Social Distancing and Face Mask Detection Using Deep Learning

K S Pavan Krishna
Department of Information Science and Engineering
Ramaiah Institute of Technology
Bangalore, India

Harshita S
Department of Information Science and Engineering
Ramaiah Institute of Technology
Bangalore, India

Nayan Deep
Department of Information Science and Engineering
Ramaiah Institute of Technology
Bangalore, India

Gautham M G
Department of Information Science and Engineering
Ramaiah Institute of Technology
Bangalore, India

Deepthi K
Assistant Professor
Department of Information Science and Engineering
Ramaiah Institute of Technology
Bangalore, India

Abstract:- The COVID-19 Pandemic caused by the new Coronavirus is the cause of this 21st-century global health crisis. It has forced the government to impose a lockdown to prevent the transmission of the virus. This led to the unprecedented shutdown of economic activities. The health care system is in crisis. Many different types of safety measures are being taken in order to reduce the risk of the spread of this disease at unprecedented times. Verified reports from renowned scientists and medial health practitioner indicated that wearing a face mask and maintaining social distance reduces the risk of transmitting the virus. Hence, we decided on an approach that is effective and economic by using deep learning techniques to create a safe environment in setups such as manufacturing plants, markets, malls, and other such places. To demonstrate our approach, the training dataset is composed of people, the images with and without the masks, which are collected from a variety of sources and use it in order to build a robust algorithm in order to measure the social distance with the help of the classic geometry methods. Our goal is to determine if a person is wearing a mask, or whether or not they maintain social distance as per protocols and guidelines which are given by leading scientists and governments in this pandemic. We hope that this study will serve as a useful tool for reducing the spread of this dangerous infectious disease in the world.

***Keywords**:- COVID-19, Face Mask Detection, Social Distancing, Deep Learning, CNN.*

I. INTRODUCTION

Covid-19 is an epidemic that is spreading across the nations like wildfire. Though the mortality rate of this viral disease is less, the rate of spreading and people getting infected is very high because of its nature. Covid-19 is becoming the talk of everyone be it in villages or cities, people are afraid of and are in a panic state because of this disease. If the disease is not controlled in the initial stages in countries like India, it will lead to a disastrous kind of situation which will lead to heavy loss of life. The earlier it's controlled, it'll be better for the whole of humanity and particularly for countries like India, which is more densely populated. The main goal of our work is to develop a deep learning model to distinguish between individuals who wear masks and to make sure social distance is being followed.

Fig. 1 People wearing face masks to protect themselves from corona virus infection in the queue in front of a CASH machine on-site, in Dharmsala, India
Source: AP (Associated Press)

I.1 Motivation

The increase in Covid-19 cases has led to numerous deaths. The Coronavirus has been spreading at a really high rate. Hence, we've got an approach by using deep learning techniques combined with data pre-processing tasks to attain effectiveness and potency in the detection of face masks and social distance. It acts as a warning system for people who are not wearing face masks and those who don't maintain social distance. This will help to limit the spread of the virus, securing the safety, and security of the people.

The lack of social distance and people not wearing masks is a major issue we need to solve. Monitoring the crowd manually is very inefficient. We need a better solution until everyone is vaccinated to save lives. Hence, we design a system that uses deep learning to monitor social distancing and classify people without facial masks or coverings.

Fig. 2. Image for representative purpose.
Source: PTI

I.2 Scope and Objectives

The scope of this project is limited to the public areas, which is typical of monitoring by closed-circuit television cameras, e.g., of a bridge or in a shopping mall. It is believed that the scene that is being monitored has to pedestrians, and other objects, such as cars and animals are nonexistent. Here it is, one should be crowds and or in pools of people. Other visual media, such as film, television, fall outside the scope of the directive. We will use the data obtained from fixed surveillance cameras in real-world environments. We will continue to build on the existing research in the area of visual surveillance.

To develop an approach to assist, decrease the spread of coronavirus by monitoring citizens that are:

1. Maintaining Social Distance
2. The wearing of masks at public places.

As a result of this, we are developing a system that makes use of a deep learning model for the classification of people who don't wear masks and don't keep social distances. For example, the output of the model within the frame of the face is shown with the green and the red color frames or boxes that indicate that a person is wearing a mask or not going to be wearing a mask, and, in conjunction with a precision of that prediction.

II. RELATED WORK

A part of the work in relation to social distancing and face detection are briefly explained in this section.

In [1] the paper is able to detect the crime, which results in real-time voice alerts. It is an efficient and cost-effective approach to AI is to create a safe environment on the factory floor. Includes a robust social distancing measurement algorithm. It uses MoibleNetV2 as the core model for person detection and haar cascade for face detection. To detect if people are following social distancing, Euclidean distance is calculated. The model is able to correctly classify the face images, even if the face is partially hidden by the door. The model does not have the ability to detect faces at the camera, the height of which is greater than that of 10-ft.

In [2] the proposed model is built by fine-tuning a pre-trained deep learning model, and InceptionV3. It is trained and tested on a Simulated Masked Face Dataset (SMFD). Image augmentation techniques have been adopted to limit the availability of the data, the better the training and testing of the model. A transfer learning-based approach is proposed, which makes use of the InceptionV3, the pre-trained model to the classification of people who do not wear a mask. The model is a lot better than the opposite; a recently proposed approach is to achieve an accuracy of up to 99.9% of the education, training, and 100% in the testing phase. The data are composed of a limited number of samples, due to the government's regulations concerning the safety and privacy of the individual. It cannot be used in a variety of locations in order to support the person's identification while wearing a mask.

In [3] the paper makes use of deep learning techniques in order to test if a person is wearing a mask or not. It features a three-class classification, i.e., the person who was wearing a mask, or a not-quite-worn a mask or no mask, has been discovered. The mask is to work, with photos and a live video stream. The Id is used for at least the complexity of the structure, and gives you instant results, and you can, therefore, can be used in the CCTV footage. Limitations include the limited availability of data. The paper has been implemented as a technique to correctly classify the face images, if the face is blurred, due to the poor quality of the camera, or should be, in the dark of night the model is not good

In [4] In this paper, it is proposed that a system that is able to limit the spread of COVID-19, and that people who do not wear a mask, in a smart city, where the public areas are monitored by CCTV cameras. If a problem or anomaly is detected, that is, the authority shall be informed, by means of the network. A deep learning architecture based on CNN, which is trained on a dataset composed of images of people with and without the use of masks, which are collected from various sources. The most important architectural training achieved a 98.7% accuracy on the unseen test data. It is the system that stands in the face of difficulties in it, categorisation of the face are covered by a single hand, because it almost looks like the person is wearing a mask. A person without a face mask, or is riding in a vehicle, it was unable to find the person who will be in these kinds of situations. In large, densely populated urban areas, it is extremely difficult to distinguish between every person's face.

In [5] features of the face are in service of a vision-based automated system to detect and recognize the subjects. A precise location, from the face of remains a challenging task nevertheless. A face is correctly detected as long as its height and width aren't greater than fourfold the space between the eyes. Here, time evaluation isn't considered.

In [6] new research from a single Deep Neural Network-Based model for the automatic people detection, tracking, and inter-people, the distance will be calculated from the crowd, with the use of the joint surveillance cameras. In the proposed model, it is equipped with YOLOv4 and is based on the framework, and the reverse perspective, the identification of the people in the right social distancing, and monitoring of adverse conditions, including the people of the partial visibility, and lighting variations at different times of the day. they are able to identify high-risk areas with the greatest potential for the virus to spread, and infection to grow. Overfitting can be an issue with such a model. There is no method or procedure in place to deal with the potential overfitting of the data.

In [7] a hybrid model with an in-depth, and a classic machine learning for face detection. The data consist of a mask and without mask, with images. It makes use of OpenCV to do that in real-time face detection on a live stream, and with the help of a website. The objective is to determine whether the person in the video, they are wearing a mask, all with the help of computer vision and deep learning. The architecture consists of a MobileNet backbone. This paper uses a large data set, it's going to require more period of time to the train the data.

In [8] paper has proposed a face mask detection using image processing which is one of the high-accuracy and efficient face mask detectors. This proposed system mainly consists of three stages: 1. Image pre-processing, 2. Face detection and crop, 3. Face mask classifier. The model might not be able to detect confusing images without a face mask.

III. PROPOSED MODEL, TESTING DESIGN & IMPLEMENTATION

Considering all the studies presented in the previous chapter, we have implemented the system architecture. The accuracy of the model depends on the training and the images considering various situations. This chapter presents the approach for the classification problem using a deep learning algorithm and the system architecture with a detailed description of each stage. The main objective of this work was to construct a model which is able to detect whether or not a person is wearing a mask or not, and if people adhere to the social distance, using a picture of people in a public place. A framework was designed using deep learning techniques to train the model on a set of images that include people with and without masks. The images were analyzed, and it worked fine. An analysis has been made to ensure the accuracy of the model by varying the hyperparameters that give the best accuracy.

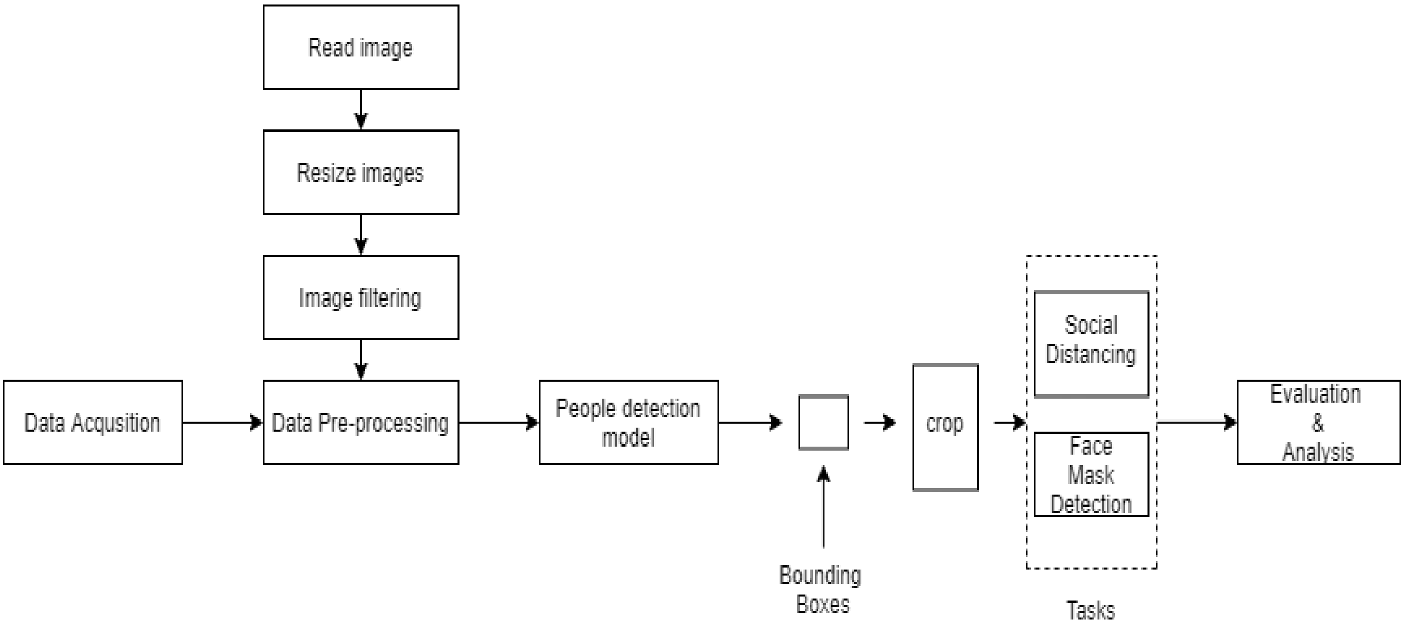

Fig. 3. System Architecture

Based on the literature study done, deep learning models require the data to be available so that the system decides. Hence, the first step of this architecture is data acquisition. This step is required for the collection, processing, and segregation that is the case where the scenarios are based on the roles that are involved in the decision-making cycle of the transfer of the data to the processing unit in order to carry out a further categorization, in the real world, the data consists of noisy, missing values, and, at times, it is in an invalid format, which can then be used directly for the design of the models. Hence, data preprocessing is done to prepare raw data and make it suitable for the models. Classification may be a technique where we categorize data into a given number of classes. The main goal of a classification problem is to spot the category/class to which a replacement data will fall into.

Classification result is the outcome we get after using the classifiers. Evaluation and Analysis is done to check the performance of the algorithm and find the best algorithm that works for our model.

III.1 Data Acquisition

In this first step, we collect images from various sources. Here, we use two classes one with one class consists of images of people with mask and the second class consists of images of people without mask. We name the files "with_mask" and "without_mask" as in the Figure 4 below.

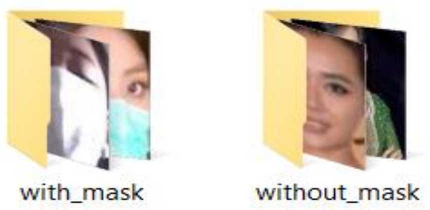

Fig. 4. Classes "with_mask" and "without_mask"

III.2 Data Pre-processing, Testing and Implementation

In the next step, we pre-process the data set so that it can be fed to the model. The aim of this is to improve the features by suppressing or enhancing the image features. The steps included are:

1. Reading the Image

At this point, on the path, we set the image dataset in a variable, and then create a function to convert images into arrays from the folder that contains the images.

2. Resizing image

We resize our image during the pre-processing phase and as the images captured by a camera and fed to our algorithm vary in size. Therefore, we should always establish a base size for all images fed into our AI algorithms. In classification tasks, we usually see images with 224×224 dimensions. It is a suitable dimension that keeps the most structure of the pictures. This doesn't keep the ratio but it's not an enormous problem thanks to resizing images while test time.

3. Image Scaling

Before passing the inputs to the model we scale the input pixels between -1 to 1.

4. Importing libraries

The first thing we do is to import libraries necessary for data preprocessing. Lots of libraries are available, but the most important and popular Python libraries for data are Matplotlib, Pandas and NumPy. For all mathematical things NumPy is generally used Pandas are the best way to import and manage the data. Matplotlib (Matplotlib.pyplot) is the best library to create charts.

5. Encoding categorical data

Category details must be coded before we can use them to match the test model. There are some ways to encode with a variety of models, although the three most common are: We use a hot encoder i.e., we use this data encoding method where the features are named (without an order). In inserting one hot coding, at each level of the category feature, we create new variables. Each category is represented by a binary variant containing 0 or 1. Here, 0 stands for non-existent, and 1 represents the presence of that category.

6. Splitting the Data

At this stage, we split the data into a training set, and the program includes a number of pictures to be trained and tested in the CNN model. Database partitioning is important for the randomized test of the performance prediction. In most cases, we only randomly split the database into three sub-systems:

1. To train the model a training set is used. For example, you can use a training set to find the appropriate weights or coefficients of a linear regression or neural networks.
2. This particular set is used for testing the neutral model. For example, if you find the total number of neurons in a neural network or two main vectors supported by a machine, you need to find different values. For each hypothetical hyperparameter setting, you need to measure the model by examining and evaluating its performance with a set of validation tools.
3. The Test set is for the use of the objective tests of the latest model.

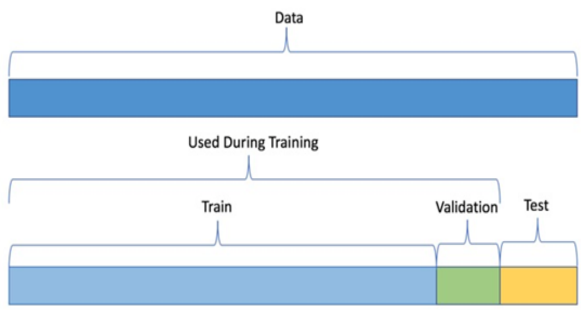

Fig. 5. Visualisation of the dataset split
Source: PATH H2O.ai

It is not difficult to do, so we do not need to configure the hyperparameters, we can only work with the training and the test sets. Before that, the algorithm and the learning of the ML can be an image of similar meaning, as capable of learning the weights and biases of the various items, photos, and be able to distinguish between those of the other.

Here the image dataset undergoes many layers of convolution and pooling which results in a flattening image and then these flattened layers are fully connected. Features are learnt or identified during the convolution and pooling layer whereas classification during latter steps. Compile and train model is done. The model is fed with images containing mask and images not containing mask for training. Once the training is over, we test and validate the model.

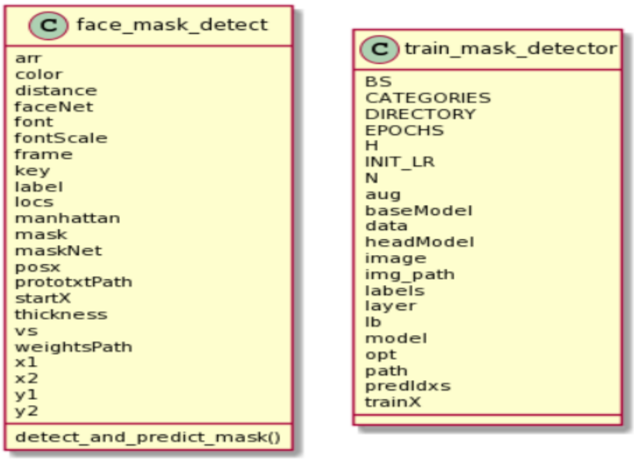

Fig. 6. Class Diagram

The validation set is used to validate the model, but this is one of the normal controls. We did this in order to learn how to use this information in order to optimize the hyperparameters of the model. As for this model, it feels sometimes that information is helpful, but have never really "read" it. We use the hyperparameters to control the settings, and the updates are at the highest level. So, the validation settings impacts the model, however, it is indirect.

The database test is the gold standard, which is used for the sample. This model is fully trained (via the train set). Some of the tests commonly used to test competing models. In the majority of cases, the validation process is the execution of a series of tests that are used, however, this is not a good practice. Some of the tests are generally are well chosen. It contains a carefully coded data of the samples which will be opened by a variety of classes that can be used in the model, when it is being used in the real world.

Data partitioning, classification and information on this is to a large extent, depends on 2 factors. First, the total number of samples in your data and second, on the models we're working on.

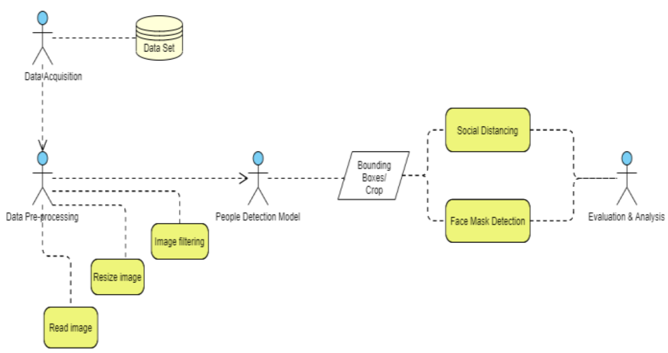

Fig. 7. Use Case Diagram

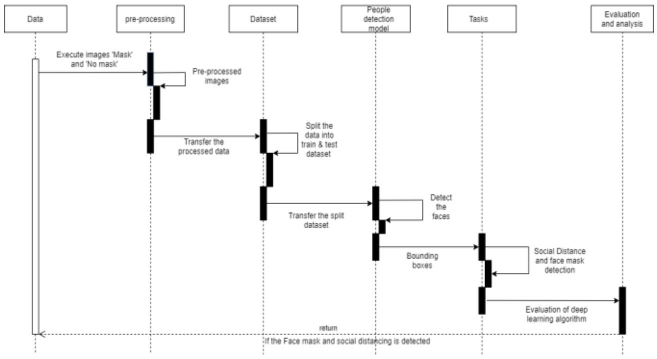

Fig. 8. Sequence Diagram

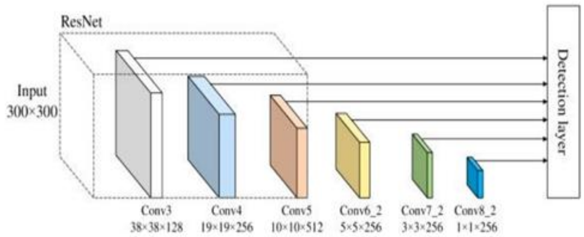

Fig. 9. Resnet and SSD
Source: ijatcse33942020

IV. RESULS

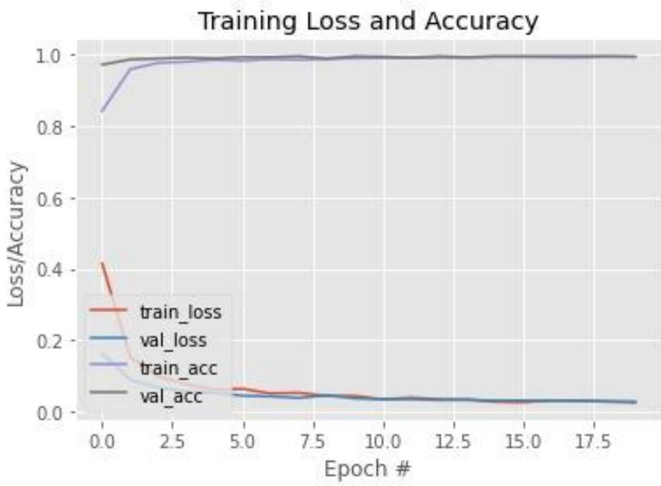

Fig. 10. Training and Validation curve

By putting together all the pieces of our architecture is that we have to rely on the observations, of which Mobile phone NetV2 system, and we are using this system. The resulting system is working, and has the potential to detect masks, and with more comprehensive different points of view.

As the technology is evolving in accordance with new trends, now we have the new detection system, and the masks detection, which can profoundly affect the work of the Department of public Health in India. The structures are built by the MobileNetV2 separator. The current continuous system looks with MobileNetV2 classifier one of the best systems to be used with alarm interface and warning system for the next generation. This plan will be integrated into a system that utilizes social exclusions that could make it a comprehensive system that could have a significant impact on the spread of. The new world will have a greater need for masks as an empty future and that will be a major security concern. Experts say, that facial coverings have been shown to be the best solution to reduce the spread of airborne pathogens such as corona, but as a major safety concern it poses a challenge to the nation as it could create a huge opportunity for people to hide their faces. We have introduced a new method in order to study the problems of face masks and carried out the experimental studies and discussed with them during the operational efficiency of the existing face recognition technology, tools, and developed methods to address the visual acuity and the use is the program of monitoring of the facial and anti-COVID applications. Our results show that all the tests of the models have a much lower performance when you're trying to discover where the faces are covered compared with the ones who have the faces not covered.

	precision	recall	f1-score	support
with_mask	0.99	0.99	0.99	383
without_mask	0.99	0.99	0.99	384
accuracy			0.99	767
macro avg	0.99	0.99	0.99	767
weighted avg	0.99	0.99	0.99	767

Fig. 11. Confusion Matrix

In addition, we recognize that it is possible to design effective face recognition techniques with well-placed masks and that the choice of model art plays a limited role in the final performance of the recognition. Finally, we have shown that existing models of face masks are only limited in real applications, as they only get face detection in images, but not how these masks are applied. Since the testing of the models work well in real-time, we plan to integrate the most efficient methods of a real system. In addition, we plan to extend our analysis to other data sets, which are available to the general public as measures to restrict the spread of COVID-19. The infection will enter into force, as of a certain group of people to use a certain type of mask, it can also be useful to develop models at a later stage which is able to distinguish between the different types of face masks. Here, we have used Manhattan distance to check social distance between people.

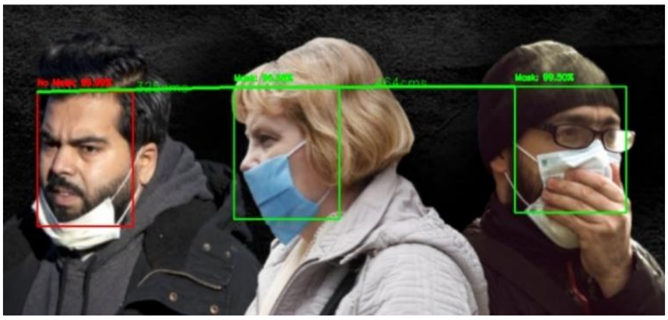

Fig. 12. Resultant Image

Limitations on data sets are a major challenge. The data sets of the face painting are usually small, and the quality of its image is not high enough, compared to standard facial data details. In addition, a variety of incorrect facial expressions increase the difficulty of identification. To overcome these challenges, FaceNet was introduced, which uses the network and transmits learning before splitting. The network solved the problem of low-quality image, while the transfer of learning solved the challenge of using a small database with a variety of face-to-face models. in these ways, performance is greatly improved. To our knowledge, there have not been many studies on colour recognition using in-depth studies. In our study, the condition of facemask wear was obtained at a healthy accuracy of 99.22%.

V. CONCLUSION

In this project, we propose an approach that uses deep learning techniques to monitor the activity of people, ensuring the safety of the people in public places. As the technology is blooming with emerging trends, the availability of new face mask detectors can contribute to public healthcare. We also discussed in detail the social distancing monitoring that helps maintain a safe environment in the post COVID world. The architecture consists of MobileNet as the backbone; it can be used for high and low computation scenarios. The model was tested with images. To get the accuracy of the model and the model optimization is continuing process. We create a high level of accuracy to solve the hyperparameters. In addition, the proposed method makes it possible for you to get the most up-to-the-minute performance of a set of data, as well as with masks and without, and more. By discovering masks we are able to determine if a person is wearing a face mask, and give them the opportunity to do so, and that would be a great help to the community. This solution has the potential to significantly reduce the bad real-world events. We believe that this will help to increase the safety of the people.

FUTURE SCOPE

Future work of this study will be based on using much more efficient models that can predict masks worn improperly. The developed system faces difficulties in classifying faces that have covered their face with hands or improperly worn masks.

A larger dataset could be used as the larger data set helps in better training of the model. Datasets with a massive number of records are good for statistical analysis. The system can be further integrated to form an application that can be used to monitor public places.

ACKNOWLEDGMENTS

We would like to thank our Principal of MSRIT for his support and encouragement. We would like to sincerely express our gratitude to all the faculty and staff of the ISE Department, our Head of Department, our Guide Deepthi K, and to all of our friends to help, with all the means possible in order for us to complete this ambitious work.

REFERENCES

[1]. Prateek Khandelwal, Anuj Khandelwal, Snigdha Agarwal, Deep Thomas, Naveen Xavier, Arun Raghuraman. Using Computer Vision to enhance Safety of Workforce in Manufacturing in a Post COVID World. arXiv 2005.05287 11/05/2020.

[2]. Jignesh Chowdary G., Punn N.S., Sonbhadra S.K., Agarwal S. (2020) Face Mask Detection Using Transfer Learning of InceptionV3. In: Bellatreche L., Goyal V., Fujita H., Mondal A., Reddy P.K. (eds) Big Data Analytics. BDA 2020. Lecture Notes in Computer Science, vol 12581. Springer, Cham. https://doi.org/10.1007/978-3-030-66665-1_6

[3]. Madhura Inamdar, Ninad Mehendale. Real-Time Face Mask Identification Using Facemasknet Deep Learning Network. SSRN 3663305 12/02/2020

[4]. M. M. Rahman, M. M. H. Manik, M. M. Islam, S. Mahmud and J. -H. Kim, "An Automated System to Limit COVID-19 Using Facial Mask Detection in Smart City Network," 2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS), 2020, pp. 1-5, doi: 10.1109/IEMTRONICS51293.2020.9216386.

[5]. R. Padilla, C. F. F. Costa Filho and M. G. F. Costa. Evaluation of Haar Cascade Classifiers Designed for Face Detection. World Academy of Science, Engineering and Technology International Journal of Computer and Information Engineering Vol:6, No:4, 2012

[6]. Mahdi Rezaei, Mohsen Azarmi. DeepSOCIAL: Social Distancing Monitoring and Infection Risk Assessment in COVID-19 Pandemic. Appl. Sci. 2020, 10(21), 7514; 26 October 2020

[7]. Dr. Vandana S. Bhat, Arpita Durga Shambavi, Komal Mainalli, K M Manushree, Shraddha V Lakamapur, 2021, Review on Literature Survey of Human Recognition with Face Mask, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 10, Issue 01 (January 2021)

[8]. Tejal Nerpagar, Sakshi Junnare, Janhavi Raut, Aarti Shah, Prof. P.C.Patil. Face Mask Recognition using Machine Learning. 150498 Volume 7, Issue 7 Page(s): 43 – 45. December 2020

[9]. S. S. Farfade, M. J. Saberian, and L. Li. Multi-view face detection using deep convolutional neural networks. In ACM ICMR, pages 643–650, 2015

[10]. Bellinger, Colin & Sharma, Shiven & Japkowicz, Nathalie. (2012). OneClass versus Binary Classification: Which and When? Proceedings - 2012 11th International Conference on Machine Learning and Applications, ICMLA 2012. 2. 102-106. 10.1109/ICMLA.2012.212.