A

Project Report

On

" A deep facial recognition system using computational intelligent algorithms "

Submitted in fullfilment for

ADVANCED MACHINE LEARNING (MIS-64061-001)

In

Master Of Science in Business Analytics (MSBA)

By

Ms. Sabiha Abubakar Mhatarnaik

Under the Guidance of :

Dr. Murali Shanker
Professor, Department of Management and Information Systems Kent State University

CERTIFICATE

It is hereby certified that the project work entitled "**A deep facial recognition system using computational intelligent algorithms**" is a bonafide work carried out by **Sabiha Abubakar Mhatarnaik** in complete fulfilment for **Advanced Machine Learning (MIS-64061-001)** in **Master of Science in Business Analytics** of **Kent State University** during the year 2022. It is certified that all corrections/suggestions indicated for Internal Assessment have been incorporated in the Report submitted. The project report has been approved as it satisfies the academic requirements in respect of project work prescribed for the said tenure.

………………………………	………………………………
Signature of Guide	Signature of Program Director
Prof. Murali Shanker	Dr. Rouzbeh Razavi

ABSTRACT

Biometric applications, such as facial recognition (FR), are becoming increasingly crucial in smart cities. Many scientists and engineers all around the world have been working to develop more reliable and precise algorithms and procedures for these systems and their applications in everyday life. FR is working on technologies with a variety of real-time applications. The purpose of this research is to create a comprehensive FR system in fog computing and cloud computing utilizing transfer learning. Because of the dominating representation, the created system uses deep convolutional neural networks (DCNN); nonetheless, occlusions, expressions, illuminations, and position can all affect the deep FR performance. To extract relevant facial features, DCNN is utilized. These traits enable us to quickly compare faces between them. By integrating the new persons, it analyzes and improving its predictions on the ones it currently has, the system may be trained to recognize a group of people and to learn via an online way. Three common machine learning methods (Decision Tree (DT), K Nearest Neighbor (KNN), and Support Vector Machine (SVM)) were used to test the suggested recognition approach. The suggested system was tested on three face picture datasets (SDUMLAHMT, 113, and CASIA) using accuracy, precision, sensitivity, specificity, and time performance criteria. The experimental results demonstrate that the suggested method outperforms competing methods across all parameters. The suggested algorithm outperforms the comparative algorithms in terms of accuracy (99.06 percent), precision (99.12 percent), recall (99.07 percent), and specificity (99.10 percent).

TABLE OF CONTENTS

LIST OF FIGURES

INTRODUCTION

The human face is regarded as the most important feature of the body.

According to studies, even a face can communicate and has different words for different emotions. It is essential for engaging with others in society. It can be used as a key for security solutions in many businesses because it conveys people's identities. Face recognition (FR) is becoming increasingly popular around the world as a highly secure and trustworthy security solution. Because of its high level of security and reliability, it is acquiring major importance and attention from hundreds of corporate and government entities. Furthermore, as compared to other biometric security methods such as palmprints and fingerprints, the FR system offers significant advantages. The technology takes biometric measurements of a person at a set distance without having to interact with them. This approach can assist many companies in identifying a person with a criminal record or other legal difficulties in crime deterrent applications. As a result, this technology is becoming increasingly important for many residential structures and businesses. This method relies on the capacity to detect a human face and then compare its many traits to previously recorded faces. This characteristic elevates the system's prominence and allows it to be widely used around the world. It is designed with user-friendly features and operations that include a variety of facial nodal points. A face has around 80 to 90 distinct nodal points. The FR system uses these nodal points to quantify important features such as the distance between the eyes, the length of the jawline, the curvature of the cheekbones, and the depth of the eyes. These points are calculated by establishing a faceprint code, which symbolizes the face's identity in the computer database. With the advancement of technology, systems based on 2D graphics can now be upgraded to 3D graphics, which improves accuracy and reliability.

Biometrics is the science and technology of measuring and analyzing biological data statistically. They are observable behavioral and/or physiological traits that can be used to confirm an individual's identity. For verification, a unique biometric could be utilized for each individual. Biometric technologies are increasingly being employed in a variety of industries, including jail security, protected access, and forensics. Biometric systems use biological traits such as the face, hand shape, iris, retina, and fingerprints to distinguish people via authentication. The FR system is a more natural biometric information processing system with greater variance than any other method. As a result, FR has emerged as a hot topic in computer science, particularly in relation to biometrics and machine learning. Machine learning is primarily concerned with developing methods for task training—it is related to the fields of computational statistics and mathematical optimization. Reinforcement learning, supervised learning, virtually supervised learning, and unsupervised learning are examples of machine learning approaches. Many things that humans believe they can only accomplish themselves, such as playing games, learning subjects, and recognition, can be automated using machine learning. **Because most machine learning algorithms take a lot of resources, it's ideal to run them in a distributed environment like cloud computing, fog computing, or edge computing.**

Cloud computing is built on the sharing of numerous resources, such as services, applications, storage, servers, and networks, in order to achieve economies and consistency, and therefore to provide the best concentration to maximize the efficiency of using shared resources. Many networks edge functions, such as data storage, processing, data provisioning, and application services for end users who can be introduced to the network edge, are included in fog computing. These environments would reduce total resource consumption, shorten work completion times, and cut prices through pay-per-use. The major objectives of this research are to develop a deep FR system in fog computing via transfer learning. This system employs deep convolutional neural networks (DCNN) and machine learning algorithms. The proposed approaches will be able to gather a person's biometric measurements from a certain distance for crime prevention reasons without having to engage with them. As a result, numerous organizations can use the recommended approaches to identify someone with a criminal past or other legal difficulties.

LITERATURE REVIEW

Many researchers have worked on pattern recognition and identification via various biometrics utilizing various constructing mining model methodologies as a result of the considerable development of machine learning, the computer environment, and recognition systems. Saliency maps are employed in the proposed model for facial expression identification to transfer knowledge from an arbitrary source to a target network by primarily "hiding" non-relevant information. Because the experience is only communicated through augmentation of the input data, the proposed method is independent of the model used. When the proposed model was pushed to focus on the parts of the input that were regarded relevant sources, the evaluation revealed that the new model was able to adapt to the new domain faster.

R. Prakash et al [1]. suggested an automatic facial recognition method based on a transfer learning methodology and a Convolutional Neural Network (CNN). CNN with weights learned from the VGG-16 pre-trained model. For classification, the collected features are passed into the Fully connected layer and SoftMax activation. The proposed approach is tested against two publicly available databases of face images: Yale and AT&T. Face recognition accuracy of 100% for AT&T database face photos and 96.5 percent for Yale database face images has been reached. In comparison to the PCA approach, the results reveal that face recognition utilizing CNN with transfer learning delivers superior classification accuracy.

Masa et al. [2] proposed to build prepared information sizes for face acknowledgment frameworks: domain explicit information development. They presented techniques to enhance realistic datasets with critical facial varieties by controlling the faces in the datasets while coordinating inquiry pictures presented by standard convolutional neural systems. They tested their framework against the LFW and IJB-A benchmarks and Janus CS2 on a large number of downloaded pictures. They reported the standard convention for unhindered, marked outside information and announced a mean grouping precision of 100% equal error rate.

A. Deep feature extraction: Network architecture

The architectures are divided into two types: backbone and assembled networks, and are based on ImageNet's success as well as traditional CNN architectures like Senet, Reset, Google Net, and Vignette. It's also used in FR as a baseline model, either fully or partially implemented. FR is still utilized as an architecture design to improve efficiency in addition to the mainstream ways. FR approaches can also be applied in assembled networks, possibly with several jobs or many inputs, using backbone networks as fundamental components. Each network is associated with a specific sort of input or job. After the results of assembled networks are collected Loss Function, higher performance is achieved during adoption. A supervisory signal uses SoftMax loss as an organizing object, which improves the variation in the features. SoftMax loss becomes ineffective for FR when intervisitations are larger than intervariations.

- **Euclidean-distance-based loss:**

  The Euclidean distance is used for intravariance compression and intervariance enlargement.

- **Angular/cosine-margin-based loss:**

  Discriminative face feature learning is done based on angular similarity, with prominent and potentially substantial angular/cosine separability between the selected representations.

- **SoftMax loss and its variations:**

  Performance is enhanced by using SoftMax loss or a modification of it.

B. Face matching by deep features.

The passed images through the networks after training the deep networks to work with huge data and an appropriate loss function. The most frequent approaches for computing feature similarity are L2 distance or cosine distance; however, for identification and verification tasks, the closest neighbor (NN) and threshold comparison are utilized. Many more methods, such as sparse representation-based classifier (SRC) and metric learning, are utilized to analyse the deep data and compute face matching with high accuracy.

Faceprocessing algorithms can manage variations in poses, expressions, and occlusions, and FR is a developed object categorization. Cross posture FR, cross-age FR, and video FR are only a few of the more intricate types of FR that are based on real-world features. To imitate scenes from reality, more realistic datasets are sometimes created.

PROPOSED SYSTEM

1. Traditional deep convolutional neural networks

Images are measured in width (W) 227, height (H) 227, and depth (D) 3 of the colors red, green, and blue; as a result, they are 2272273. The first convolutional layer filters the input color image. This layer is made up of 96 kernels (K), an 11x 11x11 filter (F), and a 4-pixel stride (s). The stride in the kernel map is the distance between adjoining neurons' receptive field centers. The output size of the convolutional layer is calculated using the mathematical formula ((W-F+2P)/S) +1, where P is the padded pixel number, which can be as low as zero. The convolutional layer's output volume size is ((227–11+0)/ 4)+1 = 55.The number of filters in this layer is 256 because the second input of the convolutional layer has a size of 5555 no of filters. Because the layers' work is split over two GPUs, the burden is shared by two across all layers in each GPU. The convolutional layer comes next, followed by the pooling layer. The dimensionality of each feature map is reduced while crucial characteristics are kept. Pooling can be sum, maximum, average, and so on. A maxpooling layer is used in AlexNet. This layer receives 256 filters (256 inputs).

AlexNet was created by Krizhevsky et al. for the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC). The input image is filtered using AlexNet's first layer. The input image has dimensions of 2272273 in height (H), width (W), and depth (D); D = 3 accounts for the colors red, green, and blue. The input color image is filtered using the first convolutional layer, which contains 96 kernels, an 11x11x11 filter (F), and a fourpixel stride (s). The stride is the distance between neighboring neurons' receptive field centers in the kernel map. The convolutional layer's output size is computed using the formula ((W-F+2P)/S) +1, where P is the padded pixel number, which can be as low as zero.The output volume of the convolutional layer is ((227–11+0)/4)+1 = 55. The convolutional layer's second input has a size of 5555number of filters, and this layer likewise has 256 filters. Because these layers' work is split between two GPUs, each layer's load is divided by two. The convolutional layer comes next, followed by the pooling layer. The dimensionality of each feature map lowers, while significant features are kept. Max, sum, average, and more pooling methods are available. AlexNet makes use of a maxpooling layer. This layer has a total of 256 filters as input. Each filter has a size of 55256 pixels with a two-pixel stride.The task is divided into 55/255/2256/2=2727128 inputs for each GPU when two GPUs are used. The third layer, which has 384 kernels with a size of 33, is connected to the second convolutional layer's normalized output. There are 384 kernels of size 33 in the fourth convolutional layer, which are split over two GPUs, resulting in a load of 331192 per GPU. The fifth convolutional layer contains 256 kernels, each of size 33, distributed across two GPUs, giving each GPU a load of 33128. The last three convolutional layers are built without the use of any pooling layers or normalization.The outputs of these three layers are fed into two completely connected layers with a total of 4096 neurons each. The architecture utilized in AlexNet to categorize distinct classes using ImageNet as a training dataset is shown in Figure. DCNNs can learn hierarchically from features. A DCNN improves picture classification accuracy, particularly when dealing with huge datasets. Because a DCNN requires a large number of images to achieve high classification rates, a lack of color images among the subjects' identifying photos presents an additional obstacle for recognition systems.A DCNN is made up of neural networks with convolutional layers that extract and classify features from images [37]. By employing a training set with various sizes or scales but the same features, the disparity between the information utilized for testing and the original data used to train the DCNN is minimized.

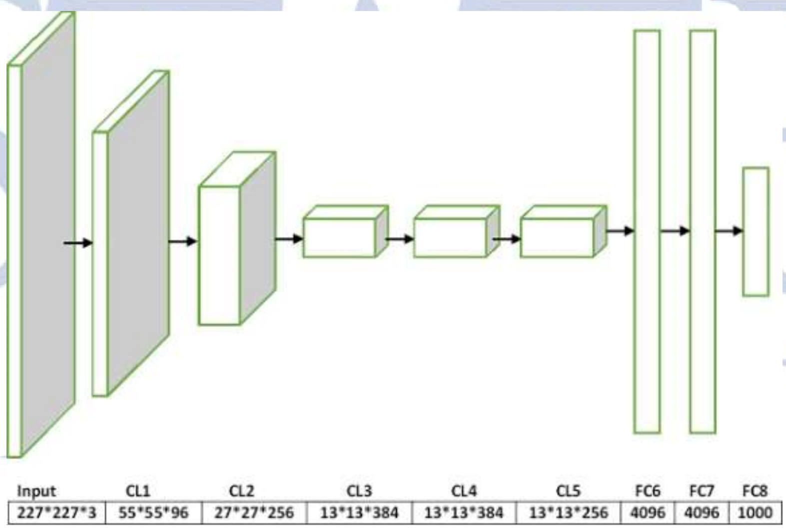

Fig 1. AlexNet architecture

2. Fundamentals of transfer learning

The center's transfer learning (TL) information can be found in. The center employs a relatively complex and fruitful pre-prepared model derived from a huge data source, such as ImageNet, a large visual database designed for visual object recognition research. It has about 14 million carefully annotated photographs, with one million of them having bounding boxes. There are around 20,000 classifications in ImageNet. Pretrained models are typically built using a 1,000-class subset of ImageNet. To eliminate a little amount of private material, we "shifted" the scholarly information to the moderately reorganized assignments (e.g., characterizing liquor misuse and nonliquor addiction).Two characteristics are required to support the exchange: 1)With the exhausting hyperparameter tuning of fresh endeavors, the achievement of the pretrained model can advance the prohibition of client mediation; 2) Pretrained models' early layers can be resolved as highlight extractors, which help separate low-level highlights such as edges, tints, hues, and surfaces. TL retrains the new layers as is customary. The pre-trained model is used first, and then the full neural system structure is reconstructed. Importantly, the global learning rate is constant, therefore moving layers will have a low factor, while freshly added layers will have a high factor. Figure 1 depicts TL's core knowledge.

3. Adaptive deep convolutional neural networks (the proposed face recognition system)

The proposed system consists of three essential stages, including

- **Preprocessing,**
- **Feature Extraction**
- **Recognition ,and identification.**

The frame begins to capture photos that must feature a human face as the topic of insertion during preprocessing. The face detecting module receives this image. The face detector does not detect human faces and instead uses a segment bit as a region of interest. The preprocessing processes are continued once the ROI is obtained. To ensure proper alignment, it is scaled to the preretinal size.

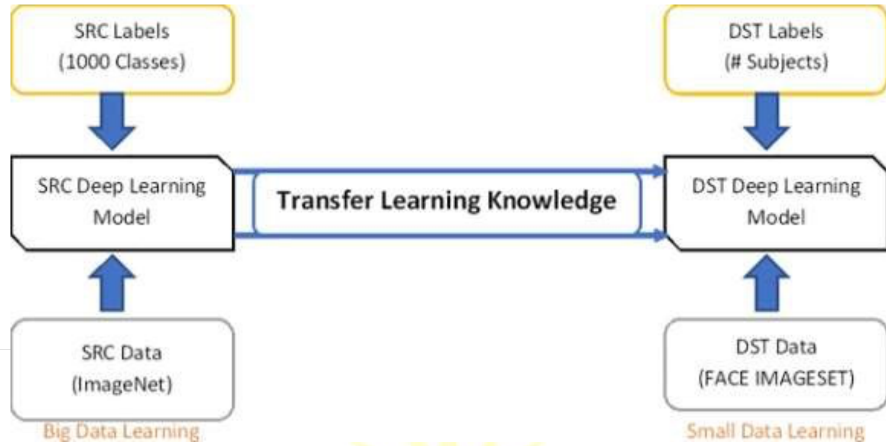

Fig 2. Core knowledge of transfer learning.

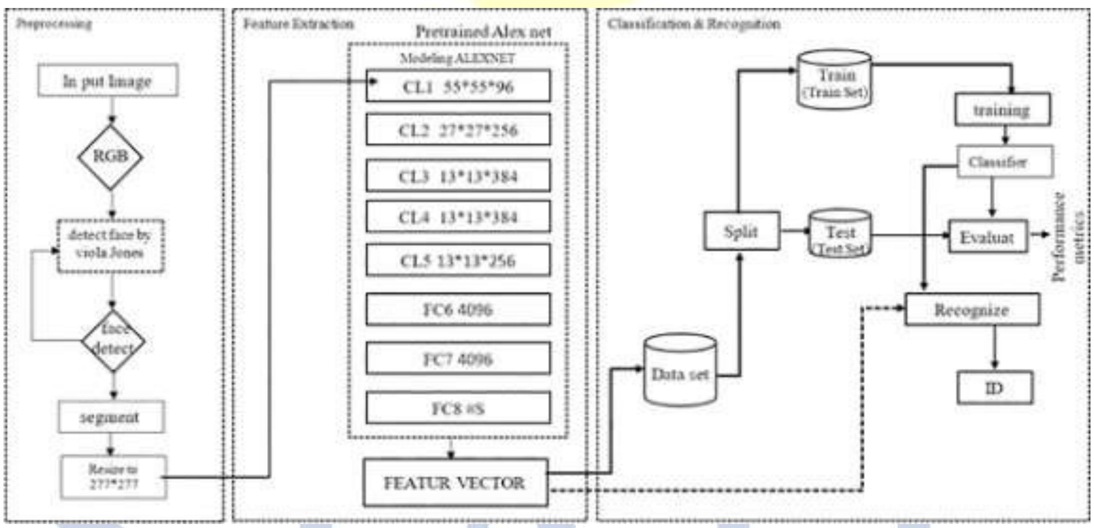

Fig 3. The general overall view of the proposed face recognition system.

The preprocessed ROI is used to extract the feature vector using a modified version of AlexNet during feature extraction. The extract vector is a representation of the image's important information.

Finally, the recognition and identification process includes determining which feature vector in the system's database belongs to which enrolled subject. Every new feature vector reflects a new or previously registered subject. The system recognizes the related ID for the feature vector of a ready register subject. The system creates a new record in the connected database for the feature vector of a new registered subject. The proposed facial recognition system is depicted in the following diagram. To extract the distinguishing traits of each face, the system performs the following procedures on the face images:

- **Pre-processing Phase:**

The system begins to ensure that the input image is the RGP image in the preprocessing step, as illustrated in Fig . Align the image to the same size. The facial detection procedure is then carried out. This stage employs the Viola-Jones detection technique, a well-known face detection mechanism. Viola-Jones detection's appeal arises from its ability to function in real-time while maintaining excellent accuracy. This face detector scans the input image using detection windows of various sizes to detect faces in a specific image.

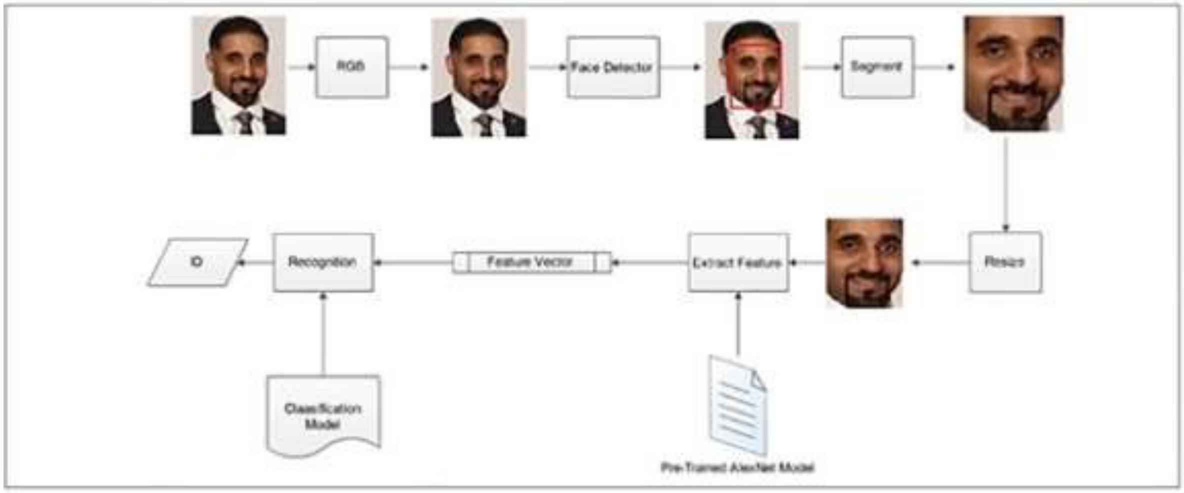

Fig 4. Block diagram of the proposed biometric system (images from dataset published)

The presence of a face window is determined during this phase. Simple local characteristics are derived using Haar-like filters and applied to face window candidates. The feature values of Haar-like filters are easily determined by calculating the difference between the total light intensity of the pixels. After that, crop and resize the facial image to 227 227 to segment the problem area.

- **Features Extraction using Pre-trained Alex Network**

The available dataset size is insufficient to create a new deep model from the beginning, and in any case, this is impossible due to the high number of prepared images. To keep the test objective, we applied the exchange learning hypothesis to AlexNet's pre-pared engineering in three different ways. We expected to change the structure first. Since the first FCLs were established to complete 1,000 classifications, the last fully-connected layer (FCL) has been updated. The scale, hairdresser chair, lorikeet, small poodle, Maltese dog, dark-striped cat, beer bottle, work station, necktie, trombone, protective crash helmet, cucumber, letterbox, pomegranate, Appenzeller, gag, snow panther, mountain bike, lock, and Diamondback were among the twenty classes randomly chosen.We noticed that the facial recognition system failed to identify any of them. As a result, we couldn't use AlexNet as an element extractor. As a result, the calibration was crucial. We expected to have to change the relating softmax layer and arrangement layer, as shown in Fig. 1, because the length of yield neurons (1000) in standard AlexNet is not equal to the number of classes in our challenge .

Another randomly introduced entirely associated layer with a number of accessible subjects in the used dataset(s), a softmax layer, and another characterisation layer with a comparable number of competitors were used in our exchange learning strategy. Figure 8 depicts the many types of activation functions accessible; we chose softmax because we had varied information and options based on the most extreme scores of various outputs. The training options are then set. Before training, three attributes were examined. For exchange learning, the total number of training iterations should be kept low. The number of training iterations was initially fixed to six.

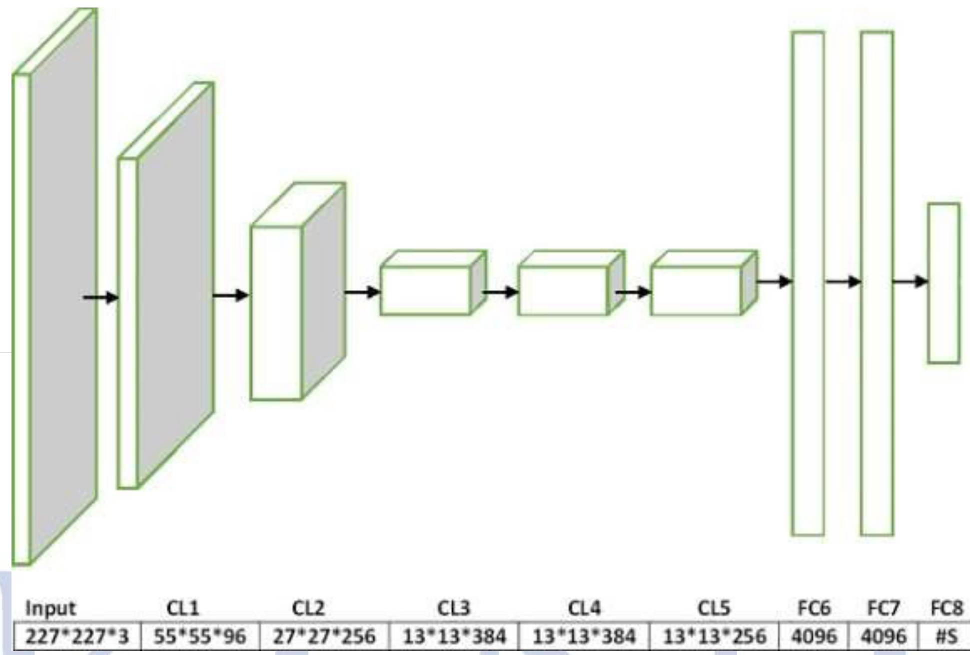

Fig 5. The schema of the modified AlexNet, where (#S) is the number of subjects in the dataset used during training.

Fig 6. Different types of activation functions for classification.

Second, because the early layers of this neural system were preprepared, the global learning rate was set to a low estimated value of 104 to turn learning off. Third, because the transfer layers had pre-prepared loads and weights and the new layers had irregularly instated loads and weights, the learning rate of new layers was several times that of the transfer layer. Third, we experimented with different transfer layer quantities and settings. AlexNet has five Conv layers (CL1, CL2, CL3, CL4, and CL5), as well as three fully associated layers (FCL6, FL7, and FL8).The proposed method's pseudocode is provided in algorithm 1. For the participants who were enrolled in the recognition systems, it starts with the original AlexNet architecture and image dataset. Viola-Jones detection is used to detect the subject's face in each image in the dataset. For transfer learning, the new face dataset is employed. We tweak AlexNet's architecture to transfer learning. The redesigned architecture is then trained using the face dataset. The trained model is used to extract features.

As shown in the pseudocode of the suggested calculation, we anticipate to overhaul the related SoftMax layer and arrangement layer (Algorithm 1).

Algorithm 1: Transfer Learning using AlexNet model
Input original AlexNet Net, ImageFaceSet imds
Output modified trained AlexNet FNet, features FSet

```
Begin
// Preprocessing Face image(s) in imds
For i = 1: length(imds)
img read(imds,i)
face detectFace(img)
img resize(face,[227, 227])
save(imds,I,img)
End for
// Adapt AlexNet Structure
FLayers Net.Layers(1:END-3)
FLayers.append(new Convolutional layer)
FLayers.append(new SoftMax layer)
FLayers.append(new Classification layer)
// Train FNet using options
Options.set(SolverOptimizer stochastic gradient descent with momentum)
Options.set(InitialLearnRate 1e-3)
Options.set(LearnRateSchedule Piecewise)
Options.set(MiniBatchSize 32)
Options.set(MaxEpochs 6)
FNet trainNetwork(FLayers, imds, Options)
//Use FNet to extract features
```

```
FSet empty
For j = 1: length(imds)
img read(imds,j)
F extract(FNet, img, 'FC7')
FSet FSet U F
End for
End
```

4. Face recognition Phase using Fog and Cloud Computing

The fog computing face recognition framework is shown in Figure 1. Client devices, cloud nodes/servers, and distributed computing environments all make up fog systems. The following are the main differences from the traditional distributed computing process:

- The distributed computing community supervises and controls a large number of cloud nodes and servers.
- Fognodes/servers at the system's edge, between the community and the client, have a procurement device that can execute preprocessing and highlight extraction operations, as well as securely exchange biometric data with client devices and cloud nodes.
- User devices are distributed and include advanced mobile phones, personal computers (PCs), hubs, and other networkable terminals.

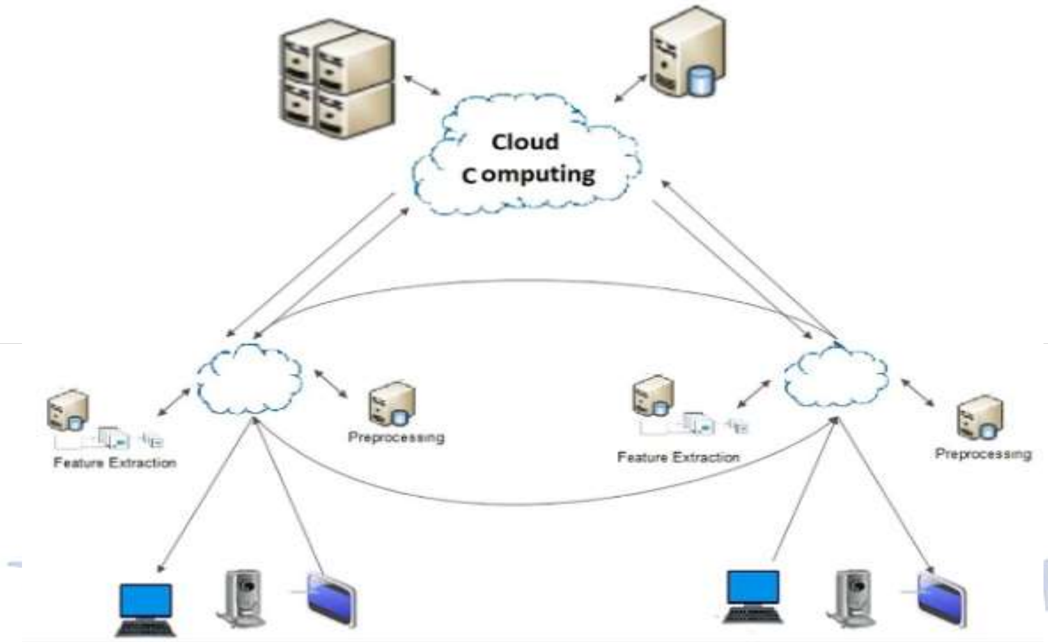

Fig 7. General block diagram of the fog computing FR system.

The communication strategy serves several functions.

- If FR information is transmitted to anode, the system communication cost will increase since all data must be delivered to and prepared by the cloud server. Additionally, the cloud server's calculating burden will grow.
- The cloud community, as the central hub of the entire system, will become a target for attacks from the standpoint of recognition security. Information obtained via fog nodes/servers becomes exposed if the focal hub is breached.
- If a neural system is used for recognition, face recognition datasets are necessary for training. Preparing datasets takes time, and if the training is done solely by the nodes, the training duration will be considerably increased, putting the training quality at jeopardy.

We propose a general engineering methodology for cloud-based face recognition frameworks because the connection between a fog node and client devices is quite unreliable. This strategy takes advantage of the processing power and capacity of fog nodes/servers and cloud servers.

Preprocessing, such as extraction, facial recognition, and recognition-based security, are all included in the design. The plan is divided into six layers, as shown in the fog architecture information stream in Figure

- **User equipment layer:**

  Client devices for FC/MEC are diverse, including PCs and smart terminals. Through various conventions, these devices can use various fog nodes/servers.

- **Network layer:**

  Various fog architecture protocols are used to link administration. It can receive information from the system and client device layer and compress and transfer it,

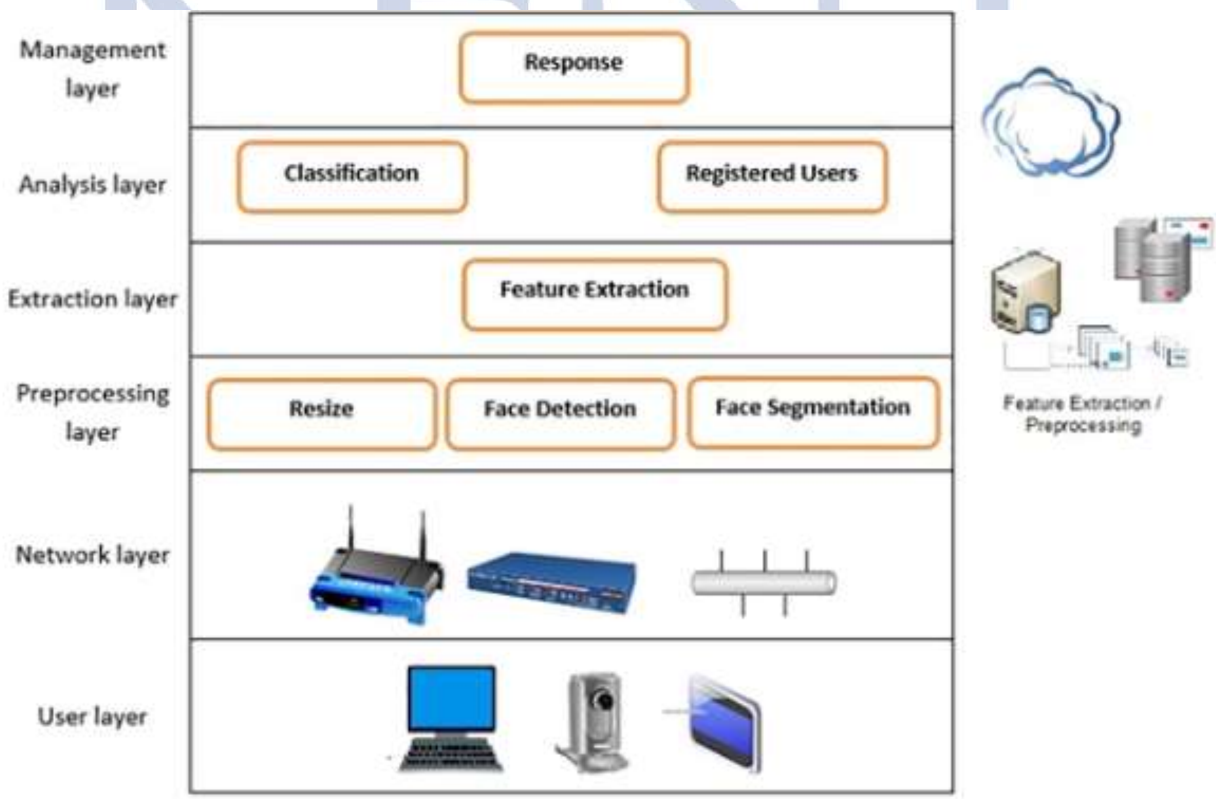

Fig 8. General architecture of the fog computing FR system.

- **Data processing layer:**

This layer's primary function is to preprocess image(s) provided from client hardware, which includes information cleaning, filtering, and preprocessing. This layer's function is carried out on cloud nodes.

- **Extraction layer:**

  The extraction layer uses the associated AlexNet to remove the highlights once the image(s) have been preprocessed.

- **Analysis layer:**

  The cloud is used to communicate with the analysis layer. Its main function is to group the eliminated element vectors discovered by fog nodes/servers. It has the ability to coordinate data across registered clients and respond to queries.

- **Management layer:**

  The cloud server's management is primarily responsible for (1) the face recognition framework's choices and replies, and (2) the information and logs of the fog nodes/servers that may be accessed.

As shown in Fig, the recognition classifier of the Analysis layer is the most significant piece of the framework for data preparation. It is identified with the resulting cloud server response to guarantee the legitimacy of the framework. Relatedly, our work centres around recognition and authentication. Classifiers on fog nodes/servers can utilize their calculation ability and capacity limit for recognition. In any case, much of the scope information cannot be handled or stored because of the restricted calculation and capacity of fog nodes/servers a distinct system.

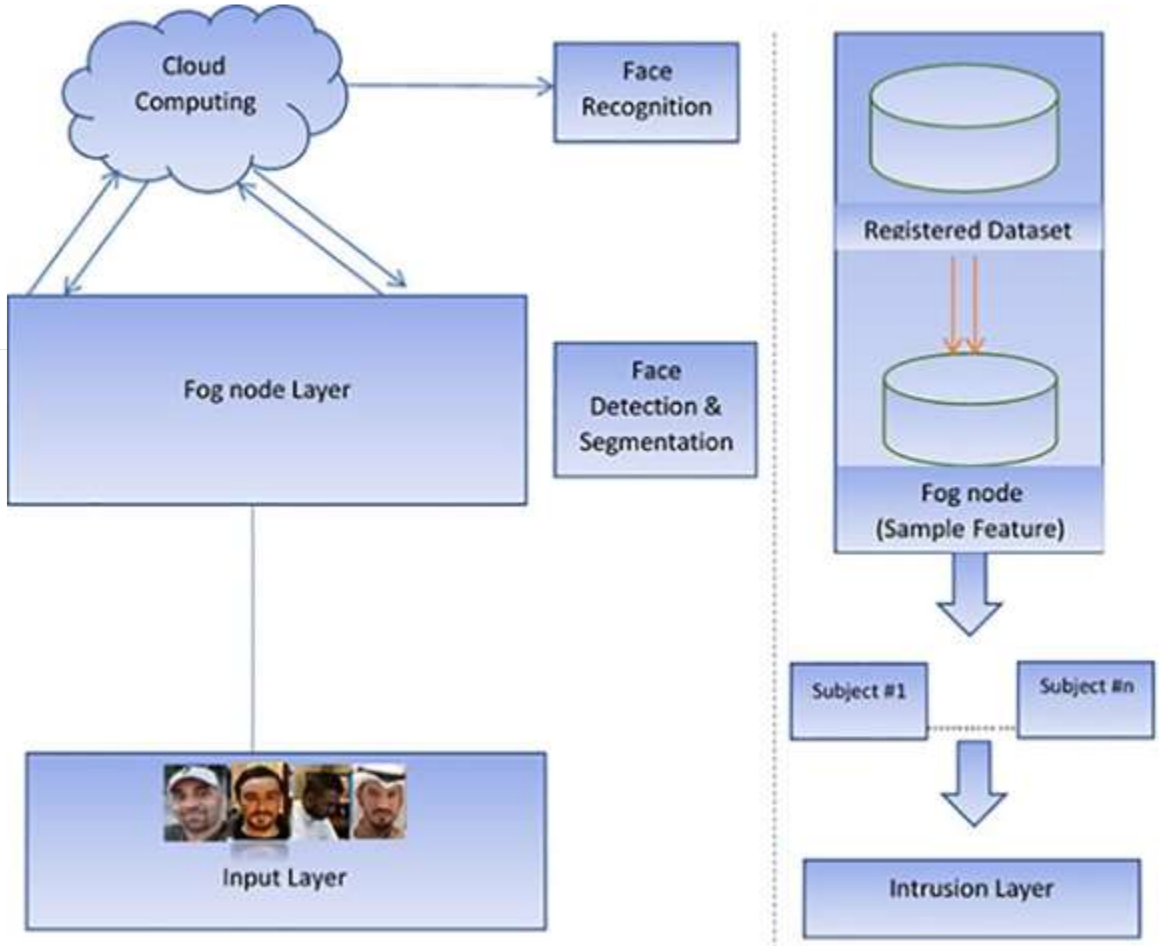

Fig 9. Fog computing network for the face recognition scheme.

Because the cloud server has more storage capacity than the fog nodes/servers, it can store and process a large number of training sets. It can deliver training sets to fog nodes/servers in stages, with the purpose of sending relevant sets to diverse fog nodes/servers.

5. Performance measure

Clearly, independent of the latest technologies and computing environment, researchers have recently focused on improving face recognition systems from accuracy measures. Cloud computing and fog computing are now accessible to improve face recognition performance and reduce time complexity. We will address these concerns in the proposed framework, which has been carefully reviewed. The classifier accuracy is classified as true positive (TP), false negative (FN), false positive (FP), and true negative (TN) by the classifier performance assessor (TN). Precision is the most interesting and sensitive metric for comparing the essential individual classifiers and the proposed system over a large range.

The parameter matrixes can be defined as follows: Accuracy;

recognition rate *TP* + *TN/ P*+ N

error rate; misclassification rate FP+ FN / P+N

specificity; true negative rate *TN* / *N* sensitivity;

TP rate; recall *TP /P*

where,
True Negative (TN): These are the negative tuples that were correctly labeled by the classifier.

True Positive (TP): These are the positive tuples that were correctly labeled by the classifier.

False Positive (FP): These are the negative tuples that were incorrectly labeled as positive.

False Negative (FN): These are the positive tuples that were mislabeled as negative.

RESULTS AND DISCUSSION

Experiments were carried out to assess the proposed system against the evaluation criteria. All trials begin with the color photos being loaded from the data source and then sent to the segmentation process. According to AlexNet, the input image size cannot exceed 227 227 pixels, and the image depth cannot exceed 3. As a result, after segmentation, we performed a check step to ensure that the image size was appropriate. If the image's size exceeds the size limit, a resizing process to 2272273 for width, height, and depth is required.

Parameter	Value
Feature Vector Size	4096
Mini Batch Size	32
Training ratio	80%
Testing ratio	20%

https://doi.org/10.1371/journal.pone.0242269.t002

The scenario presents the experimental results of the proposed FR system as well as comparisons to other methodologies. The suggested algorithm's results were found to surpass most of its competitors, particularly in terms of precision.

1. **Recognition time results**

The four algorithms: decision tree (DT), KNN classifier, SVM, and the proposed DCNN powered by the pre-trained AlexNet classifier are compared in Fig. The link between two comparison parameters, observation/sec and recognition time in seconds per observation. According to the results, the proposed DCNN outperforms other machine learning algorithms in terms of observation/sec and recognition time.

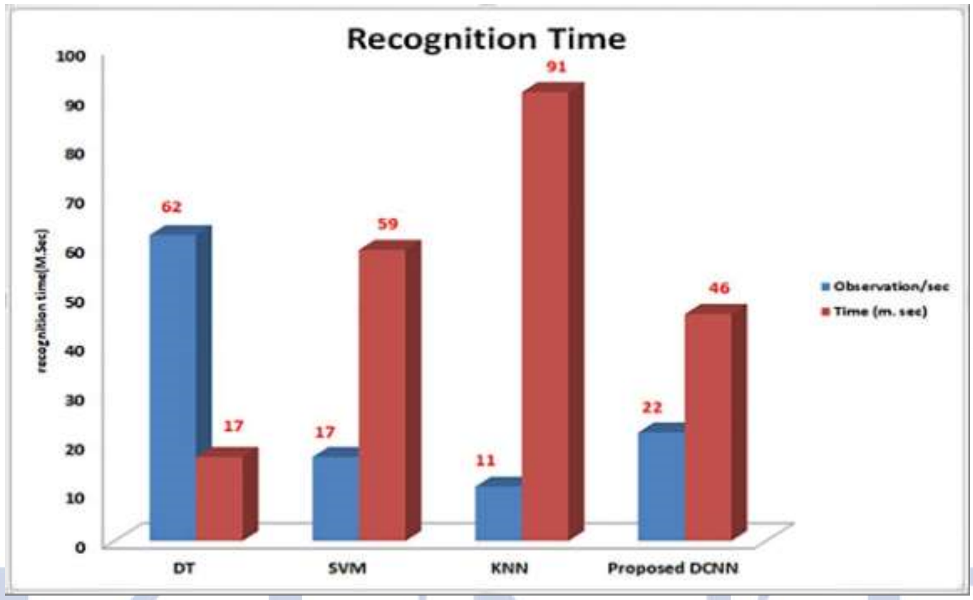

Fig 10. Recognition time of the proposed FR system and individual classifiers.

2. **Precision results**

The precision of the four algorithms is depicted in Figure 1 using the three datasets SDUMLA-HMT, 113, and CASIA. The findings reveal that the proposed DCNN outperforms other machine learning algorithms for the second and third datasets, and that SVM produces the best results for the first dataset.

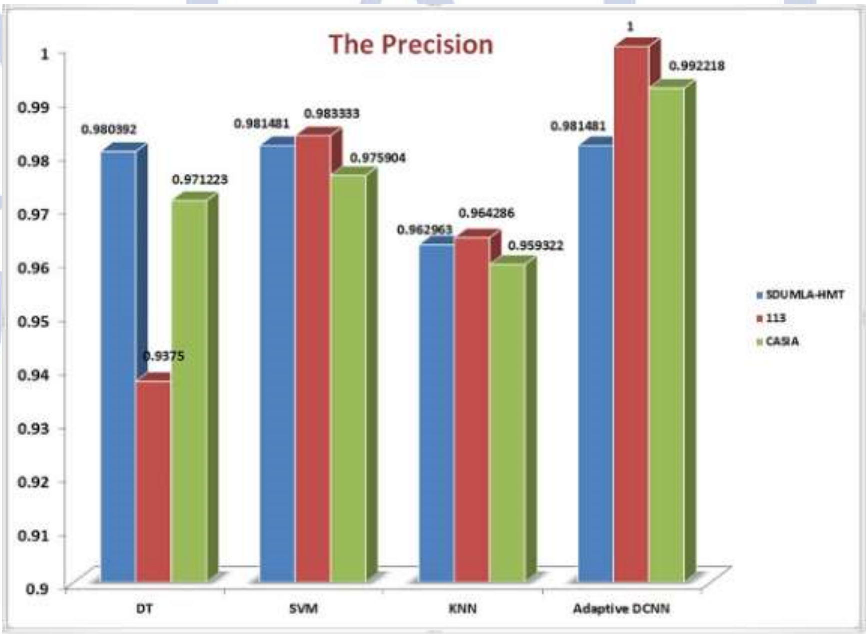

Fig 11. Precision of our proposed system and the three comparison systems.

3. **Recall results**

The recall of the four methods is depicted in Figure using the three datasets SDUMLAHMT, 113, and CASIA. The results reveal that, in terms of Recall parameters, the proposed DCNN outperforms other machine learning algorithms.

Fig 12. Recall of the proposed system and the three comparison systems.

4. **Accuracy Results**

The accuracy of our proposed system of four algorithms is shown in Figure 1 utilizing three datasets: SDUMLA-HMT, 113. The findings reveal that the suggested DCNN outperforms existing machine learning methods in terms of accuracy parameters.

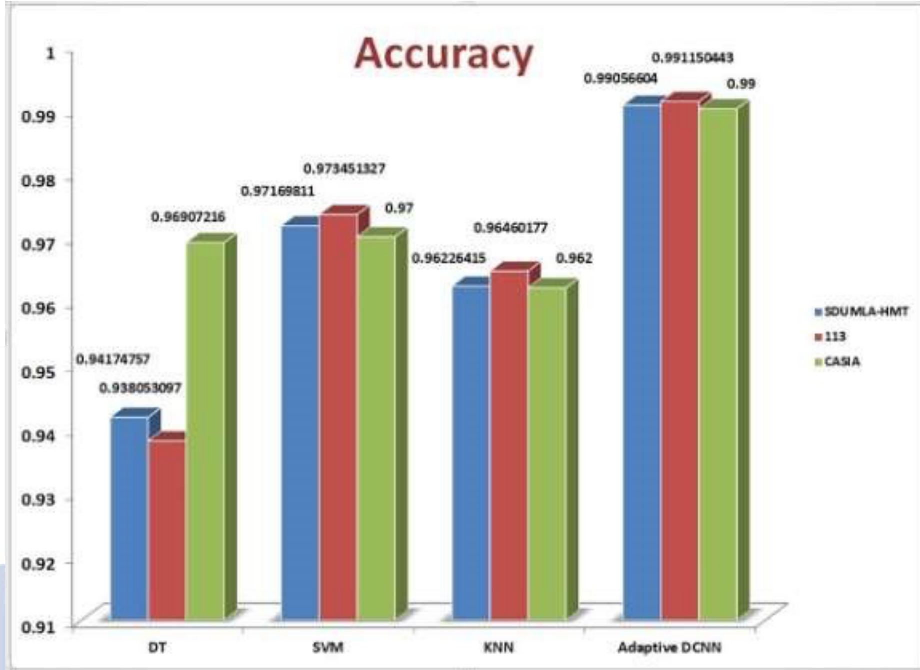

Fig 13. Accuracy of our proposed system and the three comparison systems.

5. Specificity Results

The specificity of our proposed system is compared to other four methods in Fig using three datasets: SDUMLA-HMT, 113, and CASIA.

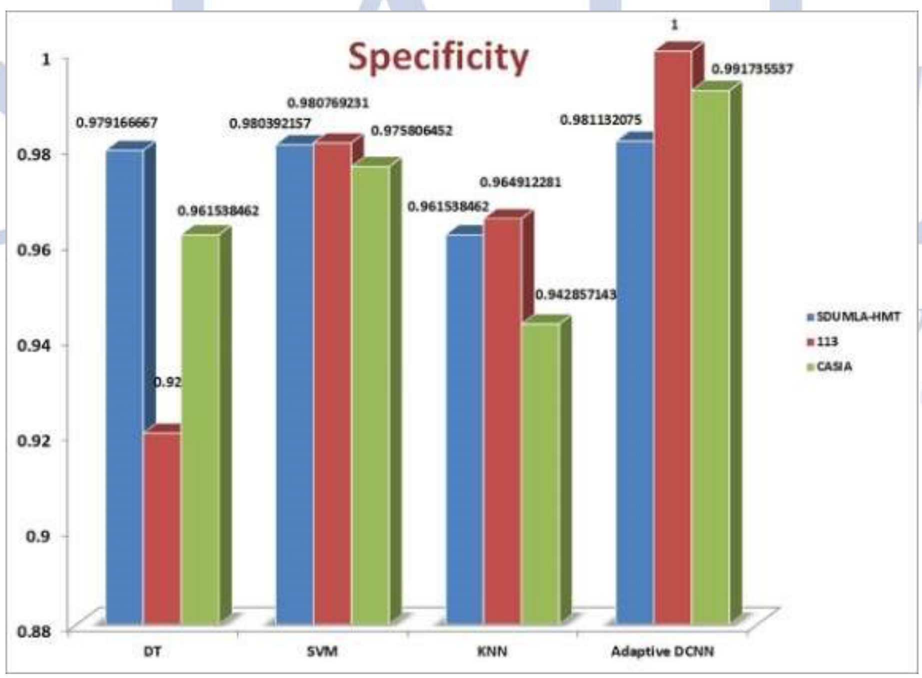

Fig 14. Recall of the proposed system and the three comparison systems.

Table below shows the average results for precision, recall, accuracy, and specificity of the four algorithms using the three datasets SDUMLA-HMT, 113, and CASIA.

Algorithm	Precision	Recall	Accuracy	Specificity
DT	96.30%	94.54%	94.96%	95.36%
SVM	98.02%	96.50%	97.17%	97.90%
KNN	96.22%	96.77%	96.30%	95.64%
Adaptive DCNN	99.12%	99.07%	99.06%	99.10%

https://doi.org/10.1371/journal.pone.0242269.t003

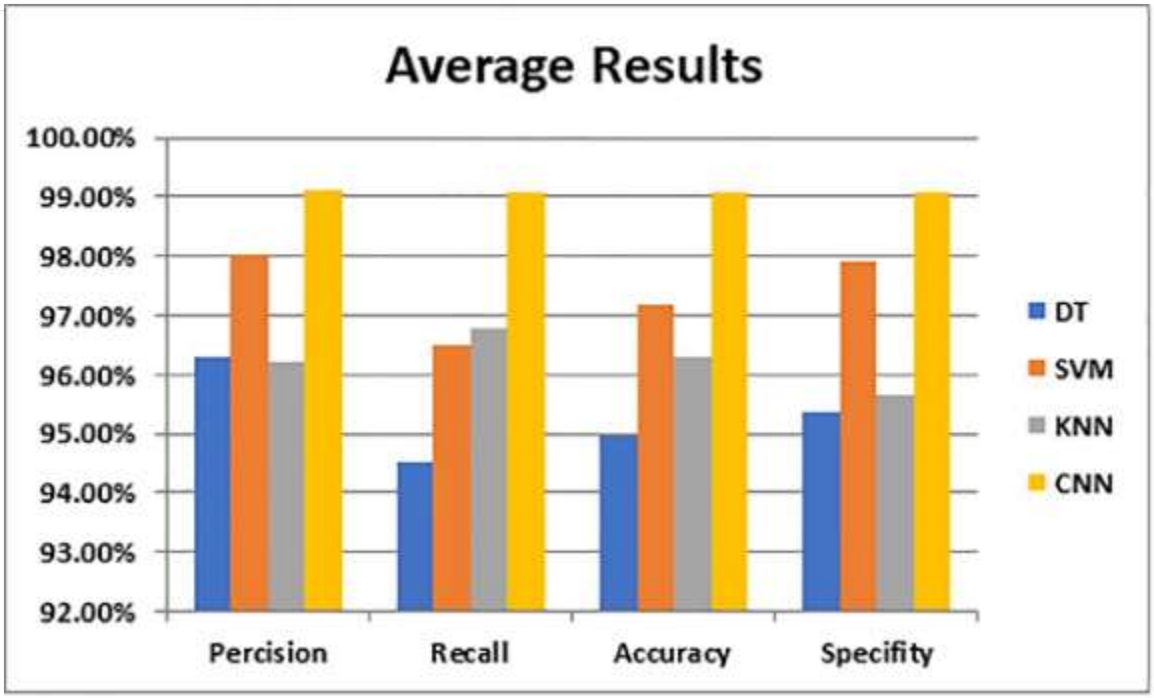

Fig 15. Average results of our proposed system and the three comparison systems.

Classifier	Jonnathann et al.	Our Proposed
KNN	74.83	94.8
SVM	86.41	97.5
DCNN	97.26	99.4

https://doi.org/10.1371/journal.pone.0242269.t004

Fig 16. Comparative accuracy details of KNN, SVM and DCNN using the SDUMLA dataset.

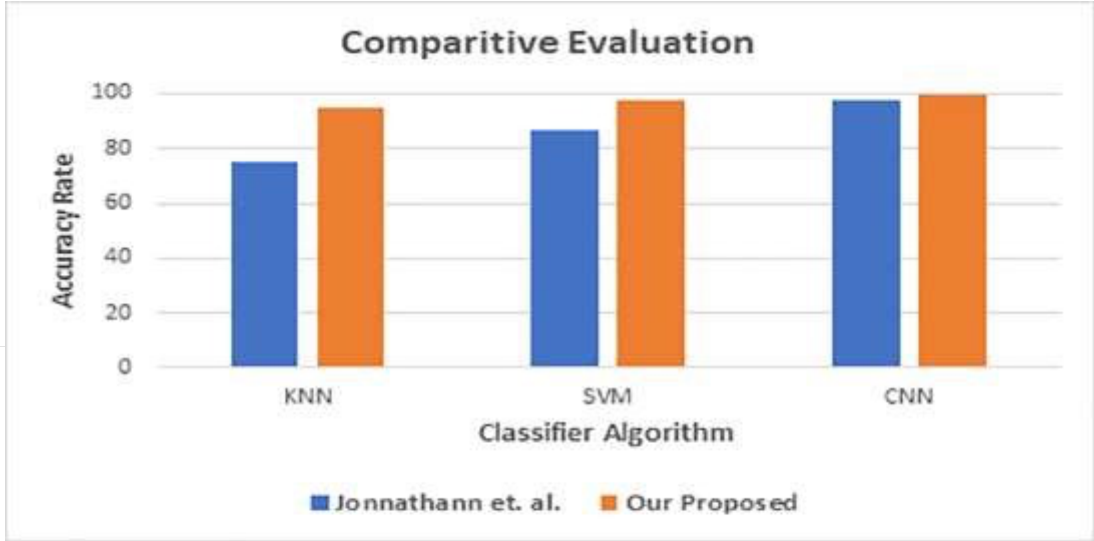

Fig 17. Comparative evaluation of the proposed FR system vs recent literature.

CONCLUSION

FR is a more natural biometric information process than previous systems offered, and it must deal with greater variety than any other method. It is one of the most well-known issues in combinatorial optimization. An effective optimization strategy is required to solve this problem in an acceptable amount of time. In terms of the input image, FR may confront a variety of issues and hurdles, including diverse facial expressions, persons wearing hats or spectacles, and varying brightness levels. This research uses Alex- Net, an adaptive version of the most recent DCNN algorithm. In fog computing, this work developed a deep FR learning approach based on TL.The proposed DCNN method is based on a series of steps for processing face photos and extracting facial features. Preprocessing, face detection, and feature extraction are the three processes that make up this method. The proposed method enhances the solution by altering the parameters in order to find the ultimate best answer. The new approach, as well as other prominent machine learning techniques such as the DT, KNN, and SVM algorithms, were tested on three common benchmark datasets to show that the proposed DCNN is efficient and effective at handling the FR problem.There were varied numbers of photos in these databases, including males and females. The proposed algorithm and other algorithms were tested on various images in the first dataset, and the results showed that the DCNN algorithm was more effective than the other algorithms in terms of achieving the optimal solution (i.e., the best accuracy) with reasonable accuracy, recall, precision, and specificity. In comparison to Jonnathann et al. [18], the proposed DCNN achieved the highest accuracy. The proposed method's accuracy was 99.4 percent, compared to 97.26 percent for Jonnathann et al. [18]. The proposed method outperforms the comparison algorithms in terms of accuracy (99.06 percent), precision (99.12 percent), recall (99.07 percent), and specificity (99.10 percent).

The results of the experiments and performance analysis of various test images (i.e., 30 images) demonstrated that the suggested algorithm could be used to identify an optimal solution in a fair amount of time when compared to other common algorithms. We intend to improve this algorithm in two ways in the future. The first step involves comparing the proposed algorithm to various contemporary metaheuristic algorithms and testing the strategies on the remaining examples from each dataset. The second method involves applying the suggested technique to real-world FR problems in a particular domain.

REFERENCES

1.Prakash, R. Meena, N. Thenmoezhi, and M. Gayathri. "Face Recognition with Convolutional Neural Network and Transfer Learning." In 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT), pp. 861–864. IEEE, 2019

2. Masi I, Tran AT, Hassner T, Leksut JT, Medioni G. Do we really need to collect millions of faces for effective face recognition? In: Leibe B, Matas J, Sebe N, Welling M, editors. European conference on computer vision (ECCV). Cham, Switzerland: Springer; 2016. pp. 579–596.

3. Sivalingam T, Kabilan S, Dhanabal M, Arun R, Chandrabhagavan K. An efficient partial face detection method using AlexNet CNN. SSRG Int J Electron Commun Eng. 2017: 213–216.

4. Deng J, Guo J, Xue N, Zafeiriou S, ArcFace: Additive angular margin loss for deep face recognition. In: 2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR). Long Beach, CA: IEEE; 2019. pp. 4685–4694.

5. Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other kernelbased learn- ing methods. Cambridge, UK: Cambridge University Press; 2000.

6. Power Jonathan D., Plitt Mark, Gotts Stephen J., Kundu Prantik, Voon Valerie, Bandettini Peter A., and Martin Alex. "Ridding fMRI data of motion-related influences: Removal of signals with distinct spatial and physical bases in multiecho data." Proceedings of the National Academy of Sciences 115, no. 9 (2018): E2105–E2114. https://doi.org/10.1073/pnas.1720985115 PMID: 29440410.

7. Power Jonathan D., Plitt Mark, Gotts Stephen J., Kundu Prantik, Voon Valerie, Bandettini Peter A., and Martin Alex. "Ridding fMRI data of motion-related influences: Removal of signals with distinct spatial and physical bases in multiecho data." Proceedings of the National Academy of Sciences 115, no. 9 (2018): E2105–E2114. https://doi.org/10.1073/pnas.1720985115 PMID: 29440410.