JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

FACE RECOGNITION TECHNOLOGY AND FACE MASK DETECTION USING DEEP LEARNING MODEL

[1] BAGGU MANASA
[1] M.Tech Student, CST with CN, Department of Computer Science and Systems Engineering
Andhra University College of Engineering, Visakhapatnam, AP, India

ABSTRACT

Face recognition technology may be a biometric technology that is predicated on the identification of face expression of a person. Folks collect the face images, and therefore the recognition instrumentality mechanically processes the images. The paper describes the event stages and the connected technologies of face recognition and social distancing. We have a tendency to introduce the analysis of face recognition for real conditions, and that we introduce the overall analysis standards and the general databases of face recognition. We have a tendency to provide an innovative read of face recognition and therefore the projected technique detects the face from the image properly then identifies if it's a mask thereon or not. As a police investigation task performer, it can even find a face in conjunction with a mask in motion. The tactic attains accuracy up to 95.77% and 94.58% severally on 2 completely different datasets. We have a tendency to explore optimized values of parameters by exploiting the successive Convolutional Neural Network model to detect the presence of masks properly while not inflicting over-fitting. The paper presents a technique for social distancing detection using deep learning to judge the gap between folks to mitigate the impact of this coronavirus pandemic. The detection tool was developed to alert people to take care of a secure distance with every other by evaluating a video feed. The video frame from the camera was used as input, and the ASCII text file object detection pre-trained model supported the YOLO v3 algorithmic program was used for pedestrian detection. Later, the video frame was remodeled into a top-down read for distance measuring from the second plane. The gap between people will be calculable and any noncompliant combination of individuals within the show will be indicated with a red frame and red line. The projected technique was valid on a pre-recorded video of pedestrians walking on the street. The result shows that the proposed method is in a position to determine the social distancing measures between multiple folks within the video and if the mask isn't detected in the video frame, with the face recognition and trained model we have a tendency to determine the person's name. The developed technique will be additional developed as a detection tool in real time application

.

Keywords: Convolutional Neural Networks, RNN, YOLO v3, Face recognition, Face Mask Detection

1. INTRODUCTION

Face recognition is a section problematic of visual pattern recognition. Humans are recognizing visual patterns all the time, and we obtain visual information through our eyes. This material is recognized by the brain as expressive concepts. For a computer, whether it is a image or a video, it is a matrix of several pixels. The machine must find out what impression a certain part of the data signifies in the data. This is a rough classification problem in visual prototypical recognition. For face recognition, it is compulsory to distinguish who the face belongs to in the part of the data that all machines think of the face. This is a subdivision problem. Face recognition in a comprehensive sense includes linked technologies for building a face recognition system. It contains face detection, face position, identity recognition, image pre-processing, etc. Face detection procedure is to find out the coordinate system of all faces in one image [1]. This is the progression of scanning the entire image to determine whether the candidate part is a face. The output of the face synchronize system can be square, rectangular, etc. The face position is the coordinate situation of the face feature in the face detection coordinate scheme. The deep learning framework fundamentally implements some present good positioning technologies. Compared with face detection, the calculation time, the face positioning algorithm is much smaller. The associate editor synchronizing the analysis of this manuscript and approving it for journal was Thomas CanhaoXu. The principle of chess playing and face recognition is to find suitable transform function [2]. Although their integrity are the same, the complication of face recognition transformation is far greater than the difficulty of finding the optimal

solution in the chessboard. We expect to find the ideal conversion function so as to achieve the optimal recognition effect, but the search process is very tough. From the presentation layout of face recognition technology, it is most widely used in attendance, access control, security and finance, while logistics, retail, smartphone, transportation, education, real estate, government management, entertainment advertising, network information security and other fields are starting to get involved. In the field of security, both the early warning of suspicious situations and the trace of suspects can be finished with the support of face recognition. It represents a great progress of artificial intelligence technology, which means that we require more accurate, more flexible and faster recognition technology [3].

This paper will describe the development stages and related technologies of face recognition, including early algorithms, artificial features and classifiers, deep learning and According to the World Health Organization (WHO)'s official Position Description – 205, coronavirus disease 2019 (COVID-19) has globally diseased over 20 million people causing over 0.7million deaths. Individuals with COVID- 19 have had a wide probability of indicators reported – going from mellow manifestations to serious illness. Respiratory complications like shortness of breath or difficulty in breathing is one of them. Elder people having lung disease can possess serious complications from COVID-19 illness as they appear to be at higher risk. Some common human coronaviruses that infect the public around the world are 229E, HKU1, OC43, and NL63. Before debilitating individuals, viruses like 2019-nCoV, SARS-CoV, and MERS-CoV infect animals and evolve to human coronaviruses [4-5]. Persons having respiratory problems can expose anyone (who is in close contact with them) to infective beads. Surroundings of a tainted individual can cause contact transmission as droplets carrying virus may arrive on his adjacent surfaces. To curb certain respiratory viral ailments, including COVID-19, wearing a clinical mask is very necessary. The public should be aware of whether to put on the mask for source control or aversion of COVID-19. Potential points of interest of the utilization of masks lie in reducing vulnerability of risk from a noxious individual during the "pre-symptomatic" period and stigmatization of discrete persons putting on masks to restrain the spread of virus [6]. WHO stresses on prioritizing medical masks and respirators for health care assistants. Therefore, face mask detection has become a crucial task in present global society. Face mask detection involves detecting the location of the face and then determining whether it has a mask on it or not. The issue is approximately cognate to general object detection to detect the classes of objects. Face identification categorically deals with distinguishing a specific group of entities i.e. face. It has numerous applications, such as autonomous driving, education, surveillance, and so on [7]

The World Health Organization (WHO) has declared Covid-19 as a pandemic due to the increase in the number of cases reported around the world. To contain the pandemic, many countries have implemented a lockdown where the government forces the citizens to stay at home during this critical period. The public health bodies such as the Centers for Disease Control and Prevention (CDC) had to make it clear that the most effective way to slow down the spread of Covid-19 is by avoiding close contact with other people. To flatten the curve on the Covid-19 pandemic, the citizens around the world are practicing physical distancing [8-9]. To implement social distancing, group activities and congregations such as travel, meetings, gatherings, workshops, and praying had been banned during the quarantine period. The people are encouraged to use phone and email to manage and conduct events as much as possible to minimize the person-to-person contact. To further contain the spread of the virus, people are also informed to perform hygiene measures such as frequently washing hands, wearing masks and avoiding close contact with people who are ill. However, there is a difference between knowing what to do to reduce the transmission of the virus and putting them into practice. The world has not yet fully recovered from this pandemic and the vaccine that can effectively treat Covid-19 is yet to be discovered. However, to reduce the impact of the pandemic on the country's economy, several governments have allowed a limited number of economic activities to be resumed once the number of new cases of Covid-10 has dropped below a certain level [10]. As these countries cautiously restarting their economic activities, concerns have emerged regarding workplace safety in the new post-Covid-19 environment. To reduce the possibility of infection, it is advised that people should avoid any person-to-person contact such as shaking hands and they should maintain a distance of at least 1 meter from each other. In Malaysia, the Ministry of Health Malaysia (MOHM) has recommended several disease prevention measures for workplaces, individuals, and families at home, schools, childcare centers, and senior living facilities. These measures include implementing social distancing measures, increasing physical space between workers at the workplace, staggering work schedules, decreasing social contacts in the workplace, limiting large work-related gatherings, limiting non-essential work travel, performing regular health checks of staff and visitors entering buildings, reducing physical activities especially for organizations that have staff in the high-risk category, and conducting company events or activities online. Individuals, communities, businesses healthcare organizations are all part of a community with their responsibility to mitigate the spread of the Covid-19 disease [11]. In reducing the impact of this coronavirus pandemic, practicing social distancing and self-isolation have been deemed as the most effective ways to break the chain of infections after restarting the economic activities. In fact, it has been observed that there are many people who are ignoring public health measures, especially with respect to social distancing. It is understandable that given the people's excitement to start working again, they sometimes tend to forget or neglect the implementation of social distancing. Hence, this work aims to facilitate the enforcement of social distancing by providing automated detection of social distance violation in workplaces and public areas using a deep learning model. In area of machine learning and computer vision, there are different methods that can be used for object detection. These methods can also applied to detect the social distance between people [12].

Evaluating the similarity of images and their descriptors by employing discriminative learners has proven itself to be an effective face recognition paradigm. In this paper we show how "background samples", that is, examples which do not belong to any of the classes being learned, may provide a significant performance boost to such face recognition systems. In particular, we make the following contributions. First, we define and evaluate the "Two-Shot Similarity" (TSS) score as an extension to the recently proposed "One-Shot Similarity" (OSS) measure.

Recent face recognition experiments on the LFW benchmark show that face recognition is performing stunningly well, surpassing human recognition rates. In this paper, we study face recognition at scale. Specifically, we have collected from Flickr a \textbf{Million} faces and evaluated state of the art face recognition algorithms on this dataset. We found that the performance of algorithms varies--while all perform great on LFW, once evaluated at scale recognition rates drop drastically for most algorithms.

Deeper neural networks are more difficult to train. We present a residual learning framework to ease the training of networks that are substantially deeper than those used previously. We explicitly reformulate the layers as learning residual functions with reference to the layer inputs, instead of learning unreferenced functions. We provide comprehensive empirical evidence showing that these residual networks are easier to optimize, and can gain accuracy from considerably increased depth.

In this work we investigate the effect of the convolutional network depth on its accuracy in the large-scale image recognition setting. Our main contribution is a thorough evaluation of networks of increasing depth using an architecture with very small (3x3) convolution filters, which shows that a significant improvement on the prior-art configurations can be achieved by pushing the depth to 16-19 weight layers. These findings were the basis of our ImageNet Challenge 2014 submission, where our team secured the first and the second places in the localisation and classification tracks respectively.

The current superresolution (SR) methods based on deep learning have shown remarkable comparative advantages but remain unsatisfactory in recovering the high-frequency edge details of the images in noise-contaminated imaging conditions, e.g., remote sensing satellite imaging. In this paper, we propose a generative adversarial network (GAN)-based edge-enhancement network (EEGAN) for robust satellite image SR reconstruction along with the adversarial learning strategy that is insensitive to noise. In particular, EEGAN consists of two main subnetworks an ultradense subnetwork (UDSN) and an edge-enhancement subnetwork (EESN). In UDSN, a group of 2-D dense blocks is assembled for feature extraction and to obtain an intermediate high-resolution result that looks sharp but is eroded with artifacts and noises as previous GAN-based methods do.

Although tremendous strides have been recently made in face hallucination, existing methods based on a single deep learning framework can hardly satisfactorily provide fine facial features from tiny faces under complex degradation. This article advocates an adaptive-threshold-based multimodal fusion network (ATMFN) for compressed face hallucination, which unifies different deep learning models to take advantage of their respective learning merits. First of all, we construct CNN-, GAN- and RNN-based underlying super-resolvers to produce candidate SR results. Further, the attention subnetwork is proposed to learn the individual fusion weight matrices capturing the most informative components of the candidate SR faces.

A probe face image may contain noise due to environmental conditions, incorrect use of sensors or transmission error. The performance of face recognition severely depletes when the probe image is contaminated with noise. Denoising techniques can improve recognition performance, provided the correct parameters are used. In this paper, a parameter selection framework is presented. In the proposed framework, the optimal parameter set is selected for denoising using quality assessment algorithms with low complexity. Quality score based parameter selection is evaluated on the AR face dataset.

2. EXISTING SYSTEM

In the 1950s, people began to study how to make machines recognize faces. In 1964, the applied research of face recognition engineering officially began, mainly using face geometry for recognition. But it has not been applied in practice. Principal component analysis (PCA) is the most widely used data dimensionality reduction algorithm. In face recognition algorithms, PCA implements feature face extraction. In 1991, Turk and Pentland of MIT Media Laboratory introduced principal component analyses into face recognition. PCA is usually used to preprocess the data before other analyses. In the face of data with more dimensions, it can remove redundant information and noise, retain the essential characteristics of data, greatly reduce the dimensions, improve the processing speed of data, and save a lot of time and cost. Therefore, this algorithm is usually used for dimensionality reduction and the multi-dimensional data visualization. In PCA based feature extraction algorithms, the Eigen face is one of the classical algorithms. A simple process of feature extraction where PCA is combined with face recognition by using K-Nearest-Neighbor (KNN) algorithm. We get the eigenvalues and the eigenvectors of the covariance matrix from sampling data, and select the principal component, which is the eigenvector with the largest eigenvalue. At the same time, the feature matrix of the testing data is obtained by the same dimensionality reduction process. Finally, the face image category of

the testing set is detected by the KNN classifier. Although PCA is efficient in dealing with large data sets. Its biggest drawback is that its training data set must be large enough. For example, the number of original photos in the face recognition system must be at least thousands, so the results of principal component analysis are meaningful. However, when the persons' facial expressions are different, there are obstacles blocking the face, or the light is too strong or too weak, and it is difficult to get good low dimensional data.

Disadvantages

- In some cases it is not sufficient for the security of the dataManual Approach
- Requires More Equipment's
- More Expensive
- Lack of user satisfaction
- Less Efficient
- Less Accurate

3. PROPOSED SYSTEM

The proposed method consists of a cascade classifier and a pre-trained CNN which contains two 2D convolution layers connected to layers of dense neurons. A. Data Processing 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 relationships between different entities. The proposed method deals with image and video data using Numpy and OpenCV

ADVANTAGES

- Empirically, fine-tuning from pre-trained models has at least two advantages. First, there are numerous state of- the-art pre-trained CNN models publicly available.
- Our proposed CNN is built on framework and thus it inherits the speed and accuracy advantages, while produces more compact and flexible models.
- Our proposed approach has very appealing advantages over existing solutions. We will elaborate and validate the method in the following sections.
- We present CNN, to the best of our knowledge, world first framework that can train object detection networks from scratch with state-of-the-art performance.
- We introduce and validate a set of principles to design efficient object detection networks from scratch through step-by-step ablation studies in face mask detection.
- We show that our method can achieve state-of-the-art performance on three standard benchmarks (MS COCO datasets) with real time processing speed and more compact models.

COMMON EVALUATION CRITERIA OF FACE RECOGNITION

Accuracy (ACC), Receiver Operating Characteristic (ROC) curve and Area Under Curve (AUC) value are important indexes to evaluate the performance of the face recognition algorithm. In face recognition tasks, ACC is a common index. Assuming that the testing set contains N images and the number of correctly recognized images is

M. The definition of ACC is given as follows

$$ACC = M/N$$

The higher the ACC value is, the better the algorithm performance is. In the face recognition task, in order to determine whether two images (also known as sample pairs) come from the same person, ROC first calculates the distance measurement or the similarity between images, and then completes the recognition according to the threshold. The abscissa of the ROC curve

represents false positive rate (FPR), and the ordinate represents recall rate or true positive rate (TPR). The definitions of FPR and TPR are given as follows

$$TPR = TP/(TP + FN)$$
$$FPR = FP/(FP + TN)$$

TP refers to the positive sample pair correctly predicted by the model, FN refers to the positive sample pair wrongly predicted by the model, TN refers to the negative sample pair correctly predicted by the model, and FP refers to the negative sample pair wrongly predicted by the model. By changing different thresholds, different TPR values and FPR values can be obtained, and ROC curves can be generated (https://blog.csdn.net/). As is shown in Fig. 1, the red curve and blue curve respectively represent the TPR − FPR curve of two different classifiers, and the point on the curve corresponds to a threshold value, which is the ROC curve. The closer the ROC curve is to the upper left corner, the better the performance of the algorithm is. In other words, it can achieve

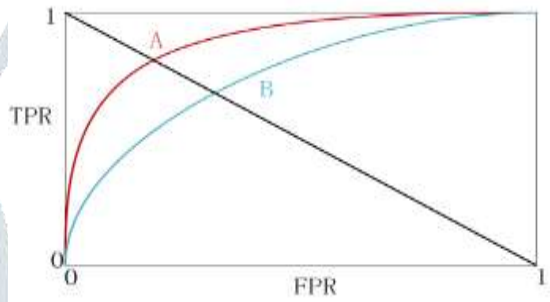


Figure 1. TPR − FPR curve of two different classifiers

4. IMPLEMENTATION

Face recognition technology is a biometric technology, which is based on the identification of facial features of a person. People collect the face images, and the recognition equipment automatically processes the images. The paper introduces the related research of face recognition from different perspectives. The paper describes the development stages and the related technologies of face recognition. We introduce the research of face recognition for real conditions, and we introduce the general evaluation standards and the general databases of face recognition. The proposed method consists of a cascade classifier and a pre-trained CNN which contains two 2D convolution layers connected to layers of dense neurons. This paper presents a simplified approach to achieve this purpose using some basic Machine Learning packages like TensorFlow, Keras, OpenCV and Scikit-Learn. The proposed method detects the face from the image correctly and then identifies if it has a mask on it or not. As a surveillance task performer, it can also detect a face along with a mask in motion. The method attains accuracy up to 95.77% and 94.58% respectively on two different datasets safe distance with each other by evaluating a video feed. The video frame from the camera was used as input, and the open-source object detection pre-trained model based on the YOLOv3 algorithm was employed for pedestrian detection. Later, the video frame was transformed into a top-down view for distance measurement from the 2D plane. The distance between people can be estimated and any noncompliant pair of people in the display will be indicated with a red frame and red line. The proposed method was validated on a pre-recorded video of pedestrians walking on the street. The result shows that the proposed method is able to determine the social distancing measures between multiple people in the video.

DEEP LEARNING

Deep learning is a branch of machine learning. Deep learning can find out the features needed for classification automatically in the training process without feature extraction steps. That is to force network learning to obtain more effective features for distinguishing different faces. The fifield of face recognition has been completely transformed by deep learning. Deep learning is widely used in face recognition and is divided into the following aspects.A face recognition method based on convolutional neural networks (CNN) is the first aspect. CNN uses the locality of data and other features to optimize the model structure by combining local perception areas, shared weights, and down-sampling of face images. CNN is very similar to ordinary neural networks. They consist of neurons with learnable weights and bias values. A dot product calculation for each neuron is performed after receiving input data.

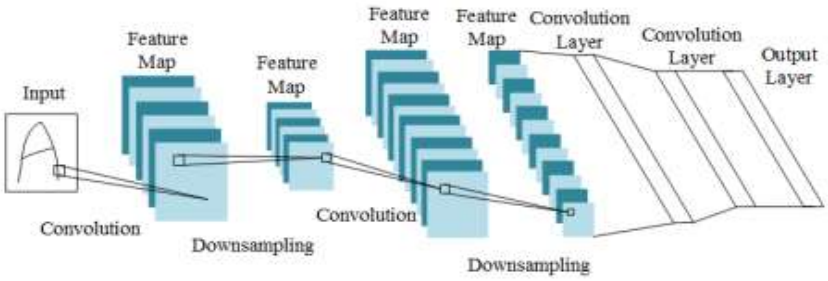


Figure 2. (Color online) The structure of CNN

CNN is composed of input layer, convolution layer, pooling layer (lower sampling layer), full connection layer and output layer. And the convolution layer and the pooling layer are alternately set. Deep nonlinear face shape extraction method is the second aspect. Face shape extraction or face alignment plays a very important role in tasks such as face recognition, expression recognition, and face animation synthesis. The difficulty in face recognition lies in the high complexity of face shape and texture. In order to further improve the nonlinear regression ability of the algorithm to obtain robustness to changes such as shape, Zhang et al. proposed a deep nonlinear face shape extraction method from course to fine (coarse-to-fine auto- encoder networks, CFAN). Face recognition based on deep learning video surveillance is the third aspect. In an intelligent monitoring environment, the identification of suspicious characters is an important use of face recognition. Recognizing the identity of people in video accurately and quickly is very important for video search and video surveillance. Schofifield et al. proposed a deep convolutional neural network method, which could automatically detect, track and record human faces in video, and could be used to study animal behavior. Low-resolution face recognition based on deep learning is the fourth aspect. In practical applications, the collected face images have a variety of posture changes, and the image resolution is low, causing the face image recognition performance to decline rapidly. In, the low-resolution face data set was studied, the most advanced supervised discriminant learning method was adopted, and the generative confrontation network pre-training method and full convolution structure were introduced to improve the low-resolution face recognition effect. Many deep learning models focus on the optimization of training methods and processes. However, the accuracy of low-resolution face recognition is constantly improved, and the running time is also reduced accordingly, so that it can be better put into practical applications. With the development of more comprehensive deep learning models, there are not only deep models that can adapt to large-scale data, but also processing methods that can adapt to the small data set in some specific scenarios. One method is to use synthetic data, the other one is to use the current.

YOLO

One of these algorithms is “YOLOv3”. “You Only Look Once”, version 3 (YOLOv3) is a real-time object detection system that recognizes specific things in videos, live feeds, and images Artificial Intelligence (AI) is utilized in object categorization systems programs to perceive certain objects in a class as subjects of interest. Objects in images are categorized into groups, with comparable qualities being grouped together and others being ignored until otherwise told.

DATASETS

Two datasets have been used for experimenting the current method. Dataset 1 consists of 1376 images with 690 images with people wearing face masks and the remaining 686 images with people who do not wear face masks. Fig. 3 mostly contains front face pose with single face in the frame and with same type of mask having white color only

Figure 3. Samples from Dataset 1 including faces without masks and with masks

Dataset 2 from Kaggle consists of 853 images and its contents are clarified either with a mask or without a mask. In Fig. 4 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.

Figure 4. Samples from Dataset 2 including faces without masks and with masks

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'. The statement categories=os.listdir(data path) categorizes the list of directories in the specified data path. The variable categories now look like ['with mask', 'without mask'] Then to find the number of labels, we need to distinguish those categories using labels=[i for i in range(len(categories))]. It sets the labels as [0, 1] Now, each category is mapped to its respective label using label dict=dict(zip(categories,labels)) which at first returns an iterator of tuples in the form of zip object where the items in each passed iterator is paired together consequently. The mapped variable label dict looks like {'with mask' 0, 'without mask' 1}

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 color to-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.

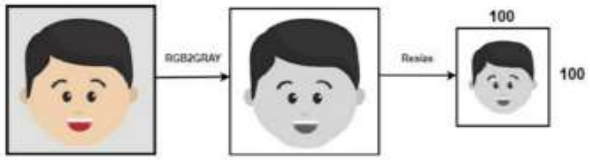


Figure 5. Conversion of a RGB image to a GrayScale image of 100x100 size

We use the function cv2.cvtColor(input image, fl flag) for changing the color space. Here the flag determines the type of conversion. In this case, the flag cv2.COLOR BGR2GRAY is used for gray conversion. Deep CNNs require a fixed-size input image. Therefore we need a fixed common size for all the images in the dataset. Using cv2.resize() the gray scale image is resized into 100 x 100.

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 models. 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.

Convolutional Neural Networks (CNNs)

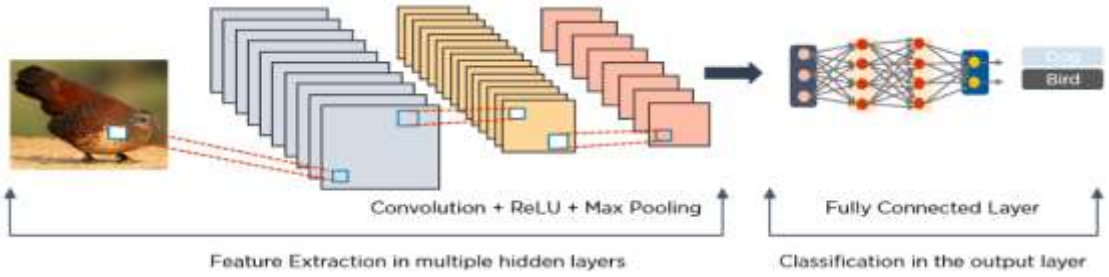
CNN's, also known as ConvNets, consist of multiple layers and are mainly used for image processing and object detection. Yann LeCun developed the first CNN in 1988 when it was called LeNet. It was used for recognizing characters like ZIP codes and digits.CNN's are widely used to identify satellite images, process medical images, forecast time series, and detect anomalies.

- **How Do CNNs Work?**

CNN's have multiple layers that process and extract features from data: Convolution Layer:CNN has a convolution layer that has several filters to perform the convolution operation.Rectified Linear Unit (ReLU):CNN's have a ReLU layer to perform operations on elements. The output is a rectified feature map. Pooling Layer:The rectified feature map next feeds into a pooling layer. Pooling is a down-sampling operation that reduces the dimensions of the feature map. The pooling layer then converts the resulting two-dimensional arrays from the pooled feature map into a single, long, continuous, linear vector by flattening it.

- **Fully Connected Layer**

A fully connected layer forms when the flattened matrix from the pooling layer is fed as an input, which classifies and identifies the images.Below is an example of an image processed via CNN.



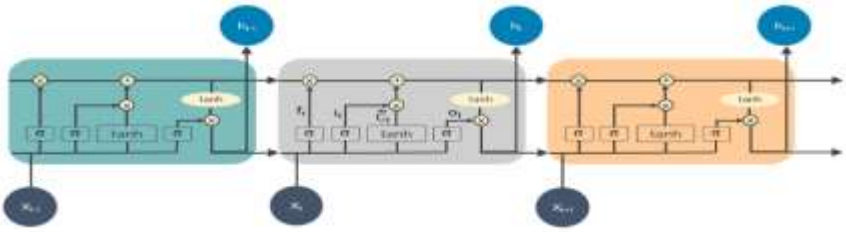
Fully Connected Layer

Long Short Term Memory Networks (LSTMs)

LSTMs are a type of Recurrent Neural Network (RNN) that can learn and memorize long-term dependencies. Recalling past information for long periods is the default behavior. LSTMs retain information over time. They are useful in time-series prediction because they remember previous inputs. LSTMs have a chain-like structure where four interacting layers communicate in a unique way. Besides time-series predictions, LSTMs are typically used for speech recognition, music composition, and pharmaceutical development.

- **How Do LSTMs Work?**

First, they forget irrelevant parts of the previous state Next, they selectively update the cell-state values Finally, the output of certain parts of the cell state Below is a diagram of how LSTMs operate:



LSTMs operate

Recurrent Neural Networks (RNNs)

RNNs have connections that form directed cycles, which allow the outputs from the LSTM to be fed as inputs to the current phase. The output from the LSTM becomes an input to the current phase and can memorize previous inputs due to its internal memory. RNNs are commonly used for image captioning, time-series analysis, natural-language processing, handwriting recognition, and machine translation.

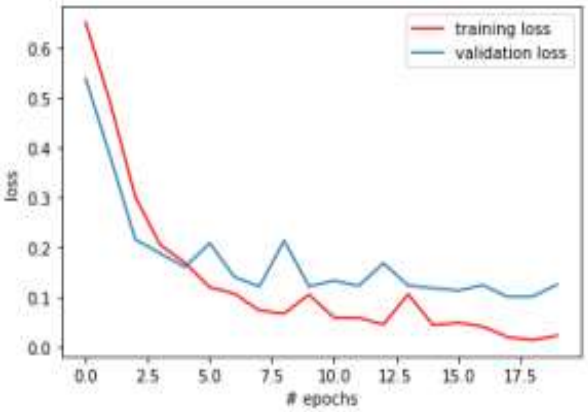
An unfolded RNN looks like this:

Unfolded RNN

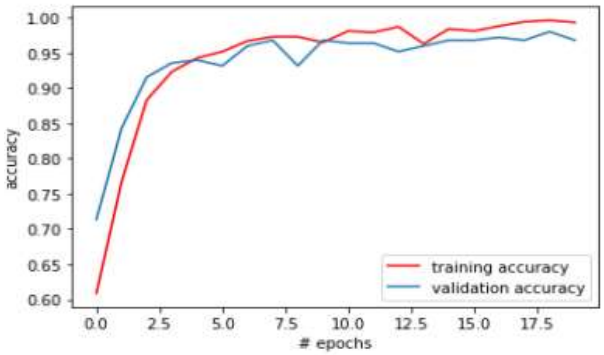
5. RESULT AND ANALYSIS

The paper describes the development stages and the related technologies of face recognition. We introduce the research of face recognition for real conditions, and we introduce the general evaluation standards and the general databases of face recognition. Datasets. Corresponding to dataset 1, the method attains accuracy up to 95.77% 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 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.

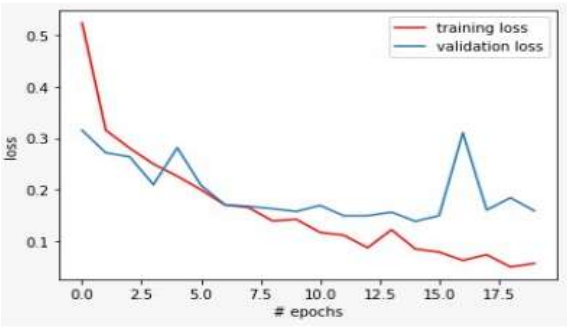


Epochs vs loss corresponding to dataset 1

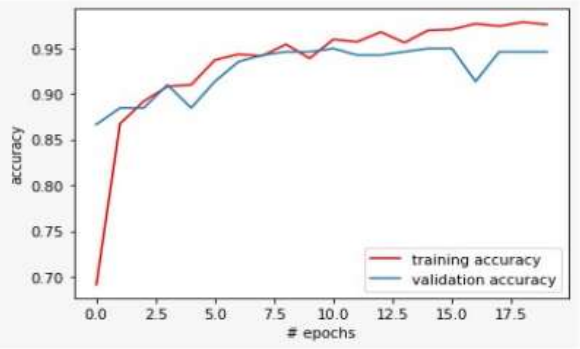


epochs vs accuracy corresponding to dataset 1

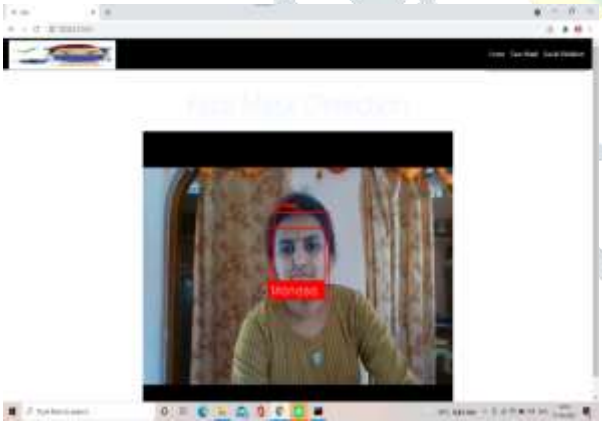
The system can efficiently detect partially occluded faces either with a mask or hair or hand. It considers the occlusion degree of four regions – nose, mouth, chin and eye to differentiate between an annotated mask or face covered by hand. Therefore, a mask covering the face fully including nose and chin will only be treated as "with mask" by the mode.



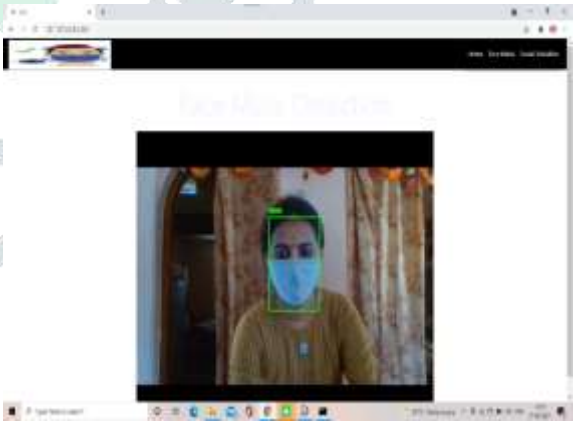
epochs vs loss corresponding to dataset 2



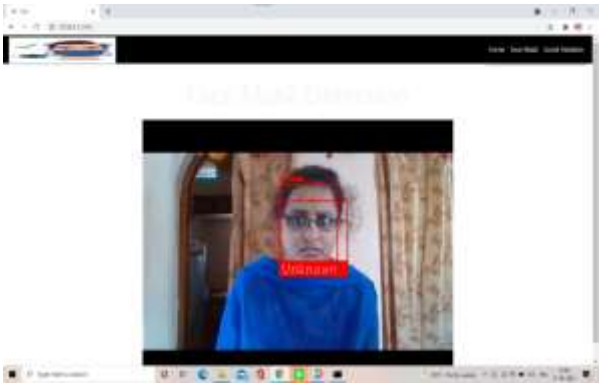
Epochs vs accuracy corresponding to dataset 2



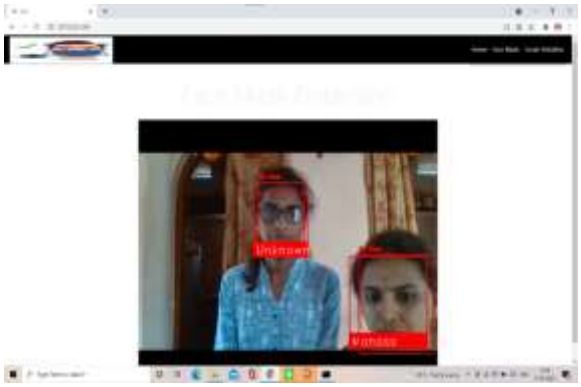
Known person manasa with no mask



Known person Manasa with mask



Unknown person with no mask



Unknown person with no mask and known person Manasa with no mask

The main challenges faced by the method mainly consist 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".

CONCLUSION

With the growth of science and technology, face recognition technology has made great successes, but there is still room for its improvement in practical application. Face Mask Detection with Face Recognition. With the help of the Res10-Caffe model we are going to identify the mask in the frames of the video. With help of Opencv we can control the laptop or surveillance Camera using Video Capture attribute, Face had been detected using hear-cascade face detect model, Mask detection our model detect the person by checking the eyes, nose and mouth is visible or not if nose mouth is visible the person is not wearing the mask and if not visible the person is wearing the mask the model is trained with 1000 person images with & without mask. The deployed model will donate immensely to the public health care scheme. In this paper to detect if a person is wearing the mask properly or not. And methodology of social distancing detection tool using a deep learning model also is proposed. By expending computer vision, the distance between people can be projected and any noncompliant pair of people will be indicated with a red frame and a red line. The proposed method was validated using a video showing pedestrians walking on a street. The visualization results showed that the proposed method is capable of determining the social distance measures between people which can be further developed for use in other environments such as office, restaurant, and school. Wearing a mask may be mandatory in the near prospect, considering the Covid-19 crisis. Several public service providers will ask the customers to wear masks correctly to avail of their facilities. The model can be additional developed to detect if the mask is virus prone or not i.e. the kind of the mask is medical, N95 or not.

REFERENCES

[1] L. Wolf, T. Hassner, and Y. Taigman, ''Similarity scores based on background samples,'' in Proc. Asian Conf. Comput. Vis., 2009, pp. 88–97.

[2] D. G. Miller, I. Kemelmachershlizerman, and S. M. Seitz, ''MegaFace A million faces for recognition at scale,'' in Proc. Comput. Vis. Pattern Recognit., 2015. [Online]. Available https//arxiv.org/abs/1505.02108

[3] K. He, X. Zhang, S. Ren, and J. Sun, ''Deep residual learning for image recognition,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 770–778.

[4] K. Simonyan and A. Zisserman, ''Very deep convolutional networks for large-scale image recognition,'' 2014, arXiv1409.1556.

[5] P. J. Phillips, A. N. Yates, Y. Hu, C. A. Hahn, E. Noyes, K. Jackson, J. G. Cavazos, G. Jeckeln, R. Ranjan, S. Sankaranarayanan, and J. C. Chen, ''Face recognition accuracy of forensic examiners, super recognizers, and face recognition algorithms,'' Proc. Nat. Acad. Sci. USA, vol. 115, no. 24, pp. 6171–6176, 2018.

[6] C. Xu, Y. Zhao, and J.-F. Zhang, ''Information security protocol based system identification with binary-valued observations,'' J. Syst. Sci. Complex., vol. 31, no. 4, pp. 946–963, Aug. 2018.

[7] K. Jiang, Z. Wang, P. Yi, G. Wang, T. Lu, and J. Jiang, ''Edge-enhanced GAN for remote sensing image superresolution,'' IEEE Trans. Geosci. Remote Sens., vol. 57, no. 8, pp. 5799–5812, Aug. 2019.

[8] K. Jiang, Z. Wang, P. Yi, G. Wang, K. Gu, and J. Jiang, ''ATMFN Adaptive-threshold-based multi-model fusion network for compressed face hallucination,'' IEEE Trans. Multimedia, early access, Dec. 18, 2019,

[9] S. Bharadwaj, H. Bhatt, M. Vatsa, R. Singh, and A. Noore, ‘‘Quality assessment based denoising to improve face recognition performance,’’ in Proc. CVPR Workshops, Jun. 2011, pp. 140–145.

[10] E. Zangeneh, M. Rahmati, and Y. Mohsenzadeh, ‘‘Low resolution face recognition using a two-branch deep convolutional neural network architecture,’’ Expert Syst. Appl., vol. 139, Jan. 2020, Art. no. 112854.

[11] Z. Wang, G. Wang, B. Huang, Z. Xiong, Q. Hong, H. Wu, P. Yi, K. Jiang, N. Wang, Y. Pei, H. Chen, Y. Miao, Z. Huang, and J. Liang, ‘‘Masked face recognition dataset and application,’’ 2020, arXiv2003.09093.

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

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

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

[15] 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.