International Conference on Artificial Intelligence and Machine Learning
In association withInternational Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com) | Volume 8, Issue 5, July-August-2021

Face Mask and Social Distance Recognition using Deep Learning

Fathima Sholapur[1], S A Quadri[2]
[1]Department of Computer Science and Engineering, VTU/ Secab Institute of Engineering and Technology, Vijayapura, Karnataka, India
[2]Department of CNE, VTU / Secab Institute of Engineering and Technology, Karnataka, India

ABSTRACT

As the COVID-19 pandemic has caused a great despair in a variety of industries all over the ecosphere. The World Health Organization (WHO) certifies wearing a face mask and the practice of corporeal distances in order to reduce the virus's blowout. In this development, a workstation vision arrangement is developed for the automatic recognition of the violation of a mask to wear, and the corporeal distance among employees in the organization. For the face recognition, the broadside is collected and remarked on for over 1000 illustrations; set of data is obtained as an input, maximum up to 1853 photos. Then, it is trained and tested with a multi-Tensor Flow using state-of-the-art object recognition replicas on the aspect concealment to the set of data, and opt for the Nearer R-CNN Inception, ResNetV-2 to a system, which is supplied with an accuracy of up to 99.8%. The Euclidean distance is used to calculate remoteness among various objects under study. A barrier of six feet was kept as a safe distance between the objects. The corporeal distance between two or more than two objects is recognized using the R-CNN network. A real- time video of students entering the campus was shot in SECAB engineering campus and data is fed for learning and training of the proposed model. The proposed system is developed to monitor and improve safety measures by providing information about working masses in the organization by distinguishing them for wearing masks and having social distancing.

I. INTRODUCTION

The spread of COVID-19 has resulted in more than 1,841,000 global deaths and more than 3, 51,000 deaths in the US by Dec. 31, 2020. The spread of virus can be avoided by mitigating the effect of the virus in the environment or preventing the virus transfer from person to person by practicing corporeal distance and wearing façades. WHO defined corporal unsociability as keeping at least six feet or two meters distance from others and recommended that keeping the corporeal distance and impairment of a face mask can remarkably turn down transmission of the COVID-19 virus. Like other sectors, the construction industry has been affected, where unnecessary projects have been suspended or mitigated people's interaction. However, many infrastructure projects cannot be suspended due to their crucial role in people's life. Therefore, bridge maintenance, street widening, highway rehabilitation, and other essential infrastructure projects have been activated again to keep the transportation system's serviceability. Although infrastructure projects are activated, the safety of construction workers cannot be overlooked. Due to the high density of workers in construction projects, there is a high risk of the infection spread in construction sites. Therefore, systematic safety monitoring in infrastructure projects that ensure maintaining the somatic detachment and wearing façade masks can enhance construction workers' safety. In some cases, safety officers can be assigned to infrastructure projects to inspect workers to detect cases that either social distancing or face mask wearing is not satisfied. However, once there are so many workers on a construction site, it is difficult for the officers to determine hazardous situations. Also, assigning safety officers increases the number of people on-site, raising the chance of transmission even more, and putting workers and officers in a more dangerous situation. Recently, online video capturing in construction sites has become very common.

Drones are used in construction projects to record online videos to manage worksites more efficiently. The current system of online video capturing can be used for safety purposes. An automatic system that uses computer vision techniques to capture real-time safety violations from online videos can enhance infrastructure project workers' safety. This study develops a model using Faster R-CNN to detect workers who either don't wear a face mask or don't maintain the corporeal distance in road projects. Once a safety violation occurs, the model highlights who violates the safety rules by a red box in the video.

II. METHODS AND MATERIAL

This research obtains a facemask dataset available on the internet and increases the number of data by adding more images. Then, the project trains multiple Faster R-CNN object recognition models to choose the most accurate model for face mask recognition. For the corporeal distance recognition, the paper uses a Faster R-CNN model to detect people and then uses the Euclidian distance to obtain the people's distance in reality based on the pixel numbers in the image. Transfer learning is used to increase accuracy. The exemplary was functional on multiple videos of students entering in SECAB institute of engineering and technology, to show the performance of the model.

A part of the dataset of face masks was obtained from Make ML website that contains 853 images that each image includes one or multiple normal faces with various illuminations and poses. The images are already explicated with faces with a mask, without mask, and incorrect mask wearing. To increase the training data 1,000 other images with their annotations were added to the database. The total of 1,853 images was used as the facemask dataset. Some samples of images with their annotations are illustrated in

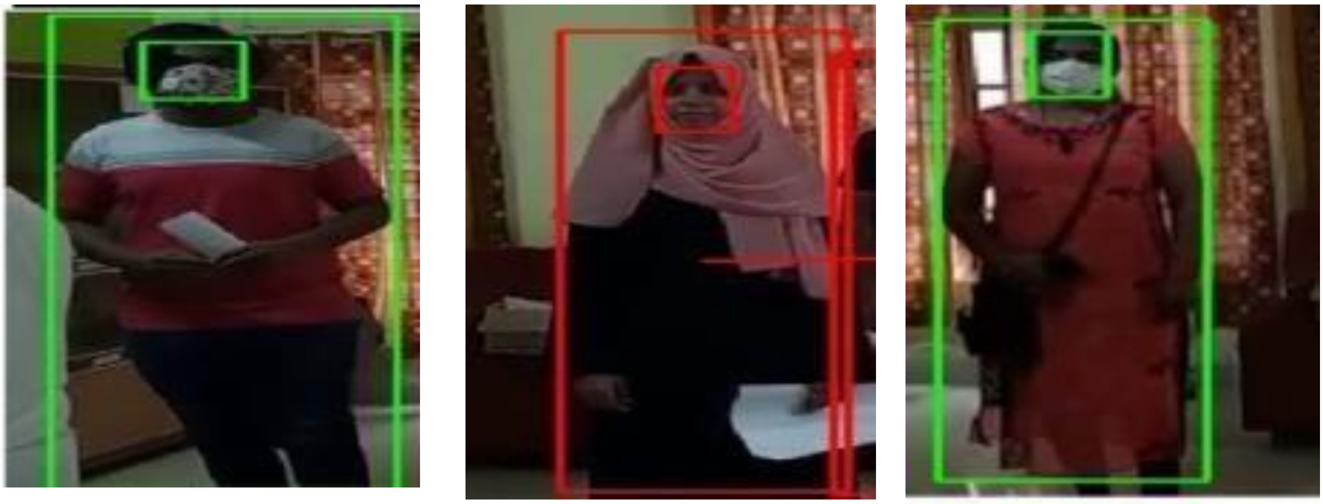

Figure1, where types of face mask wearing are with mask and without mask.

Figure-1 Corporeal distance recognition

The project used the corporeal distancing detector model developed by Roth. The model detects the corporeal distancing in three steps; people recognition, picture transformation, and distance measurement. Roth trained models available on the Tensor Flow object recognition model Zoo on the COCO data set that includes 1,20,000 images. Among all the models, the Faster R-CNN Inception V2 with coco weights was selected as people recognition through model evaluation due to its highest Camera from an arbitrary angle to the bird's eye view. Figure 2 shows the original image captured from a perspective to the vertical view of bird's eye, where the dimensions in the picture have a linear relationship with real dimensions. The rapport amid pixel of (x, y) in the bird's eye picture and pixel of (u, v) in the original picture is defined as:

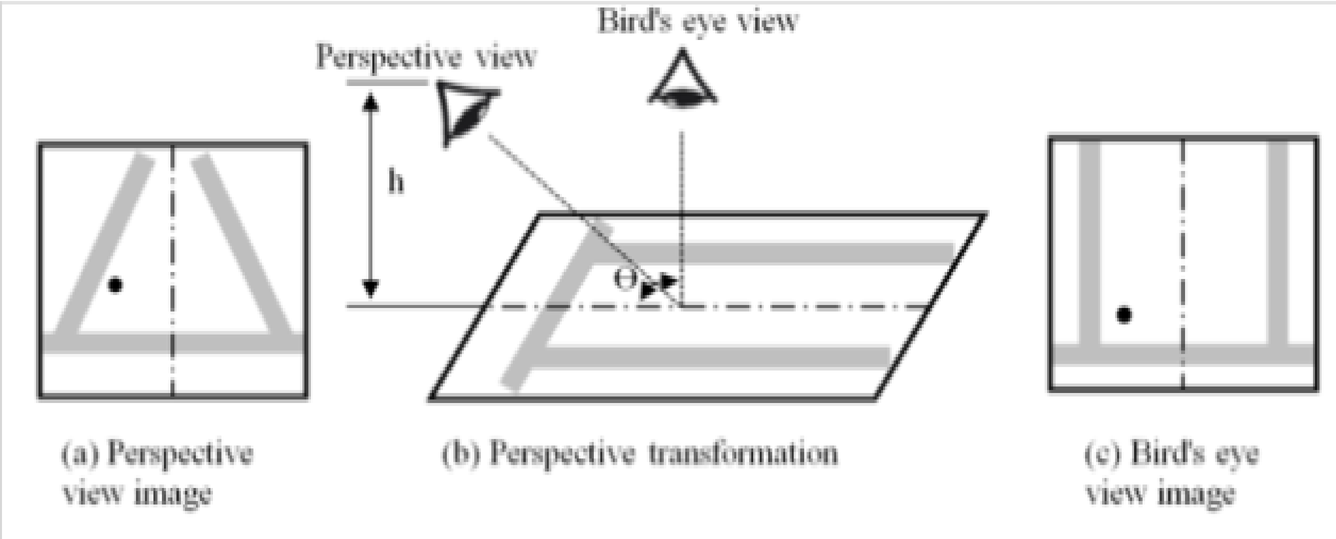

Figure 2. A perspective image transformation to a bird's eye image

Where x=x'/w' and y=y'/w'. For the transformation matrix, the Open CV library in Python was used. Finally, the distance between each pair of people is measured by estimating the distance between the bottom-center points of each boundary box in the bird's eye view. The actual corporeal distance, i.e., two feet, was approximated as 120 pixels in the image.

Faster R-CNN model

Figure 3 shows a schematic architecture of the Quicker R- CNN model. The Quicker R-CNN includes the Region Proposal Network (RPN) and the Fast R-CNN as detector network. The response image is passed through the Convolutional Neural Networks (CNN) Mainstay to extract the features. The RPN then suggests bounding boxes that are used in the Section of Interest (SOI) pooling layer to execute pooling on image's structures. Then, the network passes the output of the ROI pooling layer through two Fully Connected (FC) layers to provide the response of a pair of FC layers that one of them determines the session of each entity and the other one accomplishes a deterioration to improve the anticipated boundary line boxes.

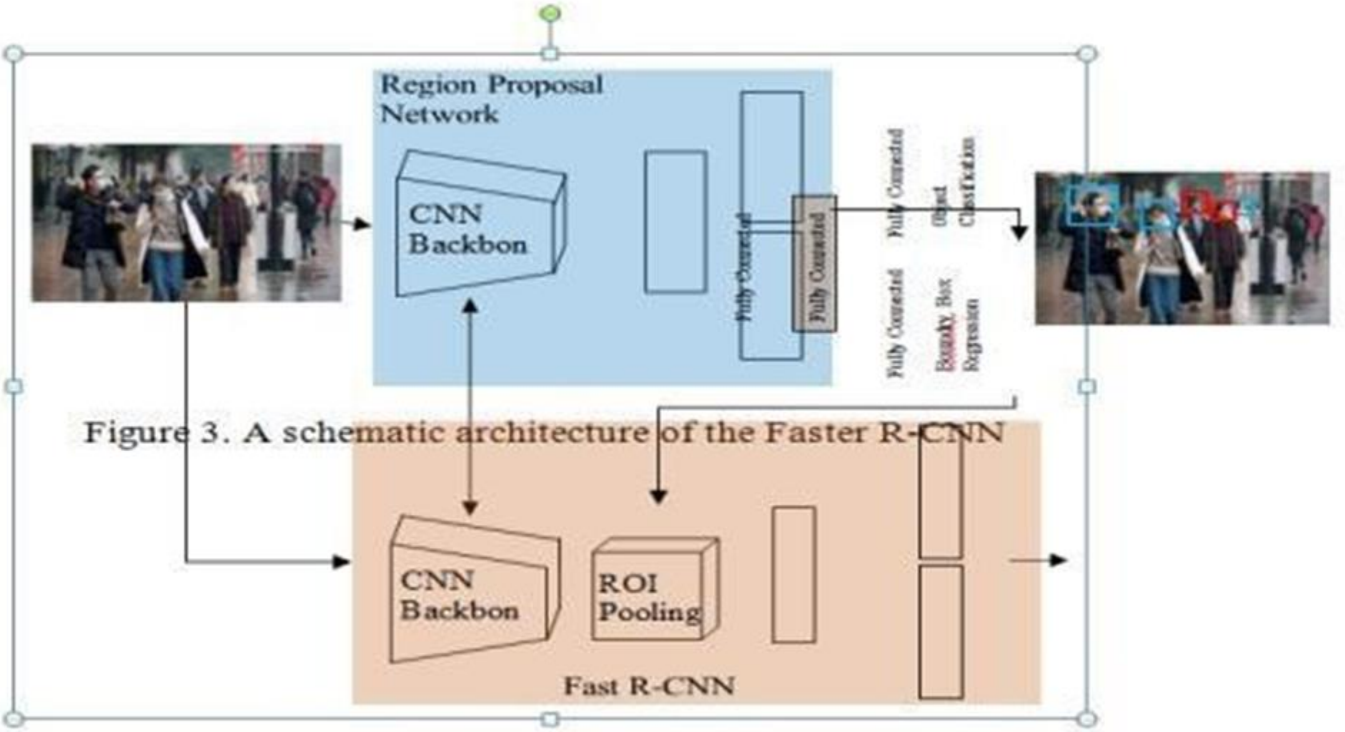

Figure-3.Faster R-CNN model

III. RESULTS AND DISCUSSION

For both networks of face mask recognition and corporeal distance recognition the Google Colaboratory was used. Google Colaboratory is a cloud service developed by Google Research that provides python programming environment for executing machine learning and data analysis codes and provides free access to different types of GPUs including Nvidia K80s, T4s, P4s and P100s and includes leading deep learning libraries. In this experiment, for the face recognition, we used the batch size of 1, the momentum optimizer of value 0.9 with cosine decay learning rate (learning rate base of 0.008), and the image size of 800*1333. The maximum number of steps was 200,000 and the training of the model was stopped when the classification loss reached below 0.07 that happened in near the 42,000th step (Figure 4). For the corporeal distance recognition model, we used the batch size of 1, the total number of steps of 200,000, and the momentum optimizer of value 0.9 with manual step learning rate. The first step was from zero to 90,000, where the learning rate was 2e-4, the second step was from 90,000 to 120,000, where the learning rate was 2e- 5, and the third step was from 120,000 to 200,000, where the learning rate was 2e-6.

Observed Result -1 (2 are safe and 3 are unsafe and 4 are wearing mask

- In this green color anchor box indicates the individuals is safe and individual is wearing mask and maintains social distance.
- Red color anchor box and RED color line indicates individuals are not wearing mask and not maintained social distance.

Observed Result -2(2 are safe and 0 are unsafe and 0 wearing mask)

- In this green colour anchor box indicates the individuals are safe and individual maintained social distance.
- Red colour anchor box indicates individuals are not wearing mask.

IV. CONCLUSION

This project developed a model to detect and differentiate two types of people; one with the face mask and safe distance and the other with without facemasks and social distance, this would be great information for security personnel to enforce safety measures in this COVID-19 pandemic. The project processed dataset including images of people with mask, without mask, and incorrect mask wearing. To increase the training dataset, 1,000 images with different types of mask wearisome were composed and additional to the dataset to create a dataset with 1,853 images. A Faster R-CNN Inception Res Net V2 network was chosen mid different models that generated the accuracy of 99.8% for face mask recognition. For corporeal detachment recognition, the Faster R-CNN

Inception V2 was used to detect people and a transformation matrix was used to remove the effect of the camera angle on actual distance. The Euclidian distance converted the pixels of the transformed image to people's real expanse. The classic set a threshold of six feet as a criterion for distance violation. Transfer learning was used for training models. Four videos of actual students entering in SECAB engineering campus Vijayapur, were used to evaluate the combined model. The output of the four gears indicated an average of more than 90% accuracy in detecting different types of mask tiring in the students. Also, the model accurately detected students/staff that were too close and didn't practice the corporeal distancing. The security personnel can use the model results to monitor students/staff to avoid infection and enhance campus safety. Future studies can employ the model on other public domain organizations such as construction companies and factories. Furthermore, the accuracy of recognition can also be increased by tuning various object parameters.

V. REFERENCES

[1]. M. Afkhamiaghda and E. Elwakil, "Preliminary modeling of Coronavirus (COVID-19) spread in construction industry," J. Emerg. Manag., vol. 18, no. 7, pp. 9–17, Jul. 2020, doi: 10.5055/JEM.2020.0481.

[2]. M. Kamari and Y. Ham, "Automated Filtering Big Visual Data from Drones for Enhanced Visual Analytics in Construction," in Construction Research Congress 2018, Mar. 2018, vol. 2018- April, pp. 398–409, doi: 10.1061/9780784481264.039.

[3]. M. Kamari and Y. Ham, "Analyzing Potential Risk of Wind-Induced Damage in Construction Sites and Neighboring Communities Using Large-Scale Visual Data from Drones," in Construction Research Congress 2020: Computer Applications - Selected Papers from the Construction Research Congress 2020, 2020, pp. 915–923, doi: 10.1061/9780784482865.097.

[4]. Y. Ham and M. Kamari, "Automated content-based filtering for enhanced vision-based documentation in construction toward exploiting big visual data from drones," Autom. Constr., vol. 105, p. 102831, Sep. 2019, doi: 10.1016/j.autcon.2019.102831.

[5]. L. Liu et al., "Deep Learning for Generic Object Recognition: A Survey," Int. J. Comput. Vis., vol. 128, no. 2, pp. 261–318, Feb. 2020, doi: 10.1007/s11263-019-01247-4.

[6]. Z. Q. Zhao, P. Zheng, S. T. Xu, and X. Wu, "Object Recognition with Deep Learning: A Review," IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 11. Institute of Electrical and Electronics Engineers Inc., pp. 3212–3232, Nov. 01, 2019, doi: 10.1109/TNNLS.2018.2876865.

[7]. L. Jiao et al., "A survey of deep learning-based object recognition," IEEE Access, vol. 7, pp. 128837–128868, 2019, doi: 10.1109/ACCESS.2019.2939201.

[8]. R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object recognition and semantic segmentation," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Sep. 2014, pp. 580–587, doi: 10.1109/CVPR.2014.81.

[9]. Z. Zou, Z. Shi, Y. Guo, and J. Ye, "Object Recognition in 20 Years: A Survey," 2019, Accessed: Aug. 02, 2020. [Online]. Available: http://arxiv.org/abs/1905.05055.

[10]. R. Girshick, "Fast R-CNN," 2015. Accessed: Aug. 03, 2020. [Online].Available: https://github.com/rbgirshick/.