Journal of Theoretical and Applied Information Technology
31st January 2022. Vol.100. No 2
© 2022 Little Lion Scientific

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

FACE RECOGNITION USING DEEP LEARNINGXCEPTION CNN METHOD

PALLAVARAM VENKATESWAR LAL[1] , UPPALAPATI SRILAKSHMI[2] ,D.VENKATESWARLU[3]

[1] Research Scholar of VFSTR Deemed to be University & Associative Professor of CSE Department, Narayana Engineering College, Guduru, A. P.

[2] Assistant Professor of CSE Department, VFSTR Deemed to be University, Vadlamudi, Guntur, A.P.

[3]Professor Of CSE Department, VFSTR Deemed to be University, Vadlamudi, Guntur, A.P.

E-mail: [1] venkateswarlal@gmail.com , [2]druppalapati2019@gmail.com, [3]drvenkateswarlud43@yahoo.co.in

ABSTRACT

The continual development of computer vision technology is one of the main research paths in the area of computer vision during recent years. It detects, tracks, recognizes, or authenticates human appearances from any picture or video taken through a digital camera and provides accurate and quick enough recognition functions for commercial use. Widely utilized in mobile payments, safe cities, criminal investigations, and other areas. Much research has progressed into face detection, identification and safety, the main problems are the consideration of those objects that had "different sizes" and "different aspects ratios" in a single framework that prevented or exceeded human level accuracy in human face appearance, such as noise in face images, opposite illumination,,Haar cascade was found to produce optimum accuracy while analyzing the multi focus faces using FERET database and the LFW database, and utilizes Xception (Depth Wise Separable). CNN is used for extracting the feature. Finally, classification is carried out. The suggested approach obtained FERET accuracy by about 96.73%, and LFW data by approximately 98.45%. The research findings have shown that the predicted technique exceeds existing methods.

Keywords: *Face Recognition, Haar cascade, LBP, CNN, Deep Learning, Artificial Neural Network*

1. INTRODUCTION

Due to the wide range of potentially sensitive applications, such as access control for physical and online sites in both commercial and military organizations, including ATM cash distributors, e- learning, information security, smart monitoring, etc., high security well-designed face identification systems have grown in focus. Despite the significant improvement in face recognition over the past few decades, facial images in unconstrained conditions continue to challenge the research community as the result of large intrapersonal variations such as changes in face expression, poses, light, old age and small interpersonal differences are changing. In addition to its incredible probable application, face recognition (FR) is extensively explored due to its potential usefulness [2,3,4]. Without further technology, people could identify face in a scenario using their inherent skills. It is extremely difficult to develop an automatic identifying system. The advancement of computer technology hardware and software removes the limit of difficulty. The difficulty of searching for face patterns is a significant concern due to the broad variety of aberrations. Face recognition is extensively utilized for hands-on equipment such as internet image exploration, access control, security enforcement, security, theater, home safety, individual e- commerce and health care. Many difficulties with the usage of such gadgets with various features such as variations in illumination, size, location, direction and position incoherence arise. In addition, the face appearance, face ornaments, masquerade, partial occlusion conditions alter the whole look, making it harder to identify faces [5]. Face recognition systems include two basic steps: extraction of features and categorization. The second phase depends on the first. The challenge of extracting and learning relevant and highly discriminating face characteristics is thus difficult in order to reduce intrapersonal variations and enhance interpersonal differences. Recently, an appealing, straightforward solution to the difficulties encountered is to change the position of the face in the pictures by generating new and unique points of view. This improves its characteristics and reduces the incoherence to be addressed by recognition systems. In the method suggested, instead of concentrating on form, the primary emphasis is on texture and color for efficient identification of the face. Color offers the visual properties for indexing and recovery of pictures and gives textual information on structural surface organization and image objects. To this aim, the descriptors of texture and color are retrieved from the pre-processed face pictures, according to which an effective classification is carried out utilizing supporting vector machines. The descriptors of texture and color are thus extracted that the dominant color, orientation, texture patterns and pictures are produced

2. RELATED WORKS

This section deals with some of the most significant facial recognition techniques introduced during the last five decades. A variety of methods to address all these limitations and issues in the face recognition system were suggested, implemented and improved in this respect. These methods may be split into two different categories: local approaches to design and profound learning. Moreover, local hand-made descriptor methods may be classified into four groups: holistic, educational and hybrid, characteristically based approaches[6]. Ouanan et al.[2,3,4], a FERET Facial Illustration dataset, has been proposed to use Gabor (GFs) and Zernike moments (ZMs) to minimize the ZM's texture and format characteristics. The geometric vector is derived by measuring and calculating the position and geometric links between the face properties such as the lips, eyes, nose and classifier of the structure. The EBGM system is an example of a functional approach which utilizes Gabor filter responses to extract a collection of local characteristics from each face feature in various orientations and frequencies [7,8].

In contrast to characteristic techniques, holistic methods typically extract the vector instead of the local geometric characteristics by working across the whole face. Autonomous approaches are the most famous examples of such methods: the major component analysis (PCA), the independent component analysis (ICA), and others[9]. The third learning approach learns from workout examples utilizing labeling techniques. Two or more of these groups are ultimately based on hybrid methods. Examples of category 3 and 4 may be found in [10– 11]. Previous studies have shown that local, expert methods employed as robust and selective face recognition detectors succeed even if there is little training per individual, as shown in [12-13]. In uncontrolled circumstances, however, the performance of the local handcrafted Desktop Descriptors drops significantly since built facial representations are extremely vulnerable to highly unilinear variables such as intra-personal expression, lighting, location and occlusion[14].

In tackling these challenges, an emphasis was put on the employment of profound learning methods (e.g. deep neural networks) for the automated learning of a number of effective representations of the hierarchy of nonlinear mapping which are able to handle nonlinear changes in face pictures. Moreover, systems employing deep learning approaches may readily extend to other new fields in contrast to handmade descriptor techniques[15]. DBN is one of the most widely utilized in-depth learning methods in a number of fields, including facial acknowledgement [16], language acknowledgement [17], audio categorization [18] and understanding of natural languages [19], to enhance hierarchy via unlabeled data. The main drawback of DBN is that the DBN function representations are sensitive to local input translations, when pixel intensity values are given directly to visible units. This may result in local characteristics that are deemed essential for face recognition being ignored in the picture. Furthermore, it is computationally costly and difficult to use the DBN for pictures of actual dimensions (e.g. 128 pp128).

A new framework to integrate the benefit of local, handmade functional descriptors with DBN is created to address the issue of facial identification under uncontaminated circumstances to enhance theability to generate and decrease the cost of DBN computation. The

Journal of Theoretical and Applied Information Technology
31st January 2022. Vol.100. No 2
© 2022 Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

use of DBN as a technique of guiding the learning process by using pictures pre- treated rather than by pixel intensity (raw data) greatly improves DBN's ability to learn more discriminating features with less time to achieve the final model. In literature, relatively few research address the possibility to integrate DBN to the best possible knowledge of the authors in addition to pre-processed image feature representations. Huang et al. [20] have demonstrated that thecoevolutionary DBN and LBP output may increase the accuracy of the final system. Li et al.[21] have similarly ended up applying the DBN symmetrical local binary pattern to the center (CS-LBP). The work in [20] was solely used for the face control job, work in [22] was measured on a very limited facial data set utilizing facial pictures collected in controlled settings.

Kamenca et al. (2017),[23] check three renowned imaging techniques, such as the PCA, the Histogram for Local Binary Patterns (LBPH), and the KNNis, as shown by the neural network (CNN).PCA, LBPH, KNN and recommended CNN detection is proven to be fully accurate. The whole experiment was carried out using an ORL dataset and the findings were presented and computed. Thisset of data consists of 400 separate entities (40 categories/10 pictures per category). The data showed that LBPH delivers superior results than PCA and KNN. The test results from the ORL data sets show the effectiveness of the proposed face identification system. The maximum identification accuracy of 98.3 percent is obtained for the given CNN. This approach is dependent on CNN, which goes beyond previous techniques. A novel way of utilizing a deep neural facial recognition system is to use Gupta et al. (2018)[24]. This method provides just the recovered face characteristics instead of raw pixel data. This reduces the complexity and gives Yale faces 97.05 percent accuracy. Ong et al. (2018)[25], a virtual sample for the closest generation of virtual samples (kNVSG), which improves intra- class sample variance data. We also offer a multi-faceted discriminatory learning technique (ISMMDL) to utilize sample data produced by the method of kNNVSG. By using kNNVSG and ISMMDL, we are presenting the closest virtual image learning k for single samples facial recognition (SSFR) tasks based on discriminating (kNNMMDL).

As a result of the above-mentioned existing methods, the most significant issues occurred with objects that had "different sizes" and "different aspect ratios" in a single framework that prevented or exceeded human-level accuracy in the appearance of human faces, such as noise in face images and the opposite illumination. It is possible to overcome this drawback by using the Haar cascade to achieve the highest accuracy possible while evaluating the multi-focus faces using the FERET database and the LFW database, as well as by utilizing the Xception software (Depth Wise Separable). The characteristics are extracted with the help of CNN. After that, the categorization process begins. FERET accuracy was achieved using the provided technique.

1. PROPOSED APPROACH

At this part, a suggested approach for robust face recognition is presented in every step utilizing important methods. Depending on these descriptors the combination of various descriptors of the picture characteristics is extracted and categorized.

The suggested approach is primarily divided into four major phases to strengthen the ensemble. The four distinct steps include picture pre-processing, multi focus face recognition, extraction and categorization of features. The diagram of the proposed approach is shown in Fig. 1. The detailed explanation of each method utilized in each stage is described briefly in the following paragraphs

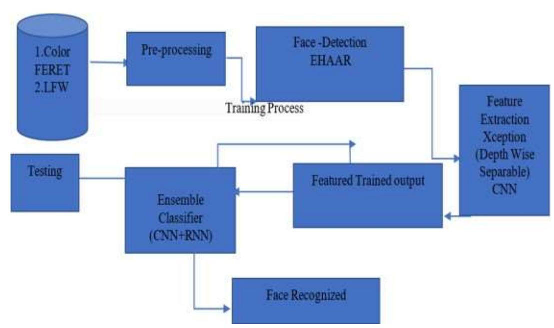

Fig.1. Proposed Framework EEHAAR_Xcla(Embedding Enhanced HAAR Cascade With Xception Classifier) For The Face Recognitionapproach

1.1 Preprocessing

Before a multi-focus face detection phase, the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

problem of lighting discrepancy must be solved by using various techniques of improvement presently introduced: Normalization of isotropic smoothing (ISN) (Heusch et al. 2005) [26] The problem of face verification in all radiance via isotropic normalization is used in this approach (a dissemination phase which basically updates each pixel using its neighboring intensity value mean, regardless of the picture data near the region under discussion). Faces of gradient (GFs) (Zhang et al. 2009) [27] This is not an adequate improvement method, however, an illuminating impermeable measure derived from the picture gradient, which isstrong enough in unrestricted natural lighting atmospheres to illuminate variances. In this article gradient faces are used to indicate a picture in the gradient space comparable to preprocessing technology.

1.2 Multi focus Face Detection

Classifier Haar Cascade Object detection is an object detection technique by Paul Viola and Michael Jones using a Haar-based cascade classifier. In 2001, they presented a study called & quot ;Rapid Object Detection using a Simple Feature Boosting Cascade." Haar Cascade is a series of Haar-like features integrated in a classifier. The feature is the sum of white pixels removed from the dark region pixels. The basis of the facial sensor is 24 to 24. There are around 160,000 Haar-like features from that basic facial detector. However this functionality is not completely utilized. The Haar cascading image is shown in Figure 2.

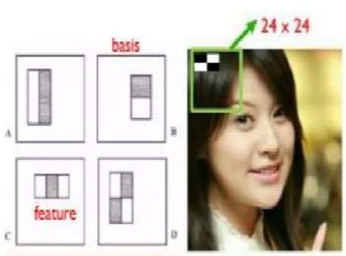

Fig 2. Illustration of Haar Cascade

Figure shows that the characteristics (A) and (B) are two rectangles, whereas the characteristics (C) are three rectangles, and the characteristics of (D) are four rectangles. In the meanwhile, the Haar Cascade method is an ML-based technique, with many positive and negative pictures forming a cascade function. It is then utilized in other pictures to detect things. The method consists of four phases: haar selection, integrated photo creation, AdaBoost training, cascade classifications, as illustrated in Figure 2 and in [3].

Pseudo code for. Pseudo code for. Algorithm for haar-cascade

1.3 Phases Of Recognition

- Pick f (maximum acceptable false positive rate per layer) and d (minimum acceptable detection rate per layer)
- Lets F_{target} is target overall false positive rate
- Lets P is a set of positive examples
- Lets N is a set of negative examples
- Lets $F_0 = 1$, $D_0=1$, and i=0 (F_0: overall false positive rate at layer 0, D_0: acceptable detection rate at layer 0, and i: is the current layer)
- While $F_i > F_{target}$ (F_i: overall false positive rate at layer i):
  - i++ (layer increasing by 1)
  - $n_i=0$; $F_i = F_{i-1}$ (n_i: negative example i):
  - While $F_i > f*F_{i-1}$:
    - n_i ++ (check a next negative example)
    - Use P and N to train with AdaBoost to make a xml (classifier)
    - Check the result of new classifier for F_i and D_0
    - Decrease threshold for new classifier to adjust detection rate r >= $d*F_{i-1}$
  - N = empty
  - If $F_i > F_{target}$, use the current classifier and false detection to set N

Detection, feature extraction and comparison are the three major steps performed by a face recognition system.

1.3.1 Face detection
1.3.2 The Viola Jones algorithm is a useful technique for
1.3.3 face detection. In general, this method may be used
1.3.4 for various stiff structured object identification
1.3.5 problems not just for the facial detection. The
1.3.6 Viola-Jones method consists of three major ideas
1.3.7 for developing a real time face-detector: haar-like
1.3.8 characteristics, integrated images, Adaboost
1.3.9 training and Cascading classifier. The system can
1.3.10 detect the presence or absence of a human face by
1.3.11 applying these characteristics.

1.3.12 Haar-like features

Haar-like characteristics are utilized for human face recognition using Haar cascade classifier. There are three Haar -like structures. The first format in Fig.4 is the edge functions, the second type is the line function, and the third type is the four-rectangle function. Using the integrated picture, the concept of haar like provides quick

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

calculation. It is referred to as haar-like characteristics [32].

The algorithm searches for a particular facial haar characteristic. This detection transforms the picture into a 24X24 window and pixels each haar feature into that window. The method first needs a large number of positive (facial pictures) and negative (facial images) images to train the classifier. These characteristics are then extracted. Features are numerical values derived from pictures that differentiate between one image and another, each feature is a value gained by subtracting the sum of the pixels under the white rectangle from the amount of the pixels under the black rectangle[33]. rectangle, and ii(1) is the integral image of the

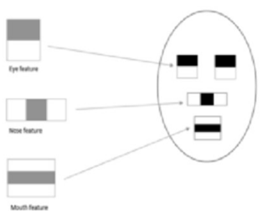

Fig 3. Types Of Haar-Like Features

Feature = Σ (pixels in black area dark) - Σ (pixels in white area white)

A great many characteristics are calculated in all conceivable sizes and positions of each kernel. Over 160,000 characteristics arise from a 24-fold window. For each function computation, the total of the pixels beneath the white and black rectangles must be found. To this end, the idea of the integral picture and adaboost algorithm is used, reducing 160000 to 6000 [32] features.

1.3.13 Integral Image

Rectangle characteristics may be quickly identified via an intermediate picture image called the integral image. The integral image includes the depiction of a particular picture in tiny components.

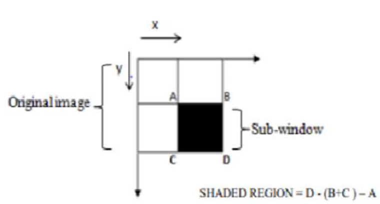

Fig 4. Integral Image Schematic Diagram

For instance, at position 1 the value of this integral picture is the total of pixels in rectangle A. At position 2 the value is A + B and so on. The sum ofrectangular D pixels is:

$$S(D) = ii(4) - (ii(3) + ii(2)) + ii(1) \qquad (2)$$

Where S(D) is a total of pixels of rectangular D only - ii(2) is an integral image of the A+B A+C rectangle (ii(3) is an integral image of A+B), and ii(1) is an integral picture of the A rectangle (addition is performed because the A region is subtracted twice in ii(3) and ii(2)).

The integral image is outlined as:

$$ii[x, y] = i[x', y'] \qquad (3)$$

Where, ii[x, y] represents integral image, and i [x', y'] represents original image. [32] The pixel value of integral images at any (x,y) location is the sum of all pixel values displayed before the current pixel. The integral value of an individual pixel is the sum of pixels on the top and the pixel towards the left. For example

5	9	12	20	23
8	21	25	35	44
17	36	40	55	71
24	46	56	76	101
25	49	61	89	117

Fig 5. Input Image

5	4	3	8	3
3	9	1	2	6
9	6	0	5	7
7	3	6	5	9
1	2	2	8	3

Fig 6. Integral Image

The picture is integrated in less pixels since the traverse starts from the top left to the bottom right. This makes the computation of the addition of all pixels utilizing just four values inside each defined rectangle. These values are in the integral picture the pixels similar to the edges of the rectangle of the image

3.3.4. AdaBoost Learning

AdaBoost is an adequate boosting algorithm which combines weak classifiers while reducing significantly not only the training error but also the more elusive

© 2022 Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

generalized error. The main idea of Boosting lies in connecting the simple classifiers which are known as weak classifiers. Since the weak classifiers do not even anticipate the best

one feature in order to readily connect the haar characteristics with a weak classifier. The haar-like feature is utilized as a threshold in Viola and Jones' AdaBoost learning method. The Haar Classifier is the most powerful because it has the most powerful features. The positive and negative samples are best divided into the function. In order to build a strong final classifier AdaBoost is used [34]. It reduces the features from 160000 to 6000, thus making the computation simpler and hence it is less in computational complexity.

3.3.5 Cascade Classifier

Cascade classification is a cascade of weak devices intended to enhance the face recognition process and decrease computer complexity. Each node in the series includes a week haar classifier and filter. AdaBoost gives the weights of the nodes and usually reaches the weighted node. If parts of the picture are not permitted in a filter, the sub windowof the particular image is removed. It is thus regarded a non-face without the face identified in the picture areas analyzed. This is extremely important for classifier performance since in the first phase all or all negative picture sub-windows are removed. On the other hand, if picture areas have crossed the filter successfully, they proceed with a more sophisticated filter to the next step. Only areas that pass through all filters are treated as facial matches. This shows that the regions of the image contain the subject of the detected face. The aim of the multi-stage categorization is to rapidly and efficiently eliminate non-facewindows. The classifier is used to refuse additional false positive sub-windows (non-face regions). The number of false positive rates has been significantly decreased after several processing phases [34].

FEATURE EXTRACTION

Depth wise separable convolutionregular convolutions:

- Examine concurrently at both channel andspatial correlations

Deep, Separable Convolution:

- Examine in consecutive stages at canal andspatial connections separately
- Spatial convergence: 3x3 convolutions perchannel

classification function to correctly categorize the data, they are termed weak classifiers. A classifier is coupled here with

- Depth wise convolution: 1x1 concatenationof channels

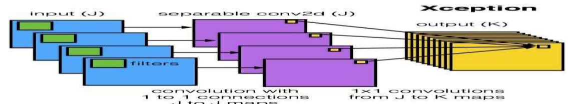

Fig 7. Depth Wise Separable Convolution

Example: Take 3x3 layers of convolution on 16 input and 32 output channels.

- normal convolution: 16x32x3x3 = 4608.
- Depth wise convolution separable:
  (Spatial conv + depth wise conv) = (16x3x3 + 16x32x1x1) = 656 parameters
- much decreased count of parameters
- more efficient complexity
- Cross-channel features are maintained

Xception architecture

- A neural architecture based only on deep, separable convolution layers.
- Basic hypothesis: the mapping of cross-chain correlations and the separation of spatial correlations.
- Consists of 36 convergence layers that provide the foundation for network extraction.
- It consists of 14 modules with linear residual connections, excluding the first and last modules.

We propose a neural architecture based only on profoundly knowledgeable convolution layers. We postulate that the mapping of cross-chain correlation and spatial correlation in feature maps of convolutionary neural networks may be fully decoupled. Since this is a stronger version of the theory behind the design of Initiation, we call our suggested architecture Xception, which is known as "Extreme Inception." A full explanation of the network requirements is provided in Figure 8.

© 2022 Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

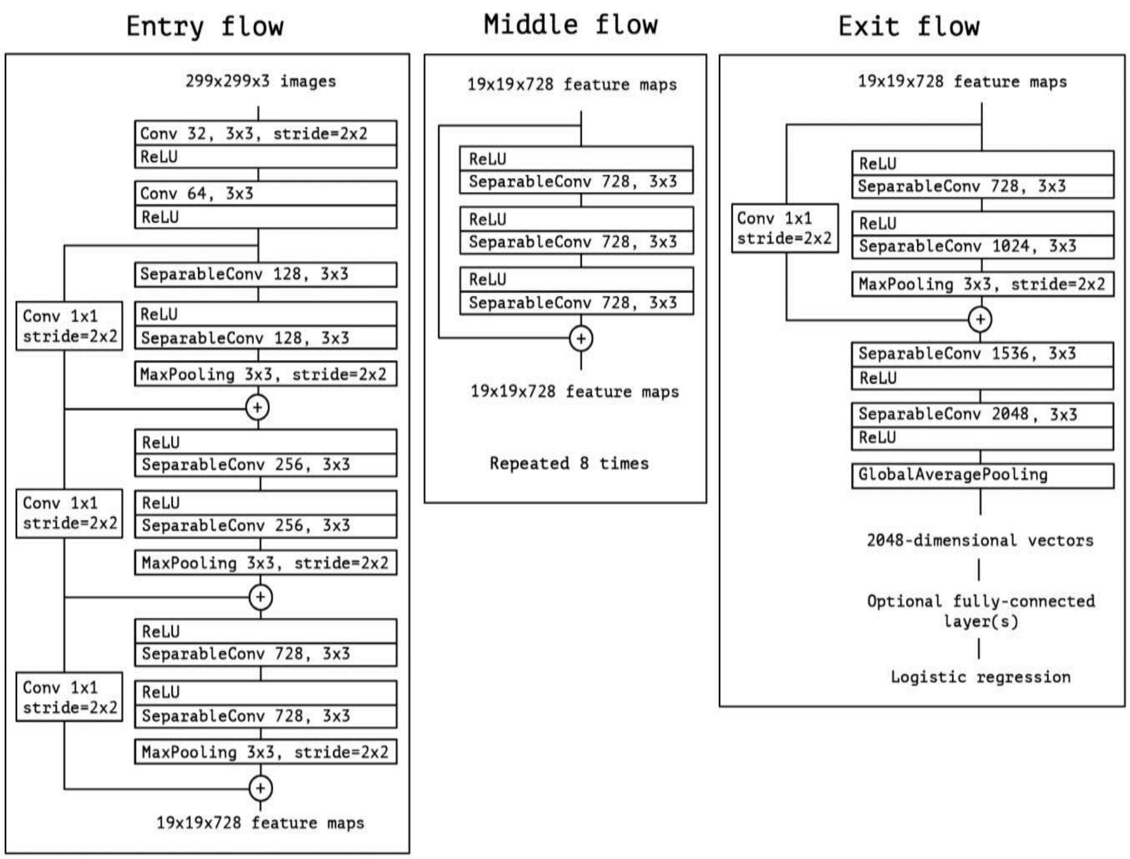

Fig 8. Architecture Of Xception

The Xception architecture includes 36 convolutionary layers that constitute the network's functional extraction foundation. In our experimental assessment, we will examine solely picture classification and thus a logistic regression layer follows our convolutionary foundation. Optionally, completely connected layers may be added in the experimental assessment phase before the logistic regression layer. The 36 convolution layers include 14 modules, all of which have a linear residual link with the exception of initial and final modules. Xception architecture[36] is a linear stack of convolutionary layers that are profoundly distinct and have long-lasting connections. This makes it easy to build and architecture: you only need 30 to 40 lines of codes utilizing high- level libraries like Keras[2] and TensorFlow-Slim[17], which are no different from the VGG-16[35], but are more complex than designs like Inception V2 or V3. The Keras Applications Module2 MIT license offers the open-source Xception implementation using Keras and TensorFlow..

3.4 Image Classification

3.4.1 The CNN Module

It is capable of extracting functions. The problem of diminishing gradients becomes increasingly apparent as the depth of the CNN structure increases. One of the better solutions to this issue is dense DenseNet connections [38]. However, many connections may increase the channel size. To address too many parameters issues, the author uses a 1 ∗ 1 convolution structure at the front of each dense block layer, thus the output of each layer is the same, regardless of how many channel inputs exist. Then the network parameters of DenseNet are substantially reduced. This dense connection is also similar to the fact that each layer directly connects input and loss to decrease the loss function in calculating the disappearance problem. In this study we used DenseNet as the RGB medical image extraction module on three channels.

3.4.2. Switchable Normalization

Switchable standardization technology, newest method to standardization. The standardization technique is not developed, but includes normalizing of current instances, layer standardization and batch normalization. If the model's map format is N, C, H, W, where N is the batch size, C is the number of channels, H is the map height, and W is the map width. Each pixel has the hncij and the ncij subscript shows the index of each dimension. Each pixel is then standardized utilizing switchable normalization and normalizing of the pixel value

$$\hat{h}_{ncij} = \gamma \frac{h_{ncij} - \sum_{k \in \Omega} \omega_k \mu_k}{\sqrt{\sum_{k \in \Omega} \omega'_k \sigma_k^2 + \epsilon}} + \beta \quad (4)$$

Where γ represents the coefficients of scaling while β denotes the coefficient of deviation that is comparable to other

$$Accuracyrate = \frac{TruePositive + TrueNegative}{TotalInstances} * 100$$

normalizing formulations. In switchable normalization the major distinction is the mean μ and the variance σ 2. These techniques are not computed at layers or channels; nevertheless, in a set of Ω, the set of Ω is BN, LN and IN three standardization methods. The weighted average approach is similar to three standardization techniques with a focus mechanism. A switchable standardization method may enable a model during the training phase to learn the optimum standardization for each layer[42]. We have used this unique, switchable standardization approach in the

$$Recall = \frac{TruePositive}{TruePositive + Falsenegative}$$

CNN module instead of a traditional batch normalizing methodology.

3.4.3 The module RNN

The spatial properties of a picture are widely understood. From the pixel level point of view, the pixels in the picture are the time sequence. If the width and height of the picture are measured as the time stage, then the time relation for each pixel row in each image may be investigated. Therefore, we considered both the spatial connection and the temporal sequence between the pixels while constructing the model. For removing pixel sequence characteristics from the RNN module, we utilized LSTM stacked[37].

4. EXPERIMENTAL RESULTS

Experiments with Python 3.6.5 version on two benchmark data sets, such as FERET (Phillips et al. 2000)[28] and LFW (Huang et al. 2007)[29], are recommended as the face recognition system. A comparison of the proposed facial recognition method is made utilizing Sparse (Previous Paper), Texture Ensemble (Lumini et al. 2017) [30] and Face Recognition based on Improved Robust Sparse Coding Algorithm (Jun-Kai 2015). [31].

The performance analysis of the method suggested is shown in the next paragraph. The assessment parameters are the accuracy, the precision, the reminder and finally the F1 score. The confusion matrix is used to compute the different performance measurements. The data sets utilized for the evaluation of the technique proposed are FERET and LFW. The model performance is evaluated randomly by selecting the test data from the data set as the result data.

Accuracy

It properly displays the proportion of categorized occurrences during categorization. It is assessed as

Precision

This metric is defined as the number of genuine positive plus the number of false positive. This is used to assess the quality and accuracy of the FERET and LFW data as indicated below:

$$Precision = \frac{Truepositive}{Truepositive + FalsePositive}$$

Recall

It is the percentage of actual positives which is accurate and defined as projected positives

F1 Score

It is calculated by accuracy and the test results are recalled. The F1 score is a test accuracy indicator to evaluate the binary categorization. Where the precise results are split by the number of positive results, the number of true positive results is recalled, divided by the number of positive results. It is computed according to:

$$F1 - Measure = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

Confusion Matrix

The table is often used to calculate the output of a classification model from the collection of test data known as the Confusion matrix for the true values. The predicted performance of our

Journal of Theoretical and Applied Information Technology
31st January 2022. Vol.100. No 2
© 2022 Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

confusion-matrix classification model is thus displayed in Figure 9.

4.1 Descriptions of data set

Database FERET This FERET sample consists of 5 data samples: Fc (194 pictures), Fb (1195 images), Fa (1196pictures), Dup2, and Dup1 (722 images). The standard FERET estimate method includes pictures matched to each image within a dataset in the validation group. In this experiment, the complete pictures of the FERET Grey scale are linked with 110 to 1110 pixels in real-life views. Database LFW This [54] collection includes 13,233 pictures of 5,749 online celebrities.

In all, 1680 faces appear in more than two pictures. In the LFW data sample, two perspectives are provided. The first view is a validating group of 2200 face pairings and a validating group of 1000 facial pairs that are used just for selecting the pattern. View 2 consists of 10 non- overlapping 600 matches and is designed to report performance.

Comparison of results

Tables 1 and 2 experiments were conducted to estimate different descriptors when combined with multifocal face detection and function extractor methods. It demonstrated that the combination with each function descriptor in a similar column of the whole technique of multifocal facial detection by sum rule. The findings in Tables 1 and 2 show clearly that the integration accomplished with each feature descriptor method by merging the whole multi focus face detection provides superior results compared to a good single multi focus face detection with every feature descriptor strategy for each descriptor. The second interesting result is that this is a good descriptor The experiment reviewed in Table 2 with FERET data sets and the experiment evaluated in Table 3 using LFW data sets are discussed. Table 4 illustrates the average comparative classification accuracies. The findings in Table 4 clearly indicate that from the methods already discussed, the suggested methodology is more accurate in its categorization for both datasets compared to the current one. It is concluded that the recommended pre-processing methods and color descriptors have improved face recognition efficiency. Figure 2 shows the average CNN RNN Ensemble classification accuracy of both datasets. The proposed technique, in combination with the current Ensemble texture descriptors and the Dominant Color structure for facial identification, provides better precision values

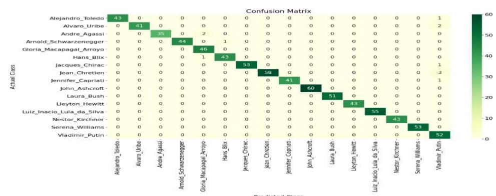

Fig 9. Confusion Matrix For The Proposed EEHAAR_Xcla Approach For COLOR FET Dataset

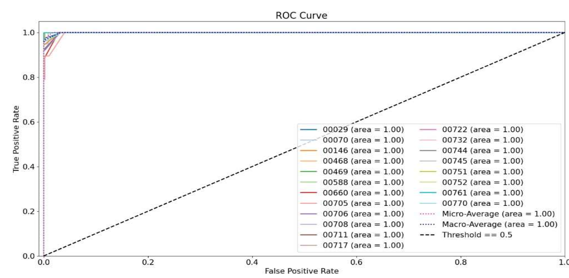

Fig 10.ROC curve for the proposed EEHAAR_XClA approach for COLOR FET dataset

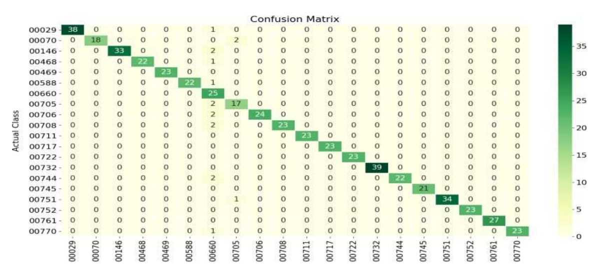

Fig 11. Confusion matrix for the proposed EEHAAR_XClA approach for LFW dataset

© 2022 Little Lion Scientific
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

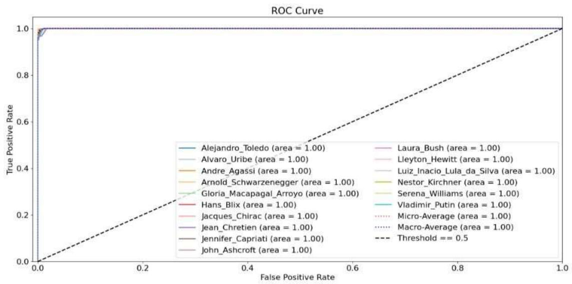

Fig 12. ROC curve for the proposed EEHAAR_XClA approach for LFW dataset

Table 1: Performance metrics of the Proposed EEHAAR_XClA approach

Dataset	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F-Measure(%)**	**AUC**
ColorFERET	96.73	97.73	96.73	96.99	99.94
LFW	98.45	98.61	98.45	98.48	99.99

Table.2: Accuracies of combining the pre-processing approaches with or without multi focus face detection for LFWdatasets

Pre-processing approaches (without face detection)	**Feature Descriptors**				
AS,AR,DG,DCT,OLHE,MSR,ISN,PN,GF	HOG	ICS_LBP	SHIFT	Dominant Color structure	EEHAAR_XClA
	87.6	86.4	83.8	87.3	
Pre-processing approaches (with face detection)	-	-	-	-	98.45

Table.3. Accuracies of combining the pre-processing approaches with or without face detection for FERET data sets

AS,AR,DG,DCT,OLHE,MSR,ISN,PN,GF	HOG	ICS_LBP	SHIFT	Dominant Color structure	EEHAAR_XClA
	88.3	85.6	82.6	85.5	
Pre-processing approaches (with face detection)	-	-	-	-	96.73

Table.4. Comparison for the Average Classification for the specific dataset

Reference	LFW	FERET
Juefei-Xu et al. 2015[31]	87.55	-
VenkateswarLal et al.2019[5]	92.5	91.23
Proposed work	**98.45**	**96.73**

From the comparative analysis of the above tables, it is clear that the proposed method for face recognition using the deep learning xception-cnn method shows the better results.

5.CONCLUSION

In this work, an ensemble of descriptor-based facial recognition technology suggests that facial pictures are classified in any natural setting. The form, texture and color of face pictures often play an important part in the extraction of important features for precise and effective categorization. For this aim, the suggested approach covers descriptors such as HOG, ICS LBP, SHIFT and color-dominant descriptors for color, texture, and orientation. Wild data sets display face pictures using rigorous formalization methods that always try to capture the frontal element of a picture.

In addition, several pre-processing methods are used to get a picture without noisy components. Color and texture characteristics are used to identify the face picture using a vector machine. The test results are obtained by utilizing FERET and LFW data sets and the findings show that the method provided produces effective results in conjunction with prior approaches and has an exceptional mean classification accuracy. Both samples have an average classification accuracy of 99% and 94% respectively. In future work, it needs to focus on the accuracy of different data sets.

REFERENCES:

[1] Bhowmik, M.K., Bhattacharjee, D., Nasipuri, M., Basu, D.K., Kundu, M.: Quotient based multiresolution

[2] Ouanan H, Ouanan M, Aksasse B (2015a) Gabor–Zernike features based face recognition scheme. Int J Imaging Robot 16(2):118–131

[3] Ouanan H, Ouanan M, Aksasse B (2015b) Face recognition by neural networks based on

Gabor filters and random projection. Int J Math Comput 26(12):30–42
[4] Ouanan H, Ouanan M, Aksasse B (2015c) Gabor-HOG features based face recognition scheme. TELKOMNIKA Indones J Electr Eng 15(12):331–33.
[5] VenkateswarLal, Pallavaram, Gnaneswarar Rao Nitta, and Ande Prasad. "Ensemble of texture and shape descriptors using support vector machine classification for face recognition." Journal of Ambient Intelligence and Humanized Computing (2019): 1-8.
[6] Parmar, D.N., Mehta, B.B.: Face recognition methods & applications. Comput. Technol. Appl. 4(1), 84–86 (2013)
[7] Jafri, R., Arabnia, H.R.: A survey of face recognition techniques. J. Inf. Process. Syst. 5(2), 41–68 (2009)
[8] Imtiaz, H., Fattah, S.A.: A curvelet domain face recognition scheme based on local dominant feature extraction. ISRN Signal Process. 2012, 1–13 (2012).
[9] Shreeja, R., Shalini, B.: Facial feature extraction using statistical quantities of curve coefficients. Int. J. Eng. Sci. Technol. 2(10), 5929–5937 (2010)
[10] Zhang, B., Qiao, Y.: Face recognition based on gradient gabor feature and efficient Kernel Fisher analysis. Neural Comput. Appl. 19(4), 617–623 (2010)
[11] Cao, Z., Yin, Q., Tang, X., Sun, J.: Face recognition with learningbased descriptor. In: Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2707–2714 (2010)
[12] Simonyan, K., Parkhi, O., Vedaldi, A., Zisserman, A.: Fisher vector faces in the wild. In: Proceedings of British Machine Vision Conference 2013, pp. 8.1–8.11 (2013)
[13] Al Ani, M.S., Al-Waisy, A.S.: Face recognition approach based on wavelet–curvelet technique. Signal Image Process. Int. J. 3(2), 21– 31 (2012)
[14] Chai, Z., Sun, Z., Mendez-Vazquez, H., He, R., Tan, T.: Gabor ordinal measures for face recognition. IEEE Trans. Inf. Forensics Secur. 9(1), 14–26 (2014)
[15] Li, Z., Gong, D., Li, X., Tao, D.: Learning compact feature descriptor and adaptive matching framework for face recognition. IEEE Trans. Image Process. 24(9), 2736–2745 (2015)
[16] Sun, Y., Wang, X., Tang, X.: Deep learning face representation by joint identification–verification. In: Advances in Neural Information Processing Systems, pp. 1988–1996 (2014) 13.
[17] Lee, H., Grosse, R., Ranganath, R., Andrew, Y.N.: Unsupervised learning of hierarchical representations with convolutional deep belief networks. Commun. ACM 54(10), 95–103 (2011)
[18] Liu, J., Fang, C., Wu, C.: A fusion face recognition approach based on 7 -layer deep learning neural network. J. Electr. Comput. Eng. 2016, 1–7 (2016). doi:10.1155/2016/8637260
[19] Fousek, P., Rennie, S., Dognin, P., Goel, V.: Direct product based deep belief networks for automatic speech recognition. In: Proceedings of IEEE International Conference on Acoustics Speech and Signal Processing ICASSP 2013, pp. 3148 –3152 (2013)
[20] Lee, H., Pham, P., Largman, Y., Ng, A.: Unsupervised feature learning for audio classification using convolutional deep belief networks. In: Advances in Neural Information Processing Systems Conference, pp. 1096–1104 (2009).
[21] Huang, G.B., Lee, H., Learned-Miller, E.: Learning hierarchical representations for face verification with convolutional deep belief networks. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2518 – 2525 (2012)
[22] Li, C., Wei, W., Wang, J., Tang, W., Zhao, S.: Face recognition based on deep belief network combined with center-symmetric local binary pattern. Adv. Multimed. Ubiquitous Eng. Springer Singap. 354, 277–283 (2016).
[23] Kamencay P, Benco M, Mizdos T, Radil R (2017) A new method for face recognition using convolutional neural network. Digital Image Process Comput Graphics 15(4):663–672.
[24] Gupta P, Saxena N, Sharma M, Tripathia J (2018) Deep neural network for human face recognition. IJ Eng Manuf 1:63–7.
[25] Dong X, Wu F, Jing XY (2018) Generic training set based multimanifold discriminant learning for single sample face recognition. KSII Trans Internet Inf Syst 12(1):368–39.
[26] Heusch G, Cardinaux F, Marcel S (2005) Lighting normalization algorithms for face verification. IDIAP, p 9.
[27] Zhang T et al (2009) Face recognition under varying illumination using gradient faces. IEEE Trans Image Process 18(11):2599–

2606.

[28] Phillips J et al (2000) The feret evaluation methodology for face recognition algorithms. IEEE Trans Pattern Anal Mach Intell 22:1090–1104.

[29] Tan X, Triggs B (2007) Enhanced local texture feature sets for face recognition under difficult lighting conditions. Anal. Model. Faces Gestures, vol 4778. LNCS, pp 168–182.

[30] Lumini A, Nanni L, Brahman S (2017) Ensemble of texture descriptors and classifiers for face recognition. Appl Comput Inform 13:79–91.

[31] Juefei-Xu F, Luu K, Savvides M (2015) Spartans: single-sample periocular-based alignment-robust recognition technique applied to nonfrontal scenarios. IEEE Trans Image Process 24(12):4780–4795.

[32] Willberger. (Deep learning Haar – cascade explained.) Internet: www.willberger.org/cascade-haar-explained, Jan 13, 2018.

[33] Ayman A. Wazwaz, Amir O. Herbawi, Mohammad J. Teeti, Sajed Y.Hmeed. “Raspberry-Pi and Computers- Based Face Detection and Recognition System”, 4th International Conference on Computer and Technology Applications, pp. 01- 03, 2018.

[34] S. Guennouni, A. Ahaitouf and A. Mansouri, "Face detection: Comparing Haar-like combined with cascade classifiers and Edge Orientation Matching," 2017 International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS), 2017, pp. 1-4, doi: 10.1109/WITS.2017.7934604

[35] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

[36] Chollet, François. "Xception: Deep learning with depthwise separable convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[37] Wang, J.; Peng, B.; Zhang, X. Using a Stacked Residual LSTM Model for Sentiment Intensity Prediction. Neurocomputing 2018, 322, 93–101. [CrossRef].

[38] Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.Cancers 2019, 11, 1901 14 of 14.

[39] Luo, P.; Ren, J.; Peng, Z.; Zhang, R.; Li, J. Differentiable Learning-to-Normalize via Switchable Normalization. In Proceedings of the International Conference on Learning Representation (ICLR), New Orleans, LA, USA, 6–9 May 2019.