FACE RECOGNITION USING HYBRID GABOR FILTER AND STACKED SPARSE AUTO ENCODERS (SSAE) WITH DEEP NEURAL NETWORK MODEL

Abdullah Ghanim Jaber , Assoc. Prof. Dr. Ravie Chandren Muniyandi

Fakulti Teknologi dan Sains Maklumat, Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor Malaysia.

p98659@siswa.ukm.edu.my, ravie@ukm.edu.my

ABSTRACT

Face recognition has been getting greater attention especially since most recognition systems are easily penetrated or copied. However, the accuracy of face recognition is very important in determining the success in identifying an individual. One of the hurdles that hampered the accuracy of face recognition is when the dataset is limited. Especially in traditional machine learning algorithms that are used for face recognition based on images requires sufficient training data to enable high levels of face recognition accuracy. Based on the afore-mentioned problem, a study was carried out to enhance face recognition technique by using hybrid Gabor Filter and deep learning Stacked Sparse Auto Encoders (SSAE). The experimental evaluation was carried out using two datasets which are Olivetti Research Laboratory OLR and Extended Yale-B databases. All face images are greyscale, and the resolution 92 × 112 for OLR database while 192×168 resolution for Extended Yale-B database. The result of the evaluation showed that the accuracy of face recognition has been improved by using proposed method and get best accuracy among all experimented State-of-the-Art methods on both of OLR and Extended Yale-B databases.

Keyword: Hybrid, Gabor Filter, Classifier of Face Recognition**,** Deep Neural Network.

I. INTRODUCTION

Face recognition is a task so familiar to humans that the individual does not even notice the extensive number of times it is performed every day. Although research in automated face recognition has been conducted since the 1960s, it has only recently caught the scientific community's attention. Many face analysis and face modeling techniques have progressed significantly in the last decade (He et al., 2005).

However, the reliability of face recognition schemes still poses a significant challenge to the scientific community.

Face recognition has been an active area of research in the past several decades. Initially, a branch of artificial intelligence to enable robots with visual perception is now part of a more general and more extensive discipline of computer vision. Computer vision applications can process images from a wide range of the electromagnetic spectrum (Alzubaidi & Kalita 2016).

Falsification of identity cards or intrusion of physical and virtual areas by cracking alphanumerical passwords frequently appear in the media. These problems of modern society have triggered a real necessity for reliable, user-friendly, and widely acceptable control mechanisms for the identification and verification of the individual (Heindl 1922).

In computer vision, face recognition applications are confined to the narrow band of visible light where surveillance and biometrics authentication can be performed. Biometrics, which is based on authentication on the intrinsic aspects of a specific human being, appears as a viable alternative to more traditional approaches (such as PIN codes or passwords).

In addition, biometrics is the term used to describe human characteristics metrics such as iris, fingerprint, or hand geometry. These metrics are used to identify and access control of individuals under surveillance (Alzubaidi & Kalita 2016). Among the oldest biometric techniques is fingerprint recognition. This technique was used in China as early as 700 AD for official certification of contracts. Later on, in the middle of the 19th century, it was used to identify persons in Europe (Heindl 1922). A currently developed biometric technique is iris recognition (J. Daugman 2002). This technique is now used instead of passport identification for frequent flyers in some airports in United Kingdom, Canada, and the Netherlands. As well as for access control of employees to restricted areas in Canadian airports and the New York JFK airport. These techniques are inconvenient due to the necessity of interaction with the individual to be identified or authenticated.

The face is becoming the preferred metric over current biometrics simply because it is a natural assertion of identity, and its non-intrusive nature provides more convenience and ease of verification. For example, in a fingerprinting system, the subject is required to interact with the system by placing a finger under a fingerprint reader, and an expert must verify the results. In contrast, using the subject's face as a metric requires no intervention, and a non-expert can verify the results. This is one of the reasons why this technique has caught an increased interest from the scientific community in the recent decade.

Facial recognition holds several advantages over other biometric techniques. It is natural, non-intrusive, and easy to use. In a study considering the compatibility of six biometric techniques (face, finger, hand, voice, eye, signature) with machine-readable travel documents (MRTD) (Heitmeyer 2000), facial features scored the highest percentage of compatibility, see Figure 1.1. In this study, parameters like enrollment, renewal, machine requirements, and public perception were considered. However, facial features should not be considered the most reliable biometric.

However, automated facial recognition can be used in many areas other than security-oriented applications (access-control/verification systems, surveillance systems), such as computer entertainment and customized computer-human interaction. Customized computer-human interaction applications will in the near future be found in products such as cars, aids for disabled people, buildings, etc. The interest in automated facial recognition and the number of applications will most likely increase even more in the future. This could be due to increased penetration of technologies, such as digital cameras and the internet, and due to a greater demand for different security schemes.

II. LITERATURE REVIEW

A. FACE RECOGNITION METHODS

There are four categories of face recognition methods (Choi et al., 2012; Dubey & Tomar, 2016; Yuille et al., 1992).

1 Knowledge-based approaches.
2 Features invariant approaches.
3 Template matching approaches.
4 Appearance-based approaches.

1. Knowledge-Based Approaches

They are known as rule-based methods. They mainly depend on a set of rules in the detection process (Rowley et al., 1998). This method considers our knowledge of face images and translates them into a set of regulations or rules. An example of a rule is a face having two eyes, a nose, a mouth, or a face having the eye area darker than the cheeks. These features are within specific distances in relation to each other.

There is a limitation to this method which is the ability to build an appropriate set of rules because very general rules might result in false positives, and too detailed a set of rules might result in false negatives. This actually could be resolved by using hierarchical knowledge-based methods, which are efficient with simple inputs. This method is generally limited since it cannot locate many faces in a complex image.

(J. Wang & Tan 2000) presented a system which was for images with uncomplicated backgrounds. The system's approach was to find the knowledge surrounding the geometry of face images and the design of the various features. They set rules that defined how to identify the face from the complicated background. These regulations provide an estimate of the geometry which is later used for recognition.

This method does not work effectively under varying positions or orientations. There is also the need for a method that can define human facial framework into clearly defined and meaningful regulations.

2. Feature Invariant Method

Method of Feature invariant identifies faces via used structural face characteristics to extracting features (Kjeldsen & Kender 1996; Leung et al. 1998; Yow & Cipolla 1996). This idea was developed to overcome the limitations of our instinctive knowledge of face images.

One of the earliest algorithms developed was by (Han et al. 1997). Normally, a statistical classifier, models, or edge detector is trained and then utilized to distinguish between non-face and face sections.

The method seeks to find distinct characteristics of an image of a face despite the position or angle. Its focus is to locate systemic characteristics like the fiducial points, the skin texture, and the colour of a face even though there are changes in the head pose, lightning variations, and viewpoint. Facial recognition uses different facial characteristics such as the mouth, cheekbones, the contour of the eye socket, nose, zone near the cheekbones, and the eyes.

Research has shown that the colour of the skin is reviewed to be one of the significant characteristics for face recognition since every person has a unique skin colour and recognition is explicit when the race is a criterion for detection (Sharifara et al. 2014). Feature invariant has a challenge when the characteristics of a face image are altered by noise, occlusion, or illusion. It is also demanding if there is a need for feature extraction.

3. Template Matching Approaches

Template matching methods use parameterized face templates to locate and detect faces. It compares test images with the template of images stored for detection. It can be computed using the relationship between the various characteristics obtained from input image and the predetermined estimate. Each feature could be defined independently, for instance, the hairline can be distinguished using filters or edge detectors.

This technique has limitations if face images are anterior or differences in scale, shape, and pose. Nevertheless, (Dubey & Tomar 2016) proposed the use of deformable templates to solve these limitations.

However, according to (Brunelli & Poggio 1993), template matching can be likened to a test face image which constitutes a two-dimensional array and can be compared with the use of acceptable measures like the Euclidean distance with one framework showing the whole face image.

The human face can also be represented by more than one framework. (Brunelli & Poggio 1993) chose a group of four feature templates: the mouth, eyes, nose, and the whole face, of 188 images of 47 individuals. They compared it to a geometrical matching algorithm, and it came out superior with a recognition rate of 100%. One limitation of the template matching approach is that it is computationally complex.

4. Appearance-Based Method

Appearance-based methods depend on delegated trained images to discover facial models (Osuna et al. 1997; Rowley et al. 1998; Viola & Jones 2001). It captures the representative variability of faces. In effect, appearance-based techniques have exhibited greater performance compared to other techniques (C. Zhang & Zhang 2010).

Generally, these methods depend on skills from machine learning and analysing statistically to search for the significant features of facial images. The method explains face recognition as an issue of two class image classification, that is the face or non-face class.

There are different recognition techniques that are based on appearance-based methods. They include eigen faces, LDA, SVM, PCA, etc. Whatever is considered face is contained in the face class, and all non-face characteristics are also grouped into the non-face. This method requires a large database and a high-quality image for the detection process

B. GENERIC FACE RECOGNITION

As stated in Chapter I, four fundamental elements are included in a generic face recognition system, i.e., face alignment, detection of face, classification, and feature extraction. Representation of feature and recognition classifier plays a crucial role in face recognition and thus attracts much interest.

Extraction of Feature In several applications, for face images, representation of the pixel raw is adopted, i.e., images of the face are represented as a 2-D composed matrix of intensity pixel values. There is much redundancy in the representation of the raw pixel, and it is susceptible to differences in modalities like occlusion, illumination, pose, etc. The raw image face is, therefore, typically pre-processed by different methods for the extraction of features.

A very significant aspect for systems of face recognition is feature extraction. The goal of this step is to acquire representations of faces that are appropriate for classification. Therefore, it is expected that the extracted characteristics emphasize the distinction among classes and minimize irrelevant variations inside each denomination. The optimal extraction of features would lead to a slight classification mission. This concept optimal is hard to achieve because many variants can exist together in images of the face.

This phase, however, describes a face with a collection of features characteristics called a "signature" vector that defines features prominent features of the face image with their geometry distribution, like eyes, nose, and mouth (Kortli et al., 2018; Smach et al. 2007). Each face is distinguished by its shape, size, and structure, allowing it to be recognized. Many methods include extracting the outline of the nose, eyes, or mouth shape and employ the size and distance to define the face (Napoléon & Alfalou 2017). (Q. Wang et al. 2019) HOG, (Turk & Pentland 1991) Eigenface, analysis of linear discriminant (LDA), analysis of independent component (ICA) (Annalakshmi et al. 2019; Seo & Milanfar 2011), scale-invariant feature transform (SIFT) (Vinay et al. 2015), quantization of local phase (LPQ), Gabor filter (Hussain et al. 2012), Haar-wavelets, Fourier transforms (Smach et al. 2007), and pattern of local binary

(LBP) (HajiRassouliha et al. 2013; Napoléon & Alfalou 2017) methods are vastly utilized to extract the characteristics of the face. A perfect exemplification of feature should be sturdy for both local and holistic variations while maintaining identity-preserving data with physical memory minimal.

C. CLASSIFICATION FOR FACE RECOGNITION

The purpose of classification is to learn a model for making forecasts in un-visual data depend on former observations. Several classifiers are taught in a supervised way, i.e., samples or prior observations are presented with tags of the ground-truth. Consequently, the generalization efficiency of the supervised classifier depends heavily on the adequacy of the named samples of training. In the following, several of the model supervised classifiers are presented.

The classifier of (NN) Nearest-Neighbor: The 1-N-N classifier in machine learning is possibly one of the easiest classifiers. The mark of the un-visual sample dot is expected to be the same as the closest adjacent sample in the 1-N-N classification. The classifier of K-N-N is a common various1-N-N, allowing vote prediction for the preponderance dependent on the mark distribution of the closest adjacent points K rather. The notion of "NEAR" is generally described in terms of gauge distance in a feature space. Consequently, the efficiency of classifier of the nearest neighbor depends to a large degree on the usefulness of the exemplification of features.

Forest Classification: Forest classification (Breiman 2001) is a special situation from random forest used to the classification trouble. Includes Forest Random with the notion of ensemble learning by preparation decision-making-trees in the form of a huge number, each serving as a poor classifier. The implementation of bagging results in a certain degree of freedom between trees, leading to a substantial enhancement after the results merge for all trees. During the process of experimenting, a sample of an un-visual is tested versus a chain of easy divides rules for the tree nodes over the track, and eventually dropping into a node of the leaf. The distribution of the category in the identical leaf node is used to measure the posterior likelihood. Eventually, the forecasting is made by calculating the outcomes mean of every tree. The mark of the category with the greatest likelihood is then appointing to the un-visual sample. Forest classification (Bosch et al. 2007; Fette et al. 2007; Gall & Lempitsky 2013) is greatly scalable for the classification trouble of scale wide and has been exceedingly applied for different research areas.

Support Vector Machine (SVM):(Boser et al. 1992) they first present SVM as a classifier of a non-probabilistic classification of the binary. Later then suggested classification of multi-category based on SVM extension (Vapnik 1999). The SVM stimulus is to know the decision-making surface, which distinguishes the samples of training by optimizing the decision-making margin. The kernel hoax allows SVM to be commonly implemented to the classification of non-linear troubles as well. makes the performance of Strong generalization SVM one of the most common classifiers. The maximal margin notion is also stretched to the field Learning of Semi-Supervised.

In several former studies, the selects both of classifiers and features are specified empirically or through experiments comprehensive. The convoluted fusing for these two key crucial elements remains a challenging problem in the implementation of facial recognition in particular scenarios. To fix this problem, researchers are working and trying together to learn about the classifier and feature in a common manner way. One of the most exemplifications approaches that pursue this theory is Deep Neural Network (DNN)

D. ARTIFICIAL NEURAL NETWORK (ANN) AND DEEP NEURAL NETWORK FOR FACE RECOGNITION

The essential variation among machine learning and deep learning is how data is viewed in the framework system. Networks of deep learning depend on layers of ANN (artificial neural network), while algorithms of machine learning nearly always demand data structured. Deep learning in practice is a subgroup of machine learning which accomplishes large both of flexibility as well as power via learning to exemplify the world as an overlapping hierarchy of notions, with every notion described in connection to simpler notions and abstract more exemplification calculated in expressions of ones abstract less (Parkhi et al. 2015; Wu et al. 2018).

The fast exponential growth of Internet-based technologies applications, such as networks of social, engines of search, and websites of sharing video, has performed in a data influx in quantity. The sources large of data transfer community of the human to the epoch of data huge, and huge data fetches both opportunities as well as challenges in machine learning and computer vision for research areas.

Deep learning is simulated by such direction, also renowned as (DNN) the Deep Neural Network, which has arisen as a new science of interest in recent years and has drawn a great deal of concern from both academia and industry. The study (Hinton & Salakhutdinov 2006) has motivated many studies and simulated implementations in several areas. (Farabet et al. 2012; G. B. Huang et al. 2012; Krizhevsky et al. 2012; Nair & Hinton 2010; Sun et al. 2013).

Compared to traditional facial recognition systems, deep learning merges education feature exemplification and classifier learning commonly. The learning process is performed straight with respect to the goal of the trouble; consequently, the learning feature is appropriate to the target.

In addition, DNN embraces a hybrid structure with several extraction layers of features-the lower layer's intermediate exemplification as inputs are forwarded for the toper layer. In many instances, the non-linear activation function is implemented to obtain superior generalization for every layer's outputs.

In comparison for learning metric with shoaly linear transformation framework, the deep hybrid of projections of non-linear offers DNN with a higher degree of abstraction and discrimination ability, that has been shown to be efficient in several modern works (Cui et al. 2013; Davis et al. 2007; Guillaumin et al. 2009).

DNN proposes a generic solution for the classification trouble. It does not involve pre-processing of trouble--specific of information or extraction of feature. The raw data is used as the network's input by researchers in most instances. The network's learning is done via continuously alternating among back propagation and forward propagation frequently until convergence.

The inputs are moved via layer by layer of the network during the forward propagation, as well as, objective or function of a cost is determined to depend on the carryout of the definitive layer in the present iteration. The cost gradients are then measured with respect for intermediate inputs and the weights for every layer and further propagated for the previous layer with respect to the law of the chain. While, for updating parameters in back-spreading measures, Stochastic Gradient Descent (SGD) is used as a popular alternative.

In recent years, have been suggested different architectures for deep learning. (Schroff et al. 2015) introduced a structure, called FaceNet, in their work that easily takes a mapping from images of face to junior space of Euclidean that relates to face similarity percentage. The merit of their strategy is a lot further popular descriptive efficacy.

The purpose of this study is to confession confront from either indeed a solitary picture or a series of appearances accompanied in a recording "video" (Parkhi et al., 2015). Delayed progress has been attributed due to two factors elements:

(i) beginning to the end learning on an errand using the algorithm system of (CNN) convolutional-neural

(ii) the usability of large-scale dataset planning.

The automatic recognition of expression facial (FER) suggested in (Mollahosseini et al. 2016) has stayed intriguing and testing, the problem with respect for PC vision. Despite the efforts made to generate various FER approaches, current methodologies have required to be popularized when related to indistinct images or those captured in the wild.

A new approach for recognition of face was suggested (Kamencay et al. 2017) with three popular methods of picture recognition, such as (LBPH) histograms of local binary pattern , (KNN) histograms of K-nearest neighbor (K-nearest neighborhood), and (PCA) analysis of principal component, by using the coevolutionary-neural-network. In their presentation, they utilized KNN, LBPH, PCA, and CNN for precision. In this article, 98 percent were effective in achieving accuracy by utilizing CNN for superior results.

(Rothe et al. 2018) this study suggested a deep learning scheme to infer age from a solitary face image via not utilizing spots of facial tourism and to display dataset IMDB-WIKI, the largest free dataset of gender and age labelled for face images. On the other hand, the investigation of the simple age prediction or age as seen by various people from the face image is an ongoing undertaking that manages the two assignments with our the VGG-16 (CNNs) coevolutionary neural networks engineering that on ImageNet are pre-prepared for features of the image. They portray the topic of estimation age as a deep issue of characterization trouble address through the softmax foreseeable amelioration of esteem.

III. RESEARCH MODEL AND RESEARCH QUESTIONS

The objective of this study is to develop a robust system for recognizing the face with high accuracy. this research includes the new face recognition pattern for two datasets, which was developed based on hybrid Gabor Filter and Stacked Sparse Auto Encoders (SSAE) Deep Neural Network Model to reduce the time-consuming for face recognition.

IV. METHODS: PARTICIPANTS AND DATA COLLECTION

The collecting of facial databases for purposes benchmarking was an integral section of the fixed growing made in automatic recognition expression and facial expression. In the 1990s, for automated face recognition, new methods have been motivated due to major developments in sensor and computer technology. There are presently many databases utilized for facial recognition that range in expressions, conditions, pose, lighting, occlusions, size, and the number of subject's image.

The two datasets were used in this thesis. The first dataset is the ORL database containing a compilation of faces taken at the Olivetti Research Laboratory between April 1992 and April 1994 in Cambridge, United Kingdom (Cambridge 1994).

There are 40 different distinct subjects, and with each subject are 10 different photographs. On certain subjects, the photographs were taken at various times. There are differences in detail of facial (no glasses/glasses) and appearance of facial (non-smiling/smiling, closed eyes /open eyes). All photographs were taken versus a dark homogeneous backdrop with subjects in a frontal posture, upright, tolerance for any rotation, and tilting up to around 20 degrees. There is some variance in the scale range of up to around 10%. Figure 3.2 presents the sample images of the OLR dataset that. The pictures are greyscale and have a resolution of 92 × 112. In this work, resize the OLR images to half in order to increase the time computation.

However, the second database used in this thesis is Extended Yale-B, consisting of 2,432 frontal-face pictures with dimentions192×168 across through 38 various people (Georghiades et al. 2000). Thus, there are 64 photographs per subject, and they differ in illumination.

The photographs were taken under different expressions of facial and varying conditions of lighting. Face photographs differ widely in lighting across objects, so much so which only a tiny piece of the face is apparent at times. The version of the close-cropped dataset is utilized, where each photograph is cropped to have a face with no hair or backdrop. Also, in this thesis, resize the Extended Yale-B database pictures to half to reduce the time computation.

1. *The Hybrid Proposed Model of Face Recognition*

The hybrid proposed method is to combine two algorithms using Gabor Filter and Stacked Sparse Auto Encoders (SSAE) for face recognition. The first step is to resize the input image to reduce the execution time. Then, initial features of the face region are extracted by implementation the Gabor filter that considers as the input to SSAE deep neural network, such as shown in Figure 3.6. The main contribution of this research is to improve the accuracy of the results after applying this proposed hybrid model is to reduce the time-consuming in face recognition that comes from the effect of several types of deformations and noise.

2. Gabor Filters-Based Feature Extraction Method

Method of feature extraction-based Gabor filters (Haghighat et al. 2015) for face picture is utilized to extract and locate initial features from the face region. The most significant merit of Gabor filters is their invariance to translation, rotation, and scale.

Moreover, they are robust against disturbances of photometric, like illumination variations and noise of images (Kamarainen et al. 2006; Liu & Wechsler 2002; Meshgini et al. 2013; Shen et al. 2007).

3. Stacked Sparse Auto Encoders (Ssae) Deep Neural Network Model

Autoencoder is a feature learning algorithm of an unsupervised that seeks to enhance better feature exemplification of high-dimensional input data via determining the correlation between the data. An auto-encoder is essentially a neural network of multi-layer feed-forward that has been learned to exemplify the input utilizing back-propagation. The autoencoder uses back-propagation to reduce the difference among input and reconstruction as far as possible via learning a decoder and an encoder. The stacked autoencoder essentially is a neural network made up of several basic SAE layers, each layer outcomes wired to the next layer inputs. In this study, we build two layers of SSAE from two simple SAE.

V. RESULTS & DISCUSSION

The two data sets utilized were the Olivetti Research Laboratory (OLR) database and versions cropped of the Extended Yale-B database. In this thesis, each of these databases is divided to testing and training dataset; however, uses 2 images per subjects for testing data and remainder images as a training dataset for both OLR and Extended Yale-B databases. This split is used to obtain a perception of how the proposed face recognition method would perform in a practical application, e.g., identification person.

Firstly, features of the face are extracted using 2D Gabor filters. Then the considered network Stacked Sparse Auto Encoders (SSAE) were trained using the features face for both OLR and Extended Yale-B databases. Finally, stacked Sparse Auto Encoders (SSAE) were implemented using two hidden layers.

The designed Stacked Sparse Auto Encoders (SSAE) Deep Neural Network model was trained on 2356 samples initial feature of face images for Extended Yale-B databases, while were trained 320 samples initial feature of face images for OLR database. The input initial feature of face images was of size 5280 pixels for Extended Yale-B databases, while the input initial feature of face images was 1680 pixels for OLR database. Training parameters for Stacked Sparse Auto Encoders (SSAE) Deep Neural Network model are shown below in Table 1.

Table 1 Training parameters for Stacked Sparse Auto Encoders (SSAE).

Parameters	Stacked Sparse Auto Encoders (SSAE) for OLR	Stacked Sparse Auto Encoders (SSAE) for Extended Yale-B
Number of training samples	320	2356
Size of Hidden layer 1	1200	1200
Size of Hidden layer 2	800	800
	1’st Auto Encoder	
Activation function	Log-Sigmoid	Log-Sigmoid
sparsity parameter	0.15	0.15
sparsity weight	4	4
weight decay parameter	0.004	0.004
Max. iterations	400	400
	2’nd Auto Encoder	
Activation function	Log-Sigmoid	Log-Sigmoid
sparsity parameter	4	4
sparsity weight	0.1	0.1
weight decay parameter	0.002	0.002
Max. iterations	200	200
	Final-Soft-max	
Activation function	Soft-max	Soft-max
Max. iterations	200	200
Learning rate for pre-training	0.000001	0.000001
Learning rate for fine-tuning	0.000001	0.000001
Max. iterations for fine-tuning	100	100

The two trained networks firstly proposed hybrid Gabor filter with Stacked Sparse Auto Encoders (SSAE) Deep Neural Network model and second trained network is Stacked Sparse Auto Encoders (SSAE) Deep Neural Network model for OLR and Extended Yale-B databases were tested with 80 sample face images for OLR database and 78 sample face images for Extended Yale-B databases that were not part of the training data to obtain the performance of the networks on classifying new cases.

In below table 2 shows the computation time of the proposed Gabor Filter and Stacked Sparse Auto Encoders (SSAE) method and only network of Stacked Sparse Auto Encoders (SSAE) for OLR database.

Table 2 Execution time of proposed method for OLR database

Number of images	Name of images	Execution Time of Proposed method	Execution Time of only Stacked Sparse Auto Encoders (SSAE)
1	01_I01	0.2973252	1.6227453
2	02_I02	0.2422379	0.4341096
3	03_I01	0.2412173	0.3973739
4	04_I02	0.2551962	0.2599288
5	05_I01	0.2518286	0.5701508
6	06_I02	0.2445681	0.2350624
7	08_I02	0.2546507	0.2208476
8	10_I01	0.2550419	0.2275963
9	12_I01	0.2448359	0.2136016
10	13_I02	0.245745	0.227358
11	15_I01	0.2505601	0.2269222
12	16_I02	0.2495898	0.2143724
13	17_I01	0.243543	0.2196226
14	18_I02	0.2411174	0.2242768
15	19_I01	0.2409688	0.2173489
16	21_I02	0.2445355	0.2295809
17	23_I01	0.2492679	0.2278315
18	24_I02	0.2642704	0.2073777
19	25_I01	0.2477056	0.2160182
20	29_I01	0.2517592	0.2011851
21	31_I01	0.2459486	0.224065
22	31_I02	0.2419242	0.2084263
23	32_I02	0.2494917	0.2226422
24	33_I01	0.2417259	0.2219115
25	34_I02	0.2469378	0.2150088
26	35_I01	0.2532889	0.2265672
27	36_I01	0.2419414	0.2074002
28	38_I01	0.2472856	0.2138572
29	39_I01	0.2515196	0.2170947
30	40_I01	0.2482124	0.2141638
Average Time		0.249474687	0.29214825

However, table 3 below shows the computation time of the proposed Gabor Filter and Stacked Sparse Auto Encoders (SSAE) method and the network of Stacked Sparse Auto Encoders (SSAE) Extended Yale-B database.

Table 3 Execution time of proposed method for Extended Yale-B database.

Number of images	Name of images	Execution Time of Proposed method	Execution Time of only Stacked Sparse Auto Encoders (SSAE)
1	01_IB_01	1.08259	0.5950821
2	02_IB_01	0.623138	0.5850558
3	03_IB_02	0.603067	0.5713341
4	04_IB_02	0.555351	0.5623317
5	05_IB_02	0.72517	0.5687663
6	06_IB_01	0.521113	0.5678378
7	07_IB_01	0.507061	0.5625726
8	08_IB_02	0.490572	0.5635912
9	10_IB_01	0.510978	0.5907825
10	10_IB_02	0.502858	0.5717226
11	14_IB_01	0.499462	0.5665855
12	15_IB_02	0.491067	0.5693121
13	16_IB_01	0.555758	0.5772575
14	18_IB_01	0.491597	0.5866595
15	19_IB_01	0.493095	0.574229
16	20_IB_02	0.493669	0.5866195
17	22_IB_02	0.511596	0.5715895
18	23_IB_01	0.548506	0.5794412
19	24_IB_02	0.516106	0.5671572
20	25_IB_02	0.498472	0.5666622
21	27_IB_02	0.501304	0.5764131
22	28_IB_02	0.490185	0.5709971
23	30_IB_02	0.497125	0.5775362
24	31_IB_02	0.489262	0.6135733
25	33_IB_02	0.511908	0.5658698
26	35_IB_01	0.673285	0.569165
27	35_IB_02	0.615203	0.5742719
28	36_IB_02	0.492643	0.559551
29	37_IB_01	0.500021	0.5697548
30	38_IB_02	0.4935	0.5753255
Average Time		0.549522	0.574568253

Moreover, accuracy comparison of face recognition between the proposed method (hybrid Gabor Filter and Stacked Sparse Auto Encoders (SSAE) method) and only Stacked Sparse Auto Encoders for both OLR and Extended Yale-B databases are showing in tables (4) and (5), respectively.

Table 4 Accuracy comparison for OLR database.

Network	Proposed method	Only Stacked Sparse Auto Encoders
Number of test face samples	80	80
Mean Squared Error (MSE)	0.0000	0.0009
Correctly classified face samples	80	79
Recognition rate	100%	98.75

Table 5 Accuracy comparison for Extended Yale-B database.

Network	Proposed method	Only Stacked Sparse Auto Encoders
Number of test face samples	76	76
Mean Squared Error (MSE)	0.0000	0.0055
Correctly classified face samples	76	71
Recognition rate	100%	93.4211

Table (4) and (5) illustrates the compared the accuracy rates of face recognition between two data networks proposed method (Gabor Filter and Stacked Sparse Auto Encoders (SSAE)) and (Only Stacked Sparse Auto Encoders) for OLR database and Extended Yale-B database.

From the table (4), The Mean Squared Error (MSE) by using the OLR database for the proposed method (Gabor Filter and Stacked Sparse Auto Encoders) is lower than Only Stacked Sparse Auto Encoders. The

Mean Squared Error (MSE) for the proposed method is 0.0000, while the Mean Squared Error (MSE) for Only Stacked Sparse Auto Encoders is 0.0009.

Also, from the table (4) above, it will be seen that by using OLR database, the proposed method achieved higher recognition rates on the test data than Only Stacked Sparse Auto Encoders. The accuracy rate for the proposed method is 100%, while the accuracy rate for Only Stacked Sparse Auto Encoders is 98.75%.

In addition, from the table (5), The Mean Squared Error (MSE) by using Extended Yale-B database for proposed method (Gabor Filter and Stacked Sparse Auto Encoders) is lower than Only Stacked Sparse Auto Encoders. The Mean Squared Error (MSE) for the proposed method is 0.0000, while the Mean Squared Error (MSE) for Only Stacked Sparse Auto Encoders is 0.0055.

Also, from the table (5) above, it will be seen that by using Extended Yale-B database, the proposed method achieved higher recognition rates on the test data than Only Stacked Sparse Auto Encoders. The accuracy rate for the proposed method is 100%, while the accuracy rate for Only Stacked Sparse Auto Encoders is 93.4211%.

Table (4) and (5) is the result accuracies of the proposed method (hybrid Gabor Filter and Stacked Sparse Auto Encoders (SSAE)) against state-of-the-art approaches on both the OLR database and Extended Yale-B databases.

In below table 6 shows the computation accuracy of the proposed Gabor Filter and Stacked Sparse Auto Encoders (SSAE) method and state-of-the-art approaches for OLR database.

Table 6 Accuracy comparison with state-of-the-art methods for OLR database.

Method	Accuracy
(Kamencay et al. 2017)	98.3%
(X. Tan et al. 2006)	74.6%
(Rejeesh 2019)	96%
(Zafaruddin & Fadewar 2019)	93%
Proposed Method	100%

However, table 7 below shows the computation accuracy of the proposed Gabor Filter and Stacked Sparse Auto Encoders (SSAE) method and the state-of-the-art approaches for Extended Yale-B database.

Table 7 Accuracy comparison with state-of-the-art methods for Extended Yale-B database.

Method	Accuracy
(Fernandes & Bala 2013)	97.50%
(Cai et al. 2006)	95.17%
(Kumar et al. 2012)	99.6%
Proposed Method	100%

From the table (6), it will be shown that by using OLR database, the proposed method achieved a higher result than other methods, where the accuracy rate for the proposed method is 100%, while the accuracy rate by the following: (Kamencay et al. 2017) is 98.3%, (Tan et al. 2006) is 74.6%, (Rejeesh, 2019) is 96%, and (Zafaruddin & Fadewar, 2019) is 93%.

However, from the table (7), it will be shown that by using Extended Yale-B database, the proposed method achieved a higher result than other methods, where the accuracy rate for the proposed method is 100%, while the accuracy rate by the following: (Fernandes & Bala) is 97.50%, (Cai et al.) is 95.17%, and (Kumar et al.) is 99.6%.

Finally, from above can be concluded that the proposed method achieved a higher result than other methods for both OLR and Extended Yale-B databases

VI. CONCLUSION

In this work, a newly proposed method for hybrid face pattern recognition technique has been on hybrid Gabor Filter and Stacked Sparse Auto Encoders (SSAE) method this our contributions for this research.

The new recognition system dealt with feature extraction by comparing our proposed with only Stacked Sparse Auto Encoders and comparing our proposed method with state-of-the-art methods of face recognition. Also, the two datasets used in this thesis are the first one is Olivetti Research Laboratory (OLR) database and the second dataset is versions cropped of Extended Yale-B database.

A face recognition based on hybrid Gabor Filter and Stacked Sparse Auto Encoders (SSAE) method has been presented in this study. The Gabor Filter feature extraction method was adopted as extracted initial face features. Then these initial features are input to Stacked Sparse Auto Encoders (SSAE) in order to reduce the time-consuming in face recognition that comes from the effect of several types of deformations and noise. The results verify the novelty of this proposed, where it is the best to describe simple and complex faces simultaneously independent of the effect of several variations such as scaling, noise, and rotation. Finally, suggestions have put in some points which are recommended and proposed by the researcher to carry out any work in this area in the future.

ACKNOWLEDGEMENT

The authors would like to thanks, Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia by giving the authors an opportunity to conduct this research.

REFERENCE

Alzubaidi, A. & Kalita, J. 2016. Authentication of Smartphone Users Using Behavioral Biometrics. *arXiv* 18(3): 1998–2026.

Breiman, L. 2001. Random forests. *Machine learning* 45(1): 5–32.

Brunelli, R. & Poggio, T. 1993. Face recognition: Features versus templates. *IEEE transactions on pattern analysis and machine intelligence* 15(10): 1042–1052

Bosch, A., Zisserman, A. & Munoz, X. 2007. Image classification using random forests and ferns. *2007 IEEE 11th international conference on computer vision*, hlm. 1–8. Ieee.

Cai, D., He, X., Han, J. & Zhang, H.-J. 2006. Orthogonal laplacianfaces for face recognition. *IEEE transactions on image processing* 15(11): 3608–3614.

Cambridge, A. L. 1994. The Database of Faces (ORL Face Database).

Choi, J. Y., Plataniotis, K. N. & Ro, Y. M. 2012. Face feature weighted fusion based on fuzzy membership degree for video face recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)* 42(4): 1270–1282.

Daugman, J. 2002. How Iris Recognition Works.

Davis, J. V, Kulis, B., Jain, P., Sra, S. & Dhillon, I. S. 2007. Information-theoretic metric learning. *Proceedings of the 24th international conference on Machine learning*, hlm. 209–216.

Dubey, D. & Tomar, G. S. 2016. Deep Perusal of Human Face Recognition Algorithms from Facial Snapshots. *International Journal of Signal Processing, Image Processing and Pattern Recognition* 9(9): 103–112.

Farabet, C., Couprie, C., Najman, L. & LeCun, Y. 2012. Learning hierarchical features for scene labeling. *IEEE transactions on pattern analysis and machine intelligence* 35(8): 1915–1929.

Fernandes, S. & Bala, J. 2013. Performance Analysis of PCA-based and LDA-based Algorithms for Face Recognition. *International Journal of Signal Processing Systems* 1(1): 1–6.

Fette, I., Sadeh, N. & Tomasic, A. 2007. Learning to detect phishing emails. *Proceedings of the 16th international conference on World Wide Web*, hlm. 649–656.

Gall, J. & Lempitsky, V. 2013. Class-specific hough forests for object detection. *Decision forests for computer vision and medical image analysis*, hlm. 143–157. Springer.

Guillaumin, M., Verbeek, J. & Schmid, C. 2009. Is that you? Metric learning approaches for face identification. *2009 IEEE 12th international conference on computer vision*, hlm. 498–505. IEEE.

Haghighat, M., Zonouz, S. & Abdel-Mottaleb, M. 2015. CloudID: Trustworthy cloud-based and cross-enterprise biometric identification. *Expert Systems with Applications* 42(21): 7905–7916.

HajiRassouliha, A., Gamage, T. P. B., Parker, M. D., Nash, M. P., Taberner, A. J. & Nielsen, P. M. F. 2013. FPGA implementation of 2D cross-correlation for real-time 3D tracking of deformable surfaces. *2013 28th International Conference on Image and Vision Computing New Zealand (IVCNZ 2013)*, hlm. 352–357. IEEE.

Hinton, G. E. & Salakhutdinov, R. R. 2006. Reducing the dimensionality of data with neural networks. *science* 313(5786): 504–507.

Heitmeyer, R. 2000. Biometric identification promises fast and secure processing of airline passengers. *ICAO journal* 55(9): 10–11.

He, X., Yan, S., Hu, Y., Niyogi, P. & Zhang, H. J. 2005. Face recognition using Laplacianfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 27(3): 328–340. doi:10.1109/TPAMI.2005.55

Hussain, S. U., Napoléon, T. & Jurie, F. 2012. Face recognition using local quantized patterns. *British machive vision conference*, hlm. 11-pages.

Kamencay, P., Benčo, M., Miždoš, T. & Radil, R. 2017. A new method for face recognition using convolutional neural network.

Kjeldsen, R. & Kender, J. 1996. Finding skin in color images. *Proceedings of the Second International Conference on Automatic Face and Gesture Recognition*, hlm. 312–317. IEEE.

Kortli, Y., Jridi, M., Al Falou, A. & Atri, M. 2018. A novel face detection approach using local binary pattern histogram and support vector machine. *2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET)*, hlm. 28–33. IEEE.

Kumar, R., Banerjee, A., Vemuri, B. C. & Pfister, H. 2012. Trainable convolution filters and their application to face recognition. *IEEE transactions on pattern analysis and machine intelligence* 34(7): 1423–1436.

Leung, T. K., Burl, M. C. & Perona, P. 1998. Probabilistic affine invariants for recognition. *Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No.*

98CB36231), hlm. 678–684. IEEE.

Liu, C. & Wechsler, H. 2002. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Transactions on Image processing* 11(4): 467–476.

Meshgini, S., Aghagolzadeh, A. & Seyedarabi, H. 2013. Face recognition using Gabor-based direct linear discriminant analysis and support vector machine. *Computers & Electrical Engineering* 39(3): 727–745

Mollahosseini, A., Chan, D. & Mahoor, M. H. 2016. Going deeper in facial expression recognition using deep neural networks. *2016 IEEE Winter conference on applications of computer vision (WACV)*, hlm. 1–10. IEEE.

Napoléon, T. & Alfalou, A. 2017. Pose invariant face recognition: 3D model from single photo. *Optics and Lasers in Engineering* 89: 150–161.

Osuna, E., Freund, R. & Girosit, F. 1997. Training support vector machines: an application to face detection. *Proceedings of IEEE computer society conference on computer vision and pattern recognition*, hlm. 130–136. IEEE.

Parkhi, O. M., Vedaldi, A. & Zisserman, A. 2015. Deep face recognition.

Pasandi, M. E. M. 2014. Face, Age and Gender Recognition Using Local Descriptors. University of Ottawa.

Rejeesh, M. R. 2019. Interest point based face recognition using adaptive neuro fuzzy inference system. *Multimedia Tools and Applications* 78(16): 22691–22710.

Rowley, H. A., Baluja, S. & Kanade, T. 1998. Neural network-based face detection. *IEEE Transactions on pattern analysis and machine intelligence* 20(1): 23–38.

Schroff, F., Kalenichenko, D. & Philbin, J. 2015. Facenet: A unified embedding for face recognition and clustering. *Proceedings of the IEEE conference on computer vision and pattern recognition*, hlm. 815–823.

Sharifara, A., Rahim, M. S. M. & Anisi, Y. 2014. A general review of human face detection including a study of neural networks and Haar feature-based cascade classifier in face detection. *2014 International Symposium on Biometrics and Security Technologies (ISBAST)*, hlm. 73–78. IEEE.

Shen, L., Bai, L. & Fairhurst, M. 2007. Gabor wavelets and general discriminant analysis for face identification and verification. *Image and Vision Computing* 25(5): 553–563.

Smach, F., Miteran, J., Atri, M., Dubois, J., Abid, M. & Gauthier, J.-P. 2007. An FPGA-based accelerator for Fourier Descriptors computing for color object recognition using SVM. *Journal of Real-Time Image Processing* 2(4): 249–258.

Tan, X., Chen, S., Li, J. & Zhou, Z.-H. 2006. Learning non-metric partial similarity based on maximal margin criterion. *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, hlm. Vol. 1, 145–168. IEEE.

Vapnik, V. N. 1999. An overview of statistical learning theory. *IEEE transactions on neural networks* 10(5): 988–999.

Vinay, A., Hebbar, D., Shekhar, V. S., Murthy, K. N. B. & Natarajan, S. 2015. Two novel detector-descriptor based approaches for face recognition using sift and surf. *Procedia Computer Science* 70: 185–197.

Wang, J. & Tan, T. 2000. A new face detection method based on shape information. *Pattern Recognition Letters* 21(6–7): 463–471.

Wu, X., He, R., Sun, Z. & Tan, T. 2018. A light cnn for deep face representation with noisy labels. *IEEE Transactions on Information Forensics and Security* 13(11): 2884–2896.

Yow, K. C. & Cipolla, R. 1996. A probabilistic framework for perceptual grouping of features for human face detection. *Proceedings of the second international conference on automatic face and gesture recognition*, hlm. 16–21. IEEE.

Yuille, A. L., Hallinan, P. W. & Cohen, D. S. 1992. Feature extraction from faces using deformable templates. *International journal of computer vision* 8(2): 99–111.

Zafaruddin, G. M. & Fadewar, H. S. 2019. Face recognition using eigenfaces. *Computing, communication and signal processing*, hlm. 855–864. Springer.

Zhang, C. & Zhang, Z. 2010. A survey of recent advances in face detection.