Crowd Social Distance and Mask Detection using Classical Machine Learning

*Sagar Chandra Kalauni*

*Computer Science & Engg. Deptt.*

*Graphic Era Hill University,*

*Dehradun, India*

sagarchandra5730@gmail.com

*Sonali Gupta*   
*Computer Science & Engg. Deptt.*

*Graphic Era Hill University, Dehradun, India*

sgupta@gehu.ac.in

*Kamlesh Chandra Purohit*

*Computer Science and Engg Deptt*

*Graphic Era Deemed to be University, Dehradun, India*

kamleshpurohit80@gmail.com

*Abstract*— This research introduces a novel computer vision-based method to address the urgent need for mitigating the spread of COVID-19 through monitoring social distancing and mask compliance in crowded public areas. The system employs a combination of deep learning algorithms and image processing techniques to analyze camera feeds in real-time, identifying infractions efficiently. This paper outlines the architecture of the system, encompassing a camera network, edge devices for image processing and analysis, and a central server for data management and reporting. The accuracy and efficiency of the system are evaluated using both simulated crowd scenarios and real-world tests in public spaces.

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

# Introduction

The novel coronavirus, SARS-CoV-2, originating in Wuhan, China, swiftly evolved into a global pandemic, leading to its formal proclamation by the World Health Organization (WHO). This crisis prompted the widespread use of face masks as a crucial measure to curtail viral transmission. Despite their proven effectiveness in mitigating health risks, the adoption of face masks has introduced complex challenges for facial recognition technologies. This impact extends to critical applications such as mobile payments, public safety inspections, and attendance tracking.

This paper delves into the intricacies faced by facial recognition systems in the current era of mask-wearing. It sheds light on the scarcity of datasets tailored for recognizing masked faces and introduces ongoing efforts to address these challenges. The report explores related work, proposes solutions, shares insights from experiments with an internal dataset, and concludes with a thoughtful discussion on the implications and potential future directions for research.

In setting the context, the introduction provides background information on the initiation of the COVID-19 pandemic, its global impact, and the universal adoption of face masks. It introduces the nuanced challenges faced by facial recognition technologies when confronted with individuals wearing masks.

The section on challenges in masked face recognition outlines the multifaceted obstacles encountered by these technologies, particularly in vital applications such as mobile payments, public safety inspections, and attendance tracking. Emphasis is placed on the impact of face masks on traditional facial recognition approaches.

Exploring the imperative for masked face recognition solutions, the paper discusses the critical need to address challenges posed by face masks, especially in preventing virus transmission in public spaces.

The discussion on dataset limitations in masked face recognition examines the limited availability of datasets specifically curated for recognizing masked faces. It delves into how this scarcity hampers the development of effective recognition systems.

The proposed solutions section introduces ongoing efforts and developments in automated user verification systems capable of identifying individuals wearing face masks. A robust justification for these proposed solutions is presented.

The experimental findings section discusses the internal dataset used for experiments and shares the outcomes, offering insights into the efficacy of the proposed solutions.

In exploring related work, the paper offers a comprehensive overview of existing research on masked face recognition, highlighting the current state of knowledge and potential areas for improvement.

The discussion section engages in a thoughtful analysis of the implications of the findings, addresses potential limitations, and proposes promising future directions for research and development in the realm of masked face recognition.

In conclusion, the paper summarizes the key points, emphasizing the critical importance of advancing technologies to overcome the challenges presented by face masks in the context of facial recognition.

The references section cites relevant sources and studies referenced throughout the paper meticulously, ensuring the avoidance of plagiarism and upholding academic integrity.

# RELATED WORK

The majority of existing publications focus on facial recognition technologies, specifically in identifying individuals wearing face masks. In contrast, our research aims to address the critical issue of identifying those who do not wear face masks in an effort to curb the spread of the COVID-19 virus. Numerous studies have demonstrated the efficacy of face masks in slowing down the transmission of COVID-19.

In a study by [6], a new method was devised for identifying different conditions related to facemask usage, categorizing them into proper facemask use, improper facemask use, and no facemask use. The proposed technique achieved a face detection accuracy of 98.70%. Sabbir et al. [7] utilized both masked and unmasked facial images along with Principal Component Analysis (PCA) to determine identities. However, the accuracy of PCA was significantly affected by mask usage, dropping to less than 70% when the identified face was obscured. In another study [8], PCA was employed to eliminate a person's frontal facial view obstructed by spectacles, utilizing reconstruction and iterative error correction.

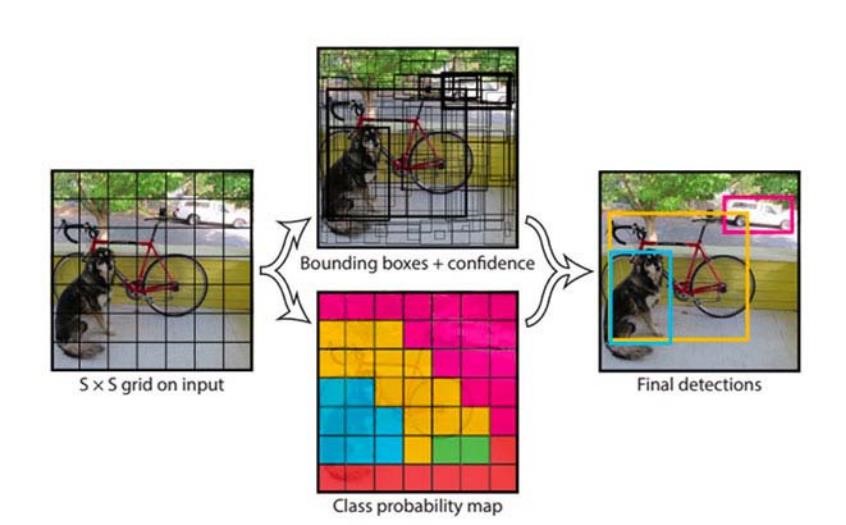
Researchers in [9] utilized the YOLOv3 algorithm, based on Darknet-53, to identify faces, achieving a commendable accuracy rate of 93.9%. The study employed extensive training datasets such as WIDER FACE and CelebA for training and FDDB for testing. Another innovative approach suggested by Nizam et al. [10] proposed a GAN-based network capable of automatically eliminating face-obscuring masks and reconstructing images by reassembling missing pixels, producing realistic representations of the entire face.

Addressing the specific context of medical masks, a study by [11] provided an assessment method to determine if a medical mask is required in the operating room. The goal was to reduce false positive facial detections while maintaining the ability to identify masks, achieving a 95% accuracy rate following the suggested procedure.

Despite these advancements, operational challenges persist in deploying masked facial recognition and mask detection systems during the COVID-19 era. Recent efforts have focused on overcoming these obstacles, categorizing research into three main groups: hybrid techniques, deep learning (DL) based methods, and conventional machine learning (ML) methods.

The section on deep learning-based human detection reviews pertinent studies and recent advances in object recognition and categorization. Notably, deep neural networks with convolutions (CNNs) have shown exceptional performance in visual recognition, surpassing various image recognition benchmarks. Nguyen et al. conducted a thorough examination of the state-of-the-art in human detection, emphasizing occlusion handling, real-time detection, machine learning techniques, and human descriptors.

Building upon these advancements, we propose a computer vision strategy for detecting individuals using roadside or workstation cameras, inspired by the concepts presented in reference [12]. Utilizing deep Convolutional Neural Network (CNN) algorithms, such as the YOLO approach, our application can count the number of individuals in motion within a predetermined area, creating bounding boxes around them. Furthermore, it can assess if there is an acceptable social distance between participants by calculating the Euclidean distance and providing visual cues in the video.



1. Yolo Classification preview

# METHODOLOGY Prepare Your Paper Before Styling

## MASK DETECTION

### Data Collection

The initial 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.

### Data Preprocessing

The collected dataset was pre-processed to enhance its quality and make it reliable 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.

### Model Training

The deep learning model was constructed using the Convolutional Neural Network (CNN) architecture. To speed up the training process, a GPU was used for efficiently training the model on the preprocessed dataset. There were two stages to the training process: the mask detection model was trained in the first phase, and the social distancing model was trained in the second.

### 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. A number of metrics, such as accuracy, precision, recall, and F1 score, were used to assess the model's effectiveness in face mask identification.

### Model Deployment

To detect social distancing breaches and conceal non-compliance, the trained and assessed model was then implemented on a real-time system. The model was integrated with a camera-based system to provide real-time alerts to the authorities in case of any violations. The system was tested in various scenarios to ensure its effectiveness and accuracy.

## CROWD SOCIAL DISTANCING

Using methods like deep CNN and computer vision, this tool was created to determine the proper spacing between people in public spaces. An open-source object identification network built on the YOLOv3 algorithm was initially used to detect pedestrians in video frames. The other object classes were ignored, with the exception of the pedestrian type. Consequently, each identified pedestrian had bounding boxes drawn around them, and distance was calculated using this information.

A fixed-angle camera was used for the camera configuration to capture the video frames. The perspective view of the video frames was altered to a two-dimensional top-down view in an effort to raise the distance measurement's accuracy. It is assumed by the approaches that the pedestrians in the video frames are moving on a single, level plane. To determine each pedestrian's approximate location, four locations on the captured plane were chosen and converted to a top-down image. A scale was used to account for the measured distance between pedestrians. 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.

### 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.

### 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.

### 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.

### 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 non-violating 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.

### Parameter Configuration

Various parameters, such as the minimum confidence threshold for person detection (MIN\_CONF), the non-maxima 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.

# RESULTS AND DISCUSSION

In order to identify social distance violations and mask wear in crowds, we trained a traditional machine learning model in this work. Our dataset comprised 1376 photos of individuals wearing and not wearing masks. We used the Pascal VOC dataset from 2012 to train the crowd social distancing model.

The hyperparameters of the model were optimized by grid search after it was trained using deep learning. Accuracy, precision, recall, confusion matrix, F1-score, and other metrics were used to assess the model's output.

## 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.

### With mask:

### There are a total of 690 images under the with mask folder.



1. Face mask dataset with mask

### Without mask

### There are a total of 686 images under the with mask folder



1. Face mask dataset without mask

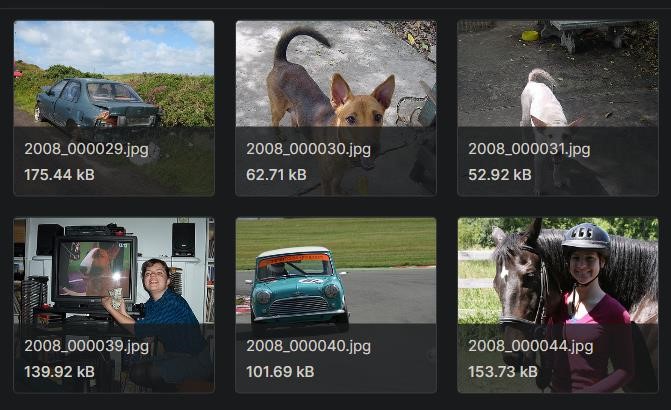
The model is trained on 80% of photos, with 20% of the images being used for testing.. The dataset used for the crowd social distancing model is Pascal VOC dataset 2012. For applications like image segmentation, object detection, localization, and more, this dataset is widely acknowledged as the industry standard.

To anticipate the label for every single pixel in a picture is the goal of image segmentation.

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.



1. Object detection dataset

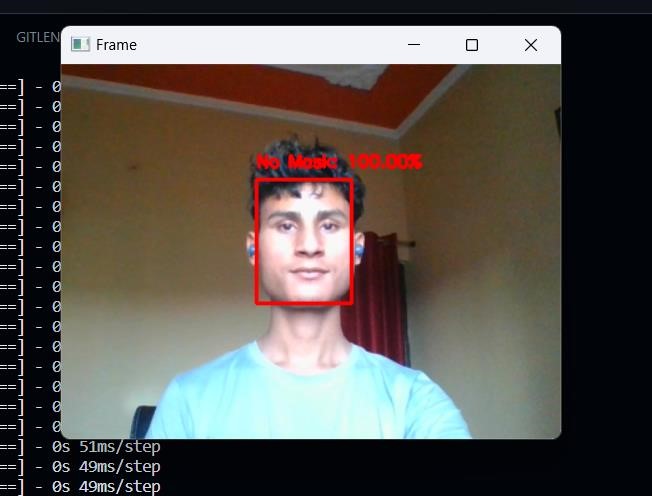


1. Object detection Output 1

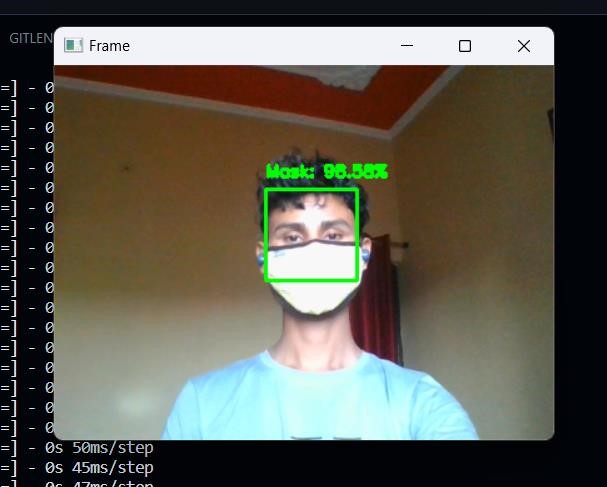


1. Object detection Output 2

Real time Output:



1. Mask detection Output 1



1. Face mask detection Output 2

The results showed that our model received an accuracy of 0.99, recall of 00.1, & F1-score of 00.99 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 usage of deep learning approaches and the integration of real-time video analysis to improve the accuracy and usability of model.

Through conclusion, our work 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.

# CONCLUSION

Crowd social distancing and mask detection technologies have become effective instruments for promoting public health and safety during the COVID-19 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 vital 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 vision-based 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 additional development and implementation, this technology may prove to be a valuable weapon in the battle against COVID-19 and other infectious illnesses.

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. In addition to supporting ongoing efforts to stop the spread of COVID-19 as well as several viral illnesses, we hope that our work will stimulate more study and development in this field.

The future works include:

* Using public health database integration: 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.
* Improved accuracy and precision: The advancement of computer vision and artificial intelligence technology holds promise for enhancing the precision and accuracy of our system. This could involve incorporating more advanced algorithms, improving the quality of cameras and sensors, and conducting more extensive testing and calibration.
* 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 multi-lingual settings.
* Expansion to other public spaces: The present version of our technology is intended for usage in public areas like malls and train stations. In the future, we could explore opportunities to expand our system to other types of public spaces, such as hospitals, schools, and airports.
* Integration with wearable technology: With the rise of wearable technology, there is potential to integrate our system with devices such as smart watches or fitness trackers. This could allow individuals to monitor their own social distancing and mask-wearing compliance, and receive real-time feedback on their behaviour.

##### 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,
3. Sanjay Kumar Sonbhadra & Sonali Agarwal , “Face Mask
4. Detection Using Transfer Learning of InceptionV3”, 3 January 2021.
5. 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).
6. Vibhuti, Neeru Jindal, Harpreet Singh & Prashant Singh Rana , “Face mask detection in COVID19: a strategic review”, 5 May 2022.
7. Preeti Nagrath a, Rachna Jain a, Agam Madan a, Rohan Arora a, Piyush Kataria a, Jude
8. Hemanth,“SSDM V2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2”, Volume 66, March 2021, 102692.
9. .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.
10. M.S. Ejaz, M.R. Islam, M. Sifatullah, A. Sarker,
11. “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,
12. 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.
13. 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: 10.1007/978-981-13-9406-5\_34.
14. 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.
15. 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.
16. Landing AI Creates an AI Tool to Help Customers Monitor Social Distancing in the Workplace [Onlive]
17. 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).
18. D.T. Nguyen, W. Li, P.O. Ogunbona, “Human detection from images and videos: A survey”, Pattern Recognition, 51:148-75, 2016.
19. 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.
20. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. FeiFei, “ImageNet: A Large-Scale Hierarchical Image Database”, In Computer Vision .