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

FACE RECOGNITION BY USING ENSEMBLE LEARNING

Name: M. Gowtam

Mobile Number: 6301861716

Abstract :

The popular deep learning network, the convolutional neural network widely used for applications like image and video recognition, image classification, medical image analysis and natural language processing. The Convolutional neural networks(CNN's) are efficient pattern recognition networks that are designed to classify and recognize the images which is mainly used in security and surveillance systems. The CNN's due to its efficient recognition ability are at the base of the security systems. The other neural networks are also able to recognize the images, but they lack the rate accuracy of recognition and low error rate. The recognition system designed to be effective must be efficient, easy to implement and produce high recognition rates. The overall performance of the security and surveillance systems depends on the accuracy, that the CNN's can produce. A brief description about face recognition systems, the challenges faced in face recognition, motivation to use the convolutional neural networks and the objectives are presented. The Feed Forward Neural Network(FFNN) when used for recognition have issues like loss of neighborhood information, it has a lot of parameters to be optimized with in turn increases the training time of the model and also affect the overall accuracy of recognition. Also there are other disadvantages like prone to over fitting, the more number of parameters which results in more converging time, large model size. While the convolutional neural network has less parameters to optimize and it doesn't loose the neighborhood data. The prominence of using the CNN is that it doesn't require any external feature extraction method, the model itself directly learns then data from input images. The learning algorithm used for training also affect the performance of system, the CNN uses sophisticated learning methods like Adam to read the features from the inputs, is implemented in this thesis. This thesis proposes a new and efficient method for face recognition using the convolutional neural networks. The main advantage of the proposed convolutional neural network model is that it provides efficient training which means the well trained network results in better testing and increases the overall accuracy of recognition. Also it reduces the external steps required for the feature extraction. The use of Adam optimization results in fast converging of the parameters by using the momentum concept in the optimization technique, this increases training accuracy and minimizes the error in model.

The proposed CNN achieved good accuracy over existing methods. The conclusion and the possible future directions to extend the proposed technique are presented in this thesis.

1. Introduction

1.1 Principles of Pattern Recognition

Pattern recognition system In the digital world, patterns are everywhere. A pattern can be seen statistically by using algorithms, or it can be seen physically. Pattern recognition system In the digital world, patterns are everywhere. A pattern can be seen statistically by using algorithms, or it can be seen physically.

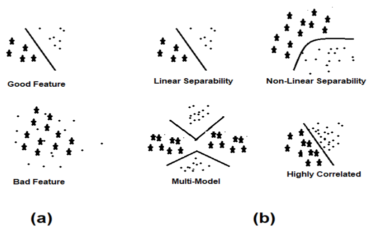
Patterns are made up of the following two essential components in pattern recognition:

- Collection of observations
- The concept behind the observation

1.1.1 Feature Vector:

Another name for the group of observations is a feature vector. To set yourself apart from the competition, your product or service has to have unique qualities, or features. A feature vector is an n-dimensional column vector that combines n features. The features values of the various classes may vary, but the features values of the same class remain constant.

Example:

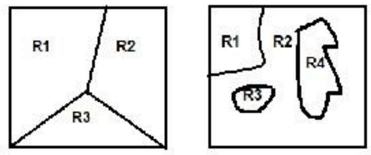


- Differentiate between good and bad features.
- Feature properties.

1.1.2 Classifier and Decision Boundaries:

The decision boundary is the hypersurface that divides the vector space into two classes in a statistical classification task. The area of a problem space where a classifier's output label is ambiguous is known as a decision boundary. A classifier is a hypothesis or discrete-valued function that is used to categorise specific data points and assign them to specific classes.

The feature space is divided into decision regions with assigned classes using a classifier. While the boundaries between decision areas are known as decision boundaries.



Classifier and decision boundaries

1.1.3 Components in Pattern Recognition System:

It is possible to disassemble a pattern recognition system into its constituent elements. For diverse pattern recognition systems, there are five standard components. These are listed as follows:

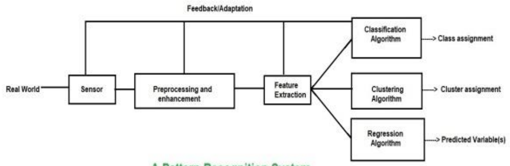
Sensor : A sensor is a tool that measures a characteristic, like pressure, position, temperature, or acceleration, and provides feedback.

Preprocessing Mechanism : The most common way of partitioning information into different fragments is known as division. Segmentation also refers to the process of breaking up a larger data set into smaller manageable pieces.

Feature Extraction Mechanism : Feature extraction is a process that takes a collection of raw measurement data and generates a set of meaningful, unique values (features) for further processing. This enhances human interpretations in specific cases and speeds up the learning and generalisation processes. It might be automated or user-controlled.

Description Algorithm : Algorithms for pattern recognition often attempt to find the "most probable" matching between inputs while also taking their statistical variability into consideration and providing an acceptable result for any given set of inputs.

Training Set : A portion of the total dataset is made up of training data, combined with a testing set. Generally speaking, the algorithm or classifier performs better the better the training data.



A Pattern Recognition System

1.1.4 Design Principles of Pattern Recognition

Two fundamental methodologies are employed in pattern recognition systems to recognise patterns or structures, and they can be applied in various ways. which are

- Statistical Approach and
- Structural Approach

Statistical Approach: The statistical evaluation of unprocessed research data involves the use of various mathematical equations, models, and techniques. A variety of tools for assessing the validity of studies may be found in statistical techniques, which are used to extract information from research data.

There are two basic types of statistics utilized :

1. Descriptive Statistics: Central measures of variation, such as mean and standard deviation, are used to collect data from a sample.

2.Inferential Statistics: From data that are prone to random variance, it draws conclusions.

3. Structural Approach: The Structural Approach is a method for teaching sentence structure to learners. The many word arrangements in one recognised style or another are known as structures.

Types of structures:

- Sentence Patterns
- Phrase Patterns
- Formulas
- Idioms

1.3FaceRecognition System

The incidence of numerous suspicious actions in open and closed contexts continues to boost the need for real-time security. Everybody's safety is at risk from everyday hazards. Though there have been many methodological advancements in this area, not all problems have been solved. In addition to its usefulness in biometrics, facial recognition also has various security and surveillance-related applications. The exponential growth of technology in recent decades has paved the way for the emergence of a "smart" civilization. The objective of a "smart" society is to reduce the amount of human involvement with machines.

The ability to monitor each individual who enters a restricted area is the primary challenge that must be met by both governmental and private security agencies. There are a number of obstacles that businesses must overcome in this respect. A person's authentication doesn't end after they've been granted access to a system; monitoring them still takes time and resources. Another significant difficulty in this field is the need to recognise and identify suspicious behaviour in the monitored region in real time. Facial recognition is, on the one hand, an automatic procedure that requires little thought or effort from the observer.Auto-facial recognition offers many useful features since it is a biometric technique. They feature the crucial benefit of doing no harm to the patient in any way.

Fingerprints, DNA, and faces are examples of physiological biometrics, whereas keystroke and voice print are examples of behavioural biometrics. Except in extreme cases of harm, the physiological methods are more permanent and dependable. The effects of stress, sickness, or weariness on an individual's general state become more apparent in their behaviour. When we use our faces to prove who we are in official documents like driver's licences and passports, facial recognition plays a crucial role. Researchers have shown that individuals differ widely in their ability to identify the face, despite the fact that it is crucial in several professions. In this thesis, I propose a technique for efficient face recognition and I also briefly examine alternative frameworks for face recognition that have been developed to far. When it comes to recognising faces, the suggested technique is more efficient, requires less time to train, and produces better results.

2. RELATED WORKS

2.1 Introduction

A facial recognition system is a piece of software or hardware that can identify an individual based on a photograph or video still from a large database of people's faces. This kind of system is often used for ID verification services to ensure the identity of its customers. In the 1960s, comparable systems started to be created as a computer programme. Facial recognition systems have been around for a while, but they've recently found new life in mobile devices and other hardware, including robots. Facial recognition systems fall under the umbrella term of bio-metrics due to the fact that they rely on analysing a person's physiological traits to identify them. Recent advances in human-computer interaction, video surveillance, and automated picture indexing have all made use of facial recognition systems. Governments and businesses alike use facial recognition technology in the modern day.

2.2 System Working

Figure 2.1 depicts the components of a typical face recognition system, and the subsections that follow describe their function: face detection; face normalisation; feature extraction; and matching.

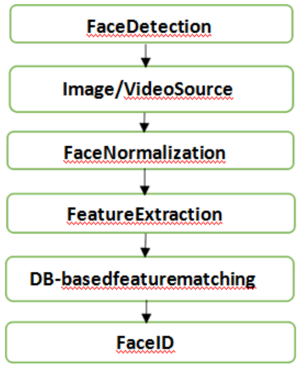


Figure2.1TypicalFaceRecognitionsystem

Image/Video sources are typically from cameras, surveillance footage/cc T.Vfootage. When a face is detected, its region is separated from the surrounding environment. When dealing with video, a face tracking component may be required to follow the identified faces over numerous frames. Whereas face detection just gives a rough estimation of the face's position and size, face landmarking pinpoints specific features of the face, such as the eyes, nose, mouth, and contour. A majority of face-detection research has concentrated on developing methods to recognise human faces from the front. It's quite similar to picture detection, where a person's likeness is compared pixel by pixel.

The purpose of face normalisation is to standardise the face's geometry and lighting. This is essential since modern identification techniques are required to work with face photos in a wide variety of positions and lighting conditions. The face is cropped as part of the geometric normalising procedure to make it fit a generic picture frame. A normalised face is the result of a photometric normalising technique that takes into account factors like lighting and grayscale. It is possible to alter the range of intensity values per pixel by a procedure called normalisation. Starting with raw measurement data, feature extraction then constructs derived values or features that are meant to be useful and non-redundant in order to aid in later learning and generalisation.

The input face's extracted features are compared to those of one or more enrolled faces. When used for 1:1 verification, the matching outputs a 'yes' or 'no,' Finding a good similarity measure for comparing facial characteristics is the primary difficulty at this point in face recognition. Precise face acknowledgment frameworks rely basically upon the highlights extricated to address the face, which thusly requires exact face confinement and standardization.

3. IMPLEMENTATION

3.1 Convolutional neural networks

In this article, we presented the idea of fully-connected layers. The input data for image processing is often quite large (for example, a chest X-ray picture can be 2330 by 2846 pixels in size). When using just fully-connected layers, the number of parameters quickly expands and becomes incredibly large, making optimization highly computationally demanding, if not impossible. Since there was a need to simplify the amount of settings, new sorts of layers were developed to do just that. By factoring in our past knowledge of the significant local correlation of surrounding pixels into the layer structure, we may do image processing with fewer parameters. As a result, neurons no longer have connections to all the other neurons in the input, but rather to a much smaller subset of the input. Therefore, most fully-connected layers in neural networks were removed and replaced with convolutional layers (see Section 3.2) [LeCun et al., 1989]; Subsequently, the expression "convolutional brain organization." Normal techniques for building such organizations these days incorporate stacking convolutional layers, pooling layers (portrayed in Segment 3.3), clump standardization layers (shrouded in Segment 3.4), and a last completely associated layer. Convolutional neural networks are introduced, and their image-data-extracting mechanisms and the various layer types are discussed below.

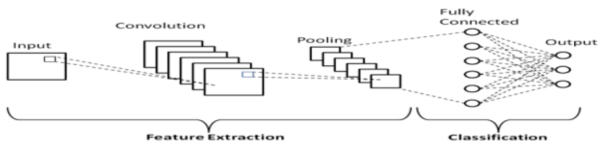


Figure 3.2.1 convolutional neural network

3.2 Pooling layer

The spatial dimensions may be decreased by using pooling layers, which are determined by three parameters: the operation performed on the filter area, the filter size (in bytes, or k lk l), and the stride (in bytes, or s l). Maximum and average pooling are the most typical procedures. . Before using the fully-connected layer, it is usual practise to utilise average pooling as a last layer to reduce the spatial dimensions. As a rule, only the width and height of the input tensor are shrunk down, but the depth remains unchanged. Figure 3.5 depicts max-pooling with a filter size of k l=2 and a stride of s l=2.

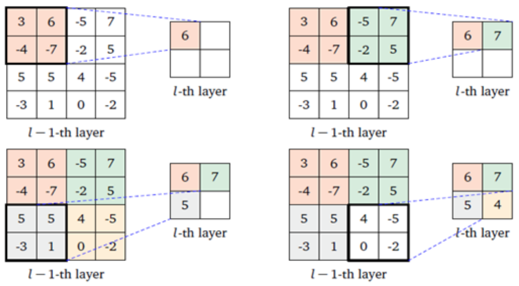


Figure 3.2.5 Illustration of a pooling layer example. The input layer (size: 4× 41) ismax-pooled with filter size $k_l = 2$ and stride $s_l = 2$ into an output layer of size 2×2×1.

3.3 Support Vector Machine

The use of SVMs for facial recognition was pioneered by Osuna et al. Support vector machines (SVMs) provide a novel training paradigm for classifiers such as neural networks, radial basis functions (RBFs), and polynomial functions. SVMs are based on an induction principle known as structural risk minimization, which aims to reduce the maximum amount by which the generalisation error may be wrong. In support vector machines (SVMs), the predicted classification error of the unseen test patterns is minimised by selecting the separation hyper plane. An effective approach for training an SVM for large-scale issues was created by Osunaet al., and they used it to solve the challenge of face identification. From the results of two test sets totaling 10,000,000 test patterns of 19 X 19 pixels, their method is marginally more accurate and runs around 30 times quicker than Sung and Poggio's. Similarly, SVMs have been used for face and pedestrian identification in the wavelet domain.

Support vector machines (SVMs) are a kind of supervised learning model that may be used in conjunction with a learning algorithm to analyse data for classification and regression purposes. Examples that have previously been assigned to one of two classes are called training examples, and the SVM algorithm utilises these labels to determine which class a new example should be assigned to. In order to classify data, it is necessary to find a separation hyperplane between the various sets depending on their characteristics.

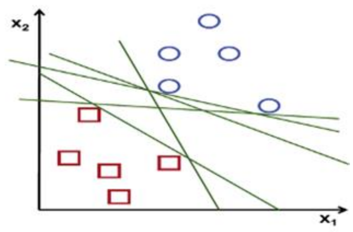


Figure 3.2.5 Support Vector Machine

3.4 Feature Extraction

Here, we employed Our convolutional neural network for face identification uses CNN to extract information from facial pictures. In the realm of computer vision, CNN serves as a feature descriptor for the purpose of locating objects. Fuzzy Neural Networks and Support Vector Machines Are Used for Classification. Using a grid-based image-slicing method, we can calculate the orientation of each pixel and store it in a histogram. At last, the histograms of all the cells are added together to get the finished product. vector.

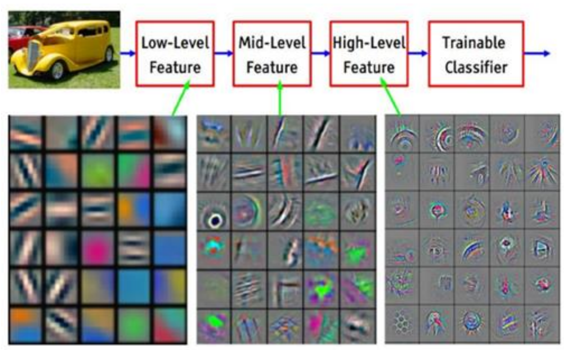


Figure 3.3Hierarchical feature extraction of a convolutional neural network

An information hierarchy is used to extract data in a convolutional neural network [Zeiler et al., 2014]. Extracting elementary characteristics like borders and colour blobs is the job of the first layers. Later layers combine features derived from earlier levels using a linear combination of those characteristics. The image's highest-level characteristics are retrieved in the last convolutional layer. The hierarchy of extracted features is shown in Figure 3.1. An example of a multi-layered convolutional neural network is shown in the top row. Some low-level characteristics extracted by each layer are shown below. Color blobs and borders are extracted in the early levels, while more complex shapes, such circles, are extracted in the intermediate layers. Following that, we pull out the things that a classifier can presumably tell apart linearly (i.e., the final fully-connected layer).

Here, we employed Our face photographs have been processed using CNN to isolate distinctive characteristics. Face recognition HOG (histogram of oriented gradients). In the discipline of computer vision, HOG is a feature descriptor used for the purpose of object detection. FNN and SVM are used for classification. Using a grid-based image-slicing method, we can calculate the orientation of each pixel and store it in a histogram. At last, the features vector is built by concatenating the histograms of each cell.

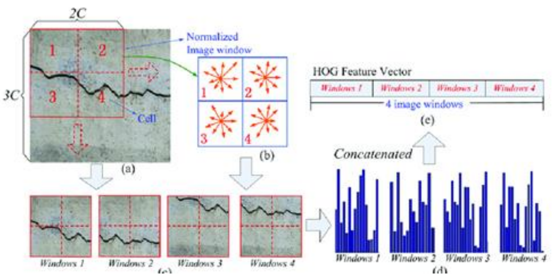


Figure 3.3.1Hierarchical feature extraction of a Histogram of oriented gradients

The information in a Histogram of directed gradients is extracted in a hierarchical fashion [Zeiler et al., 2014]. Extracting elementary characteristics like borders and colour blobs is the job of the first layers. Later layers combine features derived from earlier levels using a linear combination of those characteristics. The image's highest-level characteristics are retrieved in the last convolutional layer. The hierarchy of extracted features is shown in Figure 3.1. An example of a multi-layered convolutional neural network is shown in the top row. Some low-level characteristics extracted by each layer are shown below. Color blobs and borders are extracted in the early levels, while more complex shapes, such circles, are extracted in the intermediate layers. Then, a classifier is used in the hopes of extracting items that can be split apart linearly

3.5 Knearest-neighbor

Learning by analogy is at the heart of k-nearest-neighbor classifiers, which work by comparing a given test tuple with training tuples that are similar to it. When it comes to training tuples, n attributes do the talking. When k is equal to one, the training tuple class that is most similar to the unknown tuple in pattern space is used to label it. In the wavelet domain, KNNs have been employed for face and pedestrian detection. K-Nearest Neighbors (KNN) are supervised learning models that use corresponding learning algorithms for the purpose of conducting classification and regression analyses.

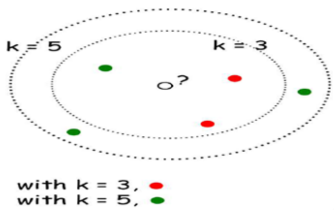


Figure4.2.2 Example of KNN Model

K-Nearest Neighbor models and their corresponding learning methods are employed in classification and regression analysis, both of which fall under the umbrella of supervised.

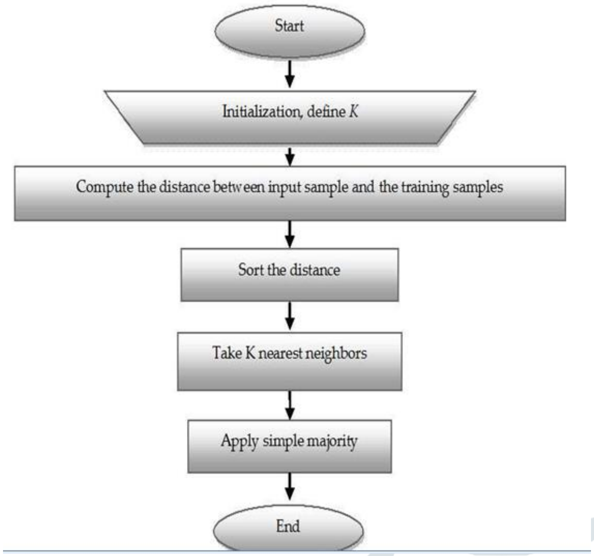


Fig4.2.3 Flowchart of KNN Classifier

3.6 Random Forest Classifier

The Random Forest Algorithm is made up of several decision trees with the same nodes but different input leading to distinct leaves. It does this by combining the results of many decision trees into a single response that is the mean of the individual results. The random forest method is an example of a supervised learning model, since it can "learn" how to categorise new data by analysing existing data that has already been classified. The Random Forest Algorithm is a versatile model used often by engineers since it can be used to a broad variety of issues ranging from regression to classification. Ensemble learning, of which Random Forest is a kind, allows for more precise results to be generated by combining the outputs of numerous models. A conclusion is reached by the algorithm based on the leaves, or the ultimate choices, of each node. Since it takes into account the outcomes of several decision trees and finds an average, the model's accuracy is improved.

Data analysis for classification problems may be performed using RFs, which are supervised learning models accompanied by learning algorithms.

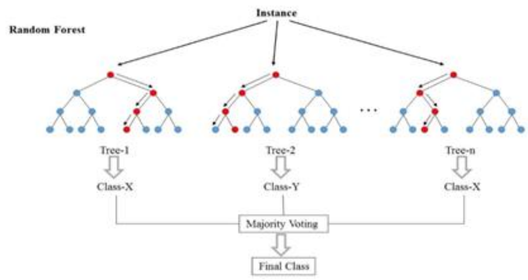


Figure4.2.4 Model of RFC

3.7 Ensemble learning

To effectively address a specific challenge in artificial intelligence, ensemble learning generates and combines numerous models (e.g., classifiers or experts) in a planned manner. You can utilise ensemble learning to become better at something. Through the use of a number of different models, ensemble learning may assist in boosting the accuracy of machine learning. By combining several models, superior predictive performance may be achieved compared to that of a single model. The fundamental concept is to get knowledge from a group of specialists (classifiers), then solicit their input.

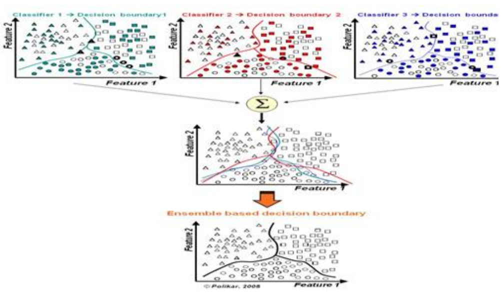
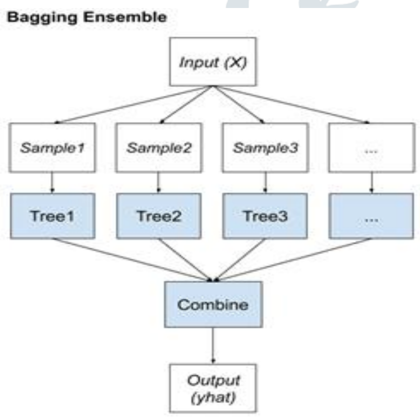


Figure4.3.1 Combining an ensemble of classifiers for reducing classification.

Ensemble learning may also be used to provide more weight to a model's conclusion, merge disparate data sets, learn from mistakes, and repair mistakes. Although the classification-related applications of ensemble learning discussed in this article are the primary emphasis of this article, the basic principles discussed below may be applied to issues involving the approximation or prediction of functions with little effort.

Given a data set and a learning algorithm that can only produce single predictive models, bagging may generate a wide variety of models by giving the learning algorithm varying uniform samples of the data set. It is crucial to have a thorough grasp of bagging, stacking, and boosting, the three major classes of ensemble learning algorithms, and to take into account each of these approaches in your predictive modelling project.



Bag-of-Gauss-Members, or "Bagging" for short, is a technique for ensemble learning in which the training data is changed to create a more varied set of ensemble members. The term "Bagging" was originally an acronym for "Bootstrap AGG regat ING." Bagging relies on the bootstrap and an aggregation technique, as the name suggests. These methods, which have become "standard" in ensemble learning, are those:

- Bagging.
- Stacking.

Boosting

4. Proposed model and Extension, Obtained results

In addition to its usefulness in biometrics, facial recognition also has various security and surveillance-related applications. In recent years, advancements have been made in the field of biometric identification technology, and one of the most promising is face recognition. When compared to other forms of biometric identification (fingerprints, irises, voices, etc.), facial recognition is more straightforward, simple to use, and practical. The purpose of face identification is to pinpoint a specific face in a picture, whereas face verification checks a person's assertion of who they are based on a photograph. In this work, we provide an unique "Gabor DCNN (GDCNN) ensemble" FR approach that uses several Gabor face representations as inputs throughout the DCNN ensemble's training and evaluation processes.

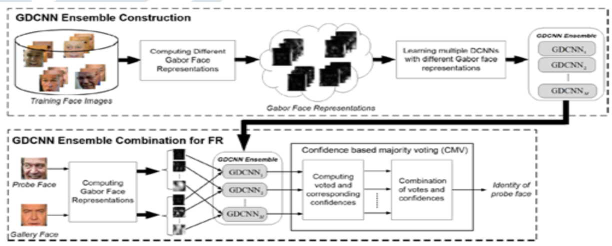


Fig4.4.1 Block diagram of proposed model

4.1Proposed MethodologyDataset used

There are 10,575 individuals in the CASIA-WebFace collection, which consists of 494,414 face photos. To achieve this goal, we assembled a large-scale Gabor face representation dataset of about 10,000 photos. In this case, we use Gabor pictures collected from 10,000 people.

If you want to save a 120x120 Gabor representation picture, you'll need around 15GB of RAM. For the training phase, we used the typical batch size of 128.

4.2 Gabor Face Representation

The Gabor filter examines the picture surrounding the analysis point to see whether it has any high-frequency material in a certain direction. The sufficiency of spatial frequencies of grayscale pixel values might be resolved utilizing Gabor channels, considering the calculation of Gabor face portrayals for each given face picture I. We may suppose, without limiting ourselves, that I is a two-dimensional grayscale picture. The following procedure is used to produce Gabor face representations with a given u and v using the Gabor filter Ψ_(u,v) (z) .

In this case, we assume that U orientations and V scales are provided, and we derive a set of UV Gabor filters, which represents faces in UV space.

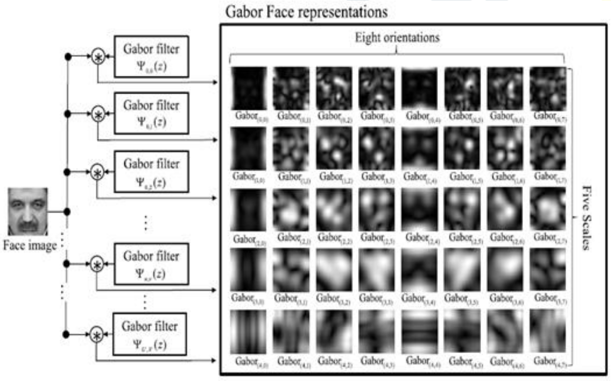


Figure4.4.2 Visual representation of gabor face filter

4.3 Deep CNN

In particular, DCNN finds widespread usage in image-related applications including classification, identification, segmentation, etc. Much of the time, a convolution layer, a pooling layer, and a completely associated (FC) layer will make up the foundation of a profound convolutional brain organization (DCNN) model. The convolution layer might separate a component map from the info map by playing out a convolution on the info map. To remove valuable elements while overlooking less helpful ones, the pooling layer use the component (actuation) map created by the former convolution layer to achieve aspect decrease and info space reflection through sub-inspecting. When computing any part of the output, it is necessary to have all other parts of the input as well, that part of the stack is called the FC layer.

The objective of the FC layer is to make benefit of all suitable appropriated portrayals (highlights) to assemble highlights with further developed capacities in the following layer. In order to extract the results of the classification or recognition, a softmax layer is commonly employed as the last layer of a deep convolutional neural network (DCNN). After a deep convolutional neural network (DCNN) model has been trained for one purpose, it may be used for another by fine-tuning it using data from different datasets. The features of this DCNN are extracted using a Residual network.

4.5 Extension using ResNet

The goal of the FC layer is to build features with enhanced capabilities in the succeeding layer using all of the features (distributed representations) in the current layer. In order to get the ultimate recognition or classification results, a deep convolutional neural network (DCNN) often employs a softmax layer as its very last layer. In machine learning, a deep convolutional neural network (DCNN) model trained for one task may be repurposed for another by fine-tuning the model using data from the new task. As for feature extraction, a Residual network is employed in this DCNN.

4.6 Construction of Gabor and DCNN

We suggest using a variety of Gabor face representations that are mutually beneficial to construct GDCNN ensemble members. Therefore, M Gabor face representations Gabor k _(k=1). A sequence of Gabor face representations with different size and orientation parameters arise from applying a transformation to an original RGB (or grayscale) face photo. Each of the M Gabor face representations has a corresponding training set. An individual GDCNN is learned from scratch using as input a training set that contains examples of faces represented by a certain Gabor representation. Weights should be linked to the full-connected layer, and the loss function in the softmax layer should be minimised.

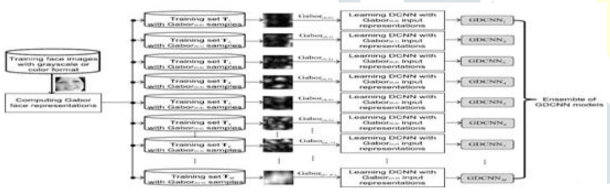


Figure4.4.5 Model of GDCNN

4.7Ensemble of GDCNN

The majority-voting based ensemble combination of GDCNNs is the current gold standard, and this is the approach that was used in developing the suggested technique. Each classifier in an ensemble casts a vote for one class label, and the ensemble then uses the class label that received the most votes in the final output. But our method takes into account the votes of identity (class) labels and the corresponding confidences to adaptively combine the outputs of all GDCNN members to improve recognition performance. One great benefit of this strategy is that it gives greater weight in the voting process to the members of the GDCNN ensemble who have shown superior FR performance

5. Experimental results

5.1 Ensemble of KNN and RFC

The system's efficiency will increase when these two methods are merged. KNN is used for the searching process in a random forest. Following that, the RFC procedure is executed. Random forest is an ensemble model where the individual model is a decision tree and the ensemble approach is bagging.

Step 1:Take n independent samples from the training data.

Step 2 :Constructed from "train" and "decision" trees When training a decision tree, just a random subset of characteristics is utilised to determine the best splits (e.g. 10 features in total, randomly select 5 out of 10 features to split)

Step 3: Each tree makes its own predictions about the test set records and candidates.

Step 4: Speculate to your heart's content.

5.2 Obtained results

Here we got the results up to Knn,Rfc classifiers.

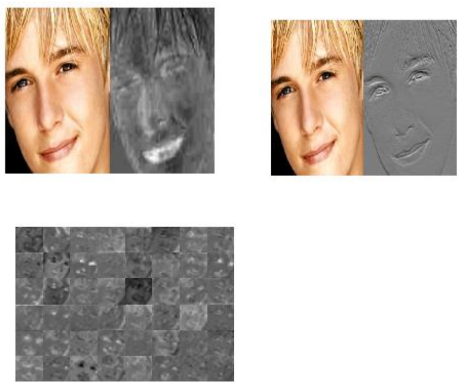
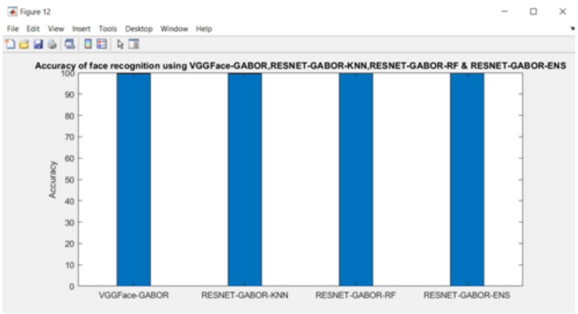


Figure4.5obtained results up to Knn,Rfc classifiers

4.5.1Accuracy of FR using different methods



5.3 Comparison Table

Technique/Parameter	Accuracy (%)
VGG-Gabor Proposed Model (GDCNN)	99.62
ResNet- Gabor (GRNCNN)- KNN	99.70
ResNet- Gabor (GRNCNN)- RFC	99.77
ResNet- Gabor (GRNCNN)- ENS	99.85

6. Conclusion and Future Scope

This evidence suggests that the high performance of our GDCNN ensemble is due to the fact that face photos collected under extreme position and lighting fluctuations still include extremely common activated neurons, which are included in the feature maps generated by our GDCNNs. In addition, Our GDCNNs capture complementary activation patterns from different Gabor face representations, suggesting they may be mutually compensational for improving FR accuracy. This finding can support the value of our GDCNN ensemble approach, which consists of combining a set of DCNNs learned with different Gabor face representations. For image analysis tasks like classification and clustering, deep learning and machine learning have surpassed previous methods as the gold standard in computer vision. However, low-level characteristics similar to Gabor filters seem to be often produced by general-purpose object identification tasks. In this project the deep CNN and gabor filter are ensemble to achieve good results. Along with this the classification techniques like KNN and RFC are used for face recognition. The results obtained using ensemble of KNN and RFC along with GRNCNN has strong generalisation and offers a competitive face recognition solution in both FR-limited and -unlimited settings. The accuracy obtained using GRNCNN and KNN-RF ensemble is 99.95%

Thefacerecognitionsystembyusingadeeplearningnetworkcalledconvolutional neural network, its architecture, different layers and their working, howto design a new CNN as per our requirements with different layers is explained . Also, the training of the network designed with a better learning algorithmalong with testing of the trained network with three different datasets are discussed.

Along with this the classification techniques like KNN and RFC are used for face recognition. The results obtained using ensemble of KNN and RFC along with GRNCNN exhibits good generalization and provides a competitive face identification solution under both constrained and unconstrained FR circumstances. The accuracy obtained using GRNCNN and KNN-RF ensemble is 99.95%

7. Future Scope

Next research will be conducted to enhance the performance of gesture authentication and to study the various gestures. In addition, we'll see how each motion performs in terms of user ease and safety. If this research is to be carried on, we advocate that the Siamese network design be expanded to include more CNNs, in particular convolutional networks that can properly extract both low- and high-level information. Future development may also entail developing a graphical user interface for use with the database to improve its usability and aesthetic appeal.

Although our work has improved the accuracy, there is still room for improvement in the future. The feature extraction by Convolution neural network(CNN), its preprocessing steps like convolution, pooling layer, advantages, disadvantages and how it determines the pattern for a face image are all discussed . Also, the classifier network FFNN, its architecture, its working, the training and testing of the network that is designed for detecting faces are also explained.

The feature extraction by Convolution neural network(CNN), its preprocessing steps like convolution, pooling layer, advantages, disadvantages and how it determines the pattern for a face image are all discussed .Implementing a database with a graphical user interface may improve usability and aesthetics, and this is a potential area of focus for next study.

Next research will be conducted to enhance gesture authentication performance and examine the various gestures. Which gesture offers the best combination of user convenience and system safety is also under investigation. Future work in this area might benefit from the incorporation of additional CNN networks into the Siamese network design, namely convolutional networks that can correctly extract both low- and high-level features. Implementing a database with a graphical user interface may improve usability and aesthetics, and this is a potential area of focus for next study.

Which gesture offers the best combination of user convenience and system safety is also under investigation. Specifically, we advocate for the use of alternative CNN networks in Siamese network design, such as convolutional networks that are able to correctly extract both low- and high-level features.

8. References

M. A. Abuzneid and A. Mahmood, "Enhanced Human Face Recognition UsingLBPH Descriptor, Multi-KNN, and BackPropagation Neural Network," IEEE Access,vol.6, pp. 20641-20651,2018.

A. Albiol, D. Monzo, A. Martin, J. Sastre, and A. Albiol, "Face recognition usingHOG-EBGM",PatternRecognitionLetters,vol.29,no.10, pp.1537–1543,2008.

Bach, F., Jordan, M.: 'Kernel independent component analysis', J. Mach. Learn.Res., 2003,3, pp. 1–48.

S. Belongie, J. Malik, and J. Puzicha, "Shape matchingandobjectrecognitionusingshapecontexts",IEEETransactionsonPatternAnalysisandMachineIntelligence,vol. 24, no. 4, pp. 509–522, 2002.

K. Bong, S. Choi, C. Kim, D. Han, and H. Yoo, "A Low-Power ConvolutionalNeural Network Face Recognition Processor and a CIS Integrated With Always-onFace Detector,"IEEEJournalofSolidStateCircuits,vol.53 ,pp.115-123,2018.

Brunelli,R.andT.Poggio,1993.Facerecognition:Featuresversustemplates.

IEEETrans.Patt.Anal.Mach.Intel.,15:1042-1052. DOI:10.1109/34.254061121 J. Canny, A computational approach to edge detection, IEEE Trans.Patt. Anal.Machine Intell. 8 (1986)679-698.

Q. Chen, H. Wu, M. Yachida, Face detection by fuzzy pattern matching, in: F'roc.5thInt.Conf. onComputerVision,MIT,Boston,1995, pp. 591-596.

S. Chintalapati and M. Raghunadh, "Automated attendance management systembased on face recognition algorithms", in Computational IntelligenceandComputingResearch(ICCIC),2013 IEEEInternationalConferenceon,pp.15,IEEE,2013 .

N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection",in IEEE Computer Society Conference on Computer Vision and Pattern Recognition,2005.CVPR2005. vol. 1, pp. 886–893, IEEE, 2005.

O. Deniz, G. Bueno, J. Salido, and F. De la Torre, "Face recognition usinghistograms of oriented gradients", Pattern Recognition Letters, vol. 32, no. 12, pp.1598–1603, 2011.

C. Geng and X. Jiang, "Face recognition using sift features",inImageProcessing(ICIP),200916thIEEE InternationalConferenceon,pp.33133316,IEEE,2009.

Hasan RI, Yusuf SM, Alzubaidi L. Review of the state of the art of deep learningforplant diseases: abroadanalysisanddiscussion.Plants. 2020;9(10):1302.

M. A. Abuzneid and A. Mahmood, "Enhanced Human Face Recognition UsingLBPH Descriptor, Multi-KNN, and BackPropagation Neural Network," IEEE Access,vol.6, pp. 20641-20651,2018.

A. Albiol, D. Monzo, A. Martin, J. Sastre, and A. Albiol, "Face recognition usingHOG-EBGM",PatternRecognitionLetters,vol.29,no.10, pp.1537–1543,2008.

Bach, F., Jordan, M.: 'Kernel independent component analysis', J. Mach. Learn.Res., 2003,3, pp. 1–48.

S. Belongie, J. Malik, and J. Puzicha, "Shape matchingandobjectrecognitionusingshapecontexts",IEEETransactionsonPatternAnalysisandMachineIntelligence,vol. 24, no. 4, pp. 509–522, 2002.

K. Bong, S. Choi, C. Kim, D. Han, and H. Yoo, "A Low-Power ConvolutionalNeural Network Face Recognition Processor and a CIS Integrated WithAlways-onFace Detector,"IEEEJournalofSolidStateCircuits,vol.53 ,pp.115-123,2018.

Brunelli,R.andT.Poggio,1993.Facerecognition:Featuresversustemplates.

IEEETrans.Patt.Anal.Mach.Intel.,15:1042-1052. DOI:10.1109/34.254061

121 J. Canny, A computational approach to edge detection, IEEE Trans.Patt. Anal.Machine Intell. 8 (1986)679-698.

Q. Chen, H. Wu, M. Yachida, Face detection by fuzzy pattern matching, in: F'roc.5thInt.Conf. onComputerVision,MIT,Boston,1995, pp. 591-596.

S. Chintalapati and M. Raghunadh, "Automated attendance management systembased on face recognition algorithms", in Computational IntelligenceandComputingResearch(ICCIC),2013 IEEEInternationalConferenceon,pp.15,IEEE,2013 .

N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection",in IEEE Computer

Society Conference on Computer Vision and Pattern Recognition,2005.CVPR2005. vol. 1, pp. 886–893, IEEE, 2005.

O. Deniz, G. Bueno, J. Salido, and F. De la Torre, “Face recognition usinghistograms of oriented gradients”, Pattern Recognition Letters, vol. 32, no. 12, pp.1598–1603, 2011.

C. Geng and X. Jiang, “Face recognition using sift features”,inImageProcessing(ICIP),200916thIEEE EInternationalConferenceon,pp.33133316,IEEE,2009.

Hasan RI, Yusuf SM, Alzubaidi L. Review of the state of the art of deep learningforplant diseases: abroadanalysisanddiscussion.Plants. 2020;9(10):1302.

Bach, F., Jordan, M.: ‘Kernel independent component analysis', J. Mach. Learn.Res., 2003,3, pp. 1–48.

S. Belongie, J. Malik, and J. Puzicha, “Shape matchingandobjectrecognitionusingshapecontexts”,IEEETransactionsonPatternAnalysisandMachineIntelligence,vol. 24, no. 4, pp. 509–522, 2002.

K. Bong, S. Choi, C. Kim, D. Han, and H. Yoo, "A Low-Power ConvolutionalNeural Network Face Recognition Processor and a CIS Integrated WithAlways-onFace Detector,"IEEEJournalofSolidStateCircuits,vol.53,pp.115-123,2018.

M. A. Abuzneid and A. Mahmood, "Enhanced Human Face Recognition UsingLBPH Descriptor, Multi-KNN, and BackPropagation Neural Network," IEEE Access,vol.6, pp. 20641-20651,2018.

A. Albiol, D. Monzo, A. Martin, J. Sastre, and A. Albiol, “Face recognition usingHOG-EBGM”,PatternRecognitionLetters,vol.29,no.10, pp.1537–1543,2008.