JOURNAL OF NORTHEASTERN UNIVERSITY
Volume 26 Issue 01, 2023 ISSN: 1005-3026 https://dbdxxb.cn/ Original Research Paper

LIGHT-WEIGHT FACE MASK DETECTION USING CENTER NET DEEP LEARNING DETECTOR

D. Asiya[1], C. Shoba Bindu[2], P. Dileep Kumar Reddy[3]
[1]Department of CSE, JNTUA College of Engineering, Ananthapuramu, Andhra Pradesh, India, dudekula.asiya@gmail.com
[2]Department of CSE, JNTUA College of Engineering, Ananthapuramu, Andhra Pradesh, India, shobabindhu.cse@jntua.ac
[3]Department of CSE, Narsimha Reddy Engineering College (Autonomous), Secunderabad,Telangana, dileepreddy503@gmail.com

ABSTRACT
The Coronavirus disease had a significant effect on the world. One significant defensive intervention for people is to put on veils in open provinces. A few districts applied a necessary veil wearing principle in open regions to forestall transmission of the infection. Hardly any examination studies have inspected programmed facial covering discovery in light of picture examination. In Covid-19 circumstance is significant thing we ought to safeguard ourselves by wearing covers in open regions. A few districts gave obligatory cover - wearing standard in open regions to keep transmission of the infection starting with one individual then onto the next individual. In our paper, by using CenterNet Deep Learning Detector, Figuring out if someone is using a mask or not. A Red colored rectangular box is accurately formed all around face of any individual in the web camera who is not using a surgical mask/face mask and a Green colored rectangular box is accurately painted all around face of anyone who is donning a mask. Our CenterNet model achieved 99% validation accuracy, 98% testing accuracy, 99% F1-Score, 99% precision accuracy and 99% recall accuracy.
INDEX TERMS: YOLO V3 (You Only Look Once), Face Mask Detection, CenterNet, Deep Learning(DL) Detector, COCO dataset,
COVID-19.

INTRODUCTION
The World Wellbeing Association (WHO) has expressed that Covid sickness 2019 (Coronavirus) had contaminated over 160million individuals and affected more than 3.4 million of deaths overall as of 2021 May [1]. Related enormous scope respiratory sicknesses, extreme intense pulmonary condition (SARS) and Middle East respiratory disorder (MERS), have happened in the last two decades [2], [3]. SARS Covid 2 (SARS-CoV-2), the viral specialist of Coronavirus, has a higher regenerative number than SARS [4]. Expanding quantities of individuals are worried about their wellbeing, and general wellbeing is a significant need of states [5]. Different AI related techniques characterised by the fact been used in medical care to help the recognition of Coronavirus cases from clinical pictures [6]. One issue that cut off points AI strategies for identifying Coronavirus cases is the absence of information. Luckily, generative antagonistic organization-based techniques can be taken on to expand the size of datasets as in. For people, facial coverings could lessen the spread of

Submitted: 03/01/2023 **Accepted**: 23/01/2023

Covid by diminishing their discharge in respiratory beads [7]. Surgical veils, clinical covers, and hand-crafted covers can obstruct around 100 percent, 97%, and 95% of infection particles [8]. Presently, the WHO suggests that individuals ought to wear facial coverings assuming they have respiratory side effects, or on the other hand they are dealing with individuals with side effects [9]. A new report brought up that most conditions and contacts are under states of infection restricted where wearing facial coverings can really forestall infection spread. Locales that had widespread wearing of facial coverings have offered more to the control of Coronavirus than those without this necessity. Numerous public specialist organizations expect clients to wear veils. In any case, certain individuals actually don't wear covers in open regions, which could prompt contamination of themselves or others. Likewise, modified recognizable proof of the wearing of facial covers could help overall society, yet research associated with this is limited.

The detection of face masks, or their wearing, refers to the localization of faces and determining whether or not masks are worn. Other acknowledgment undertakings connecting with facial coverings incorporate distinguishing their administration stage and effectiveness, as these are helpful to identify whether facial coverings can be re-utilized or their quality. These strategies could assume an integral part with facial covering location calculations to shield individuals from Coronavirus. Facial covering recognition frameworks could be sent in observation frameworks, web of things frameworks, or brilliant urban communities to assist public region chiefs with guaranteeing that all guests are wearing veils, to lessen the gamble of the spread of Coronavirus. Facial covering recognition frameworks could replace labourers who need to check the veil wearing status of guests at grocery stores, colleges, libraries, comparative areas.

A few investigations have investigated the discovery of facial coverings. One methodology is a two-step technique which right off the bat recognizes faces utilizing face locators and afterward independently orders whether a facial covering is worn in light of facial covering classifiers[10]. Two strategies might be adequate in certain conditions, but sending the outcomes starting with one phase then onto the next can definitely dial back the cycle Two strategies might be adequate in certain conditions, but sending the outcomes starting with one phase then onto the next can definitely dial back the cycle. Real-time face mask detection may benefit using face mask sensors constructed upon edge convolutional neural network (CNNs) who can both recognize people and locate masks. Even though a hysteresis network-based face mask detector utilising a YOLO theory may attain adequate classification performance, such network is sluggish and heavy for smart objects. With MobileNet as its underlying technology, Retina Facemask offered a lightweight version, however it was unable to overcome the problem of the lightweight model considerably lowering detection performance [11]. Due to the variety of in-the-wild circumstances, including non-veil ends, various types of veils, various face directions, and small or veiled features, diverse facial covering identification challenges arise (Fig. 1).

FIGURES 1. Challenges in face mask detection [12]

II. RELATED WORK

A. DETECTION OF OBJECTS:

The Viola-Jones identifier [13] accomplishes continuous recognition of items by a calculation that concentrates highlights utilizing a Haar include descriptor with a fundamental picture strategy and a flowed locator. Unless the approach makes use of fundamental pictures, it is indeed prohibitively costly. A viable component remover to identify humans, called histogram of arranged slopes (Hoard), figures the headings and extents of situated inclinations over picture cells [14]. Profound learning-based locators can two-stage perform well because of their power and high capacity to extricate highlights [15]. There are two well-known classifications, one-and item detectors. One-stage locators straightforwardly relapse the bouncing box in a solitary step. The methodology in YOLOv1 [16] separated the picture multiple units and attempted to objects in every cell, except this was not great for little items. YOLOv1 did not actually work well by just utilizing the bottom component yield, as the bottom element map has a decent open field and can notice specific regions on the first pictures. Thusly, multi-scale discovery was brought into a single shot indicator (SSD) to direct identification on a few element maps and identify countenances of various sizes [17]. Guo et al. created RetinaNet, that combined an SSD and an FPN architecture and introduced an unique focus objective functions to address the imbalanced class problems, to increase detection capability. In terms of innovation, YOLOv2 is similar to SSDs in that it uses multi-scale features, and YOLOv3 is similar to RetinaNet in that it makes use of an FPN.

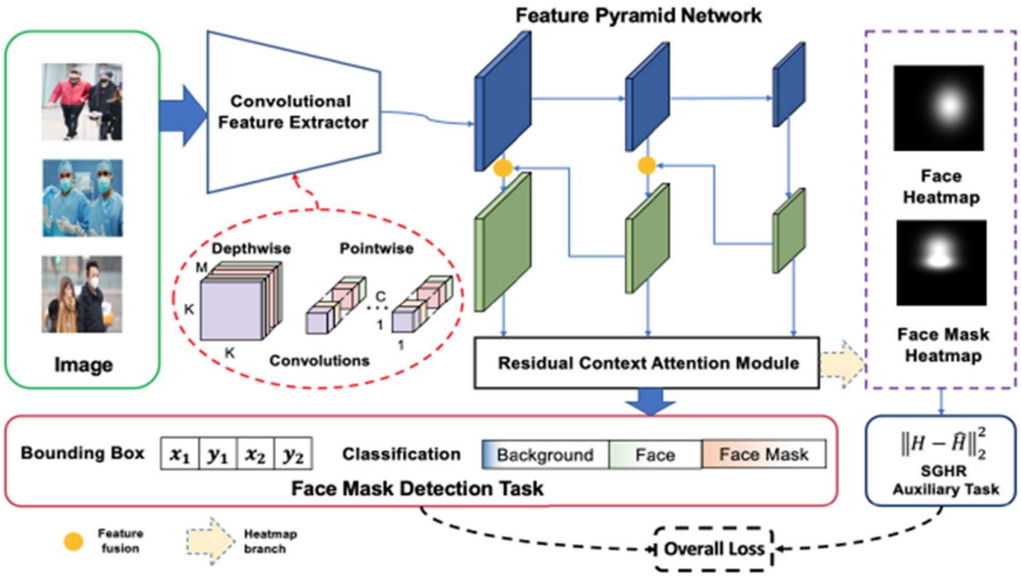


FIGURE 2. The pipeline of the SL-FMDet. [18]

Two-stage detectors. produce locale proposition in the main stage and afterward adjust these recommendations in the subsequent stage. The 2 different finder may provide high-quality distinguishing proof execution anyway at a slow-speed Locale based CNN (R-CNN) [19] utilizes particular pursuit to propose up-and-comer districts that might encompass items. The suggestions are dealt with into a CNN model to remove properties, and an assist vector with machining (SVM) is used to see classes of things. Fast R-CNN tackled this issue by presenting a district of interest (return on initial capital investment) pooling layer to enter all proposed locales immediately [20]. A locale proposition organization (RPN) presented by quicker R-CNN replaced particular pursuit, the speed restricting step of two-stage finders [21]. Quicker R-CNN coordinated every identification part, locale proposition, highlight extractor, and identifier into a start to finish brain network engineering.

B. FACE MASK DETECTION

Facial covering location calculations have become more effective as of late, since veils can assist with controlling the spread of Coronavirus during the pandemic. Among these, just go for it the far more common simulation approaches famous identifiers. ResNet based YOLOv2 was involved by to further develop highlight extraction for facial covering discovery. A distance convergence over association non-greatest suppression (DIOU-NMS) calculation was utilized to further develop the post-handling phase of YOLOv3[22]. YOLOv3 accomplished the most noteworthy Guide in an examination of YOLOv3, YOLOv3-tiny, SSD, and Quicker R-CNN on the recently settled Moxa3K facial covering location dataset. An individual global positioning framework with a three-section facial covering acknowledgment framework, an

individual finder, a tracker, and a facial covering classifier, was created to work with facial covering location applications in brilliant urban communities [23].

The supplementary inadequacies after being recognized fundamentally observing the clear design:

1. While these were a number accessible models that have been periodically re on baseline methods, a few algorithms are already capable of handling data sources relevant to the coronavirus that are hidden from view.

2.These available head covered statistics are sparse and must be reinforced with varying degrees of obstacles and interpretations around different sorts of veils.

3. Despite the fact that there must be two significant sorts of condition of craftsmanship object locators: both solitary and dual-stage sensors. However, neither one of them actually satisfies the criteria for genuine surveillance footage equipment. Lesser computational time power places limitations on these devices. They need better visual feature systems as a result, ones that can execute observation gradually with less memory use and without noticeably losing precision. While two-stage locators can easily deliver precise results for complicated data sources, they do so at the expense of computing time. Solitary identifies are wonderful for strict monitoring but are constrained by poor exactness. A coordinating paradigm for surveillance equipment must be developed in order to take advantage of the multiplicity of factors and using in terms of both time and accuracy.

III. YOLO (YOU ONLY LOOK ONCE)

YOLO Just go for it is a calculation that utilizes brain organizations to give ongoing article identification. This calculation is well known on account given its velocity and exactness. There have utilized in different applications to recognize traffic lights, individuals, stopping meters, and creatures. This is a calculation that identifies and perceives different items in an image (progressively). Object discovery in Consequences be damned is finished as a relapse issue and gives the class probabilities of the identified pictures. Consequences be damned calculation utilizes convolutional brain organizations (CNN) to recognize items in real TIMBE. The computation just needs one forward propagation through the brain organization to differentiate things, as the name suggests. This indicates that the expectation for the entire scenario is completed in a single calculation run. The CNN is used to predict various probability for each class while also bouncing boxes. There are various variants included in the Just go for its computation. Some of the common ones include YOLOv3 and tiny Consequences be damned [22]. They should go for its assessment is important for the following reasons: Speed: Because it may anticipate items in stages, this estimate focuses on the rate of identification. Excellent precision: YOLO Should go for it is a forward-thinking strategy that produces exact results with negligible foundational errors. Learning potential: The computation has excellent educational capabilities that enable it to become accustomed to different depictions and use them in machine vision.

Consequently, the complete image divides it into parts, forecasts hopping boxes and probabilities for each portion. The quintessential possibility weighs those missing rectangles. Within about one backward enlargement along through memory connection, the method "essentially looks on the double" at the image and draws inferences. After non-max disguise, it then communicates identified things (which ensures that the article area computation simply perceives everything once). YOLOV3's core design is highly quick, effective, and simple to use. YOLOV3 is very simple to use on customized objects and is ready to use right out of another screen. Whereas the design approach maintains close to YOLOv4, YOLOV3 derives a significant chunk of its appearance enhancements from PyTorch prepping procedures. The purpose is to create an image detection model that performs well (Y-axis) in comparison to the predictable times (X-axis).

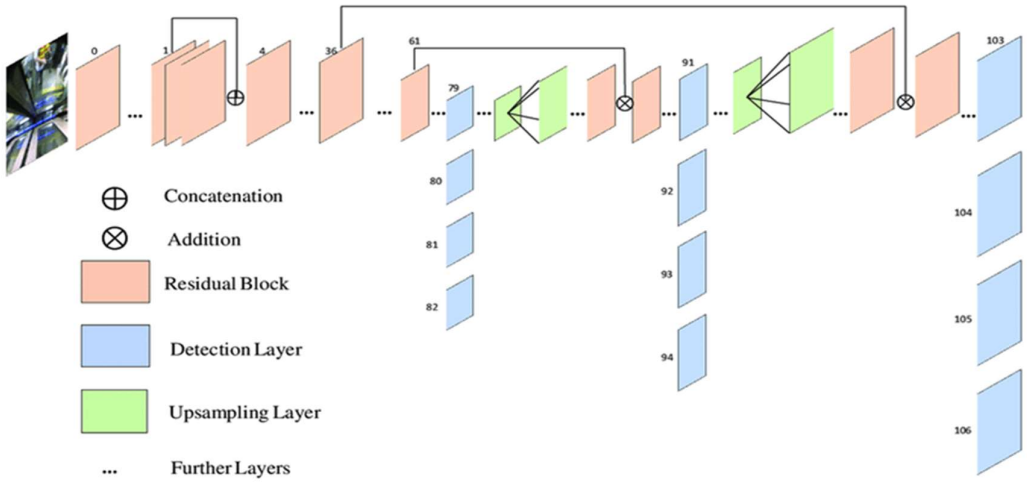


Figure 3: The building style of Yolo [25]

YOLO's architecture is illustrated above (Figure 3). Its architecture mainly consisted of three parts, namely.

1. Backbone: Model Spine is for the most part always had to distinguish value highlights from an information picture. CSP (Cross Stage Halfway Organizations) were utilized as a spine in Consequences be damned YOLO v3 to extricate wealthy in helpful qualities from an information picture

2. Neck: The Model Neck is generally utilised to make include pyramids. Highlight pyramids help models in summing up effectively with regards to protest scaling. It helps with the ID of similar article in different sizes and scales. Highlight pyramids are very helpful in helping models to perform really on beforehand concealed information. Different models, like FPN, BiFPN, and PANet, utilize different kinds of component pyramid draws near. PANet is utilized as a neck in Just go for it v3 to get highlight pyramids.

3. Head: The model Head is generally liable for the last recognition process. It utilizes box anchors to build last result class-probabilistic vectors, bouncing boxes and wretchedness scores.

The YOLO V3 model's face is identical to that of the YOLO V2 model.

- Benefits and Drawbacks of YOLO v3
- Comparing it to YOLOv2, it is 88 percent smaller (27 MB vs 244 MB).
- Compared to YOLOv2, it is around 180 percent faster (140 FPS vs 50 FPS).
- On the identical assignment, it is approximately as accurate as YOLOv2 (0.895 mAP vs 0.892 mAP).
- However, the fundamental issue is that, unlike prior YOLO versions, YOLOV2 has not had an official article published. Additionally, YOLO v3 is still being worked on, and since Ultralytics sends us frequent updates, the developers may tweak some settings in the future.

IV.DATASET Of COCO

A sizable entity identification, captioning dataset and segmentation called MS COCO [24] was released by Microsoft. The COCO database is frequently used by Computer vision and machine learning experts for a variety of design projects. Comprehension visible situations is one of computers vision's main objectives; it entails identifying the items that are visible, localizing them in 2D and 3D, figuring out their characteristics, and defining the connections between them. As a result, the dataset can be used to train systems for item object identification categorization. The image database was produced with the objective of raising image identification, hence the name COCO, which stands meaning Common Objects in Context. The COCO data comprises difficult, striking aesthetic data for computer vision, which are primarily used by cutting-edge neural networks. For instance, COCO is frequently used to test algorithms and evaluate context of physical threat detection. Advanced neural network libraries immediately understand the COCO dataset's structure.

V. PROPOSED WORK

In businesses or other places of employment where many people arrive for work, this study suggests a method to determine whether or not a face veil is worn. Convolutional brain networks have been used by us for a very similar task. The model was developed using data from the real world and was successfully tested with real-time, live video. Additionally, the subject's accuracy is completed with several persons at varying distances from the edge and varied hyper limits. As a result, the layer of the proposed CNN types of approaches face sometimes with covers and includes two neurons with Softmax initiation to characterize phenomena that are quite comparable. The unfortunate capacity is complete trans. The suggested framework has a 99 percent accuracy rate. A red rectangular box with the text "NO MASK" is placed all around face of everyone in the video stream who is not using a surgical mask, and a green rectangular box with both the text "MASK" is drawn around the face of anyone who is. If someone on the live broadcast is not wearing a gas mask for protection, a red bounding box is painted around their head and the dialogue "NO MASK" is displayed. Similar to this, a person wearing a MASK has an ecofriendly bounding box drawn over their face.

Training:

- Produce your own explanation record and class names document: Line Arrangement: image_file_path box 1, box2, … boxN.
- Box design:, x_max, y_max, x_min, y_min, class_id
- Convert the pre-prepared loads top configuration as expected by Keras.
- Freeze all layers with the exception of the last layers and train for certain ages until Level (no improvement stage) is reached.
- Thaw every one of the layers and train every one of the loads while persistently diminishing the learning rate until again level is reached.
- End

Utilizing OpenCV for test results:

Use a camera or another stored videos assessment documents to record the assessment.
Pass every video or webcam-captured contour forward through model.
Place the rectangles on the framework in accordance with the scores, classes and boxes that were received as result (color depending on the class of each box).
Just at bottom of the page, determined the number of boxes in both classes and save the amount in a variable.

Developing the model:

Following trained with Tensorflow-GPU == 2.5.0, our model has achieved 99 percent efficiency for face mask detection without the need for a mask.
We tested the effectiveness of the customized computational intelligence model to determine whether or not such a person wears a face mask using our own photographs. mask sans mouth mask

FIGURE 4: CNN Model [27]

Actual mask and social detachment identification are becoming more and more valuable as a result of the pressing need to regulate COVID-19. Our full face detection is efficient and uses no dataset of altered masked photos. Its efficiency in computation makes it simpler to implement the paradigm in embedded devices (Raspberry Pi, Google Coral, etc.). Due to the Covid-19 outbreak, this technology can therefore be employed in real-time applications that demand face-mask detection for security reasons. To strengthen public safety regulations are followed, this project can be connected with embedded technologies for use in airports, train stations, offices, schools, and public areas.

Proposing a new prioritization method for regression test cases

Face Mask Detection Using MobileNetV2 and CNN

Technologies for face mask identification are already in growing market because of the increased number of COVID-19 cases. This technology is utilized in implementations that need to identify face masks for security reasons. Here, we've constructed a convolutional neural network with MobileNetV2 for face mask detection.

Algorithm for face detection:

1: Start
2: Input: Extricate picture from the edge
3: Load face recognition model
4: If (face is identified)
apply image processing
else
return to Input.
5: Load the facial covering recognition model and recognize the veiled face.
6: If (veil identified)
display bouncing box and precision rate with title "cover"
else
display bouncing box and precision rate with title "no cover".
7: Return to Input.

CENTER NET FACE MASK DETECTION DETECTOR FOR DEEP LEARNING:

We have once more identified the face mask by using a different method called Center Net Deep Learning Detector. Although the accuracy calculation is somewhat complicated, the output is more exact and the speed is quick. It has a number of layers: - Input: the location of the photo.
Core: picture object detection.
Neck: a collection of feature maps from several levels.
Head: output that includes object classes and bounding boxes.
The AP and FPS of the YOLOV2 have increased by 10% and 12%, respectively.

ALGORITHM FOR MASK DETECTION:

1.Input
2. Input < Clone DarkNet
3. Load the assistant capability
4. Load YOLO weights
5. Load the data sets
6: Train the model
7. Test the model on images/videos

8. Loop: If (veil recognized)
Then bouncing box and precision rate with title "cover"
else
display bouncing box and precision rate with title "no cover".
9. Performed simulations with a live video stream.
10. Outcome go to: Loop
11. Repeat go to 9.

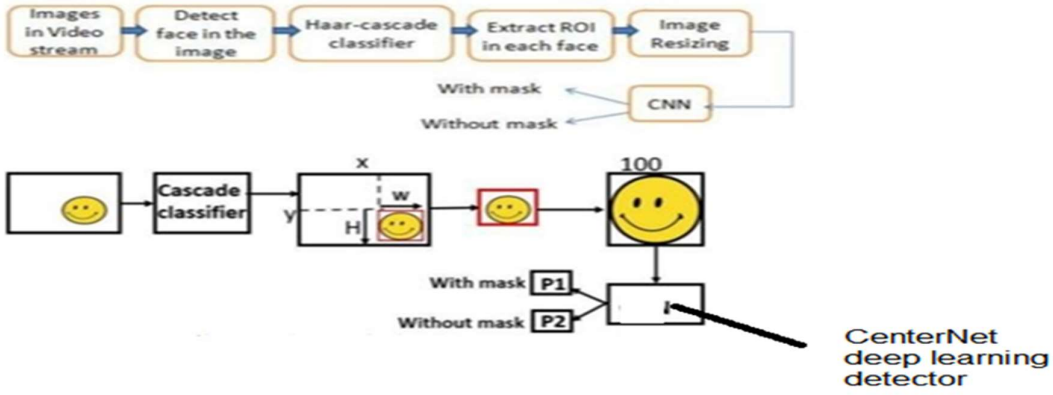


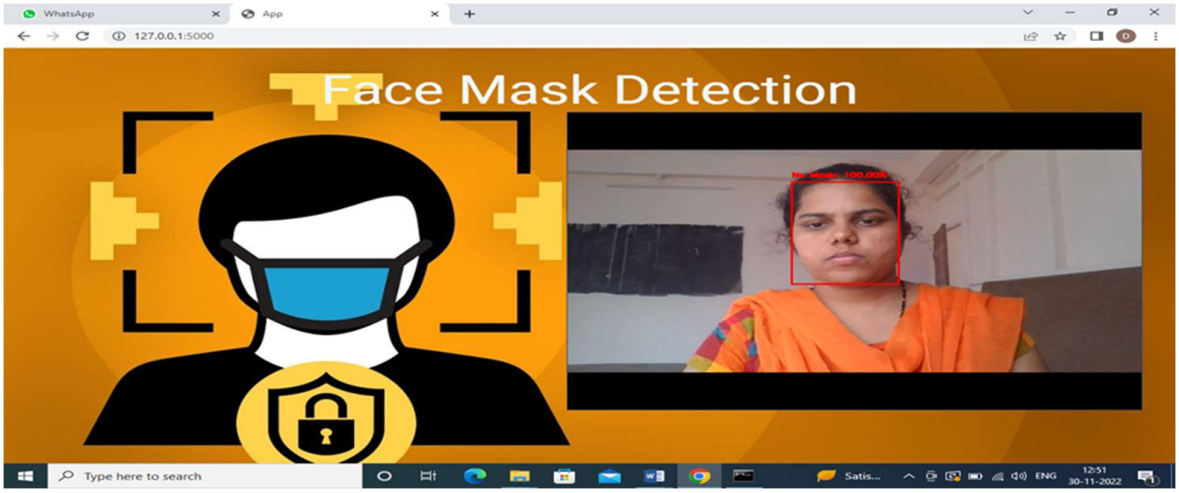
FIGURE 5. WORKFLOW OF OUR RESEARCH METHODOLOGY

VI. EXPERIMENTS AND ANANLYSIS

Center Net Deep Learning Detector in real- time application. there are many cases while wearing face masks one person wear a mask remaining not. some situations all the persons not wearing mask. many cases for wearing of masks. we are performed all case scenarios on mask detection. This part will introduce the trial aftereffects of the facial covering recognition involving CENTER NET Profound LEARNING Finder progressively application. A GUI (Graphical UI) will be made by utilizing Tkinter. The Information planning was first to stack face pictures with and without cover, then, at that point, comment on bouncing box of every vehicle in a picture and save explained pictures and names. Instating consequences be damned organization and producing preparing and testing datasets. We are utilizing a freely accessible dataset to prepare our model. Preparing in Burden pictures and clarified marks, then, at that point, Train just go for it and save prepared model. In the wake of preparing, testing of the datasets can be carried out. Test the Heap setup and weight document of prepared just go for it and Information picture to stacked model.

1.PERSON NOT WEARING MASK:

If a person not wearing mask face mask detection shows red rectangle box along with accuracy. here one person not wearing a mask and it showing red color rectangle box with the accuracy of 100% entitled as No Mask.



2.THE PERSON WEARING MASK:

If single person wears a mask, the result is green rectangle box with accuracy of 99.24% entitled as Mask

3.MASK DETECTION FOR MULTIPLE PEOPLE:

In a group of people CenterNet detects each person face whether the person wear a mask or not. In that group how many members wear a mask then their faces are showed with green rectangle along with accuracy entitled Mask as result. If one person not wear a mask then it showed with red rectangle box with accuracy entitled No Mask as a result.

4.MULTIPLE FACE MASK DETECTION:

In a group of people, the CenterNet detector detects the all the persons wear a mask or not. In our result all persons wear a mask so the output is all the faces are green rectangle box entitled as Mask with accuracy.

a) While running the model through 5000 trials, the results of face mask identification employing YOLO3 with such a database of 1925 photos that includes frames of persons wearing and not wearing masks. It is possible to tell if someone is donning a mask or not.

Average loss = 0.0327

Mean average precision = 99%

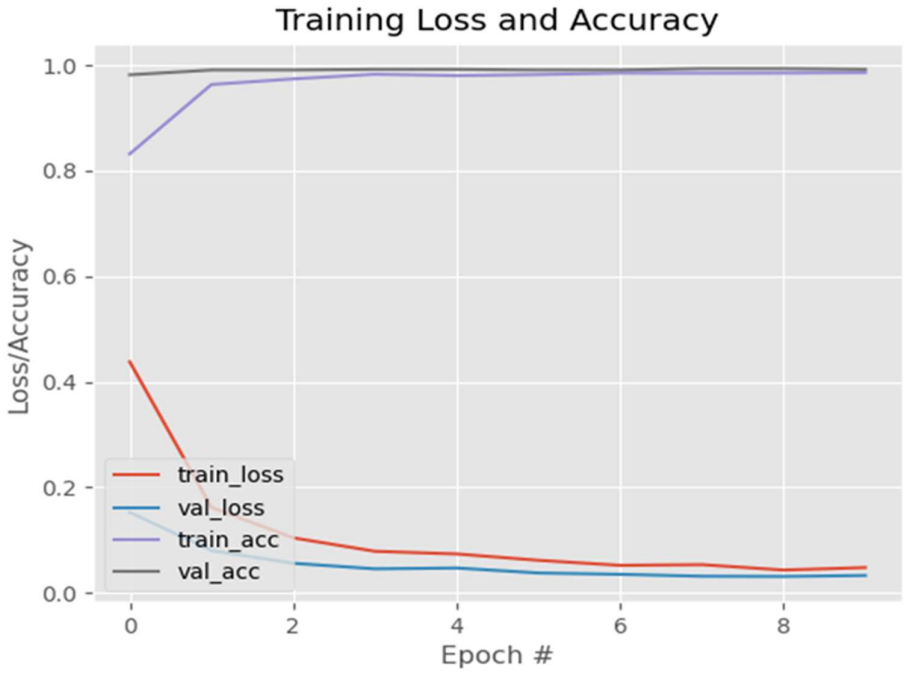


FIGURE 7. Accuracy in Training, Accuracy in Validation, and Losses in Training and Validation

Accuracy and loss features are two critical measures used to assess how well the trained and assessment modes operate. Figure 7 illustrates the loss and accuracy for the pretrained model, whereas Figure 7 illustrates the behaviors for the training of the suggested models on a particular dataset by assessing the training and a validation database for each epoch and plotting the outcomes.

VII. CONCLUSION:

The order framework utilizing profound brain organizations can be viewed as the best way to deal with accomplish high exactness and give improved results than other conventional methodologies as far as precision and misfortune capabilities. During the time of this review, there are a few notes that can be finished up as a result of this work. The proposed two models make progress results to recognize the situation with facial covering by right/mistaken or no veil. Utilizing 2D CNN furnished the chance of managing crude information, and this saved the time and exertion expected for the irritating pre-handling. The ReLU actuation capability coordinated with the CenterNet is utilized to remove remarkable highlights and disregard the feeble elements, which prompts manage commotion related with tests. The ReLU does that by eliminating all the commotion components from the succession and keeping just those conveying a positive worth. Adding the group standardization layers after the convolutional layers to obey the convolutional property, which diminishes the posture variety issue in the example. In the group standardization layer various components of a similar element map, at

various areas, are standardized similarly freely of its immediate extraordinary environmental factors.

REFERENCES

[1] Coronavirus Disease 2019 Weekly Epidemiological May 2021. World Health Organization. https://www.who.int/publications/m/item/weekly-epidemiological-update-on-covid-19---18-may-2021.
[2] Characterization of a novel coronavirus associated with severe acute respiratory syndrome Paul A Rota, M Steven Oberste, Stephan S Monroe, W Allan Nix, Ray Campagnoli, Joseph P Icenogle, Silvia Penaranda, Bettina Bankamp, Kaija Maher, Min-Hsin Chen, Suxiong Tong,
[3] Family Cluster of Middle East Respiratory Syndrome Coronavirus Infections DOI: 10.1056/NEJMoa1303729
[4]Ying Liu, Albert A Gayle, Annelies Wilder-Smith, Joacim Rocklov The reproductive number of COVID-19 is higher compared to sars coronavirus, https://doi.org/10.1093/jtm/taaa021
[5] Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis DOI: 10.1002/jmv.25750
[6] Artificial intelligence applied to chest x-ray images for the automatic detection of COVID DOI:10.1109/ACCESS.2020.3044858
[7] Respiratory virus shedding in exhaled breath and efficacy of face masks Available:https://www.researchgate.net/publication/340412133_Respiratory_virus_shedding_in_exhaled_breath_and_efficacy_of_face_masks.
[8] A two-dimensional sparse matrix profile dense net for COVID-19 diagnosis using chest CT images, Changjian Zhou, Jia Song, Sihan Zhou, Zhiyao Zhang,
doi: 10.1109/ACCESS.2021.3086229
[9] Rational use of face masks in the COVID-19 pandemic DOI: 10.1016/S2213-2600(20)30134-X
[10] Deep learning framework to detect face masks from video footage, DOI:10.1109/CICN49253.2020.9242625
[11] M. Jiang and X. Fan, Retina facemask: A face mask detector Available: http://arxiv.org/abs/2005.03950
[12]A_Deep_Learning_based_Lightweight_Face_Mask_Detector_with_Residual_Context_Attention_and_Gaussian_Heatmap Fight_Against_COVID-19 https://ieeexplore.ieee.org/document/9475521
[13] Rapid object detection using a boosted cascade of simple features, https://www.researchgate.net/publication/3940582_Rapid_Object_Detection_using_a_Boosted_Cascade_of_Simple_Features.
[14] Histograms of oriented gradients for human detection, DOI:10.1109/CVPR.2005.177
[15] Object detection in 20 years: A survey, Available: http://arxiv.org/abs/1905.05055
[16] Joseph Redmon; Santosh Divvala; Ross Girshick. ``You only look once: Unified, real-time object detection," DOI:10.1109/CVPR.2016.91 https://pjreddie.com/darknet/yolo/

[17] SSD: Single shot multibit detector Available: arXiv:1512.02325
[18]The pipeline of the SL-FMDet https://www.researchgate.net/publication/353054502_A_Deep_Learning_based_Light-weight_Face_Mask_Detector_with_Residual_Context_Attention_and_Gaussian_Heatmap_to_Fight_Against_COVID-19
[19] Focal loss for dense object detection https://doi.org/10.48550/arXiv.1708.02002
[20] Rich feature hierarchies for accurate object detection and semantic segmentation arXiv:1311.2524
[21] Faster R-CNN: Towards real-time object detection with region proposal networks arXiv:1506.01497
[22] Xiaokang Ren and Xinxiang Liu. Mask wearing detection based on YOLOv3 DOI 10.1088/1742-6596/1678/1/012089
[23] B. Roy, S. Nandy, D. Ghosh, D. Dutta, P. Biswas, T. Das MOXA: A Deep Learning Based Unmanned Approach for Real-Time Monitoring of People Wearing Medical Masks https://www.researchgate.net/publication/343217408
[24] T. Draughon. ``Implementation of a computer vision framework for tracking and visualizing face mask usage in urban environments" DOI: 10.1109/ISC251055.2020.9239012
[25] T. Y. Lin et al., Microsoft COCO: Common objects in context, in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2014, vol. 8693 LNCS, no. PART 5, pp. 740–755, doi: 10.1007/978-3-319-10602-1_48.
[26] object detection using yolo3 deep learning https://www.mathworks.com/help/vision/ug/object-detection-using-yolo-v3-deep-learning.html
[27] CNN Convolutional neural networks: an overview and application in radiology https://www.mdpi.com/2071-1050/12/22/9490/html.