IN-SITU FACE MASK DETECTION USING DEEPLEARNING TECHNIQUE

[1] A S Salma Banu
[1]Assistant professor, Department of Electronics and Communication Engineering,
Aalim Muhammed Salegh College of Engineering, AVADI, IAF Chennai

Abstract

After the prisonbreak of the worldwide COVID-19 pandemic, there arises a severe want for protection mechanisms, the mask being the first one. The project's essential aim is to see the presence of a mask on human faces on live streaming video still as in pictures. We've got used deep learning to develop our face detector model. The design used for the item detection purpose is Single Shot Detector (SSD) due to its smart performance accuracy and high speed. Aboard this, we've used basic ideas of transfer learning in neural networks to finally output the presence or absence of a mask in a picture or a video stream. Experimental results show that our model performs well on the check knowledge with 100% and ninety-nine exactitude and recall, severally.

Keywords— Machine learning, Deep learning, computer vision, face mask, OpenCV, TensorFlow, Keras, CNN

I. INTRODUCTION

The World Health Organization (WHO) reported that COVID-19 had infected over 6 million people worldwide, resulting in more than 379,941 deaths. Maintaining social distancing, improving surveillance, and strengthening health systems were key measures to control the pandemic. Recent research from the University of Edinburgh revealed that wearing face masks significantly reduces the risk of COVID-19 transmission by more than 90%. Furthermore, a study by Steffen et al. demonstrated that widespread adoption of even moderately effective masks could prevent a substantial number of projected deaths. As countries reopen following lockdowns, governments and public health agencies now recommend using face masks as an essential measure to protect individuals in public spaces. However, ensuring mask compliance requires effective techniques for face mask detection and its Challenges: Face mask detection involves determining whether a person is wearing a mask [1]. Face detection algorithms have traditionally been developed for security, authentication, and surveillance purposes. Extensive research in Computer Vision and Pattern Recognition has contributed sophisticated algorithms for face detection, primarily using handcrafted features and traditional machine learning techniques. However, these approaches suffer from high complexity in feature design and low detection accuracy. In recent years, deep convolutional neural networks (CNN) have emerged as a powerful tool for face detection, improving performance significantly, Differentiating "Detection of the Face Under Mask" and "Detection of Mask Over Face": While considerable research has focused on face detection and recognition, there remains a notable difference between detecting a face under a mask and detecting the mask itself. The literature indicates limited research in the area of detecting masks over faces. Therefore, this work aims to develop a technique that accurately detects masks over faces in public areas such as airports, railway stations, crowded markets, and bus stops, to curb the spread of Coronavirus and contribute to public healthcare. The scarcity of available datasets for detecting masks on human faces makes model training difficult [2]

Type Style and Fonts. To address this issue, we employ transfer learning, leveraging pre-trained models trained on extensive datasets for similar face detection tasks. Proposed Technique and Contributions: Our proposed technique achieves an outstanding accuracy of 98.2% using an extensive dataset containing 45,000 images.The key contributions of our work are as follows: Novel Object Detection Method: We combine one-stage and two-stage detectors to develop a real-time object detection method that leverages transfer learning. This approach enhances the accuracy of object detection in video streams. Improved Affine Transformation: We introduce an enhanced affine transformation technique to accurately crop facial areas from uncontrolled real-time images, accounting for variations in face size, orientation, and background. This step aids in better localizing individuals who violate face mask norms in public areas or offices. Creation of an Unbiased Face Mask Dataset: We curate a comprehensive face mask dataset with a balanced distribution of images featuring faces with and without masks, faces with and without masks in the same image, and confusing images without masks. This dataset helps train the model effectively. Memory-Efficient Model: Our proposed model requires less memory, making it suitable for deploying embedded devices for surveillance purposes [3].

RELATED WORK

This paper proposes a face mask detection method that utilizes a single face image. It combines colour-based segmentation and a deep learning-based face detection network to detect masks on faces accurately. The authors present a real-time face mask detection approach incorporating a modified Faster R-CNN architecture. Their method accurately and efficiently detect faces with and without masks in diverse environments. This study explores deep learning and transfers learning techniques for face mask detection [4]. The authors employ pre-trained models, such as VGG16 and InceptionV3, to achieve accurate detection results. The paper proposes an approach that utilizes the EfficientNet model for face mask detection. Transfer learning is employed to fine-tune the model on a custom dataset, effectively detecting masked faces. This survey overviews various face mask detection techniques employed during the COVID-19 pandemic. It covers traditional methods, deep learning approaches, and hybrid methods and their advantages and limitations. The recently published Retina facemask detector, the proposed model achieves 11.07% and 6.44% higher precision and recall in mask detection. The proposed model is best suited for video surveillancedevices [5].

METHODOLOGY

The existing system to detect face masks is developed using yolov3. In the existing system, they introduced a backbone network which can allocate more resources and employ Glow andfocal loss to accelerate the training process and improve performance. It produces an accuracy of 86%. The accuracy can be improved while using theapplication with different algorithms. Face mask detection in the proposed system uses Mobile NET V2. Mobile NET V2 can detect face masks in a group of people and produce better accuracy than theexisting system. Provide an image of a few people wearing a mask and not wearing a mask as the input dataset and the segmented image of the same isobtained as the output. Preprocessing is one of the common and initial steps followed in this project. Preprocessing aims to decrease unwanted distortions, improve image data, and enhance a fewimportant image features that will be used in furtherprocessing. In computer graphics and digital imaging, the term image scaling defines resizing a digital image. In video technology, the term upscaling or resolution enhancement is usedto define the magnification of digital material. The graphic primitives in a vector graphic image can be scaled without losing image quality using geometric transformations. A new image of a higher or lower number of pixels is expected while scaling a raster graphics image. A visible quality loss can be expected when decreasing the number of pixels With relation to the digital signal processor, one in all the two-dimensional samples of sample rate conversion is the scaling of formation graphics that involves the conversion of the distinct signal from one rate to a different. The conversion of RGB to grey involves in gimpiness image software package for this. Methodology the tactic the strategy} wherever the common of most distinguished and least distinguished colours used is named lightness method. Methodology the tactic the strategy} wherever an easy average of 3 colours is employed is named the common method. A lot of subtle technique is the physical property technique. During this technique, inexperience is weighted heavily because it is a lot of sense. The common is employed additionally to weighted average for human perception. The conversion of grey to black and white happens during a binary image of pixels having precisely one in all 2 colours (black and white).In binary pictures, every pel is held on as one bit, either zero or one. These pictures are usually referred to as black-and-white, monochromatic etc. The icon mode in Photoshop expression is the same as a binary image. These pictures are used as masks or thresholds and video digitizing in the digital image process. Figure 1 shows overview of system.

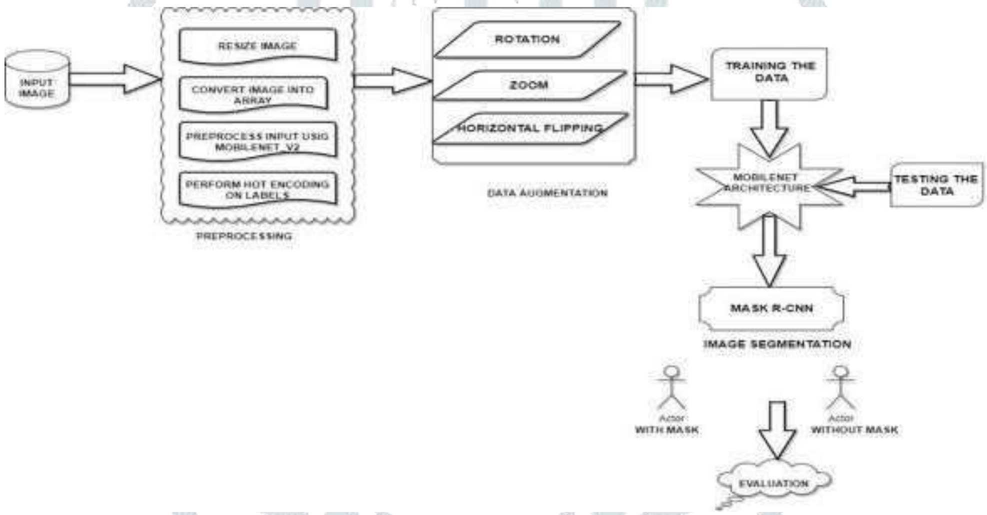


Figure 1: Overview of proposed system

MODEL TRAINING

MobileNetV2 advances MobileNetV1 in areas like classification, object identification and semantic segmentation for mobile visual recognition. As part of the TensorFlow-Slim Image Classification Library, MobileNetV2 is available. MobileNetV2 is also available as TF-Hub modules with pre-trained checkpoints. MobileNetV2 has two new. Figure 2 shows CNN architecture.

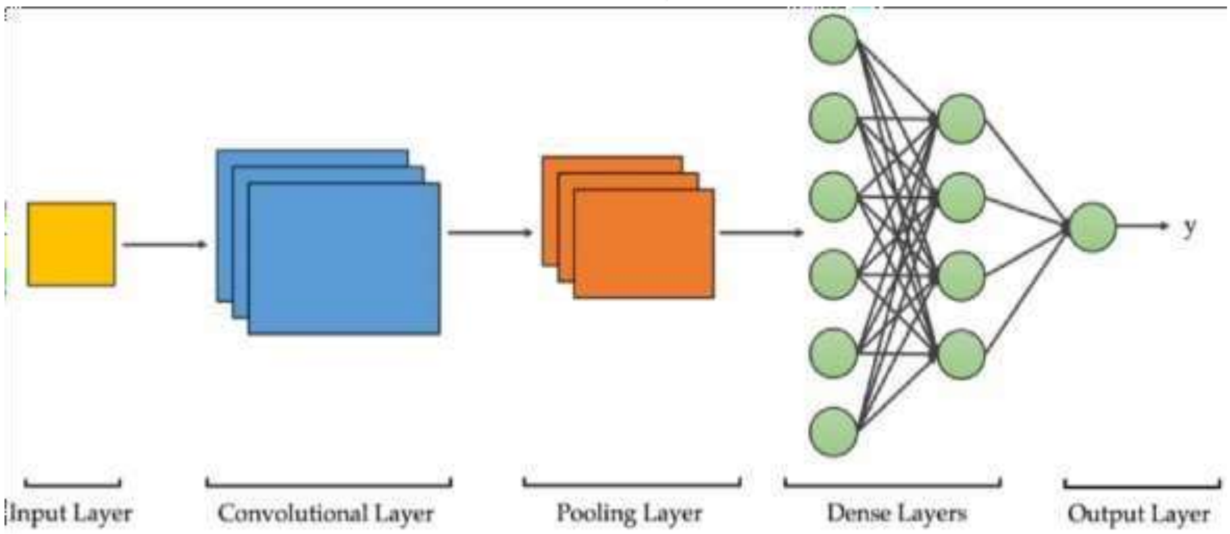


Figure 2: Convolutional Neural Network Architecture.

CONVOLUTIONAL NEURAL NETWORK

An artificial neural network optimized toprocess pixel data for image recognition andprocessing is called Convolutional Neural Network (CNN). The fundamental and building block of the computer vision task of image segmentation is the Convolutional Neural Network. ● Convolutional layer: The use of filters and kernels helps abstract the input image as a feature map. ● Pooling layer: This layer summarises the presence of features in patches of the feature map, which helps downsampling feature maps. ● Fully connected layer:Every neuron in one layer is connected to every neuron in another layer.

MACRO AVERAGE

The method is straightforward. Just take the average of the precision and recall of the system on differentsets. The Macro-average will be the average mean of Macro-average precision and macro-average recall as in figure 3.

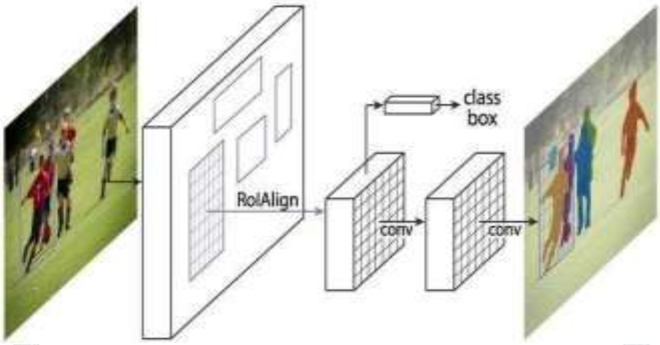


Figure 3: Macro average precision and recall.

WEIGHTED AVERAGE

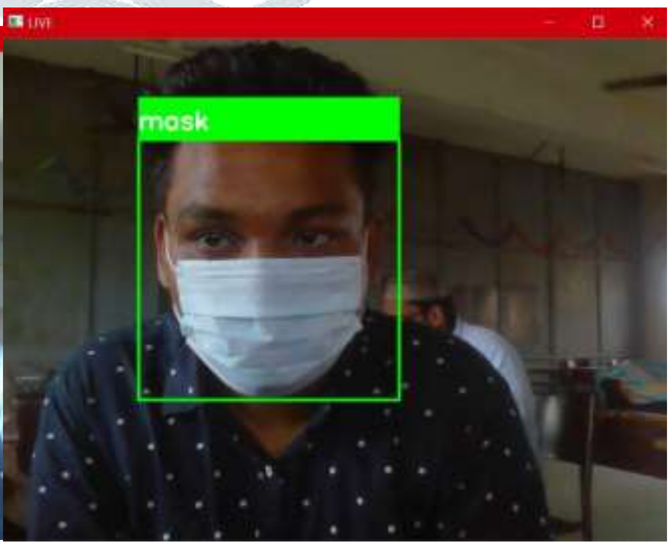
The F1 Scores area unit is calculated for every label, and then their average is weighted bysupport - that is, the range of true instances for every label. It may result in AN F1Score that's not between exactitude and recall.

IMAGE SEGMENTATION USING MASK R-CNN

A CNN for image segmentation and instancesegmentation which is developed on top of faster R-CNN used to locate objects and boundaries, is the Mask R-CNN. The precise detection of all objectsin an image and segmenting each instance is called Instance segmentation or Instance recognition. It comes out as the result of object detection, object localization and object classification. A clear difference between each object classified as similar instances can be observed in this type of segmentation. Each person is separated as a single entity in the instance segmentation process. It is also called foreground segmentation, which works on the image's subjects rather than the background. Mask R-CNN is faster than R-CNN by adding an output for object mask along with existing outputs like theclass label and bounding box as in figure 4.

RESULT:

Without mask



Without mask With mask:

Figure 4: Object recognition

CONCLUSION

With the rapid rise of Covid-19 cases, the government was forced to impose lockdowns worldwide. And wearing masks has become a necessary part of our lives, as suggested by WHO. Hence, we have implemented two Real-Time Face Mask Detection Models with MobileNetV2 and CNN in this project. Another with YOLOV4. Compared to both, the model with CNN and MobileNetV2 is much better than YOLOV4 as its accuracy is 98%, whereas YOLOV4's accuracy is 88.92. Since we used the MobileNetV2 architecture, it's also computationally efficient and thus making it easier todeploy the model to embedded systems (Raspberry Pi, Google Coral, etc.).

REFERENCES

[1] J. -M. Guo, C. -C. Lin, M. -F. Wu, C. -H. Chang and H. Lee, "Complexity Reduced Face Detection Using Probability-Based Face Mask Prefiltering and Pixel-Based Hierarchical-Feature Adaboosting," in IEEE Signal Processing Letters, vol. 18, no. 8, pp. 447-450, Aug. 2011, doi: 10.1109/LSP.2011.2146772.

[2] G. Deore, R. Bodhula, V. Udpikar and V. More, "Study of masked face detection approach in video analytics," 2016 Conference on Advances in Signal Processing (CASP), Pune, India, 2016, pp. 196-200, doi: 10.1109/CASP.2016.7746164.

[3] W. Bu, J. Xiao, C. Zhou, M. Yang and C. Peng, "A cascade framework for masked face detection," 2017 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), Ningbo, China, 2017, pp. 458-462, doi: 10.1109/ICCIS.2017.8274819.

[4] R. Raghavendra and C. Busch, "Novel presentation attack detection algorithm for face recognition system: Application to 3D face mask attack," 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 2014, pp. 323-327, doi: 10.1109/ICIP.2014.7025064.

[5] R. Raghavendra and C. Busch, "Robust 2D/3D face mask presentation attack detection scheme by exploring multiple features and comparison score level fusion," 17th International Conference on Information Fusion (FUSION), Salamanca, Spain, 2014, pp. 1-7.