FACE MASK RECOGNITION MODEL USING DEEP LEARNING AND CONVOLUTIONAL NEURAL NETWORS

Venkata Sameera Dronamraju[1], Dr.Tauheed Ahmed[2]

[1] Research Scholar, Department of Computer Science and Engineering, GITAM University, Hyderabad, Telangana, India

[2] Professor, Department of Computer Science and Engineering, GITAM University, Hyderabad, Telanagana, India

ABSTRACT

The world is battling with Covid19 pandemic. There are such countless fundamental types of gear expected to battle against Corona infection. One of such most fundamental is Face Mask. In the current circumstance manual power is being utilized to check assuming an individual has placed on facial covering or not, and there are no productive facial covering identification applications, which are presently popular for transportation implies, thickly populated regions, private locale, huge scope makers and different endeavors to guarantee wellbeing. This system can along these lines be used continuously applications which require facial covering disclosure for security purposes on account of the episode of Covid-19.In this task Deep Learning is utilized for recognition of facial coverings. CNN is utilized to recognize the presence and nonattendance of facial covering. This undertaking can identify the covered faces exceptionally quick from each conceivable point. This framework is extremely helpful in post episode period and can be introduced openly places, for example, Railway Stations, Airports, Parks, Schools, universities, workplaces and so on to follow and guarantee wearing of covers by individuals.

Key Words: Deep Learning, CNN, Face Mask, Recognition

1. INTRODUCTION

The world is battling with Covid19 pandemic. There are such countless fundamental supplies expected to battle against Corona infection. One of such most fundamental is Face Mask.[2] In the current circumstance manual power is being utilized to check assuming an individual has placed on facial covering or not, and there are no productive facial covering discovery applications, which are presently popular for transportation implies, thickly populated regions, private locale, huge scope makers and different endeavors to guarantee wellbeing.

The model in this paper uses the Convolutional Neural Network and Deep Learning[1]. It is a profound brain network model utilized for investigating any visual symbolism. It accepts the picture information as info, catches every one of the information, and ship off the layers of neurons. It has a completely associated layer, which processes the last result that addresses the forecast about the picture. The Convolutional brain

network model utilized here is the MobileNetV2 engineering. MobileNet model is an organization model involving profundity wise divisible convolution as its essential unit. Its profundity wise distinguishable convolution has two layers: profundity wise convolution and point convolution. It relies upon an irritated excess development where the extra affiliations are between the bottleneck layers. In general, the engineering of MobileNetV2 contains the underlying completely convolution layer with 32 channels, trailed by 19 leftover bottleneck layers.

The model has utilized OpenCV[3] to satisfy the motivation behind utilizing the video transfer for catching the casings in the video transfer.

2.BACKGROUND

2.1 Convolutional Neural Networks:

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm[4] which can take an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. CNN is composed of two major parts: Feature Extraction: In this part, the organization will play out a progression of convolutions and pooling tasks during which the elements are recognized[6]. Assuming that you had an image of a zebra, here the organization would perceive its stripes, two ears, and four legs. Classification: Here, the completely associated layers will act as a classifier on top of these separated highlights. They will assign a probability for the object on the image, being, what the algorithm predicts it is.
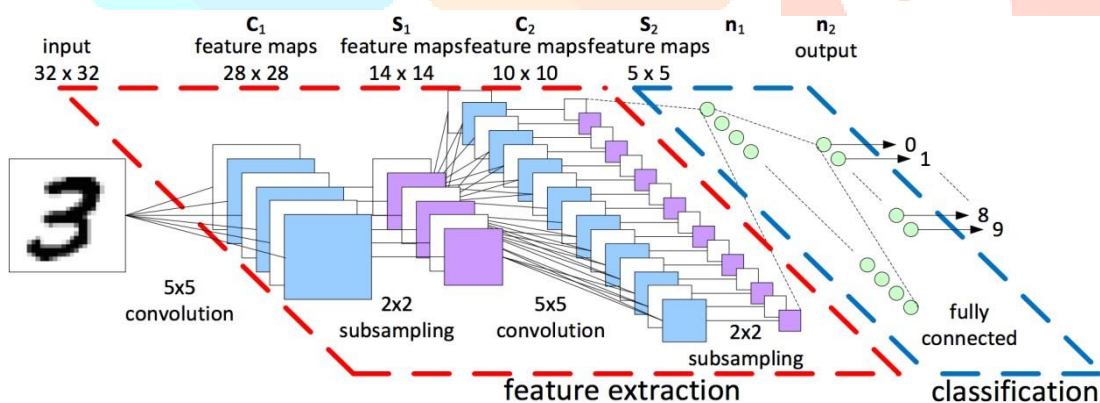

Fig 1.Layers of CNN

2.2 OPENCV:

OpenCV is the immense open-source library for the PC vision, AI, and picture handling and presently it assumes a significant part continuously activity which is vital in the present frameworks. By utilizing it, one can deal with pictures and recordings to recognize items, faces, or in any event, penmanship of a human. Whenever it coordinated with different libraries, for example, NumPy, python is fit for handling the OpenCV exhibit structure for examination. To Identify picture plan and its various components we use vector space and go through mathematical methodology on these features.

3. PROPOSED METHODOLOGY:

The model proposed here is planned and displayed utilizing python libraries in particular Tensorflow, Keras and OpenCV. The model we utilized is the MobileNetV2 of convolutional brain organization. The technique for utilizing MobileNetV2 is called utilizing Transfer Learning. Transfer learning is using some pre trained model to train your present model and get the prediction which saves time and makes using training the different models easy.

3.1 DATA IMPLEMENTATION:

The implementation of the proposed system is presented here. The steps followed as well as the frameworks used and datasets are also described in this section.

Data At Source:

The data is augmented using OpenCv. The images are taken in the form of two different datasets "with mask" and without mask". The images are of varied sizes and these tags are given in the initial phase itself.

Data Preprocessing:

Data Preprocessing is the process of making the images noise free and clear. Here few Data Preprocessing steps are performed on the data. The images are resized in the first phase of preprocessing. After that RGB colour filtering is applied all over the channels. Then the images are normalized and pixel value of 224*224*3 is applied on them. Lastly they are converted into tensors.

3.2 WORKFLOW:

First we feed the dataset withinside the version, run the education program, which trains the version at the given dataset. Then we run the detection program, which activates the video movement, captures the frames constantly from the video movement with an anchor field the usage of item detection process. This is surpassed thru the MobileNetV2 version layers which classifies the picture as without or with masks. If the individual is carrying a masks, a inexperienced anchor field is displayed and crimson if now no longer carrying a masks with the accuracy for the equal tagged at the anchor field. The go with the drift chart depicts the clean photo of the process.(Fig 2)
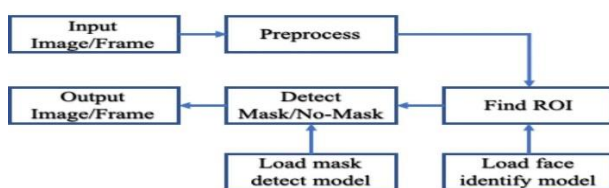

Fig 2. Flow of process

Deep Learning Frameworks:

•   Tensorflow

•   Keras

•   PyTorch

•   Caffe

•   Microsoft Cognitive Toolkit

These are some of the deep learning frameworks and there are few others which can be used accordingly.

In this paper Tensorflow and Keras are used which can be implemented using some basic python and a deep learning model is built.

The following modules are used for development of model using tensorflow and keras.

• Image Data Generator: Utilized for expanding the pictures continuously while the model is as yet running as it saves time.

• Adam: It is the optimization algorithm used to update network weights based on   training.

• Label Binarizer: It is used to represent the given labels as 0,1

• Imutils: A bundle is utilized to perform tasks like interpretation, turn, resizing and so on.

• One-hot Encoding: Converts data from one form to other.

• Blob: It is binary large object, it helps to find only what is required.

3.3 MODEL ARCHITECTURE:

There are numerous models CNN offers like Resnet, Inception, Mobilenet and so on..

In this model MobilenetV2 is used as it is light weight and efficient.

MOBILENETV2 ARCHITECTURE:

MobileNet-v2 is a convolutional brain network that is 53 layers profound. You can stack a pretrained form of the organization prepared on in excess of 1,000,000 pictures from the ImageNet information base. The pretrained organization can group pictures into 1000 article classes, like console, mouse, pencil, and numerous creatures. The MobileNetV2 design depends on a rearranged lingering structure where the information and result of the leftover block are meager bottleneck layers inverse to customary remaining models which utilize extended portrayals in the info a MobileNetV2 utilizes lightweight depthwise convolutions to channel highlights in the halfway extension layer.Moreover, we observe that it is vital to eliminate non-linearities in the thin layers to keep up with illustrative power. We exhibit that this further develops execution and give an instinct that prompted this The MobileNetV2 design depends on a rearranged lingering structure where the information and result of the leftover block are meager bottleneck layers inverse to customary remaining models which utilize extended portrayals in the info a MobileNetV2 utilizes lightweight depth wise convolutions to channel highlights in the halfway plan.

The basic building block could be a bottleneck depth-seperable convolution with residuals. The detailed structure of this block is shown in fig 3

Input	Operator	Output
$h \times w \times k$	1x1 conv2d , ReLU6	$h \times w \times (tk)$
$h \times w \times tk$	3x3 dwise s=s, ReLU6	$\frac{h}{s} \times \frac{w}{s} \times (tk)$
$\frac{h}{s} \times \frac{w}{s} \times tk$	linear 1x1 conv2d	$\frac{h}{s} \times \frac{w}{s} \times k'$

Fig 3: : Bottleneck residual block transforming from k to k 0 channels, with stride s, and expansion factor t.

The engineering of MobileNetV2 contains the underlying completely convolution layer with 32 channels, trailed by 19 leftover bottleneck layers portrayed in Fig 4. We use ReLU6 as the non-linearity in view of its strength when utilized with low-accuracy calculation. We generally use bit size 3×3 as is standard for current organizations, and use dropout and cluster standardization during preparing. Except for the primary layer, we utilize consistent development rate all through the organization.

Input	Operator	t	c	n	s
$224^2 \times 3$	conv2d	-	32	1	2
$112^2 \times 32$	bottleneck	1	16	1	1
$112^2 \times 16$	bottleneck	6	24	2	2
$56^2 \times 24$	bottleneck	6	32	3	2
$28^2 \times 32$	bottleneck	6	64	4	2
$14^2 \times 64$	bottleneck	6	96	3	1
$14^2 \times 96$	bottleneck	6	160	3	2
$7^2 \times 160$	bottleneck	6	320	1	1
$7^2 \times 320$	conv2d 1x1	-	1280	1	1
$7^2 \times 1280$	avgpool 7x7	-	-	1	-
$1 \times 1 \times 1280$	conv2d 1x1	-	k	-	

Fig 4. Each line depicts a succession of at least 1 indistinguishable (modulo step) layers, rehashed n times. All layers inside a similar succession have steady reach c of result channels. The main layer of each grouping has a step s and all others use step 1. All spatial convolutions utilize 3 × 3 parts. The expansion factor t is always applied to the input size as described in Fig 4.

Layers used in the model:

MobilenetV2 model can be implemented on any mobile device as well. This model uses 4 sequential layers: Average pooling layer which calculates the average values for the feature map.

Linear layer with ReLu activation function.

Dropout function and Soft max activation function with result of 2 values.

The Softmax layer gives the outcome as "with cover" or "without veil".

4. RESULTS AND DISCUSSIONS:

Along with the livestreaming video results a graph has been plotted which shows training loss and accuracy. The dataset has been split into training and validation set. The dataset has 1915 images with mask and 1918 images without mask. A graph is also plotted between training loss and accuracy and described in Fig 5.
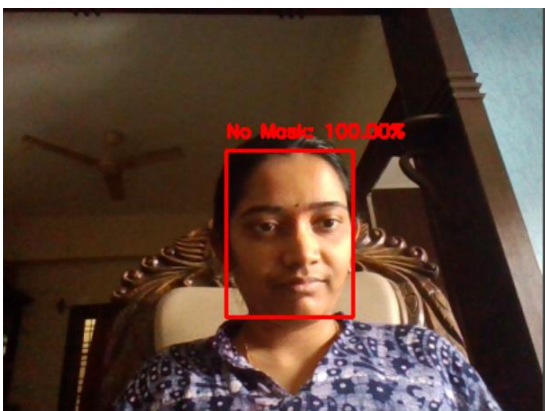
Fig 5: Graph between Training loss and Accuracy

LIVE STREAMING IMAGES:

WITH MASK:
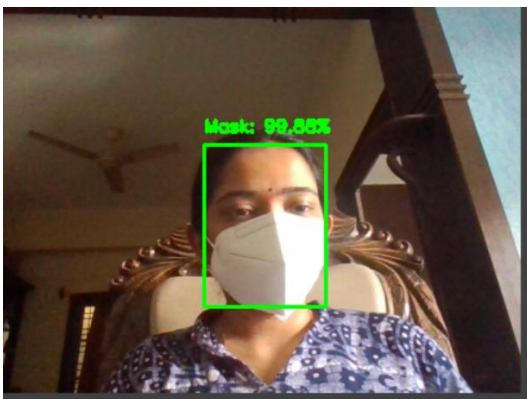
WITHOUT MASK:

5. CONCLUSION:

To reduce the spread of the COVID-19 pandemic, measures should be taken. One of the measure required is wearing a face mask. This project demonstrated a Facemask detector using Convolutional Neural Network and Deep learning techniques. To train, validate and test the model, it utilized the dataset that consisted of 1915 masked face pictures and 1918 exposed face pictures. These pictures were taken from different assets like Kaggle and RMFD datasets. The model was evoked on photos and live video transfers. It is additionally computationally efficient using MobileNetV2 which makes it simpler to introduce the model to inserted frameworks. This facial covering identifier can be sent in various areas like retail plazas, air terminals and other significant traffic spots to separate individuals general and to evade the spread of the contamination by checking who is adhering to fundamental guidelines and who isn't wearing.

6. REFERENCES:

[1] S. Feng, C. Shen, N. Xia, W. Song, M. Fan, and B. J. Cowling, "Rational use of face masks in the covid-19pandemic,"The Lancet Respiratory Medicine, 2020.

[2] Y. Fang, Y. Nie, and M. Penny, "Transmission dynamics of the covid-19 outbreak and effectiveness of government interventions: A data-driven analysis, "Journal of medical virology, vol. 92, no. 6, pp. 645-659", 2020.

[3] Z. Wang, G. Wang, B. Huang, Z. Xiong, Q. Hong, H. Wu, P. Yi, K. Jiang, N. Wang, Y.Peiet al., "Masked face recognition dataset and application arXiv preprintarXiv:2003.09093", 2020.

[4] Z.-Q. Zhao, P. Zheng, S.-t. Xu, and X. Wu, "Object detection with deep learning: A review", IEEE transactions on neural networks and learning systems, vol. 30, no. 11,pp. 3212-3232, 2019.

[5] Qin B. and Li D. , Identifying face mask wearing condition using image super-resolution with classification network to prevent COVID-19.doi: 10.21203/rs.3.rs-28668/v1

[6] Erhan, D., Szegedy, C., Toshev, A., and Anguelov, D. (2014)."Scalable object detection using deep neural networks," in Computer Vision and Pattern Recognition Frontiers in Robotics and AI

[7] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proceedings of the IEEE conference on
 computer vision and pattern recognition, 2014.

[8] W. Bu, J. Xiao, C. Zhou, M. Yang, and C. Peng, "A cascade framework for masked face detection," 2017 IEEE Int. Conf. Cybern. Intell. Syst. CIS 2017 IEEE Conf. Robot. Autom. Mechatronics, RAM 2017 - Proc., vol. 2018-January, pp. 458–462, 2017, doi: 10.1109/ICCIS.2017.8274819.

 [9] N. Ud Din, K. Javed, S. Bae, and J. Yi, "A Novel GAN-Based Network for Unmasking of Masked Face," IEEE Access, vol. 8, pp. 44276–44287, 2020, doi: 10.1109/ACCESS.2020.2977386.

[10] M. S. Ejaz and M. R. Islam, "Masked face recognition using convolutional neural network," 2019 Int. Conf. Sustain. Technol. Ind. 4.0, STI 2019, vol. 0, pp.1-6, 2019,doi:10.1109/STI47673.2019.9068044

 [11] T. Meenpal, A. Balakrishnan, and A. Verma, "Facial Mask Detection using Semantic Segmentation," 2019 4th Int. Conf. Comput. Commun. Secur. ICCCS 2019, pp. 1–5, 2019, doi: 10.1109/CCCS.2019.8888092.

 [12] Y. Heydarzadeh, A. T. Haghighat, and N. Fazeli, "Utilizing skin mask and face organs detection for improving the Viola face detection method," Proc. - UKSim 4th Eur. Model. Symp. Comput. Model. Simulation, EMS2010, pp. 174–178, 2010, doi: 10.1109/EMS.2010.38

[13] R. F. Sear et al., "Quantifying COVID-19 Content in the Online Health Opinion War Using Machine Learning," IEEE Access, vol. 8, pp. 91886–91893, 2020, doi: 10.1109/ACCESS.2020.2993967