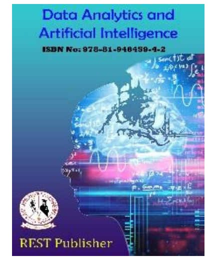
Data Analytics and Artificial Intelligence

Vol: 3(3), 2023

REST Publisher; ISBN: 978-81-948459-4-2

Website: http://restpublisher.com/book-series/daai/



Real Time Face Mask Detection Using Deep Learning

***R. Vikram, B.K. Santhosh, G. G. Maneesh, R. Mohan, K. Nagaraj**

Adhiyamaan college of engineering Hosur,Tamil Nadu, India.

*Corresponding Author Email: vikrambrv@gmail.com.

Abstract: *Globally, COVID-19 has had a negative impact. Wearing a face mask is one of the precautions to lessen the risk of viral transmission, according to studies. Additionally, a lot of public facilities and service providers only permit visitors to use their facilities or services if they are properly mask-clad. Therefore, it is impossible to manually track the customer whether or not they are wearing a mask. This technology is therefore the key in this situation. One of the most accurate and effective face mask detectors is the one we suggest in this paper, which uses image processing. The three stages of the proposed system are as follows: 1. Image pre-processing Face recognition and cropping 2. 3. A classifier for face masks. Our system has the ability to recognize masked. This system will encourage the use of face masks, track safety violations, and guarantee a secure workplace.*

1. INTRODUCTION

The COVID-19 pandemic is currently affecting the entire world. To stop the Corona virus from spreading, people are taking a variety of actions. One of the most important of the many crucial measures required to combat COVID-19 is the use of a face mask. There is still a ton of COVID-19 research and study being done. Additionally, studies have shown that using a face mask significantly lessens the issue of viral transmission. A person who is wearing a face mask can also feel protected. In our homes, we consciously take care of everything, but in public settings like offices, malls, colleges, etc., it can be more difficult to maintain people's safety. Manually determining whether or not someone is donning a face mask is not practical. Here, technology plays a role. Artificial intelligence and machine learning are two categories of technologies that offer practical answers to challenging issues across a variety of fields. We have made an effort to create a face mask recognition system using machine learning to stop the spread of the Corona virus. This method of face mask detection works well. It has the ability to identify both masked and unmasked faces. With the development of this system, it is possible to tell whether or not someone is wearing a face mask. The system will display a message like "No Mask" if the person is not wearing a face mask. Additionally, it will greatly contribute to maintaining public safety to permit those wearing face masks to enter. The system can produce some statistical information that will be useful in forecasting future COVID-19 outbreaks. In the end, we're attempting to contribute some useful technology that can identify a face mask in the fight against COVID-19.

2. RELATED WORK

Single-stage detectors: The input image is used by the single-stage detectors to learn the class probabilities and bounding box coordinates, treating the detection of region proposals as a straightforward regression problem. Early examples included Deep Multi Box and Over Feat. By demonstrating real-time predictions and achieving astounding detection speeds, YOLO (You Only Look Once) popularized the single-stage approach but suffered from poor localization accuracy when compared to two-stage detectors, particularly when small objects are taken into account [YOLO network basically divides an image into GxG grids, with each grid producing N predictions for bounding boxes. The prediction process only allows each bounding box to have one class, which prevents the network from finding smaller objects. The SSD network uses convolutional layers of various sizes, whereas YOLO uses two fully connected layers. This is the primary distinction between the two architectures. Additionally, Lin's Retina Net is a single-stage object detector that achieves remarkable accuracy and speed comparable to two-stage detectors by detecting dense objects in the image across multiple layers using the featured image pyramid and focal loss.

Two-stage detectors : Two-stage detectors, as opposed to single-stage detectors, use a sophisticated line of thought in computer vision to predict and categorize region proposals. In order to classify potential detection, they first predict proposals in an image and then apply a classifier to these regions. Researchers have previously proposed several two-stage region proposal models. Ross Girshick et al. first introduced the region-based convolutional neural network, also known as R-CNN in 2014. One of the earliest extensive CNN applications to the issue of object localization and recognition may have been this one. The model successfully demonstrated state-of-the-art results on benchmark datasets like VOC-2012 and ILSVRC-2013. In essence, R-CNN uses a selective search algorithm to extract a set of object suggestions. In order to fine-tune the model, it adds a new layer between shared convolutional layers called Region of Interest (RoI) pooling layer. Additionally, it enables the training of a detector and regressor at the same time without changing the network configurations. Even though Fast-R-CNN successfully combines the advantages of R-CNN and SPPNet, it still lags behind single-stage detectors in terms of detection speed . Additionally, Faster R-CNN is a combination of Region Proposal Network (RPN) and Fast R-CNN. By gradually integrating distinct object detection system building blocks (such as proposal detection, feature extraction, and bounding box regression) in one step, it enables almost cost-free region proposals. Although this integration achieves a breakthrough for the Fast R-CNN'sspeedbottleneck, there is computation redundancy at the following stage.

3. EXISTING SYSTEM

In the existing system, we attempt to determine whether a person is wearing a mask or not using general ML algorithms for image processing techniques. The study of computer algorithms that get better on their own with practice is known as machine learning (ML). It's considered to be a part of artificial intelligence. Without being explicitly programmed to do so, machine learning algorithms create a mathematical model from sample data, also referred to as "training data," in order to make predictions or decisions. In a wide range of applications, including email filtering and computer vision, where the development of conventional algorithms to carry out the required tasks is challenging or impractical, machine learning algorithms are used. Quite difficult to tell whether a person is wearing a face mask or not.

4. PROPOSED SYSTEM

We attempt to create a hybrid model for image analysis in the system that is being proposed. For FMD, a hybrid model combining deep learning and traditional machine learning will be presented. We will use OpenCV to perform real-time face detection from a live stream via our webcam using a FMD dataset that consists of with mask and without mask images. Using the dataset, Py, OpenCV, Tensor Flow, and Keras will be used to build a COVID-19 face mask detector. Using computer vision and deep learning, we want to determine whether the person in the image or video stream is wearing a face mask or not. Both datasets of people wearing masks and those who aren't can be used to train this system.

5. MATERIAL AND METHODS

TensorFlow: That eases the process of acquiring data, training models, solving the predictions, and refining future results. It is created by google brain team TensorFlow is an open-source library for numerical computation and large-scale Machine learning. Tensors are also a standard way of representing data in deep learning. TensorFlow helps developers to create graphs which are structures that describe how the data moves through graphs.

Keras: Keras is a deep learning framework which is easy-to-use. And also, it is a free open-source library in Python. Keras is a powerful framework and effective. Used for developing and evaluating different deep learning models. It wraps the efficient and effective numerical computation libraries Theano and TensorFlow. And also, it allows you. to define as well as train neural network models in just a few lines of code.

PyTorch: PyTorch is a Python machine learning package based on Torch, also a deep learning framework, which is an open-source package in machine learning. PyTorch has two main features: Tensor computation (like NumPy) with a strong GPU acceleration. Automatic differentiation for building as well as training neural networks.

OpenCV: The Standard definition of OpenCV (Open-Source Computer Vision Library) is an open-source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and also to accelerate the use of machine perception in the commercial products.

.

6. METHODOLOGY

The benchmark for object recognition provided is the foundation for the suggested model. This benchmark states that all tasks involved in an object recognition problem can be grouped together under the three main categories of the "backbone," "neck," and "head" shown in. A basic convolutional neural network that can extract data from images and turn them into feature maps serves as this system's backbone. The concept of transfer learning is used in the proposed architecture's backbone to use previously learned characteristics of a potent pre-trained convolutional neural network in extracting new features for the model.

Mathematical model :
System Description:
- Input: Image showing masked/unmasked faces
- Output: Detect face showing 'Mask' or 'No Mask' message.
- Functions: Extract (), Detect(), Classify(), Display().
- Mathematical Formulation:
- S = (I, F, O)
  where, Input = (I1, I2, I3,...,In)
- Function = (F1, F2, F3,...,Fn)
- Output = (O1, O2, O3,…,On)
- Success Conditions: Masked and unmasked faces are successfully detected and expected output is displayed on screen.

Failure Conditions:

1.Camera is not capturing input frame.

2.Face is not available for detection purpose

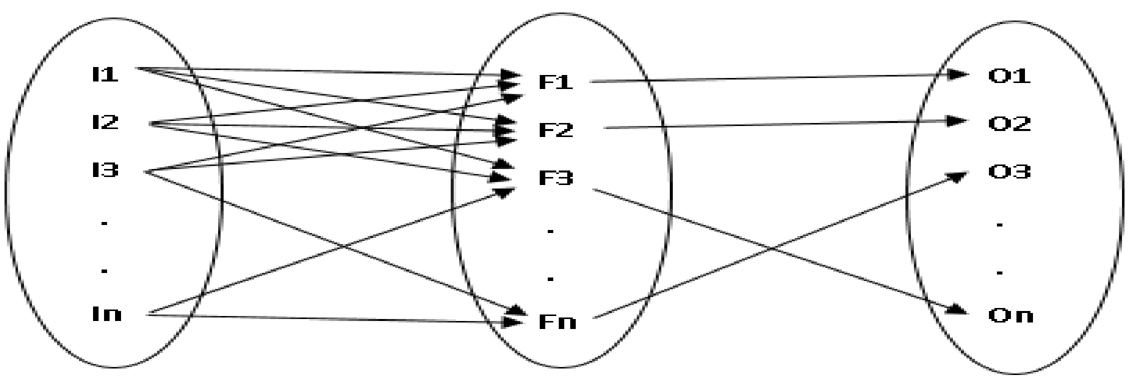


FIGURE 1. Failure Conditions:

Creation of unbiased facemask dataset: Initially, a facemask-centric dataset called MAFA with a total of 25,876 images divided into masked and non-masked classes was taken into consideration. While there are only 2018 non-masked images in MAFA, there are 23,858 masked images. It has been noticed that MAFA struggles with an extrinsic class imbalance issue that could lead to bias against the majority class. In order to compare how well the image classifier performed using the proposed dataset (unbiased) and the original MAFA set (biased), an ablation study was carried out.

Fine tuning of pre-trained model: Deep neural networks are used in the proposed work to detect facemasks because of their superior performance to other classification algorithms. But deep neural network training is costly because it takes a long time and a lot of computing power. Deep learning-based transfer learning is used in this case to train the network more quickly and affordably. Transfer learning enables the neural network's trained knowledge to be transferred to the new model in terms of parametric weights. Even when the new model is trained on a small dataset, it improves performance. The 14 million images from the ImageNet dataset were used to train a number of pre-trained models, including AlexNet, MobileNet, ResNet50, etc. ResNet50 is used in the proposed model as a pre-trained

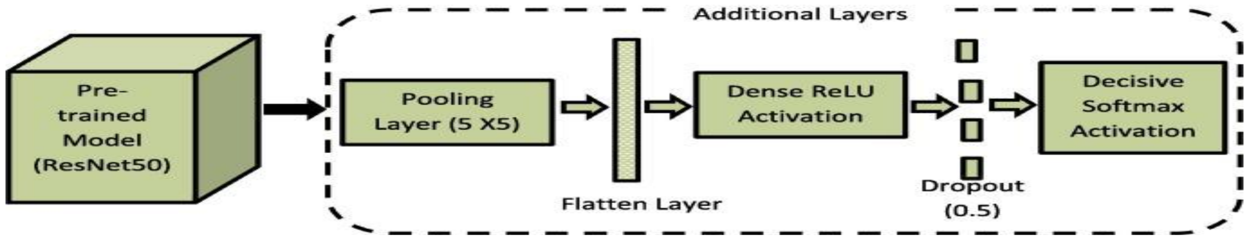


FIGURE 2.

Image complexity predictor for face detection: Its goal is to first divide the data into soft and hard images, and then later, through a facemask classifier, to classify the data into masks and non-masks. How to tell whether an image is soft or hard is the crucial question that needs to be addressed. The "Semi-supervised object classification strategy" put forth by Lonescu et al. provides the answer to this query. Our task is a good fit for the semi-supervised object classification strategy because it predicts objects without localizing them. Three sets of image samples were used to test this strategy: the first set (L) contains labelled (hard/soft) training images, the second set (U) contains unlabelled training images, and the third set (T) contains labelled (hard/soft) test images. The algorithm for image complexity predictor is outlined below:

Algorithm : Image_Complexity_Predictor ()
1. Input:
2. Image ← input image
3. Dfast ← single-stage detector
4. Dslow ← two-stage detector
5. C ← Image complexity
6. Computation:
7. If (C =Soft) R ← Dslow(Image)
8. else R ← Dfast(Image)
9. Output: 10. R ← set of region proposal

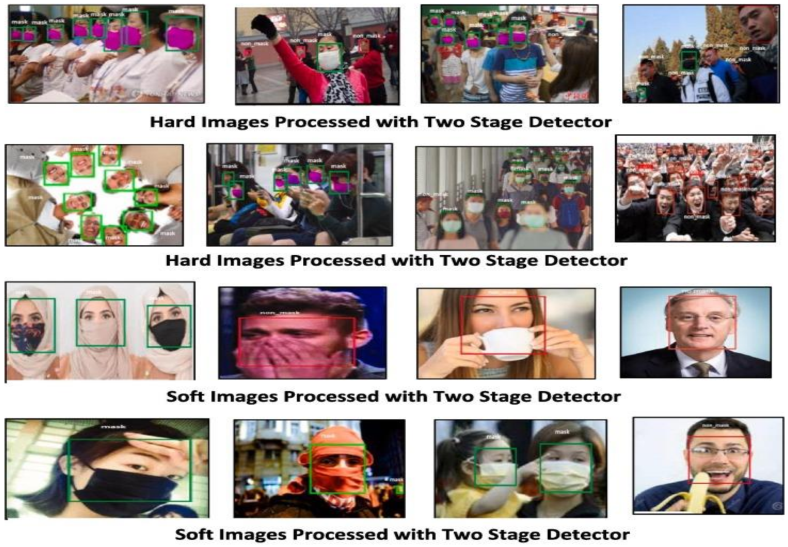


FIGURE 3.

A soft image is suggested to be processed through a quick single-stage detector while the hard image is accurately processed by a two-stage detector after the hardness of the test images has been determined using an image complexity predictor. We use the faster R-CNN based on ResNet50 for hard image prediction and the MobileNet-SSD model for soft image prediction.

Identity prediction: The non-mask faces are passed separately into a neural network for further investigation of a person's identity for being in violation of the facemask norm after faces with masks and non-masks are detected in the search proposal. The step needs an input with a fixed size. To obtain a fixed-size input, it may be necessary to reshape the face in the bounding box to 96 96 pixels. The face might be gazing in a different direction, which could be a problem with this solution. This problem is very easily resolved using affine transformation.

7. RESULT AND DISCUSSION

Two datasets are used to train, validate, and test the model. The method achieves accuracy up to 95.77% when applied to dataset 1. shows how the cost of error is reduced by this optimized accuracy. Dataset 2 is more adaptable than Dataset 1 because it includes multiple faces in the frame and various types of masks with various colors. As a result, the model on dataset 2 achieves an accuracy of 94.58%. The comparison of training and validation loss for dataset 2. MaxPooling is one of the main contributors to this accuracy. It reduces the number of parameters in the model and provides basic translation invariance for the internal representation.

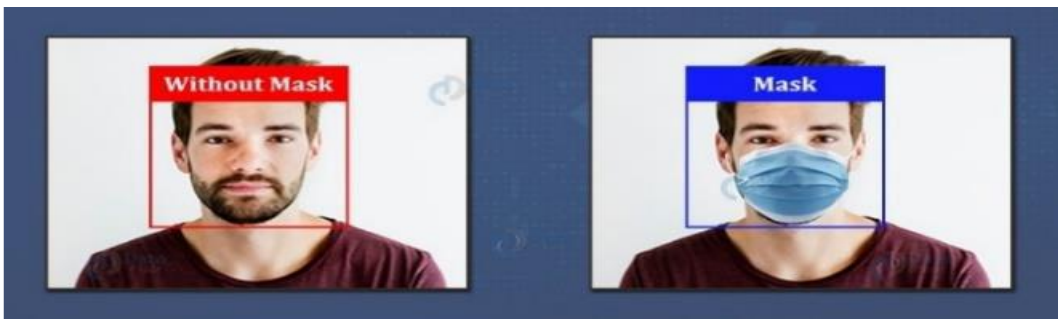


FIGURE 4. output

Advantages of Proposed System: It increases accuracy and reduces the time required for identifying mask violators. It helps in making sure that everybody is wearing mask. It helps in making public places safer. It avoids the risks and cost involved in manual detection.

Future Scope: Future Work More than fifty countries around the world have recently initiated wearing face masks compulsory. People have to cover their faces in public, supermarkets, public transports, offices, and stores. Retail companies often use software to count the number of people entering their stores. They may also like to measure impressions on digital displays and promotional screens. We are planning to improve our FMD tool and release it as an open-source project. Our software can be equated to any existing USB, IP cameras, and CCTV cameras to detect people without a mask.

8. CONCLUSION

In this paper, we first provided a brief explanation of the work's motivation. The model's learning and performance task was then illustrated. The method has produced reasonably high accuracy using simple ML tools and simplified techniques. There are numerous applications for it. Given the Covid-19 crisis, wearing a mask might soon be required. To use the services of many public service providers, customers must properly wear masks. The public health care system will benefit greatly from the deployed model. In the future, it might even be able to tell if someone is wearing their mask correctly. The model can be further enhanced to determine whether or not the mask is virus-prone. With the aid of a digital camera, this work has been a successful attempt to create a consensus configuration model to stop the spread of COVID19. It explores the risk of COVID 19 transmission by identifying people wearing masks and estimating the social distance between individuals. This model used the quick object-detection algorithm YOLOv3 to take pictures of people in an open area. After that, ResNet50V2 used those pictures to determine whether or not the people were wearing masks. The social distance between any two individuals in a moving image has been determined using the TVTM model. This research may result in a revolutionary web application that reports on the risk of "disease spreading" immediately

REFERENCES

[1]. World Health Organization et al. Coronavirus disease 2019 (covid-19): situation report, 96. 2020. - Google Search. (n.d.). https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200816-covid-19-sitrep-209.pdf?sfvrsn=5dde1ca2_2.

[2]. Social distancing, surveillance, and stronger health systems as keys to controlling COVID-19 Pandemic, PAHO Director says - PAHO/WHO | Pan American Health Organization. (n.d.). https://www.paho.org/en/news/2-6-2020-social-distancing-surveillance-and-stronger-health-systems-keys-controlling-covid-19.

[3]. Garcia Godoy L.R., et al. Facial protection for healthcare workers during pandemics: a scoping review, BMJ. *Glob. Heal.* 2020;5(5) doi: 10.1136/bmjgh-2020-002553.

[4]. Eikenberry S.E., et al. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. *Infect. Dis. Model.* 2020;5:293–308. doi: 10.1016/j.idm.2020.04.001.

[5]. Wearing surgical masks in public could help slow COVID-19 pandemic's advance: Masks may limit the spread diseases including influenza, rhinoviruses and coronaviruses -- ScienceDaily. (n.d.). https://www.sciencedaily.com/releases/2020/04/200403132345.htm.

[6]. Nanni L., Ghidoni S., Brahnam S. Handcrafted vs. non-handcrafted features for computer vision classification. *Pattern Recogn.* 2017;71:158–172. doi: 10.1016/j.patcog.2017.05.025. [CrossRef] [Google Scholar]

[7]. Y. Jia et al., Caffe: Convolutional architecture for fast feature embedding, in: MM 2014 - Proceedings of the 2014 ACM Conference on Multimedia, 2014, doi: 10.1145/2647868.2654889.

[8]. P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. Lecun, OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks, 2014.

[9]. Erhan D., Szegedy C., Toshev A., Anguelov D. *Proceedings of the IEEE conference on computer vision and pattern recognition.* 2014. Scalable Object Detection using Deep Neural Networks; pp. 2147–2154.

[10]. J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016, vol. 2016-Decem, pp. 779–788, doi: 10.1109/CVPR.2016.91.

[11]. M. Jiang, X. Fan, and H. Yan, RetinaMask: A Face Mask detector, 2020, http://arxiv.org/abs/2005.03950.

[12]. Inamdar M., Mehendale N. Real-Time Face Mask Identification Using Facemasknet Deep Learning Network. SSRN Electron. J. 2020 doi: 10.2139/ssrn.3663305

[13]. Qiao S., Liu C., Shen W., Yuille A. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2018. Few-Shot Image Recognition by Predicting Parameters from Activations.

[14]. Kumar A., Zhang Z.J., Lyu H. Object detection in real time based on improved single shot multi-box detector algorithm. J. Wireless Com. Netw. 2020;2020:204. doi: 10.1186/s13638-020-01826-x.

[15]. Morera Á., Sánchez Á., Moreno A.B., Sappa Á.D., Vélez J.F. SSD vs. YOLO for detection of outdoor urban advertising panels under multiple variabilities. Sensors (Switzerland) 2020 doi: 10.3390/s201645.