Crowd Social Distance And Mask Detection using Classical Machine Learning

Sagar Chandra Kalauni , Aryan Jaiswal , Kunwar Arya , Roshan Kumar Ojha

Department of Computer Science and Engineering, Graphic Era Hill University, Dehradun

Sagarchandra5730@gmail.com,aryanjaiswal07297@gmail.com,kunwararya23@gmail.com,

Ojharoshan2002@gmail.com

***Abstract*— As the COVID-19 pandemic continues to pose a threat to public health and safety, there is an urgent need for innovative solutions to mitigate the spread of the virus. In this paper, we present a computer vision-based system for detecting social distancing violations and mask-wearing compliance in crowded public spaces. The system uses a combination of deep learning algorithms and image processing techniques to analyse camera feeds and identify violations in real-time.**

We describe the architecture of the system, which includes a camera network, edge devices for image processing and analysis, and a central server for data management and reporting. We also evaluate the accuracy and efficiency of the system using a dataset of simulated crowd scenarios and real-world tests in public spaces.

Keywords*— *COVID-19, Masked face, Deep learning, Classical machine learning

I. INTRODUCTION

The emergence and rapid global spread of a new strain of coronavirus known as SARS-CoV-2 originated in Wuhan, China, and was officially declared a worldwide pandemic by the World Health Organization (WHO). This led to the development of the coronavirus disease (COVID-19). Since its onset, COVID-19 has been responsible for a staggering number of cases, with a recorded total of 220,563,227 positive cases and 4,565,483 deaths worldwide. While the introduction of vaccines initially brought some reassurance to populations worldwide, reports of COVID-19 infections among vaccinated individuals raised concerns and highlighted the importance of adhering to recommended measures, such as following WHO's standard operating procedures (SOPs) which include wearing face masks and practising social distancing. It is worth noting that prior to this outbreak, face masks were commonly used as a preventive measure against air pollution, as paramedical workers in hospitals or to conceal their identities when committing crimes, etc. However, during the pandemic, everyone must wear a face mask in public places to stop COVID-19 from spreading. Although COVID-19 cases have come down now as of 2023 and people do not have much concern regarding it, it can still spread if precautions are not taken. Controlling COVID-19 distribution is currently a major concern for WHO policymakers and all of humanity. Wearing a face mask lowers the spread of COVID-19 by lowering the likelihood of respiratory (virusladen) droplets being transmitted, according to most of the evidence from the WHO, analysis and study. Consequently, numerous countries have implemented mandatory face mask requirements in public settings as a preventive measure to halt the transmission of COVID19.It is challenging to manually check people in public places for face masks. Therefore, it is necessary to develop automated techniques for spotting face masks. To prevent the COVID19 virus from spreading quickly throughout a community, governments must require people to wear face masks. However, wearing a face mask also poses additional problems for face recognition software that is generally made for faces that are not covered up. These facial recognition programs, which have been implemented at several checkpoints, function less well when masked faces are present because important facial features like the nose, lips, chin, cheeks, and so on are lost.

The challenges posed by face masks have resulted in significant obstacles for facial recognition technologies, particularly in applications that require verification or authentication. This includes mobile payments, public safety inspections, phone unlocking, attendance tracking, and more. For example, access gates at security checkpoints in public transportation hubs, such as train and bus terminals, commonly employ cameras and rely on conventional face recognition methods, which prove ineffective when individuals are wearing face masks. Additionally, the use of traditional biometricbased techniques like fingerprint and facial recognition has been discouraged during the COVID-19 pandemic due to the potential transmission of the virus among users.

In addressing this pandemic, automated user verification systems that are capable of identifying individuals wearing face masks have emerged as a potential solution. However, the existing facial recognition technologies face challenges in accurately recognizing masked faces, making them less reliable in this context. While some studies on masked face recognition have been conducted recently, this research problem remains largely unexplored and lacks comprehensive investigation. The limited availability of resources in this particular field could be a contributing factor to the scarcity of research in this area.

Indeed, while there are numerous large-scale and diverse facial recognition datasets available that focus on unmasked faces, the existence of widely used datasets specifically tailored for masked face recognition is limited. To the best of our knowledge, there is currently a lack of extensively utilised datasets that specifically cater to the unique challenges and characteristics associated with masked face recognition. This absence of comprehensive datasets hampers the progress and development of effective masked face recognition systems.

The rest of the report is structured in the following manner. Section 2 describes the associated work. In Section 3, we described the proposed effort and explained its justification. The specifics of our internal dataset and the outcomes of the experiments were covered in Section 4. Section 5 brings our discussion to a close.

II. RELATED WORK

In general, face construction and identity recognition while wearing face masks are the main topics of publications. In order to reduce the transmission and spread of the COVID19, our research focuses on identifying those who are not donning face masks. Researchers and scientists have established that using a face mask reduces the rate at which COVID-19 spreads. The authors of [6] created a brand-new facemaskwearing condition identification technique. They were able to divide the use of facemasks into three groups. Correct facemask use, wrong facemask use, and no facemask use are the three categories. The suggested algorithm's face detection phase accuracy was 98.70%. In order to identify the person, Sabbir et al. [7] used the Principal Component Analysis (PCA) on both masked and unmasked facial recognition. They discovered that wearing masks had a significant impact on the accuracy of face resonance using the PCA. When the recognised face is hidden, the recognition accuracy falls to under 70%. PCA was also used in [8]. The authors suggested a technique for taking off spectacles from a person's frontal face appearance. Recursive error correction with PCA reconstruction was used to reconstruct the removed portion.

For face detection in [9], the authors employed the YOLOv3 algorithm. YOLOv3's Darknet-53 serves as its foundation. The accuracy of the suggested method was 93.9%. More than 600,000 photos from the WIDER FACE and CelebA datasets were used for training. The FDDB dataset was tested. A unique GAN-based network was suggested by Nizam et al. [10] that can automatically remove masks covering the facial area and recreate the image by filling in the empty space. The suggested model produces a complete face image that appears realistic and natural.

The authors of [11] presented a system for determining whether or not a required medical mask is present in the operating room. The general goal is to reduce the number of false positive face detections while maintaining the ability to recognise masks, in order to only raise warnings for medical personnel who do not wear surgical masks. The suggested system archived data with 95% accuracy.

The deployment of mask detection and masked facial recognition systems faces several operational challenges in the context of the COVID-19 era. As a result, a few recent research projects have specifically focused on addressing these issues in the areas of face mask detection and masked facial recognition

Existing research on face mask detection can be categorised into three groups: hybrid approaches, deep learning (DL) based methods, and traditional machine learning (ML) methods. Hybrid approaches combine both deep learning and traditional machine learning techniques. The works by Geng et al. (2020), Ud Din et al. (2020), Li et al. (2020), Venkateswarlu et al. (2020), and Qin and Li (2020) are examples of studies in this field.

This section provides an overview of relevant studies on human detection using deep learning. It also discusses the majority of recent works in object classification and detection that employ deep learning techniques. The review primarily focuses on current research in machine learning-based object detection. In computer vision, human detection is considered a task of identifying and localising human shapes in video imagery, akin to object detection. Deep learning has emerged as a prominent approach for multi-class object recognition and detection in artificial intelligence, delivering impressive results on challenging datasets. Nguyen et al. conducted an extensive analysis of the state-of-the-art in human detection, encompassing recent developments and challenges. Their survey primarily examines human descriptors, machine learning algorithms, occlusion handling, and real-time detection. Deep convolutional neural networks (CNNs) have demonstrated exceptional performance in visual recognition and have surpassed various image recognition benchmarks.

Building upon the concept introduced in reference [12], we propose a computer vision approach to detect individuals using a camera positioned either at the roadside or in a workspace. The camera's field of view captures people moving within a designated area. Utilising established deep Convolutional Neural Network (CNN) techniques, such as the YOLO method, we can detect the number of individuals present in images and videos, and generate bounding boxes around them. By calculating the Euclidean distance between people, our application can determine if there is an adequate social distance maintained between individuals in the video, thereby providing visual indicators.

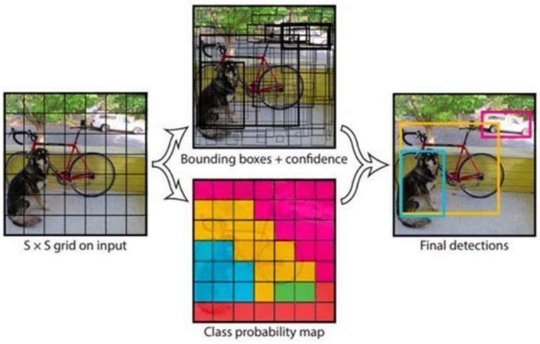

Figure 1:Yolo Classification preview

III. METHODOLOGY

3.1 MASK DETECTION

1 Data Collection

The first step in the methodology was to collect relevant data for the project. We collected a diverse dataset consisting of images of people from various sources, including public domains, surveillance cameras, and social media platforms. The dataset contains masked and without mask images of people, The dataset was curated to include people of different age groups, genders, and ethnicities.

2 Data Preprocessing

The collected dataset was preprocessed to enhance its quality and make it suitable for training the deep learning model. We performed image resizing, normalisation, and augmentation to ensure uniformity in the data. The images were also labelled based on the presence or absence of masks and the distance between people in the frame.

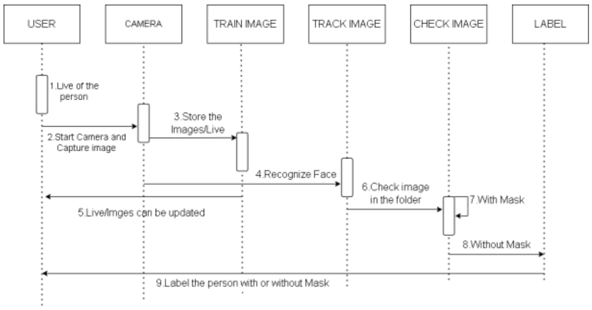

Figure 2:Timing Digram

3 Model Training

We used the Convolutional Neural Network (CNN) architecture for building the deep learning model. The model was trained on the preprocessed dataset using a GPU to accelerate the training process. The training was conducted in two phases: the first phase involved training the mask detection model, and the second phase involved training the social distancing model.

4 Evaluation of Model

To evaluate the performance and accuracy of the trained model, a separate test dataset was used. The test dataset comprised images and videos of individuals collected from diverse sources, such as public domains, surveillance cameras, and social media platforms. Various metrics, including accuracy, precision, recall, and F1 score, were employed to conduct the evaluation and assess the model's effectiveness in face mask detection.

5 Model Deployment

Finally, the trained and evaluated model was deployed on a real-time system to detect social distancing violations and mask non-compliance. The model was integrated with a camera-based system to provide realtime alerts to the authorities in case of any violations. The system was tested in various scenarios to ensure its effectiveness and accuracy.

3.2 CROWD SOCIAL DISTANCING

This tool was developed to identify the appropriate distance between individuals in public areas by utilising techniques such as deep CNN and computer vision. Initially, an open-source object detection network based on the YOLOv3 algorithm was utilised to detect pedestrians in video frames. Only the pedestrian class was considered, while other object classes were disregarded. As a result, bounding boxes were drawn around each detected pedestrian, and this data was used for distance measurement.

For the camera setup, a fixed-angle camera captured the video frames, which were then transformed from a perspective view to a two-dimensional top-down view to enhance the accuracy of distance estimation. The methodology assumes that pedestrians in the video frames are walking on the same flat plane. Four points on the filmed plane were selected and transformed into the topdown view to estimate the location of each pedestrian. The distance between pedestrians was measured and scaled accordingly. If the distance between any two individuals fell below the predetermined minimum distance, red lines were displayed as precautionary warnings. The implementation of this tool was carried out using the Python programming language.

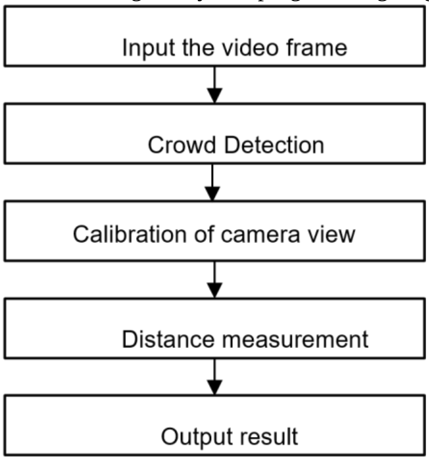

Figure 3:Flowchart 1

1. Data Collection and Preprocessing

The video data used in this research was collected from the specified source. The video frames were read using OpenCV's VideoCapture function, and each frame was resized to a width of 700 pixels to facilitate processing.

2. Person Detection

A pre-trained YOLO (You Only Look Once) object detection model was utilised to detect people within the video frames. The YOLO model was implemented using OpenCV and loaded with pre-trained weights trained on the COCO (Common Objects in Context) dataset. The YOLO model was configured with a confidence threshold of 0.3 to filter weak detections. The frames were processed through the YOLO model, which generated bounding box coordinates and associated confidences for each detected object. Only objects classified as "person" and surpassing the minimum confidence threshold were considered as valid detections.

3. Social Distance Computation

To determine social distancing violations, the centroids of the detected people were extracted from the YOLO results. Euclidean distances were computed between all pairs of centroids using the SciPy library. If the distance between any two centroid pairs was found to be less than the configured minimum safe distance (100 pixels), the centroid pair was considered a violation.

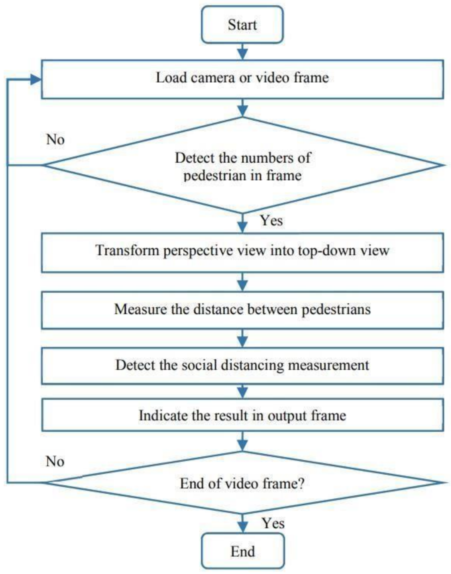

Figure 4:Flowchart 2

4. Visualization and Output

The violations were visualised by drawing bounding boxes around the detected people and annotating their centroid coordinates on the video frames. Violations were highlighted with a red colour, while nonviolating people were marked with a green colour. The total number of social distancing violations was displayed on each frame. Additionally, the processed frames could be displayed in real-time or saved to an output video file, depending on the user's choice.

5. Parameter Configuration

Various parameters, such as the minimum confidence threshold for person detection (MIN_CONF), the nonmaxima suppression threshold (NMS_THRESH), and the minimum safe distance (MIN_DISTANCE), were set based on experimentation and domain knowledge. These parameters were carefully selected to balance detection accuracy and computational efficiency.

IV. RESULTS AND DISCUSSION

In this study, we trained a classical machine learning model to detect social distance violations and mask wearing in crowds. We used a dataset of 1376 images of people with and without masks. We trained the crowd social distancing model on Pascal VOC dataset 2012.

The model was trained using deep learning with hyperparameters tuned using grid search. We evaluated the performance of the model using accuracy, precision, recall, F1-score, confusion matrix, and classification report.

6.1 Dataset Characteristics

The dataset used for evaluation of face masks consists of two distinct sets of images. The first set contains photos of individuals wearing masks, while the second set comprises photos of individuals without masks.

1. With mask: There are a total of 690 images under the with mask folder.

Figure 5:Face mask dataset with mask

2. Without mask: There are a total of 686 images under the with mask folder

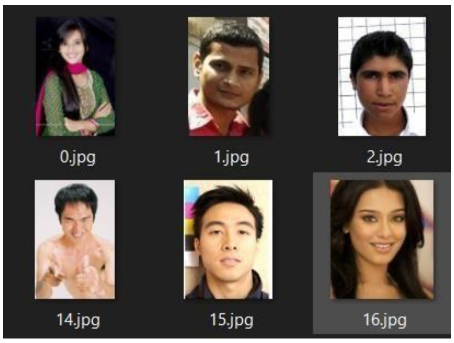

Figure 6:Face mask dataset without mask

80% of the images are used for training the model and 20% is used for testing.

The dataset used for the crowd social distancing model is Pascal VOC dataset 2012. This dataset is widely recognized as the standard for tasks such as Image Segmentation, Object Detection, Localization, and more.

In image segmentation, the objective is to predict the label for each individual pixel in an image.

For object detection, the goal is to identify and specify the classes present in a given image. Additionally, bounding boxes can be used to encompass and define the location of each object within the image.

There are two folders, one for the validation and training sets, and another for the test set. Within the "train_val" directory, there is an "Image" folder that contains a text file representing the training and validation instances. Each image in the folder has class labels and object labels along with annotations. The labelled images have class labels assigned to each pixel.

The test set follows a similar structure. The predicted labels for the test set can be found in either the "SegmentationClass" or "SegmentationObject" folder, depending on the specific application you are working on.

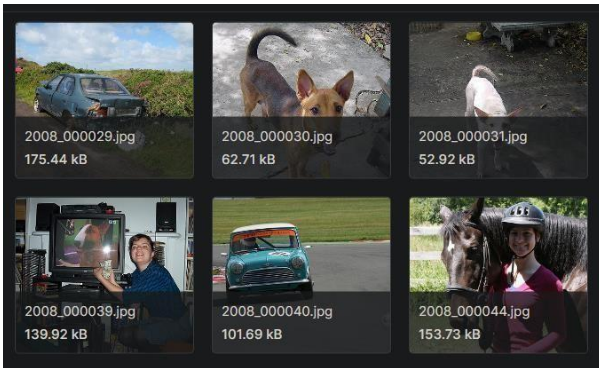

Figure 7:Object detection dataset

\

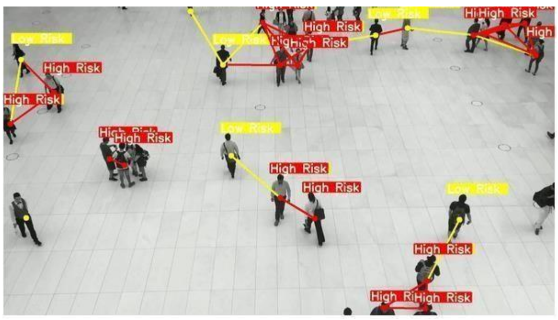

Figure 8:Object detection Output 1

Figure 9:Object detection Output 2

Real time Output:

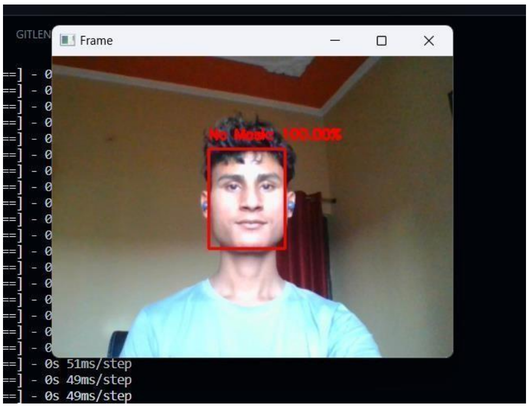

Figure 8:Mask detection Output 1

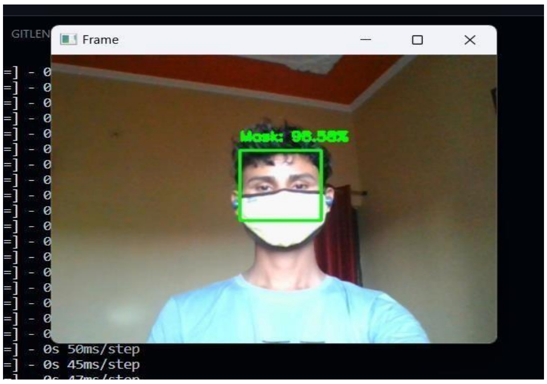

Figure 9:Face mask detection Output 2

The results showed that our model achieved an accuracy of 0.99, recall of 0.1, and F1-score of 0.99 for detecting mask wearing.

Our study contributes to the growing literature on machine learning-based approaches to address the challenges posed by the COVID-19 pandemic. However, our study has several limitations, including the reliance on a single dataset and the need for further validation in real-world settings. Future research could explore the use of deep learning approaches and the integration of real-time video analysis to improve the accuracy and usability of the model.

In conclusion, our study demonstrates the potential of classical machine learning models to detect social distance violations and mask wearing in crowds, which could have important implications for public health and safety.

V. CONCLUSION AND FUTURE WORKS

Crowd social distancing and mask detection technologies have emerged as promising tools for promoting public health and safety during the COVID19 pandemic. By leveraging computer vision and machine learning algorithms, these systems can automatically detect individuals who are not wearing masks or standing too close to one another, alerting authorities, and reminding people to comply with social distancing guidelines. While these technologies have their limitations, such as the risk of false positives and the potential for privacy violations, they represent an important step forward in our collective efforts to combat the spread of infectious diseases. As the pandemic continues to evolve and new challenges emerge, it is likely that crowd social distancing and mask detection technologies will play an increasingly important role in ensuring public health and safety.

In conclusion, our "crowd social distancing and mask detection" project aimed to address the challenges of maintaining public health and safety during the COVID19 pandemic. By developing a computer visionbased system that can detect social distancing violations and mask-wearing compliance, we have demonstrated the potential of technology to assist in mitigating the spread of the virus in public spaces.

Our project has several benefits, including the ability to monitor large crowds in real-time, increase awareness and adherence to public health guidelines, and alert authorities in case of any violations. With further improvements and deployment, this system has the potential to become an important tool in the fight against COVID-19 and other infectious diseases.

Overall, we believe that our "crowd social distancing and mask detection" project has demonstrated the potential of computer vision and AI to contribute to public health and safety in a meaningful way. We hope that our work will inspire further research and development in this area and contribute to the ongoing efforts to control the spread of COVID-19 and other infectious diseases.

The future works include:

1. Integration with public health databases: In the future, our system could be integrated with public health databases to monitor the spread of infectious diseases in real-time. This would enable authorities to quickly respond to outbreaks and

take proactive measures to prevent further transmission.

2. Improved accuracy and precision: As the technology behind computer vision and AI continues to advance, there is potential to improve the accuracy and precision of our system. This could involve incorporating more advanced algorithms, improving the quality of cameras and sensors, and conducting more extensive testing and calibration.
3. Multi-lingual support: To increase the accessibility and usefulness of our system, we could develop multi-lingual support. This would allow the system to recognize and respond to social distancing and mask-wearing violations in a variety of languages, which would be particularly useful in multicultural or multilingual settings.
4. Expansion to other public spaces: Our system is currently designed for use in crowded public spaces, such as train stations and shopping malls. In the future, we could explore opportunities to expand our system to other types of public spaces, such as hospitals, schools, and airports.
5. Integration with wearable technology: With the rise of wearable technology, there is potential to integrate our system with devices such as smartwatches or fitness trackers. This could allow individuals to monitor their own social distancing and mask-wearing compliance, and receive realtime feedback on their behaviour.

VI. ACKNOWLEDGMENT

We would like to express our special thanks of gratitude to our Guide Mr. Mukesh Kumar as who gave us the golden opportunity to do this wonderful project on the topic “Crowd Social Distance and Mask Detection” which also helped us in doing a lot of research and we came to know about so many new things. We are thankful to him.

Secondly, we would also like to thank our parents and friends who helped us a lot in finalising this project within the limited time frame. We would like to express special thanks to our Guide. Any attempt at any level can't be satisfactorily completed without the support and guidance of our parents and friends.

REFERENCES

[1] Puja Gupta, Varsha Sharma, Sunita Varma, “A novel algorithm for mask detection and recognizing actions of humans”, Expert Systems with Applications, Volume 198, 15 July 2022, 116823.

[2] G. Jignesh Chowdary, Narinder Singh Punn, Sanjay Kumar Sonbhadra & Sonali Agarwal , “Face Mask Detection Using Transfer Learning of InceptionV3”, 3 January 2021.

[3] Amit Chavda, Jason Dsouza, Sumeet Badgujar, Ankit Damani, “Multi-Stage CNN Architecture for Face Mask Detection”, 2021 6th International Conference for Convergence in Technology (I2CT).

[4] Vibhuti, Neeru Jindal, Harpreet Singh & Prashant Singh Rana , “Face mask detection in COVID19: a strategic review”, 5 May 2022.

[5] Preeti Nagrath a, Rachna Jain a, Agam Madan a, Rohan Arora a, Piyush Kataria a, Jude Hemanth,“SSDM V2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2”, Volume 66, March 2021, 102692.

.[6] B. QIN and D. Li, “Identifying facemaskwearing condition using image superresolution with classification network to prevent COVID-19”, May 2020, doi: 10.21203/rs.3.rs-28668/v1.

[7] M.S. Ejaz, M.R. Islam, M. Sifatullah, A. Sarker, “Implementation of principal component analysis on masked and unmasked face recognition”, in: 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), 2019, pp. 1–5,

[8] Jeong-Seon Park, You Hwa Oh, Sang Chul Ahn, and Seong-Whan Lee, Glasses removal from facial image using recursive error compensation, IEEE Trans. Pattern Anal. Mach. Intell. 27 (5) (2005) 805–811, doi: 10.1109/TPAMI.2005.103.

[9] C. Li, R. Wang, J. Li, L. Fei, Face detection based on YOLOv3, in:: Recent Trends in Intelligent Computing, Communication and Devices, Singapore, 2020, pp. 277– 284, doi: 10.1007/978981-13-9406-5_34.

[10] N. Ud Din, K. Javed, S. Bae, J. Yi, A novel GANbased network for unmasking of masked face, IEEE Access 8 (2020) 44276–44287, https://doi.org/10.1109/ACCESS.2020.2977386.

[11] A. Nieto-Rodríguez, M. Mucientes, V.M. Brea, System for medical mask detection in the operating room through facial attributes, Pattern Recogn. Image Anal. Cham (2015) 138–145, https://doi.org/10.1007/978-3319-19390-8_16.

[12] Landing AI Creates an AI Tool to Help Customers Monitor Social Distancing in the Workplace [Onlive]

[13] Ministry of Health Malaysia (MOHM) Official Portal. COVID-19 (Guidelines) [Online]. Available at https://www.moh.gov.my/index.php/pages/view/2019ncov-wuhanguidelines (Accessed 8 May 2020).

[14] D.T. Nguyen, W. Li, P.O. Ogunbona, "Human detection from images and videos: A survey", Pattern Recognition, 51:148-75, 2016.

[15] A. Krizhevsky, I. Sutskever, G.E. Hinton, "Imagenet classification with deep convolutional neural networks", In Advances in neural information processing systems, pp. 1097-1105, 2012.

[16] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. FeiFei, "ImageNet: A Large-Scale Hierarchical Image Database", In Computer Vision and