Surveillance System for Detecting Face-Mask Using Deep Learning Algorithms

[1]Pragatheeswari B K, [2]Pradeepa K , [3]Pavitra S, [4]Jeba Rexciya M

[1,2,3]Student, [4]Assistant professor, Sri Shakthi Institute of Engineering and Technology, Coimbatore, India. [1]bkpragati2000@gmail.com, [2]pradeepakodeeswaran@gmail.com, [3]pavitraselvaraj2831@gmail.com, [4]rexciyajeba@gmail.com,

Abstract: COVID-19 is a big pandemic that is being faced by the whole world and there arises a severe need for a protection mechanism. As a Face mask is the primary one, every person must wear a mask. The project aims to detect the presence of a face mask on human faces on live streaming video. Deep learning has been used to develop our face detector model. The Single Shot Detector (SSD) is used for object detection purposes with its good performance accuracy and high speed. Along with this, basic concepts of transfer learning in neural networks has been used to finally output the presence or absence of a face mask in an image . Experimental results show that our model performs well on the test data with 100% and 99% precision and recall, respectively.. We have used the OpenCV and CNN (Convolutional Neural Network) to detect the presence of a mask and to detect the person's identity. When we find an unauthorized person based on the absence of mask the system will be able to generate an alert e-mail and send it to the concerned e-mail address and if a person wears a mask he is free to move. The graphs here are drawn using Matplotlib library.

Keywords — COVID-19, Preventive Measure, OpenCV, Convolution Neural Network, Email-alert generation

I.INTRODUCTION

There are several countries in the world that have actually made mask-wearing mandatory by law, and it has been observed that certain private organizations in the other countries have also been following their footsteps. In a huge population, it's hard to ensure that people are adhering to these crucial social distancing rules. To easily track such violators, an automated system is an absolute need of the hour. Recognizing this need we have developed a model particularly suited to detect certain violations in real-time. Firstly our model is used to detect people's faces and to determine whether they're wearing a mask. Secondly, it is used to determine whether social distancing is being maintained between 2 individuals, in the most efficient, accurate, and simple manner, hence requiring overseeing authorities to take minimum effort. Here, a face mask detection model has been introduced that is based on computer vision and deep learning. The model, that is proposed can be integrated with surveillance cameras to impede the COVID-19 transmission by allowing the detection of people who are wearing masks not wearing face masks. The proposed model is also an integration between deep learning and classical machine learning techniques with OpenCV, tensor flow and Keras which enriches the result accuracy. Deep transfer leering has been used for feature extractions and combined with three classical machine learning algorithms for better results. A comparison between them has been introduced to find the most suitable algorithm that achieved the highest accuracy and consumed the least time in the process of training and detection.

II. MATERIALS AND METHODS

A. *LINEAR DISCRIMINATE ANALYSIS*

Linear Discriminate Analysis is a method to find a linear combination of features which characterize or separate two or more classes of objects or events.

The resultant gives the Linear classifier. Large number of pixels are used to represent face mask in computerized face mask recognition. Linear discriminant analysis is used to reduce features and makes it more manageable, before classification.

Linear combination of pixel values are new dimensions, which forms a template.

B. *LINEARITY PROBLEM*

Linear Discriminate Analysis technique is used to find a linear transformation that discriminates between different classes. However, if the classes are non-linearly separable, LDA cannot find a lower dimensional space.

In other words, LDA fails to find the LDA space when the discriminatory information is not in the means of classes.

C. *SMALL SAMPLE SIZE*

A sample size also affects the reliability of survey's results because it leads to a higher variability, which may lead to risk. Inflates false discovery rate. Inflate effect size estimation

 DOI : 10.35291/2454-9150.2022.0114

D. *PROPOSED SYSTEM*

Here we are using OpenCV and convolution neural network .By using these methods, we can detect the faces whether it is covered using mask or not. If the student is not wearing the mask ,his/her image will be captured and a mail will be sent to the student with fine .

The main advantage of OpenCV compared to its predecessors is that it automatically detects the important features without any human supervision.

It is used for recording real-time videos at public places for surveillance. Videos are used as an input which is the big advantage of this system, instead of using images.

It is used as a precaution for the human safety from the killer virus like COVID-19 or any other upcoming viruses.

III. DATA FLOW DIAGRAM

The Data Flow Diagram is a simple graphical diagram that can be used to represent a system in terms of input data. Various processing are carried out on this data, and the output data are generated by this system.

The data flow diagram is one of the important modeling tools to model the system components. These components are system process, the data used by the process, an external entity that interacts with the system and the information flows in the system.

Data Flow Diagram shows how the information moves through the system and how it is modified by a series of transformations. It is a graphical technique that depicts information flow and the transformations that are applied as data moves.

A Data Flow Diagram may be used to represent a system at any level of abstraction.

Data Flow Diagram may be partitioned into increasing information flow and functional detail.

FIG 1: Firstly the user accesses the system and them the required classifier is loaded and kept ready. Then the image of the face is extracted to determine if the mask is present or not. Then the remaining work in the side of the admin takes place.

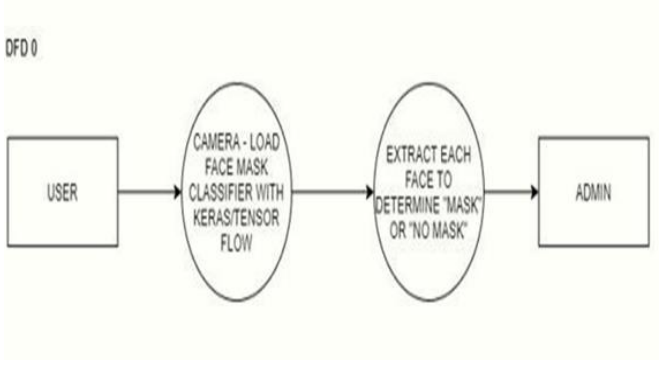


FIG 2: Firstly the user uses the camera present in the system to capture the image and make the classifier ready. Then the required classifier is loaded .Next the image of the face is extracted to determine if the mask is present or not. Finally if the mask is not identified in the face, then a mail with the image and information of the person is sent to the admin.

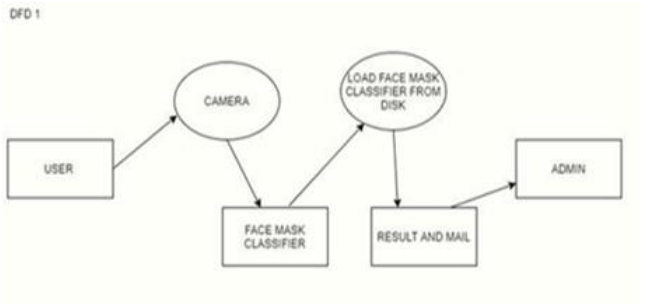
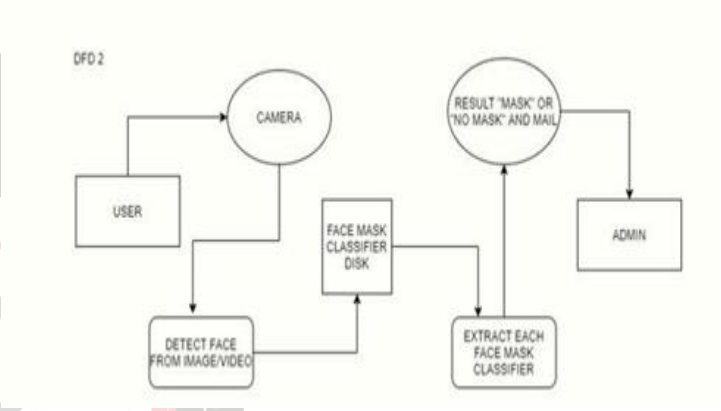
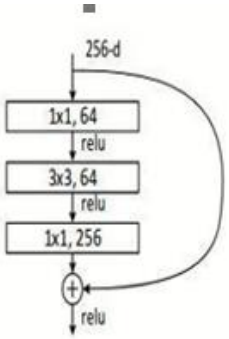


FIG 3: The user uses a camera to perform a live video and also capture image for training. Secondly the face is detected from the video or image. The images are sent to the face mask classifier disk. We then extract each face mask classifier. After the classification we find if the person is wearing a mask or not and then send the image and information of the victim to the admin.



V. ALGORITHM

Residual Networks (ResNet-50) is a convolutional neural network which is 50 layers deep. We can load a pretrained version of network, trained on more than a million images from the ImageNet database. The pretrained network can classify the images into 1000s of object categories, such as keyboard, pencil, mouse, and many animals. The network has learned rich feature representations for a wide range of images as a result. The network has an image with input size of 224 by 224. The bottleneck class implements the layered block. We can train millions of images and then load it as a pretrained version of a network, by using ImageNet database.



 DOI : 10.35291/2454-9150.2022.0114

The classification report of the work which is done is given below for reference.

classification report:

	precision	recall	f1-score	support
with_mask	0.87	0.81	0.84	384
without_mask	0.83	0.88	0.85	386
accuracy			0.85	770
macro avg	0.85	0.85	0.85	770
weighted avg	0.85	0.85	0.85	770

VI. Deep Learning based face liveness detection in videos

An important biometric quantity is a human face which can be used to access a user-based system. As human face images can easily be obtained via mobile cameras and social networks, it would be convenient to have high-resolution images for detection. A reliable face-based access system is used to determine both the identity and the liveness of the input face to get a better understanding. To enrich this, various feature-based face-mask detection methods have been proposed. These methods usually apply a series of processes against the input images in order to detect the liveness of the face. In this paper, a deep-learning-based face-mask detection is proposed. In order to achieve this, two different deep learning models are used, namely local receptive fields (LRF)-ELM and CNN. One of the recently developed models is LRF-ELM which contains convolution and a pooling layer before a fully connected layer that makes the model fast. However, CNN contains a series of convolution and pooling layers. In addition to this, CNN model may have more fully connected layers. A series of experiments were conducted on some face-mask detection databases. On comparing the obtained results we found that the LRF-ELM method yielded better results.

Focusing on Performance comparison of intrusion detection system between DBN Algorithm and SPELM Algorithm. Researchers have used this new algorithm SPELM to perform experiments in the area of face recognition, pedestrian detection, and for network intrusion detection in the area of cyber security. The scholar used the proposed State Preserving Extreme Learning Machine(SPELM) algorithm as machine learning classifier and compared its performance with Deep Belief Network (DBN) algorithm using NSL KDD dataset. The NSL- KDD dataset has four lakhs of data record; out of which 40% of data were used for training purposes and 60% data used in testing purpose while calculating the performance of both the algorithms. The experiment as performed by the scholar compared the Accuracy, Precision, recall and Computational Time of existing DBN algorithm with proposed SPELM Algorithm. The findings show better performance of SPELM; when compared its accuracy of 93.20% as against 52.8% of DBN algorithm;69.492 Precision of SPELM as against 66.836 DBN and 90.8 seconds of Computational time taken by SPELM as against 102 seconds DBN Algorithm.

A centralized server is used to store the videos collected from different surveillance cameras in a networked video surveillance system. These videos can be made available for scrutiny purposes, during an event of security threat and it also has other important uses. Real-time analysis of these surveillance videos can also be helpful in identifying people who restrict the rules in the areas under surveillance. One of the major challenges in video analytics is object detection from video frames. Identifying people in the videos has many applications like handling security through surveillance cameras, crime detection, personalized assistance for the needy, product purchase promotion, employee monitoring and checking people, etc. Accurate person re-identification from videos has huge potential that can revolutionize the way businesses work today. This paper discusses a method to identify the people who do not wear a mask and report the person to higher authorities so that a fine amount could be generated for them. The crucial tasks in this process are face detection from videos and prediction of persons using Convolutional Neural Network models developed using the cropped face images from the face detection stage and it was done perfectly.

Face mask-detection is one of the critical unsolved problems in computer science, a lot of time and energy is spent to invent a robust solution for it. Face mask detection can play a significant role in security, diseases prevention, especially in pandemic situations. The paper presents a unique model that authenticates the faces in the image, and securely shares the images. Initially, the deep learning algorithm identifies the face in the image, extracts the invariant features of each face. The feature of the face is studied to detect the mask in the face. The model captures the face of the people not wearing the mask, compares it with the trained faces, identifies the one not wearing the mask, and sends his details along with his image to the authorized person.

Face mask detection systems are commonly used for verification and security purposes but the levels of accuracy are still being improved. Errors occurring in face-mask detection due to occlusions, pose and illumination changes can be compensated by the use of hog descriptors. The most reliable way to check for a mask is by employing deep learning techniques. The final step is training a good classifier that can take in the requirements from a new test image and tell which known person is violating the rules. A python Based application is being developed to recognize faces for the same.

VII. INPUT DESIGN

The Input design is the link between information system and the user. It comprises of developing the specification and procedures, for data preparation. The processing can be

 DOI : 10.35291/2454-9150.2022.0114

achieved by inspecting the computer to read data from a document. It can also occur by having people by keying the data directly into the system. The design of input focuses on controlling the amount of input required, the errors, avoiding delay and extra steps and keeping the process simple. The input is designed in such a way to provide ease of use with retaining the privacy and security. Input Design considered the following things:

- What data should be given as input?
- How the data should be coded and arranged?
- To guide the operating personnel in providing the input dialog.
- Steps to follow when error occur and methods for preparing input validations.

VIII. OBJECTIVES

i. Input Design is the process of converting a user description of the input into the system. This design is important to show the correct direction to the management for getting correct information from the computerized system and avoid errors in the data input process.

ii. It is achieved by creating user friendly screens handle large volume of data. The goal of designing input is to make data entry free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed and also provides record viewing facilities.

iii. When the data is entered it checks for its validity. Data can be entered with the help of appropriate messages are provided as when needed so that the user will not be in confusions. Thus the objective is to create an input layout that is easy to follow up.

IX. OUTPUT DESIGN

An output design should meet the requirements of the end user and presents the information clearly. Results of processing are communicated to the users of the system and to other system through outputs. The Intelligent and Efficient Output design improves the system's relationship to help user in decision making.

1. Designing computer output should proceed in an organized ,well defined manner. While analysing

2. While analysing the computer's output design, one should Identify the specific output that is needed to meet the requirements.

3. Select methods for presenting information.

4. Create report, document, or other formats that contain information produced by the system.

The output form should accomplish one or more of the following objectives.

- Convey information about current status or projections of the Future or past activities.
- Signal important problems, or warnings, events or opportunities.
- Trigger an action.
- Confirm an action.

X. MODULES

There are four major steps:

- Image Data collection
- Data implementation
- Model Training
- Image predictions

A. IMAGE DATA COLLECTION

In this module, we are going to capture the images which are needed to train and test the model we are going to use. We can capture the images of the person through the take image function which will be stored in a folder and can be used for the tracking process.

B. DATA IMPLEMENTATION

Here the collected images are trained very well. So if a person did not wear a mask his face will be similar to the data collected and he will be recognized as an unauthorized person.

C. MODEL TRAINING

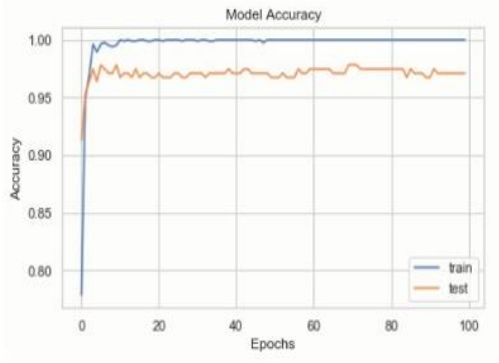
Now, the classified images are given as input to the models which are used to train the machine to detect whether the person is authorized or not.

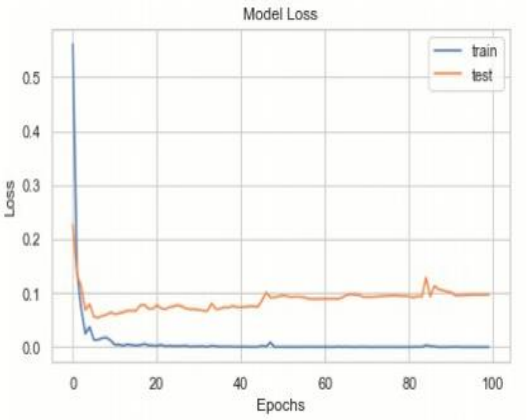
D. IMAGE PREDICTION

Only the authorized person's details will be displayed and when an unauthorized person is detected by the camera it sends an e-mail to the concerned e-mail address. Now, the machine can able to detect and send the e-mail.

The modal accuracy and modal loss diagrams are given below

 DOI : 10.35291/2454-9150.2022.0114





XI. CONCLUSION

To control the spread of COVID-19 pandemic, measures must be taken. Face mask detector was modeled using SSD architecture and transfer learning methods in neural networks. To train, validate and test the model, we used the dataset that consist of masked face images and unmasked face images. The images were taken from various resources like RMFD and Kaggle datasets. The model was inferred on live video streams and images. We evaluated the metrics like accuracy, precision and recall and selected MobileNetV2 architecture with the best performance having 100% precision and 99% recall to select a base model. It is also computationally efficient using MobileNetV2 which make it easier to install the model on embedded systems. This face mask detector can be deployed in many areas like hospitals, work places and chemical and food industries to monitor the public and check who is following basic rules and who is not to avoid the spread of the disease.

XII. REFERENCES

[1] Jeba Rexciya M, Miraclin Joyce Pamila J.C. Sentiment Analysis on Financial News Using Deep Learning Algorithm. Journal of Computer Technology & Application.2021; 12(1): 24-27p.

[2] WHO , ”Coronavirus disease (COVID-19) advice for the public”, 2020, Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public.

[3] Ge, J. Li, Q. Ye and Z. Luo, ”Detecting Masked Faces in the Wild with LLE-CNNs,” 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 426-434, doi: 10.1109/CVPR.2017.53.

[4] Shuo Yang, Ping Luo, Chen Change Loy, Xiaoou Tang, ”WIDER FACE: A Face Detection Benchmark”, 2015, arXiv:1511.06523

[5] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, ”You Only Look Once: Unified, Real-Time Object Detection,” 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.

[6] R. Girshick, J. Donahue, T. Darrell and J. Malik, ”Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation,” 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, 2014, pp. 580-587, doi: 10.1109/CVPR.2014.81.

[7] Narinder Singh Punn, Sanjay Kumar Sonbhadra and Sonali Agarwal, ”Monitoring COVID-19 social distancing with person detection and tracking via fine-tuned YOLO v3 and Deepsort techniques”, 2020, arXiv:2005.01385.

[8] Prateek Khandelwal, Anuj Khandelwal, Snigdha Agarwal, Deep Thomas, Naveen Xavier and Arun Raghuraman, ”Using Computer Vision to enhance Safety of Workforce in Manufacturing in a Post COVID World”, 2020, arxiv:2005.05287.

[9] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam, “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” arXiv:1704.04861.

[10] Alexey Bochkovskiy, Chien-Yao Wang, and Hong-Yuan Mark Liao, ”Yolov4: Optimal speed and accuracy of object detection”, 2020, arXiv:2004.10934.

[11] Joseph Redmon, Ali Farhadi, ”YOLOv3: An Incremental Improvement”, 2018, arXiv:1804.02767

[12] Chien-Yao Wang, Hong-Yuan Mark Liao, Yueh-Hua Wu, Ping-Yang Chen, Jun- Wei Hsieh, and I-Hau Yeh, ”CSPNet: A new backbone that can enhance learning capability of cnn”, IEEE Conference on Computer Vision and Pattern Recognition Workshop (CVPR Workshop), 2020.

[13] Hui, ”YOLOv4”, Medium, 2020. [Online]. Available: https://medium. com/@jonathan hui/yolov4-c9901eaa8e61

[14] Zhanchao Huang, Jianlin Wang, ”DC-SPP-YOLO: Dense Connection and Spatial Pyramid Pooling Based YOLO for Object Detection”, 2019, arXiv:1903.08589.

[15] Sangdoo Yun, Dongyoon Han, Seong Joon Oh, Sanghyuk Chun, Junsuk Choe and Youngjoon Yoo, ”CutMix: Regularization Strategy to Train Strong Classifiers with Localizable Features”, 2019, arXiv:1905.04899

[16] Golnaz Ghiasi, Tsung-Yi Lin, and Quoc V Le. DropBlock: ”A regularization method for convolutional networks”, Advances in Neural Information Processing Systems (NIPS), 2018, pages 10727–10737.

[17] Zhaohui Zheng, Ping Wang, Wei Liu, Jinze Li, Rongguang Ye, and Dongwei Ren, ”Distance-IoU Loss: Faster and better learning for bounding box regression”, Proceedings of the AAAI Conference on Artificial Intelligence (AAAI), 2020.

[18] Zhuliang Yao, Yue Cao, Shuxin Zheng, Gao Huang, and Stephen Lin, ”Cross- iteration batch normalization”, 2020, arXiv:2002.05712.

[19] Sanghyun Woo, Jongchan Park, Joon-Young Lee, and In So Kweon ”CBAM: Convolutional block attention module”, European Conference on Computer Vision (ECCV),2018, pages 3–19.

[20] Shu Liu, Lu Qi, Haifang Qin, Jianping Shi, and Jiaya Jia, ”Path aggregation network for instance segmentation”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pages 8759–8768.

[21] ”Coronavirus Disease 2019 (COVID-19)”, Centers for Disease Control and Prevention, 2020, Available: https://www.cdc.gov/coronavirus/ 2019-ncov/prevent- getting-sick/social-distancing.html

 DOI : 10.35291/2454-9150.2022.0114