Face Mask Detection Using Machine Learning

Mrunali Mahesh Wakarekar

Student

K. I. T's colleagues of engineering kolhapur

mwakareker@gmail.com

Uma Prashant Gurav

Assistant Professor

K. I. T's college of Engineering, Kolhapur

gurav.uma@kitcoek.in

Abstract— Face mask recognition has been growing quickly once corona insistent last years for its multiple uses within the areas of enforcement Security functions and alternative industrial uses Face seems spreading others to corona a unique approach to perform face printing operation detection and mask recognition is projected. The projected system to classify mask detection victimization COVID-19 precaution each in pictures and videos victimization convolution neural network. in depth experimentation on the datasets and therefore the performance analysis of the projected strategies square measure exhibited. Further, we have a tendency to created a self-made plan to preserve repose and intra category variations of mask detection victimization symbolic approach. We have a tendency to study the various classifiers like Support Vector Machine and a Symbolic Classifier. The project is developed as a example to watch temperature activity and to find mask for the folks. the primary technique is performed victimization temperature detector wont to find this temperature of the body and mechanically spray the sanitizer. Within the second technique, the work is intended to produce a security system for the folks so as to avoid COVID-19. we have a tendency to projected continuous observance of the folks conditions and store the people's knowledge within the server victimization the Deep learning construct. so as to research the performance the projected technique an intensive experimentation is conducted on fifty numerous Image dataset. we have a tendency to conducted experimentation beneath variable of coaching and testing proportion for ten random trails. From the results we have a tendency to may observe that, the results obtained for symbolic approach is healthier than the traditional approach.

Keywords—: Deep Learning, Corona Virus, Covid- 19

I. Introduction

Corona virus sickness 2019 (COVID-19) unexpectedly stone-broke call at 2019 and has seriously affected the total world. As of twenty six March 2021, COVID-19 has infected over a hundred twenty five million individuals

Worldwide and caused over two.7 million deaths. one in every of the transmission routes of COVID-19 is thru droplets of spit or nasal secretions once associate infected person coughs or sneezes, that is very infectious and will be worse in, jammed places. Since there's no specific treatment for COVID-19, infections should be restricted through hindrance strategies. Studies have shown that carrying masks will scale back the chance of corona virus transmission, which implies carrying masks is presently one in every of the effective hindrance strategies. per the globe Health Organization (WHO), the correct thanks to wear a mask is by adjusting the mask to hide the mouth, nose, and chin. The protection is greatly reduced if masks don't seem to be worn properly. At present, security guards area unit organized publicly places to cue individuals to wear masks. However, this live not solely exposes the guards to the air that will contain the virus, however conjointly ends up in overcrowding at the entrances because of its unskilfulness. Therefore, a quick and effective methodology is required to deal with matters.

Computer vision is an interdisciplinary scientific field that involves how computers Gain advanced understanding from digital images or videos. Traditional computer Vision tasks include image processing, image classification, object detection, and image Recognition. Object detection can detect instances of visual objects of a certain class in the Images, which are a proper solution for the problem mentioned above. Consequently, Mask detection has become a vital computer vision task to help the global society.

Although in most cases CNNs are utilized in the diagnosing of COVID-19, they will even be utilized in alternative applications, as a part of contagion bar measures [44]. In [45], a system is given that enables folks to be monitored once getting into and being within a definite place, and to gauge if they're obliging with the established insecurity measures. Within the event that this can be not complied with, others may be aware to exercise caution and health personnel to use the several measures. they need additionally been accustomed develop detection systems for the right use of face masks. For this reason, in [46], a system is planned that differentiates the people that use a mask or not with the algorithms RCNN, Fast RCNN, Associate in Nursing quicker RCNN with an accuracy of ninety three.4%. In , the VGG-16 CNN model is employed to implement a detection system with Associate in Nursing accuracy rate of ninety six. Similarly, in they propose the SSDMNV2 model supported the MobileNetV2 design, that has Associate in Nursing accuracy of ninety two.64% once playing the experimental tests. On the opposite hand, [49] describes a system for the detection of face masks employing a support vector machine (SVM) formula. The datasets ar the Real-World covert Face Dataset (RMFD), the Simulated covert Face Dataset (SMFD), and therefore the tagged Faces within the Wild (LFW). The results show Associate in Nursing accuracy of ninety nine.64% with SVM in RMFD, 99.49% in SMFD, and 100% in LFW. In, InceptionV3 transfer learning is employed, getting Associate in Nursing accuracy of ninety nine.92% throughout coaching and 100% throughout tests with SMFD information. In , a way to spot the right use of masks is outlined by combining classification networks and super-resolution of pictures (SRCNet).

An accuracy of ninety eight.70% is achieved, surpassing standard image classification ways of this kind. the matter of biometric identification thanks to the employment of face masks throughout the COVID-19 pandemic has caused new horizons to be explored in computing, representing a challenge for researchers, that has motivated the event of ocular recognition systems, as a parallel response. In [52], a biometric identification system mistreatment eye data and CNN trained by Image Net is given. The results gift Associate in Nursing accuracy of between 90–95%. Similarly, [53] provides a biometric identification system mistreatment SVM with 3 databases (UBIPr, Color FERET, and Ethnic Ocular). Performance tests show a yield of roughly ninety two.

II. LITERATURE REVIEW

Pattern learning and visual perception area unit the inherent tasks that a laptop vision (CV) technique should affect. Visual perception encompasses each image classification and object detection. The task of recognizing the mask over the face within the bone space are often achieved by deploying Associate in Nursing economical visual perception algorithmic rule through police investigation devices. The article recognition pipeline consists of generating the region proposals followed by category ification of every proposal into connected class . We tend to review the recent development in region proposal techniques exploitation single-stage and two-stage detectors, general technique for up detection of region proposals and pre-trained models supported these techniques.

Single-stage detectors
The single-stage detectors treat the detection of region proposals as a simple regression disadvantage by taking the input image and learning the class possibilities and bounding box coordinates. Over effort and DeepMultiBox [9] were early examples. YOLO (You entirely Look Once) popularized single-stage approach by demonstrating amount predictions and achieving outstanding detection speed but suffered from low localization accuracy as compared with two-stage detectors; notably once small objects area unit taken into thought [10]. Basically, the YOLO network divides an image into a grid of size GxG, and each grid generates N predictions for bounding boxes. each bounding box is taboo to have just one class throughout the prediction, that restricts the network from finding smaller objects. Further, YOLO network was improved to YOLOv2 that embedded batch standardization, high-resolution classifier and anchor boxes. Moreover, the event of YOLOv3 is built upon YOLOv2 with the addition of degree improved backbone classifier, multi-sale prediction and a replacement network for feature extraction. Although, YOLOv3 is dead faster than Single-Shot Detector (SSD) but does not perform well in terms of classification accuracy.

Moreover, YOLOv3 desires Associate in Nursing large amount of procedure power for logical thinking, making it not applicable for embedded or mobile devices. Next, SSD networks have superior performance than YOLO thanks to very little convolution filters, multiple feature maps and prediction in multiple scales. The key distinction between the architectures is that YOLO utilizes 2 whole connected layers, whereas the SSD network uses convolution layers of variable sizes. Besides, the RetinaNet planned by sculptor is in addition a single-stage object discoverer that uses featured image pyramid and focal loss to sight the dense objects among the image across multiple layers and achieves outstanding accuracy conjointly as speed love two-stage detectors.

Two-stage detectors
In distinction to single-stage detectors, two-stage detectors follow a lengthy line of reasoning in laptop vision for the prediction and classification of region proposals. They initial predict proposals during a image then apply a classifier to those regions to classify potential detection. Varied two-stage region proposal models are projected in past by researchers. Region-based convolution neural network in addition abbreviated as R-CNN depicted in 2014 by Ross Airsick et al. it's planning to ar one in every of the first large-scale applications of CNN to the matter of object localization and recognition. The model was successfully incontestable on benchmark datasets like VOC-2012 and ILSVRC-2013 and created state of art results. Basically, R-CNN applies a selective search formula to extract a set of object proposals at associate initial stage and applies SVM (Support Vector Machine) classifier for predicting objects and connected classes at later stage. spatial pyramid pooling SPPNet (modifies R-CNN with associate SPP layer) collects choices from varied region proposals and fed into a totally connected layer for classification. the potential of SPNN to work feature maps of the full image in Associate in Nursing passing single-shot resulted in necessary improvement in object detection speed by the magnitude of nearly twenty folds larger than R-CNN. Next, fast R-CNN is associate extension over R-CNN and SPPNet. It introduces a greenhorn layer named Region of Interest (RoI) pooling layer between shared convolutional layers to fine-tune the model. Moreover, it permits to at constant time train a detector and regress or whereas not sterilization the network configurations. tho' Fast-R-CNN effectively integrates the

benefits of R-CNN and SPPNet but still lacks in detection speed compared to single-stage detectors .

Further, faster R-CNN is Associate in nursing amalgam of fast R-CNN and Region Proposal Network (RPN). It permits nearly cost-free region proposals by step by step group action individual blocks (e.g. proposal detection, feature extraction and bounding box regression) of the article detection system throughout one step. This integration winds up within the accomplishment of break-through for the speed bottleneck of fast R-CNN but there exists a computation redundancy at subsequent detection stage. The Region-based completely Convolution Network (R-FCN) is that the exclusively model that allows complete back propagation for employment and reasoning. Feature Pyramid Networks (FPN) can discover non-uniform objects, but least used by researchers due to high computation worth and heaps of memory usage . What's additional, Mask R-CNN strengthens faster R-CNN by along with the prediction of divided masks on each RoI . Two-stage yields high object detection accuracy, but it's restricted by low reasoning speed in period for video investigating [14].

III. PROPOSED SYSTEM MACHINE LEARNING APPROCH

Machine learning strategies square measure trained on datasets and a model is made for analysis.
Based on the accuracy of the model, the machine learning technique is appropriate. The three
methods in machine learning algorithms square measure supervised learning, unattended learning and
reinforcement learning. In supervised learning, the model is trained victimization labeled information that
contains each input and results. The sections of process square measure coaching section and testing phase.
Unsupervised learning strategies don't use coaching information or labeled information. It finds the hidden structures or patterns from unlabeled information.

Supervised Learning

Supervised learning needs a well-labelled dataset to coach. supervised learning is of 2 types particularly regression and classification. Classification techniques facilitate to seek out the appropriate category labels which may predict the positive, negative and neutral sentiments. A machine learning model is developed that uses the tagged knowledge to coach, classify the tweets and predict the emotions of the tweets. call Tree, Random Forest, Bayesian belief network, Naive Thomas Bayes and KNN classifiers ar a number of the algorithms that ar employed in this method.

Unsupervised Learning

Unsupervised ways ar supported machine learning or lexicon. the necessity of the labelled datasets isn't needed in unsupervised learning. Sentiment analysis once done using unsupervised learning; it's typically supported a Sentiment Lexicon. Text classification helps to extract phrases that contain adjectives or adverbs to estimate a phrase"s linguistics

orientation. linguistics orientation is then accustomed classify the emotions.

The model planned here is supposed and sculptural model python libraries specifically Tensor flow, Keras and OpenCV. The model we've a bent to used is that the MobileNetV2 of convolurional neural network. The plan of action of victimization MobileNetV2 is termed model Transfer Learning. Transfer learning is victimization some pre trained model to educate your gift model and procure the prediction that saves time and makes victimization work the assorted models straightforward. We've a bent to tune the model with the hyper parameters: learning rate, vary of epochs and batch size. The model is trained with a dataset of images with two class, with mask and whereas not mask. The dataset has 993 footage of with mask class and 1918 footage of whereas not mask class.

(i)   Training the model with the taken dataset.
(ii)   Deploying the model

In the paper we have got developed model victimization the on high of mentioned libraries. we have got tested the model fully totally completely different for various} conditions with different hyper parameters, that the results square measure mentioned inside ensuing section. first we've a bent to feed the dataset inside the model, run the tutorial program, that trains the model on the given dataset. Then we've a bent to run the detection program, that activates the video stream, captures the frames unendingly from the video stream with academic degree anchor box model object detection technique .is this can be often skillful the MobileNetV2 model layers that classifies the image like or whereas not mask. If the person is sporting a mask, a inexperienced anchor box is displayed and red if not sporting a mask with the accuracy for the same tagged on the anchor box.
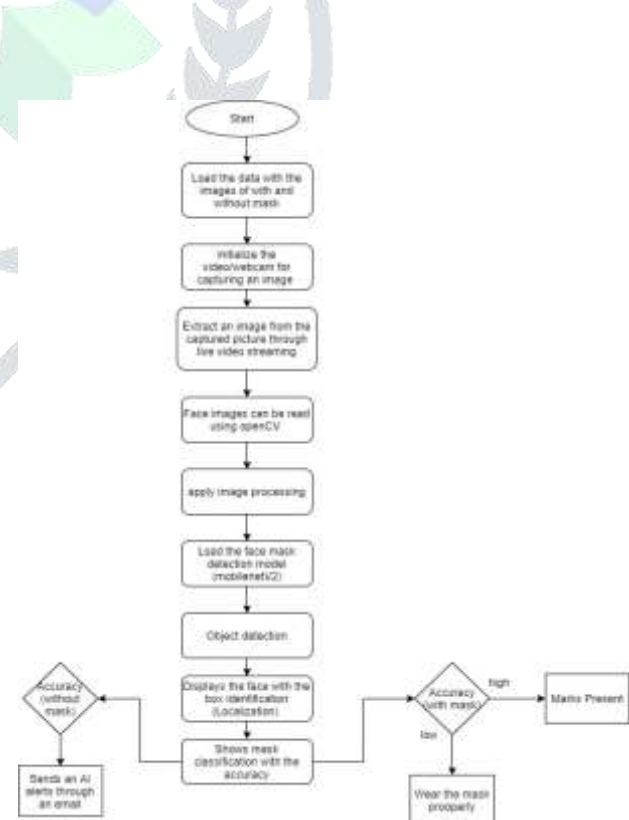

Fig. Proposed Model Architecture

IV. CONCLUSION

Due to the urgency of dominant COVID-19, the appliance price and importance of Real-time mask detection are increasing. to handle this issue, we tend to designed the PWMFD with 9205 quality covert face pictures and developed SE-YOLOv3, a quick and correct mask Detector with a channel attention mechanism that increased the feature extraction capability of the backbone network. what is more, we tend to used GIoU and focal loss and adopted the corresponding knowledge augmentation to enhance the accuracy and strength of the model.

V. REFERENCES

[1] A. G. Howard, M. Zhu, B. Chen et al., "Mobilenets: efficient convolutional neural networks for mobile vision applications," 2017, https://arxiv.org/abs/1704.04861.

[2] Wei Wang, Yutao Li, Ting Zou, Xin Wang, Jieyu You, Yanhong Luo, "A Novel Image Classification Approach via Dense-MobileNet Models", Mobile Information Systems, vol. 2020,ArticleID 7602384, 8 pages, 2020. https://doi.org/10.11 55/2020/7602384

[3] I. B. Venkateswarlu, J. Kakarla and S. Prakash, "Face mask detection using MobileNet and Global Pooling Block,"4 2020 IEEE 4th Conference on Information & Communication Technology (CICT), 2020, pp. 1-5, doi: 10.1109/CICT51604.2020.9312083.

[4] M. S. Ejaz and M. R. Islam, "Masked Face Recognition Using Convolutional Neural Network," 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI), 2019, pp. 1-6, doi: 10.1109/STI47673.2019.9068044

[5] Changjin Li, Jian Cao, and Xing Zhang. 2020. Robust Deep Learning Method to Detect Face Masks. In <i>Proceedings of the 2nd International Conference on Artificial Intelligence and Advanced Manufacture</i> (<i>AIAM2020</i>). Association for Computing Machinery, New York, NY, USA, 74–77. DOI:https://doi.org/10.1145/3421766.342176