PROJECT REPORT ON

Face Mask Detection

SUBMITTED BY

Shaikh Parvez Alam

MSc. Information Technology

Academic Year 2022-2023

A Project Report

Submitted in contentment of the requirements for the award of the degree

of

MASTERS OF SCIENCE IN INFORMATION TECHNOLOGY

By

Shaikh Parvez Alam

PRN: 2018016401053031

DEPARTMENT OF INFORMATION TECHNOLOGY

Chembur-Trombay Education Society's

N. G Acharya & D. K. Marathe College

Of Arts, Commerce, Science & Professional Courses

N.G Acharya Marg, Chembur (E), Mumbai-400071.

PROFORMA FOR THE APPROVAL PROJECT PROPOSAL

PRN No**.:** A2018016401053031

1. Name of the Student

   Shaikh Parvez Alam

2. Title of the Project

   Face Mask Detection

3. Is this your first submission? Yes No

Signature of the Student

Date: …………………

Signature of the Coordinator

Date: ……………….

Chembur-Trombay Education Society's

N. G Acharya & D. K. Marathe College

Of Arts, Commerce, Science & Professional Courses

N.G Acharya Marg, Chembur (E), Mumbai-400071

DEPARTMENT OF INFORMATION TECHNOLOGY

<u>CERTIFICATE</u>

This is to certify that the project entitled, **"**Face Mask detection**"**, is bonafide work of **Shaikh Parvez Alam** bearing Seat No: 4135828 submitted impartial fulfilment of the requirements for the award of degree of MASTER OF SCIENCE in INFORMATION TECHNOLOGY from University of Mumbai.

Internal Guide **Coordinator**

External Examiner

Date: **College Seal**

List of Tables

ACKNOWLEDGEMENT

I'm so glad to present my project report on "**Face Mask Detection**", which is prepared as a part of my Third semester project in "Masters of Science in Information Technology “This project work is the outcome of guidance provided by dedicated and experienced personnel of "**N.G. Acharya & D.K. Marathe College**". We are thankful to God and our parents for helping us to get a good education.

With profound sense of indebtedness, we use this opportunity to express our sincere thanks to our beloved principal for providing all support and necessary facilities throughout the course and for the inspiration to complete my project. We express our sincere gratitude to Head, “**Priti Palkar**” coordinator of Masters of Science in Information Technology Department for inspiration and permission to carry out our project. We wish to extend our thanks to our project guide and Subject guides "**Dr. Hiren Dand**",“**Dr.Abhijeet Kale**”.

Last but not the least we extend our sincere thanks to all our classmates for their help and support. And also we would like to extend our deep sense of gratitude to all those who have assisted us in the successful completion of this project.

Thanking You

Shaikh Parvez Alam

ABSTRACT

The coronavirus COVID-19 pandemic is causing a world health crisis. one in every of the effective protection methods is wearing a mask publically areas in step with the globe Health Organization (WHO). Nowadays advancement of synthetic brainpower is effectively creating; they open up tremendous potential outcomes before us. Investigation, gauging, detection visited another level with the employment of synthetic reasoning advancements.

Starting late, a unimaginably promising field of exploration is Computer vision. Face identification is where recognizing the countenances from the pictures or video sources. It might be used and requirement for circumstance to reinsure that each individual is wearing appropriate cover as per the rule gave by WHO. Regardless of immense assortments in visual overhauls on account of developing condition, developing, and interferences like bristles, glasses, and hair style changes, this limit is amazingly incredible.

1. Introduction

The trend of wearing face masks in public is rising due to the COVID- 19 corona virus epidemic all over the world. Before Covid-19, People used to wear masks to protect their health from air pollution. While other people are self-conscious about their looks, they hide their emotions from the public by hiding their faces. Scientists proofed that wearing face masks works on impeding COVID-19 transmission. COVID19 (known as corona virus) is the latest epidemic virus that hit the human health in the last century. In 2020, the rapid spreading of COVID-19 has forced the World Health Organization to declare COVID- 19 as a global pandemic. More than five million cases were infected by COVID-19 in less than 6 months across 188 countries. The virus spreads through close contact and in crowded and overcrowded areas. The corona virus epidemic has given rise to an extraordinary degree of worldwide scientific cooperation. Artificial Intelligence (AI) based on Machine learning and Deep Learning can help to fight Covid-19 in many ways. Machine learning allows researchers and clinicians evaluate vast quantities of data to forecast the distribution of COVID-19 to serve as an early warning mechanism for potential pandemics, and to classify vulnerable populations. The provision of healthcare needs funding for emerging technology such as artificial intelligence, IoT, big data and machine learning to tackle and predict new diseases. In order to better understand infection rates and to trace and quickly detect infections, the AI's power is being exploited to address the Covid-19 pandemic. People are forced by laws to wear face masks in public in many countries .These rules and laws were developed as an action to the exponential growth in cases and deaths in many areas. However, the process of monitoring large groups of people is becoming more difficult. The monitoring process involves the detection of anyone who is not wearing a face mask. Here we introduce a mask face detection model that is based on computer vision and deep learning. The proposed model can be integrated with surveillance cameras to impede the COVID-19 transmission by allowing the detection of people who are wearing masks not wearing face masks. The model is integration between deep learning and classical machine learning techniques with opencv, tensor flow and keras. We have used deep transfer leering for feature extractions and combined it with three classical machine learning algorithms.

Overview

Due to COVID-19 wearing a face mask is important and vital part of SOP issued by the WHO so our project is to ensure a person wearing mask or not which can be helpful to ensure the SOP is followed by the people in an concentrated areas like malls office restaurant and different areas.

Face mask detection is a simple model to detect face mask.Due to COVID-19 there is need of face mask detection application on many places like Malls and Theatres for safety.By the development of face mask detection we can calculate if the person's mask is on the face or not and allow their entry would be of great help to the society. Face Mask detection model is built using the machine learning it can be supervised learning or unsupervised learning where as in this case we use (CNN). This CNN Model is built using the TensorFlow framework and the OpenCV library which is highly used for real-time applications.

Object are normally perceived by their novel highlights. There are numerous highlights in a human face, which can be perceived between a face and numerous different articles. It finds faces by removing basic highlights like eyes, nose, mouth and so forth and afterward utilizes them to identify a face. Whereas there are many features in a very external body part, which might be recognized between a face and lots of other objects. It locates faces by extracting structural features like eyes, nose, mouth etc. so uses them to detect a face. Typically, some variety of statistical classifier qualified then helpful to separate between facial and non-facial regions. additionally, human faces have particular textures which might be accustomed differentiate between a face and other objects. Moreover, the sting of features can help to detect the objects from the face.In the coming section, we will implement a feature-based approach.

1.1 Objective:-

Image-put together approaches depend with respect to methods from factual examination and AI to discover the related qualities of face and non-face pictures. The learned qualities are as dispersion models or discriminant capacities that is thus applied for face discovery. In this method, we use different algorithms such as Neural-networks and other leaning techniques used for it like SVM and other more. In the coming section, we will see how we can detect faces with MTCNN or Multi-Task Cascaded Convolutional Neural Network, which is an Image-based approach of face detection

Face recognition is a method of identifying or verifying the identity of an individual using their face. There are various algorithms that can do face recognition but their accuracy might vary. Here I am going to describe how we do face recognition using deep learning.

1.1.2 Scope and Applicability :

The current ongoing system is gracing with MobileNetV2 classifier one of the best system which would be implemented along with the interface of alarm and alerting system in future generation. This system will be integrated with the system implementing social distancing that would make it a complete system which can bring a dramatic impact on the spread of. The new world will be well being of high demand of mask as faceless future and that will be a big security concern. Expertise say, CNN that using face mask proves to be the best solution to mitigate the spread of air borne virus like corona, but as a big security concern headed to challenge the nation as it would create a massive opportunity for people who cover their faces for nefarious reason. And also experts say the mass no of mask wearing in could complicate in crime investigation in the coming days, as facial recognition is an important part in tracking of the criminals. When the pandemic covid-19 getting over, then this system comes into play for chemical factories, bank, glass factories etc. If a person enters the bank while wearing a mask he would be not allowed to enter and also if the person does not wear masks in glass factories chemical factories and etc. then the person would not be allowed to

enter to the industry. A mind concept of human being have been proved out to be very good at recognizing familiar faces and facial recognizing familiar faces and facial recognition algorithms are getting better in identifying pattern. So thus this challenge would create a scope to new face detection algorithms which can identify aces which are covered with greater accuracies and precisions

2. Literature survey

Generally, most of the projects specialize in face construction identity recognition when wearing mask. During this projects, the focus is on recognizing the people that wearing mask, or not help in decreasing the transmission and spreading of covid-19. The scientist has proven that wearing a mask help in minimizing the spreading rate of Covid-19. The authors developed a face mask wearing condition identification method. They were ready to classify three categories of face mask-wearing. The categories are face maskwearing, incorrect face mask-wearing and no face mask-wearing. We have applied the principal component analysis on masked and unmasked face recognition to acknowledge the person. Also, PCA was utilized . The author proposed a way that's used for removing glasses from human frontal faces.The authors used the YOLOv3 algorithm for face detection. YOLOv3 uses Darknet-53 because the backbone. Proposed a completely unique GAN-based network, which will automatically remove mask covering the face area and regenerate the image by building the missing hole. The authors presented a system for detecting the presence or absence of a compulsory medical mask within the OR. The general is to attenuate the false positive face detection as possible without missing mask detection so as to trigger alarms just for medical staff who don't wear a surgical mask. Used deep learning real-time face emotion classification and recognition.

2.1 Literature Riview

2.1.1 Literature Review Introduction :

Following chapter is discussing, below mentioned are related research papers which are using mainly and some of the similar authentication process like ours or go with the hardware that we are using in the project. The purpose of studying and going through these related research papers is to get a wider perspective and the view of the existing similar projects that how researcher have tried to resolve the similar problems and through which tools, techniques and procedures which have helped them in their researches and projects. As the Smart Irrigation system is not a new idea or to have it as the project to work on, there are already many systems proposed to resolve such issues and problems. therefore, this chapter will briefly define and help us in order to understand all the related work of other researchers and students in this particular area that if there are any glitches and if there is need to improve this system including some innovations in it which we will try to accomplish with best of our abilities in order to make human effort more less and make it more easier to use.

2.2 Related Research Work

In face detection method, a face is detected from an image that has several attributes in it. According to research into face detection requires expression recognition, face tracking, and pose estimation. Given a solitary image, the challenge is to identify the face from the picture. Face detection is a difficult errand because the faces change in size, shape, color, etc and they are not immutable. It becomes a laborious job for opaque image impeded by some other thing not confronting camera, and so forth. Authors think occlusive face detection comes with two major challenges: 1) unavailability of sizably voluminous datasets containing both masked and unmasked faces, and 2) exclusion of facial expression in the covered area. Utilizing the locally linear embedding (LLE) algorithm and the dictionaries trained on an immensely colossal pool of masked faces, synthesized mundane faces, several mislaid expressions can be recuperated and the ascendancy of facial cues can be mitigated to great extent. According to the work reported in, convolutional neural network (CNNs) in computer vision comes with a strict constraint regarding the size of the input image. The prevalent practice reconfigures the images before fitting them into the network to surmount the inhibition.

Here the main challenge of the task is to detect the face from the image correctly and then identify if it has a mask on it or not. In order to perform surveillance tasks, the proposed method should also detect a face along with a mask in motion.

2.2.1 System Facial Recognition using OpenCV

The developing interest in PC vision of the former decade. Bursting by the constantgrowing pace of processing influence at regular breaks, face location and salutation has risen above from a recondite to a mainstream region of exploration in PC vision and one of the better and effective utilizations of picture examination and calculation based arrangement. In view of the inherent idea of the issue, PC vision isn't just a software engineering territory of examination, yet additionally the object of neuro-logical and mental investigations, mostly in light of the overall feeling that progresses in PC picture handling and understanding exploration will give bits of knowledge into how our cerebrum work and the other way around.

Due to the intrinsic nature of the matter, computer vision isn't only a technology area of research, but also the article of neuro-scientific and psychological studies, mainly thanks to the overall opinion that advances in computer image processing and understanding research will provide insights into how our brain work and the other way around. thanks to general curiosity and interest within the matter, the author has proposed to form an application that might allow user access to a selected machine supported an in-depth analysis of a person’s countenance.

Over the past decade face detection and recognition have transcended from esoteric to popular areas of research in computer vision and one among the higher and successful applications of image analysis and algorithm based understanding. due to the intrinsic nature of the matter, computer vision isn't only a computing area of research, but also the thing of neuroscientific and psychological studies also, mainly thanks to the overall opinion that advances in computer image processing and understanding research will provide insights into how our brain work and the other way around. A general statement of the face recognition problem (in computer vision) are often formulated as follows: given still or video images of a scene, identify or verify one or more persons within the scene employing a stored database of faces. face recognition generally involves two stages: Face Detection where a photograph is searched to search out a face, then the image is processed to crop and extract the person’s face for easier recognition. Face Recognition where that detected and processed face is compared to a database of known faces, to make a decision who that person is. Since 2002, face detection are often performed fairly easily and reliably with Intel’s open source framework called OpenCV . This framework has an inbuilt Face Detector that works in roughly 90-95% of clear photos of someone looking forward at the camera. However, detecting a person’s face when that person is viewed from an angle is sometimes harder, sometimes requiring 3D Head Pose Estimation.Selected image searched to find a face, then the image is processed to crop and extract the person’s face for easier recognition. Face Recognition where that detected and processed face is compared to a database of known faces, to decide who that person is. Since 2002, face detection can be performed fairly easily and reliably with Intel’s open source framework called OpenCV . This framework has an inbuilt Face Detector that works in roughly 90-95% of clear photos of a person looking forward at the camera. However, detecting a person’s face when that person is viewed from an angle is usually harder, sometimes requiring 3D Head Pose Estimation. Also, lack of proper brightness of an image can greatly increase the difficulty of detecting a face, or increased contrast in shadows on the face, or maybe the picture is blurry, or the person is wearing glasses, etc. Face recognition however is much less reliable than face detection, with an accuracy of 30-70% in general. Face recognition has been a strong field of research since the 1990s, but is still a far way away from a reliable method of user authentication. More and more techniques are being developed each year. The Eigenface technique is considered the simplest method of accurate face recognition, but many other (much more complicated) methods or combinations of multiple methods are slightly more accurate.

2.2.2 Face Detection and Recognition using Open CV Based on Fisher Faces Algorithm

Facial Recognition represents the event of a system which can determine the person with the help of a face using Computer Vision (Open CV). Face recognition is used within the fields of Identity Recognition, investigation and enforcement. it is a method of characteristic someone supported facial features. This method is enforced in 2 stages. They're the training stage and thus the testing stage. This study primarily consists of three elements, specifically face detection from the image, feature extraction and storing many reminder images, and recognition. Face finding rule is utilized to detect the face from the given image. The foremost helpful and distinctive options of the face image are extracted within the feature extraction part. Face Detection is also challenging thanks to pictures and video frames will contain advanced background, completely different head poses and occlusion like carrying glasses or scarf. It presents a rule for locating face recognition downside and concatenated into one feature vector that's employed to teach the system to recognise among the prevailing photos with it. Within the testing stage the system takes the face of the image of somebody for recognition. Image acquisition, preprocessing, image filtering, feature extraction is simply just like the learning stage. For classification the choices are fed to the trained system. The algorithms can determine the face image from the content and acknowledges it.

Face recognition takes a photograph from a video or a camera as input and outputs the diagnosed photo material. countenance may additionally accommodates regions inside the face, variations within the face structure, face cuts and angles which are formatted and styled. Face extraction includes grabbing of the capabilities from camera. Face detection includes the elimination of the background and that specialize in the foreground eliminating another elements aside from the face vicinity, but the device nevertheless pertains some drawbacks because it cannot stumble upon the top be counted which might be a present thanks to overlapping of faces or mistaken recognition of faces having similar facial functions.

•Find faces - no matter whether the errand of perceiving individuals in photos, or video acknowledgment, or whatever else.

•Face positioning - pics aren't regularly located on which a personal stand straightforwardly before the main focus, often the face grows to become, we face the challenge of situating it as if the image become taken legitimately.

•Defining outstanding facial capabilities - this development may be brought up as a full face acknowledgment step, it examinations the photograph and gets certainly one amongst a sort automated estimations of the face.

•Identification of an individual - we assessment a got information and therefore the information efficiently accessible to us, if the statistics are similar, we are going to show the decision of the character, if now not, in like way we've no longer recognised at this time to us character. this may analyse in element all of the means to manufacture a face acknowledgment framework and comparison their execution and also the help of assorted libraries, simply because the velocity of crafted by way of each section in various libraries of Computer vision

2.2.3 FACE DETECTION & FACE RECOGNITION USING OPEN COMPUTER VISION CLASSIFIRES

The accompanying record is a report on the small venture for Robotic visual observation and self-governance. It included structure a framework for face location and face acknowledgment utilizing a few classifiers accessible in the open PC vision library(OpenCV). Face acknowledgment is a non-obtrusive identification framework and quicker than different frameworks since numerous countenances can be examined simultaneously. The difference between face discovery and identification is, face recognition is to recognize a face from a picture and find the face. Face acknowledgment is settling on the choice "whose face is it ? ", utilizing a picture information base. In this task both are cultivated utilizing different strategies and are portrayed beneath. The report starts with a concise history of face acknowledgment. This is trailed by the clarification of HAAR-falls, Eigenface, Fisherface and Local paired example histogram (LBPH) calculations. Next, the philosophy and the aftereffects of the venture are depicted. A conversation with respect to the difficulties and the goals are depicted. At last, an end is given on the upsides and downsides of every calculation and potential executions.

A Haar may be a mathematical fiction that produces square-shaped waves with a beginning and an end and accustomed create box shaped patterns to recognise signals with sudden transformations. An example is shown in figure 1. By combining several wavelets, a cascade will be created that may identify edges, lines and circles with different colour intensities. These sets are utilized in Viola Jones face detection technique in 2001 and since then more patterns are introduced [10] for object detection as shown in figure 1. To analyse a picture using Haar cascades, a scale is chosen smaller than the target image. it's then placed on the image, and also the average of the values of pixels in each section is taken. If the difference between two values pass a given threshold, it's considered a match. Face detection on a personality's face is performed by matching a mix of different Haar-like-features. as an example, forehead, eyebrows and eyes contrast yet because the nose with eyes as shown below in figure one classifier isn't accurate enough. Several classifiers are combined on provide an accurate face detection system as shown within theblock diagram below in figure

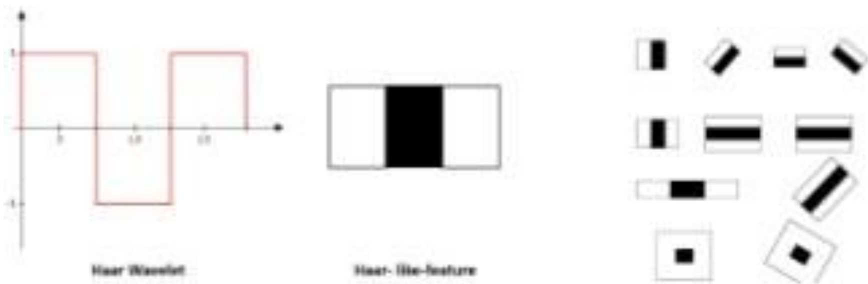
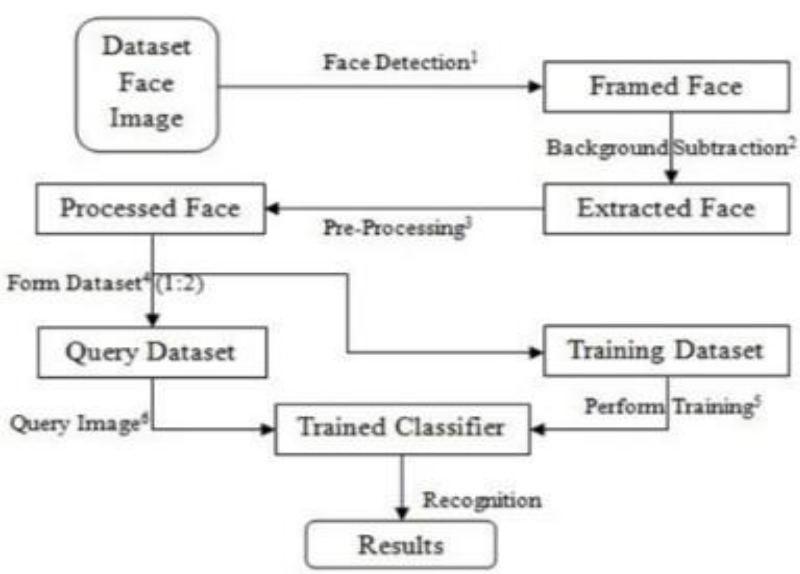


Image-based Face Detection and Recognition

Face recognition from image or video could be a popular topic in biometrics research. Many public places usually have surveillance cameras for video capture and these cameras have their significant value for security purpose. it's widely acknowledged that the face recognition have played a vital role in closed-circuit television because it doesn't need the object's cooperation. the particular advantages of face based identification over other biometrics are uniqueness and acceptance. As face may be a dynamic object having high degree of variability in its appearance, that produces face detection a difficult problem in computer vision. during this field, accuracy and speed of identification may be a main issue.

The objective of this paper is to assess different face discovery and acknowledgment strategies, give total answer for picture based face location and acknowledgment with higher exactness, better reaction rate as an underlying advance for video observation. Arrangement is proposed dependent on performed tests on different face rich information bases regarding subjects, present, feelings, race and light.



FACE ANALYSIS AND FACE RECOGNITION

Images and videos play a crucial role in shaping and reflecting life. Digitization has vastly increased the presence of such images in daily life, creating valuable new research opportunities for social scientists. It suggests show recent innovations in computer vision methods can substantially lower the costs of using images as data. It has been introduced the deep learning algorithms commonly used for object recognition, facial recognition, and visual sentiment analysis. There are a lot of guidance and specific instructions for scholars interested in using these methods in their own research. Early research study presented a hybrid neural-network for human face recognition which compares favourably with other methods .The system combines local image sampling, a self-organizing map (SOM) neural network, and a convolutional neural network. The SOM provides a quantization of the image samples into a topological space where inputs that are nearby in the original space are also nearby in the output space, thereby providing dimensionality reduction and invariance to minor changes in the image sample, and the convolutional neural network provides partial invariance to translation, rotation, scale, and deformation. The convolutional network extracts successively larger features in a hierarchical set of layers. The research present results using the Karhunen-Loeve transform in place of the SOM, and a multilayer perceptron (MLP) in place of the convolutional network for comparison. A database of 400 images of 40 individuals which contains quite a high degree of variability in expression, pose, and facial details.It analyzes computational complexity and discuss how new classes could be added to the trained recognizer. This study elaborates one of the approach: Neural Network. Figure 1 describes how CNN introduced in the early study of face recognition. There are two primary applications of machine learning that analyze images containing faces: face detection and face comparison. A face detection system is designed to answer the question: is there a face in this picture? A face detection system determines the presence, location, scale, and (possibly) orientation of any face present in a still image or video frame. This system is designed to detect the presence of faces regardless of attributes such as gender, age, and facial hair. A face comparison system is designed to answer the question: does the face in an image match the face in another image? A face comparison system takes an image of a face and makes a prediction about whether the face matches other faces in a provided database. Face comparison systems are designed to compare and predict potential matches of faces regardless of their expression, facial hair, and age. Both face detection and face comparison systems can provide an estimate of the confidence level of the prediction in the form of a probability or confidence score. For example, a face detection system may predict that an image region is a face at a

confidence score of 90%, and another image region is a face at a confidence score of 60%. The region with the higher confidence score should be more
For likely to contain a face. If a face detection system does not properly detect a face, or provides a low confidence prediction of an actual face, this is known as a missed detection or false negative. If a facial detection system incorrectly predicts the presence of a face at a high confidence level, this is a false alarm or false positive. Similarly, a facial comparison system may not match two faces belonging to the same person (missed detection/false negative), or may incorrectly predict that two faces from different people are the same person (false alarm/false positive).Confidence scores are a critical component of face detection and comparison systems. These systems make predictions of whether a face exists in an image or matches a face in another image, with a corresponding level of confidence in the prediction. Users of these systems should consider the confidence score/similarity threshold provided by the system when designing their application and making decisions based on the output of the system. For example, in a photo application used to identify similar looking family members, if the confidence threshold is set at 80%, then the application will return matches when predictions reach an 80% confidence level, but will not return matches below that level. This threshold may be acceptable because the risk of missed detections or false alarms is low for this type of use case. However, for use cases where the risk of missed detection or false alarm is higher, the system should use a higher confidence level. One should use a 99% confidence/similarity threshold in scenarios where highly accurate facial matches are highly important

2.3 Methodology

2.3.1 Feasibility Analysis

First of all explain our project on the ground of feasibility analysis means briefly describe what do you meant by facemask detection to feasibility analysis.This project provide ease in deciding or identifying the particular person without mask**.**

2.3.2 Technical Feasibility

Face mask detection system is technically feasible that it can be opted and design which is feasible according to every aspect about cost, innovative solution and economical.

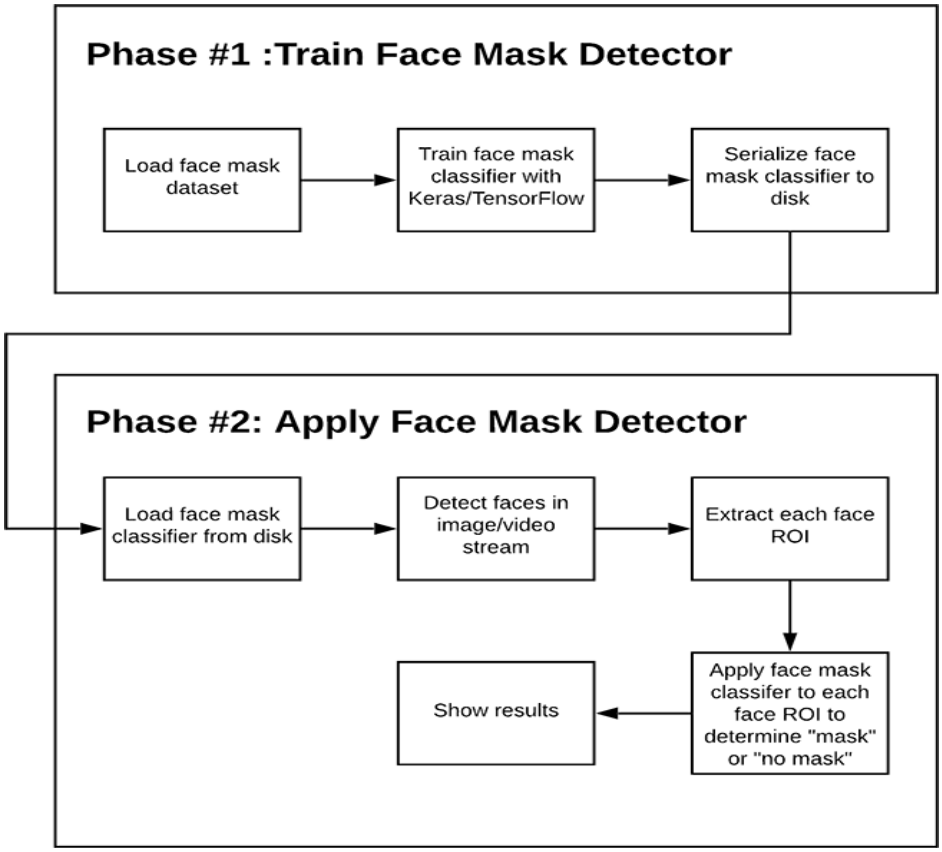
2.3.3Operational Feasibility

The proposed and designed solution of facemask detection system is operationally feasible because it has many advantages which leads to its operationally possible solution because it is economical and cost effective secure and speedy which have a great impact of it feasibility for operation.

2.3.4 Financial Feasibility

According to world health organization (WHO) wearing mask is considered as least SOP for COVID-19 to follow to in order maintain the business to our project can be done using same CCTV cameras to detect weather a person is wearing a mask or not so its financially feasible

Two-phase COVID-19 face mask detector



Phases and individual steps for building a COVID-19 face mask detector with computer vision using Python, OpenCV, and TensorFlow/Keras.

In order to train a custom face mask detector, we need to break our project into two distinct phases, each with its own respective sub-steps (as shown by Figure above):

1. **Training:** Here we'll focus on loading our face mask detection dataset from disk, training a model (using Keras/TensorFlow) on this dataset, and then serializing the face mask detector to disk

2. **Deployment:** Once the face mask detector is trained, we can then move on to loading the mask detector, performing face detection, and then classifying each face as

   with_mask
   or
   without_mask

We'll review each of these phases and associated subsets in detail in the remainder of this tutorial, but in the meantime, let's take a look at the dataset we'll be using to train our COVID-19 face mask detector.

2.4 PROPOSED APPROACH

The proposed system focuses on how to identify the person on image/video stream wearing face mask with the help of computer vision and deep learning algorithm by using the OpenCV, Tensor flow, Keras library.

Approach :

1. Train Deep learning model (MobileNetV2)
2. Apply mask detector over images / live video stream

Data at Source :

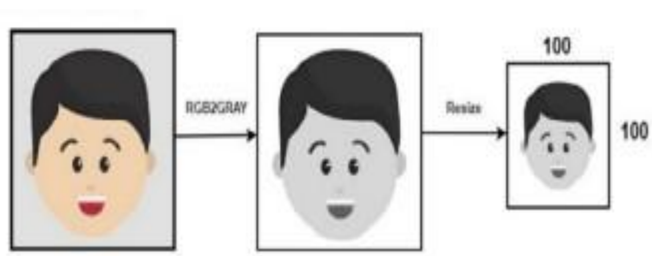
The majority of the images were augmented by OpenCV. The set of images were already labeled “mask” and “no mask”. The images that were present were of different sizes and resolutions, probably extracted from different sources or from machines (cameras) of different resolutions.

Data Preprocessing :

Preprocessing steps as mentioned below was applied to all the raw input images to convert them into clean versions, which could be fed to a neural network machine learning model. Data preprocessing involves conversion of data from a given format to much more user friendly, desired and meaningful format. It can be in any form like tables, images, videos, graphs, etc. These organized information fit in with an information model or composition and captures relationship between different entities . The proposed method deals with image and video data using Numpy and OpenCV.

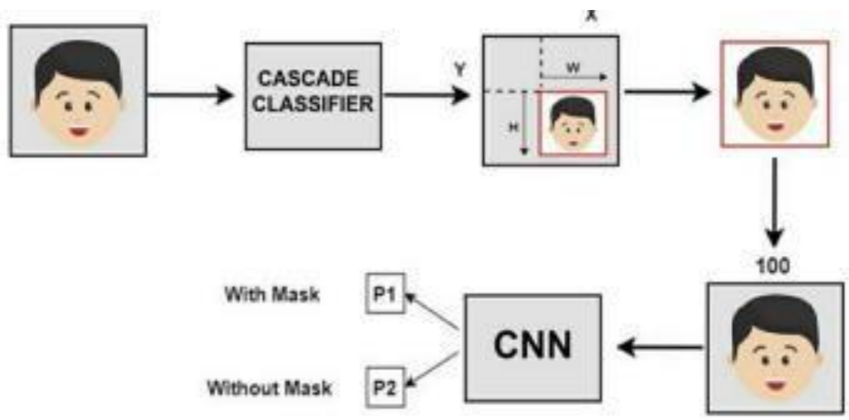
Data Visualization: Data visualization is the process of transforming abstract data to meaningful representations using knowledge communication and insight discovery through encodings. It is helpful to study a particular pattern in the dataset . The total number of images in the dataset is visualized in both categories – ‘with mask’ and ‘without mask’

Conversion of RGB image to Gray image: Modern descriptor-based image recognition systems regularly work on grayscale images, without elaborating the method used to convert from color-to-grayscale. This is because the colorto-grayscale method is of little consequence when using robust descriptors. Introducing nonessential information could increase the size of training data required to achieve good performance. As grayscale rationalizes the algorithm and diminishes the computational requisites, it is utilized for extracting descriptors instead of working on color images instantaneously.



Conversion of RGB image to Gray image of 100 x 100 size

Splitting the data and training the CNN model: After setting the blueprint to analyze the data, the model needs to be trained using a specific dataset and then to be tested against a different dataset. A proper model and optimized train test split help to produce accurate results while making a prediction. The test size is set to 0.1 i.e. 90% data of the dataset undergoes training and the rest 10% goes for testing purposes. The validation loss is monitored using ModelCheckpoint. Next, the images in the training set and the test set are fitted to the Sequential model. Here, 20% of the training data is used as validation data. The model is trained for 20 epochs (iterations) which maintains a trade-off between accuracy and chances of overfitting. Fig show depicts visual representation of the proposed model



Overview of model

Image Reshaping: The input during relegation of an image is a three-dimensional tensor, where each channel has a prominent unique pixel. All the images must have identically tantamount size corresponding to 3D feature tensor. However, neither images are customarily coextensive nor their corresponding feature tensors. Most CNNs can only accept fine-tuned images. This engenders several problems throughout data collection and implementation of model. However, reconfiguring the input images before augmenting them into the network can help to surmount this constraint. The images are normalized to converge the pixel range between 0 and 1. Then they are converted to 4 dimensional arrays using data = np.reshape(data,(data.shape[0], img size,img size,1)) where 1 indicates the Grayscale image. As, the final layer of the neural network has 2 outputs – with mask and without mask i.e. it has categorical representation, the data is converted to categorical labels.

Training of Model : Building the model using CNN architecture: CNN has become ascendant in miscellaneous computer vision tasks . The current method makes use of Sequential CNN. The First Convolution layer is followed by Rectified Linear Unit (ReLU) and MaxPooling layers. The Convolution layer learns from 200 filters. Kernel size is set to 3 x 3 which specifies the height and width of the 2D convolution window. As the model should be aware of the shape of the input expected, the first layer in the model needs to be provided with information about input shape. Following layers can perform instinctive shape reckoning. In this case, input shape is specified as data.shape which returns the dimensions of the data array from index 1. Default padding is "valid" where the spatial dimensions are sanctioned to truncate and the input volume is non-zero padded. The activation parameter to the Conv2D class is set as "relu". It represents an approximately linear function that possesses all the assets of linear models that can easily be optimized with gradient-descent methods. Considering the performance and generalization in deep learning, it is better compared to other activation functions . Max Pooling is used to reduce the spatial dimensions of the output volume. Pool size is set to 3 x 3 and the resulting output has a shape (number of rows or columns) of: shape of output = (input shape - pool size + 1) / strides), where strides has default value (1,1)

The learning process needs to be configured first with the compile method [13]. Here "adam" optimizer is used. categorical crossentropy which is also known as multiclass log loss is used as a loss function (the objective that the model tries to minimize). As the problem is a classification problem, metrics is set to "accuracy".

Support vector Machine

One of the most popular and spectacular supervised learning techniques with related learning algorithms for treatment classification and regression tasks in patterns is SVM. SVM is a classification machine learning algorithm based on hinge function as shown in Eq. (1), where z is a label from 0 to 1, w.I−b is the output, w and b are coefficients of linear classification, and I is an input vector

Decision tree

The decision tree is the classification model of computation based on entropy function and information gain. Entropy computes the amount of uncertainty in data. Where D is current data, and q is a binary label from 0 to 1, and p(x) is the proportion of q label. To measure the difference of entropy from data, we calculate information gain (I) as illustrated in eq. Where v is a subset of data

Ensemble methods

Ensemble methods are algorithms of machine learning that create a collection of classifiers. An ensemble of classifiers is a collection of classifiers whose individual decisions (usually by weighted or unweighted voting) are merged in one way or another to identify new instances . The used Ensemble methods are K-Nearest Neighbors Algorithm (k-NN) , Linear Regression and Logistic Regression . The steps of the ensemble method are: 1) generate M classifiers 2) Train each classifier alone 3) merge the M classifiers and average their output. We improve our ensemble by using complex weight (α) to get better results as illustrated.

DATASET

Two datasets have been used for experimenting the current method. Dataset 1 [consists of 1376 images in which 690 images with people wearing face masks and the rest 686 images with people who do not wear face masks. Fig. 1 mostly contains front face pose with single face in the frame and with same type of mask having white color only.

Sample for dataset 1 including faces without mask and with mask

Dataset 2 from Kaggle consists of 853 images and its countenances are clarified either with a mask or without a mask. In fig. 2 some face collections are head turn, tilt and slant with multiple faces in the frame and different types of masks having different colors as well.

3. Hardware and software specification

The experiment setup is carried out on a computer system which has the different hardware and software specification as given in the table

3.1 HARDWARE DETAILS:-

Processor	Intel Pentium
HDD	200 GB
RAM	4 GB

3.2 SOFTWARE DETAILS:-

Operating system	Windows 10
Programming language	Python using Tensorflow and OpenCV

3.3 Problem Definition:

If you look at the people in videos captured by CCTV cameras, you can see that the faces are small, blurry, and low resolution. People are not looking straight to the camera, and the face angles vary from time to time. These real-world videos are entirely different from the videos captured by webcams or selfie cameras, making the face mask detection problem much more difficult in practice.

In this, we will first explore mask/ no mask classification in webcam videos, and next, shift to the mask/ no mask classification problem in real-world videos as our final goal. Our reported model can detect faces and classify masked faces from unmasked ones in webcam videos as well as real-world videos where the faces are small and blurry and people are wearing masks in different shapes and colors.

4.1 Application

4.1.1 MACHINE LEARNING : Machine learning (ML)is the study of computer algorithms that improve automatically through experience. It is seen as a subset of artificial intelligence. Machine learning algorithms build a mathematical model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so. Machine learning algorithms are used in a wide variety of applications, such as email filtering and computer vision, where it is difficult or infeasible to develop conventional algorithms to perform the needed tasks. Machine learning is closely related to computational statistics, which focuses on making predictions using computers. The study of mathematical optimization delivers methods, theory and application domains to the field of machine learning. Data mining is a related field of study, focusing on exploratory data analysis through unsupervised learning.

4.1.2 COMPUTER VISION : Computer vision is an interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos. From the perspective of engineering, it seeks to understand and automate tasks that the human visual system can do, Computer vision tasks include methods for acquiring, processing.

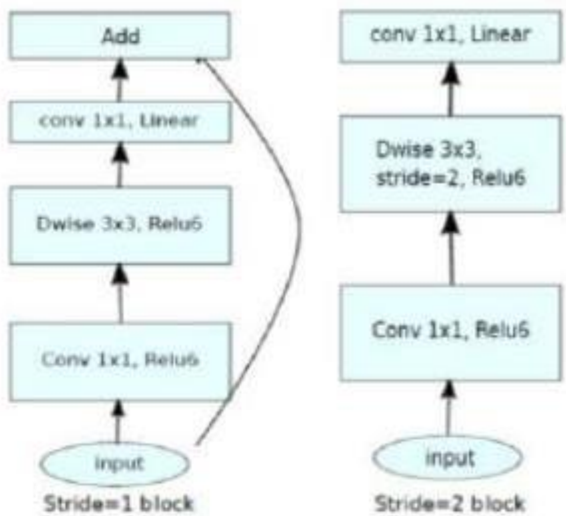
4.1.3 DEEP LEARNING : Deep learning methods aim at learning feature hierarchies with features from higher levels of the hierarchy formed by the composition of lower level features. Automatically learning features at multiple levels of abstraction allow a system to learn complex functions mapping the input to the output directly from data, without depending completely on human-crafted features. Deep learning algorithms seek to exploit the unknown structure in the input distribution in order to discover good representations, often at multiple levels, with higher-level learned features defined in terms of lower-level features

4.1.4 OpenCV (Open Source Computer Vision Library) : OpenCV is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products. Being a BSD-licensed product, OpenCV makes it easy for businesses to utilize and modify the code. The library has more than 2500 optimized algorithms,which includes a comprehensive set of both classic and state-ofthe-art computer vision and machine learning algorithms. These algorithms can be used to detect and recognize faces, identify objects, classify human actions in videos, track camera movements, track moving objects, extract 3D models of objects, produce 3D point clouds from stereo cameras, stitch images together to produce a high resolution image of an entire scene, find similar images from an image database, remove red eyes from images taken using flash, follow eye movements, recognize scenery and establish markers to overlay it with augmented reality, etc.

4.1.5 TENSORFLOW : TensorFlow is a free and open-source software library for dataflow and differentiable programming across a range of tasks. It is a symbolic math library, and is also used for machine learning applications such as neural networks. Itis used for both research and production at Google, TensorFlow is Google Brain's second-generation system. Version 1.0.0 was released on February 11, While the reference implementation runs on single devices, TensorFlow can run on multiple CPUs and GPUs (with optional CUDA and SYCL extensions for general-purpose computing on graphics processing units). Tensor Flow is available on 64-bit Linux, macOS, Windows, and mobile computing platforms including Android and iOS. Its flexible architecture allows for the easy deployment of computation across a variety of platforms (CPUs, GPUs, TPUs), and from desktops to clusters of servers to mobile and edge devices. The name Tensor Flow derives from the operations that such neural networks perform on multidimensional data arrays, which are referred to as tensors. During the Google I/O Conference in June 2016, Jeff Dean stated that 1,500 repositories on GitHub mentioned TensorFlow, of which only 5 were from Google. Unlike other numerical libraries intended for use in Deep Learning like Theano, TensorFlow was designed for use both in research and development and in production systems, not least Rank Brain in Google search and the fun Deep Dream project

4.1.6 MobileNetV2 :

MobileNetV2 is an architecture of bottleneck depth-separable convolution building of basic blocks with residuals. It has two types of blocks. The first one is a residual block with stride of 1.Second one is also residual block with stride 2 and it is for downsizing

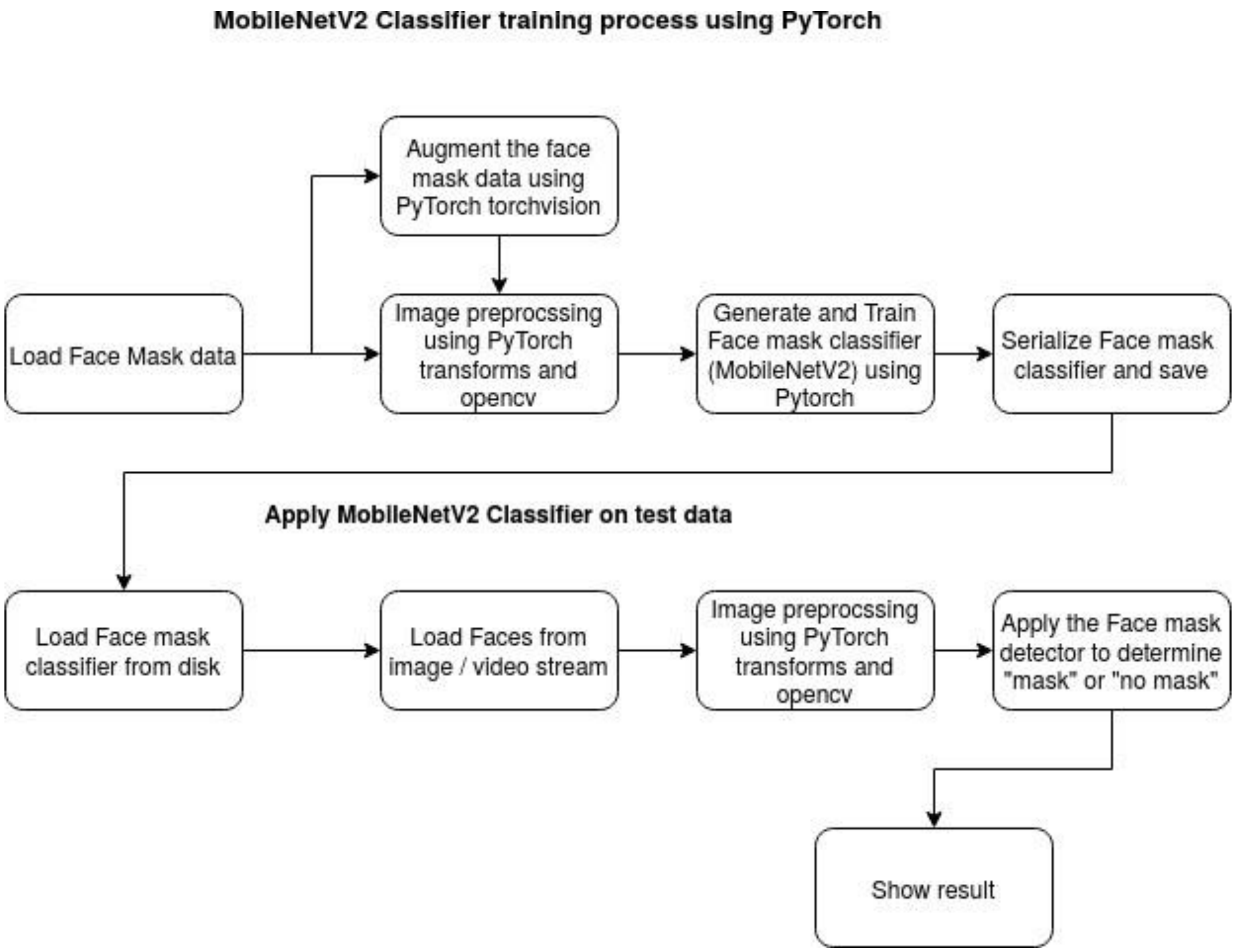


Convolutional Block of MobileNetV2

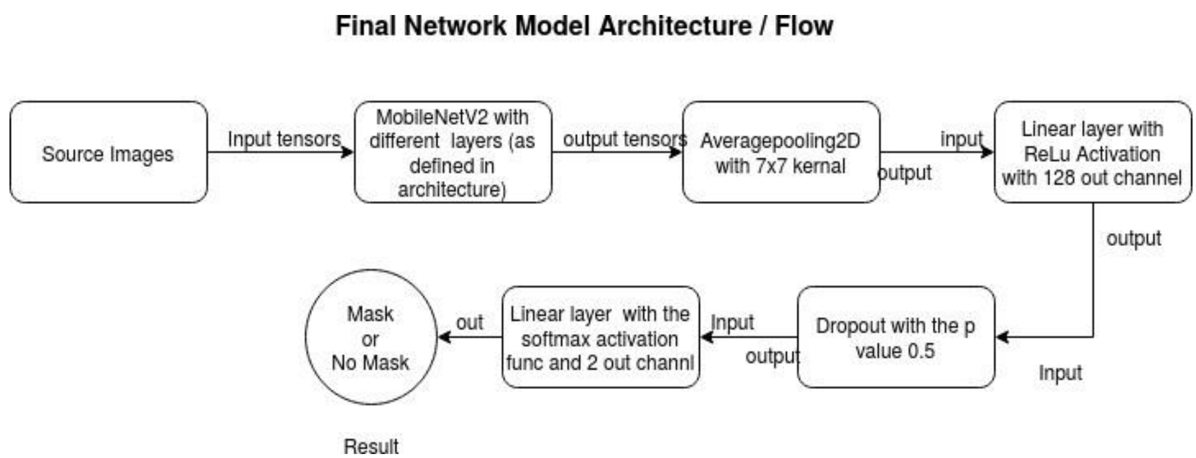
In the discussion of the layer part, there are three layers for both blocks. First one is 1x1 convolution with ReLU6. Depth wise convolution is in the second layer and again in the third layer there is a1x1 convolution but without any non-linearit

4.2 System Design

4.2.1 Flowchart Diagram

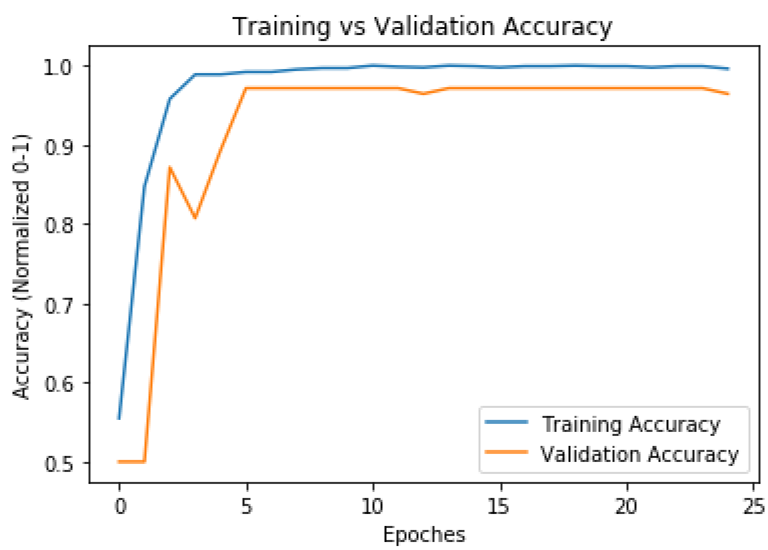


4.2.2 General Flow Design:



4.2.3 Accuracy Overview :

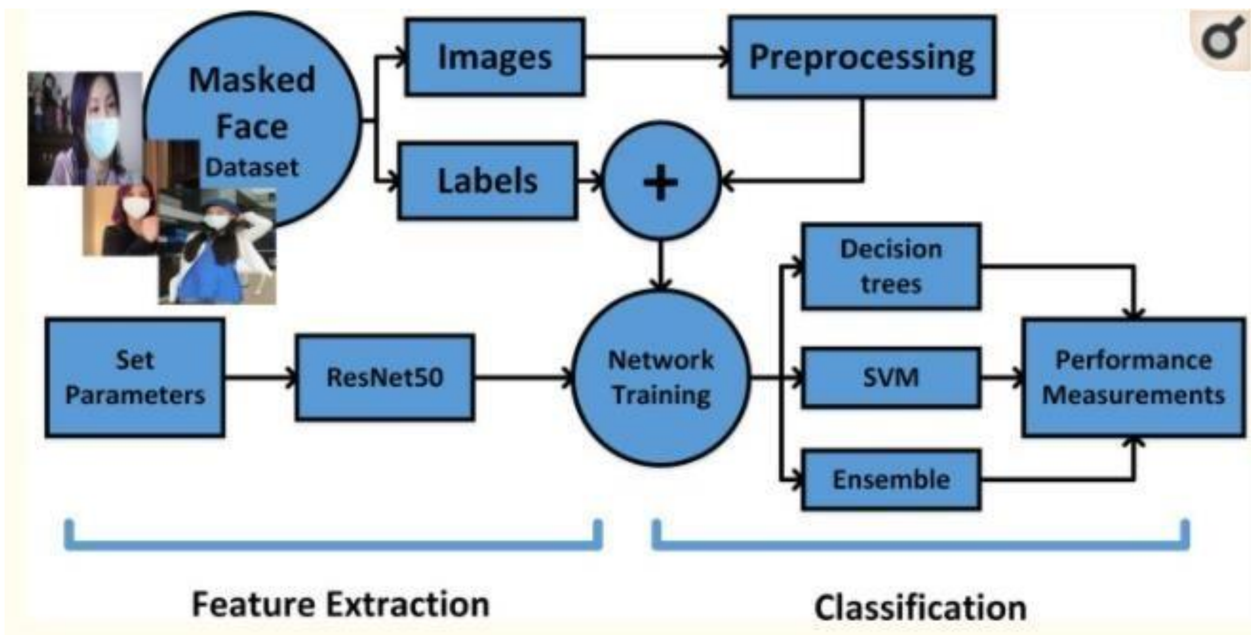
The data set has been divided into two sets, likely a training and validation set. The accuracy of image classifier over the training set vs validation.



Training vs Validation Accuracy

4.2.4 Procedural Design:

Two main components, the first component is deep transferring learning as feature extractor and the second component is a classical machine learning like decision trees, SVM, and ensemble. It has achieved better results when it is used as a feature extractor. Illustrates the proposed classical transfer learning model. Mainly used for the feature extraction phase while the traditional machine learning model used in the training, validation, and testing phase.



4.2.5 Test Sample :

For the reference I have attached a single image prediction result.

Test sample Image

5. Implementation:

The algorithm is trained on thousands of positive (with a face) and negative images (without a face), and uses Haar Features to calculate the difference between different regions of an image. These calculations are made by subtracting pixel values from different regions within a specified area.

First, make a new directory for the project files. Inside of the directory, let's make a virtual environment to download the necessary packages. If you do not have virtualenv you should run the first line of code, otherwise, skip the first line.

```
python3 -m pip install --user -U virtualenv
python3 -m virtualenv your_env
```

Inside of our virtual environment, lets download the necessary packages:

```
pip install tensorflow
pip install opencv-python
import warnings
warnings.filterwarnings('ignore')
import numpy as np
from keras.models import load model
```

5.1 Coding Details:

Data Collector.py

```
import cv2

video=cv2.VideoCapture(0)

facedetect=cv2.CascadeClassifier('haarcascade_frontalface_default.xml')

count=0

while True:
   ret,frame=video.read()
   faces=facedetect.detectMultiScale(frame,1.3, 5)
   for x,y,w,h in faces:
      count=count+1
      name='./images/face_without_mask/'+ str(count) + '.jpg'
      print("Creating Images........." +name)
      cv2.imwrite(name, frame[y:y+h,x:x+w])
      cv2.rectangle(frame, (x,y), (x+w, y+h), (0,255,0), 3)
   cv2.imshow("WindowFrame", frame)
   cv2.waitKey(1)
   if count>500:
      break
video.release()
cv2.destroyAllWindows()
```

Test.py

```
import warnings
warnings.filterwarnings('ignore')
import numpy as np
import cv2
from keras.models import load_model
facedetect = cv2.CascadeClassifier('haarcascade_frontalface_default.xml')
threshold=0.90
cap=cv2.VideoCapture(0)
cap.set(3, 640)
cap.set(4, 480)
font=cv2.FONT_HERSHEY_COMPLEX
model = load_model('MyTrainingModel.h5')

def preprocessing(img):
    img=img.astype("uint8")
    img=cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
    img=cv2.equalizeHist(img)
    img = img/255
    return img
```

```python
def get_className(classNo):
   if classNo==0:
      return "Mask"
   elif classNo==1:
      return "No Mask"


while True:
   sucess, imgOrignal=cap.read()
   faces = facedetect.detectMultiScale(imgOrignal,1.3,5)
   for x,y,w,h in faces:
      # cv2.rectangle(imgOrignal,(x,y),(x+w,y+h),(50,50,255),2)
      # cv2.rectangle(imgOrignal, (x,y-40),(x+w, y), (50,50,255),-2)
      crop_img=imgOrignal[y:y+h,x:x+h]
      img=cv2.resize(crop_img, (32,32))
      img=preprocessing(img)
      img=img.reshape(1, 32, 32, 1)
      # cv2.putText(imgOrignal, "Class" , (20,35), font, 0.75, (0,0,255),2,
cv2.LINE_AA)
      # cv2.putText(imgOrignal, "Probability" , (20,75), font, 0.75,
(255,0,255),2, cv2.LINE_AA)
      prediction=model.predict(img)
      classIndex=model.predict_classes(img)
      probabilityValue=np.amax(prediction)
      if probabilityValue>threshold:
         if classIndex==0:
            cv2.rectangle(imgOrignal,(x,y),(x+w,y+h),(0,255,0),2)
            cv2.rectangle(imgOrignal, (x,y-40),(x+w, y), (0,255,0),-2)
            cv2.putText(imgOrignal, str(get_className(classIndex)),(x,y-10),
font, 0.75, (255,255,255),1, cv2.LINE_AA)
         elif classIndex==1:
            cv2.rectangle(imgOrignal,(x,y),(x+w,y+h),(50,50,255),2)
            cv2.rectangle(imgOrignal, (x,y-40),(x+w, y), (50,50,255),-2)
            cv2.putText(imgOrignal, str(get_className(classIndex)),(x,y-10),
font, 0.75, (255,255,255),1, cv2.LINE_AA)

         # cv2.putText(imgOrignal,str(round(probabilityValue*100, 2))+"%"
,(180, 75), font, 0.75, (255,0,0),2, cv2.LINE_AA)
   cv2.imshow("Result",imgOrignal)
   k=cv2.waitKey(1)
   if k==ord('q'):
      break


cap.release()

cv2.destroyAllWindows()
```

5.1 Results and Discussions:

5.1.1 Result:

The model is trained, validated and tested upon two datasets. Corresponding to dataset 1, the method attains accuracy up to 95.77% (shown in fig. 7). Fig. 6 depicts how this optimized accuracy mitigates the cost of error. Dataset 2 is more versatile than dataset 1 as it has multiple faces in the frame and different types of masks having different colors as well. Therefore, the model attains an accuracy of 94.58% on dataset 2 as shown in Fig. 9. Fig. 8 depicts the contrast between training and validation loss corresponding to dataset 2. One of the main reasons behind achieving this accuracy lies in MaxPooling. It provides rudimentary translation invariance to the internal representation along with the reduction in the number of parameters the model has to learn. This sample-based discretization process down-samples the input representation consisting of image, by reducing its dimensionality. Number of neurons has the optimized value of 64 which is not too high. A much higher number of neurons and filters can lead to worse performance. The optimized filter values and pool size help to filter out the main portion (face) of the image to detect the existence of mask correctly without causing over-fitting. The system can efficiently detect partially occluded faces either with a mask or hair or hand. It considers the occlusion degree of four regions – nose, mouth, chin and eye to differentiate between annotated mask or face covered by hand. Therefore, a mask covering the face fully including nose and chin will only be treated as “with mask” by the model. The main challenges faced by the method mainly comprise of varying angles and lack of clarity. Indistinct moving faces in the video stream make it more difficult. However, following the trajectories of several frames of the video helps to create a better decision – “with mask” or “without mask”. This section will present the experiment results of the face mask detection in real-time application and that has already been installed at Politeknik Negeri Batam. The first experiment which is depicted on Fig. 4 has been done as the first trial before it is implemented for the moving person. Fig. 4(a) illustrates the face detector that detected the single user wearing a face mask accurately even it has some disturbance in the area. As for Fig. 4(b), the user was added slowly from below the camera and the detector was able to detect the mask properly. When the users are standing close to each other as seen on Fig. 4 (c)-(f), this system was also able to detect the face mask even if the user was surrounded by many objects with similar color.

After did the trial with no error, we are ready to verify this device with more user. As seen on Fig. 5(a)-(c), we added the user into three people with different types of face mask such as surgical and fabric face mask. Each person was standing in different position to verify the performance of face mask detection. From the picture it is verify that the face detected remains steady in detecting face mask of the users even the lighting was in different brightness. To make different in brightness, we turned off and on the lamp at our lab as to test the feature of this system.

On the other hand, the experiment of detecting a nonwearing mask is presented on Fig. 6(a)-(c). On Fig. 6(a) the first user who wear a white T-shirt attempted to pull off his mask, and the mask detector was able to distinguish the nonwearing mask and mask-wearing user. Fig. 6(b) also presented the mask detection precisely. Moreover, on Fig. 6(c) the first and third users tried to take off their mask, and the mask detector detected the face mask condition steadily.

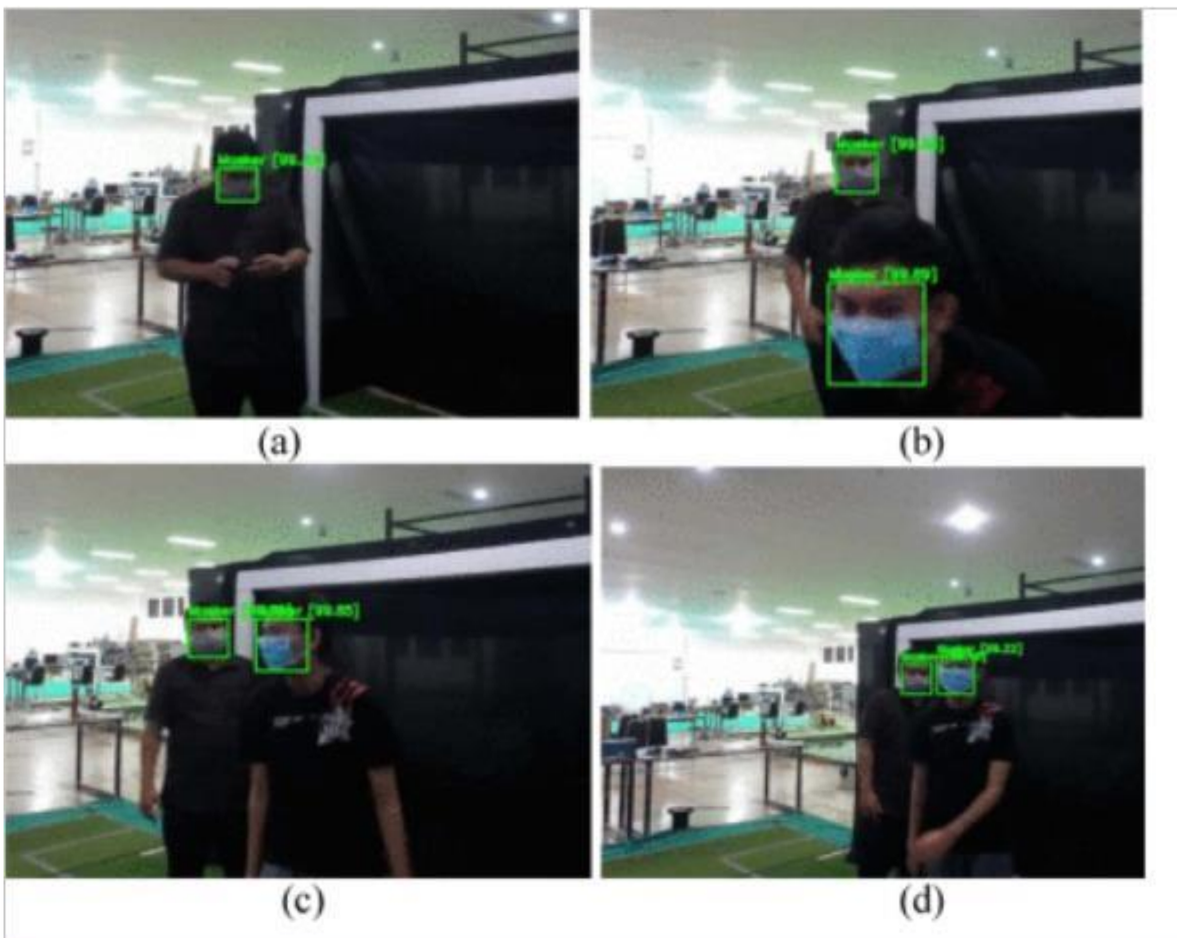


Fig. 4.
The face mask detector detected the user who wear the face mask (a) alone in the frame, (b) detected the new user, (c)-(f) detected the face mask where the user was close to each other.

5.1.2 Discussions:

The system works great, but it required some effort to overcome all these challenges!

Using pose estimation was definitely an interesting approach for this use case. As mentioned, it made sense by the fact that we have really small faces, hard to detect and even harder to track. We need tracking to classify people in the right moment, as they walk through the sequence of frames looking in different directions.

What's the requisite for a video stream to be correctly processed? Faces need to be visible at some point (we filter images below 16x16 pixels). And in this case we're using the body to estimate the head bounding box, but that's very easy to change for other cameras and use facial keypoints instead if they're visible, still using pose estimation.

If you have some particular video stream where you want to test this software, let us know! It's always great to have comments and feedback from readers.

Work on this kind of real world video analytics is far from over. There are several lines of work to tackle next, which can both improve the current solution in terms of accuracy, and also make it faster. Our review of the literature offers evidence in favor ofwidespread mask use as source control to reduce communitytransmission: non-medical masks use materials that obstructdroplets of the necessary size; people are most infectious inthe initial period post-infection, where it is common to havefew or no symptoms (10–16); non-medical masks have beeneffective in reducing transmission of influenza; and places andtime periods where mask usage is required or widespread haveshown substantially lower community transmission.The available evidence suggests that near-universal adop-tion of non-medical masks when out in public, in combination with complementary public health measures could successfullyreduce Re(effective-R) to below 1, thereby reducing commu-nity spread if such measures are sustained. Economic analysissuggests that the impact of mask wearing could be thousands of US dollars saved per person per mask (102).Interventions to reduce COVID-19 spread should be priori-tized in order of their expected multiple on effective R dividedby their cost. By this criterion, experimentation with anddeployment of universal masks look particularly promising.When used in conjunction with widespread testing, contacttracing, quarantining of anyone that may be infected, handwashing, and physical distancing, face masks are a valuabletool to reduce community transmission. All of these mea-sures, through their effect on Re, have the potential to reducethe number of infections. As governments talk about relax-ing lockdowns, keeping transmissions low enough to preservehealth care capacity will be critical until a vaccine can be de-veloped. Mask wearing may be instrumental in preventing asecond wave of infections from overwhelming the health caresystem – further research is urgently needed here.UNESCO states that "when human activities may lead tomorally unacceptable harm that is scientifically plausible butuncertain, actions shall be taken to avoid or diminish thatharm" (103). This is known as the "precautionary principle."The World Charter for Nature, which was adopted by the UNGeneral Assembly in 1982, was the first international endorse-ment of the precautionary principle. It was implemented in aninternational treaty in the 1987 Montreal Protocol. The lossof life and economic destruction that has been seen alreadyfrom COVID-19 is a "morally unacceptable harm." The pos-itive impact of public mask wearing on this is "scientificallyplausible but uncertain". This notion is reflected in Figure1 - while researchers may reasonably disagree on the magni-tude of transmissibility reduction and compliance, seemingly modest benefits can be massively beneficial in the aggregate due to the exponential character of the transmission process. Therefore, the action of ensuring widespread use of masks inthe community should be taken, based on this principle (104).Models suggest that public mask wearing is most effectiveat reducing spread of the virus when compliance is high (96).We recommend that mask use requirements are implementedby governments, or when governments do not, by organizations that provide public-facing services, such as transit seer-vice providers or stores, as "no mask, no service" rules. Such mandates must be accompanied by measures to ensure access to masks, possibly including distribution and rationing mechanisms so that they do not become discriminatory, but remain focused on the public health benefit. Given the value of the source control principle, especially for presymptomatic people, it is not good enough for only employees to wear masks, customers must wear masks as well. It is also important for health authorities to provide clear guidelines for the production, use and sanitization or re-use of face masks, and consider their distribution as shortages allow. A number of countries have distributed surgical masks (South Korea, Taiwan) from early on, while Japan, Singapore and Belgium are now distributing cloth masks to their entire populations. Clear and implementable guidelines can help increase compliance, and bring communities closer to the goal of reducing and ultimately stopping the spread of COVID-19.

6 Conclusion and Future Work:

6.1 Conclusion:

In this paper, we briefly explained the motivation of the work at first. Then, we illustrated the learning and performance task of the model. Using basic ML tools and simplified techniques the method has achieved reasonably high accuracy. It can be used for a variety of applications. Wearing a mask may be obligatory in the near future, considering the Covid-19 crisis. Many public service providers will ask the customers to wear masks correctly to avail of their services. The deployed model will contribute immensely to the public health care system. In future it can be extended to detect if a person is wearing the mask properly or not. The model can be further improved to detect if the mask is virus prone or not i.e. the type of the mask is surgical, N95 or not.A method is designed for checking the correct wearing of face protection mask from a video selfie. Different analysis scenarios have been experimented using diverse types of conventional mask and varied acquisition conditions. The performance of the designed method relies on the efficiency of the exploited face and face-feature detectors. In the present study, wearing glasses had no negative effect. The use of rigid masks seems preferable because they reduce possibilities of wrong positioning on the face. For this latter, the designed prototype can particularly be efficient. Hence, a promising application "CheckYourMask" has been proposed. A proof of concept as well as a development base are provided towards reducing the spread of COVID-19 by allowing people to validate the wearing of their masks via their smartphones (m-health). Moreover, this self-checking of the correct mask wearing could be exploited by monitoring-related applications as a conformity attribute. Corporate giants from various verticals are turning to AI and ML, leveraging technology at the service of humanity amid the pandemic. Digital product development companies are launching mask detection API services that enable developers to build a face mask detection system quickly to serve the community amid the crisis. The technology assures reliable and real-time face detection of users wearing masks. Besides, the system is easy to deploy into any existing system of a business while keeping the safety and privacy of users' data. So the face mask detection system is going to be the leading digital solution for most industries, especially retail, healthcare, and corporate sectors. Discover how we can help you to serve the communities with the help of digital solutions. Future works may investigate the Figure 6: Tests of some users with different types of mask. (a) Unvalid wearing. (b) Valid wearing. (c) Unvalid wearing. (d) Valid wearing. (e) Unvalid wearing. (f) Valid wearing. (g) Unvalid wearing. (h) Valid wearing CMES, 2020, vol.124, no.3 1057development of highly robust detectors by training a deep learning model with respect to specific face features or more globally to correctly/incorrectly worn mask classes. A new facemask-wearing condition identification method was proposed, which combines an SR network with a classification network (SRCNet) for facial image classification. To identify facemaskwearing condition, the input images were processed with image pre-processing, facial detection and cropping, SR, and facemask-wearing condition identification. Finally, SRCNet achieved a 98.70% accuracy and outperformed traditional end-to-end image classification methods by over 1.5% in kappa. Our findings indicate that the proposed SRCNet can achieve high accuracy in facemask-wearing condition identification, which is meaningful for the prevention of epidemic diseases including COVID-19 in public. There are a few limitations to our study. Firstly, the Medical Masks Dataset Sensors 2020, 20, 5236 19 of 23 we used for facemask-wearing condition identification is relatively small, where it cannot cover all postures or environments. In addition, the dataset does not contain video, where the identification result on a video stream cannot be tested. As for the proposed algorithm, the identification time for a single image is a little long, where an average of 10 images can be identified in a second, which does not meet the basic video frame rate of 24 frames per second (fps). In future studies, a more extensive facemask-wearing data set including images and videos will be collected and labelled with more details, in order to improve the performance of SRCNet. The data set shall contain faces with different postures, environments, and lighting conditions. In addition, SRCNet will be improved, based on either single image or video with IoT technologies, and a more efficient and accurate algorithm will be explored, which can contribute to the practical application of identifying facemask-wearing condition

6.2 Future Work:

There are a number of aspects we plan to work on shortly: Currently, the model gives 5 FPS inference speed on a CPU. In the future, we plan to improve this up to 15 FPS, making our solution deployable for CCTV cameras, without the need of a GPU. The use of Machine Learning in the field of mobile deployment is rising rapidly. Hence, we plan to port our models to their respective Tensor Flow Lite versions. Our architecture can be made compatible with Tensor Flow Runtime (TFRT), which will increase the inference performance on edge devices and make our models efficient on multithreading CPUs. Stage 1 and Stage 2 models can be easily replaced with improved models in the future that would give better accuracy and lower latency

7. Reference

1. W.H.O., “Coronavirus disease 2019 (covid-19): situation report, 205”. 2020

2. Faizaan Ahmad,AaimaNajam,Zeeshan Ahmed “Image-based Face Detection and Recognition” Department of Computer Science & Engineering, Beijing University of Aeronautics & Astronautics (2018)

3. Lahirudinalankara“FACE DETECTION & FACE RECOGNITION USING OPEN COMPUTER VISION CLASSIFIRES”, Robotic Visual Perception and Autonomy Faculty of Science and Engineering Plymouth University4August2017.

4. J. Manikandan, S. Lakshmi Prathyusha, P. Sai Kumar, Y. Jaya Chandra, M. Umaditya Hanuman, “Face Detection and Recognition using Open CV Based on Fisher Faces Algorithm”, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-5, January 2020

5. Shervin EMAMI1, Valentin Petruț SUCIU “Facial Recognition using OpenCV”, Department of Economic Informatics and Cybernetics Bucharest University of Economic Studies ROMANIA on 12 August 2015

6. “Coronavirus Disease 2019 (COVID-19) – Symptoms”, Centers for Disease Control and Prevention, 2020. [Online]. Available: https://www.cdc.gov/coronavirus/2019 ncov/symptomstesting/symptoms.html. 2020.

7. “Coronavirus — Human Coronavirus Types — CDC”, Cdc.gov, 2020. [Online]. Available: https://www.cdc.gov/coronavirus/types.html. 2020.

8. W.H.O., “Advice on the use of masks in the context of COVID-19: interim guidance”, 2020.

9. M. Jiang, X. Fan and H. Yan, “RetinaMask: A Face Mask detector”, arXiv.org, 2020. [Online]. Available: https://arxiv.org/abs/2005.03950. 2020.

10. B. Suvarnamukhi and M. Seshashayee, “Big Data Concepts and Techniques in Data Processing”, International Journal of Computer Sciences and Engineering, vol. 6, no. 10, pp. 712-714, 2018. Available: 10.26438/ijcse/v6i10.712714.

11. F. Hohman, M. Kahng, R. Pienta and D. H. Chau, “Visual Analytics in Deep Learning: An Interrogative Survey for the Next Frontiers,” in IEEE Transactions on Visualization and Computer Graphics, vol. 25, no. 8, pp. 2674-2693, 1 Aug. 2019, doi: 10.1109/TVCG.2018.2843369.

12. C. Kanan and G. Cottrell, “Color-to-Grayscale: Does the Method Matter in Image Recognition?” PLoS ONE, vol. 7, no. 1, p. e29740, 2012. Available: 10.1371/journal.pone.0029740.

13. Opencv-python-tutroals.readthedocs.io. 2020. C

14. M. Hashemi, “Enlarging smaller images before inputting into convolutional neural network: zero-padding vs. interpolation”, Journal of Big Data, vol. 6, no. 1, 2019. Available: 10.1186/s40537-019-0263-7 . 2020.

15. S. Ghosh, N. Das and M. Nasipuri, “Reshaping inputs for convolutional neural network: Some common and uncommon methods”, Pattern Recognition, vol. 93, pp. 79-94, 2019. Available: 10.1016/j.patcog.2019.04.009.

16. R. Yamashita, M. Nishio, R. Do and K. Togashi, “Convolutional neural networks: an overview and application in radiology”, Insights into Imaging, vol. 9, no. 4, pp. 611-629, 2018. Available: 10.1007/s13244- 018-0639-9. [13] “Guide to the Sequential model - Keras Documentation”, Faroit.com, 2020. [Online]. Available: https://faroit.com/keras-docs/1.0.1/gettingstarted/sequential-model-guide/. 2020.

17. Nwankpa, C., Ijomah, W., Gachagan, A. and Marshall, S., 2020. Activation Functions: Comparison Of Trends In Practice And Research For Deep Learning. [online] arXiv.org. Available at: [https://arxiv.org/abs/1811.03378. 2020](https://arxiv.org/abs/1811.03378.2020).

18. K. Team, “Keras documentation: MaxPooling2D layer”, Keras.io, 2020. [Online]. Available: https://keras.io/api/layers/pooling layers/ max pooling2d/. 2020.

19. “Face Mask Detection”, Kaggle.com, 2020. [Online]. Available: https://www.kaggle.com/andrewmvd/face-mask-detection. 2020. [18]

20. “TensorFlow White Papers”, TensorFlow, 2020. [Online]. Available: https://www.tensorflow.org/about/bib. 2020. [19] K. Team, “Keras documentation: About Keras”, Keras.io, 2020. [Online]. Available: https://keras.io/about. 2020.

21 “OpenCV”, Opencv.org, 2020. [Online]. Available: https://opencv.org/. 2020.

22. D. Meena and R. Sharan, “An approach to face detection and recognition,” 2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE), Jaipur, 2016, pp.

23. S. Ge, J. Li, Q. Ye and Z. Luo, “Detecting Masked Faces in the Wild with LLE-CNNs,” 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, 2017, pp. 426-434, doi: 10.1109/CVPR.2017.53.

24. Yang, J.; Wright, J.; Huang, T.S.; Ma, Y. Image super-resolution via sparse representation. IEEE Trans. Image Process. 2010, 19, 2861–2873. [CrossRef] [PubMed]

25. Timofte, R.; De Smet, V.; Van Gool, L. A+: Adjusted Anchored Neighborhood Regression for Fast Super-Resolution. In Computer Vision—ACCV 2014; Cremers, D., Reid, I., Saito, H., Yang, M.H., Eds.; Springer: Cham, Switzerland, 2015; Volume 9006, pp. 111–126.

26. Kim, J.; Lee, J.K.; Lee, K.M. Deeply-Recursive Convolutional Network for Image Super-Resolution. arXiv 2015, arXiv:1511.04491.

27. Dong, C.; Change Loy, C.; Tang, X. Accelerating the Super-Resolution Convolutional Neural Network. arXiv 2016, arXiv:1608.00367.

28. Johnson, J.; Alahi, A.; Fei-Fei, L. Perceptual Losses for Real-Time Style Transfer and Super-Resolution. In Computer Vision—ECCV 2016; Leibe, B.,

Matas, J., Sebe, N., Welling, M., Eds.; Springer: Cham, Switzerland, 2016; Volume 9906, pp. 694–711.

29. Kim, J.; Lee, J.K.; Lee, K.M. Accurate Image Super-Resolution Using Very Deep Convolutional Networks. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 1646–1654.

30. Shi, W.; Caballero, J.; Huszar, F.; Totz, J.; Aitken, A.P.; Bishop, R.; Rueckert, D.; Wang, Z. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 1874–1883.

31. Chen, Y.; Tai, Y.; Liu, X.; Shen, C.; Yang, J. FSRNet: End-to-End Learning Face Super-Resolution with Facial Priors. arXiv 2017, arXiv:1711.10703.

32. Lai, W.-S.; Huang, J.-B.; Ahuja, N.; Yang, M.-H. Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 5835–5843.

33. Lim, B.; Son, S.; Kim, H.; Nah, S.; Lee, K.M. Enhanced Deep Residual Networks for Single Image Super-Resolution. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017; pp. 1132–1140. Sensors 2020, 20, 5236 21 of 23

34. Dong, C.; Loy, C.C.; He, K.; Tang, X. Image Super-Resolution Using Deep Convolutional Networks. IEEE Trans. Pattern Anal. Mach. Intell. 2016, 38, 295–307. [CrossRef]

35. Dong, C.; Loy, C.C.; He, K.; Tang, X. Learning a Deep Convolutional Network for Image Super-Resolution. In Computer Vision—ECCV 2014; Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T., Eds.; Springer: Cham, Switzerland, 2014; Volume 8692, pp. 184–199.

36. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv 2014, arXiv:1409.1556.

37. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going Deeper with Convolutions. In Proceedings

of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.

38. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

39. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A.A. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning. arXiv 2016, arXiv:1602.07261.

40. Chollet, F. Xception: Deep Learning with Depthwise Separable Convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 21–26 July 2017; pp. 1800–1807.

41. Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. arXiv 2016, arXiv:1608.06993.

42. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv 2017, arXiv:1704.04861.

43. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. MobileNetV2: Inverted Residuals and Linear Bottlenecks. arXiv 2018, arXiv:1801.04381.

44. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 6848–6856.

45. Iandola, F.N.; Han, S.; Moskewicz, M.W.; Ashraf, K.; Dally, W.J.; Keutzer, K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters.

46. Hearst, M.A.; Dumais, S.T.; Osuna, E.; Platt, J.; Scholkopf, B. Support vector machines. IEEE Intell. Syst. Appl. 1998, 13, 18–28. [CrossRef]

47. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Li, F.F. ImageNet: A Large-Scale Hierarchical Image Database. In Proceedings of the Cvpr: 2009 IEEE Conference on Computer Vision and Pattern Recognition, Vols 1–4, Miami, FL, USA, 20–25 June 2009; IEEE: New York, NY, USA, 2009; pp. 248–255.

48. Krizhevsky, A.; Hinton, G. Learning Multiple Layers of Features from Tiny Images. 2009. Available online: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.222.9220&rep=rep1&type=pdf (accessed on 10 September 2020).

49. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A Simple Way to Prevent Neural Networks from Overfitting. J. Mach. Learn. Res. 2014, 15, 1929–1958.

50. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 6000–6010.

51. Smirnov, E.A.; Timoshenko, D.M.; Andrianov, S.N. Comparison of Regularization Methods for ImageNet Classification with Deep Convolutional Neural Networks. Aasri Procedia 2014, 6, 89–94. [CrossRef]

52. Lin, M.; Chen, Q.; Yan, S. Network In Network. arXiv 2013, arXiv:1312.4400.

53. Zhang, F.; Yang, F.; Li, C.; Yuan, G. CMNet: A Connect-and-Merge Convolutional Neural Network for Fast Vehicle Detection in Urban Traffic Surveillance. IEEE Access 2019, 7, 72660–72671. [CrossRef]

54. Hao, S.; Wang, W.; Ye, Y.; Li, E.; Bruzzone, L. A Deep Network Architecture for Super-Resolution-Aided Hyperspectral Image Classification With Classwise Loss. IEEE Trans. Geosci. Remote Sens. 2018, 56, 4650–4663. [CrossRef] Sensors 2020, 20, 5236 22 of 23

55. Lu, P.; Barazzetti, L.; Chandran, V.; Gavaghan, K.; Weber, S.; Gerber, N.; Reyes, M. Highly Accurate Facial Nerve Segmentation Refinement From CBCT/CT Imaging Using a Super-Resolution Classification Approach. IEEE Trans. Biomed. Eng. 2018, 65, 178–188. [CrossRef]

56. Na, B.; Fox, G.C. Object Detection by a Super-Resolution Method and a Convolutional Neural Networks. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 2263–2269.

57. Cai, D.; Chen, K.; Qian, Y.; Kämäräinen, J.-K. Convolutional low-resolution fine-grained classification. Pattern Recognit. Lett. 2019, 119, 166–171. [CrossRef]

58. Zou, W.W.; Yuen, P.C. Very low resolution face recognition problem. IEEE Trans. Image Process. 2012, 21, 327–340. [CrossRef]

59. Uiboupin, T.; Rasti, P.; Anbarjafari, G.; Demirel, H. Facial image super resolution using sparse representation for improving face recognition in surveillance monitoring. In Proceedings of the 24th Signal Processing and Communication Application Conference (SIU), Zonguldak, Turkey, 16–19 May 2016; pp. 437–440.

60. Jing, Z.; Mariani, R. Glasses detection and extraction by deformable contour. In Proceedings: Pattern Recognition and Neural Networks, Proceedings of the 15th International Conference on Pattern Recognition, Barcelona, Spain, 3–7 September 2000; Sanfeliu, A., Villanueva, J.J., Vanrell, M., Alquezar, R., Jain, A.K., Kittler, J., Eds.; IEEE: Piscataway, NJ, USA, 2000; Volume 2, pp. 933–936.

61. Fernández, A.; Casado, R.; Usamentiaga, R. A Real-Time Big Data Architecture for Glasses Detection Using Computer Vision Techniques. In Proceedings of the 3rd International Conference on Future Internet of Things and Cloud, Rome, Italy, 24–26 August 2015; pp. 591–596.

62. Fernández, A.; García, R.; Usamentiaga, R.; Casado, R. Glasses detection on real images based on robust alignment. Mach. Vis. Appl. 2015, 26, 519–531. [CrossRef]

63. Du, S.; Liu, J.; Liu, Y.; Zhang, X.; Xue, J. Precise glasses detection algorithm for face with in-plane rotation. Multimed. Syst. 2015, 23, 293–302. [CrossRef]

64. Shao, L.; Zhu, R.; Zhao, Q. Glasses Detection Using Convolutional Neural Networks. In Biometric Recognition; You, Z., Zhou, J., Wang, Y., Sun, Z., Shan, S., Zheng, W., Feng, J., Zhao, Q., Eds.; Springer International Publishing: Cham, Switzerland, 2016; pp. 711–719.

65. Xie, Z.; Liu, H.; Li, Z.; He, Y. A convolutional neural network based approach towards real-time hard hat detection. In Proceedings of the IEEE International Conference on Progress in Informatics and Computing (PIC), Suzhou, China, 14–16 December 2018; pp. 430–434.

66. Finlayson, G.; Hordley, S.; Schaefer, G.; Yun Tian, G. Illuminant and device invariant colour using histogram equalisation. Pattern Recognit. 2005, 38, 179–190. [CrossRef]