An Efficient Masked-Face Recognition Model during COVID-19 using Artificial Intelligence Techniques

Graceline Jasmine S
School of Computer Science and Engineering
Vellore Institute of Technology,
Chennai, INDIA
graceline.jasmine@vit.ac.in

L. Jani Anbarasi
School of Computer Science and Engineering
Vellore Institute of Technology,
Chennai, INDIA
janianbarasi.l@vit.ac.in

Rukmani.P
School of Computer Science and Engineering
Vellore Institute of Technology,
Chennai, INDIA
rukmani.p@vit.ac.in

Vergin Raja Sarobin M
School of Computer Science and Engineering
Vellore Institute of Technology,
Chennai, INDIA
verginraja.m@vit.ac.in

Dhanya D
Department of Computer Science and Engineering,
Mar Ephraem College of Engineering and Technology,
Elavuvilai, INDIA
dhanvis@gmail.com

Prabhu Natarajan
Department of Information Technology University of Technology and Applied Sciences - Al Mussanah,
Al Muladdah, Sultanate of Oman
lord.n.prabhu@gmail.com

Febin Daya J L
Electric Vehicles Incubation, Testing And Research Center,
Vellore Institute of Technology,
Chennai, INDIA
febindaya.jl@vit.ac.

Abstract—

The COVID-19 is a modern-day crisis unparalleled in its effects leading to an enormous number of sufferers and security problems. People frequently use masks to guard themselves and lessen the spread of the virus. This makes face recognition a very challenging task since certain portions of the face, vital for recognition, are unseen. Face detection has become an important aspect with respect to safety and security and is also widely used in Image processing and Computer Vision. Several new algorithms are being analyzed and researched upon using various convolutional architectures to make the algorithm as efficient and truthful as possible. The crucial attention of researchers during the current COVID-19 situation is to devise means to manage this problem through quick and efficient solutions. These convolutional architectures have made it possible to bring out even the pixel details. The objective is to design a face classifier that can spot any face present in the frame regardless of its alignment, detect the unmasked facial regions, and enhance the recognition accuracy of different masked faces.

Keywords— *COVID-19, Masked Face, Artificial Intelligence, CNN, Convolution Neural Network*
DOI NUMBER: 10.14704/NQ.2022.20.12.NQ77334 **NEUROQUANTOLOGY2022;20(12): 3288-3294**

I. INTRODUCTION

COVID-19 pandemic can spread through contact or through contaminated surfaces. The traditional method of using the biometrics which is based on the person's fingerprint is not safe anymore. Physical touch of any device poses to be a risk in acquiring this virus. To prevent the spread of these viruses, various measures are asked to be taken for the proper health and well-being of the person. Measures that are recommended to be taken are washing hands regularly, maintaining a social attendance, avoid attending large gatherings etc. But among all of these, one of the most important measures that need to be taken when going anywhere is to wear a mask. Research indicated that maintaining social distance between co-workers as well as wearing face masks were effective means of reducing this risk.

While wearing a mask proves to be a benefit, it has its issues also. Wearing a mask causes problem such as fraudsters or thieves take advantage of these masks, stealing of valuable items without being identified and the present face recognition approaches are not effective when wearing a mask and does not offer the whole face image. The above problems prove to be a challenge for facial recognition [1, 2].

In order to attack the problems, we aim to recognize a face with mask based on the unmasked regions. We also apply a deep learning technique to extract features that are beneficial for facial recognition and improve the recognition accuracy of diverse masked faces.

II. LITERATURE SURVEY

Occlusion detection

Face recognition is a computer technology that decides the position and dimension of human face in digital image. This is a vital technology in face information processing [3]. The difference between face detection and face recognition is that in face detection we need to find if there is a face in that particular image and for recognition, we know whose face it is [4].

A. Appearnce-Based Face Recognition Approach

Rather than the local illustration of the complete image for face identification, this method is founded on representing the total face. One of the main algorithms used in dimensionality reduction is the Principal Component Analysis. This technique is deliberated as a arithmetical process that practices an orthogonal transformation to translate a set of observations of probably interrelated variables into a set of values of linearly uncorrelated variables named principal components [5]. PCA is for discovery of patterns in data of high measurement. On compressing the data, it will reduce the number of dimensions minus loss of information [6].

B. 3D based Face recognition

Here a 3D sensor is used to catch information about the shape of a face. The data collected are then used to find distinguishing points on the exterior of the face, such as the shaper of the eye, nose, and chin. Data collected is then subject to pre-processing to get clear images. Feature extraction algorithm is used to extract features and is stored in feature data base. When we want to test a face, the input image is matched with the face images stored in the data base. By applying 3-dimensional based approach, the thorough information has the potential to generate more sound, judicial features specifically with facial occlusions or missing part [7]. The 3 dimensional based is considered to be faster and need less memory. 3-dimensional model can reduce the difficulty of head posture, brightness and preserve all information's about the face.

C. Template Based Approach

This technique applies the relation between the pattern of the input image and the defined pattern of the face or its features. The method based on template is humble but works only for frontal portion of the face lacking any occlusion. Likewise, face needs to be of similar size as that of the preset templates.

D. Haar-cascade features approach

This approach is grounded on attributes mined from images used in pattern identification. This approach will train a model for spotting faces and objects in a picture. The aim is to dig-out the combination of these features that represent a face. These features include left eye, right eye, nose, lips and forehead. These features are detected using the Haar-features or also called as Haar-wavelets. Edge based features are capable of detecting edges quite effectively. Line based features have the ability to detect lines effectively.

The cascade classifier picks up positive and negative images. This classifier contains a group of stages. They are named as group of weak learners or decision stumps. The technic called Boosting is used to

train these individual stages. Based on objects are found or not found, regions are named. If an object is found, it is labelled as positive and in case the object is not found, it is labelled as negative. For the purpose of detection, a sliding window may also be used. The window will slide to the following location, in case the label is negative.

E. Convolutional Neural Network (CNN)

CNN is a kind of neural network, widely applied to recognize, classify [8] and analyze images. It is known that CNN has great precision as well. Each layer takes a multi-dimensional range of numeric values as input and produces one more multi-dimensional range of numeric values as output. CNN network consists convolutional layer, hidden layer and a classification layer [9].

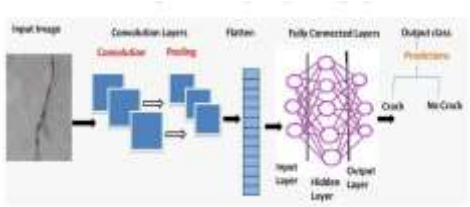

Fig. 1. Convolution Neural Network Architecture [9]

One chief advantage of using CNN is that it studies or learns the highest imported features in the absence of any human support and is shown in Fig. 1.

Feature Extraction

Facial feature abstraction is way to identify the nose, mouth, eyes and other parts existing in the human face. These features are very vital in facial identification, face matching, recognition of different facial expressions, classification and tracking of faces. The system takes image as input. Processing is then with either cloud-based dataset or local capabilities. It then mines over twenty-seven facial information. This information includes the features such as mood, gender, visual age, color [10], quality of picture [11] etc.

The proposed system is to discover the face in the specified image by representing borders in a geometrical four-sided shape on the sensed face. This will highlight the landmarks points which then leads to removing the face characteristics along with precise scores from input [12].

The image loaded should have at least a face size of 36 x 36 pixels. This image is to be in any standard supported image formats such as jpg, png, bmp. First a lawful API key is acquired. This is a matchless subscription key to enter authorized services presented by the vendor. Next the input image is uploaded to the system. Then comes the face recognition stage. In this stage, the system senses the faces in the image and then draws a four-sided box shaped structure above the detected faces. Along with this a series of landmarks are also detected. Using these landmarks face direction is determined.

A. Feature Extraction Algorithm using Mixed Transforms

Feature extraction algorithm have changed over the years. For facial recognition, a key step is to mine the unique and reliable features that support in identification. Here, a Discrete Wavelet Transform (DWT) and Discrete Cosine Transform (DCT) are used iteratively to find the best features for facial recognition [13]. In DCT there is an image which exists in spatial domain. DCT will help in converting the image from spatial domain to frequency domain. ORL, Yale and Ferret-Fc databases are used. These databases are used for comparison with their proposed system.

The entire data used have a definite extent of people with a minimum of 10 postures per individual.6 There are two tools used for the entire training & testing of these postures. First tool is DWT and the second tool will be DCT. In the DWT tool, entire postures will be converted using the DWT and the DCT. DCT is a part of the diverse face identification system. This diverse system uses DCT as a strainer that trails DWT and is repeatedly iterated till a threshold is arrived. The energy variation in the output of DCT and DWT will determine the threshold.

In the suggested system each posture, A, is transformed first through the DWT and half of the coefficients, C0. Along with this the uppermost values are retained. A weight matrix denoted by W is used to multiply each of these coefficients. The first weight matrix W0 that is equal to [α0] which contains a starting value of 0.8 is multiplied using the Hadamard Product. The product obtained is changed back into the spatial domain, D.

In YALE database, the suggested system was nearby to the regular and highest values. Comparatively, highest recognition rate at 6 training images was improved. So all-in-all the suggested system accomplished quite good in the bigger database equated to the systems tested.

B. Feature Vector Selection using Genetic Algorithm

The biometrics which are currently used for uniquely identifying a person facial recognition are speech, iris, fingerprint and facial recognition. Facial

identification will compare the shapes founded on facial structures of a subject. A face recognition method includes several steps. This process involves face detection, conversion of raw pixel values into a feature vector, lower the characteristics essential for identification, and identifying the face founded on this feature vector. Among all these steps, feature mining is a key step. It is essential that these facial vectors are denoted in a best means so that the time needed for calculating is the lowest. Main constituent study has frequently been used for dimensionality drop. Nevertheless, using a progressive algorithm like GA confirms that only the best features are designated to upsurge effectiveness and exactness whereas guaranteeing overfitting do not occur [14].

C. Local Binary Pattern (LBP)

The above picture displays outline of activities in face recognition procedure. The face is detected from raw images in the first stage which is called pre-processing stage and the detected images will have facial characters like corner of eyes, line forehead, cheeks, nose, chin and lips. To extract the feature, it will be sent for processing after obtaining such image. Followed by that matching and classification are performed. The face data based is used for matching. The feature model then compared with the feature got in the above process and given the result as a recognized face. And if it fails, we get the result as unrecognized face.

LBP is used to describe the texture of a rectangular block as shown in figure 2. It assigns the value of a pixel by associating its value with the adjacent pixels. A binary number is formed from each of these pixel values. In LBP, we train the algorithm. We use a set of data which contains the facial images of people to be identified. An ID is set for each image. This ID can be name or a number. This will help the algorithm to recognize the input image. After that, an intermediate image is created that will describe the novel features in an improved way. This is done by emphasizing the facial characteristics.

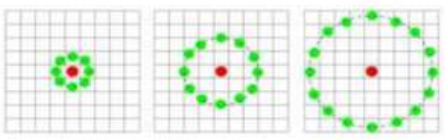

Fig.2. Algorithm of Local Binary Pattern [15]

The image is then divided using multiple grids which are represented in the form of histograms. The images of training dataset are symbolized by using these histograms. Therefore, to discover the image that matches the input image, both histograms are compared and return the image to the nearby histogram. For the purpose of classification this histogram is used. LBP histogram uses fundamentally 4 parameters and these are known as radius, neighbors, grid X and Y.

Let us assume an image which is in grayscale. A portion of the image can be obtained as blocks that have the size of 3 x 3 pixels. It may otherwise be called as a Matrix. Out of these pixel values, let us take the middle value as a threshold. This specific value will govern all the other new neighboring values. If the values are greater than the threshold, let us give it as One and if it is less than the threshold, let us give the value as Zero. Therefore, the Matrix contains Zeroes and Ones which are called binary values.

To get a new binary value from the Matrix, we bring together all binary values starting from left hand side top corner. We can change this new binary value to a decimal value. This new decimal value may be used to arrive the central value of the Matrix. As a result, a new image is obtained which is far improved than the original image.

D. Genetic Algorithm (GA)

Genetic algorithm is used to detect the finest traits for face recognition. In this method 4096 traits are considered and these are extracted from LBP. Chromosomes are represented as 4096-bit binary number. Each bit will indicate the use of a trait from 4096 traits. Initially, thirty chromosomes are produced, each having a randomly produced value that ranges from 0 to 4096. Next, a fitness function is applied to each chromosome. K-NN classifier is used and it is trained. For this selected LBP traits are used. Correctness is determined on the traits in the test set.

First a stable selection process is done. Four appropriate chromosomes are straight away upheld to the offspring. The rest twenty-six chromosomes experience crossover and transformation. Due to one-to-one relation between parents, the crossover takes place. Then mutation is performed where mutation means the reversal of a bit in chromosomes. This is done by controlling bits in an arbitrarily produced range of indices. The Genetic Algorithm ends if any of the subsequent situations are fulfilled: a) The correctness rises by fifteen percentage of the correctness attained by considering all 4096 features. b) The required number of features are lowered by approximately sixty-five to seventy percentage. GA has decreased the rate by around fifty-five percentage

although maintaining correctness to remain as large as ninety-six percentage.

III. Face Recognition

One of the prime worries when it comes to facial recognition are occlusions. Occlusions can be of various forms. It can be of a person wearing a mask, covering the mouth with hand, part of the hair falling on the face etc. The major apprehension to this work is with regards to facial masks. This can boost the accuracy when faces with different masks or different facial features, are recognized. One practicable approach that has been suggested comprises of first identifying the facial areas. The Multi-Task Cascaded Convolutional Neural Network (MTCNN) has been used to detect the occluded face problem. Then Google FaceNet embedding model is implemented to perform facial feature extraction. In final stage, for the classification task Support Vector Machine (SVM) has been used [16, 17]. Facial recognition method involves of 3 major components: To detect a face from a given image, extracting the features and lastly the identification.

A. *Facial Image Acquisition*

For this first a collection of masked and unmasked faces from AR, Delhi-IIIT face database has been collected. To deal with the issue of scarcity of huge number of images, data augmentation process is done in order to enlarge the dataset images. This makes it more efficient.

B. *Masked Face Detection Using MTCNN*

A pre-trained MTCNN model is sued for face detection. There are three networks in this model. These are P-Net, R-Net and O-Net. This model is identified as a MTCNN [18]. An image is passed to the model and an image pyramid is created in order to perceive face of all diverse sizes. It will create copies of the same image in different sizes. For each scaled copy, there is a 12x12 stage kernel that will go through all part of the image and scan for faces. It starts from the top-left corner to the rest of the image. If a face is found, this part of the image is passed to the P-Net which yields the coordinates of a bounding box. Coordinates of the bounding box are saved and passed to the next layer which is the R-Net layer.

In the R-Net layer, it holds the coordinates of the new and more correct bounding box. Sometimes the images can exist partly out of the frame. Then we copy the pixel values of the image inside the bounding box into an array and fill the rest with zeros. After that we resize them to 24x24 pixel and normalize them to value ranging between -1 to 1. After standardizing the coordinates, it passes to the O-net. The O-Net contains the coordinates of the bounding box, coordinates of five facial landmarks.

C. *Image Post-processing*

After a face is detected an input image is provided upon which image trim and resizing technics are applied. An MTCNN model is used to produce the bounding box that is found at the face detection level.

D. *Feature Extraction using FaceNet*

This is an advanced model for face recognition for confirmation and clustering [19]. FaceNet advances on twenty-two deep CNN layers. The image of an individual's face is given as input and the output obtained is straightaway trained on different layers to get compressed into 128 numbers of vector. The entire major details of an image are embedded within a vector. Once the modification stage is finished, the fully linked layers are utilized as the face descriptor [20, 21]. These descriptors turn out to be similarity-based descriptor using the embedding module. To formulate an exceptional feature vector from a template, the operator used is max operator as shown in figure 3.

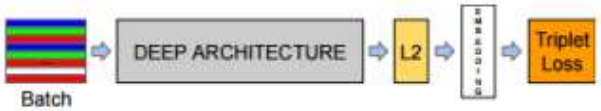

Fig.3. Pipeline of FaceNet Model [22]

E. *Face Verification Using SVM*

A unified SVM is used to fulfill the verification process. The process is done within the SVM by classifying and recognizing the candidate's face. The SVM algorithm is efficiently used in diverse classification linked complications subsequently it was submitted [23]. SVM catches a hyperplane to do classification tasks of a maximization problem. The boundary is maximized between the two classes for a given input-target pair. The margin represents the class separation efficiency.

IV. Results & Discussion

Using the above mentioned datasets, the accuracy for these masked faces are measured. Mask On (masked face) and Mask Off (un-masked) facial images of various combinations attained a better recognition accuracy of around 80% to 90%.

a B

Fig. 4(a) Person predicted as wearing "Mask" (b) Person predicted as not wearing mask

If the person is wearing a mask, the label assigned is "mask". A bounding box in green color appears when the label is "mask" as shown in Fig. 4(a). If the person is not wearing a mask, the label assigned is "no mask". In this case a red coloured bounding box appears around the face as shown in Fig. 4(b).

During the covid pandemic situation, in the public it is necessary to wear mask. However amids multiple warning and instructions, people find ways to avoid wearing mask. This situation worse the situation even more. In order to avoid this behaviour, an automated mechanism of identifying the people who were not wearning mask is being developed by many researchers. Such research outputs were thouroughly analyzed and disscussed in this paper.

V. CONCLUSION

Facial recognition techniques have always been under constant research. Various methods are being used for the same. A primary concern also has been with regards to occlusions. This reduces the chance of accurate recognition. Different approaches for occlusion detection such as template based, 3D based face recognition, Haar-cascade features approach, convolutional neural network etc. have proven to be efficient for occlusion removal. Feature extraction is another important aspect for face recognition. Algorithms such genetic algorithm and local binary pattern are used for extracting the feature vectors.

ACKNOWLEDGMENT

The authors would like to thank Vellore Institute of Technology, Chennai for supporting this research work.

References

[1] Sharadhi AK, Gururaj V, Shankar SP, Supriya MS, Chogule NS. Face Mask Recogniser Using Image Processing and Computer Vision Approach. Global Transitions Proceedings. 2022 Apr 3.

[2] Hussain GJ, Priya R, Rajarajeswari S, Prasanth P, Niyazuddeen N. The face mask detection technology for Image analysis in the covid-19 surveillance system. InJournal of Physics: Conference Series 2021 May 1 (Vol. 1916, No. 1, p. 012084). IOP Publishing.

[3] Wang YQ. An analysis of the Viola-Jones face detection algorithm. Image Processing On Line. 2014 Jun 26;4:128-48.

[4] Annagrebah S, Maizate A, Hassouni L. Real-time Face Recognition based on Deep neural network methods to solve occlusion problems. In2019 Third International Conference on Intelligent Computing in Data Sciences (ICDS) 2019 Oct 28 (pp. 1-4). IEEE.

[5] Kassambara A. Practical guide to principal component methods in R: PCA, M (CA), FAMD, MFA, HCPC, factoextra. Sthda; 2017 Aug 23.

[6] Ejaz MS, Islam MR, Sifatullah M, Sarker A. Implementation of principal component analysis on masked and non-masked face recognition. In2019 1st international conference on advances in science, engineering and robotics technology (ICASERT) 2019 May 3 (pp. 1-5). IEEE.

[7] Zhang L, Verma B, Tjondronegoro D, Chandran V. Facial expression analysis under partial occlusion: A survey. ACM Computing Surveys (CSUR). 2018 Apr 18;51(2):1-49.

[8] Jawahar M, Ravi V, Prassanna J, Jasmine SG, Manikandan R, Sekaran R, Kannan S. CovMnet–Deep Learning Model for classifying Coronavirus (COVID-19). Health and Technology. 2022 Aug 4:1-6.

[9] Reghukumar, Arathi, et al. "Vision based segmentation and classification of cracks using deep neural networks." International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 29.Supp01 (2021): 141-156.

[10] Jawahar M, Anbarasi LJ, Jasmine SG, Narendra M. Diabetic Foot Ulcer Segmentation using Color Space Models. In2020 5th International Conference on Communication and Electronics Systems (ICCES) 2020 Jun 10 (pp. 742-747). IEEE.

11] Fu B, Kirchbuchner F, Damer N. The effect of wearing a face mask on face image quality. In2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021) 2021 Dec 15 (pp. 1-8). IEEE.

[12] Patil PR, Pramod R, Sandhya S. A General Approach on Facial Feature Extraction and Face Attributes. In2018 3rd International Conference on Computational Systems and Information Technology for Sustainable Solutions (CSITSS) 2018 Dec 20 (pp. 151-155). IEEE.

[13] Sapijaszko G, Alobaidi T, Mikhael WB. Adaptive Feature Extraction Algorithm using Mixed Transforms for Facial Recognition. In2018 IEEE 61st International Midwest Symposium on Circuits and Systems (MWSCAS) 2018 Aug 5 (pp. 226-229). IEEE.

[14] Yerremreddy S, Talele KT, Kokate Y. Genetic Algorithm for Optimal Feature Vector Selection in Facial Recognition. In2019 IEEE 5th International Conference for Convergence in Technology (I2CT) 2019 Mar 29 (pp. 1-5). IEEE.

[15] https://en.wikipedia.org/wiki/Local_binary_patterns

[16] Rahman MM, Islam MS, Rahman MH, Sassi R, Rivolta MW, Aktaruzzaman M. A new benchmark on american sign language recognition using convolutional neural network. In2019 International Conference on Sustainable Technologies for Industry 4.0 (STI) 2019 Dec 24 (pp. 1-6). IEEE.

[17] Vignesh Baalaji S, Vergin Raja Sarobin M, Jani Anbarasi L, Graceline Jasmine S, Rukmani P. An Extensive Survey of Applications of Advanced Deep Learning Algorithms on Detection of Neurodegenerative Diseases and the Tackling Procedure in Their Treatment Protocol. Advanced Analytics and Deep Learning Models. 2022 May 10:207-30.

[18] Z hang K, Zhang Z, Li Z, Qiao Y. Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Processing Letters. 2016 Aug 26;23(10):1499-503.

[19] Schroff F, Kalenichenko D, Philbin J. Facenet: A unified embedding for face recognition and clustering. InProceedings of the IEEE conference on computer vision and pattern recognition 2015 (pp. 815-823).

[20] Jawahar M, Prassanna J, Ravi V, Anbarasi LJ, Jasmine SG, Manikandan R, Sekaran R, Kannan S. Computer-aided diagnosis of COVID-19 from chest X-ray images using histogram-oriented gradient features and Random Forest classifier. Multimedia Tools and Applications. 2022 May 10:1-8.

[21] Basu A, Ali MF. COVID-19 Face mask recognition with advanced face cut algorithm for human safety measures. In 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT) 2021 Jul 6 (pp. 1-5). IEEE.

[22] Schroff F, Kalenichenko D, Philbin J. Facenet: A unified embedding for face recognition and clustering. InProceedings of the IEEE conference on computer vision and pattern recognition 2015 (pp. 815-823).

[23] Byun H, Lee SW. Applications of support vector machines for pattern recognition: A survey. InInternational workshop on support vector machines 2002 Aug 10 (pp. 213-236). Springer, Berlin, Heidelberg.