I.J. Image, Graphics and Signal Processing, 2023, 5, 60-76
Published Online on October 8, 2023 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijigsp.2023.05.05

Face Recognition Using Modified Histogram of Oriented Gradients and Convolutional Neural Networks

Raveendra K*
Government Engineering College, K R Pet, India
E-mail: raveendrakit@gmail.com
ORCID iD: https://orcid.org/0000-0003-3629-2407
*Corresponding Author

Ravi J
Global Academy of Technology, Bangalore, India
E-mail: gnkereravi@gmail.com
ORCID iD: https://orcid.org/ 0000-0002-1488-9547

Khalid Nazim Abdul Sattar
Majmaah University, Al Majmaah 11952, Saudi Arabia
E-mail: k.sattar@mu.edu.sa
ORCID iD: https://orcid.org/0000-0002-0759-0512

Received: 10 February 2023; Revised: 15 April 2023; Accepted: 11 May 2023; Published: 08 October 2023

Abstract: We are aiming in this work to develop an improved face recognition system for person-dependent and person-independent variants. To extract relevant facial features, we are using the convolutional neural network. These features allow comparing faces of different subjects in an optimized manner. The system training module firstly recognizes different subjects of dataset, in another approach, the module processes a different set of new images. Use of CNN alone for face recognition has achieved promising recognition rate, however many other works have showed declined in recognition rate for many complex datasets. Further, use of CNN alone exhibits reduced recognition rate for large scale databases. To overcome the above problem, we are proposing a modified spatial texture pattern extraction technique namely modified Histogram oriented gradient (m-HOG) for extracting facial image features along three gradient directions along with CNN algorithm to classify the face image based on the features. In the preprocessing stage, the face region is captured by removing the background from the input face images and is resized to 100×100. The m-HOG features are retrieved using histogram channels evenly distributed between 0 and 180 degrees. The obtained features are resized as a matrix having dimension 66×198 and which are passed to the CNN to extract robust and discriminative features and are classified using softmax classification layer. The recognition rates obtained for L-Spacek, NIR, JAFFE and YALE database are 99.80%, 91.43%, 95.00% and 93.33% respectively and are found to be better when compared to the existing methods.

Index Terms: Face recognition, Convolutional Neural Network, softmax, deep learning, HOG, L-Spacek, NIR, JAFFE and YALE Feature extraction ·Image database Preprocessing ·Recognition accuracy.

1. Introduction

The issue of face recognition has been intensively researched in the fields of computer vision and pattern recognition. Automatic face recognition is utilized in a broad variety of practical applications, such as security, surveillance, identity verification, intelligent vision monitoring and immigration cleansing systems. Depending on the application environment, it may be divided into two functions: facial examination and facial recognition. The first objective is to assess whether or not a pair of facial images belongs to the same person, while the second objective is to learn about the individual by examining a gallery of facial images and identifying identical ones. However, face recognition in real-world applications is still a challenge [1]. The main reason for this is because the face is not stable and has different facial features, ages, multiple angles, and, most importantly, changing light intensity. Furthermore, several parameters, such as closure and posture, have an impact on face recognition performance.

In the last few years, computer-based learning has become more popular. Over the last decade, many perspectives have been gained from computer vision issues and discriminating solutions, such as image classification, object identification and face recognition [2, 3].

CNN can be trained using effective features derived from recurrent pooling and convolution operations on a large database. Currently, CNN is emerging as a powerful algorithm and has become most sufficient for dense problem predictions, especially for face recognition problems. From the past decade many algorithms have been proposed which are based on CNNs [2, 3] shows accuracy improvements. The size of training database images for CNNs should be increased for optimal accuracy.

Data addition (DA) is a proposed technique that uses domain-specific synthesis to deliberately infect a database in order to add more unchanged cases [4]. Furthermore, it is a group of low-cost computer algorithms that were previously utilized to reduce the usage of CNN [5]. However, large data sets contain high noise signals or variations, especially if they are automatically collected from image or from movie sources.

The main objective of the work is to develop smart and efficient face biometric authentication system by hybridizing texture feature extraction technique (HOG) with deep learning algorithm (CNN) and thereby increasing recognition rate. Use of CNN alone for face recognition has achieved promising recognition rate, however many other works have showed declined in recognition rate for many complex datasets. Further, use of CNN alone exhibits reduced recognition rate for large scale databases. The authors of some of the articles have followed DA methods to improve training data to create more examples of deep training demand.

To overcome the above problem, our method consists of hybrid approach which contains application of CNN for spatial feature matrix obtained using m-HOG a feature extraction technique. The m-HoG features are extracted for input face detected image. We present a collection of face images from different databases. In fact, we apply our model with different datasets having color variations, fading, contrast and brightness change. Performance analysis is performed by assessing parameter like recognition rate. Based on the results of the experimental analysis, our proposed model performed well compared to the existing models.

In this work, an explanation of the system description and methodology consists of three parts: m-HOG feature extraction, the use of CNN and the extraction of features and their comparisons have all been explored. The remaining of the manuscript is organized as follows: Section 2 presents the literature review. Section 3 describes the suggested technique in detail. The experimental results are summarized and discussed in Section 4. Section 5 concludes the manuscript.

2. Literature Review

Gaili Yue and Lei Lu [6] developed a face recognition approach that makes use of histogram equalization and a CNN. Preprocessing of the face image is carried out by histogram equalization. CNN is constructed with Google deep learning framework TensorFlow1.3.0, whose structure is based on LeNet-5, for which the preprocessed face images are fed for training. The efficacy of this method in terms of recognition rate is assessed using the ORL database.

Muhtahir O. Oloyede et al. [7] introduced a technique for recognizing face images which utilizes efficient image enhancement approach for preprocessing face images, as well as a unique set of hybrid features and CNN. The approach uses a metaheuristic optimization strategy for robust face image enhancement in an unconstrained environment. When compared to the original image, recognition performance improves when more features are added to the facial image. The derived hybrid feature is meant to enhance the classification efficiency of advanced CNN architecture. Experimentation on typical face datasets have been performed to demonstrate an increase in the efficiency of recognition which accounts for all the limitations in the face dataset. When compared to other facial conditions, lower face occlusion has a significant influence on the system.

Wenqi Wu et al. [8] presented a Faster R-CNN-based face detection with variable scales. In this methodology, the network employs a number of methodologies such as multitask learning, feature pyramiding, and feature fusion. An efficient multitask region based network with improved facial recognition is built to obtain human face region of interest. In order to build a human face proposal, the anchor is combined with facial landmarks. The proposal scale is then used to propose a parallel-type Fast R-CNN network. The three networks differ in the weight of feature map fusion and are characterized by their proposal scales. The experimental findings demonstrate that the proposed technique outperforms common benchmarks such as FDDB, PASCAL and WIDER FACE when compared to UnitBox, HyperFace, and FastCNN algorithms.

Xianzhang Pan [9] proposed a method for acquiring comprehensive characteristics for video-based facial expression identification using CNNs and HOG. It retrieves deep features from video frames using a set of convolutional kernels in CNNs with displacement, scaling, and deformed invariance. The HOG is then used to generate discriminative features from CNN deep features that are highly associated with face emotions. Classification of expressions is carried out using the Support Vector Machine (SVM) technique. The performance of this technique was assessed using the RML, CK+, and AFEW5.0 datasets.

Yanhong Zhang et al. [10] used a patch method in CNN architectures to build a novel technique for learning enough effective features for face recognition. The network has enabled the cropping of a face region into patches while requiring no additional storage capacity for face patches. Furthermore, using the patch approach, a multi-branch CNN is trained to learn the characteristics of every cropped patch and the patch features are then fused to generate a full face representation. The performance of this method was evaluated by conducting experiments on the LFW and YouTube platforms.

Feng Cen and Guanghui Wang [11] proposed a deep feature dictionary model for recognizing occluded faces. The deep features are retrieved using a CNN and then linearly coded using a dictionary. In addition to the deep features from training samples, the dictionary includes mapping vectors from people inside or outside the training set that are correlated with occlusion patterns from testing face samples.. To normalize the coding coefficients, a squared Euclidean norm is utilized. The performance of this approach was evaluated by conducting experiments on AR, FERET and CelebA databases.

Pengfei Ke et al. [12] used LBP and CNN to examine and assess the impact of posture, lighting, mood, and other variables on face recognition. A CNN is made up of 4 convolution layers, 2 max-pooling layers, an activation layer, a fully linked layer, and an output layer. The LBP descriptor is used to generate LBP coded images, which are then fed into a CNN to extract spatial information from images while minimizing feature dimensionality. Batch normalization is performed to the convolution layer to improve the network structure. Experimenting with the CMU-PIE face database demonstrated that the proposed approach considerably improves face recognition rates.

Fanzhi Kong [13] proposed a method for recognizing facial expressions using a CNN and Local Binary Pattern (LBP) to extract face information. The fully connected layer interprets image abstract characteristics by feeding the original image into a deep convolutional neural network, which eliminates the inherent inaccuracy of image preprocessing and artificially selected features. A standard LBP feature extractor for facial expression images is then constructed at the full connection layer. The enhanced LBP facial expression texture attributes are integrated with the abstract face expression features learnt by a deep convolutional neural network. This will provide unique facial expression features and enhance classification accuracy.

Erfan Zangeneh et al. [14] exploited deep convolutional neural networks (DCNN) to construct a novel linked mapping strategy for recognizing low resolution facial images. The system is made up of two DCNN branches that use nonlinear transformations to map high and low quality images into a common space. The branch dealing with high-resolution image processing includes fourteen layers, while the branch dealing with low-resolution facial image mapping has a five-layer super-resolution network coupled with a fourteen-layer network. Through back propagation, the distance between two attributes that correspond to high- and low-resolution images is used to train networks. The recognition rate of this approach was evaluated on different databases to measure its overall performance.

Raj Silwal et al. [15] propose an approach that combines deep learning, customized architecture, and an enhanced loss function to recognize individuals in unconstrained environments. Multi-Block Local Binary Pattern modules retrieve customized features, whereas CNN modules extract high-level unique features in the system. Individuals are identified by fusing the characteristics from both modules and sending them through a fully connected layer using a softmax classifier. The accuracy and processing time are measured to test its effectiveness.

Ashok Kumar Rai et al. [16] present a hyper spectral face recognition method that combines bands using the Firefly algorithm and classifies using the CNN. The architecture for categorization and band fusion, consisting of four major components the formulation of the hyper spectral imaging (HSI) problem as an Image-Set, the alignment of several posture images inside the same person, the building of an HSI Face ConvNet to train and classify the images, and finally, the application of an upgraded firefly approach to get fused images for classifying intra-person of hyper spectral images. The effectiveness of this method was evaluated on CMU-HSFD & UWA-HSFD databases.

Leslie Ching Ow Tiong et al. [17] suggested the multimodal facial biometrics recognition for identifying faces in surveillance applications. It employs multi-feature fusion layers in dual-stream CNNs, which results in considerable and useful data learning. The network consists of two progressive parts with unique fusion algorithms for aggregating RGB information and texture descriptions for multimodal face biometrics. The performance of this technique was evaluated on four different datasets.

Tripti Goel and R Murugan [18] proposed a method for recognizing face images based on deep convolutional networks and Kernel Extreme Learning Machine (KLEM) classifier. To get robust features from an image, the residual network (ResNet) uses a set of convolution and pooling layers. These layers work together to get features that aren't affected by lighting, posture, or expression. These collected features are then learnt and classified using the polynomial function KELM, the parameters of which are also tuned using the particle swarm optimization process. To test the efficiency of this technique, experiments are conducted on the AT&T, Yale, CMU PIE, and UMIST datasets.

Shuai Peng et al. [19] present a new technique using the Inception-ResNet network to reduce the complexity of training such deep convolutional neural networks while simultaneously improving their performance. A fixed residual scaling factor is employed in Inception-ResNet. Model training stability is improved by converting the value to a trainable parameter and setting it to a modest value. The Inception-ResNet module uses Leaky ReLU and PReLU instead of ReLU activation function, increasing the amount of training parameters while improving training stability and performance. The performance is evaluated by conducting extensive tests on the datasets VGGFace2, MS1MV2, IJBB, and LFW.

Xi Yin and Xiaoming Liu [20] demonstrated the Multitask CNN for recognizing pose invariant faces, with identity classification as the primary task and pose, illumination, and expression estimations as subtasks. The dynamic-weighting approach is used to dynamically allocate loss weights to each sub task, hence resolving the issue of task balance during training, when a bigger loss weight is provided to a simpler sub task. During training, CNN learns certain identification features, as well as a stochastic routing approach for feature fusion during testing. The approach outperforms the current state of the art on the LFW, CFP, and IJB-A datasets when tested on wild datasets.

Yanan Wang et al. [21] suggested a bi-directional Collaborative representation-based classification system for identifying facial images using CNN features. Deep convolutional neural networks were used to obtain the face characteristics from the original gallery and query sets, resulting in an effective reverse representation strategy for extracting more information between training and test samples. To prevent bi-directional optimization, the input sample is characterized by a forward and a backward linear combination. For robust classification, the remainders from forward representation and reverse modelling are combined at a low cost. Experiments using well-known face datasets including AR, FERET and ORL demonstrate the algorithm's validity and resistance to small sample size issues.

Hana Ben Fredj et al. [22] developed a CNN framework to recognizing faces in unconstrained environment. In the framework, aggressive data augmentation is used for learning. Further adaptive fusion of softmax loss and center loss as supervised signals is added to improve the performance. The effectiveness of the model was tested on the LFW and You Tube datasets.

M. Chandrakala and P. Durga Devi [23] proposed a two stage classifier using HOG features to recognize face images. The desired features from the preprocessed images are extracted using HOG. At first, the k-NN classifier was used. Then, unrecognized face images were tested with the SVM classifier, which led to more accurate face recognition.

Htwe Pa Pa Win et al. [24] proposed method for face recognition using CNN. The deep learning strategies of the CNN are used for detecting face, extracting face features and for recognition. The effectiveness of this method was tested by experimentation on FEI dataset resulting in better accuracy and reduced time complexity.

Peng Lu et al. [25] developed a method for face recognition that combines CNN with augmented dataset. The small dataset is augmented to a larger one by transformations of the face images, such as flip, shift, scaling, and rotation. Face features extracted from these augmented datasets resulted in higher recognition rate. This method was tested on the ORL dataset.

Walid Hariri [26] introduced an efficient technique for recognizing masked faces. In this technique, first the masked face region is removed, then three pre-trained deep CNN namelyVGG-16, Alex Net, and ResNet-50 is applied to extract deep features. The classification is performed by multilayer perception. Experiment was conducted on Real-World-Masked-Face-Dataset to evaluate the performance.

The survey identifies several methodologies and their associated experimental outcomes for various datasets. It is discovered that despite the development of several algorithms, high-accuracy face recognition remains a difficult problem owing to a variety of situations such as variations in intensity, lighting, direction, lightning, occlusion, and changing position and facial expressions. Today's deep learning algorithms are quite successful and yield promising results for applications such as pattern recognition and picture classification, among others. Convolutional Neural Networks (CNNs) are feed forward networks comprised of several hidden layers that automatically learn features. While using CNNs alone for face recognition has demonstrated a promising recognition rate, several other studies have demonstrated a drop in recognition rates for a variety of complex datasets. Additionally, using CNNs alone results in a lower recognition rate for big scale databases. We propose a hybrid approach in this study effort that combines modified Histogram oriented gradient (m-HoG) for extracting facial image characteristics in three gradient directions with a deep learning CNN algorithm for classifying the face image based on the features. The study indicates that the suggested approach outperforms existing methods in terms of recognition rates.

3. Methodology

The proposed work consists of two different approaches, a spatial texture pattern extraction technique for extracting facial image features and deep learning algorithm for further feature extraction and classification. Figure 1 shows the proposed model block diagram.

3.1 Database description

L-Spacek Database: This database is being explored to test the algorithm because its wide range of lightning variants, different orientations and expressions. It contains 113 male subjects, each with 20 face images. Out of 113 subjects first 100 subjects are considered for training by choosing first 10 samples of each subject, with a total of 1000 images. 15th face image to 19th face image of the same 100 subjects are considered to compute validation accuracy. The sample images from L-Space k dataset are shown in Figure 2(a).

NIR Database: The NIR database was chosen to test the algorithm due to its variability in terms of intensity, light change and blurring effects, as well as diverse poses and face expressions. It consists of 115 subjects having 14 samples per subject. Out of 115 subjects first 40 subjects are considered for training by choosing first 5 samples of each subject, with a total of 200 images. The 6th face image of the same 40 subjects is considered for compute the recognition rate. The sample images from NIR dataset are shown in Figure 2(b).

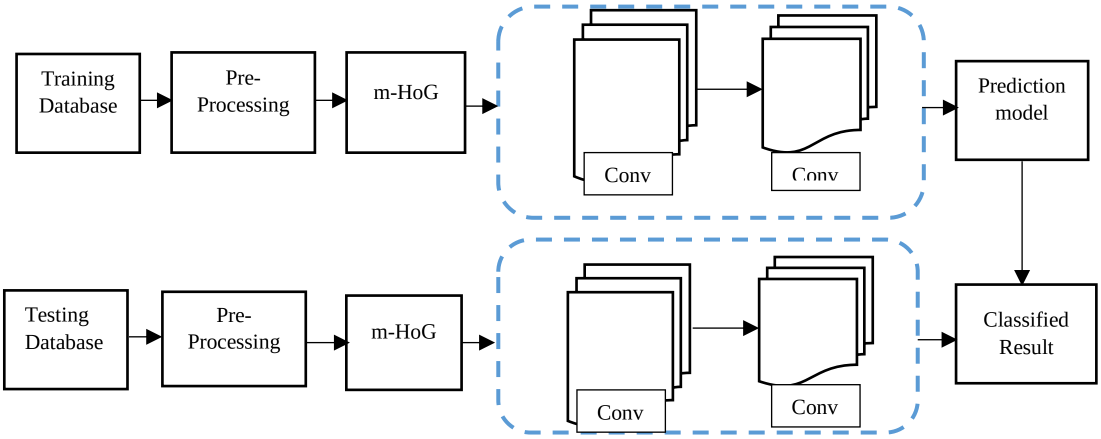


Fig. 1. Proposed model block diagram

JAFFE Database: The JAFFE Database is taken into account since all of the individuals have identical/similar facial images. The database contains

213 images of 10 Japanese female models posing in seven different face expressions. All 10 models face images are considered for training by choosing first 10 samples of each model, with a total of 100 images. The 16th and 17th face images of the same 10 models are considered to evaluate the performance. The sample images from JAFFE dataset are shown in Figure 2(c).

YALE Database: This database is taken into consideration because of the changes in intensity, changes in illumination, varied poses and face expressions. It has 15 subjects containing 11 samples per subject with a total of 165 images. All 15 subjects are considered for training by choosing first 9 samples of each subject, with a total of 135 images. The 10th face image of the same 15 subjects is considered to compute the recognition rate. The sample images from Yale dataset are shown in figure 2(d).

3.2 Preprocessing

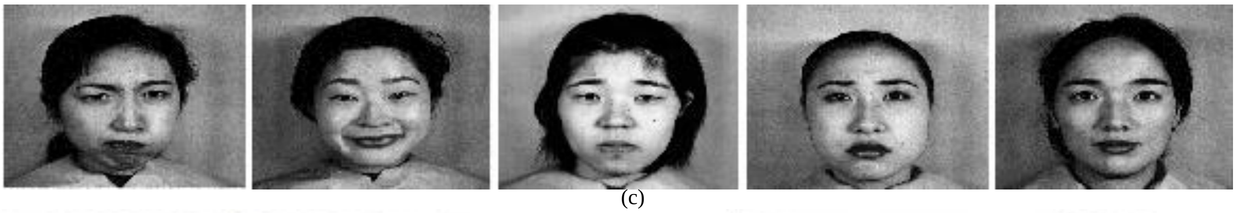
The suggested approach employs the Viola Jones algorithm [27] to locate the face in a complex environment. This algorithm extracts characteristics from both face and non-facial areas by using Haar-like features. Haar features are tiny kernels with varying forms and sizes that are used to identify the existence of a feature in a given image. The Adaboost learning method was used to reduce any redundancies in the collected features. Adaboost is a machine learning technique that determines the best among all characteristics, however these features are referred to as weak classifiers, while strong classifiers are constructed by linearly combining weak classifiers. Finally, the cascade classifier incorporates powerful classifiers at several levels that are utilized to recognize the desired face in the provided image. For all databases, the region of interest is determined by cropping and resizing to 100×100.

3.3 Modified Histogram of Oriented Features

The gradient of the image in traditional HoG is computed using the intensity change only in *x-y* direction. This results in ignoring the change in intensity along *z* direction, which leads to the lack of discriminative features from the image. To overcome this drawback, a modified HoG (m-HoG) is proposed, in which the gradient is computed from the change intensity levels along x-y, x-z and y-z directions which is as shown in Figure 3, making it more descriptive of the image and hence contains discriminative characteristics. The m-HoG gradient is used to generate three distinct HoG versions.

(a)

(b)



(c)

(d)

Fig. 2. Sample images of (a) L-Spacek (b) NIR (c) JAFFE and (d) YALE database

$$\mathit{Modified\ HoG} = \left[HoG1, HoG2, HoG3\right]$$

HoG1, HoG2, and HoG3 are the m-HoG characteristic features computed using the gradient in three directions. The feature vector in this m-HoG version is three times as long as the standard HoG.

The face detected regions using Viola Jones Algorithm have been resized for fixed dimension of 100×100. Resizing process is done to make sure that the number of m-HoG features generated for the subsequent images are uniform. After resizing the image, gradient calculation is performed, which comprises computing the gradient values by using a 1D centered point discrete derivative mask in *x-y, y-z and x-z* directions.

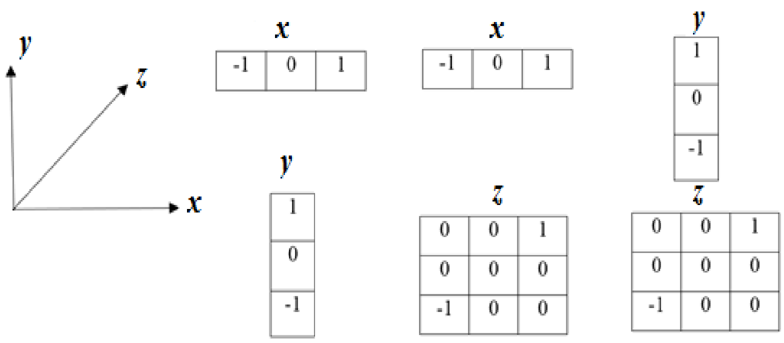


Fig. 3. Shows Gradient estimation along three directions

a. x-y direction

This method entails filtering the grayscale image with the following *x and y* filter kernels, which are represented mathematically as

$$D_x = [-1\ 0\ 1]\ and\ D_y = [1 - 1] \tag{1}$$

So, given an image I , we can use a convolution technique to obtain the *x and y* derivatives as

$$I_x = I * D_x \quad and\ I_y = I * D_y \tag{2}$$

Then the magnitude of the gradient is given by:

$$|G_{x-y}| = \left(I_x{}^2 + I_y{}^2\right)^{0.5} \tag{3}$$

And angular orientation of the gradient is given by:

$$\theta_{x-y} = tan^{-1}\left(I_y / I_x\right) \tag{4}$$

b. x-z direction

$$D_x = [-1\ 0\ 1]\ \ and\ \ D_z = [001; 000; -100] \tag{5}$$

So, given an image I , we can use a convolution technique to obtain the *x and z* derivatives as

$$I_x = I * D_x \quad and\ I_Z = I * D_Z \tag{6}$$

Then the magnitude of the gradient is given by:

$$|G_{x-Z}| = \left(I_x{}^2 + I_Z{}^2\right)^{0.5} \tag{7}$$

And angular orientation of the gradient is given by:

$$\theta_{x-z} = tan^{-1}\left(I_z / I_x\right) \tag{8}$$

c. y-z direction

$$D_y = [1\ 0 - 1] and\ \ D_z = [001; 000; -100] \tag{9}$$

So, given an image I , we can use a convolution technique to obtain the *y and z* derivatives as

$$I_y = I * D_y\ \ and\ \ I_Z = I\ \ D_z \tag{10}$$

Then the magnitude of the gradient is given by:

$$\left|G_{y-z}\right| = \left(I_y^2 + I_z^2\right)^{0.5} \tag{11}$$

And angular orientation of the gradient is given by:

$$\theta_{y-z} = tan^{-1}\left(I_y / I_z\right) \tag{12}$$

The m-HOG returns the total features in a row vector and consists of 13068 features for each face image. Which is three times larger than conventional HoG features. All such row vectors for the subsequent images are stacked one below the other as shown in Figure 4 to create the trained dataset that would be used for feature matching. The obtained 13068 features are resized for size 66 $\times$198 matrix and which is given as input for CNN.

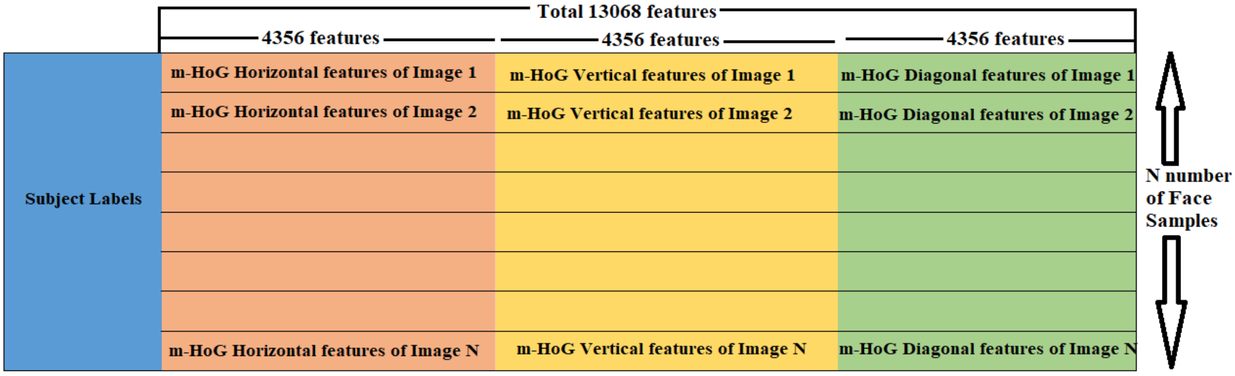


Fig. 4. Shows the m-HoG features values for the training dataset considering N number of face samples.

3.4 Convolution Neural Network

CNNs are the class of deep neural networks that are very effective for applications like pattern recognition and classification. These CNNs are made up of many hidden layers and are of type feed forward network. CNNs contains kernels that have trained weights and biases. Each kernel holds some input and performs convolution operation non-linearly. A CNN architecture can be shown in Figure 5 and it consists of different layers like Convolution, pooling, ReLU and fully connected layer.

Following the extraction of m-HoG features, CNN is employed in the study to extract further face features. The proposed CNN architecture is shown in Figure 6 and it is made up of two convolutional layers followed by batch normalization layers, two maximum pooling layers, and a final fully connected layer.

3.4.1 Convolutional layer

This layer consists of one or more convolutional feature extractors. Each feature extractor consists of an input matrix $\boldsymbol{x}$ of size N×N, convolutional kernel $\boldsymbol{w}$ of size K×K and the bias b .The output of convolutional feature extractor is a matrix $\boldsymbol{h}$ and it can be mathematically expressed as

$$\boldsymbol{h} = f\left(conv\left(\boldsymbol{x}.\boldsymbol{w}\right)+b\right)$$

In which, $f(*)$ is an activation function, $conv(*)$ is a convolution.

In the proposed work, the output feature matrix of size 66×198 obtained by m-HoG technique for the input face image is used as the input matrix x for the CNN model. The model is composed of two convolutional layers, the first of which contains four convolutional kernels of size 5×5 and the second of which contains sixteen convolutional kernels of size 7×7.

3.4.2 Pooling layer

The pooling layer is also known as the subsampled layer. This layer is usually placed after the convolutional layer, can collect features and minimize input image size. Averaging and Max Pooling are two common types of pooling layers. The proposed CNN model employs Max Pooling to down sample the input image. The mathematical expressions for the pooling layer are as follows:

$$F = D_{max}^{l,l}\left(P\right)$$

P Is the input vector, $D_{max}^{l,l}$ denotes the maximum pooling of the l and Fd enotes the output pooling face. The proposed model implements 2×2 pooling filter and the size of feature extractor could be reduced to half of its original size.

3.4.3 Rectified Linear Unit (ReLU)

A popular activation function, the rectified linear unit (ReLU), yields unsigned data type values. The ReLU layer eliminates the negative values present in the feature maps by using a threshold value. Mathematically it can be expressed as

$$f(u) = \begin{cases} 0 \; for \; u < 0 \\ u \; for \; u \geq 0 \end{cases}$$

Where u represents the value of input matrix.

3.4.4 Fully-connected layer

This layer is similar to a classifier in that it is made up of neuron layers, with each neuron in one layer coupled to every neuron in the layer before it. It can be stated mathematically as follows:

$$\boldsymbol{J} = f\left(\boldsymbol{r} \times \boldsymbol{a} + b\right)$$

where r denotes input vector, a and b denotes the fully-connected layer's weight and bias, respectively, J is the output of fully-connected layer, and $f(*)$ is the activation function.

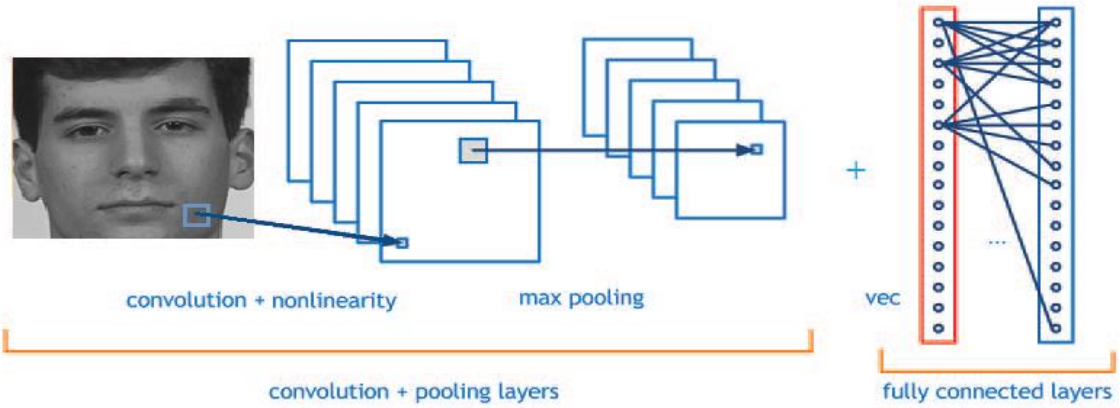


Fig. 5. Typical Convolutional Neural Network architecture

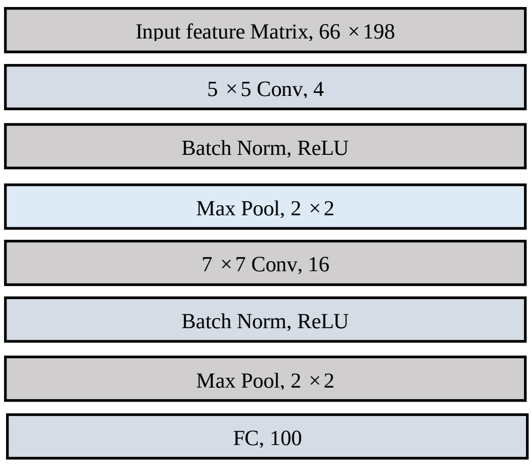


Fig. 6. Proposed Convolutional Neural Network architecture

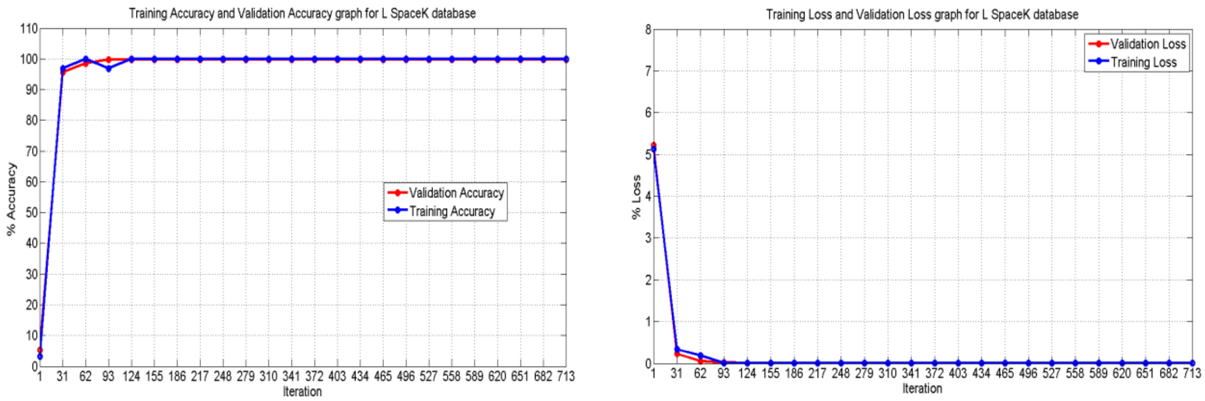
4. Results and Discussions

After extracting the m-HoG features from the input face image, the size of the each image feature matrix is changed to 66×198. We implemented different tests by making selecting different batch size. The proposed method is implemented on MATLAB 2016a since it has built-in CNN library. The minimum batch size is set to 16, 32 and 64 and the recognition rate is evaluated for all the batch sizes. The learning rate of the network is set to 0.001. The gradient decent optimization algorithm is used. The performance of proposed face recognition system has been assessed on L Spacek, JAFFE, NIR and YALE databases by computing recognition rate.

4.1 L-Spacek database

The Epoch is set to 50 to increase the reliability of the experimental findings, where one epoch is equal to training all 1000 database images once. To enhance the accuracy of training samples of this database, all 1000 images are trained 50 times in the experiment. The experiment is carried out using batch sizes of 16, 32, and 64. It is found that better recognition accuracy is achieved for a batch size of 32.

Figure 7 (a) and (b) shows the Accuracy and loss (error) graphs obtained for both training and testing (validation) features. Table 1 shows the values tabulated for validation accuracy and validation loss for different epochs. It is noticed that the system accuracy is over 90% from epoch 1 and reaches 99.80% with minimal validation loss of 0.0098 for 23 epochs. After a specified number of training sessions, the training accuracy reaches 100% and the test recognition rate remains consistent. This model's optimal test recognition rate for trained L-Spacek database is 99.80%.



(a) (b)

Fig. 7. (a) Shows Accuracy graphs for L-Spacek database (b) Shows Loss graphs for L-Spacek database

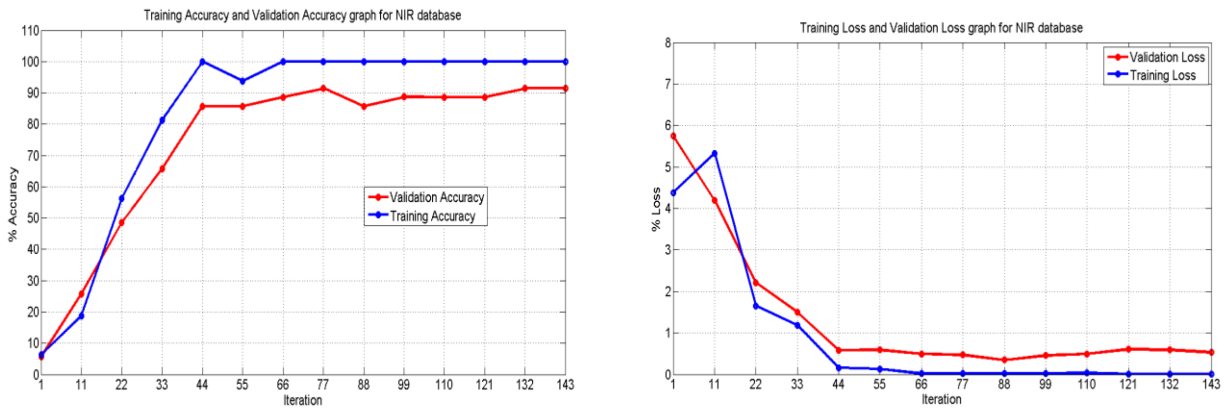
Table 1. shows the values tabulated for L-Spacek database for different epochs

Epoch	Iteration	Validation Accuracyin %	Validation Loss	Mini-Batch Accuracy in %	Mini-Batch Loss
1	1	5.20	5.2135	3.13	5.1352
1	31	95.60	0.2272	96.88	0.3361
2	62	98.60	0.0556	100.00	0.1938
3	93	99.80	0.0155	96.88	0.0053
4	124	99.80	0.0146	100.00	0.0039
5	155	99.80	0.0137	100.00	0.0019
6	186	99.80	0.0133	100.00	0.0019
7	217	99.80	0.0127	100.00	0.0019
8	248	99.80	0.0123	100.00	0.0014
9	279	99.80	0.0120	100.00	0.0021
10	310	99.80	0.0118	100.00	0.0024
11	341	99.80	0.0115	100.00	0.0010
12	372	99.80	0.0113	100.00	0.0011
13	403	99.80	0.0111	100.00	0.0008
14	434	99.80	0.0110	100.00	0.0021
15	465	99.80	0.0108	100.00	0.0018
16	496	99.80	0.0106	100.00	0.0013
17	527	99.80	0.0104	100.00	0.0007
18	558	99.80	0.0103	100.00	0.0008
19	589	99.80	0.0102	100.00	0.0035
20	620	99.80	0.0101	100.00	0.0012
21	651	99.80	0.0100	100.00	0.0006
22	682	99.80	0.0099	100.00	0.0009
23	713	99.80	0.0098	100.00	0.0010

4.2 NIR dataset

The Epoch is set to 50 to increase the reliability of the experimental findings, where one epoch is equal to training all 175 database images once. To enhance the accuracy of training samples of this database, all 175 images are trained 50 times in the experiment. The experiment is carried out using batch sizes of 16, 32, and 64. It is found that better recognition accuracy is achieved for a batch size of 16.

Figure 8 (a) and (b) shows the Accuracy and loss (error) graphs obtained for both training and testing (validation) features. Table 2 shows the values tabulated for validation accuracy and validation loss for different epochs. It is noticed that the system accuracy is over 90% from epoch 12 and reaches 91.43% with minimal validation loss of 0.5321 for 13 epochs. After a specified number of training sessions, the training accuracy reaches 100% and the test recognition rate remains consistent. This model's optimal test recognition rate for trained NIR database is 91.43%.



(a) (b)

Fig. 8. (a) Shows Accuracy graphs for NIR database (b) Shows Loss graph for NIR database

Table 2. Shows the values tabulated for NIR database for different epochs.

Epoch	Iteration	Validation Accuracy in %	Validation Loss	Mini-Batch Accuracy in %	Mini-Batch Loss
1	1	5.71	5.7488	6.25	4.3789
1	11	25.71	4.1943	18.75	5.3290
2	22	48.57	2.2104	56.25	1.6501
3	33	65.71	1.4960	81.25	1.1888
4	44	85.71	0.5860	100.00	0.1583
5	55	85.71	0.5887	93.75	0.1311
6	66	88.57	0.4909	100.00	0.0196
7	77	91.43	0.4682	100.00	0.0211
8	88	85.71	0.3466	100.00	0.0166
9	99	88.71	0.4563	100.00	0.0204
10	110	88.57	0.4924	100.00	0.0312
11	121	88.57	0.6041	100.00	0.0063
12	132	91.43	0.5879	100.00	0.0064
13	143	91.43	0.5321	100.00	0.0117

4.3 JAFFE database

The Epoch is set to 50 to increase the reliability of the experimental findings, where one epoch is equal to training all 100 database images once. To enhance the accuracy of training samples of this database, all 100 images are trained 50 times in the experiment. The experiment is carried out using batch sizes of 16, 32, and 64. It is found that better recognition accuracy is achieved for a batch size of 32.

Figure 9 (a) and (b) shows the Accuracy and loss (error) graphs obtained for both training and testing (validation) features. Table 3 shows the values tabulated for validation accuracy and validation loss for different epochs. It is noticed that the system accuracy is over 90% from epoch 14 and reaches 95% with minimal validation loss of 0.0025 for 30 epochs. After a specified number of training sessions, the training accuracy reaches 100% and the test recognition rate remains consistent. This model's optimal test recognition rate for trained JAFFE database is 95%.

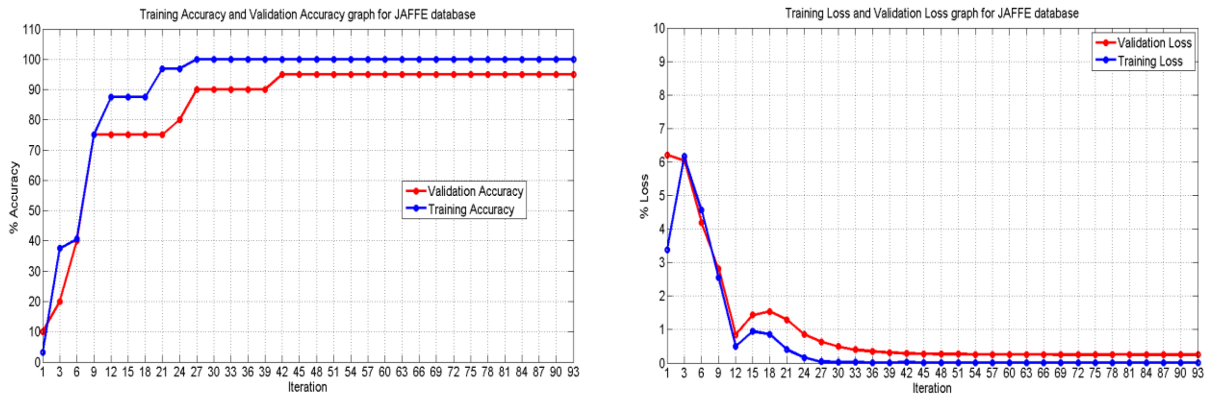


Fig. 9. (a) Shows Accuracy graphs for JAFFE database (b) Shows Loss graphs for JAFFE database

Table 3. shows the values tabulated for JAFFE database for different epochs

Epoch	Iteration	Validation Accuracy in %	Validation Loss	Mini-Batch Accuracy in %	Mini-Batch Loss
1	1	10.00	6.2005	3.13	3.3891
1	3	20.00	6.0339	37.50	6.1608
2	6	40.00	4.1914	40.63	4.5667
3	9	75.00	2.8158	75.00	2.5437
4	12	75.00	0.8419	87.50	0.4917
5	15	75.00	1.4347	87.50	0.9478
6	18	75.00	1.5400	87.50	0.8661
7	21	75.00	1.2882	96.88	0.4005
8	24	80.00	0.8498	96.88	0.1649
9	27	90.00	0.6335	100.00	0.0454
10	30	90.00	0.4934	100.00	0.0266
11	33	90.00	0.4013	100.00	0.0243
12	36	90.00	0.3455	100.00	0.0140
13	39	90.00	0.3114	100.00	0.0181
14	42	95.00	0.2904	100.00	0.0221
15	45	95.00	0.2775	100.00	0.0157
16	48	95.00	0.2700	100.00	0.0078
17	51	95.00	0.2652	100.00	0.0148
18	54	95.00	0.2623	100.00	0.0062
19	57	95.00	0.2601	100.00	0.0086
20	60	95.00	0.2589	100.00	0.0051
21	63	95.00	0.2574	100.00	0.0048
22	66	95.00	0.2556	100.00	0.0066
23	69	95.00	0.2542	100.00	0.0077
24	72	95.00	0.2533	100.00	0.0032
25	75	95.00	0.2528	100.00	0.0031
26	78	95.00	0.2525	100.00	0.0028
27	81	95.00	0.2525	100.00	0.0053
28	84	95.00	0.2529	100.00	0.0038
29	87	95.00	0.2532	100.00	0.0040
30	90	95.00	0.2534	100.00	0.0025
31	93	95.00	0.2537	100.00	0.0038

4.4 *YALE database*

The Epoch is set to 50 to increase the reliability of the experimental findings, where one epoch is equal to training all 135 database images once. To enhance the accuracy of training samples of this database, all 135 images are trained 50 times in the experiment. The experiment is carried out using batch sizes of 16, 32, and 64. It is found that better recognition accuracy is achieved for a batch size of 16.

Figure 10 (a) and (b) shows the Accuracy and loss (error) graphs obtained for both training and testing (validation) features. Table 4 shows the values tabulated for validation accuracy and validation loss for different epochs. It is noticed that the system accuracy is over 90% from epoch 6 and reaches 93.33% with minimal validation loss of 0.2614 for 14 epochs. After a specified number of training sessions, the training accuracy reaches 100% and the test recognition rate remains consistent. This model's optimal test recognition rate for trained YALE database is 93.33%.

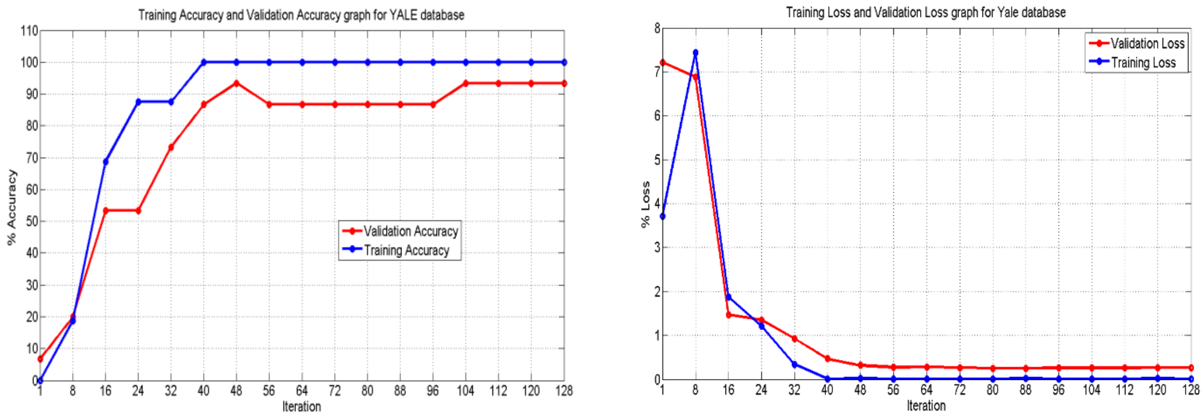


Fig. 10. (a) Shows Accuracy graphs for YALE database (b) Shows Loss graphs for YALE database

Table 4. Shows the values tabulated for YALE database for different epochs.

Epoch	Iteration	Validation Accuracy in %	Validation Loss	Mini-Batch Accuracy in %	Mini-Batch Loss
1	1	6.67	7.2130	0.00	3.7148
1	8	20.00	6.8748	18.75	7.4419
2	16	53.33	1.4699	68.75	1.8690
3	24	53.33	1.3547	87.50	1.2084
4	32	73.33	0.9257	87.50	0.3331
5	40	86.67	0.4679	100.00	0.0134
6	48	93.33	0.3167	100.00	0.0298
7	56	86.67	0.2778	100.00	0.0041
8	64	86.67	0.2818	100.00	0.0011
9	72	86.67	0.2663	100.00	0.0052
10	80	86.67	0.2551	100.00	0.0013
11	88	86.67	0.2552	100.00	0.0208
12	96	86.67	0.2568	100.00	0.0114
13	104	93.33	0.2620	100.00	0.0006
14	112	93.33	0.2614	100.00	0.0013
15	120	93.33	0.2708	100.00	0.0192
16	128	93.33	0.2680	100.00	0.0009

From Table 5, it is noticed that the proposed face recognition system using CNN along with m-HoG feature extractor achieves better recognition rate. It results 99.80% accuracy for L-Spacek, 93.33% Yale dataset, 95% for JAFFE dataset and 91.43% for NIR dataset respectively.

Table 5. Recognition Rate evaluated for different datasets for different epochs.

Batch Size	Training Accuracy	Recognition Rate	Learning Rate
Jaffe Dataset			
16	100	90.00	
32	100	95.00	0.001
64	91.15	70.00	
NIR Dataset			
16	100	91.43	
32	97.25	88.57	0.001
64	90.11	82.46	
L-Spacek Dataset			
16	100	99.40	
32	100	99.80	0.001
64	99.90	98.80	
YALE Dataset			
16	99.26	93.33	
32	99.26	73.33	0.001
64	98.52	80.00	

Table 6 shows the comparison of the recogntion rate achieved by the conventinal methods with that of the proposed method for different datasets.It is noticed that the proposed work results in better recognition rate compare to these conventional methods. The proposed method implements modified Histogram oriented Gradient (m-HoG) for extracting facial image features and deep learning CNN algorithm to further extract features and classify the face image. As a result, our suggested method performs better in rejecting unknown faces and accepting known faces with the lowest possible error rate

Table 6 Comparison of recognition rates (in %) for different database with previous studies

Author-Year	Database	Technique	Classifier	Recognition Rate
Surbhi Gupta et al. [28] 2021	L-Spacek(Face94)	SIFT+SURF		87.3%
Deepali Virmani et al. [29] 2019	L-Spacek(Face94)	EmbedNet	SVM	98.03%
			KNN	97.83%
Srinivas Halvi et al. 30]	NIR	1DTD	FFBP NN	72%
Shreyas Saxena et al. [31] 2016	NIR	CNN	Softmax	85.9%
Srijith Rajeev et al., [32] 2019	IIT Delhi NIR	Circular with HOG	SVM	80.41%
Sateesh Kumar H C et al. [33] 2017	NIR	STWT	ED	86.67%
		DTCWT		90.24%
	YALE	STWT		87.5%
		DTCWT		87.5%
Farid Ayeche et al. [34] 2021	YALE JAFFE	HDGG	SVM	92.12% 91.43%
Rahul Ravi et al. [35] 2020	JAFFE	CNN	Softmax	77.27
Sujay H S et al.[36] 2019	YALE	LBP	SVM	86.66%
Shucheng Huang et al. [37] 2016	YALE	EDLPP	-	88.56%
V. Betcy Thanga Shoba et al. [38] 2020	YALE	SURF+HOG+MSER	-	91.2%
Proposed Work	L-Spacek	m-HOG+CNN	Softmax	99.40%
	NIR			91.43%
	JAFFE			95.00%
	YALE			93.33%

4.5 Work Limitations and Future scope

Face recognition rate will decline as the face database expands. More public datasets can be found in the face databases. As a result, the key direction of future research will be to maintain recognition rate stability in large databases. Compression techniques can be used to compress the feature dimensions for the datasets containing large set of face images and thereby increasing the speed of recognition. Face biometric algorithms can be developed to recognize the face images present in the video sequences. A real time face recognition system can be developed using the hardware module for portability.

5. Conclusion

In this paper, we propose a model based on architectures of convolutional neural networks (CNN) to human face recognition problems. This model has 2 convolutional neuron layers (CONV) and 1 fully connected neuron layers (FC). Modified Histogram oriented gradient (m-HoG) for extracting facial image features and deep learning. The face images are classified using the CNN algorithm based on their features. The proposed hybrid approach overcomes the performance degradation problem of using CNN models alone. Further, m-HOG features are extracted to avoid loss of features along three direction and the obtained histogram features are uniformly spread over 0 to 180 degrees. The obtained features are resized as a matrix having dimension 66×198 and which are passed to the four layers CNN structure made up of convolutional, max pooling, convolutional, max pooling layers respectively for training and face image classification is carried out by softmax layer. We applied our model with different datasets having color variations, fading, contrast and brightness change. Performance analysis is measured in terms of recognition accuracy. It is found that L-Spacek and JAFFE datasets yields better recognition rate of 99.80% and 95.00% respectively. Whereas the datasets like YALE and NIR having similarity, pose and brightness variations, yields recognition rate of 93.33% and 91.43% respectively.

Acknowledgment

The research was supported by Visvesvaraya Technological University, Jnana Sangama, Belagavi – 590018, Karnataka, India.

Conflicts of Interest

The authors have no conflicts of interest to declare.

Author's Contribution Statement

Raveendra K: Conceptualization, Investigation, Data collection, Writing – original draft, Analysis and Interpretation of results and and Draft manuscript preparation. **Ravi J:** Study Conception, Design, Supervision and Investigation on challenges.

References

[1] Yong Xu, Qi Zhu, Zizhu Fan, David Zhang, Jianxun Mi and Zhihui Lai, "Using the idea of the sparse representation to perform coarse to-fine face recognition," Information Sciences. Vol. 238, pp. 138–148, July 2013

[2] Yi Sun, Xiaogang Wang, Xiaoou Tang, "Deep learning face representation by joint identification-verification," Advances in Neural Information Processing Systems, Vol. 27, pp. 1988–1996, 2014.

[3] Jae Young Choi, "Spatial pyramid face feature representation and weighted dissimilarity matching for improved face recognition," The Visual Computer: International Journal of Computer Graphics, 34(11), 1535–1549 (2018)

[4] Larry S. Yaeger, Richard F. Lyon, Brandyn J. Webb, "Effective Training of a Neural Network Character Classifier for Word Recognition," Advances in Neural Information Processing Systems, pp. 807–813 (1996)

[5] Alex Krizhevsky, Ilya Sutskever and Geoffrey E. Hinton, "ImageNet classification with deep convolutional neural networks," Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)

[6] GailiYue and Lei Lu, "Face recognition based on histogram equalization and convolution neural network," International Conference on Intelligent Human-Machine Systems and Cybernetics, pp. 336-339, 2018.

[7] Muhtahir O. Oloyede, Gerhard P. Hancke and Herman C. Myburgh, "Improving Face Recognition Systems Using a New Image Enhancement Technique, Hybrid Features and the Convolutional Neural Network," IEEE Access, vol. 6, pp. 75181-75191, 2018.

[8] Wenqi Wu, Yingjie Yin, Xingang Wang and De Xu, "Face Detection With Different Scales Based on Faster R-CNN," IEEE Transactions on Cybernetics, vol. 49, no. 11, pp. 4017-4028, Nov. 2019.

[9] Xianzhang Pan, "Fusing HOG and convolutional neural network spatial-temporal features for video-based facial expression recognition," IET Image Processing, vol. 14, Issue 1, pp. 176 – 182, January 2020.

[10] Yanhong Zhang , Kun Shang, Jun Wang, Nan Li and Monica M.Y. Zhang. "Patch strategy for deep face recognition," IET Image Processing, Vol.12, Issue 5, pp. 819- 825, May 2018.

[11] Feng Cen and Guanghui Wang, "Dictionary Representation of Deep Features for Occlusion-Robust Face Recognition," IEEE Access, vol. 7, pp. 26595-26605, 2019.

[12] Pengfei Ke, Maoguo Cai, Hanmo Wang and Jialong Chen, "A novel face recognition algorithm based on the combination of LBP and CNN," International Conference on Signal Processing, pp. 539-543, 2018.

[13] Fanzhi Kong, "Facial expression recognition method based on deep convolutional neural network combined with improved LBP features," Personal and Ubiquitous Computing, vol. 23, pp. 531–539, 2019.

[14] Erfan Zangeneh, Mohammad Rahmati and Yalda Mohsenzadeh, "Low resolution face recognition using a two-branch deep convolutional neural network architecture," Expert Systems with Applications, vol. 139, pp. 1-11, January 2020.

[15] Raj Silwal, Abeer Alsadoon, P. W. C. Prasad, Omar Hisham Alsadoon & Ammar Al-Qaraghuli, “A novel deep learning system for facial feature extraction by fusing CNN and MB-LBP and using enhanced loss function,” Multimedia Tools and Applications, vol, 79, pp. 31027–31047, 2020.

[16] Ashok Kumar Rai , Radha Senthilkumar and Aswin Kumar R, “Combining pixel selection with covariance similarity approach in hyper spectral face recognition based on convolution neural network,” Microprocessors and Microsystems, vol. 76, pp. 1-8, 2020.

[17] Leslie Ching Ow Tiong, Song Tae Kim and Yong Man Ro, “Multimodal Facial Biometrics Recognition: Dual-stream Convolutional Neural Networks with Multi-feature Fusion Layers,” Image and Vision Computing, vol. 12, pp. 2020.

[18] Tripti Goel and R Murugan, “Deep Convolutional - Optimized Kernel Extreme Learning Machine Based Classifier for Face Recognition,” Computers and Electrical Engineering, vol. 85, pp. 1-12, 2020.

[19] Shuai Peng, Hongbo Huang, Weijun Chen, Liang Zhang and Weiwei Fang, “More trainable inception-ResNet for face recognition,” Neurocomputing, vol. 411, pp. 9-19, 2020.

[20] Xi Yin and Xiaoming Liu “Multi-Task Convolutional Neural Network for Pose-Invariant Face Recognition,” IEEE Transactions on Image Processing, vol. 27, no. 2, pp. 964-975, Feb. 2018

[21] Yanan Wang, Tian Na, Xiaoning Song and Guosheng Hu, “Bi-directional CRC algorithm using CNN based features for face classification,” Asian Conference on Artificial Intelligence Technology, Vol.2018 ,Issue 16 p. 1457 – 1462, November 2018

[22] Hana Ben Fredj, Safa Bouguezzi and Chokri Souani, “Face Recognition in Unconstrained Environment with CNN,” The Visual Computer, vol. 37, pp. 217–226, 2021.

[23] M. Chandrakala and P. Durga Devi, “Two-stage Classifier for Face Recognition using HOG Features,” Materials Today: Proceedings, vol. 47, pp. 5771-5775, 2021.

[24] Htwe Pa Pa Win, Phyo Thu Thu Khine and Khin Nwe Ni Tun, “ Face Recognition System based on Convolution Neural Networks,” International Journal on Image, Graphics and Signal Processing, vol. 6, pp. 23-29, 2021.

[25] Peng Lu, Baoye Song and Lin Xu, “Human face recognition based on convolutional neural network and augmented dataset,” Systems Science & Control Engineering, vol. 9, pp. 29-37, 2021.

[26] Walid Hariri, “Efficient masked face recognition method during the COVID-19 pandemic,” Signal, Image and Video Processing, vol. 16, pp. 605–612, 2021.

[27] Paul Viola and Michael J. Jones, “ Robust Real-Time Face Detection,” International Journal of Computer Vision, Vol. 57, No. 2, pp. 137-154, 2004.

[28] Surbhi Gupta, Kutub Thakur and Munish Kumar, “2D-human face recognition using SIFT and SURF descriptors of face’s feature regions,” The Visual Computer, vol. 37, pp. 447–456, 2021.

[29] Deepali Virmani, Palak Girdhar, Prateek Jain and Pakhi Bamdev, “FDREnet: Face Detection and Recognition Pipeline,” Engineering, Technology & Applied Science Research, vol. 9, pp. 3933-3938, 2019.

[30] Srinivas Halvi, Nayina Ramapur , K B Raja and Shanti Prasad, “Fusion Based Face Recognition System using 1D Transform Domains,” International Conference on Advances in Computing and Communications, Procedia Computer Science, Vol. 115, pp.383–390, 2017

[31] Shreyas Saxena and Jakob Verbeek, “Heterogeneous Face Recognition with CNNs,” ECCV Workshops, pp. 483-491, 2016.

[32] Srijith Rajeev, Shreyas Kamath K.M, Qianwen Wan and Karen Panetta, “Illumination Invariant NIR Face Recognition using Directional Visibility,” Electronic Imaging, pp. 273-1-273-7, 2019.

[33] Sateesh Kumar H C, C Chowda reddy, Raja K B and Venugopal K R “Face Recognition based on STWT and DTCWT using two dimensional Q-shift Filters,” International Journal of Engineering Research and Application, vol. 7, pp. 64-69, 2017.

[34] Farid Ayeche and Adel Alti, “HDG and HDGG: an extensible feature extraction descriptor for effective face and facial expressions,” Pattern Analysis and Applications, vol. 24, pp. 1095–1110, 2021.

[35] Rahul Ravi, Yadhukrishna S.V and Rajalakshmi prithviraj, “A Face Expression Recognition Using CNN & LBP,” International Conference on Computing Methodologies and Communication, pp. 684-689, 2020.

[36] Sujay S N and H S Manjunatha Reddy, “Extended Local Binary Pattern Features based Face Recognition using Multilevel SVM Classifier,” vol.8, pp. 4123-4128, 2019.

[37] Shucheng Huang and Lu Zhuang, “Exponential Discriminant Locality Preserving Projection for Face Recognition,” Neurocomputing, vol. 208, pp. 373-377, 2016.

[38] V. Betcy Thanga Shoba and I. Shatheesh Sam, “Hybrid Features Extraction on Face for Efficient Face Recognition,” Multimedia Tools and Applications, vol. 79, pp. 1-22, 2020.

Authors’ Profiles

Dr. Raveendra K is a Assistant Professor and HOD in the Department of Electronics and Communication Engineering, Government Engineering College, K R Pet. He obtained his BE Degree in Electronics and Communication Engineering from Kuvempu University, Shimoga. His specialization in Master degree was Digital Communication Engineering from VTU, Belagavi. He was awarded PhD in Electrical and Electronics Engineering Sciences from VTU Belagavi. His area of interest is in the field of Image Processing, Signal Processing and Biometrics.

Dr. Ravi J is a Professor in the Department of ECE, GAT, Bengaluru. He received his BE Degree in Instrumentation Technology from Bangalore University, Bengaluru. He obtained his Master degree in Digital Electronics from VTU, Belagavi. He was awarded PhD in Electronics and Communication Engineering from JNTU Anantapur. He has over 30 peer-reviewed research articles in international journals and conference proceedings. His research interests include Image Processing, Biometrics VLSI and computer networks. He is life member of IETE, ISTE, and ISAMP.

Dr. Khalid Nazim Sattar Abdul is a Assistant Professor in the Computer Science and Information Department, College of Science, Al-Zulfi Campus, Majmaah University. He received his BE Degree in Computer Science & Engineering from Ghousia College of Engineering, Bangalore University, Bangalore. He obtained his master's degree in computer science & Engineering from VTU, Belagavi. He was awarded PhD in Computer Science & Engineering from School of Computer Science, Singhania University, Rajasthan. He has over 33 peer-reviewed research articles in international journals and conference proceedings. His research interests include Image processing, Software Engineering, Information Systems, Computer Networks.

How to cite this paper: Raveendra K, Ravi J, Khalid Nazim Abdul Sattar, "Face Recognition Using Modified Histogram of Oriented Gradients and Convolutional Neural Networks", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.15, No.5, pp. 60-76, 2023. DOI:10.5815/ijigsp.2023.05.05