A Stochastic Model for Face Detection

A.A. Ojugo., D. Allenotor, and D.A. Oyemade

Department of Math/Computer Science Federal University of Petroleum Resources Effurun, Delta State, Nigeria. ojugo.arnold@fupre.edu.ng, allenotor.david@fupre.edu.ng, oyemade.david@fupre.edu.ng

ABSTRACT

Abstract—Biometric systems offer several merits over other means of authentication – making them more popular. Designed to withstand many attacks, they are still circumvented. Our study is an evolutionary model with deployment ease as employed in security-critical apps. This will help it withstand frequent attacks that aim to circumvent it. The model's strengths and tradeoffs on its resolution are noted, with emphasis on face recognition.

Keywords- Stochastic, elitists, hybrid, biometric, optimization, conflict, fitness function.

Aims Research Journal Reference Format:

A.A. Ojugo., D. Allenotor, and D.A. Oyemade (2016): A Stochastic Model for Face Detection. Advances in Multidisciplinary Research Journal. Vol. 2. No. 2, Issue 1 Pp 101-114

1. INTRODUCTION

Personal identification is the process that associates to a user, an identity that grants physical access of a system to such user. With various complexities, identification is grouped into: (a) verification/authentication (confirms/denies a user's identity), and (b) recognition (helps establish the identity of an authorized user from a set of known or unknown identity templates stored in a database). Biometric system use personal physiological (fingerprint, facial etc), or behavioral (voice, signature etc) traits or characteristics [6, 11]. Thus, they are designed to allow its sensor scans/acquires images of trait(s) to be verified through its feature extractor. The extracted images are matched against user identity or profile template (i.e. the stored database of user identity in the embedded system). If the user's identity exists as authorized, access to the system is granted; else, it is denied. Where such identity is for a newly authorized user whose profile is not stored in the template, a new identity profile is created and consequently stored in the template. A merit of biometric is that extracted traits cannot be misplaced or forgotten - as it represents a tangible component of the user. Use of physiological or behavioral feats provides the following desirable properties [1-3]: (a) universality (all users have the trait), (b) permanence (they do not change or vary in time), (c) collectability (are quantitatively measured), (d) acceptability (high acceptance rate), (e) performance (are resources with high identification accuracy and are not easily affect by environmental factors, (f) uniqueness (two persons cannot have the same exact feat), and (g) circumvention (how easily is the system fooled).

For effective authentication and recognition (both for known and unknown identity), biometrics and biometric systems use statistically-based heuristics and techniques (like component analysis and soft computing) – all of which, measures the underlying probability in its corresponding percentage match of the extracted trait(s). The use of soft-computing technique in biometrics, aims to reduce authentication task with greater certainty and at the shortest or fastest time possible [1-3].

2. TYPES OF BIOMETRIC SYSTEMS

[3-4, 12-14] note biometric types as to include:

- 1. Voice is not a sufficiently unique identification means, as its quality can be degraded by communication channel or device. Its signals are extracted, normalized and are decomposed via frequency or time domain channels, into several band pass. It uses Fourier Transforms Logarithm, vector quantization, dynamic time warp and Markov model as its match-strategy. Its input is adversely affected by user's health, stress and emotions, and extracting such invariant feats in such cases can be difficult. The system is circumvented by mimic or reproduction of a recorded voice. To combat this, system must prompt user to utter a different phrase each time as authentication procedure.
- 2. Thermograms are images obtained via infrared radiations with the gray levels at each pixel characterizing radiation magnitude and heat patterns that are specific to each individual. The absolute values of the heat radiations are dependent on many extraneous factors and are not completely invariant to the identity of an individual; the raw measurements of heat radiation need to be normalized with respect to heat radiating from a landmark feature of body. It is used for covert identification for signature and distinguishes twins. Its demerits: (a) heat source of nearby space in uncontrolled environment affects acquisition, and (b) expensive nature of infrared sensors. It is applied in facial (non-contact and non-invasive) sensing technique.
- 3. Fingerprints are graphic-ridges whose formations depend on the embryonic development as they are unique to each person. As the most widely used in forensic, its image is captured in one-of-two ways: (i) scan an inked impression of a finger, and (ii) scan via a fingerprint scanner. Its representations are based either on an entire image, finger ridges, or salient feats of the ridges (minutiae). It employs four basic methods of identification: (a) invariant feats of gray scale profiles, (b) global ridge patterns or fingerprint classes; (c) ridge patterns, and (d) fingerprint minutiae feats from ridge endings and bifurcations.
- 4. Face has proven to be non-intrusive. It has two methods: (a) Transform image is stored in orthonormal vectors as Eigen-face, each is a covariance analysis of the image population. Two faces are identical if they are sufficiently close in their Eigen-face feat space, and (b) Attribute facial attributes like nose, eyes, etc. are extracted from the face image and invariance of geometric properties among in face landmark features is used for recognizing feats. A major issue is facial disguise and it is a huge task to develop recognition method that can tolerate the effects of aging, facial expressions, and slight variations in pose with respect to camera 2- or 3-D rotations).
- 5. Iris: Visual texture determined by chaotic morphogenetic process, known to be unique for each person and each eye. Its image is captured via non-contact camera of high resolution to register image from predetermined distance in a camera's focal plane. Identification error is extremely small, and constant length position invariant code permits an extremely fast method of iris recognition.

System	Universal	Unique	Perform	Accept	Circumvent
Face	high	medium	low	high	low
Finger	medium	high	high	medium	high
Iris	high	High	High	Low	High
Retinal	high	High	High	Low	High
Voice	medium	Low	Low	High	Low
Thermos	high	High	Medium	High	High
Gait	medium	Low	Low	High	High
DNA	high	High	High	Low	Low

Table 1: Comparison of all Biometric systems

Pattern Matching Algorithms

Biometric matching, recognition, classification and pattern grouping is an important task employed in many fields. A pattern is an image, handwritten word or speech signal, which can be classified as: (a) Supervised (conditions of the system is defined by the designed so that the system learns via use of discriminant analysis where inputs are identified, as members of a predefined class, and (b) Unsupervised (system learns from similarity of patterns via clustering and grouping of the patterns assigned to known and/or unknown class [23]. Its application are in data mining, multimedia database retrieval, financial forecasting and in biometrics. Biometric design has 4-phases: (a) data acquisition / preprocess, (b) representation, (c) train and test, and (d) decision making.



Fig. 1: A functional pattern Recognition Model

The recognition model involves data sensing, preprocessing of data, training the system to identify user profiles, data scheme representation using the template, and decision model. A well-defined, sufficiently constrained recognition task (with small or large interclass variations) often leads to a compact pattern representation with simple decision-making strategy. Learning from a training set of examples is important.

[6-7] notes four best known approaches as:

- Matching establishes the similarity between two entities of same type. A template is prototype of the pattern to be recognized. An inputted pattern is matched against stored pattern, noting allowable pose (rotation and translation) and scale changes. Similarity correlation is optimized via training data. Also, a template itself is learned from it. Matching is computationally demanding and its few demerits makes it not the most effective.
- Statistical Groupings: Decision and estimation theories based on vectors defines a family of classconditional probability density functions with Pr (x|c_i) (probability of feat vector x given class c_i) are arranged in optional order not considering relations between such feats.
- 3. Syntactic/Structure Match is a complex method component patterns and their relationship. Its strategy for learning defines structures. It is difficult for the model to compensate for noise.

4. Neural Networks attempts to model biological neural systems to solve practical pattern recognition tasks as it compare models with various learning process, predicts the neural processes in such system, based on either direct data from statistical and geometric approach; Or on a higher level symbolic data from structural approach. Template matching is compared with learning, by storing all facts as data without understanding them.

Recognition Algorithms

Various algorithms have been developed, to acquire images via device of high resolution (camera) with feats like weight matrix extracted from them, and stored in template database. Obtained weight matrix is identified using same mean image as obtained in the training phase, and compared against those stored in the template to get a match (check if such image exists in database). [5] The algorithms are:

- a. Principal Component Analysis determines vectors of lower dimension that best approximates to a given data by taking S-dimensional vector representation of a face in a training dataset as input, and determines a T-dimensional subspace whose basis vector is maximum corresponding to the original image (with the dimension of this new subspace usually lower than the original as t<<s). If the original image element(s) is seen as random variables, the principle components are along its Eigen vectors and will corresponds to a larger Eigen values of the correlation matrix and error minimization done in a least square.</p>
- b. Independent Component Analysis (ICA) extracts the statistics of a random variable with its secondorder and higher-order dependencies input minimized, and its basis along which the data is statistically, independent found.
- c. Elastic Bunch Graph: Faces share similar topological structure and are represented as graphs, with nodes positioned at fiducially points (eyes, nose, etc) and edges labeled with 2D vectors. Each node is a set of 40 complex Gabor wavelength coefficients at different scales and orientations (phase, amplitude). Recognition is based on labeled graphs (nodes connected by edges; each node is a jet; while edges are distances).
- d. Trace Transform: is a generalization of Randon transform for image processing (for recognizing objects under transformations via translation, rotation and scaling) to yield a trace. It computes a functional along tracing lines of an image. Different Trace transforms can be produced from an image using different trace functional.
- e. Evolutionary Approach is an Eigen space-based approach that searches for the best set of projection axes in order to maximize a fitness function, measuring at the same time the classification accuracy and generalization ability of the system. Because the dimension of the solution space is too big, it is involved in using a special kind of genetic algorithm called Evolutionary Pursuit.

3. MATERIALS AND METHODS

Motivation / Statement of Problem

Computational models do not address vital issues such as: (a) face pose problem with the same set of traits (eyes, nose, mouth) and same position, (b) many face detection systems do not perform well in realtime scenarios with factors such as scaling, rotation, pose and lighting (which have proved to be limiting factors in their performance), and (c) Preprocessing is required in order to obtain satisfactory results.

Rationale for Study Design and Framework Adopted

- 1. Face (or other trait) forgery has made it expedient to train systems using various orientations so that an arbitrarily search the database can output a user's identity based on extracted/inputted image. The system has four modules namely: detection, alignment, feature extraction and matching.
- 2. Detection aims to segment face area from its background, which will yield a coarse estimate of the localization scale of each detected face. Alignment achieves a more accurate localization normalizing face from the extracted feats, located based on location points, normalized with respect to geometrical feats, transformation and morphing such as size and pose, and improved further with respect to photometrical properties such as illumination and gray scale. After geometric and photometric evaluation, extraction will yield effective data to distinguish between two or more faces of



different users, with respect to the variations. Matching helps match the extracted feats vectors of the input face to those stored in the database template to output a face's identity if match is found with sufficient confidence; Else, indicates unknown where a match is not found.

4, PROPOSED GENETIC ALGORITHM TRAINED NEURAL NETWORK MODEL

[9] Evolutionary models lends itself to resolve complex task by exploiting historic data whose underlying probability feats are found as optimal solution. It explores domain knowledge and symbolic reasoning expressed as mathematical model to yield a heuristic method that is tolerant to noise, uncertainty and ambiguities; And, exhibits: (a) robustness, (b) flexibility, and (c) continuous adaptation, as feat in the model in its bid to resolves a constraints within the system. It also mimics agents in search of optimal solution (food and survival) in a domain space, and is mostly inspired by behavioral pattern and natural laws of biological evolution.

It yields output feats with uncontrolled parameters as input (that are not explicitly present from the outset); But, modeled therein via boundary values in domain space confined to real parameters [8] and derives via experience, ability to recognize behavioral feats from observed data, and can suggest optimal solution of high quality that is void of over-fitting, irrespective of modification made to the model via other approximations with multiple agents. These constantly affect the quality of the solution. Also, many of such heuristics are combined to create hybrid(s) that seeks to explore the structural differences in the statistical-method(s) used, and help resolve the implications of the conflict constrained on the model by such a multi-agent populated system – as the agents can often create their own behavioral rules based on historic dataset.

Artificial Neural Network

[29] ANN is inspired by biological neurons. Its major feat is its learning ability via examples cum prediction – to make them universal approximators. Each node send/receive signal via Dendrites, which resends the signals to axons as converted by Synapse so that learning occurs by adjusting the its weight. Weight inputs are summed depending on task at hand by its activation function that modulates a system's nonlinear feats to yield an output [14]. Learning adjusts its weights to define its output as induced local field given by:

$$\emptyset = f(net) = f \sum_{i=1}^{m} X_i * W_i j$$
 (1)

Encoded, ANN has 3-basic layers: input, hidden and output – with two configurations: feed-forward (signal flows via input to output without feedback and processing extends over its multiple layers) and recurrent (with feedback of dynamic feats and activation values that undergoes relaxation to evolve the network to a stable state where its values change no more. A change in dynamic feat constitutes its output). Architecture chosen depends on application area and system requirement. Its connections are set via explicit apriori knowledge or via implicit training that teaches the network patterns that changes its weight based on learning rule.

Learning is divided into: supervised, unsupervised and reinforcement [23]. [8-10] notes that in supervised learning, an input vector with a set of desired responses, one for each node, is relayed to output. With forward pass, errors between desired and actual response for each node in the output is found, and used to determine weight changes in learning algorithm. Thus, desired output is provided by catalyst or external teacher. Examples include back-propagation, delta and perceptron rule. In unsupervised (self-organized), output is trained to respond to clusters of pattern at its input so that the system discovers statistically, salient feats of its input sample – with no prior knowledge unto which patterns are classified; Rather, the model develops its representation of input.

In Reinforcement, it learns what to do, maps states to actions to help maximize a numeric reward. The network discovers what actions yield best state via trial/error. The actions may affect, not only the immediate reward, but also the next state and all subsequent rewards. These two feats, trial/error search and delayed reward are its two distinguishing properties [16-17].

Genetic Algorithm Model

[8-10] GA as inspired by Darwinian evolution and genetics (survival of fittest), consists of a population (data) chosen for natural selection with potential solutions to a specific task. Each potential solution is an individual (genes) for which an optimal is found via four operators: initialization, selection, crossover and mutation. Individual with a gene combination close to the optimal is described as being fit. A new pool is created by mating two individuals from current pool. Fitness function is applied to determine how close an individual is to the optimal solution. GA has four steps:

- a. Initialize encodes data into format suitable for selection. Each encodings has its merit and demerit. The fitness function evaluates how close a solution is to its optimal – after which they are chosen for reproduction. If solution is found, function is good; else bad and not selected for crossover. Fitness function has knowledge of task. If more solutions are found, the higher its fitness value.
- b. Selection Fit individuals close to optimal are chosen to mate. The larger the number selected the better chances of yielding fitter individuals. It continues till last two/three remaining solutions are selected to become parents to new offspring. Selection that only mates the fittest is elitist and often leads to converging at local optima.
- c. Crossover ensures gene exchange of fitter individuals, to yield a new, fitter pool.
- d. Mutation alters chromosomes by changing its genes or its sequence, to ensure that a new pool converges to global minima. Algorithm stops if optimal is found or after number of runs (though computationally expensive) if a number of new pools are created. Genes may change based on probability of mutation rate. Mutation improves the much needed diversity in reproduction.

Network Training and Calibration

At training, image Eigen weights are used as inputs and its corresponding binary ID is our desired output. Training is repeated until network can identify all images in the dataset with error function reduced to acceptable value 0.2. Weights and threshold obtained at training are stored in a file to be used during recognition as in fig. 3 [16-17]. With new images, its feature vectors are computed from Eigen faces found before, and the image gets new descriptors, as inputted. Model is then simulated with these descriptors, outputs are compared and if the maximum output exceeds the predefined threshold level, this new face is decided to belong to person with this maximum output; else, if maximum output does not exceed, face is treated as unknown [30-32]. The model adopts BP with momentum learning and trained with GA. The model learns the weight of the multilayer net via back-propagation and minimizes the squared error between desired output and target output propagated back into the network as [19-22] using the algorithm thus:

- a. In each output k, error $\delta_k = \delta_k \leftarrow o_k (1-o_k)(t_k o_k)$ (2)
- b. In each hidden layer h, error $\delta_h = \delta_k \leftarrow o_k (1-o_k) \sum w_{kh} \delta_{k}$, where k ϵ outputs
- c. Updating the weights w_{ij} , we have that: $\Delta w_{ij} = \eta \delta_j x_{ji}$ $W_{ii} \leftarrow W_{ij} + \Delta w_{ij}$ (where $\Delta w_{ij} = \eta \delta_j x_{ji}$)

With Back-Propagation Algorithm, we have:

- 1. Initialize weights to random values, randomly choose inputs (μ)
- 2. Propagate the signal forward through the network
- Compute iL in the output layer (oi = yiL) diL = g'(hiL) [dui-yiL], where hiL is input to the *i*-th unit in *I*-th layer, and g' is activation function derivative g.
- 4. Compute Deltas for newer layer propagating errors backwards; $\delta i = g'(hil) \sum w_i |l+1| \delta i|+1$, for l = (L-1)..1
- 5. Update weights using $\Delta w | j 1 = \eta \delta i | y| 1 j$
- 6. Go to step 2, repeat for next pattern until error in output layer is below pre-specified threshold, or maximum number of iterations is reached.

Since BP is often stuck at local minima in its search where a global optima exists; But, the smaller the learning-parameter η , the smaller the changes to synaptic weights in the network from one move to next, and the smoother will be the trajectory in the weight space. The improvement is attained at the cost of slower learning rate. Else, if learning-rate parameter is large so as to speed up rate of learning, a resultant large change in the synaptic weights yields an unstable network. Learning rate is increased yet avoiding danger of instability by modifying the Delta rule to include a momentum as:

 $\Delta w_{ii}(n) = \alpha \Delta w_{ii}(n-1) + \eta \delta_i(n) yi(n) \quad (3)$

a is positive number or momentum constant, added to cross such local optima. Once feedforward network is trained, it can be tested on a database of images. The output neuron which fires corresponds to the class of the image.

5. EXPERIMENTAL FRAMEWORK FOR GANN

Dataset used is the Olivetti Research Laboratory with 10 orientations of each face. The framework is implemented as thus:

- a. Acquisition/Enhancement: Images are acquired via 480 x 640 pixels camera with proper lighting, and is subjected to skin color and gray scale to remove background noise, isolate skin portion and minimize contrasts effects such as texture, lighting etc. It is then scaled and any image taken via varying capture device is reduced to same resolution to apply the Eigen algorithm efficiently. Enhancements like filtering, clipping and edge detection are made.
- b. Feat Extraction After preprocessing, enhanced image passes through to find key feat used for classification. Training file has m Eigen weights for each m image input. Each image name followed by its id and Eigen weight. Thus, module composes the feature vector that represents the image via a standard algorithm [12] as:

//Algorithm for feature extraction

- a. Assume training sets: $G_1 \dots G_m$ m = number of image b. Find Mean image: $\Psi = \left(\frac{1}{m}\right) \sum G_i$...For i = 1 to m
- c. Calculate mean-subtracted given by: $\mathbf{F} = \mathbf{G}_i - \mathbf{\Psi} \text{ for } \mathbf{i} = 1 \text{ to } \mathbf{m} (4)$
- Mean subtracted matrix $A = [F_{1}, F_{2}, ..., F_{m}]$ d. Covariance matrix = $A_{mp} \times A_{pm}^{\dagger}$
- e. Find Eigen values λ_m and Eigen vectors V_{mm}
- f. Eigen faces $U_k = \sum F_n V_{kn}$, k=n=1,2,...m
- g. Eigen weights $W_k = U_k T (G \Psi) k = 1, 2, ..., m$
- ANN Learning: BP uses error derivative propagated back to contributing neurons with weights С updated as [24-28]:
 - 1. Set all weights to small random values.
 - 2. Input to each node these: x_i input from previous node, w_i is weight so that, with sigmoid function we compute:

Input
$$a = \in X_i W_i$$
 (5)
Output $Y = f(x) = \frac{1}{(1 + e - x)}$ (6)

3. The Errors, desired and actual output are sent back to nodes with updated weights via Eq. 6 (wii is weight from node i to j at t, n is learning rate, oj is output of j and μ is error for node j). Thus, $W_{ij}(t+1) = W_{ij} + \eta \mu j o j \qquad (7)$

Output node yields: $\mu j = koj(1 - oj)(tj - oj)$ (8) Hidden nodes with μ_k as next nodal error term:

 $\mu \mathbf{j} = k \mathbf{o} \mathbf{j} (\mathbf{1} - \mathbf{o} \mathbf{j}) \in \mu \mathbf{k} . \mathbf{w} \mathbf{j} \mathbf{k} \quad (9)$



d. Crossover/Mutation – The Time-Lag Recurrent Net (MLP with short-memory) with local recurrent connections is used as it requires a smaller network to learn temporal task, it is more plausible and computationally more powerful than other adaptive models. It uses BP-in-time and momentum supervised learning for training so that its output at t is used along with a new inputs to compute output in t+1, as in response to dynamism.

[8-10] notes the GANN model adopts a solution space (with 15-initialized images, corresponding to a non-fixed face orientation) that conforms to its belief space as thus: (1) Normative (individuals have face genes from 1-to-15), (2) Domain (individual have integer as genes), (3) Spatial (1-to-15 once of same image with varied orientation, and (4) Temporal (mutation must not alter values of fixed data). An additional influence function ensures the belief space is adhered to and its rounding function ensures all values are integer and random numbers generated are not repetitions of fixed numbers in the pool [9, 24, 26, 28].

With GA, model is initialized with data whose fitness is determined and moved over from GA training to ANN via sub-pool selection of 10-images to form a new pool via tournament method as it is easier, more efficient to code, best suit for parallel architectures and selection is easily adjusted as it allots random numbers to individuals.



Both crossover (multi-point) and mutation is applied – to distort images (random generate Gaussian distribution) that corresponds to crossover points. The new pool also undergoes mutation from which 3-genes are selected for mutation and are allocated new random values that still conforms to belief space. New individuals replace old ones with low fitness values (creating a new pool) [18]. Number of mutation applied depends on how far GA has progressed (how fit is the fittest individual in pool). Thus, knowledge of solution has direct impact on how algorithm is employed. It stops if best individual has a fitness of 0. Thus, solution is found. New individual also undergoes mutation so that 3-random genes are selected and allotted random value that still conforms to its belief space. New individuals replace older ones of low fitness values (creating new pool). This continues till a gene with fitness of 0 is found. Initialization/selection ensures first 3-beliefs are met; while mutation ensures the fourth belief is met. In addition, an influence function (best fitness) helps influence how many mutations takes place.

. 111



- e. Testing GA is applied to modified or yield new weights and threshold values, stored and used to rettrain via the cross-validation and testing phase. So as to help with further deformation of the images for recognition.
- f. Recognition Eigen weights of the image to be identified is passed as input to the already trained network and the outputs obtained. The outputs of individual neurons of the output layer are then rounded off to the nearest 0 or 1 to form a valid binary ID. This binary ID is then checked against the database to validate the authenticity and display the details of the face if identified.

6. DATA FINDINGS

Our modification to ORL database was used with additional faces samples. The different images for each subject provide variation such as lighting, facial features (such as glasses), and slight changes in head orientation etc. This dataset contains a bias as to type of face representation. Males outweigh females in the subject population by a factor greater than 3-to-1 as well as its bias now with face images of Blacks, Caucasian and other(s) for fair representation. Despite shortcomings in the dataset such as age representation, ORL provides a good starting point for face recognition as long as we understand that test performance is an optimistic estimate [15]. Face recognition accuracy for varying number of hidden neurons is summarized in Table 2-3. For larger image sizes, there is little change in recognition accuracy for hidden nodes more than 40. But for smaller images, the plateau is reached when number of neurons is 20. Thus, ANN requires more nodes to extract hidden feats from large image sizes. Also, with GANN, recognition of more than 90% can be achieved [27-28]. Appendix is Java Program to accompany the study and the ORL pool.

Number of Eigen Faces	Neurons in Hidden Layer	Recognition Rate (%)	
	5	40	
40	10	58.3	
40	15	72	
	20	74.5	
	5	41	
45	10	73.5	
	15	80.8	
	20	88.5	
	5	54.8	
50	10	73	
	15	87.5	
	20	90.8	
	5	50	
55	10 81.8	81.8	
	15	88.5	
	20	91.5	
	5	48.8	
60	10	81.3	
80	15	89.5	
	20	92.8	
	5	52	
65	10	81.8	
	15	91.5	
	20	94.3	
	5	50	
70	10	86.3	
	15	92.3	
	20	93.3	

Table 2: Accuracy using a number of Eigen faces and neurons

Resolution Hidden Nodes	and	10	20	30	40	50	60
50x40		44	67	71.5	91.5	93.5	92
20x20		10	79	87.5	89	92.5	93.5
10x10		6	91.5	89.5	88	90	88.5

Table 2: Network Recognition Accuracy

REFERENCES

- [1] Atick, J., Griffin, P and Redlich, A., "Statistical approach to shape from shading: Reconstruction of 3D face surfaces from single 2D images" Neural Computation, 1996, Vol. 8, pp. 1321-1340.
- [2] Batruni, R., "Multilayer network with piecewise-linear structure and BP-learning", IEEE Transaction on Neural Network, 1991, Vol 2, p395
- [3] Cardinaux, F., Sanderson, C and Bengio, S., "User Authentication via Adapted Statistical Models of Face Images", IEEE Transaction on Signal Processing, 2006, 54(1), pp. 361 373.
- [4] Cunado, D., Nixon, M and Carter, J., "Automatic extraction and description of gait models for recognition purposes", Computer Vision and Image Understanding, 2003, 9(1), 1–14.
- [5] Fleming, M., and Cottrell, G., "Face categorization using unsupervised feature extraction", International Journal of Computers and Neural Networks, 1990, 90(2), pp 99-106.
- [6] Jain, A.K., Ross, A and Pankanti, S., "Biometric: tool for data security", IEEE Transaction Information Forensics and Security, 2006, 1(2), pp. 125–144.
- [7] Marin, C., Penedo, M., Penas, M., Carreira, J and Gonzalez, F., "Personal authentication using digital retinal images", Springer Journal of Pattern Analysis and Application, 2006, 9(1), pp. 21– 33.
- [8] Ojugo, A.A., and Yoro, R.E., "Computational intelligence in stochastic solution for Toroidal Queen problem", PICA: Progress in Intelligence Computing and Applications, 2(1), pp 46 56, 2013.
- [9] Ojugo, A.A., Emudianughe, J., Yoro, R.E., Okonta, E.O and Eboka, A., "Hybrid artificial neural network gravitational search algorithm for rainfall runoff simulation and modeling in Hydrology", Progress in Intelligence Computing and Applications, 2(1), 2013, pp 22 – 33.
- [10] Ojugo, A.A., Eboka, A.O., Okonta, E.O., Yoro, R.E and Aghware, F., "Genetic algorithm rule-based intrusion detection system", J. Emerging Trends in Computing and Information Systems, 2012, 3(8), p1182.
- [11] Pham, D and Karaboga, D., "Training Elman and Jordan networks for system identification using genetic algorithms", Artificial Intelligence in Engineering, 1999, 13, pp.107–117.
- [12] Rizon, M and Hashim, F.,"Face recognition using Eigen faces in neural networks", American J. of Applied Sciences, 2006, 3(6), pp51.
- [13] Ross, A., Dass, S and Jain, A.K., "Deformable model for fingerprint match", Elsevier Journal of Pattern Recog., 2005, 38(1), pp. 95–103.
- [14] Sontag, E.D., "Learning for continuous-time recurrent neural networks", Systems and Control Letters, 1998, 34, pp. 151-158.
- [15] Swets, D and Weng, J., "Using discriminant Eigen features for image retrieval" IEEE Transaction Pattern Analysis and Machine Intelligence, 1996, 18, pp. 831-836.
- [16] Turk, M and Pentland, A., "Eigen faces recognition", Neuroscience, 1991, 3(1):71-86.
- [17] Volkan, A., "Face Recognition via Eigen faces in Neural Networks", Masters' Thesis, EEE Department, Graduate School Of Natural And Applied Sciences Of The Middle East Technical University, 2003.
- [18] Weaver, A., "Biometric Authentication", IEEE Computer Society, 2006, 39(2), pp. 96-97.
- [19] Wen, C and Ma, X., "A max-piecewise-linear neural network for function approximation", Neurocomputing, 2008, 71, pp. 843–852.
- [20] Wright J., Yang A.Y., Ganesh A., Sastry S and Ma, Y., "Robust Face Recognition via Sparse Representation" Pattern Analysis and Machine Intelligence, 2009, 31(2), pp.210-227.

- [21] Yun, L and Haubler, A., "Artificial evolution of neural networks and its application to feedback control", 1996, Artificial Intelligence in Engineering, 10(2), pp. 143–152.
- [22] Zhujie, Y and Yu, Y., "Face Recognition with Eigen face", published Masters' thesis of the Hong-Kong University of Science and Technology, 2010, Kowloon, pp. 434-438.
- [23] Kumar, S., "Neural Networks: Classroom Approach", McGraw-Hill, Singapore, 2005
- [24] Li, S and Jain, A.K., "Handbook of Face Recognition", 2004, NY: Springer.
- [25] Maltoni, D., Maio, D., Jain, A.K and Prabhakar, S., "Handbook of Fingerprint Recognition", 2003, Springer, New York.
- [26] Wayman, J., Jain, A.K, Maltoni, D and Maio, D., "Biometric Systems: Technology, Design and Performance Evaluation", 2005, NY: Springer
- [27] Dabbah, M., Woo, W and Dlay, S., "Secure Authentication for Face Recognition" Proceeding of IEEE on Computational Intelligence in Image and Signal Processing, 2007, pp. 121 126.
- [28] Dmitry, B and Starovoitov, V., "Access Control by Face Recognition Using Neural Networks", In Proceeding of 2nd International Conference on AI, 2002, pp 428–436
- [29] Hopfield, J.J., "Neural networks and systems with emergent collective computational abilities", In Proceedings of National Academy of Sciences, 1979, pp.2554–2558.
- [30] Phillips, P, Rauss, J and Der, S., "Face Recognition Technology: evaluation methodology," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1997, Puerto Rico
- [31] Ross, A and Govindarajan, R., "Feature level fusion using hand and face biometrics", Proceeding of SPIE Conference Biometric Technology for Human Identification, 2005, pp. 196–204.
- [32] Zhang, J., Yan, Y and Lades, M., "Face Recognition: Eigen face, Elastic Matching, and Neural Nets" Proceedings of IEEE on Memetic Algorithm, 1997 85(9):1423–1435.

<u>4</u>		
	Select Trainning Im	age
Train Subject	0921/2011 👻	
Select Image	roject/att_faces/s9/Training	Browse
	Train	
	Triaining in Progre	ss

Appendix 1: Training Image selection and processing

Appendix 2: Testing phase and matched patterns processing

	Reco	gnition Report			
cognition Sta	tus: Subject Identified				
	Input Image		Database Match		
Details: Name	iboi Jonathan	Matric Number	0921/2011		
Department	Petroleum Engineering	RAddress	12. Oloko Street, Mushin, Lagos		
Email	jzjoboj@jahos.com				
		Ok			





Appendix 3: Face recognition accuracy for matched patterns

Appendix 4: The Original ORL

