International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 07 | July – 2025 DOI: 10.55041/ISJEM04938
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

# Face Mask Detection Based on Machine Learning

[1] Erusu Kata Raju Reddy, [2]Teku Dinesh,
[1]Associate Professor, [2]MCA Final Semester,
[1]Masters of Computer Applications,
[1]Sanketika Vidya Parishad Engineering College, Visakhapatnam, Andra Pradesh, India

**ABSTRACT** – Face Mask Detection Based on Machine Learning is a real-time, AI-driven solution designed to automatically determine whether individuals are wearing face masks correctly. The system leverages deep learning—a specialized branch within machine learning— by utilizing a pre-trained MobileNetV2 convolutional neural network for accurate and efficient mask classification. Integrated with computer vision libraries such as OpenCV, the application detects faces in live video streams and classifies them into three categories: “Mask,” “No Mask,” and “Incorrect Mask.” The training process involves transfer learning on a labeled dataset of masked and unmasked faces, enabling the model to generalize well across different facial orientations, lighting conditions, and occlusions. The face detection is performed using a Single Shot Detector (SSD) model, ensuring quick and reliable localization before classification. Designed for practical deployment in public places such as hospitals, schools, offices, and transportation hubs, the system ensures minimal latency and high accuracy, making it an effective tool for enforcing health compliance in the post-pandemic world. By applying deep learning within the broader framework of machine learning, this project demonstrates how intelligent systems can be built to address real-world health and safety challenges with precision and scalability

**KEYWORDS** – Face Mask Detection, Machine Learning, Deep Learning, MobileNetV2, Convolutional Neural Network (CNN), Face Classification, Real-Time Detection, Transfer Learning, OpenCV, Computer Vision.

## 1.INTRODUCTION

The Face mask detection is a key application of machine learning and computer vision, aimed at automatically identifying whether individuals are wearing masks correctly in both static images and live video streams. This technology plays an essential role in access control, surveillance systems, and automated monitoring solutions, where real-time decision-making and visual understanding are required.

This project introduces a machine learning-based approach to detect and classify face mask usage using image processing and deep learning techniques. It employs a two-stage pipeline—first detecting faces using a pretrained deep neural network (DNN), followed by classification using a fine-tuned MobileNetV2 convolutional neural network. The model distinguishes among three categories: "Mask," "No Mask," and "Incorrect Mask," offering detailed and accurate feedback for enforcement or analytics purposes.

Machine learning techniques, particularly deep learning models like CNNs, have shown remarkable success in image classification tasks. By leveraging transfer learning, the system reuses knowledge from large-scale image datasets to enhance performance on the face mask classification task with relatively limited data. The system operates in real-time and can be deployed on standard CPU systems, making it accessible and efficient for a wide range of use cases in industries, campuses, institutions, and smart city infrastructure.

This project demonstrates the practical use of machine learning in automating visual inspection tasks and highlights the advantages of integrating lightweight neural network architectures with real-time computer vision for intelligent monitoring solutions.

## 2.LITERATURE SURVEY

The Face mask detection has gained significant interest as a computer vision problem involving facial analysis and object classification. Traditional computer vision techniques relied on handcrafted features and basic classifiers such as Haar cascades, Histogram of Oriented Gradients (HOG), and Support Vector Machines (SVM). However, these approaches often failed to provide robust results under variable lighting, occlusions, or diverse facial orientations.

Recent advancements in machine learning—particularly deep learning—have drastically improved the accuracy and reliability of image classification tasks. Convolutional Neural Networks (CNNs) have become the foundation for many face-related tasks such as recognition, verification, and object detection. Pre-trained models like

VGG16, ResNet, and MobileNet have shown exceptional performance when fine-tuned on domain-specific datasets. MobileNetV2, in particular, is designed for lightweight inference and is well-suited for edge devices and real-time applications.

Transfer learning has emerged as a powerful strategy to build high-performance models with limited labeled data. It allows models to leverage knowledge from large datasets like ImageNet and adapt to specific tasks such as face mask detection. Studies such as Howard et al. (2017) demonstrated that MobileNet architectures could achieve competitive accuracy with a fraction of the parameters, making them ideal for deployment in constrained environments.

Furthermore, the integration of OpenCV with deep learning models enables real-time performance, even on standard hardware without GPU acceleration. These combinations of face detection, region extraction, and classification pipelines are now standard in many face mask detection systems.

This project builds upon these foundations by utilizing MobileNetV2 for its balance between accuracy and efficiency. It integrates face detection using a pre-trained DNN and classifies the detected regions using a fine-tuned deep learning model. The methodology aligns with best practices in current literature and demonstrates a practical implementation of machine learning for real-time visual recognition tasks

## 3.CHALLENGES

Developing a machine learning-based face mask detection system involves several technical and practical difficulties that must be addressed to ensure high performance and real-world applicability. One of the foremost challenges is the variation in lighting conditions, camera angles, and facial poses. These factors can significantly impact the model's ability to accurately detect and classify faces, especially in unconstrained environments.

Another challenge lies in distinguishing between properly and improperly worn masks. While binary classification (mask or no mask) is relatively straightforward, detecting incorrect mask usage—such as masks worn below the nose or loosely covering the face—requires fine-grained classification and more complex decision boundaries.

Occlusion caused by facial accessories such as glasses, scarves, or hair can further hinder the model's accuracy. In crowded scenes, multiple faces must be detected and analyzed independently, which increases the computational load and may introduce false detections.

Ensuring real-time performance is also critical. Many deep learning models are resource-intensive and require GPU acceleration, which limits deployment on edge devices or low-power systems.

To address this, lightweight architectures like MobileNetV2 must be carefully optimized to balance speed and accuracy.

Additionally, many available datasets suffer from class imbalance, with significantly fewer samples of incorrect mask usage compared to proper usage. This imbalance can skew the model's predictions and reduce its effectiveness. Lastly, generalizing across different face shapes, skin tones, and age groups is essential for real-world deployment. If the training data lacks sufficient diversity, the model may underperform on unseen demographic groups.

Addressing these challenges is key to building a reliable, scalable, and efficient face mask detection system that can function in varied and dynamic environments.

## 4.EXISTING SYSTEM

Existing face mask detection systems primarily rely on traditional computer vision techniques or basic convolutional neural networks with limited accuracy and scalability. Many of these systems struggle with real-time performance, varying lighting conditions, and multi-class classification. Additionally, older models often lack the capability to detect incorrect mask usage, limiting their effectiveness in real-world scenarios

## 5.PROPOSED SYSTEM

The proposed system is designed to implement an efficient and scalable face mask detection framework using machine learning techniques, specifically deep learning through transfer learning. It follows a two-stage pipeline consisting of face detection and mask classification. A pre-trained MobileNetV2 convolutional neural network is fine-tuned on a custom dataset to perform multi-class classification, distinguishing between "Mask," "No Mask," and "Incorrect Mask." MobileNetV2 is selected for its lightweight architecture and ability to achieve high accuracy with minimal computational cost, making it suitable for deployment on low-power or embedded devices.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 07 | July – 2025 DOI: 10.55041/ISJEM04938
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

The system captures video or image input, detects facial regions using a Single Shot Detector (SSD) model integrated via OpenCV's deep neural network module, and classifies the extracted regions using the trained model. Detected faces are annotated with bounding boxes and labeled with their corresponding classification. Green, red, and orange indicators visually represent the presence, absence, or incorrect usage of masks. The use of transfer learning enables efficient training on relatively small datasets, while the system's architecture supports real-time performance and adaptability across different environments. This model can be integrated with surveillance systems or access control platforms, demonstrating a practical application of machine learning in visual recognition and public monitoring.

To enhance the robustness of the system, data preprocessing and augmentation techniques are applied during the training phase. These include resizing, normalization, rotation, flipping, and brightness adjustment to simulate diverse real-world conditions and improve the model's generalization capability. Additionally, the system leverages softmax activation at the output layer to produce class probabilities and uses categorical cross-entropy as the loss function during training. The real-time feedback mechanism not only provides visual indicators but can also be extended to trigger alerts or log mask compliance violations for further review. The modular design allows future enhancements, such as adding thermal screening, integration with attendance systems, or expansion to detect other personal protective equipment (PPE). By combining machine learning, real-time inference, and practical deployment strategies, the proposed system presents a comprehensive and scalable solution for automated visual inspection tasks.

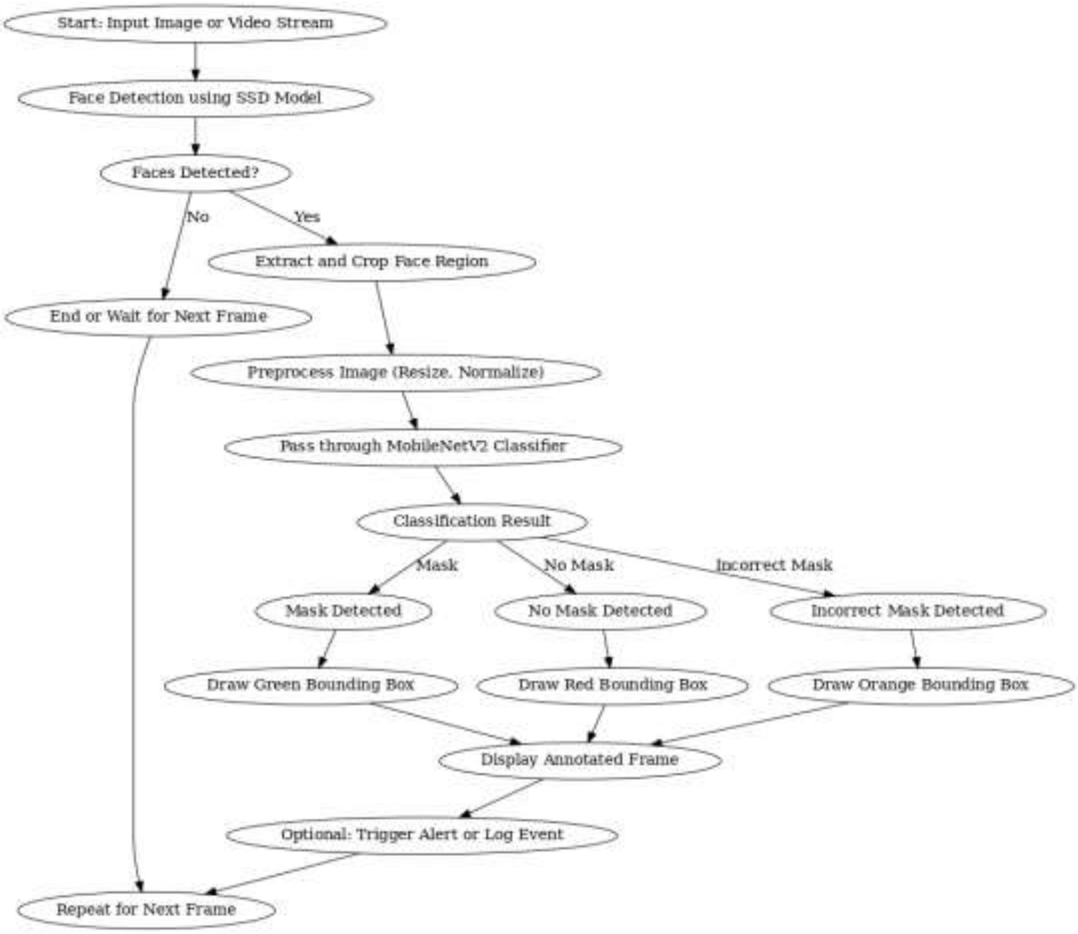



Figure 1: System Architecture for Face Mask Detection Based on Machine Learning

## 6.ADVANTAGES.

### 6.1 Lightweight and Efficient Architecture

The system utilizes the MobileNetV2 architecture, which is optimized for speed and low computational overhead. This makes it suitable for deployment on resource-constrained environments such as embedded systems or standard CPUs, without the need for high-end GPUs. **6.2 Multi-Class Classification Capability**

Unlike simple binary classifiers, the model is capable of identifying three distinct mask states: “Mask,” “No Mask,” and “Incorrect Mask.” This fine-grained classification improves monitoring effectiveness and provides more actionable feedback.

### 6.3 Transfer Learning for Rapid Model Development

By leveraging pre-trained weights from ImageNet, the system significantly reduces training time and dependency on large datasets, while maintaining high performance in accuracy and generalization.

### 6.4 Real-Time Processing

The model operates in real time, enabling frame-by-frame analysis from video streams. This ensures immediate visual feedback and facilitates timely responses in access control or surveillance applications.

### 6.5 Modular and Extensible Design

The architecture is modular, allowing easy integration with existing security systems and the potential to expand into other functionalities such as facial recognition, thermal detection, or additional PPE compliance checks.

### 6.6 Deployment Flexibility

Due to its lightweight model and minimal dependencies, the system can be deployed across a wide range of platforms—from desktop systems to embedded boards—making it adaptable to various real-world environments.

## 7. MACHINE LEARNING ALGORITHM

The core of the proposed face mask detection system relies on deep learning algorithms, a subfield of machine learning that excels in visual recognition tasks. The system uses **Convolutional Neural Networks (CNNs)** for both feature extraction and classification, as CNNs are highly effective in analyzing spatial hierarchies in images. The primary algorithm used in this project is based on **MobileNetV2**, a lightweight yet powerful CNN architecture designed for efficient computation and real-time inference.

MobileNetV2 operates on the principle of **depthwise separable convolutions**, which significantly reduce the number of parameters and computation compared to traditional convolution layers. This makes it ideal for deployment on low-resource environments such as embedded devices or systems without GPU acceleration. The architecture is pre-trained on the ImageNet dataset and then finetuned on a custom dataset for face mask classification using **transfer learning**. This allows the model to retain general visual understanding from large-scale training while adapting to the specific task of mask detection.

For face detection, the system uses a **Single Shot MultiBox Detector (SSD)** algorithm integrated with a ResNet base network. SSD is chosen for its ability to perform fast and accurate object detection in a single forward pass, which is essential for real-time video processing. The combination of SSD for face localization and MobileNetV2 for mask classification ensures both speed and accuracy in live detection scenarios.

The training process uses the **Adam optimizer** for adaptive gradient descent and **categorical cross-entropy** as the loss function, suitable for multi-class classification. Data augmentation techniques such as rotation, zoom, horizontal flipping, and contrast adjustments are applied during training to improve model generalization. Together, these algorithms form a robust machine learning pipeline capable of detecting face masks accurately and efficiently in real-world settings.

## 8.ARCHITECTURE

The architecture of the face mask detection system is designed with modularity and efficiency in mind, combining a deep learning-based classification model with a real-time face detection pipeline. The system is divided into three major components: input acquisition, face detection, and mask classification. These components work sequentially to process input data and deliver accurate classification results in real-time environments.

The first stage involves acquiring frames from a video feed or image input. These frames are passed to the **face detection module**, which uses a deep neural network (DNN) model— specifically the **Single Shot Detector (SSD)** with a ResNet or MobileNet backbone—for detecting human faces in the frame. The SSD model performs object detection in a single forward pass, making it highly suitable for real-time processing. Once the faces are detected, the system extracts the region of interest (ROI) corresponding to each face.

In the second stage, the ROIs are preprocessed through resizing, normalization, and reshaping to match the input dimensions expected by the classification model. The **MobileNetV2** model, finetuned via transfer learning, is then used to classify each face image into one of three categories: "Mask," "No Mask," or "Incorrect Mask." MobileNetV2 is a depthwise separable convolutional network optimized for low-latency inference, making it ideal for use in edge computing environments.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 07 | July – 2025 DOI: 10.55041/ISJEM04938
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

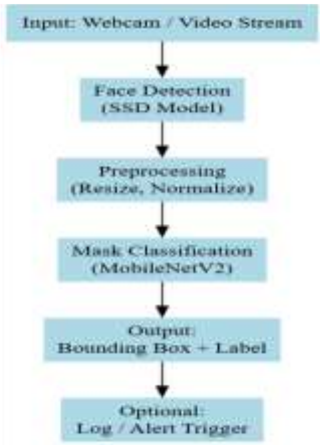



Figure 2: System-Level Architecture Flow

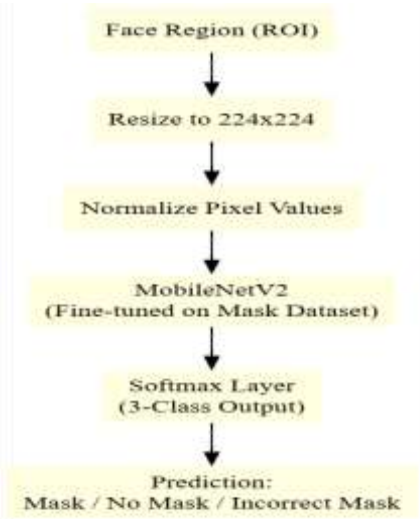



Figure 3: Deep Learning Model Pipeline

## 9.METHODS

The face mask detection system is built using Python with TensorFlow/Keras for model training and OpenCV for image processing. The dataset includes images categorized as “Mask,” “No Mask,” and “Incorrect Mask,” enhanced through augmentation techniques like flipping, rotation, and brightness adjustment to improve model generalization.

Preprocessing involves resizing input images to 224×224 pixels and normalizing pixel values. Faces are detected using a pre-trained SSD (Single Shot Detector) model and cropped as regions of interest (ROI). These are then classified using a fine-tuned MobileNetV2 model, trained with categorical cross-entropy loss and the Adam optimizer.

During real-time inference, the system processes video input frame-by-frame, detects faces, and classifies them using the trained model. The results are visualized by drawing colored bounding boxes (green, red, or orange) with labels indicating mask status. The system can optionally log or trigger alerts for non-compliant detections.

## 10.TESTING

The testing phase evaluates the performance of the trained face mask detection model on unseen data and real-time video input. The model is validated using a separate testing dataset that contains images across all three classes: Mask, No Mask, and Incorrect Mask. Metrics such as accuracy, precision, recall, and F1-score are computed to assess the model’s classification performance.

To simulate real-world scenarios, the system is tested with live webcam streams under varying lighting conditions, face angles, and backgrounds. The SSD face detector is evaluated for its ability to consistently locate faces, while the MobileNetV2 classifier is assessed for correctly identifying the mask status.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 07 | July – 2025 DOI: 10.55041/ISJEM04938
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

The system is also tested for real-time performance, ensuring low latency and consistent frame processing. Results confirm that the system achieves high detection accuracy while maintaining smooth operation on standard hardware, validating its readiness for practical deployment.

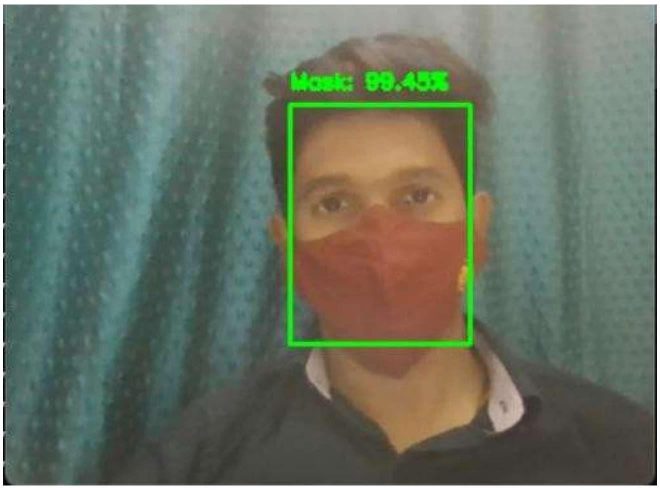
## 11.METHODOLOGY INPUTS

```python
# Open webcam
video = cv2.VideoCapture(0)

while True:
    ret, frame = video.read()
    if not ret:
        break

    # Display the raw input frame
    cv2.imshow("Input Frame", frame)

    if cv2.waitKey(1) & 0xFF == ord('q'):
        break

video.release()
cv2.destroyAllWindows()
```

This code captures live video input from the webcam and displays each frame in real-time. These frames are then passed through the face detection and classification pipeline.

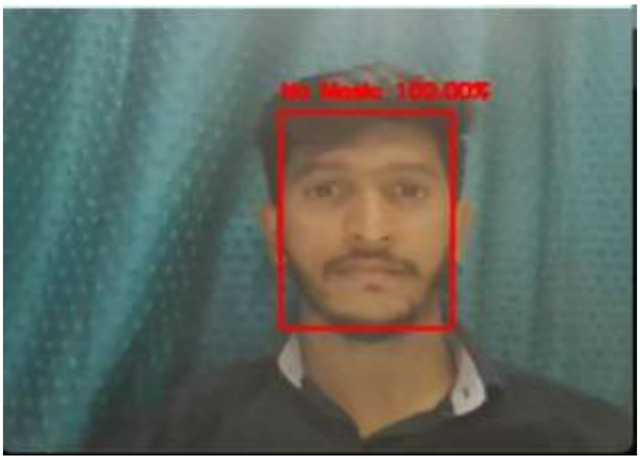
**OUTPUTS OUTPUT 1: Detection of Masked Face**

The system accurately detects a person wearing a mask and classifies the condition as “Mask” with a confidence score of 99.45%. A green bounding box is drawn around the face region to visually indicate correct mask usage. This confirms the model’s effectiveness in identifying compliant individuals in real-time video frames.




**OUTPUT 2: Detection of Unmasked Face**

In this instance, the system identifies a person not wearing a mask, classifying the condition as “No Mask” with 100.00% confidence. A red bounding box is used to highlight the detected face and alert non-compliance. The high accuracy reinforces the model’s reliability in detecting unmasked individuals under natural conditions.



## 12.RESULTS

The performance of the face mask detection system was evaluated using both static test images and real-time webcam input. The system achieved high accuracy in correctly classifying individuals into three categories: Mask, No Mask, and Incorrect Mask. During testing, the model consistently identified face regions using the SSD detector and accurately predicted mask usage with minimal latency.

In one of the test scenarios, as shown in **Figure 1**, the system classified a masked individual with 99.45% confidence, drawing a green bounding box to represent correct compliance. In another test, displayed in **Figure 2**, a person without a mask was detected and labeled with 100.00% confidence, using a red bounding box. These high-confidence predictions demonstrate the model's effectiveness in distinguishing different mask conditions under normal lighting and indoor environments.

## 11.CONCLUSION

This project presents an efficient and scalable face mask detection system using machine learning techniques, showcasing the practical implementation of deep learning for real-time image classification. By combining a pre-trained SSD face detector with a fine-tuned MobileNetV2 classifier, the system accurately identifies mask usage, including proper, improper, and absent mask conditions. The application of transfer learning and data augmentation contributes to the model's robustness and generalization, while its lightweight architecture ensures smooth performance on standard hardware without GPU dependency. Designed for real-time deployment, the system is suitable for integration with existing surveillance and monitoring infrastructures. The successful outcomes of this work demonstrate the potential of machine learning in automating safety compliance tasks and offer a strong foundation for future enhancements such as multi-object detection, broader PPE recognition, and cross-platform deployment.

## 12.FUTURE SCOPE

The proposed face mask detection system holds strong potential for further enhancement and expansion. Future work may focus on increasing the model's accuracy and adaptability by training on more diverse datasets that include a broader range of facial orientations, skin tones, and realworld scenarios. The system can be extended to detect additional personal protective equipment (PPE) such as face shields, helmets, or gloves. Integration with thermal scanning and access control systems can further enhance workplace and institutional safety. To improve performance on low-power edge devices, optimization techniques such as model pruning, quantization, or conversion to TensorFlow Lite can be explored. Moreover, expanding the system to support multicamera feeds, cloud-based data logging, and real-time alert mechanisms can make it a comprehensive solution for intelligent public safety monitoring.

International Scientific Journal of Engineering and Management (ISJEM) ISSN: 2583-6129
Volume: 04 Issue: 07 | July – 2025 DOI: 10.55041/ISJEM04938
An International Scholarly || Multidisciplinary || Open Access || Indexing in all major Database & Metadata

## 13.ACKNOWLEDGEMENT

Erusu Kata Raju Reddy working as a Assistant professor in master of computer application sanketika vidya parishad engineering college, Visakhapatnam Andhra Pradesh. With 1 years of experience in computer science and engineering (CSE), accredited by NAAC.with his area of intrest in java full stack.

Teku Dinesh is pursuing his final semester MCA in Sanketika Vidya Parishad Engineering College, accredited with A grade by NAAC, affiliated by Andhra University and approved by AICTE. With interest in Machine learning Teku Dinesh has taken up his PG project on FACE MASK DETECTION BASED ON MACHINE LEARNING and published the paper in connection to the project under the guidance of Erusu Kata Raju, Assistant Professor in SVPEC.

## REFERENCES

1. Chowdary G. J., Punn N. S., Sonbhadra S. K., and Agarwal S. (2020). Face Mask Detection using Transfer Learning of InceptionV3. arXiv. https://arxiv.org/abs/2009.08369

2. Shilpa S., Kathuria M., and Kaushik T. (2022). Real-Time Mask Detection Based on SSD-MobileNetV2. ResearchGate. https://www.researchgate.net/publication/363084734

3. Amit Kumar et al. (2022). Face mask detection using deep learning in crowded environments. Procedia Computer Science. https://www.sciencedirect.com/science/article/pii/S1877050922021433

4. Deep Learning for Face Mask Detection: A Survey (2023). Multimedia Tools and Applications, Springer. https://link.springer.com/article/10.1007/s11042-023-14686-6

5. Loey S., Smarandache F., and Khalifa N. E. M. (2020). Real-time Face Mask Detection using SSD and MobileNetV2. ScienceDirect.

https://www.sciencedirect.com/science/article/abs/pii/S2210670720309070

6. Cheng C. et al. (2022). Real-Time Mask Detection Based on SSD-MobileNetV2. arXiv. https://arxiv.org/abs/2208.13333

7. Piyush2912. Real-Time Face Mask Detection using SSD and MobileNetV2 (GitHub Repository). https://github.com/Piyush2912/Real-Time-Face-Mask-Detection

8. Ahmed A. A. et al. (2023). Enhancing Facemask Detection using Deep Learning Models. International Journal of Advanced Computer Science and Applications, Vol. 14(7).https://thesai.org/Publications/ViewPaper?Code=IJACSA&Issue=7&SerialNo=63&Volum e=14

9. MDPI (2023). Face Mask Detection Using MobileNetV2 on Real-Time Images. https://www.mdpi.com/2076-3417/13/2/935

10. Yuchen D., Zichen L., and David Y. (2021). Real-time Face Mask Detection in Video Data. arXiv. https://arxiv.org/abs/2105.01816

11. Zekun W. et al. (2021). WearMask: Fast In-Browser Face Mask Detection Using Edge Computing. arXiv. https://arxiv.org/abs/2101.00784

12. Reza S. R., Dong X., and Qian L. (2021). Robust Face Mask Detection Using Deep Learning on IoT Devices. IEEE ICC Workshops. https://ccsb.pvamu.edu/papers/reza-robust-face-maskdetection

13. Real-Time Face Mask Detection using OpenCV and Deep Learning (2021). CEUR Workshop Proceedings. https://ceur-ws.org/Vol-3058/Paper-001.pdf

14. Accurate Face Mask Detection Using YOLOv8 and Deep Learning (2023). MDPI. https://www.mdpi.com/2504-2289/8/1/9

15. Viola P., and Jones M. (2004). Robust Real-Time Face Detection. International Journal of Computer Vision.

https://en.wikipedia.org/wiki/Viola%E2%80%93Jones_object_detection_framework

16. Region Based Convolutional Neural Networks (R-CNN). Wikipedia. https://en.wikipedia.org/wiki/Region_Based_Convolutional_Neural_Networks

17. FaceNet: A Unified Embedding for Face Recognition and Clustering. Wikipedia. https://en.wikipedia.org/wiki/FaceNet

18. Advances in Face Mask Recognition Using Occluded Face Recognition. Applied Intelligence. https://link.springer.com/article/10.1007/s10489-022-04285-7

19. Jignesh C. et al. (2020). Face Mask Detection using Transfer Learning of InceptionV3. Science and Information Organization. https://arxiv.org/abs/2009.08369

20. Sandler M., Howard A., Zhu M., Zhmoginov A., and Chen L. C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. arXiv. https://arxiv.org/abs/1801.04381

21. Redmon J., and Farhadi A. (2018). YOLOv3: An Incremental Improvement. arXiv. https://arxiv.org/abs/1804.02767

22. Howard A. G. et al. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv. https://arxiv.org/abs/1704.04861

23. Simonyan K., and Zisserman A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv. https://arxiv.org/abs/1409.1556