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

CONVOLUTIONAL NEURAL NETWORK BASED FACE MASK DETECTION USING GLCM AND PCA FEATURE EXTRACTION ALGORITHM

[1]Simranjeet Kaur, [2]Priyanka Jarial, [3]Navroz Kahlon

[1]Student, [2]Assistant Professor, [3]Assistant Professor
[1]Computer Engineering
[1]Punjabi University, Patiala, Punjab, India

Abstract: As the entire world battles COVID19, masks have been made mandatory in public places, offices, and bus stops, among other places. Although recognizing masked faces with various head pose angles is a huge difficulty for professionals. A number of techniques for identifying faces hidden behind masks have been developed in recent years. These technologies were difficult to utilize, and the feature extraction approaches used were ineffective. This research proposes an effective way to overcome the constraints of conventional models. The suggested model chooses images from the available dataset and detects faces from these images. to detect edges and extract more features, these discovered faces images are then transformed to grey scale images. Moreover, the proposed model employs PCA and GLCM feature extraction schemes to extract additional informative features from the images. The network's dimensionality is reduced by PCA, and the GLCM extracts four texture-based features: contrast, correlation, energy, and homogeneity. After that, all of these features are merged into a single feature matrix. The suggested CNN network is then given the feature matrix for training and testing purposes. Finally, the suggested model's performance is evaluated in terms of several dependent variables using the MATLAB simulation software. The results of the simulations demonstrated that the suggested model is more effective at detecting masked faces.

***Index Terms*- COVID-19, Deep learning (DL), mask detection**

I. INTRODUCTION

Computer vision field is the branch of artificial intelligence that usually focuses on making computers to imitate human vision, including learning, making decisions and performing necessary actions based on visual inputs, i.e. Images. Computer vision also plays a major role in pattern recognition. Human computer interaction (HCI) becomes more effective if computer can predict about emotional state of a person and hence mood of a person from supplied images on the basis of facial expressions can be classified by computer [1]. Since the outbreak of COVID-19 in late December 2019, face masks are regularly worn by people to protect themselves from the deadly virus. The performance of face-related algorithms is challenged because they are not able to detect a person's facial features under the mask. The facial expressions are reported to play primary role in reviling the actual emotional status of an individual. Ekam et al. classified the facial expressions into seven universal expressions: anger, contempt, surprise, happiness, sadness, fear, and disgust [2]. Meanwhile, a crucial step in facial expression recognition is the feature

extraction task that involves obtaining features from the face for further processing. There are various approach for extracting facial features, and the geometric method has been widely applied. The geometric-based feature extraction method considers information relating to the eyes, mouth, nose, eyebrows, other facial components [3]. These components are very important for face recognition, detection and attribute extraction. Many methods are available and developed for feature extraction in the intention of reliability, complexity, efficiency, speed and accuracy of the results. All these methods are classified into three categories, namely Holistic Based Approach, Feature Based Method and Template Based Approach [4].

a. Holistic Based Approach

In holistic approach, the whole face region is taken into account as input data into face detection system. Examples of holistic methods are eigenfaces (most widely used method for face recognition), probabilistic eigenfaces, fisher faces, support vector machines, nearest feature lines (NFL) and independent-component analysis approaches. They are all based on principal component-analysis (PCA) techniques that can be used to simplify a dataset into lower dimension while retaining the characteristics of dataset [5].

b. Feature-based Approach

In feature-based approaches tries to extract features present in the image to match it against the knowledge of face features. Feature-based approach is divided into three steps Low-level analysis: Low-level analysis, Feature analysis, and Active shape model. The Low-level analysis manages with segmentation of visual components using some properties such as edge detection, gray scale analysis, motion, color information etc [6]. The objective of feature analysis algorithms is to search structural features that exist also with varying pose, viewpoint, or lighting conditions. The structural features are used to locate faces. These methods are proposed generally for face localization. Various feature analysis algorithms are Viola Jones method, Gabor feature method and Constellation method etc. Active shape models (ASMs) concentrate on complex non-rigid features such as real physical and higher level appearance of features. ASMs are expected to automatically locate landmarks that characterize the shape of any statistically modelled object & facial features such as the eyes, nose, lips, mouth and eyebrows in an image. In the training stage of an ASM, the development of a statistical facial model from training set containing images with manually annotated landmarks. ASM is classified into three groups i.e. Snakes, Point Distribution Model (PDM), Deformable templates [7].

c. Geometric/Template-based Approach

The template based methods compare the input image with a set of templates. The set of templates can be constructed using statistical tools like Support Vector Machines, Principal Component Analysis, Linear Discriminant Analysis, Independent Component Analysis, Kernel Methods, or Trace Transforms. The geometry feature-based methods analyze local facial features and their geometric relationships. This approach is sometimes called feature based approach. Examples of this approach are some Elastic Bunch Graph Matching algorithms [8].

Face recognition pipeline consists of four main phases: face region detection, alignment, feature extraction, and classification where the most crucial phase is feature extraction. Hand-crafted features have achieved reasonable results for constrained environment. However, the recognition of unconstrained face images is an evolving and challenging field in the context of the real-world issues such as varying poses, expressions, illumination, and quality of images [9]. Classification is the mechanism by which the class of variables is predicted. Occasionally classes are referred to as labels or levels. For example, a classifier is utilized to categorize certain characteristics extracted from a face picture and to produce a label that is used to identify a person [10]. In recent years, with the development of machine learning, deep learning as a new research direction caused widespread concern in the field of artificial intelligence. The machine learning techniques for facial recognition have provided decent results; these techniques do not perform well under unconstrained environments. This is mainly because machine learning approaches rely on hand-crafted features or representations selected by human experts that may work for one scenario and fail for other situations [11]. On the other hand, deep learning (DL)-based approaches have proven to be most suitable as the representations and features are discovered automatically from data by the back-propagation learning technique. Other than this to detect and identify faces, various researchers have performed various techniques in the recent years that have been described in the next section.

II. LITERATURE REVIEW

During past decades various methods were proposed by different experts to detect the occluded faces effectively, some of them are defined here; This paper addresses the problem of head poses estimation in order to infer a non-intrusive feedback from users about gaze attention**. Shuang Li et al. [12],** this paper proposes a head pose classification method based on the algorithm of line portrait generation and convolutional neural network (CNN). **Tsun-Yi Yang et al. [13],** This paper proposes a method for head pose estimation from a single image. Our method is based on regression and feature aggregation. For having a compact model, we employ the soft stagewise regression scheme. **Shahenda Sarhan et al. [14]** this paper, a combined adaptive deep learning vector quantization (CADLVQ) classifier is proposed. **Jirui Lin et al. [15],** This study presents a method to estimate the poses of image set by applying nonlinear least squares to facial landmarks. **Moein Razavi et al. [16],** this paper developed a computer vision system to automatically detect the violation of face mask wearing and physical distancing among construction workers to assure their safety on infrastructure projects during the pandemic. **I. Q. Mundial, et al. [17],** In this paper, we present a methodology that can enhance existing facial recognition technology capabilities with masked faces. We used a supervised learning approach to recognize masked faces together with in-depth neural network-based facial features. **M. Z. Khan et al. [18],** This paper proposes an algorithm for face detection and recognition based on convolution neural networks (CNN), which outperform the traditional techniques. **Ji-Hae Kim et al. [19],** In this paper, we propose a new scheme for FER system based on hierarchical deep learning. In this paper, facial landmark detection has been used for feature extraction and Convolutional Neural Network Classifier. **Xiao Liu et al. [20],** This paper presents a facial emotion recognition technique using two newly defined geometric features, landmark curvature, and vectorized landmark.

From the literature survey it was analyzed that in the field of AI, various researchers have proposed efficient techniques to detect and identify face while wearing a mask and with different head pose angles. In traditional methods, the experts used CNN to extract features from the images. These techniques provided optimal outputs but had several shortcomings. These techniques were time-consuming and were unable to detect and identify face when the individual was wearing a skin color mask. In addition to this, the traditional techniques were unable to detect edge effectively, which made pre-processing of the images complex and difficult. Other than this, the traditional models were unable to recognize the face by using the HSV channel, resulting in a decrease in overall system efficiency and an increase in complexity. These issues must be addressed, which motivates the current work to develop a system that can extract features from images with different head pose angles.

III. PRESENT WORK

To overcome issues related to the conventional models, it became important to create a model that could smoothly extract features from photos to detect masked faces. In this research, a model is developed for identifying and detecting masked faces with various head pose angles. Furthermore, to extract features from the images, the suggested model would utilize GLCM (Gray Level Co-Occurrence Matrix) and PCA (Principal Component Analysis). The main motive behind using GLCM and PCA technique in the proposed work is explained in this section. The GLCM is a statistical approach that extracts a large number of features from the RGB color images. In addition to this, the GLCM approach is simple to use and by reducing processing time and complexity this approach shows optimal outcomes in a variety of applications. The PCA is a dynamic data analysis methodology that reduces the dimensions of datasets to boost their usability while reducing information loss. The PCA technique removes the correlated features, improves visualization, reduces overfitting, and enhances the performance of the algorithm. Hence, by deploying the GLCM and PCA together, the classification of head pose images can be performed with high performance.

3.1 Methodology

The suggested methodology is carried out in several steps, which are briefly described in this section.

Step 1: The suggested model's first step is to collect data from the datasets. The information gathered can come from publicly available datasets online or be acquired using a camera lens. The MAFA dataset, which is freely accessible online has been utilized in the suggested model.

- **Dataset used**

The MAFA dataset will be used in the suggested method to collect the various RGB-colored images. This dataset comprises a collection of 23,845 pictures, with 20,139 being used for training and 3,706 being used to test the suggested model. Fig. 1.1 depicts the photos taken from the chosen dataset.

Figure 1.1 Images obtained from the dataset

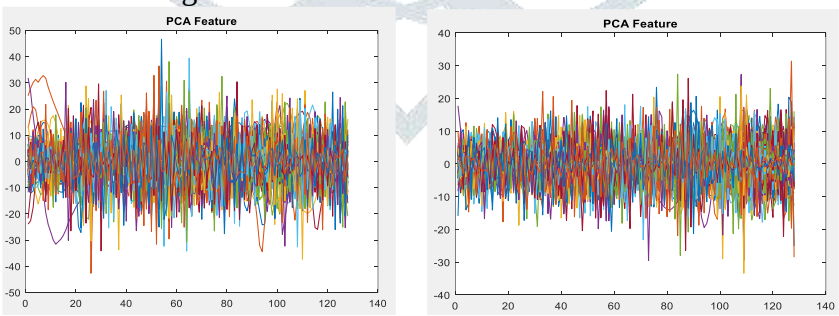
Step 2: After the images from the accessible dataset have been collected, the next step is to detect faces from the images.

Step 3: Then the RGB face detection photos are transformed to gray scale images. The gray scale images, which are illustrated in Fig. 1.2, can easily detect the edges of the images, making the feature extraction method simple and easy.

(a) (b)

Figure 1.2 gray scale converted images

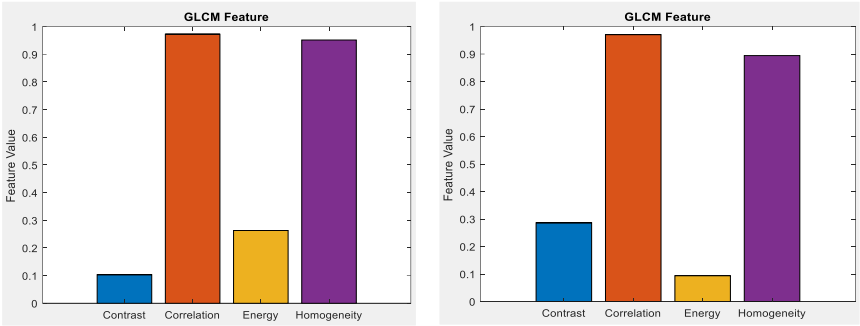
Step 4: the suggested model's next step is to extract only essential features from the selected RGB images to minimize the system's dimensionality and complexity. For this purpose, the suggested system utilizes the PCA feature extraction technique. PCA's major task is to identify and choose only those features that can be used for identifying a person. Fig. 1.3 (a) and (b), represents the PCA features, which are extracted from the selected RGB images.



(a) (b)

Figure 1.3 PCA feature extracted images

Step 5: The next step is to extract GLCM-based texture features from the chosen RGB. Contrast, correlation, energy, and homogeneity are four GLCM features retrieved from the given images in the suggested work. Fig. 1.4(a) and (b) illustrate the extracted features of the GLCM



(a) (b)

Figure 1.4 GLCM features of selected images

Step 6: After extracting all the features from the given images, the next step is to integrate the features of gray scale, PCA features, and GLCM features into a single feature matrix. This feature matrix, together with its configurational collection of distinct CNN properties, is subsequently sent to the proposed CNN network for training purposes. Table 1 represents the exact values of CNN parameters.

Table 1: Specific parameter values

CNN Parameter	Value
Max Epochs	30
Initial Learn Rate	0.001
Learn Rate Drop Factor	0.1
Learn Rate Drop Period	20
No of Layer	18
Min. Batch Size	32
Input Layer	13 x 13
Conv. Filter Window size	3 x 3
No of stride	2 x 2
No of Pool Size	2 x 2
Dropout Layer Probability	0.2

Step 7: Finally, the suggested model's performance is evaluated by supplying testing data to it. The simulation outcomes are calculated by using a variety of performance parameters such as RMSE and accuracy. The next section contains a full discussion of the simulated outcomes.

IV. RESULTS AND DISCUSSION

In the MATLAB simulation software, the suggested PCA-GLCM-CNN model performance is tested and validated. The simulation results are expressed in terms of the RMSE curve and accuracy and are briefly discussed in this section.

4.1 Performance Evaluation

The suggested PCA-GLCM-CNN system's performance is evaluated and analyzed in terms of the RMSE value obtained. The majority of data in traditional models is lost in the network during the training process, although this is not the case with the suggested approach. The RMSE curve should decrease as the number of iterations increases. The RMSE achieved during the training process of the proposed PCA-GLCM-CNN model is shown in Fig. 2.1.

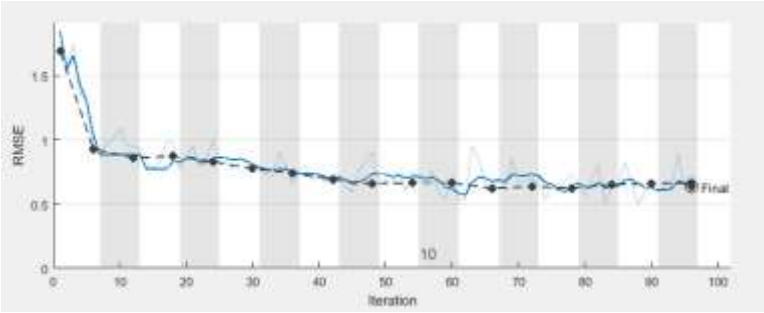


Figure 2.1 RMSE achieved during CNN training process

The line graph for RMSE, which is acquired in the developed model during the training process, is shown in Figure 2.1. The number of iterations is represented on the x-axis, while the RMSE value is represented on the y-axis. It can be seen from the graph that the proposed model's RMSE value is initially quite high. The RMSE value, on the other hand, rapidly declines as the number of iterations increases. Furthermore, the RMSE value has fallen to about 0.6 in the final iteration.

Finally, in figure 5.3, the performance of the suggested PCA-GLCM-CNN model is assessed and observed in terms of accuracy.

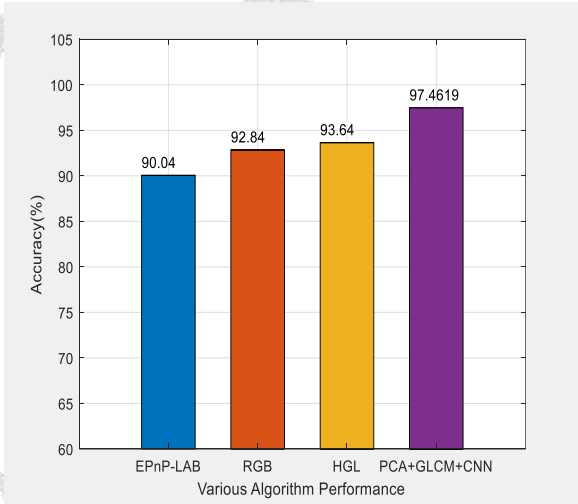


Figure 2.2 Comparison Graph for Accuracy

The accuracy comparison graph for classic EPnP-LAB, RGB, and HGL models, as well as the new PCA-GLCM-CNN model, is shown in Figure 5.3. The blue, orange and yellow colored bar represents the performance traditional EPnP-LAB, RGB, and HGL models, respectively. The purple-colored bar represents the proposed PCA-GLCM-CNN model performance. The value of accuracy obtained by the traditional EPnP-LAB model is equivalent to 90%, whereas the value of accuracy obtained by the traditional RGB and HGL models is 92.8 % and 93.6 %, respectively, according to the graph. However, when the accuracy value for the suggested PCA-GLCM-CNN model was calculated, it came out to be 97.46%. This demonstrates that the suggested PCA-GLCM-CNN model is more accurate and efficient at detecting the faces of masked people with various head position angles. The confusion matrix, illustrated in Figure 2.2, depicts the accuracy of the proposed PCA-GLCM-CNN model. Table 2 shows the specific accuracy reached by the traditional EPnP-LAB, RGB, and HGL models, as well as the proposed PCA-GLCM-CNN model.

Table 2: specific values for Accuracy

Algorithm	**Accuracy (%)**
EPnP-LAB	90.04
RGB	92.84
HGL	93.64
PCA+GLCM+CNN	97.46

The suggested PCA-GLCM-CNN model is more easy, efficient, and accurate in finding masked faces with diverse head pose angles, as shown in the graphs and tables.

V. CONCLUSION

In this research, a technique is presented to detect masked faces and undergo various head pose angels. The primary goal of the proposed model is to lower the overall complexity of traditional models. To achieve the desired goal, the suggested model included gray scale features, PCA features, and GLCM features with CNN. In the MATLAB simulation software, the proposed PCA-GLCM-CNN technique performance is evaluated and validated. In terms of accuracy, the simulation results were compared to standard EPnP-LAB, RGB, and HGL. The value of accuracy obtained from the traditional EPnP-LAB is 90.04%, HGL is 93.64%, and RGB is 92.82%. The accuracy obtained in the suggested PCA-GLCM-CNN system was 97.46 %. Furthermore, throughout the training phase, the value of RMSE is near 0.6, and the loss curve value is virtually minimal to zero. All of these findings show that the suggested PCA-GLCM-CNN technique is more efficient and effective in identifying masked faces.

REFERENCES

[1] Y. Muttu and H. G. Virani, "Effective face detection, feature extraction & neural network based approaches for facial expression recognition," 2015 International Conference on Information Processing (ICIP), pp. 102-107, 2015.

[2] M. Owayjan, R. Achkar and M. Iskandar, "Face Detection with Expression Recognition using Artificial Neural Networks," 2016 3rd Middle East Conference on Biomedical Engineering (MECBME), pp. 115-119, 2016.

[3] M. I. Rusydi, R. Hadelina, O. W. Samuel, A. W. Setiawan and C. Machbub, "Facial Features Extraction Based on Distance and Area of Points for Expression Recognition," 2019 4th Asia-Pacific Conference on Intelligent Robot Systems (ACIRS), pp. 211-215, 2019.

[4] A. A. Ewees, H. A. ElLaban and R. M. ElEraky, "Features Selection for Facial Expression Recognition," 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), pp. 1-6, 2019.

[5] E. Bagherian and R. W. O. K. Rahmat, "Facial feature extraction for face recognition: a review," 2008 International Symposium on Information Technology, pp. 1-9, 2008.

[6] H. Bahonar and N. M. Charkari, "Facial feature detection and extraction using symmetry and region-based deformable template matching," 2009 14th International CSI Computer Conference, pp. 664-669, 2009.

[7] N. Singh, A. K. Daniel and P. Chaturvedi, "Template matching for detection & recognition of frontal view of human face through Matlab," 2017 International Conference on Information Communication and Embedded Systems (ICICES), pp. 1-7, 2017.

[8] S. R. Benedict and J. S. Kumar, "Geometric shaped facial feature extraction for face recognition," 2016 IEEE International Conference on Advances in Computer Applications (ICACA), pp. 275-278, 2016.

[9] Y. Taigman, M. Yang, M.A. Ranzato, L. Wolf, “Deep face: Closing the gap to human-level performance in face verification,” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1701, 2014

[10] G. Klambauer, T. Unterthiner, A. Mayr, S. Hochreiter, “Self-normalizing neural networks,” In Neural Information Processing Systems, pp. 971, 2017.

[11] S. Li, L. Sun, X. Ning, Y. Shi and X. Dong, "Head Pose Classification Based on Line Portrait," 2019 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS), pp. 186-189, 2019.

[12] T. Yang, Y. Chen, Y. Lin and Y. Chuang, "FSA-Net: Learning Fine-Grained Structure Aggregation for Head Pose Estimation From a Single Image," 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1087-1096, 2019.

[13] Shahenda Sarhan, et al. “Multipose Face Recognition-Based Combined Adaptive Deep Learning Vector Quantization,” Computational Intelligence and Neuroscience, pp. 1-11, 2020.

[14] Lin, J., Xiao, L., Wu, T. et al.,, “Image set-based face recognition using pose estimation with facial landmarks,” Multimed Tools Appl., pp.19493–19507, 2020.

[15] Moein Razavi et al., “An Automatic System to Monitor the Physical Distance and Face Mask Wearing of Construction Workers in COVID-19 Pandemic,” arXiv:2101.01373, 2021.

[16] I. Q. Mundial, M. S. Ul Hassan, M. I. Tiwana, W. S. Qureshi and E. Alanazi, "Towards Facial Recognition Problem in COVID-19 Pandemic," 2020 4rd International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM), pp. 210-214, 2020.

[17] M. Z. Khan, S. Harous, S. U. Hassan, M. U. Ghani Khan, R. Iqbal and S. Mumtaz, "Deep Unified Model For Face Recognition Based on Convolution Neural Network and Edge Computing," in IEEE Access, vol. 7, pp. 72622-72633, 2019.

[18] J. -H. Kim, B. -G. Kim, P. P. Roy and D. -M. Jeong, "Efficient Facial Expression Recognition Algorithm Based on Hierarchical Deep Neural Network Structure," in IEEE Access, vol. 7, pp. 41273-41285, 2019.

[19] X. Liu, X. Cheng and K. Lee, "GA-SVM-Based Facial Emotion Recognition Using Facial Geometric Features," in IEEE Sensors Journal, vol. 21, no. 10, pp. 11532-11542, 2021.