International Journal of Intelligent Computing and Information Science

A Face Recognition System Based on Deep Learning (FRDLS) to Support the Entry and Supervision Procedures on Electronic Exams

A.E. Amin

Department of Computer Science,
Mansoura University, Mansoura
35516, Egypt
ahmedel_sayed@mans.edu.eg

Abstract:
The novelty of this paper is represented in using some artificial intelligence techniques in the entry control to the electronic exams (E-exam) in addition to monitoring students and distinguish the situation they are during the E-exam. Therefore, the proposed system divides into two main parts, the first part to support E-exams to handle some of the weaknesses points such as validation from students' entry by using deep learning. The Self-Organized Maps (SOM) neural network was used to recognition on students' faces. SOM is characterized by its efficient for faces' image data management, as well as it's the closest technique to match inputted untrained faces' images with a database of trained faces' images accurately. On the other part, the Bag of Words model (BoWM) is used to discriminate the status of students during the exam process. The BoWM is based on Speeded-Up Robust Features (SURF) that building on the strengths of the leading existing detectors and descriptors by using a Hessian matrix. Then extracts a report showing the status of the student such as confusion, concentration, cheating ... etc.
From the experimental results, the proposed system was verified images of students' faces with high accuracy and execution time have a significant indication. Determining the status of the student during the exam by adopting the technique of retrieving documents known as the bag of word model, which proved the accuracy of determining the status of the student arrived in some cases to 100%.

Keywords: *Machine learning techniques; Face Recognition; Self-organize maps neural network; Bag of Words model.*

1. Introduction:

Developing countries are now interested in automating the educational process to save time, effort and money. Among the procedures of the educational process is the testing phase, which aims to assess the students' knowledge. Traditional exam systems require more effort on the part of lecturers, where there are several methods of setting and selecting questions as well as the correction of exams and examination works. In traditional examination systems, student affairs departments are suffering to identify and verify students during exams.

So researchers began to study the automation of exams and the use of modern computer techniques in the effective evaluation of students' knowledge electronically without any impact on the traditional examination procedures. In addition, use biometrics that support security control, authentication and integrity in the examination process as well as the use of digital cameras and identification devices, which helps to prevent cheating, and identify the students during the exam.

Due to the increasing sophistication of information technology and the impressive success of artificial intelligence applications in various fields, this field has been used to propose the integration of some artificial intelligence techniques with exam systems to deal with problems resulting from traditional exam systems.

The idea of integration is based on the electronic exam system, and some artificial intelligence techniques on the interaction between the data sets within the educational institution, such student data, faculty data, courses data, organizational structure data, and the available knowledge within the exam range such the student identification, Choose the name of the course, choose the questions, the type of the questions... etc.

There are five main models that have been classified for the integration process [1]: Stand-alone models [2], Transformational models [3], Loose Coupling models [4], Tightly Coupled Models and Fully Integrated Models [5].

Loose coupling model is one of the first real models for integrating artificial intelligent systems. This model allows interaction between systems with different characteristics. The typical configuration of this model consists of several phases are preprocessing, post- processing, co-processes and user interface. The pre-processing phase is responsible for handling data such as removing noise and errors, identify objects and recognize patterns, before transferring them to a knowledge-based system. The post-processing phase is to classify inputs, process data and prepare them as outputs. While the co-processing manipulates allows for interaction and cooperative behavior between systems and knowledge base system, which helps to data refinement, iterative problem solving and decision-making. Finally, the user interfaces phase that increases the flexibility of user's interaction with knowledge-based systems.

The rest of the paper is structured as follows. In Section 2, we provide Materials and Methods used in this paper such as neural networks, and some techniques for machine learning. Section 3 introduces in detail the proposed method to image retrieval approaches that serviced the aim of paper. Several comprehensive experiments are performed to evaluate the effectiveness of our approach in Section 4. The experimental settings, evaluation metrics, benchmark datasets, and results are provided in Section 5. Finally, it concludes this paper in Section 6.

2. Materials and Methods:

2.1. Self-Organize Maps Neural Network:

There are many types of artificial neural networks (ANN) and their methods of learning [6]. One from ANN is Self-Organizing Maps (SOM) that can be called Kohonen maps also, relative to the Finnish professor Teuvo Kohonen [7]. SOM is subject to an unsupervised learning method that acts as a low-dimensional representation of training samples while maintaining the topological properties of the inputted image area.

The main aim of SOM [8] is transforming an input arbitrary dimension signal pattern to a one or two-dimensional discrete map then performing of this transformation is adaptive in an ordered fashion of the topology. Learning the SOM to recognize the group of image features' vector in such a way that reshaping the positions of neurons physically where converge with each other in the neuron layer to be similar with input vectors. This providing dimensionality reduction with maintaining slight changes in the image sample [9].

The architecture of SOM consists of single layer feedforward networks while the output layer that is usually 2D or 3D are arranged in a low dimensional grid [10]. All output neurons are connected by each input. Each neuron is attached with the same dimensionality as the input vectors by a weight vector. The output grid dimension is usually a lot lower than the number of input dimensions. The SOM algorithm stages are summarized as follows [11].

SOM is commonly used as a classification tool, where it's easy to see the relationships between huge amounts of data. SOM is typically applied to represent central dependencies within map data, so it is easy to use to recognize faces images.

1. **Initialization**:
   Choose random values for the initial weight vectors W_j.
2. **Sampling**:
   Draw a sample training input vector X from the input space.
3. **Matching**:
   Find the winning neuron I(X) that has weight vector closest to the input vector; the minimum value of:
   $$d_j(X) = \sum_{i=1}^{D}\left(x_i - w_{ij}\right)^2.$$
4. **Updating**:
   Apply the weight update equation:
   $$\Delta w_{ij} = \eta(t)T_{j,I(x)}(t)\left(x_i - w_{ji}\right)$$
   Where:
   $T_{j,I(x)}(t)$ ⟶ A Gaussian neighborhood and;
   $\eta(t)$ ⟶ The learning rate.
5. **Continuation**:
   Keep returning to step 2 until the feature map stops changing.

2.2. Discriminate State of Face:

The concept of classification in the documents classification field based on the separate word count vector was exploited and applied in the classification of images where the features of images were treated as words. So, the Bag of Words model (BoWM) [12] is used to treat the image as the document, therefore, the visual words can be defined as a vector of local image features vocabulary that occurrence counts. To achieve this aim the Speeded-Up Robust Features (SURF) [13] is used, which is a fast and robust algorithm for local, similarity invariant representation and comparison of images. The SURF algorithm relies on the local descriptor approach [14], that divided into three main steps are: defined the interest points, built the orientation invariant descriptors and feature matching.

2.2.1. The interest points defined:

The Second Gaussian Derivative Operators (SGDO) [15] is used to detect features. The face images of students are represented as linear scale-space by SGDO at several scales (L). The SURF approach used box filters that considered uniform kernels with separable rectangular to approximate the Gaussian kernels and its spatial derivatives. As a consequence, the time-consuming procedure of SURF is speeding up.

In the students' face image (Im), preprocessing is achieved by using the filters with finite support (Ω) to approximate the kernels to produce integral image (Im_{INT}) as:

$$\forall(x,y)\epsilon\mathbb{Z}^2,\ (f*g)(x,y)=\sum_{(i,j)\in\Omega} f(x-i,y-j)\,g(i,j) \quad ..... (1)$$

SURF used the scale-normalized determinant of Hessian ($|H^L(...)|$) [16] as:

$$|H^L(Im)|=\frac{1}{L^4}\left(D_{xx}^L Im\cdot D_{yy}^L Im-\left(wD_{xy}^L Im\right)^2\right) \quad ..... (2)$$

Where: $L \rightarrow$ scaling relation, $w \rightarrow$ constant weighting factor, $\frac{1}{L^4} \rightarrow$ the normalization factor, D_{xx}^L& $D_{yy}^L \rightarrow$ the second order operators at scale L where:

$$D_{xx}^L Im=(B_{\Gamma 1}-3B_{\Gamma 2})*Im$$

$$D_{yy}^L Im=(B_{\Gamma 3}-3B_{\Gamma 4})*Im$$

$$where,\begin{cases}\Gamma 1=\left[\!\left[-\frac{3L-1}{2},\frac{3L-1}{2}\right]\!\right]\times[\![-(L-1),(L-1)]\!],\\ \Gamma 2=\left[\!\left[-\frac{L-1}{2},\frac{L-1}{2}\right]\!\right]\times[\![-(L-1),(L-1)]\!]\subset\Gamma 1\end{cases}$$

$$D_{xy}^L Im=\left(B_{\Gamma_{NE}}+B_{\Gamma_{SW}}-B_{\Gamma_{NW}}-B_{\Gamma_{SE}}\right)*Im$$

Where; $B_{\Gamma\_\_} \rightarrow$ Ball of radius $\Gamma\_\_$, $NE \rightarrow$ North-East, $SW \rightarrow$ South-West, $NW \rightarrow$ North-West, $SE \rightarrow$ South-East.

$$\begin{cases}\Gamma_{NE}=[\![1,L]\!]\times[\![1,L]\!]\\ \Gamma_{SW}=[\![-L,-1]\!]\times[\![-L,-1]\!]\\ \Gamma_{NW}=[\![-L,-1]\!]\times[\![1,L]\!]\\ \Gamma_{SE}=[\![1,L]\!]\times[\![-L,-1]\!]\end{cases}$$

To detect the key points in images that represents the highly discriminant of interesting features in images. So, the local maxima of the $|H^L(...)|$operator, which detected by performing an exhaustive comparison voxel of the discrete box-space with its 26 nearest-neighbors, is considered the interest points or key points. In each local maxima, the corresponding interest point location in the box space is refined using quadratic fitting [17] as shown in the following algorithm.

After applying the algorithm, a set of (K) interest points are obtained as:

$$\{X_n: (x_n, y_n, L_n) \in [0, M-1] \times [0, N-1] \times [0,65]\}_{n=1,2,\ldots K}$$

2.2.2. The interest points description:

After obtaining a set of interest points in the area of the box-space, which represents the most salient features of the image of the face of the student, the observed scale-space coordinates are continued as a result of the refinement step. The matching between different images needs not only geometric invariant representation but also to eliminate noise, illumination or contrast change. So, the local descriptor of each interest point neighborhood is encoded. To achieve the similarity invariance, there are three main phases are interest point scale (i_P_{Sc}), dominant orientation of an interest point (i_P_{DO}), and SURF descriptors as shown in figure 1.

2.2.3. Matching of Feature:

There are two techniques are used to the matching of students' faces image features are Euclidean distance and Near Neighbor Distance Ratio thresholding [18]. Both of query and original images $(Im_Q, Im_O)$are represented by sets of interest points $(i\_p_Q^k, i\_p_O^k)$ with the SURF descriptors that corresponding to insert points $(\{SURF_Q^k\}\ and\ \{SURF_O^k\})$.

- Input: the detected point (X_0) at scale(L_0): $X_0: (x_0, y_0, L_0)$
- Procedure:

1. Step parameter for finite different scheme: $p \rightarrow 2^o i + 1$
2. Hessian matrix H_0:

$$H_{xx}(X_0) = \frac{1}{p^2}\left(|H^{L_0}(Im)|(x_0+p, y_0) + |H^{L_0}(Im)|(x_0-p, y_0) - 2|H^{L_0}(Im)|(x_0, y_0)\right),$$

$$H_{xy}(X_0) = \frac{1}{4p^2}[|H^{L_0}(Im)|(x_0+p, y_0+p) + |H^{L_0}(Im)|(x_0-p, y_0-p) - |H^{L_0}(Im)|(x_0-p, y_0+p) - |H^{L_0}(Im)|(x_0+p, y_0-p)].$$

$$H_{xL}(X_0) = \frac{1}{8p^2}[|H^{L_0+2p}(Im)|(x_0+p, y_0) + |H^{L_0-2p}(Im)|(x_0-p, y_0) - |H^{L_0+2p}(Im)|(x_0-p, y_0) - |H^{L_0-2p}(Im)|(x_0+p, y_0)].$$

$$H_{LL}(X_0) = \frac{1}{4p^2}[|H^{L_0+2p}(Im)|(x_0, y_0) + |H^{L_0-2p}(Im)|(x_0, y_0) - 2|H^{L_0}(Im)|(x_0, y_0)].$$

3. Gradient Vector (d_0):

$$d_x(X_0) = \frac{1}{2p}\left(|H^{L_0+2p}(Im)|(x_0+p, y_0) - |H^{L_0+2p}(Im)|(x_0-p, y_0)\right),$$

$$d_L(X_0) = \frac{\partial i}{\partial L} d_i(X_0) = \frac{1}{4p}\left(|H^{L_0+2p}(Im)|(x_0, y_0) - |H^{L_0+2p}(Im)|(x_0, y_0)\right),$$

4. Max. Refinement (ξ):

$$X = X_0 + \xi, \quad where \quad \xi = (\xi_x\ \ \xi_y\ \ \xi_L) = -H_0^{-1} d_0.$$

5. Check precision improvement:

$$if\ max\left(|\xi_x|, |\xi_y|, \frac{1}{2}|\xi_L|\right) < p\ \ Then$$

   5.1. Refinement using 2nd order Taylor expansion,

$$X = X_0 + \xi, \quad f(X) = f(X_0) + \xi^T \cdot Df(X_0) + \frac{1}{2}\xi^T \cdot D^2 f(X_0) \cdot \xi + O(\|\xi\|^3)$$

6. Return True, $X: (x, y, L)$

   6.1. Else Return False

- Output: True/False: $X: (x, y, L)$

Where (k) is a unitary Euclidean normal vector contained on 64 elements as:

$$V \in [-1,1]^{64}, \quad \|V\| = 1.$$

So, the Euclidean distance(E_d) is computed as:

$$(E_d)_{Im_Q,Im_O} = \sqrt[2]{\left| i\_p_Q^k - i\_p_O^k \right|_{k=1}^{64}}$$

There are several million from interest points, therefore the nearest neighbor distance ratio threshold matching $\left(nnd_R^{Th}\right)$ is used to validate the correct matches from the fake ones as:

$$nnd_R^{Th} \in arg \min_{nnd_R^{Th}} (E_d)_{Im_Q,Im_O}$$

Initialization:

Box-Space parameter: $L, \sigma \rightarrow Scale\ Variable.$

$\mathcal{B}_{6\sigma_k} \rightarrow disk\ of\ radius\ 6\sigma_k\ with\ center\ (x_k, y_k),\ u \rightarrow numerical\ image,$

$U \rightarrow discrete\ image, \Omega \rightarrow pixel\ grid, \Phi \rightarrow Score\ Vector, \theta \rightarrow Orientation.$

$\angle\phi_k(x,y) \rightarrow angle\ bet.\ vector\ \emptyset \in \mathbb{R}^2 and\ canonical\ vector\ (1,0)^T.$

$(u,v) \rightarrow gradient\ samples\ coordinates\ according\ to\ the\ discriptor\ grid.$

$(x,y) \rightarrow pixel\ grid\ (\Omega) coordinates.$

$G_\sigma \rightarrow Gaussian\ kernal\ with\ standard\ deviation\ \sigma$

Insert live image($\mathcal{U}$)

Discretized image($\mathbf{u}$)

Calculate (i_P_{Sc}) by:

$$\boldsymbol{\sigma(L) = \frac{1.2}{3}(2^O \times i + 1) = 0.4L}$$

Where:

$$\boldsymbol{L = 2^O i + 1}$$

$$\boldsymbol{L \rightarrow Scale, O \rightarrow Octave, i \rightarrow level}$$

(i_P_{DO}) Calculation:

- Weighted gradient computation:

$$\boldsymbol{\forall (x,y) \in \Omega \cap \mathcal{B}_{6\sigma_k},}$$

$$\boldsymbol{\emptyset_k(x,y) = \begin{pmatrix} D_x^{L_k} \\ D_y^{L_k} \end{pmatrix} \circ U(x,y) \cdot G_1\left(\frac{x - x_k}{2\sigma_k}, \frac{y - y_k}{2\sigma_k}\right)}$$

$$\boldsymbol{where:\ D_x^L U = \left(B_{[\![-l,-1]\!]\times[\![-l,l]\!]} - B_{[\![1,l]\!]\times[\![-l,l]\!]}\right) * U}$$

$$\boldsymbol{D_y^L U = \left(B_{[\![-l,l]\!]\times[\![-l,,-1]\!]} - B_{[\![-l,l]\!]\times[\![1,l]\!]}\right) * U}$$

$$\boldsymbol{G_\sigma(i,j) = \frac{1}{C_K} G_\sigma(i,j), \qquad C_K = \sum_{i,j=-K}^{K} G_\sigma(i,j)}$$

- Orientation Score function:

$$\Phi_k(\theta) = \sum_{(x,y)\in B_{6\sigma_k}(x_k,y_k)} \phi_k(x,y) \times 1_{[\theta - \frac{\pi}{6}, \theta + \frac{\pi}{6}]}(\angle\phi_k(x,y))$$

- Dominant orientation:

$$\theta_k = \angle\Phi_k(\theta^\star)$$

$$where:\ \theta^\star \in \underset{\theta\in\Theta}{\operatorname{argmax}} \|\Phi_k(\theta)\|$$

$$\{X_n : (x_n, y_n, L_n) \in [0, M-1] \times [0, N-1] \times [0,65]\}_{n=1,2,\ldots K}$$

SURF Descriptors:

- Input: oriented interest point $X_k : (x_k, y_k, L_k, \theta_k)$.
- Scale Normalized Sampling:

  $\mathcal{R}_{i,j} \rightarrow Square\ subregion\ with\ 5\sigma_k = 4\ length$

  $\sigma_k \rightarrow the\ step\ of\ regularly\ gradient.$
- Change of coordinates:
  - Similarity Bet. $(u,v)\ and\ (x,y)$ is:

$$S_k : \begin{pmatrix} u \\ v \end{pmatrix} \mapsto \begin{pmatrix} x \\ y \end{pmatrix} = \sigma_k \cdot \mathcal{R}_{\theta_k} \begin{pmatrix} u \\ v \end{pmatrix} + \begin{pmatrix} x_k \\ y_k \end{pmatrix}$$

$$where,\ \mathcal{R}_\alpha = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

$$for\ approximate:$$

$$(x,y) = \lfloor S_k(u,v) \rceil = (\lfloor x \rceil, \lfloor y \rceil) \in \Omega$$

- Gradient Normalization:

$$\forall (u,v) \in \mathcal{R}, \begin{pmatrix} d_x(u,v) \\ d_y(u,v) \end{pmatrix} = \mathcal{R}_{-\theta_k} \begin{pmatrix} D_x^{L_k} \\ D_y^{L_k} \end{pmatrix} u(x,y) \times G_1\left(\frac{u}{3.3}, \frac{v}{3.3}\right)$$

- Gradient Statistics:

  $for\ \mathcal{R}_{i,j}, The\ Statistical\ Vector\ is:$

$$\forall (i,j) \in [\![1,4]\!]^2,\ \mu_k(i,j) = \begin{pmatrix} \sum_{(u,v)\in\mathcal{R}_{i,j}} d_x(u,v) \\ \sum_{(u,v)\in\mathcal{R}_{i,j}} d_y(u,v) \\ \sum_{(u,v)\in\mathcal{R}_{i,j}} |d_x(u,v)| \\ \sum_{(u,v)\in\mathcal{R}_{i,j}} |d_y(u,v)| \end{pmatrix}$$

$$Then:\ \mu_k = \left(\mu_k(i,j)\right)_{1\le i,j\le 4}$$

- Descriptor Normalization:

  $By\ using\ Euclidean\ Norm:$

$$SURF(X_k) = \frac{\mu_k}{\|\mu_k\|_2}$$

- Laplacian at Scale (L):

$$\Delta^L(U) = \frac{1}{L^2}(D_{xx}^L + D_{yy}^L)u$$

SURF Descriptor & Insert Points Orientation

Fig. 1: The similarity invariance phases.

To measure the quality of matching between the query and the most similar candidate, the ratio between the closest first and second nearest neighbors is used as:

$$nnd_R^{Th} \in arg \min_{E_d, s.t.(E_d)_{Im_Q,Im_O} \geq (E_d)_{Im_Q,Im_{O_1}}} (E_d)_{Im_Q,Im_O}$$

Then the corresponding ratio of distances is compared with a fixed threshold (T_f) (the best experimental results for T_f is 0.8) as:

$$if \ \frac{(E_d)_{Im_Q,Im_{O_1}}}{(E_d)_{Im_Q,Im_{O_2}}} \leq T_f \quad then \ (i_{p_Q}^{\ k}, i_{p_{O_1}}^{\ k}) \ is \ validated$$

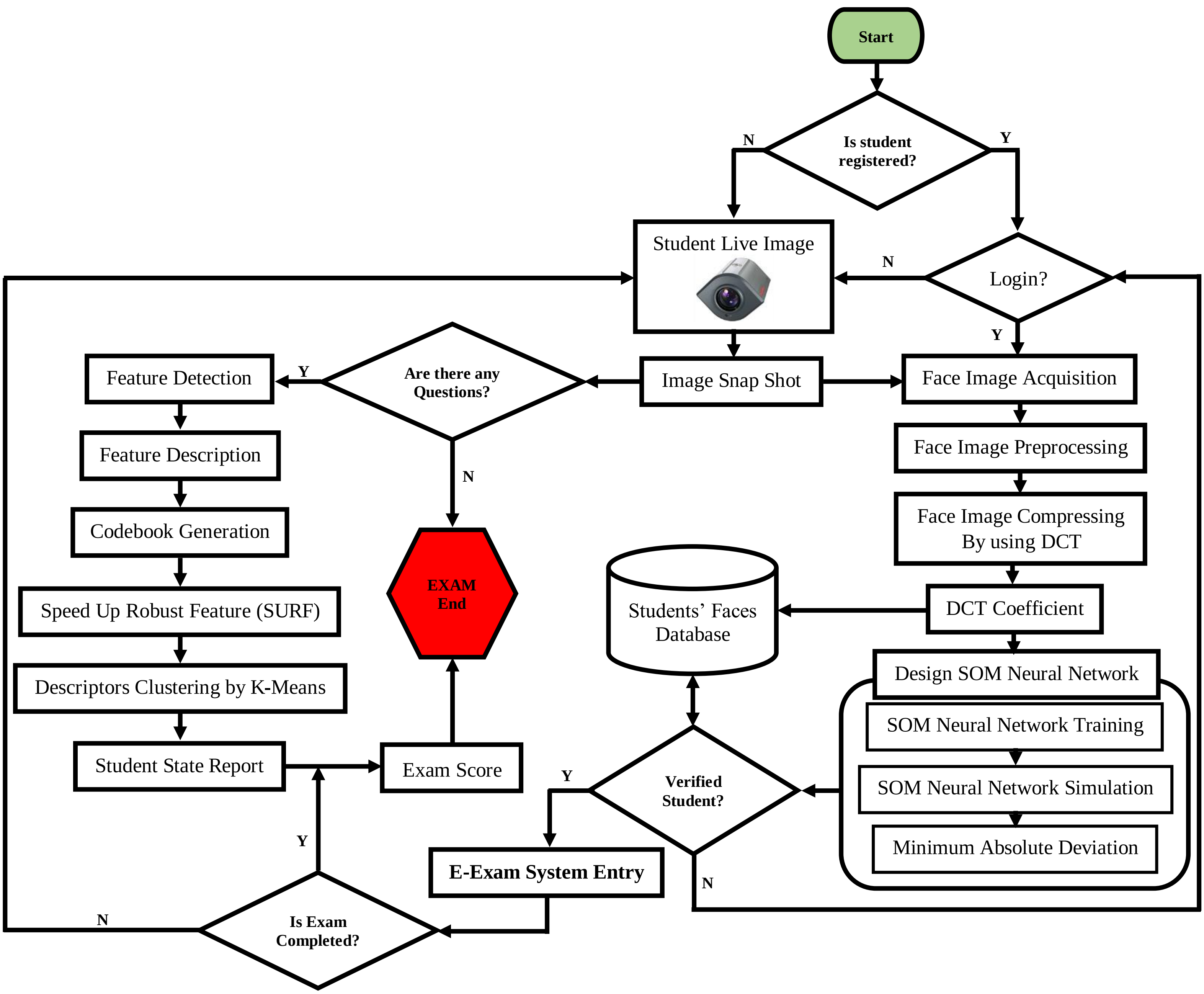


Fig. 2: The proposed method phases.

3. The Proposed Method:

The proposed system architecture for development E-exam systems is divided into two main phases: verification student's face based on self-organize map neural network and Student State Discernment during exam time based on (BoWM). Figure 2 illustrates the two phases and its components.

3.1. Students' Face Verification:

The student's identity will be verified when he/she enters the exam by taking a picture of his / her face and performing the pre-processing stage through several steps are adjusting brightness and contrast, converting to gray level, resizing, compressing by discrete wavelet transform and image data reshaping respectively. Then entering the processed data on a self-organized neural network that verifies the student's identity by comparing it with the database of the previously stored students as shown in figure 3.

3.2. Student State Discernment:

During the examination, the digital camera takes random images of the student's face at intervals controlled by the system. The images are processed, initialized and compared to the student database for follow-up during the examination period to determine the student's concentrate - fear - cheating ... etc. as shown in figure 4.

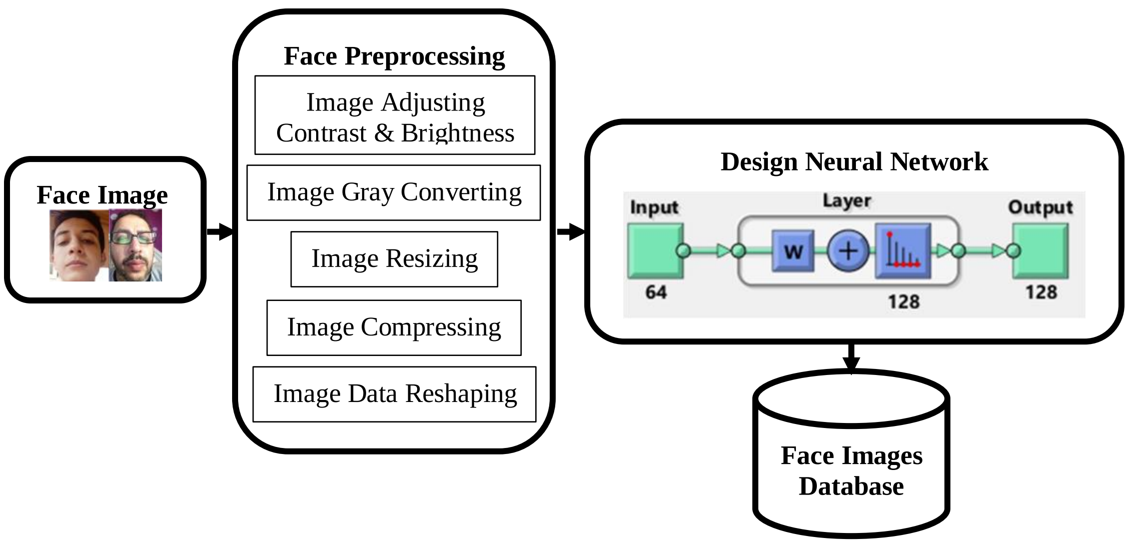


Fig. 3: Self-Organized neural network for Students identification.

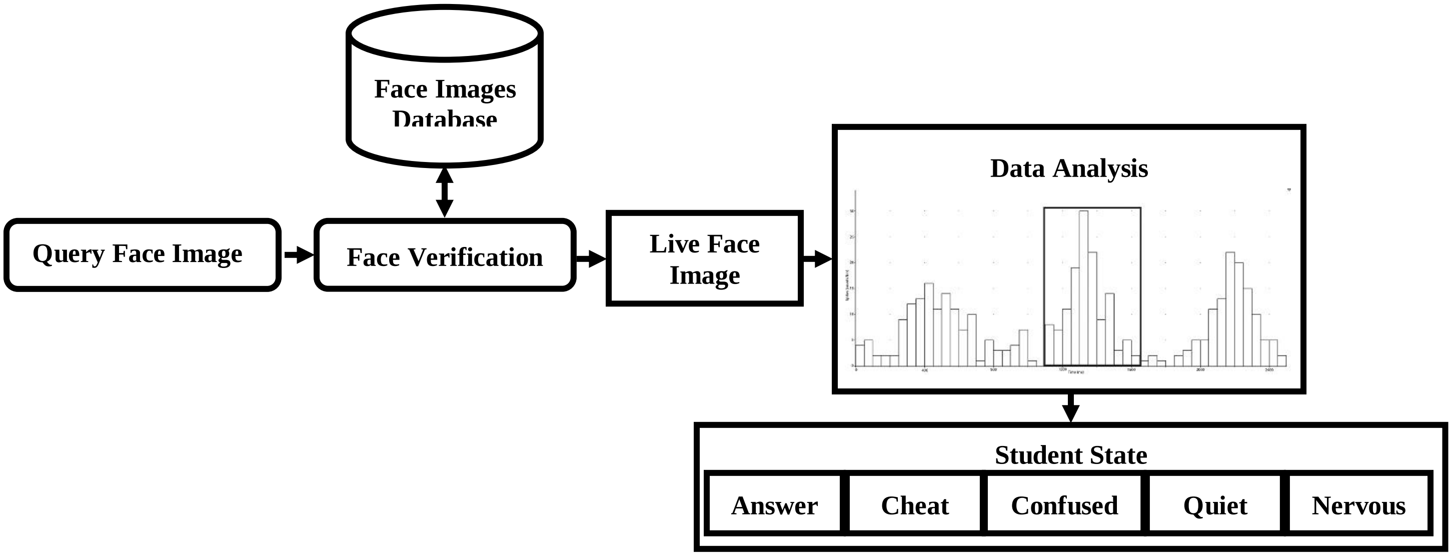


Fig. 4: Students' state discernment.

4. Experimental work:

4.1. Data Set:

The extended Cohn-Kanade (CK+) [19] is used to evaluate the proposed system. CK+ included two types of images, one used to face verification and the other used to emotion recognition. All of the types are extracted from live images videos in a controlled laboratory environment. Although the CK+ dataset is small, it provides well defined facial.

4.2. Students' face image verification:

The proposed system took advantage of self - organize map (SOM) characteristics such as to improve data management and accuracy of neural networks in the students' face verification. CK+ database was made compatible to fit the SOM design, a set of 100 image data, 10 different face images with 10 different facial expressions for the training database was loaded to MATLAB 2016a after preprocessing for this data. Created SOM neural network; as shown in figure 5; was trained for 1000 epochs (optimal epochs for training) as shown in Table 1.

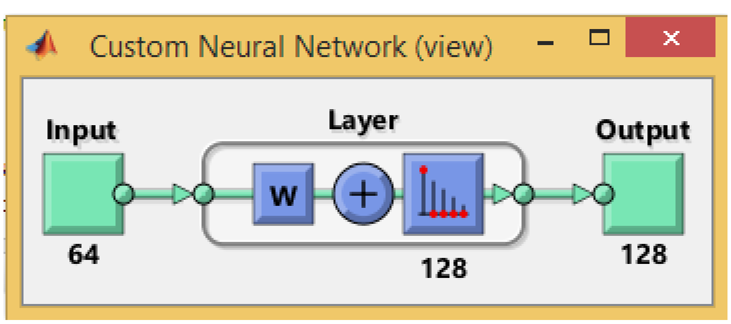


Fig. 5: Architecture of SOM neural network

Table 1: Determination the optimal No. of epochs.

No.	No. of epochs	Training time (sec)	Program execution time (sec)	Performance (Mean Squared Error)
1	100	2.0285	2.4882	0.0148
2	200	4.3395	4.8421	0.0148
3	300	5.6599	6.5182	0.0156
4	400	7.5263	8.0321	0.0148
5	500	10.2319	10.9896	0.0148
6	600	10.8802	11.2776	0.0156
7	700	13.2535	14.0098	0.0148
8	800	14.7450	15.2316	0.0148
9	900	17.9612	18.7353	0.0148
10	1000	17.9075	18.3983	0.0141
11	1100	22.0253	22.6074	0.0141
12	1200	23.9163	24.4563	0.0141
13	1300	25.6334	26.0868	0.0141
14	1400	26.6135	27.0838	0.0141
15	1500	28.9296	29.4181	0.0141
16	1600	31.0422	31.5464	0.0141
17	1700	33.5961	34.0843	0.0141
18	1800	34.1210	34.6419	0.0141
19	1900	35.6482	36.1272	0.0141
20	2000	38.2806	38.7659	0.0141

For instance, the weight of SOM layers for 20 students' face images in the training database is shown in figure 6. While, figure 7 shows the SOM weight vector.

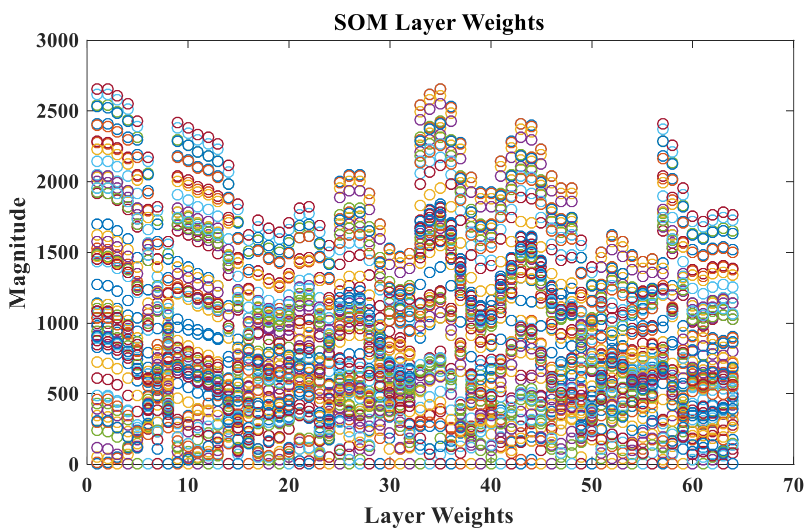


Fig. 6: The weights of SOM layer for 20 face images.

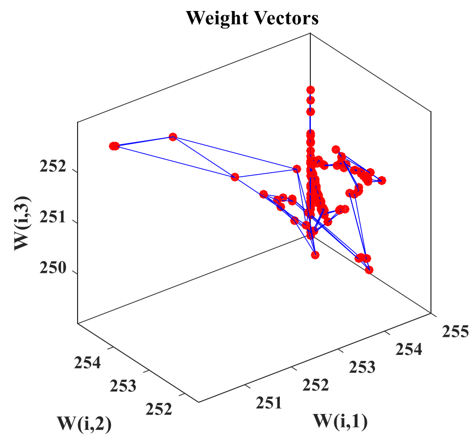


Fig. 7: The SOM weight vector.

The SOM neural network was validated by training the network on students' faces and determining the optimal Number of Epochs as shown in Table 1. Then, the SOM was tested using an untrained image and verified the result, as shown in Figure 8.

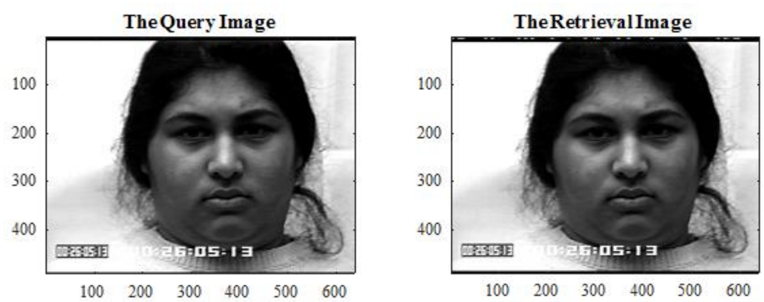


Fig. 8: The trained SOM testing.

4.3. Students State Discrimination:

To discriminate the student's state during the exam, the computer vision system algorithms [20] were used, which is to detect the key-points of interest regions in the face of the student such as eye, eyebrows, nose, and mouth. Thus, for each area in the student's face, their key-points data are described as a vector. To classifier of the students' face images, a Bag-of-Words (BoW) model [21] was used which treating image features as predefined words. This is done in three main steps are Feature detection [22], Feature description [23], and Codebook generation [24].The codebook is defined as the centers of the learned clusters, that clustered by k-means [25], of all vectors that represent all interest region of the face. The size of the codebook the same size as the word dictionary. Thus, through the clustering process, each interest region in the face image of a student is mapped to a certain code-word, then the face image is represented by a code-word histogram as shown in figure 9.

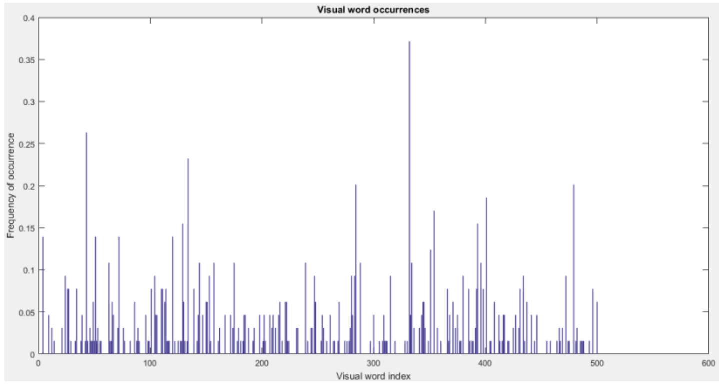


Fig. 9: The Code word histogram.

5. Results and Discussion:

During the examination process, the student's face is recorded as a live image. A Kanade-Lucas-Tomasi (KLT) algorithm [26] was used to track the movement of a student's face. A BoW model was used to classify the student's facial images according to the categories that the proposed model was trained like Normal, Happy, Confuse, Answer, Quite, and Cheat as shown in figure 10.

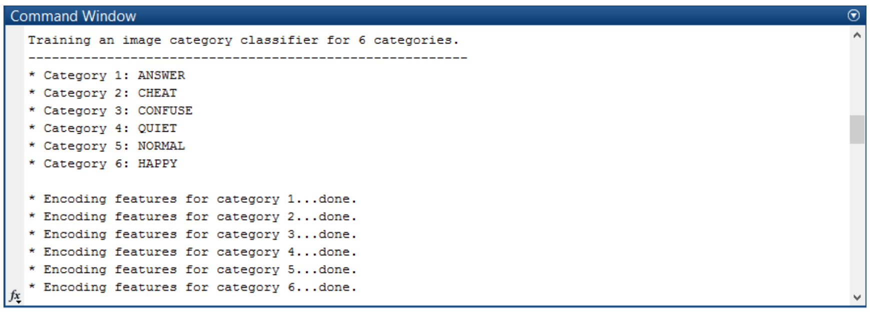
Fig. 10: States of students images.

The local image features are considered the better choice to extract the students' faces image features around the key points in the face. The Speeded-Up Robust Features (SURF) is used to detect and extract the feature points as shown in figure 11.

Fig. 11: Encoding Result for Student Face Image.

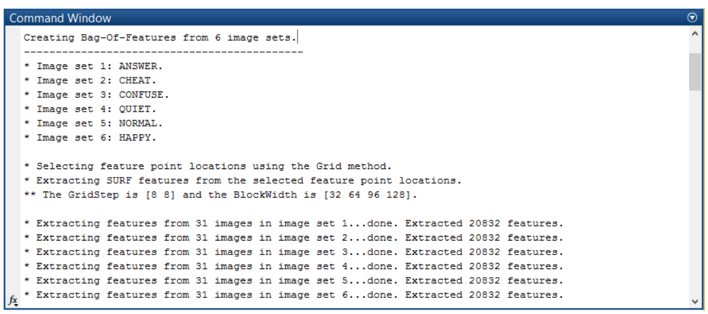
So, Student Face Image features are detected by finding the key points (KPs) that contain meaningful and semantical structures. To achieve this aim, the difference of Gaussian [27] is used to comparing each location in the image under deferent scales. Then, for each KP the scale-invariant descriptor is constructed.

The classifier is trained based on the bag of features (BoF) as shown figure 12 that represented the histogram of visual word occurrences in the face image as shown figure 9. Bag of features created from image categories (No. of Categories in experimental is 6) is shown in figure 13. The predictive model is generated by feeding each category of encoded training images into the classifier training process as shown in figure 14.


```
Command Window
Training an image category classifier for 6 categories.
--------------------------------------------------------
* Category 1: ANSWER
* Category 2: CHEAT
* Category 3: CONFUSE
* Category 4: QUIET
* Category 5: NORMAL
* Category 6: HAPPY

* Encoding features for category 1...done.
* Encoding features for category 2...done.
* Encoding features for category 3...done.
* Encoding features for category 4...done.
* Encoding features for category 5...done.
* Encoding features for category 6...done.
```

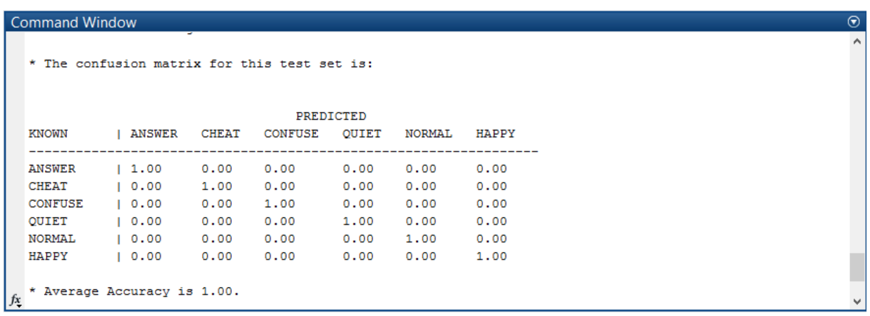
Fig. 12: The training classifier based BoF.


```
Command Window
Creating Bag-Of-Features from 6 image sets.
--------------------------------------------
* Image set 1: ANSWER.
* Image set 2: CHEAT.
* Image set 3: CONFUSE.
* Image set 4: QUIET.
* Image set 5: NORMAL.
* Image set 6: HAPPY.

* Selecting feature point locations using the Grid method.
* Extracting SURF features from the selected feature point locations.
** The GridStep is [8 8] and the BlockWidth is [32 64 96 128].

* Extracting features from 31 images in image set 1...done. Extracted 20832 features.
* Extracting features from 31 images in image set 2...done. Extracted 20832 features.
* Extracting features from 31 images in image set 3...done. Extracted 20832 features.
* Extracting features from 31 images in image set 4...done. Extracted 20832 features.
* Extracting features from 31 images in image set 5...done. Extracted 20832 features.
* Extracting features from 31 images in image set 6...done. Extracted 20832 features.
```

Fig. 13: Bag of features of image categories.


```
Command Window
* The confusion matrix for this test set is:


                                  PREDICTED
KNOWN      | ANSWER   CHEAT   CONFUSE   QUIET   NORMAL   HAPPY
-----------------------------------------------------------------
ANSWER     | 1.00     0.00    0.00      0.00    0.00     0.00
CHEAT      | 0.00     1.00    0.00      0.00    0.00     0.00
CONFUSE    | 0.00     0.00    1.00      0.00    0.00     0.00
QUIET      | 0.00     0.00    0.00      1.00    0.00     0.00
NORMAL     | 0.00     0.00    0.00      0.00    1.00     0.00
HAPPY      | 0.00     0.00    0.00      0.00    0.00     1.00

* Average Accuracy is 1.00.
```

Fig. 14: The predictive model.

To evaluate the proposed system was tested image taken during the electronic exam was processed and then entered on the system to determine the status of the student, the test has identified the status of the student who was in a state of cheat with an average accuracy of 100% as shown in Figure 15.

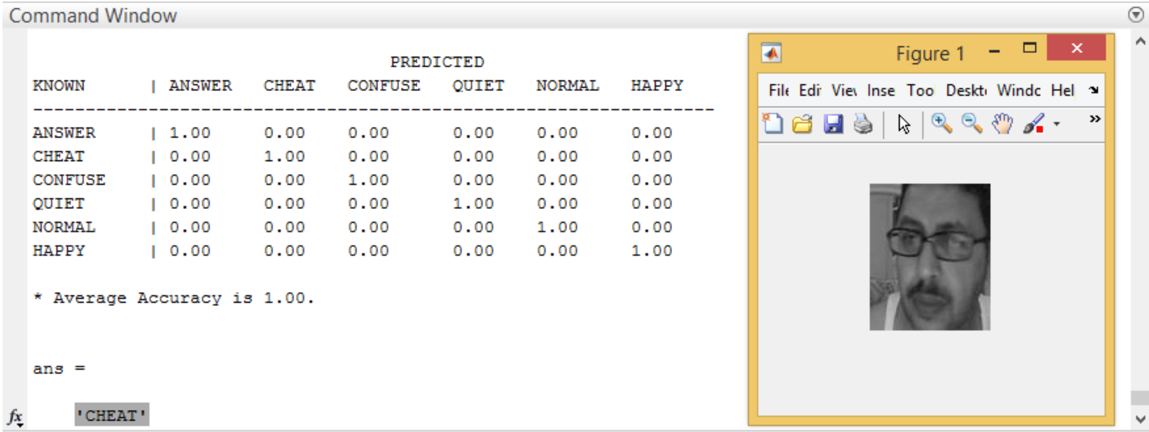


Fig. 15: Evaluated of the proposed method.

The summarization of the procedure that achieved to discriminate the status of students during the examination is shown in Table 2.

Table 2: summarization of proposed method.

Input Image	Output Image	Matching Feature	Decision
			Confused

6. Conclusion:

One of the most important problems facing student assessment in e-learning, which has become a common option for academic institutions, is the security problems related to e-learning systems. One of the biggest challenges that are considered a major challenge in the e-learning environment is the entry of unauthorized students to enter the electronic exams in addition to follow-up students during the exam process. Based on the advances in information technology systems and the introduction of machine learning techniques and artificial intelligence in the fields of education, methods and techniques have been integrated with e-learning to control and follow-up during exams.

One of the methods and techniques used in this paper is the SOM neural network because it will- suited pattern recognition problems for developing artificial intelligence systems like face recognition. SOM efficiency test performance is determined by the tolerance of the noise of face images. To obtain a high-speed SOM training the SOM efficiency was optimized the number of training epochs that rising the efficient face recognition system and handling with the tolerance of noise in input to the SOM before giving incorrect output as shown in Table 1. The test results showed that the SOM has been used successfully to obtain the optimal high-speed efficient face recognition system.

In this paper also adapted a technique used in retrieving documents and used it to distinguish the student's status during electronic exams. The Bag of Words model (BoWM) is used to treat the image as the document, therefore, the visual words can be defined as a vector of local image features vocabulary that occurrence counts. Speeded up Robust Features (SURF) is an algorithm based on the local descriptor approach. Where the interest points are defined as salient features from a scale-invariant representation in the input student face image. This aim is achieved by a multiple-scale analysis that consists of box filters that aggregate of the initial image with discrete kernels at several scales. Then the descriptors of orientation invariant are built by local gradient statistics.

Experiments have proven the ability of the proposed system to distinguish between cases whose features were extracted from the images taken for students during the electronic exam with an average accuracy of 100%.

References:

[1] Girolami, M., & Zhong, M. (2007). Data integration for classification problems employing Gaussian process priors. In *Advances in Neural Information Processing Systems* (pp. 465-472).

[2] Hansen, A. D., Sørensen, P. E., Hansen, L. H., & Bindner, H. W. (2001). Models for a stand-alone PV system.

[3] Institute of Medicine (US). Committee on the Robert Wood Johnson Foundation Initiative on the Future of Nursing. (2011). *The future of nursing: Leading change, advancing health*. Washington, DC: National Academies Press.

[4] Clarke, K. C., & Gaydos, L. J. (1998). Loose-coupling a cellular automaton model and GIS: long-term urban growth prediction for San Francisco and Washington/Baltimore. *International journal of geographical information science*, *12*(7), 699-714.

[5] Corchado, J. M., & Lees, B. (1998). Integration ai models. In *WORKSHOP ON KNOWLEDGE DISCOVERY AND DATA MINING. PML-NERC, PLYMOUTHLONDON, UK*.

[6] Karlik, B., & Olgac, A. V. (2011). Performance analysis of various activation functions in generalized MLP architectures of neural networks. *International Journal of Artificial Intelligence and Expert Systems*, *1*(4), 111-122.

[7] Basu, J. K., Bhattacharyya, D., & Kim, T. H. (2010). Use of artificial neural network in pattern recognition. *International journal of software engineering and its applications*, *4*(2).

[8] Wlodarczyk-Sielicka, M., & Stateczny, A. (2015, June). Selection of SOM parameters for the needs of clusterization of data obtained by interferometric methods. In *2015 16th International Radar Symposium (IRS)* (pp. 1129-1134). IEEE.

[9] Liu, Y., Li, H., Chen, D., & Niu, H. (2016). A Combined SOM/SVM Learning Algorithm for Vibration Recognition in Nanometer Imaging Systems. *International Journal of Simulation--Systems, Science & Technology*, *17*(36).

[10] Schneider, P., Müller, A. T., Gabernet, G., Button, A. L., Posselt, G., Wessler, S. ... & Schneider, G. (2017). Hybrid network model for “deep learning” of chemical data: application to antimicrobial peptides. *Molecular informatics*, *36*(1-2), 1600011.

[11] Hertz, J. A. (2018). *Introduction to the theory of neural computation*. CRC Press.

[12] El-Din, D. M. (2016). Enhancement bag-of-words model for solving the challenges of sentiment analysis. *International Journal of Advanced Computer Science and Applications*, *7*(1).

[13] Boulkenafet, Z., Komulainen, J., & Hadid, A. (2016). Face antispoofing using speeded-up robust features and fisher vector encoding. *IEEE Signal Processing Letters*, *24*(2), 141-145.
[14] Oyallon, E., & Rabin, J. (2015). An analysis of the SURF method. *Image Processing On Line*, *5*, 176-218.
[15] Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2018). Numerical Gaussian processes for time-dependent and nonlinear partial differential equations. *SIAM Journal on Scientific Computing*, *40*(1), A172-A198.
[16] Tareen, S. A. K., & Saleem, Z. (2018, March). A comparative analysis of sift, surf, kaze, akaze, orb, and brisk. In *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)* (pp. 1-10). IEEE.
[17] Yang, Z., Shen, L. Y., Yuan, C. M., & Gao, X. S. (2015). Curve fitting and optimal interpolation for CNC machining under confined error using quadratic B-splines. *Computer-Aided Design*, *66*, 62-72.
[18] Chen, G. H., & Shah, D. (2018). Explaining the success of nearest neighbor methods in prediction. *Foundations and Trends® in Machine Learning*, *10*(5-6), 337-588.
[19] Lucey, P., Cohn, J. F., Kanade, T., Saragih, J., Ambadar, Z., & Matthews, I. (2010). The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. Proceedings of the Third International Workshop on CVPR for Human Communicative Behavior Analysis (CVPR4HB 2010), San Francisco, USA, 94-101.
[20] Steger, C., Ulrich, M., & Wiedemann, C. (2018). *Machine vision algorithms and applications*. John Wiley & Sons.
[21] Gabryel, M. (2018). The Bag-of-Words Method with Different Types of Image Features and Dictionary Analysis. *J. UCS*, *24*(4), 357-371.
[22] Chu, W., & Cai, D. (2018). Deep feature based contextual model for object detection. *Neurocomputing*, *275*, 1035-1042.
[23] Lee, M. H., Cho, M., & Park, I. K. (2015). Feature description using local neighborhoods. *Pattern Recognition Letters*, *68*, 76-82.
[24] Besbes, B., Rogozan, A., Rus, A. M., Bensrhair, A., & Broggi, A. (2015). Pedestrian detection in far-infrared daytime images using a hierarchical codebook of SURF. *Sensors*, *15*(4), 8570-8594.
[25] Agrawal, S., & Agrawal, J. (2015). Survey on anomaly detection using data mining techniques. *Procedia Computer Science*, *60*, 708-713.
[26] Mstafa, R. J., & Elleithy, K. M. (2016). A video steganography algorithm based on Kanade-Lucas-Tomasi tracking algorithm and error correcting codes. *Multimedia Tools and Applications*, *75*(17), 10311-10333.
[27] Nychka, D., Bandyopadhyay, S., Hammerling, D., Lindgren, F., & Sain, S. (2015). A multiresolution Gaussian process model for the analysis of large spatial datasets. *Journal of Computational and Graphical Statistics*, *24*(2), 579-599.