JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Facial Mask Detection using Deep Neural Networks

[1]Jagadishwari V, [2]Lakshmi Narayana, [3]Kesavamoorthy R
[1]Department of ISE, [23]Department of CSE
[1 2 3] CMR Institute of Technology, Bengaluru, India

Abstract : The Novel coronavirus has badly affected lives and livelihoods of people in the last two years, the spread of the virus continues to cause havoc on mankind. The medical fraternity have managed to come up with vaccine for the disease, but there no proper cure. Hence it has become very important to contain the disease. One of the various measures recommended by WHO to contain the disease is to wear face mask in public places. People tend to violate the rules and do not wear a face mask, monitoring of people without masks is a tedious job. The work in this paper proposes a methodology to detect people without masks. This can be deployed in the crowded places within cities where cameras are placed to capture images of people. A deep learning architecture is trained with the data set consisting of masked images, unmasked images and images where people have hands are used to cover the face. The novelty of this work is to identify if people are using their hands instead of a mask. This trained architecture provides the accuracy of 98.7% on finding out people with masks, without masks and people covering the face with their hands.

IndexTerms **- Covid 19, Facial Mask , Deep Neural Network**

I. INTRODUCTION

The novel corona virus spreads from one person to another. One of the measures taken to control the spread is the use of a facial mask in public places. The spread of this particular corona virus can be reduced with the help of maintaining a social distancing along with using a facial mask. The covid appropriate behaviour is not followed by the people. It becomes a tedious task for the law enforcing agencies to identify people without masks especially in crowded public places like Supermarkets, the possibility of infection of the coronavirus is very high. Though there are security available at the entrance of the supermarket to make sure people are with masks still few people tend to go in without masks which is a threat to the public safety. This kind of behaviour leads to rise in the infection rates. Therefore, this system focus on detecting people without covid-19 mask, and once detected covid appropriate behaviour can be enforced.

A facial mask detection is analogous to detecting any object [1] [2] from the scene. Though there are many systems which helps in facial mask detection [3] [4], but the proposed system also finds people covering their face with their hand. There are many deep learning techniques which are highly used in the medical imaging applications. The deep learning architectures [5] [6] have a better role in an object detection. This can be done only in the smart cities where public CCTV cameras are available to collect the data. The proposed system will work effectively in malls and shopping complexes where cameras are installed to monitor people. The proposed system to will work effectively , only if CCTV cameras are used to capture the videos of the people in the crowded places. These captured videos are used to extract images containing only the facial information. The Convolutional neural network learning algorithm is used for the feature extraction from the images and the features of the images can be learned from hidden layers of the convolutional networks. When the architecture which is developed identifies the people without mask and the people who closed their face with hand the information will be transferred through the network to their corresponding authorities so that they can take necessary actions. The proposed systems helped to get an promising output on the data being collected from the different sources. A lot of work is been done in this direction ever since the breakout of covid 19. CNN is used by Xinqi et al , the architecture proposed in this paper [7] contains a single stage facial mask detector this copes with different scenes along with that ResNet50 is used for feature extraction which is used
as a backbone network. These features maps with the generated by convolution with the distinct receptive fields and that allows for the detection of objects of varying sizes. They have used a multi-scale dimensional model for the particular point. A a single shot detector has been used by Mayuri S, the paper as gives a better performance with high speed and accuracy. They have introduced some transfer learning algorithms [8] in neural networks to output presence or absence of the face mask in an image or a video stream, the system results showed 100 percent and 99 percent precision and recall respectively. They have designed it using python and AI techniques which will be quick and safe. At this point they recognized person with or without mask or following social distancing. Shilpa Sethi et al [9] have a model where they have popular baseline models that is ResNet50, Alex-Net and Mobile-Net. They have explored mentioned models to plug-in with the proposed models so that the results will be accurate.

Figure 1. Sample Images of face detection

The model proposed has an got accuracy of 98.2 percent when implemented with ResNet50. CNN architectures has been used they have used by [10] which performed classification tasks using available dataset spublically. Here the first CNN is used to decide whether a given chest X-ray images of patient contains covid-19 or not with an average accuracy of 98.92%. the second CNN architecture will e used to divide a given chest x-ray image of a patient into three classes. Experimental results on large clinical data sets show the effectiveness of the proposed architectures

II. METHODOLOGY

The proposed system helps in screening the people with mask, without mask and the people who covered their face with hands. The workflow is shown in Fig 4. Most of the cities nowadays are monitored with the CCTV cameras. These cameras were installed for security purposes, but these images can also be used in facial mask detection. These cameras helps to capture the videos of the people in the public, then the images are extracted from videos are fed into the developed system which identifies person without mask. If a person without the mask then it is sent to the corresponding authority so that they can take necessary actions. Images captured by surveillance camera require prepossessing, making them into a next step. With help of preprocessing step the images are converted into the gray scale images since RGB colour images contains so much of redundant information which isn't required for the Facial Mask detection.

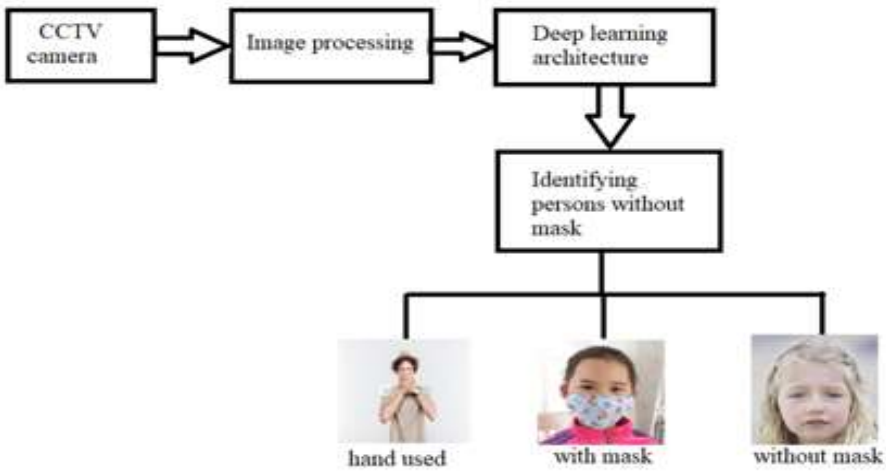

Figure 2. Work flow Diagram

.Every image in the image processing step should be of 24-bit pixel size. And the converted gray scale image has to be in the 8bit for each pixel as it contains the specific required information required for classification. Then the images must be reshaped into 64*64size to have uniformity of images.The workflow is shown in Fig.2, the first step Data preprocessing where the images are processed to get it to the required formats, the second step is to split the data into train and testing sets, about 80% is used to train the model and 20% is utilized in testing the models. The model is later built and trained and tested. The details of the model is discussed in the subsequent sections. Finally any image can be given to the model which will classify it into people wearing masks and not wearing masks.

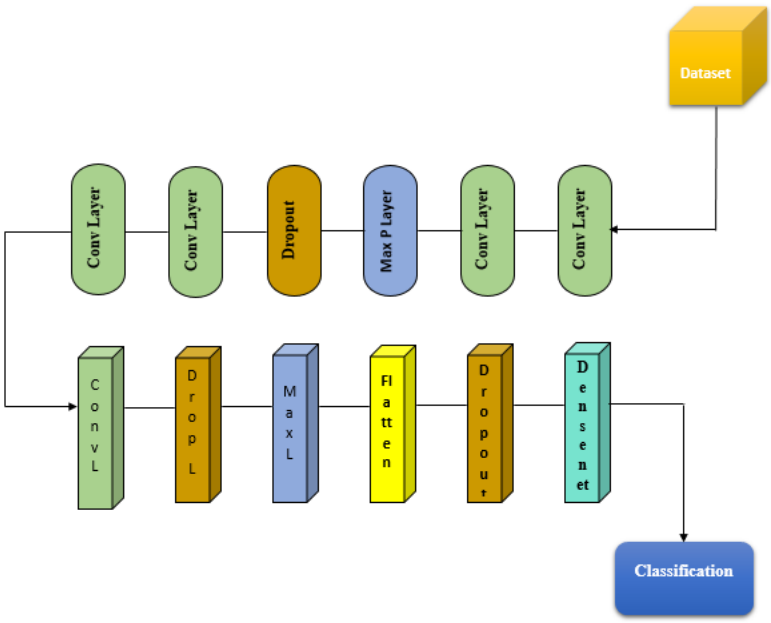

Figure 3. CNN Architecture diagram

Many image processing models use deep learning architectures because they learns many important nonlinear features from the samples. And these learned architectures is used to predict the previous unseen samples. To train the deep learning architecture, images are collected from the various sources. This architecture which is developed highly depends on CNN.

- Data set collection:Dataset consists of images these images which are faces of people which are masked, Unmasked and people covering faces with their hands. Since there was no publicly available data sets these 3 category of images are taken from various sources and a dataset was constructed. A total of 950 images were collected out of which 350 were masked, 400 unmasked and 200 were images with hands used to cover the face. The dataset of 83% is used a training dataset and 17% is used for testing purposes.
- Architecture development: As shown in the block diagram of Fig 3 a CNN architecture is constructed. CNN has several layers such as input layer, many hidden layers and the output layer. The hidden layers has a few convolution layers, they help in feature extraction of the images. Feature extracted images will be used for the later classification purpose. One of the convolution layer is Max pooling layer, this layer is used to decrease the spatial representation size and reduces the number of parameters. So reducing the number of parameter scan result in simplified computation for network. The flatten layer helps reshaping the information into a vector format and will be given into dense network. There are 3 pairs of dropout and dense layers used in the classification purpose. The dense layer helps in non linear features and dropout layer helps in prevents the network from overfitting by dropping the units. Lastly, the classification of the three classes is done using dense layer which has two neurons. The three classes are people with mask, without mask and people covering the face with their hands.

III. Results and Discussion

The problem at hand is a three class classification problem ,a reasonable proportions of all the three classes are collected in the data set. The data set is separated into training set and testing set. The data set have a set of samples where 83% is used in training phase and 17% is used for testing phase. The developed system or architecture is trained for 10 epochs because, training results cause over fitting in the training data. Over fitting appears when a model get to know the unwanted patterns from the training samples. Hence, training accuracy increases as test accuracy decreases. The metrics used to evaluate the model are confusion matrix and accuracy-loss graph. The Confusion matrix shown in Figure 4 shows that there are only 28 miss-classifications for 96 images that were tested. The accuracy-loss graph shown in Figure 5 is a plot of Loss and accuracy against the number of epochs, the figure shows that the accuracy increases with the epochs and loss decreases with the increasing epochs. The graph shows up to 98% accuracy at the tenth epoch and the loss is low which implies that the model is performing well.

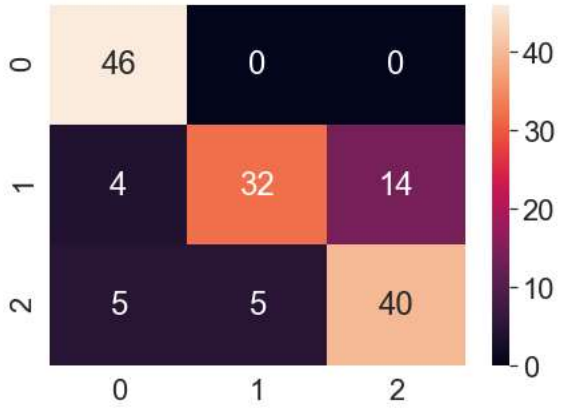

Figure 4. Confusion Matrix

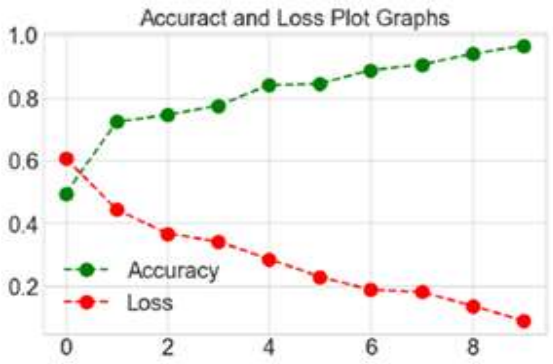

Figure 5. Accuracy-Loss Graph of CNN Model

IV. *Conclusion*

The proposed work in the paper is to identify people not wearing a face mask using the available images. Since peopletend to violate covid norms of wearing a mask in public places, this work will help the competent authorities to findpeople violating the covid appropriate behavior. The system contains a FMD(face mask detection) architecture wherea deep learning algorithm is used so that it detects the mask on the face. To train the model, image data are used whichare images with masks, without a mask and hand used to cover the face. The developed system detects with face maskor without mask with an accuracy of 99.8%.

V. References

1. M. Shah and R. Kapdi, “Object detection using deep neural networks,” in 2017 International Conference on Intelligent Computing and Control
Systems (ICICCS) (IEEE, 2017) pp. 787–790.
2. 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 30, 3212–3232 (2019).
3. 3. A. Sedik, O. S. Faragallah, H. S. El-sayed, G. M. El-Banby, F. E. A. El-Samie, A. A. Khalaf, and W. El-Shafai, “An efficient cybersecurity
framework for facial video forensics detection based on multimodal deep learning,” Neural Computing and Applications 34, 1251–1268
(2022).
4. S. Sethi, M. Kathuria, and T. Kaushik, “Face mask detection using deep learning: An approach to reduce risk of coronavirus spread,” Journal
of biomedical informatics 120, 103848 (2021).
5. J. Ieamsaard, S. N. Charoensook, and S. Yammen, “Deep learning-based face mask detection using yolov5,” in 2021 9th International
Electrical Engineering Congress (iEECON) (IEEE, 2021) pp. 428–431.
6. M. Koklu, I. Cinar, and Y. S. Taspinar, “Cnn-based bi-directional and directional long-short term memory network for determination of face
mask,” Biomedical Signal Processing and Control 71, 103216 (2022).
7. X. Fan and M. Jiang, “Retinafacemask: A single stage face mask detector for assisting control of the covid-19 pandemic,” in 2021 IEEE
International Conference on Systems, Man, and Cybernetics (SMC) (IEEE, 2021) pp. 832–837.
8. G. Jignesh Chowdary, N. S. Punn, S. K. Sonbhadra, and S. Agarwal, “Face mask detection using transfer learning of inceptionv3,” in
International Conference on Big Data Analytics (Springer, 2020) pp. 81–90.

9. M. Shi, L. Xu, and X. Chen, “A novel facial expression intelligent recognition method using improved convolutional neural network,” IEEE
Access 8, 57606–57614 (2020).
E. Irmak, “Implementation of convolutional neural network approach for covid-19 disease detection,” Physiological Genomics 52, 590–601
(2020).
10. E. Ryumina, D. Ryumin, D. Ivanko, and A. Karpov, “A novel method for protective face mask detection using convolutional neural networks
and image histograms.” International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences (2021).
11. M. Mandal, S. Jana, S. K. Nandi, A. Khatua, S. Adak, and T. Kar, “A model based study on the dynamics of covid-19: Prediction and control,”
Chaos, Solitons & Fractals 136, 109889 (2020).