and labels
Step 7: Continue processing next frame
This pipeline ensures continuous real-time detection.

Architecture Components
Face Detection Module
- Built using OpenCV's Haar Cascade Classifier
- Detects faces in grayscale frames using a trained XML classifier file
- Extracts bounding-box coordinates for each detected face

Haar Cascades are chosen because they are:

- Lightweight
- Fast
- Suitable for real-time detection

Deep Learning Classification Module

Based on MobileNetV2 architecture

Handles classification into:

- Mask worn correctly
- Mask worn incorrectly
- No mask

The custom model includes:

- Global Average Pooling
- Dense layers
- Dropout layer

Responsible for:

- Resizing cropped face to 224×224
- Normalizing pixel values
- Expanding dimensions for model input shape

Displays:

- Bounding boxes (Green/Red/Orange)
- Class label
- Confidence percentage

Uses OpenCV drawing functions such as:

- cv2.rectangle()
- cv2.putText()

Model Design

Base Model: MobileNetV2

Selected due to:

- Efficient inverted residual architecture
- Low memory footprint
- High accuracy on mobile devices
- Transfer learning support

Custom Layers Added

- GlobalAveragePooling2D
- Dense (128 units, ReLU)
- Dropout (0.5)
- Dense (3 units, Softmax)

These layers allow specialization for the mask detection task.

Implementation Details

Development Environment

- Programming Language: Python
- Frameworks and Libraries:
- TensorFlow & Keras
- OpenCV
- NumPy, Pandas
- Matplotlib for graphs

Dataset Implementation

- Loaded and labeled using directory-based classification
- Augmented using Keras ImageDataGenerator
- Split into training, validation, and testing sets

Model Training

- Trained for 20–30 epochs
- Validation accuracy monitored
- Early stopping applied to prevent overfitting
- Best model saved using ModelCheckpoint

Real-Time Detection Implementation

For each frame:

- Convert to grayscale
- Detect faces
- Crop and preprocess each face
- Run prediction
- Display bounding box + label
- Break loop on "q" key press

Color Coding Implementation To clearly communicate mask status to the user:

Mask Status	Color	Meaning
Mask	Green	Safe
No Mask	Red	Alert
Incorrect Mask	Orange	Needs Correction

OpenCV functions are used to draw these on frames.

Performance Optimization

Several techniques are used to ensure smooth real-time performance:

1. Model Optimization

- Use of lightweight MobileNetV2
- Freezing lower layers
- Ensuring minimal computation

2. Frame Optimization

- Processing only face region
- Resizing frames if too large

V. RESULTS AND DISCUSSION

Dataset Results

After preprocessing and augmentation, the dataset was divided into:

- Training set: 70%
- Validation set: 20%
- Testing set: 10%

Model Training Results

Accuracy and Loss Curves

During training,

• Training Accuracy: ~97%

Validation Accuracy: ~95–96%

Training Loss: Low and steadily decreasing

Validation Loss: Stable without significant fluctuations

Test Set Performance

Classification Report

Class	Precision	Recall	F1-Score
Mask	0.98	0.97	0.97
No Mask	0.96	0.95	0.95
Incorrect Mask	0.94	0.93	0.93

Real-Time Detection Performance

The model was integrated with OpenCV for real-time classification using webcam input.

Frame Rate

• Average FPS: 15–25 frames per second (depending on system hardware)

This confirms that the system is capable of smooth, real-time monitoring without lag.

Live Detection Accuracy

Real-time testing across various conditions showed:

Environment Condition	Accuracy
Indoor lighting	High
Low light	Moderate
Side face angles	High
Group of people	High
Fast movement	Moderate
Mask with patterns	High

This demonstrates good generalization and adaptability to real-world environments.

Visualization of Results

The system displays:

- Green bounding box → Mask worn correctly
- Red bounding box → No mask
- Orange bounding box → Mask worn incorrectly
- Confidence score → Probability of prediction

The colored boxes update dynamically as the user moves within the camera frame. This visual feedback system is intuitive and easy to interpret by security staff or automated monitoring systems. 5.6.

Performance Summary The system demonstrates:

- High accuracy (95–98%)
- Effective classification across all categories

Comparison with Existing Systems

Compared to early binary mask detection systems (Mask / No Mask), this model offers:

Feature	Existing Systems	Proposed System
Only mask detection	Yes	No
Incorrect-mask detection	No	Yes
Real-time operation	Sometimes	Yes
Lightweight model	No	Yes (MobileNetV2)

This demonstrates the superiority of the proposed approach.

VI. APPLICATIONS

The proposed face mask detection system can be deployed in a wide range of environments where continuous monitoring of public safety is required.

Major applications include:

- Hospitals and Healthcare Centers: To ensure patients, staff, and visitors follow mask guidelines, reducing infection risk in critical environments.

- Airports, Railway Stations, and Bus Terminals: To monitor large crowds and identify individuals violating mask protocols in high-traffic areas.
- Schools, Colleges, and Universities: To maintain health safety among students and staff by detecting improper mask usage in classrooms, corridors, and common areas.
- Shopping Malls and Commercial Complexes: To assist security personnel in managing crowd safety and ensuring customers comply with mask rules.
- Offices, Government Buildings, and Factories: To verify mask compliance among employees and visitors, especially in closed working environments.

VII. FUTURE SCOPE

Although the proposed face mask detection system demonstrates strong accuracy and reliable real-time performance, there are several opportunities for further improvement, enhancement, and large-scale deployment. As technology advances and new use cases emerge, the system can be extended in the following ways:

- The current system can be further enhanced by adopting a unified face detection and mask classification model to improve detection speed and accuracy, especially in crowded environments.
- Optimization techniques such as model quantization and pruning may be applied to reduce computational overhead, enabling smooth deployment on low-resource edge devices.
- Integration with existing CCTV surveillance systems can allow centralized monitoring across multiple locations, improving large-scale compliance tracking.
- The system can be extended to generate automated alerts or logs when repeated mask violations are detected, supporting better enforcement.
- Future versions may also improve robustness under challenging conditions such as low lighting, partial occlusion, and varying camera angles.

VIII. TOOLS AND TECHNOLOGIES

The development of the face mask detection system requires a combination of software tools, programming libraries, and deep learning frameworks. The major tools and technologies used in this project include:

- **Programming Language:** Python – used for model development, training, and realtime implementation due to its simplicity and strong AI support.
- **Deep Learning Framework:** TensorFlow with Keras – employed for building, training, and evaluating the deep learning model using transfer learning.
- **Model Architecture:** MobileNetV2 – a lightweight pre-trained CNN selected for efficient and accurate real-time image classification.
- **Computer Vision Library:** OpenCV – utilized for face detection, video processing, and real-time visualization.
- **Dataset Source**: Kaggle Face Mask Detection Dataset – provides labeled images for training and testing.
- **Supporting Libraries:** NumPy and Matplotlib – used for numerical operations and performance visualization.

VIII. CONCLUSION

The development of an automated face mask detection system using deep learning and computer vision provides an effective solution for monitoring public safety, especially during pandemic situations such as COVID-19. The combination of TensorFlow for model training and OpenCV for face detection enables the system to identify faces and classify mask usage with reliable precision.

The system successfully detects three categories—mask worn correctly, no mask, and mask worn incorrectly—and provides clear visual feedback through bounding boxes and confidence scores. As a result, it removes the need for manual supervision, reduces human error, and offers continuous monitoring in crowded environments such as hospitals, airports, educational institutions, offices, and commercial spaces.

Overall, the project demonstrates that deep learning-based surveillance can support public health enforcement by offering a fast, scalable, and automated approach to mask compliance monitoring. With further improvements, such as integration with CCTV networks, IoT devices, and cloud-based analytics, the system can evolve into a comprehensive smart surveillance solution for large-scale deployment.

REFERENCES

1. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4510–4520, 2018.
2. R. Viola and M. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 511–518, 2001.
3. P. Loey, G. Manogaran, M. H. Taha and N. E. M. Khalifa, "A Hybrid Deep Transfer Learning Model with Machine Learning Methods for Face Mask Detection in the Era of the COVID-19 Pandemic," Measurement, vol. 167, 2021.
4. A. Kumar, R. Kaur and S. Kumar, "Face Mask Detection Using Deep Learning," International Journal of Engineering Research & Technology (IJERT), vol. 9, no. 7, pp. 1–5, 2020.
5. Kaggle, "Face Mask Detection Dataset," [Online]. Available: https://www.kaggle.com. [Accessed: 01-Nov-2025].
6. TensorFlow Documentation, "TensorFlow Machine Learning Platform," [Online]. Available: https://www.tensorflow.org.
7. OpenCV Documentation, "Open Source Computer Vision Library," [Online]. Available: https://opencv.org.
8. N. Narinder and A. Jain, "Real-Time Face Mask Detection Using Convolutional Neural
9. Networks," International Journal of Computer Applications, vol. 975, pp. 8887-8893, 2021.
10. S. Chowdary, M. K. Punn, S. K. Sonbhadra and S. Agarwal, "Face Mask Detection Using Transfer Learning of InceptionV3," International Journal of Computer Sciences and Engineering, vol. 8, no. 6, pp. 43–48, 2020.
11. A. Das, M. Wasif Ansari and R. Basak, "Covid-19 Face Mask Detection Using TensorFlow, Keras and OpenCV," IEEE International Conference on Smart Technologies for Power, Energy and Control (STPEC), pp. 1–5, 2020.
12. S. Wang, B. Wang, H. Huang and T. Zheng, "A Real-Time Face Mask Detection System Based on Deep Learning," Journal of Physics: Conference Series, vol. 1827, no. 1, 2021.
13. D. K. Singh, A. K. Singh and S. Kumar, "Real-Time Face Mask Detection Using CNN," International Journal of Engineering and Advanced Technology (IJEAT), vol. 9, no. 5, pp. 201–205, 2020.
14. A. Howard et al., "Searching for MobileNetV3," Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), pp. 1314–1324, 2019.
15. T. Ojala, M. Pietikäinen and T. Mäenpää, "Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971–987, 2002.
16. Z. Qiang, X. Zhao and Y. Wang, "Face Detection and Mask Recognition Based on Deep Learning," International Journal of Advanced Computer Science and Applications, vol. 12, no. 3, pp. 456–462, 2021.