Face mask detection

Fathima Noohu Nisha M
20MIS1147
VIT, Chennai

Srinivasan H
20MIS1078
VIT, Chennai

Suganeshwaran M
20MIS1093
VIT, Chennai

***Abstract* - Corona virus sickness has become a big public health issue in 2019. Because of its contact-transparent characteristics, it is rapidly spreading. The use of a face mask is among the most efficient methods for preventing the transmission of the Covid-19 virus. Wearing the face mask alone can cut the chance of catching the virus by over 70%. Consequently, World Health Organization (WHO) advised wearing masks in crowded places as precautionary measures. Because of the incorrect use of facial masks, illnesses have spread rapidly in some locations. To solve this challenge, we needed a reliable mask monitoring system. Numerous government entities are attempting to make wearing a face mask mandatory; this process can be facilitated by using face mask detection software based on AI and image processing techniques. For face detection, helmet detection, and mask detection, the approaches mentioned in the article utilize Machine learning, Deep learning, and many other approaches. It will be simple to distinguish between persons having masks and those who are not having masks using all of these ways. The effectiveness of mask detectors must be improved immediately. In this article, we will explain the techniques for face mask detection with a literature review and drawbacks for each technique.**

Keywords: Corona virus disease 2019, Face mask detection, CNN, YOLO, Object Detection.

1. Introduction

Because of the global pandemic of Covid-19, wearing a face mask in public is becoming increasingly popular. Since the people of Covid-19 refuse secure their health by wearing masks against air pollution [7]. Others, on the other hand, are demure about their appearance and conceal their feelings from broader public by moving their own faces. Putting face masks helps to prevent Covid-19 transmission, according to someone. Covid-19 is most modern epidemic virus to poke human well-being in twentieth centenary [15]. The quick spread of Covid-19 has prompted WHO to announce it an international pandemic in 2020. Covid-19 infected more than 5 million patients in 188 countries in less than six months. Close interaction is how the virus is spreading and in densely populated places. The corona virus outbreak has resulted in unprecedented levels of international scientific cooperation [19]. Machine learning and deep learning [1] powered by computer science will aid in fight for Covid-19 in number of different ways. Machine learning analyses massive amounts of data to estimate Covid-19 dispersion, behaves as a short notification system for prospective pandemics, and categorize sensitive populations. Many countries have regulations requiring people to keep face masks in crowd. We have a tendency to produce these regulations and laws as activity to exponential development of occurrences and deaths in numerous domains [5].

As a result, detecting face masks is a difficult process. Because of the expansion of the corona virus sickness, it has gotten a lot of attention lately since numerous countries have adopted policies such as ”No admission without a mask.” Face mask detection is a critical issue in security and the prevention of Covid-19 [4]. In the medical industry, a mask minimizes the danger of infection from an infected individual, whether or not they show symptoms. Face mask detection is employed in a variety of settings, including airports, hospitals, offices, and educational institutions. Face recognition with no a mask is simpler, but faces recognition with just a mask is more difficult since masked face feature extraction is more difficult than conventional face feature extraction. Many facial characteristics, such as the nose, lips, and chin, are missing from the covered face. The figure 1 shows the architecture model for Face Mask Detection. There are eight steps are used for the detection of face with mask.

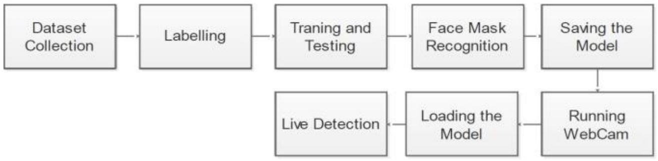

Figure 1: Model Architecture for Face Mask Detection

2. *Techniques for Face Mask Detection*

Convolutional Neural Networks (CNNs) are a type of deep neural network motivated by bio- logical phenomena. A CNN is composed of several components, including one with convolutional layer, pooling layer, as well as then fully connected layer, and it learns the spatial patterns of data autonomously and fluidly using the backpropagation method. The CNN kernels are common across entire image positions, making it incredibly parameter-efficient. The CNN is a strong option for computer vision problems because of these properties. Because of major advancements in GPU com- puter capability, deep learning technologies have blossomed in recent years [2]. Throughout computer vision, object recognition seems to be a critical task that has attracted a lot of attention. According to the recommended recommendations and tactics for improvement, figure 2 shows us the current object detection methods:

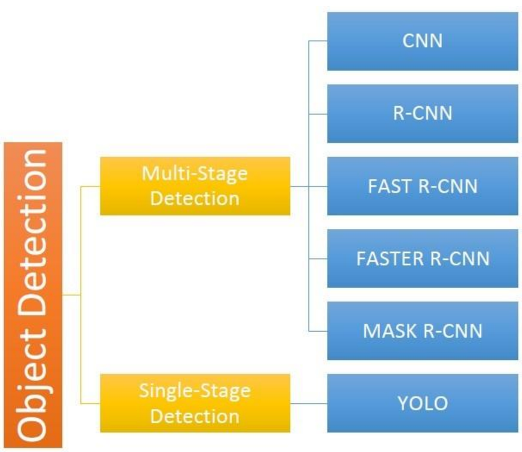

Figure 2: Object Detection Methods

2.1. *Multi-Stage Detectors*

The two-stage method uses a heuristic algorithm as CNN to generate a huge number of region recommendations for every image, next classifies and stagnates these eligible provinces. As first object detection technique, deep learning was applied.The figure 3 shows the model architecture of retina face detection. There are many stages in this detection for the detection of masked face is shown in the figure.

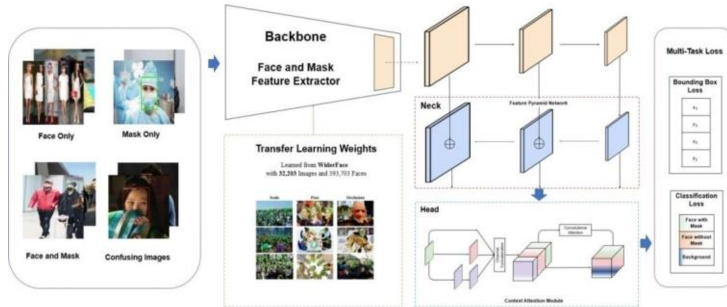

Figure 3: Architecture of Retina Facemask [20]

2.1.1. *Convolutional Neural Network (CNN)*

Because of its low computational cost and ability to extract spatial information, CNN plays ancrucial role in computer vision tasks like pattern recognition. In order to eliminate top-level features, CNN uses convolution parts to mix with primary images.

S.Shivaprasad, 2021 [24] Deep learning, OpenCV, TenserFlow, as well as Keras are used in the study strategy to aid in the faces detection having masks. With assistance of this technique, safety is ensured. The technique for face detection has employed the MobileNetV2 and CNN framework as classifier; it is lightweight, has fewer parameters, whichmay be utilized in embedded devices (Onion Omega2 and Raspberry Pi) to perform authentic mask identification. The accuracy of the approach utilised in this study is 0.96, and the F1-score is 0.92. The data was gathered from a variety of sources and several scientists can use it to create more sophisticated models, like face recognition, facial patterns, and facial characteristics for detection method.

Prathmesh Deval, 2021 [6], contemplate developing a detection system for face masks connected with digital healthcare services. By used OpenCV, toget access to the live video stream and also for image pre-processing. For facedetection, Haar-Cascade will be used, as it is a very effective face detection method. Figure 4 depicts the system’s design, which demonstrates how it works automatically to avoid the expansion of COVID19. The research uses advanced learning algorithms to recognise various facial features and determine whether or not person is employing the face mask. The system deals with Face Mask detection in real-time and also helps in reducing the transmission rate. The system also provides few digital facilities for receptionists and doctors.

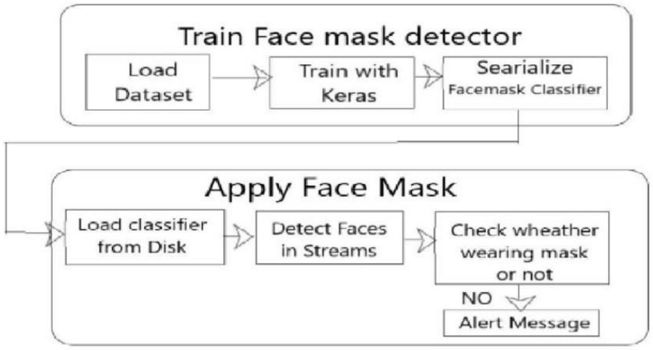

Figure 4: The system architecture [6]

Jansi Rani Sella Veluswami 2021 [23], proposed model was trained on a database of over 11,000 images of faces either with or without masks, employing numerous deep learning techniques. A SSDNET model is being deployed for face detection, the output of which is passed to a custom- made Lightweight CNN for mask detection. On two distinct testing datasets, the model obtains a remarkable accuracy of 96% and 96%. The model will help government agencies and health officials fight the global pandemic.
The disadvantages of CNN include the need for huge datasets (i.e., hundreds as well as thousands of images) and their accurate annotation, which can be a sensitive technique that requires domain specialists.

2.1.2. Region-based Convolutional Neural Network (R-CNN)

The very first CNN-focused two-stage object detection algorithm was Region focused convolu- tional neural network (R-CNN) [17]. R-design CNN's is comprised of three distinct blocks, as seen in Figure 5. The authors utilized a selective search technique to create about 2000.0 class-independent regional ideas by every source image during the first step. To retrieve feature vectors of specific length from every zone suggestion, they utilized CNN having five convolutional layers (Conv) as well as double fully connected layers (FC) inside the second block. Following that, each region proposal is fed into a separate CNN to retrieve specific length feature vectors. The final block uses a linear support (SVM) vector machine for a given category to classify each region proposition.
Chenchen Zhu 2017 [31], suggested Contextually Multi-Scaled Region-focused Convolutional Neu- ral Networks (CMS-RCNNs) to address various hard issues such as significant facial occlusions, in- credibly lower resolutions, intense illumination, highly pose changes, video or image compression errors. Proposed networks, like region-based CNNs, have a component of regional proposal and a

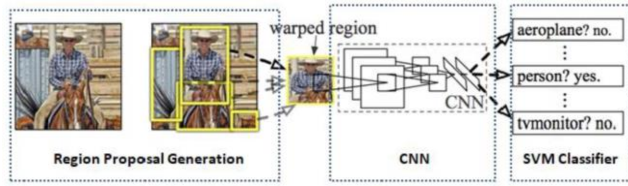

Figure 5: Architecture of R-CNN [8]

region-of-interest (RoI) detecting components. However, outside of those networks, two major shares for improving state-of-the-art facial recognition performance.
The multi-scaled information is aggregated including both feature map and RoI traced to concil- iation with narrower face region. Secondly, our suggested network, which is based on intuition of the human vision system, permits explicit body spatial reasoning. By combining WIDER FACE Datasets with Face Detection Datasets and Benchmarking, and the Face Detection Dataset as well as Bench- mark, we were able to create a dataset with a high degree of variability (FDDB). The investigation results reveal that the suggested methodology, when trained on the WIDER FACE Dataset, consis- tently Upon that WIDER FACE Data, it surpasses strong benchmarks and consistently produces competitive results on FDDB when compared to latest state face detection approaches.
One of R-shortcomings CNN's is that each image must classify 2000 region recommendations. As a result, training the network requires a long duration, it takes 49 seconds to recognise objects in image, and it consumes a lot of disc space for storing extracted features of region proposal.

2.1.3. Fast Region-based Convolutional Neural Network (FAST R-CNN)

Fast R-CNN is a one-phase training method that learns to categorise region ideas while also correcting their spatial positions. Fast R-CNN could train an extremely deep detection network in 9 times the time it takes R-CNN. The full image and a collection of object ideas are sent into the fast R-CNN. The mechanism of FAST R-CNN is depicted in Figure 6.

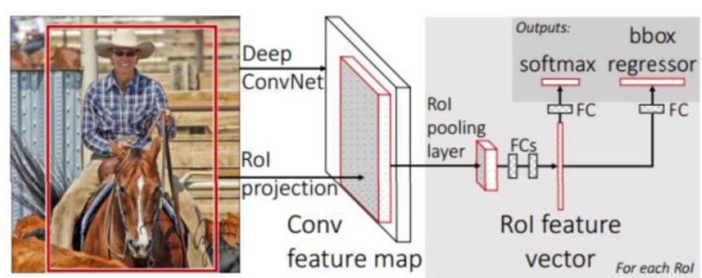

Figure 6: Mechanism FAST R-CNN [8]

K. Wang, 2016 [26],proposed deep cascade convolutional network that use Fast-R-CNN, in the research they used Detection Date Set and Benchmark (DDSB) and Annotated Face in-the-Wild (AFW)as a dataset for testing part.for image input, the first stage starts by Lower Stage Classifi- cation Networks (L-Cls-Net) which scan the entire image in various scales with the 95% reject of detection windows, resize image and put it in Lower Stage Calibration Networks (L-Cal-Net) form size adjusttion and nearby faces. For high abandon place those into Higher Stage Classification Net- works (H-Cls-Net), afterwards adjust by Higher stage Calibration Networks (H-Cal-Net) for spatial locations uses Fast-R-CNN Networks (FR-Net). Rongqiang Qian, 2016 [18], suggested an existing road proposed detection system using CNN, detection system comprises of two phases: first, there are a huge amount of candidate zones that can comprise intended objects and all these objects shall be passed to Fast R-CNN having the process steps shown in figure 7.

Figure 7: Fast R-CNN process steps

The dataset was recorded by set up camera on the vehicle having a pre-determined angle of shooting. The recordings were donewith 1280 720 pixels resolution. We decoded movies and gathered all of frames with traffic indications. There are entire numbers of 2223 images taken, with 4204 road signs incorporated. The results of experiments shows that the Fast R-CNN either having or without having regression with bounding boxes achieves the maximum recall rates of 90.73% as well as 88.33%, respectfully, whereas the comparable rates of accuracy are 14.49% and 71.23%. When compared to the region proposal stage recall rate of 89.74%, Fast R-CNN using regression of bounding boximproves by 0.99%, while Fast R-CNN with no regression losesof bounding box 1.41% positive sectors. Table 1 shows the Average Precisions.

Table 1: Detection Results [18]

Method	AP
Fast-RCNN plus bbox	85.580%
Fast-RCNN with no bbox	83.990%

Qihang Wang 2018 [27], suggested face identification algorithm relying on Fast R-CNN that uses three algorithms to discover the candidate area of such face that might exists in image (CNN strategy, Haar-Adaboost method, and candidate search heuristics). The candidate zone is fed into trained convolution neural network, which produces a final convolution attribute (ROI) depending on Fast R-CNN network architecture after a sequence of convolution and pooling procedures. The ROI is then fed into the two full interconnections (ROI). Table 2 shows the effectiveness of various techniques.

Table 2: Performance algorithms [27]

	Detection rate(%)	False rate(%)	Missed rate(%)	Detection time(s)
Traditional CNN	91.260	15.500	8.760	1.548
PCA+SVM	93.400	13.650	7.670	0.980
Adaboost	96.100	12.340	7.370	0.810
Proposed algorithm	98.200	8.690	1.200	0.240

Lin Jiang 2021 [11],To recognize human faces autonomously, it uses sophisticated deep learning algorithm relying on machine vision. A multiscale rapid RCNN procedure depending on upper or lower layers (UPL-RCNN) is suggested to purelyrecognize a range of human faces. The +e network is made up of components that perform spatial affine transformations and feature region constituents (ROI). Face detection relies heavily on this technology. To begin with, multiscale data can be bundled in detection to handle small portion of face [16]. The approach may then execute contextual sophismand spatial transformations, such as zooming, trimming, and rotating, using the inspirations of the visual perception. The results of comparative studies reveal that this technology can not only accurately recognize different faces but also outperforms fast R-CNN in terms of performance.
The drawbacks with FAST R-CNN most time taking during detection is done by the selective search region.

2.1.4. Faster Region-based Convolutional Neural Network (FASTER R-CNN)

Ren et. al. suggested Faster-R-CNN [21], yet They've switched to a region proposal network as a replacement for the proposal approach (RPN) [8]. The RPN is just a fully convolutional (FCN) net- work that gets every size image as an input as well as outputs a bunch of rectangular aspirant object recommendations. Every object concept has an objectness scoring rate that determines whether or not the proposal comprises an object. Figure 8 depicts the structure of FASTER-R-CNN.

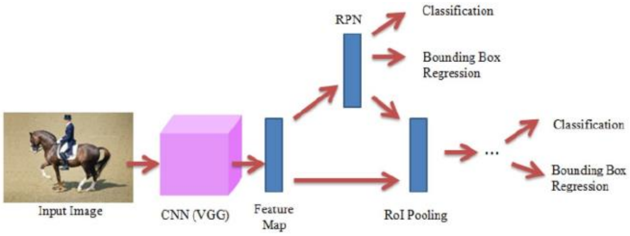

Figure 8: Structure of FASTER R-CNN [21]

Huaizu Jiang 2017 [12],the Faster R-CNN has beensuggested. It is comprised of two modules. Regional Proposal Networks (RPN), first is a fully convolutional networks that generates item sug- gestions for the second module. The Fast-R-CNN detector seems to be second one, and its aim is to refine ideas. They employed the WIDER face database, which comprises of 12,880 images with 159,424 faces. On three benchmark datasets, the outcome indicates cutting-edge face identification performance, and through RPN, several convolutional layers can be used without increasing the computational overhead.
Wenqi Wu 2018 [28], proposed a face detector with variable scales (DSFD) relying on Faster R- CNN. The novel network can enhance facial detection accuracy while also acting as a real-time Faster R-CNN. To acquire the human face ROI, first an effective multitask region proposal network (RPN) is created, which is paired with enhancing face detection. The anchor is used with facial landmarks to derive a human face concept. Then, depending on the proposal scale, a Fast R-CNN of parallel-type network is presented. By FDDB dataset, 500 images generated randomly, the tested experiments in Table 3 contrast DSFD, baseline technique actual Faster R-CNN, MXnet, and UnitBox. The suggested method obtains a 96.69% recall rate with 130.0ms for processing aimage frame, whereas actual Faster R-CNN as well as MXnetutilise 140.0ms and 230.0ms, respectively, underneath the 700 false positives discrete score.

Table 3: Performance algorithms [27]

Methods	Recall rate (%)	Runtime (ms/frame)
Proposed method	96.69	130
Faster R-CNN	96.05	140
MXnet	96.10	230
UnitBox	94.61	110

Mosab Rezaei 2019 [22], investigated on the robustness of two networks Single Shot-Multi-Box- Detector SSD and Faster-R-CNN in impulse noise, Gaussian blur, and JPEG2000 compression are all used in this image. The tests are carried out using Wider Face database. The technique in the first investigation is standard; that is, the network do trained on actual training samples with no any extra distortion or compressing, and effectiveness is yet assessed on standard test samples Table 4. These tests reveal that Faster-R-CNN is to be more resistant to Gaussian blur, whilst the SSD is significantly more perceptive to edges. SSD, on the other hand, is more resistant to JPEG compressed photos of lower quality. The explanation for this should be related to Faster R-sensitivity CNN's to texturing of objects.

Table 4: MAP of FASTER R-CNN and SSD [22]

Network	Hard	Easy	Medium
SDD	0.63	0.58	0.43
Faster-R-CNN	0.82	0.81	0.56

One deficit of Faster R-CNN [3] would be that RPN is trained utilizing single image to retrieve all anchors inside mini-batch with size 256. Even though all samples by a single image might well be associated (i.e. characteristics seems to be identical), network's convergence could take a long time.

2.1.5. Mask Region-based Convolutional Neural Network (MASK FAST R-CNN)

Kaiming He [9],For generating region proposals, Mask R-CNN uses exactly regional proposal networks (RPN) by Faster-R-CNN. Instead of using a RoI pooling layer, the researchers used a RoI Align layer on region suggestions to align retrieved features with object's input position. After that, the aligned RoIs are input into the Mask R-final CNN's stage, which produces three outputs: a bounding box offset, a class label, and binary object mask. Each RoI is masked using a tiny fully convolutional neural network. The structure of Mask-R-CNN is shown in Figure 9.

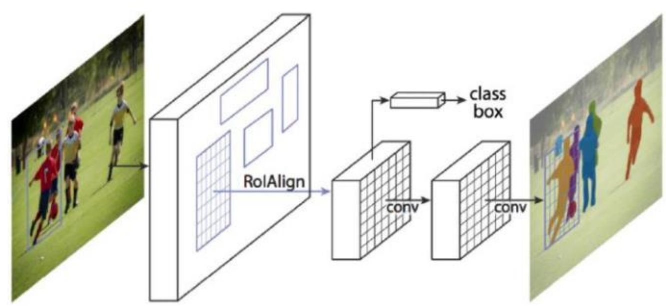

Figure 9: MASK R-CNN [9]

Kaihan Lin [14], proposed G-Mask, a face identification and segmentation approach focused on Mask R-CNN that uses generalised intersection across union. RoI Align is used in this procedure to keep the appropriate spatial

locations. To fragment the face imageby either background image, RoIs are generated using the ResNet-101 network and even the RPN network, and appropriate binary face mask is constructed using Fully Convolution Network. For dealing with multi-scale face problems, GIoU was created as loss function depending on basic Mask R-CNN model. To train the model, a novel face dataset including segmentation annotation info is created. The results of the experiments reveal that our proposed technique performs well on the well-known FDDB and AFW standards.
Kaihan Lin [13],For face detection and segmentation, the G-Mask approach was suggested. ResNet-101 can retrieve features, RPN can produce ROIs, RoIAlign can keep precise spatial lo- cation, and the full convolutional network (FCN) can produce binary masks. As a result, the sug- gested scheme is capable of effectively recognising and segmenting each individual facial image. Even though to its comparatively simple structure, Faster-R-CNN has the lowest running time, whereas the suggested approach has aelapsed time equivalent to Mask-R-CNN.

Table 5: Elapsed time of various region-focused methods [13]

Mechanism	FDDB Time running (s)	AFW	ChokePoint
R-CNN	14.750	15.320	14.510
FastR-CNN	3.120	3.080	2.840
FasterR-CNN	0.300	0.320	0.280
MaskR-CNN	0.320	0.350	0.330
G-Mask	0.350	0.420	0.330

The drawbacks of MASK R-CNN are It only works with still images, thus it can’t look at tempo- rary information about object of interest, also including energetic hand gestures, and it often misses objects that are motion blurred at low resolution, like hands.

2.2. Single Stage Detection

Classification as well as regression are completed in single shot employing consistent and dense sampling having regard to locations, sizes, and aspect ratios in the one-stage technique:
Zhi Tian 2019 [25], FCOS is a one-stage detector that is both anchor-free as well as proposal- free. In tests, FCOS outperforms common anchor-depending one-stage detectors such as RetinaNet, YOLO, and SSD, yet with far less design complication. FCOS entirely skips all anchor box calculation and hyper-parameters, solving object detection in a per-pixel predictions manner, identical to those other dense prediction problems like semantic segmentation. Between one detectors, FCOS also reaches state-of-the-art effectiveness. We also demonstrate that FCOS can be employed as RPNs in the two-stage Faster R-CNN detector and surpasses its RPNs by a significant margin.
Hongtao, W. 2020 [10], optimized algorithm to improve efficiency of object detection as the accuracy of the one-stage detector often lags behind of the two -stage detector, the model was trained on VOC 2007 and 2012 trainval with total of 16.551 pictures, for enhancing part the data flipped left and right and random sampling is used. The results show that the one-stage detector gets high accuracy on SSD. By comparing the results, it can draw that first the improved algorithm for objects with similar categories has higher detection accuracy and reduce false detection, second improved algorithm has higher detection confidence for the same object detected.
Guanhao Yang 2020 [30], To effectively apply YOLOV5 figure 10 revealed that the network structure illustration, the most effective objection detection technique at the moment, in the real world, particularly in the guidance of donning masks in crowded locations, it was suggested to substitute manual inspection with only a deep learning approach and utilize YOLOV5 figure 10 exhibited the network structure illustration, the most influential objection detection technique at the moment. Figure 11 depicts the system. When visitors enter the mall, they will capture images with the camera, which will then be transferred to interactfor face mask identification. The mall gate would be opened and exhibited to pass if face detected within two seconds is indeed a face have a mask; else, it will be reverted to face mask identification until success is achieved. The experimental findings suggest that the suggested algorithm can efficiently recognize face masks and enable staff surveillance.

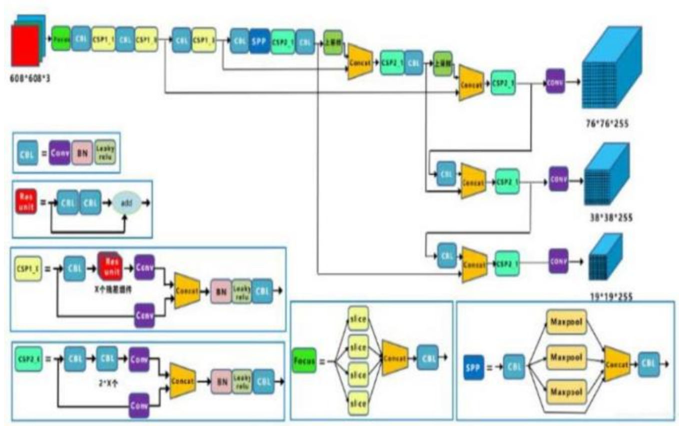

Figure 10: YOLO V5 Network [30]

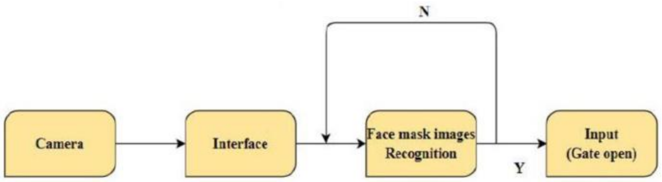

Figure 11: Proposed system [30]

Xing, C. 2020 [29], proposed a modified network based on YOLOv3-Tiny for water conservancy facility inspection which introduced residual network to merge deep features with lower features into the original YOLOv3-Tiny network. All images (3576) are taken by camera mounted on flying drones during daytime and captured above 80 meters, most images are token at or nearby the sites of water conservancy facilities. The outcomes of the experiments reveal that precision and speed are high in the proposed algorithm see table 6.

Table 6: Comparison of Speed and Accuracy [29]

Network	mAP(%)	FPS
SSD	63.8	38.7
YOLOv3	62.2	41.6
YOLOv3-Tiny	50.3	53.8
Proposed	53.2	48.5

When contrasted to Faster R-CNN, YOLO has weaker recall and greater error in localization, fails to detect close items because every grid could only suggest two bounding boxes, and suffers to identify small objects.

3. *Conclusion*

In this paper, we have discussed some research papers about facial mask detection. As we know nowadays mask detection is a very challenging task. The applications of Facial Mask Detection are used especially for the prevention of spreading Corona Virus, tracking and identifying criminals and anti-spoofing etc. Each of these papers uses a different kind of algorithms, different techniques, different approaches (see Table 7) but their goal is the same to detect a face, facial features like eyes, nose, eyebrows and to find out whether the face of a person is covered with a mask or not. After doing a deep study of all the algorithms we have concluded that each of these techniques have their own pros and cons but as compared to the other algorithms YOLO algorithm give better results with more accuracy and are more successful in real life.

References

[1] A.S. Abd-Alzhra and M.S.H. Al-Tamimi, Lossy image compression using hybrid deep learning autoencoder based on k-mean clustering, Design Engin. (2021) 7848–7861.

[2] M.S.H. Al-Tamimi, Combining convolutional neural networks and slantlet transform for an effective image re- trieval scheme, Int. J. Electr. Comput. Eng. 9(5) (2019) 4382–4395.

[3] M.S.H. Al-Tamimi, I.A. Abdulmunem and S.S. Sulaiman, Improved merging multi convolutional neural networks framework of image indexing and retrieval, Int. J. Adv. Sci. Tech. 29(8) (2020) 1884–1901.

[4] M. Dasgupta, O. Bandyopadhyay and S. Chatterji, Automated helmet detection for multiple motorcycle riders using CNN, IEEE Conf. Inf. Commun. Technol. Allahabad, India, 2019, p. 1–4.

[5] G. Deore, R. Bodhula, V. Udpikar and V. More, Study of masked face detection approach in video analytics, 2016 Conference on Advances in Signal Processing (CASP), Pune, (2016) 196–200.

[6] P. Deval, A. Chaudhari, R. Wagh, A. Auti and M. Parma, CNN based face mask detection integrated with digital hospital facilities, Int. J. Adv. Res. Sci., Commun. Tech. 4(2) (2021) 492–497.

[7] M.S. Ejaz, M.R. Islam, M. Sifatullah and A. Sarker, Implementation of principle component analysis on masked and non-masked face recognition, IEEE 1st Int. Conf. Adv. Sci. Engin. Robotics Technol. 2019, pp. 1–5.

[8] R. Girshick, Fast R-CNN, Proc. IEEE Int. Conf. Comput. Vision 2015, pp. 1440–1448.

[9] K. He, G. Gkioxari, P. Doll´ar and R. Girshick, Mask R-CNN, Proc. IEEE Int. Conf. Comput. Vision 2017, pp. 2961–2969.

[10] W. Hongtao and Y. Xi, Object detection method based on improved one-stage detector, 5th Int. Conf. Smart Grid Electric. Autom. (ICSGEA), IEEE, 2020, pp. 209–212.

[11] L. Jiang, J. Chen, H. Todo, Z. Tang, S. Liu and Y. Li, Application of a fast RCNN based on upper and lower layers in face recognition, Comput. Intell. Neurosci. 2021 (2021).

[12] H. Jiang and E. Learned-Miller, Face detection with the faster R-CNN, 12th IEEE Int. Conf. Automatic Face Gesture Recogn. (FG 2017), IEEE, 2017, pp. 650–657.

[13] K. Lin, H. Zhao, J. Lv, C. Li, X. Liu, R. Chen and R. Zhao, Face detection and segmentation based on improved mask R-CNN, Discrete Dyn. Nature Soc. 2020 (2020).

[14] K. Lin, H. Zhao, J. Lv, J. Zhan, X. Liu, R. Chen, C. Li and Z. Huang, Face detection and segmentation with generalized intersection over union based on mask R-

CNN, International Conference on Brain Inspired Cognitive Systems, Springer, Cham. (2019) 106–116.
[15] T. Meenpal, A. Balakrishnan and A. Verma, Facial mask detection using semantic segmentation, IEEE 4th International Conference on Computing, Communication and Security, (2019) 1–5.
[16] N.A. Mohamed and M.S.H. Al-Tamimi, Image fusion techniques: A review, Int. J. Psych.Rehabil. 24(10) (2020).
[17] N.A. Mohamed and M.S.H. Al-Tamimi, Image fusion using a convolutional neural network, Solid State Technol. 63(6) (2020).
[18] R. Qian, Q. Liu, Y. Yue, F. Coenen and B. Zhang, Road surface traffic sign detection with hybrid region proposal and fast R-CNN, 12th Int. Conf. Natural Comput. Fuzzy Syst. Knowledge Discov. IEEE, 2016, pp. 555–559.
[19] B. Qin and D. Li, Identifying facemask-wearing condition using image super-resolution with classification network to prevent COVID-19, Sensors 20(18) (2020) 5236.
[20] J. Redmon, S. Divvala, R.B. Girshick and A. Farhadi, You only look once: Unified, real-time object detection, In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, (2016) 779–788.

[21] S. Ren, K. He, R. Girshick, X. Zhang and J. Sun, Object detection networks on convolutional feature maps, IEEE Trans. Pattern Anal. Machine Intell., 39(7) (2016) 1476–1481.
[22] M. Rezaei, E. Ravanbakhsh, E. Namjoo and M. Haghighat, Assessing the effect of image quality on SSD and faster R-CNN networks for face detection, 27th Iran. Conf. Electric. Engin. IEEE, 2019, pp. 1589–1594.
[23] J.R. Sella Veluswami, S. Prakash and N. Parekh, Face mask detection using SSDNET and lightweight custom CNN, Proc. Int. Conf. IoT Based Control Networks and Intelligent Systems-ICICNIS 2021, (2021).
[24] S. Shivaprasad, M.D. Sai, U. Vignasahithi, G. Keerthi, S. Rishi and P. Jayanth, Real time CNN based detection of face mask using mobilenetv2 to prevent Covid-19, Ann. Roman. Soc. Cell Bio. 25(6) (2021) 12958–12969.
[25] Z. Tian, C. Shen, H. Chen and T. He, Fcos: Fully convolutional one-stage object detection, Proc. IEEE/CVF Int. Conf. Comput. Vision 2019, pp. 9627–9636.
[26] K. Wang, Y. Dong, H. Bai, Y. Zhao and K. Hu, Use fast R-CNN and cascade structure for face detection, Visual Commun. Image Process. (VCIP), IEEE, (2016) 1–4.
[27] Q. Wang and J. Zheng, Research on face detection based on fast R-CNN, Recent Developments in Intelligent Computing, Communication and Devices, Springer, Singapore, (2019) 79–85.
[28] W. Wu, Y. Yin, X. Wang and D. Xu, Face detection with different scales based on faster R-CNN, IEEE Trans. Cyber. 49(11) (2018) 4017–4028.
[29] C. Xing, X. Liang and R. Yang, Compact one-stage object detection network, IEEE 8th Int. Conf. Comput. Sci. Network Technol. (ICCSNT), IEEE, (2020) 115–118.
[30] G. Yang, W. Feng, J. Jin, Q. Lei, X. Li, G. Gui and W. Wang, Face mask recognition system with YOLOV5 based on image recognition, IEEE 6th Int. Conf. Comput. Commun. (ICCC), IEEE, 2020, pp. 1398–1404.
[31] C. Zhu, Y. Zheng, K. Luu M. Savvides, CMS-R-CNN: contextual multi-scale region-based CNN for unconstrained face detection, Deep Learn. Biometrics, Springer, Cham. (2017) 57–79.