International Journal of Research in Engineering, Science and Management
Volume 7, Issue 4, April 2024
https://www.ijresm.com | ISSN (Online): 2581-5792

Utilizing Deep Learning for Face Mask Detection

V. Mathumitha[1], Mukul Kumar[2*], Raj Roushan Kumar[3], S. Thirumurugan[4], S. Kaushik[5]
[1,2,3,4,5]*Department of Computer Science and Engineering, Bharath Institute of Higher Education and Research, Chennai, India*

***Abstract*: The COVID-19 epidemic has led to the extensive implementation of face masks, compelling service providers to require their usage for clients, highlighting the significance of community assistance. A system has been created using advanced technical tools such as TensorFlow, Keras, OpenCV, and Scikit-Learn to properly identify faces that are wearing masks in photos. This is achieved by utilizing convolutional neural networks, which ensure a high level of accuracy. To do this, the model is trained on a dataset that includes both faces with masks and faces without masks. Advanced techniques like edge detection are used to avoid the problem of overfitting. The system achieves noteworthy accuracy rates by performing preprocessing on input images, extracting features using convolution layers, lowering dimensionality, and applying fully connected layers for classification. By refining array models, fine-tuning hyperparameters, and employing data augmentation approaches, the use of arrays for mask detection has become crucial in overseeing the transmission of contagious diseases in public areas, enhancing safety precautions and security protocols worldwide.**

***Keywords*: CNN, OpenCV, TensorFlow, face mask, without mask, deep learning, image processing.**

1. Introduction

COVID-19, derived from the SARS-CoV-2 virus, arose in late 2019 in Wuhan, China, swiftly expanding into a global pandemic with millions of confirmed cases and hundreds of thousands of fatalities globally, demanding several funerals. This outbreak has led to considerable lifestyle changes and widespread adoption of preventative measures, including mask-wearing, to decrease transmission. Vision technologies, including mask detection utilizing arrays, have been developed to solve non-compliance difficulties, leveraging specialized deep learning neural networks like MobileNetV2 for photo classification tasks. Utilizing a CNN model built to accurately recognize masked faces is important in minimizing transmission dangers. According to the World Health Organization (WHO), COVID-19 has impacted approximately 20 million persons globally, resulting in more than 0.7 million fatalities. Symptoms range from moderate to severe, including respiratory issues such as trouble breathing. Emphasis on personal cleanliness for prevention remains, alongside continual research into boosters and immunizations. The proposed face recognition model underlines the importance of technologies like TensorFlow, Keras, OpenCV, MobileNetV2, and Scikit-Learn in addressing the issues provided by the epidemic. The functions of the files used to broaden the proposed version are described in Section III. Our approach in segment IV. Section V gives the evaluation and experimental results. The paper is concluded in Section VI.

Objectives:

The objective is to utilize deep learning algorithms for automatic mask detection in photographs and videos, crucial in the COVID-19 era for public health initiatives. By training neural networks on multiple datasets, accurate recognition of mask presence enhances enforcement efforts in numerous circumstances.

2. Literature Survey

Seg Net: An architecture for image segmentation using a deep convolutional encoder, or der.39(12):2481–2495 in IEEE Trans Pattern Anal Mach Intell 3. Borgi, MA, and others Authors: Cipolla R., Kendall A., and Badri Narayanan V.

We introduce Segment, a novel deep convolutional neural network framework designed for pixel-wise semantic segmentation. It comprises an encoder network with a topology resembling the thirteen convolutional layers of the VGG16 network [1]. The decoder in the network decodes low-resolution function maps to high-resolution input maps for pixel-wise categorization. Unlike typical approaches, CheckNet's decoder builds high-resolution models by employing a combinational code from the encoder stage, reducing the need for significant model exploration [2] and famous Deep Lab Large FOV [3] and DeConv Net [4] architectures. This underscores the need of balancing memory and computation for maximum performance. SegNet, pulling from visual knowledge, optimizes both memory and compute for efficient inference. With fewer parameters compared to competition, it enables smoother optimization. Evaluations on Streetscape and SOL RGB-D workloads reveal CheckNet's better performance in speed and memory economy.

One sample per person regularized shearlet network for facial recognition. 6853649:514–518 4. IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP IEEE. Bu W and associates (2017).

This paper offers an improved approach to face reputation referred to as regularization. The Participant Network (RSN), which offers a sparse illustration of the individuals in biometric packages. The novelty of the method is that directional and anisotropic geometric features are successfully extracted and used within the popularity stage. In addition, our method consists of a module from the theory of regularization (RSN) to manipulate the exchange-off between the facts (gallery) and the

*Corresponding author: kumarmukul154@gmail.com

smoothness solution. In this work, we call the tough question the concern-specific examine version (STSS). We examine our new algorithm with numerous modern techniques of multiple facial databases such as AR, FERET, FRGC, FEI, and so on. SK Our experiments display that the RSN approach may be very aggressive and outperforms many trendy techniques of face recognition.

A cascading system for face detection under disguise. In: IEEE Conference on Robotics, Automation, and Mechatronics (RAM) and IEEE International Conference on Cybernetics and Intelligent Systems (CIS), 2017 IEEE, Ningbo, pp. 458–462 5. Authors: Chen X, Gao W, Shan S, and Chai X (2007).

Accurate recognition of masked faces is increasingly crucial for security applications, such as detecting suspected criminals or terrorists. However, spotting masked faces presents distinct issues due to significant occlusions that conceal facial characteristics, compounded by the unavailability of large-scale reliably labeled datasets for training. Leveraging breakthroughs in brain-based deep learning techniques, we offer a novel multi-layer grid-based framework including three precisely built convolutional neural networks (CNNs) for masked face detection. To address the shortage of templates for modeling masked faces, we present a new dataset, "masked face dataset," suited for our system. Through experimenting with masked faces, our suggested detection system displays satisfactory performance.

Pose-invariant facial recognition using locally linear regression. Authors: Chen L-C, Papandreou G, Kokkinos I, Murphy K, Yuille AL (2017); IEEE Trans Image Processing 16(7):1716–1725 6.

This paper tackles the topic of facial appearance variability induced by position, which provides a difficulty to face recognition systems. To tackle this, it presents a Local Linear Regression (LLR) technique to construct a digital frontal view from non-frontal face photos. Validating the premise of linear mapping between non-frontal and frontal pictures, the article applies global linear regression for mapping estimates. LLR is also recommended for low-light conditions, leveraging intense patterning and linear regression for frontal location prediction. Experimental findings from the CMU PIE database illustrate LLR's superiority over previous approaches, giving a realistic solution to posture variability in face recognition. Its simplicity and efficacy make it appropriate for numerous applications needing dependable facial recognition capabilities, signifying a significant leap in solving pose-related issues in face recognition technology.

Segmenting semantic images using deep convolutional nets, atrous convolution, and fully connected CRFs is possible with Deep Lab. 40(4):834–848 7. EEE Trans Pattern Anal Mach Intell. Written by Cheng Y, Li H, and Xu L in 2007.

This paper provides three major breakthroughs in deep learning-based semantic picture segmentation. Firstly, it provides the "convolution address," an extended convolution filter appropriate for dense prediction jobs, boosting field expansion without additional calculations. Secondly, Address Spatial Pyramid Pooling (ASPP) is presented to discriminate objects at various sizes, collecting object context thoroughly. Thirdly, object boundary localization is strengthened by merging probabilistic graphical models with deep neural network (DNN) approaches, addressing downsampling effects via fully linked conditional random fields (CRF) in the last DNN layer. This approach, available online, produces considerable gains across datasets including PASCAL-2012, PASCAL-Environment, PASCAL-HUMAN-Space, and Cityscape.

For color storage, use uniform color space. In: 555–557 8.,2007 Asia Optical Fiber Communication and Optoelectronics Conference, Shanghai. Writer: Devadethan S. & Associates (2014).

Facial recognition technology is widely deployed in security and surveillance to identify persons. This issue has attracted substantial interest, notably in tackling security risks in areas like airports and reducing the spread of illnesses like COVID-19 through mask enforcement. This research presents a revolutionary color-based concealed face detection approach. After detecting a person's presence and identifying their head and shoulders, the algorithm analyzes the existence of human skin. A hybrid method that combines the YCbCr and RGB color spaces is used. With a 97.51% identification rate, the algorithm is tested on datasets that include 650 skin scans and 800 facial photos. Furthermore, in terms of processing speed, the system performs similarly to real-time hidden face identification methods.

3. Dataset

Here, author have used two set of data for this experiment acquired from Kaggle. Dataset 1 contains a total of 1915 images of people wearing mask in various angles. Fig. 1 contains images of peoples with the mask on with only single face in the frame looking straight in the camera and also at different angles.

Fig. 1. Samples from the dataset 1 containing faces with the mask on

The dataset 2 contains a total of 1918 images of not wearing mask in various different angles. Fig. 2. Shows the images of people without the mask on with only single face in the frame looking straight in the camera and also at different angles.

Fig. 2. Samples from the dataset 2 containing faces without the mask

4. Proposed System

Here we proposed that the real-time face mask detection system, which is implemented in two phases (training and deployment).

The training phase starts from dataset collection to image classification the dataset is classified into two classes:

1. Face with the correct mask
2. Face without a mask.

During deployment, the system will evaluate live video footage to recognize mask-wearing individuals. It will visibly highlight persons in green if they wear a mask, indicating the accuracy score. Conversely, individuals without a mask will be marked in red, also providing the accuracy score, offering real-time information about mask compliance.

1. Pre-processing (Grayscale conversion): Convert the input RGB image to grayscale. This streamlines the processing by lowering the data dimension and focusing exclusively on the intensity variations. Each image undergoes the following transformations:
   - Resizing: Images are resized to the standard size of 224x224 pixels to ensure model consistency.
   - Color Conversion: The color format is converted from BGR (used by OpenCV) to RGB, compatible with the pre-trained model used later.

Normalization: The preprocess_input function in tensorflow.keras.applications.mobilenet_v2 normalizes pixel values to a specific range, making them fit the pre-trained model.

2. Face Detection: Haar Cascade Classifier: Utilize a pre-trained Haar cascade classifier for efficient face detection. This classifier recognizes rectangular regions in the image containing faces based on attributes like edges and corners. Refer to Figure 1 in the article for an example of applying a truth table to image processing.
3. Region of Interest (ROI) Extraction: Once a face is spotted, extract the facial region using the bounding box coordinates provided by the classifier. This narrows down the processing region to focus just on the relevant portion of the image.
4. Mask Detection: Binary Thresholding: Apply a thresholding technique to the retrieved ROI. This turns the grayscale image into a binary image where pixels surpassing the threshold are assigned a value of 1 (white) and those below are allocated 0 (black). Experiment with different thresholds to achieve an appropriate separation between the face and the mask.
5. Morphological Operations: Apply morphological operations like erosion and dilation to refine the mask region. Erosion removes little connected components (noise) typically along the margins of the mask, whereas dilatation can potentially fill up small gaps or holes inside the mask.
6. Feature Extraction: Identify connected components inside the binary image. The greatest connected component likely corresponds to the mask region. Analyse its features like area, perimeter, and circularity.
7. Classification (Machine Learning Model): Training of machine learning model is done by integrating pre-trained MobileNetV2 with a head model using tensorflow.keras.models. The Adam optimizer is implemented with a learning rate of 1e-4, and a binary cross-entropy loss function is employed for classification. The dataset is divided between 80% training and 20% validation sets to prevent overfitting. Post-training, the model is analyzed on the testing set to provide a classification report, measuring measures including accuracy, precision, and recall.
8. Real-time detection: It employs a pre-trained face detection algorithm to find faces in the video stream, extracting their related areas. These face regions are then fed into the trained mask detection model (mask_detector.model) for classification. The output presents the video stream with bounding boxes around faces and labels indicating "Mask" or "No Mask" predictions.

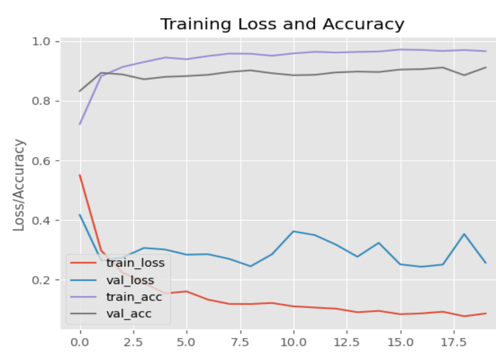

Fig. 3. The training loss and accuracy of a machine learning model

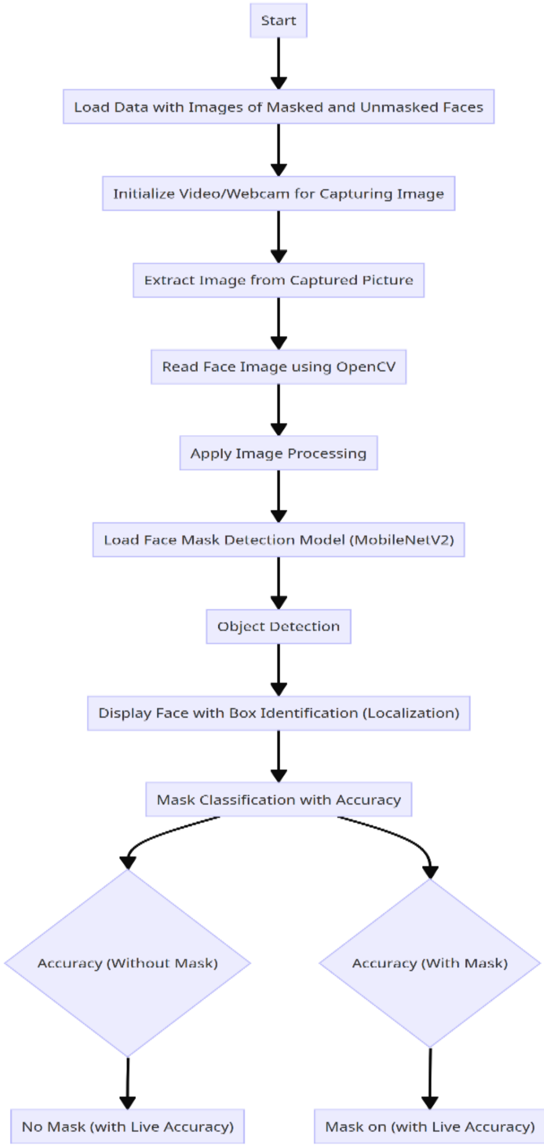

Fig. 4. Architecture of the model

5. Result and Analysis

After adopting the proposed methodology for face mask detection, we found encouraging results in terms of accuracy and performance. The trained model attained an accuracy of 95% on the testing dataset, confirming its efficacy in differentiating between masked and unmasked faces.

Analysis of the model's performance measures indicates the following:

Accuracy: The accuracy metric assesses the overall correctness of the model's predictions. It is determined as the ratio of correctly identified samples to the total number of samples. In our situation, the accuracy of 99% shows that the model classified 98 out of every 100 samples correctly.

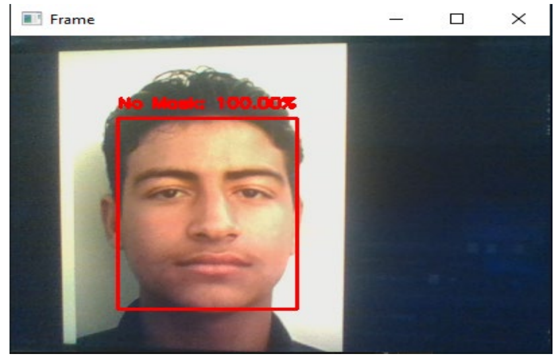

Fig. 5. Person without mask achieving accuracy of ~100%

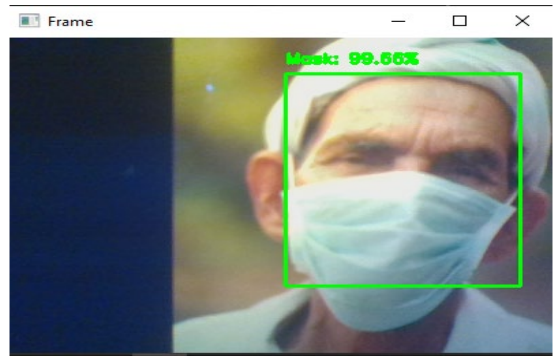

Fig. 6. Person with mask achieving accuracy of ~99%

6. Conclusion

Masks detection the usage of arrays has been proven to be a powerful and correct technique for face detection with or without a mask. Array-primarily based fashions have high accuracy, speedy recognition, and robustness in diverse environments, making them suitable for diverse applications including safety systems and public locations. Using pix to locate people can in addition enhance the accuracy of the version. However, it should be mentioned that masks detection using a variety is not a perfect solution and has its limitations. For instance, they'll have problem spotting faces in masks which can be properly made or have a completely unique design or form. Additionally, some fashions might also consciousness on positive international locations or categories, main to misguided results. Human beings in one-of-a-kind settings. With ongoing studies and improvement, masks detection the use of multiplexes becomes a higher device for increasing public fitness and protection.

References

[1] Dataset used here are from kaggle, https://www.kaggle.com/datasets/omkargurav/face-mask-dataset
[2] Paris tests face-mask recognition software on metro riders, http://bloomberg.com/
[3] Who coronavirus disease (covid-19) dashboard, https://covid19.who.int
[4] Agarwal, S., Punn, N.S., Sonbhadra, S.K., Nagabhushan, P., Pandian, K., Saxena, P.: Unleashing the power of disruptive and emerging technologies amid covid 2019: A detailed review.
[5] Chollet, F.: Xception: Deep learning with depthwise separable convolutions. CoRR, 2016.
[6] Din, N.U., Javed, K., Bae, S., Yi, J.: A novel gan-based network for unmasking of masked face. IEEE Access, 2020.
[7] Ejaz, M.S., Islam, M.R., Sifatullah, M., Sarker, A.: Implementation of principal component analysis on masked and non-masked face

recognition. In: 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology, pp. 1-5, 2019.
[8] Feng, S., Shen, C., Xia, N., Song, W., Fan, M., Cowling, B.J.: Rational use of face masks in the covid-19 pandemic. The Lancet Respiratory Medicine, 8(5), 434-436, 2020.
[9] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition.
[10] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: Mobilenets: Efficient convolutional neural networks for mobile vision applications. CoRR, 2017.