Facial Images Quality Assessment based on ISO/ICAO Standard Compliance Estimation by HMAX Model

Article · May 2020

3 authors, including:

Azamossadat Nourbakhsh
Islamic Azad University-Lahijan Branch

Mohammad-Shahram Moin
Iran Telecommunication Research Center

Facial Images Quality Assessment based on ISO/ICAO Standard Compliance Estimation by HMAX Model

Azamossadat Nourbakhsh
Science and Research Branch, Islamic Azad University, Tehran, Iran
a.nourbakhsh@srbiau.ac.ir
Mohammad-Shahram Moin*
IT Research Faculty, ICT Research Institute, Tehran, Iran
moin@itrc.ac.ir
Arash Sharifi
Science and Research Branch, Islamic Azad University, Tehran, Iran
a.sharifi@srbiau.ac.ir

Received: 20/Aug/2019 Revised: 24/Oct/2019 Accepted: 20/Nov/2019:

Abstract
Facial images are the most popular biometrics in automated identification systems. Different methods have been introduced to evaluate the quality of these images. FICV is a common benchmark to evaluate facial images quality using ISO / ICAO compliancy assessment algorithms. In this work, a new model has been introduced based on brain functionality for Facial Image Quality Assessment, using Face Image ISO Compliance Verification (FICV) benchmark. We have used the Hierarchical Max-pooling (HMAX) model for brain functionality simulation and evaluated its performance. Based on the accuracy of compliancy verification, Equal Error Rate of ICAO requirements, has been classified and from those with higher error rate in the past researches, nine ICAO requirements have been used to assess the compliancy of the face images quality to the standard. To evaluate the quality of facial images, first, image patches were generated for key and non-key face components by using Viola-Jones algorithm. For simulating the brain function, HMAX method has been applied to these patches. In the HMAX model, a multi-resolution spatial pooling has been used, which encodes local and public spatial information for generating image discriminative signatures. In the proposed model, the way of storing and fetching information is similar to the function of the brain. For training and testing the model, AR and PUT databases were used. The results has been evaluated by FICV assessment factors, showing lower Equal Error Rate and rejection rate, compared to the existing methods.

Keywords: Facial Images Quality; ISO/IEC19794 Standard; ICAO; FICV; HMAX Model

1- Introduction

Facial recognition quality assessment is one of the most important factors in the automatic face recognition accuracy. Face recognition has been announced by the International Civil Aviation Organization (ICAO), as a biometric feature of machine-verified. The International Institute for Standardization (ISO) has proposed ISO / IEC19794-5, which includes face image information requirements, environmental conditions and shooting features [1]. Since there are many testing requirements, it is difficult to determine the compliancy of a face image with ISO / ICAO standards. Fully automating of a face image compliancy detection with ISO / ICAO standards has many benefits such as no need to human experts and accelerating the document production process. Researches about the quality assessment of commercial systems have shown that their performance needs to be improve for standard compliancy verifying and has not still reached the human accuracy level [2], [3]. Therefore, automated facial images quality assessment, is one of the most challenging issues in automated document production process. To evaluate the quality of produced algorithms, and comparing their performance, the FICV's Biolab benchmark is provided by the University of Bologna Biometrics Research Group. The Face Image ISO Compliance Verification (FICV) test, which includes assessing the requirements introduced in the face recognition standard, performs face evaluation and recognition. This benchmark includes a ground truth database, a well-defined testing protocol, and baseline algorithms for all ISO / ICAO requirements.
Some of the 24 requirements (looking away, unnatural skin tone, hair across eyes, head rotation (roll/pitch/yaw

Greater than 8◦), red eyes, shadow across face, frame too heavy, frame across eyes and mouth open) were used as quantitative variables of the problem. The compliancy of each of these requirements in the image is returned with a 1, 0, 1 score by the proposed model which shows three logically compliance, noncompliance, and dummy modes [4].

The Hierarchical Maximum (HMAX) pooling model is used to encode the properties of the noticeable face components. HMAX acts like a MAX operator and extracts location and scale-independent features for detection. This model expresses a hierarchy of brain regions through which object recognition is performed in the cerebral cortex. The purpose of this model is the cognitive phenomena describing in terms of simple and complex computational processes in an acceptable physiological model. To perform computational processes, two layers are embedded in this model:

• Simple "S" layers are derived from local filters convolution to compute higher order features from the different types of units in the previous layer.

• Complex "C" layers are stabilized by fetching units and the number of units is reduced by sub-sampling and all position and scale information are deleted simultaneously.

Object detection in the cerebral cortex extends through the feedforward ventral visual pathway. It travels through the primary visual cortex (V1) and reaches the InferoTemporal cortex (IT) by passing other visual areas V2 and V4. Different layers of HMAX model used for feedforward brain ventral visual pathway simulation in this research, are shown in Table 2 (which are driven from [5]).

Table 2: different HMAX layers for brain Visual pathway simulating (driven from [5])

Layers of the HMAX model	Ventral Visual Pathway of the Cerebral Cortex
S_1 (The first Simple layer)	V1 (The primary Visual Cortex)
C_1 (The first Complicated layer)	V2 (The second Visual cortex)
S_2 (The Second Simple layer)	V4 (The fourth Visual cortex)
C_2 ((The Second Complicated layer)	IT (The InferoTemporal cortex)

In this work, we have proposed, for the first time, a new integrated system for ISO/ICAO face image compliancy, which is based on HMAX method, inspired from the human brain functionality. The main contribution of our work is using an integrated method for different requirement compliancy assessment, compared to the existing methods, which use their specific features for each requirement compliancy assessment, separately. We have also approved the superiority of our approach over the existing methods, using the experimental results.

2- Related Works

Many researches have performed on facial images quality assessment based on ISO / ICAO standard requirements [6-12]. From the results of these studies, it can be concluded that this area of research still needs further development.

The Hierarchical Max-pooling model (HMAX) is a feedforward model mimicking the structures and functions of V1 to posterior inferotemporal (PIT) layer of the primate visual cortex, which could generate a series of position and scale- invariant features.

In [13] to mimic the attention modulation mechanism of V1 layer, a bottom-up saliency map is computed in S1 layer of the HMAX model, which can support the initial feature extraction for memory processing. Also to mimic the short-term memory to long-term memory conversion abilities of V2 and IT, an unsupervised iterative clustering method is used for clusters learning with multiscale middle level patches. Simulation results show that the enhanced model with a smaller memory size, exhibits higher accuracy than the original HMAX model and other unsupervised feature learning methods in multiclass categorization task.

An object recognition model by extracting features temporally and utilizing an accumulation to bound decision-making model is introduced in [14]. This model accounted recognition time and accuracy. In face recognition, for temporally extracting informative features, a hierarchical spiking neural network, called spiking HMAX is modified. In the decision making part of the model, the extracted information accumulates over time using accumulator units. The input category is determined if any of the accumulators reaches a threshold, called decision bound. Testing Results showed that the model follows human accuracy in a psychophysics task better than the classic spiking HMAX model.

For Image classification, a method based on ontology and HMAX features performed by integrating clusters [15]. This method relied on training visual-feature classifiers according to the taxonomic relationships between image categories. Using the HMAX model, the visual features and the concepts were extracted from the image categories. The taxonomic relationship between visual features and concepts were created to make an ontology that represents the semantic information associated with the training images. Using ontology-based HMAX and Bag-of-Visual-Words (BoVW) models, superior performance achieved over baseline methods. To evaluate this method, the Inception-v3 deep learning network was used, and the classification results performed better in some image classes than Inception-v3.

Bottom-up attention is crucial to primary vision and helps reducing computational complexity. In [16], a bottom-up attention model was presented based on the C1 features of

HMAX model. Attention modeling in layer C1 of the HMAX model showed better results than Graph-Based Visual Saliency (GBVS).

In [17], a face recognition model was presented that used the visual attention model using skin color features to find saliency maps of the face candidate areas and the C2 texture features in the visual cortex of the HMAX model for face recognition. After finding candidate face areas, C2 texture features were extracted for face or non-face areas classification using a support vector machine classifier. Experimental results on the Caltech Face Database with background, showed that the proposed model was reliable against variations in face brightness, expression and cluttered backgrounds.

In [18] Binary HMAX model (B-HMAX) was introduced. In the C1 layer of this model, using image patches selection instead of random usage in the standard model at the training phase increased the accuracy and decreased the calculating costs. Also using Hamming distance instead of Euclidean distance for calculating the distance between patches, increased the speed.

In [19], the original HMAX model [5] was used and the end-of-network filters, which integrated local filters, were modified for producing complex filters to cover larger and more complex areas of the image. To better discriminate the image content, they trained the coefficients of each filter in the last layer. This increased the discrimination and also the invariance.

A flexible multilayer radial method for the outputs' pooling of the filters in the image was presented. Neurons in the inferior temporal visual cortex (IT) are known for regions by varying sizes accepting [20]. This is called local areas of various sizes pooling in the visual field, causing slight variability with respect to spatial location. The multi-resolution pooling introduced in the study was equivalent to applying a specific filter in a spatial neighborhood with different radial pooling that caused different levels of invariance. The optimal level of invariance with a single classifier was obtained by training at higher levels of the network [21]. The classification in this research showed better results than previous architectures. Since this method achieved very good results with increasing discrimination and invariance, it has been used in the present study.

As it is mentioned above, there are many researches that payed attention to the face recognition subject using HMAX model. However, none of them has worked on facial image quality assessment by this model. Thus, in this study, we evaluated the suitability and effectiveness of using HMAX model for the facial image quality assessment.

In the following sections, first, ICAO requirements selection process has been descried and after creating key and non-key patches from face components by Viola-Jones algorithm, HMAX model is introduced and employed in FICV process. The results of executing the proposed model on AR and PUT databases have been evaluated using standard Facial image quality verification factors.

3- Requirements Selection

Considering the scope and content of the assessment factors operations; which are introduced in ISO / IEC19794-5, first, some requirements from 24 FICV requirements should be selected. According to Ferrara et al. [9], the error rate obtained for each requirement has been divided into three categories. Table 1 can be used for identifying the need for further research on requirements and selecting new research areas for the requirements assessment results improvement.

Table 1: Biolab's ICAO Requirements' classification based on their Accuracy Rates (Driven from [9])

ICAO Requirements Difficulty Diagnosing	Accuracy Rate	Name of Requirement
Easy Diagnosing Requirements	$EER < 3\%$	ICAO08(pixelation),ICAO10 (Eye Closed),ICAO13(Flash Reflection on Skin), ICAO15 (shadow behind Head), ICAO17(Dark Tinted lenses) , ICAO18(Flash Reflection on Lenses), ICAO22(Veil over Face)
Middle rate Diagnosing Requirements	$3\% \leq EER \leq 7\%$	ICAO02(Blurred) , ICAO04 (Ink Marked/Creased), ICAO05 (Unnatural Skin Tone), ICAO06 (Too Dark/Light), ICAO11 (Varied Background) , ICAO14(Red Eyes) , ICAO19 (Frames too Heavy) , ICAO20(Frame Covering Eyes) , ICAO23(Mouth Open)
Hard Diagnosing Requirements	$EER > 7\%$	ICAO01(Eye Location), ICAO03 (Looking Away) , ICAO07(Washed Out) , ICAO09(Hair Across Eyes) , ICAO12 (roll/pitch/yaw Greater than 8◦),ICAO16 (Shadow Across Face), ICAO21 (Hat/CAP), ICAO24 (Presence of other Faces or Toys too Close to Face)

Requirements selection in this study, are based on the results of Table 1. Requirements with less than 3% error rates, which are not challenging, are ignored. Image and background features have also been excluded, and we have focused on nine following facial requirements:

Look Away (ICAO03), Unnatural Skin Tone (ICAO05), Hair Across Eyes (ICAO09), Head Rotation (roll/pitch/yaw Greater than 8◦) (ICAO12),Red Eyes (ICAO14), Shadow Across Face (ICAO16), Frame Too Heavy (ICAO19), Frame Across Eyes (ICAO20), Mouth Open(ICAO23).

4- Proposed HMAX model for Face Image Quality Assessment

The proposed model for face image quality assessment using the HMAX method is shown in Figure 1. In the following, its different parts has been described:

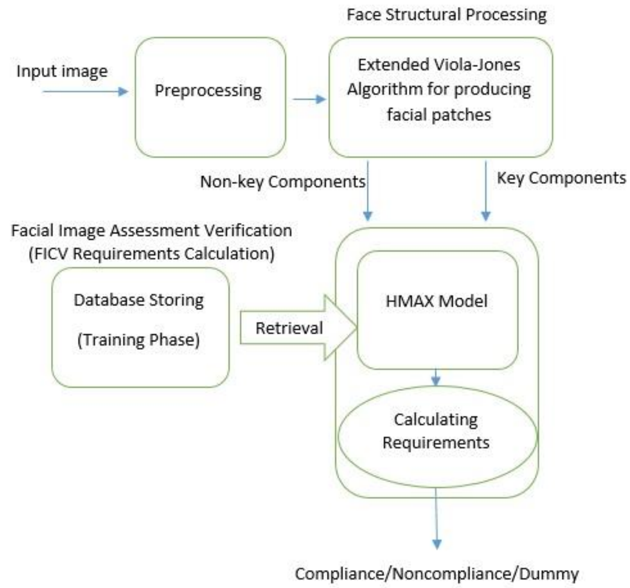


Fig. 1 Proposed method for Facial Image Quality Assessment using HMAX Model

4-1- Pre-Processing

In this section, basic image processing is performed to produce a suitable image. First the image is examined for possibility of being tokenizable (without padding). So the database images should be corrected as follows:

- Distance between eyes (E_{Dist}) be at least 60 pixels.
- Rectangular area be W * H (with W = 4. E_{Dist} and H = W.4 / 3) size. Eyes aligned horizontally, centered at C_E = (W.1 / 2, W.3 / 5) which is generally enclosed in the original image (Figure 2).

Basic tasks of this part are face detection (ROI), face alignment and normalization (which can be used for reducing illumination effects).

4-2- Face Structural Processing

In this part, face semantic patches of the key and non-key components are obtained using elemental images based on a component-based approach.
Based on biological evidences, the diagnosing operation is performed in two operation of location estimation and semantic division. In the estimation section, the center of four face key components locations (left eye, right eye, nose, and mouth) are estimated. To implement this section, a hybrid method using Viola-Jones and skin color pixel detection is used; that causes more accurate detection of the facial components location and increases detection speed [22]. Different semantic patches are formed by segmentation based on the location estimation results and by component-based method. In this way, the primary image is divided into 8 patches, 4 of which consist of key components of the face (left eye, right eye, nose, and mouth) and 4 patches containing non-key facial components (left cheek, right cheek, forehead and chin). For being almost every patch component-based, the size of divisions must be specific to each individual. Hence, the size of each patch is obtained based on a constant rate of the distance between the two eyes centers.

Fig. 2 Geometric properties of the obtained image format [1]

4-2-1- Producing Facial Patches using Viola-Jones algorithm

Michael Jones and Paul Viola [23] developed the famous Viola-Jones Algorithm for face detection in 2003. In this algorithm, learning is performed by measuring the similarity of two sample images. A set of computationally efficient rectangular features (Haar features), are described and operate on a pair of input images. The features compare inside the input images areas in different locations, scales and directions. To quickly evaluating these features from integral images and performing the training of facial similarity by features, the Adaboost algorithm is used. Finally, a hierarchical classifier is considered for rejecting windows that have not been recognized as a face. As this method is very accurate but time consuming, a very fast detection algorithm based on skin tone pixel detection has been merged to it in [22]. In the case of eyes and mouth detection, physical location approximation is made in detected face to locate the eyes and mouth. This method increased the accuracy of system and decreased its consumed time.

Since the introducing of the Viola-Jones algorithm, many researchers have used this method in their researches, which [24], [25] and [26] are amongst them.

4-2-2- Feature Selection

Selecting Features for Multiclass Classification is an essential step in pattern recognition and machine learning applications. Specially, a big challenge is an optimal subset selecting from high-dimensional data, which has much more variables than observed and contains noise or outliers. We used the feature selection method presented in [27]. In that research, a feature selector named Fisher-Markov is presented to identify the features; which are more important in describing the essential differences around possible groups.

It is a systematic method of factors optimizing for the best feature subset selecting, to identify factors for sparsity and separability in the high dimensional scenarios. Since the introduced method is linear in number of features and quadratic in number of observations, it operates very quickly. In pattern recognition and model selection view, in the proposed model, it is easily possible to select the most discriminable subset of variables by solving an objective function without constraint.

In supervised classification, with the training data $\{(x_k, y_k)\}_{k=1}^{n}$, where $x_k \in R^p$ are the p dimensional feature vectors and $y_k \in \{w_1, \ldots, w_m\}$ are classes' tags, the most important features should be selected for the most separable representation of the multi-class classification with m Ci class, where $i = 1,2, \ldots, m$. Each Ci class has n_i observation. With a new test observation, the selected features are used for predicting an unknown class tag for each observation. In order to global optimization and effective feature selection by Fisher-Markov method, in the feature subset, for large p, some specific kernels including polynomial kernels k have been considered [28] [29].

$$k(x_1, x_2) = (1 + \langle x_1, x_2 \rangle)^d \quad (1)$$

where d is the parameter degree, alternatively:

$$k'(x_1, x_2) = (\langle x_1, x_2 \rangle)^d \quad (2)$$

4-3- Face Image Assessment Confirmation Module (FICV Attribute Detection)

Inspired by biological evidences, facial image compliancy verification of the proposed model, for simulating memory structure such as brain functionality, includes code generation, storage, retrieval and final decision. The inputs of this section are the face key and non-key components patches (and thus it is a component-based model) and the output of this section is the result of ICAO requirements recognition.

4-3-1- HMAX Model

As illustrated in [5] and shown in Figure 3, a general HMAX model is designed of frequency of pooling and convolution layers. Each convolutional layer has a series of feature maps and each pooling phase produces changing resistance against these feature maps. In the following, different layers of HMAX model are described. S_1, C_1, S_2 and C_2 layers of HMAX model are named L1, L2, L3 and L4, respectively.

- First layer

Each feature map $L1_{б,\Theta}$ is produced by the input image convolution against a set of Gabor filters $g_{б,\Theta}$ (Eq. 3), with orientation Θ and scale $б$. These filters are used to simple cell activation in the V1 region of the visual cortex modeling [5].

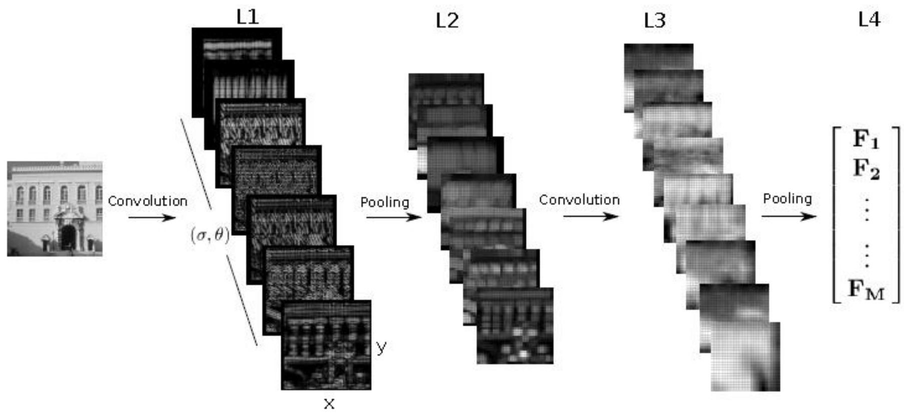


Fig. 3 A general convolution network: this network alternates feature mapping layers (convolution) and feature pooling layers alternately. The convolution layers produce information of a particular feature, and the pooling layers create invariance by relaxing the configuration of these features [5].

$$g_{\sigma,\theta}(x,y) = \exp\left(\frac{x_0^2+\gamma y_0^2}{2\sigma^2}\right).\cos\left(\frac{2\pi}{\lambda}\,x_0\right) \qquad (3)$$

where $x_0 = x\,cos\theta + y\,sin\theta$ and $y_0 = -x\,sin\theta + y\,cos\theta$. Parameter γ shows the aspect ratio of the filter and λ is its wavelength.

With image *I*, the first layer in orientation Θ and the scale *б*, can be expressed as an absolute value convolution product, as follows:

$$L1_{\sigma,\theta} = \left| g_{\sigma,\theta} * I \right| \qquad (4)$$

- Second layer

Each feature map of $L2_{б,\Theta}$ is a dimension reduction of $L1_{б,\Theta}$, that is obtained by the maximum number of local neighborhoods selecting. Maximum pooling impact on local neighborhoods is the invariance of local conversions and global transformations [30].
The second layer divides each $L1_{б,\Theta}$ map into small neighborhoods $u_{i,j}$ and finds the maximum value inside each $u_{i,j}$ such that:

$$L2_{\sigma,\theta}(i,j) = max_{u_{i,j} \epsilon L1_{\sigma,\theta}} u_{i,j} \qquad (5)$$

By keeping only the maximum output at two scales adjacent to each point (i, j), scale invariance can be achieved to some extent.

- Third layer

The *L3* layer at the *б* scale is obtained by the α^m filters convolution against the $L2_{б,\Theta}$ layer, which are called HL filters.

$$L3_\sigma^m = \alpha^m * L2_\sigma \qquad (6)$$

HL filters are visual descriptors of mid-level regions in the image that combine low-level Gabor filters with multiple orientation in one scale.

- Fourth layer

To achieve general invariance, the final step (last signature) is calculated by the maximum $L3_б^m$ output selecting in all location conditions and scales. Thus the last layer is a vector of *M ~ 1000* dimension, which determines each coefficient of each HL filter maximum output on the scale *б* and location *(x, y)*.

$$L4 = \begin{bmatrix} max_{(x,y),\sigma} L3_\sigma^1(x,y) \\ \vdots \\ max_{(x,y),\sigma} L3_\sigma^M(x,y) \end{bmatrix} \qquad (7)$$

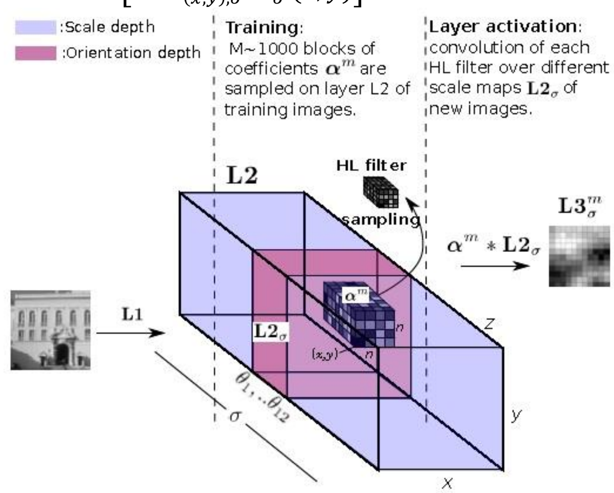


Fig. 4 Third level of HMAX: in training Phase, M~1000 HL filters is defined by L2 coefficients of sampling blocks. Layer activation in each image is obtained by convolving each HL filter on all positions of each $L2_\sigma$ scale map [5].

HMAX model has been used in several researches including [19], [20], [24] and [25]. HMAX method in the model proposed in this work is based on [19], where the first level local filters are merged with more sophisticated filters at the previous level, producing a flexible descriptor of the object regions and combining local information across multiple scales and orientation. These filters are invariant and discriminative, making them more suitable for visual classification. It also introduces a multi-resolution spatial pooling that encodes local and public spatial information for generating image discriminative signatures. In Figure 5, each HL filter is convolved simultaneously on several scales that focus on the scale *б*. In training phase, the coefficients associated with weak scales and orientations, receive zero values; which makes the filter more discriminative and ignores weaker scales and orientations during the test phase.

4-3-2- Code Generation

Real code creating in the human brain, means information sensing and receiving from the environment in the form of physical and chemical stimuli. Especially when looking at a particular face, the brain encodes various facial features with many patterns. It is assumed that there are 9 coding patterns (for each face, 9 ICAO requirements would be stored in long-term memory). To mimic this fact functionally, the HMAX descriptor can be used to encode these requirements.

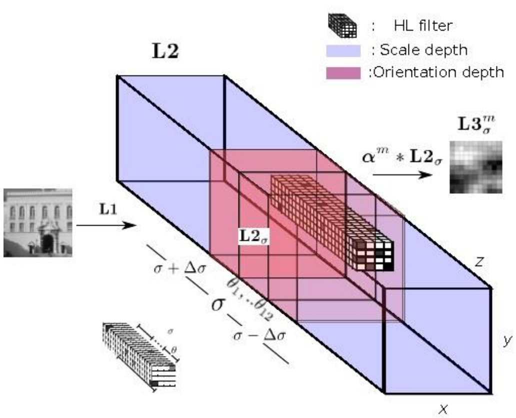


Fig. 5 Third level operations - Each HL filter measures simultaneously on several scales [19].

To prepare and extract the *C1* level from the *HMAX* model, split face patches $cpath_{ij}$ from known individuals can be used where $i = 1,2, ..., 8$ (for 8 Key and non-Key face components patches) And $j = 1,2,... .N$, where N represents the number of known people (known people refers to those whose ICAO attribute recognition tags exist in the database). *C1* patches of an ICAO requirement, produce a patch cluster C_i where $C_i \epsilon\ cpath_{ij}$. For each C_i, at the *C2* level of the HMAX model for each face image, the $f_i \epsilon R^N$ feature is extracted. f_i is the final consideration feature that introduces a face component for ICAO requirement recognition (Figure 6).

4-3-3- Storing

In the mammals' long-term memory, different features of a known object are regularly stored in distributed areas, and common features of various known objects are stored through aggregating. Labeled faces are database images for which ICAO requirements assessment are labeled. As shown in Figure 7, the same strategy was chosen to store the ICAO requirements of labeled faces. For example, the first studied ICAO requirement of different individuals f_{1j} $(j = 1, 2, ..., N)$ are stored together and constitute a subspace storage of an ICAO requirement. The first person's attributes f_{i1} (where i is one of the 9 case study requirements) are stored separately in the distributed subspace. The storing phase is similar to a training procedure in a general requirements estimation method, and does not include unlabeled faces.

4-3-4- Decision Making

This step identifies and assesses the ICAO requirements of a new face image from labeled faces images. To identify a person's requirements, it is needed to retrieve requirements for all labeled faces before decision making, which is called retrieval. R_{ij} is used for the requirement assessment of a new face image based on the j[th] labeled faces requirements by using a notable face feature f_i . R_{ij} can be estimated using a support vector machine. There are 9 Binary SVM classifier for assessment of each studied requirements. Final classification results shows 1 for compliance, 0 for non-compliance and -1 for dummy classes. For 1 and 0 classes, the compliancy of each requirement in the image is returned with a score in range of zero to 100 by regression.

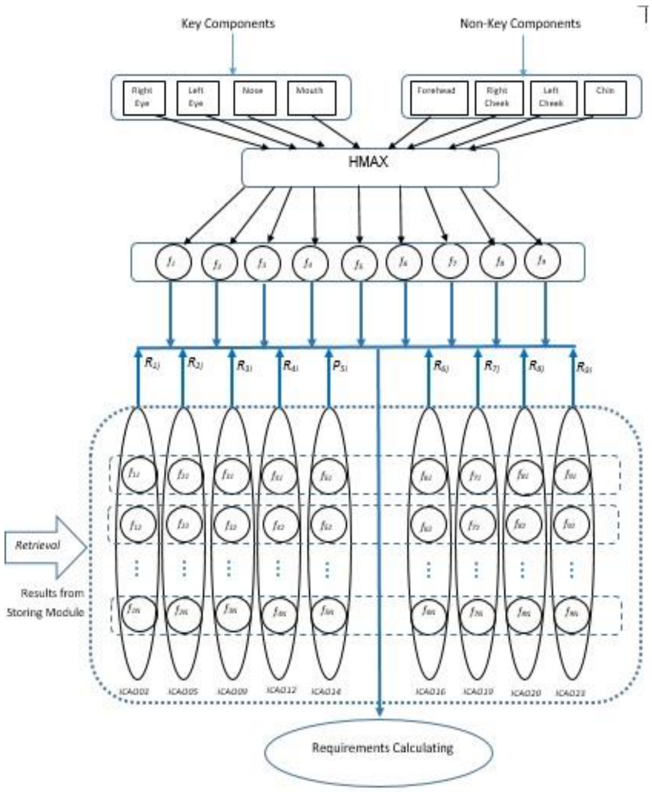


Fig. 6 Proposed Face Image Assessment Confirmation Module containing HMAX model for detecting compliancy with ICAO requirements.

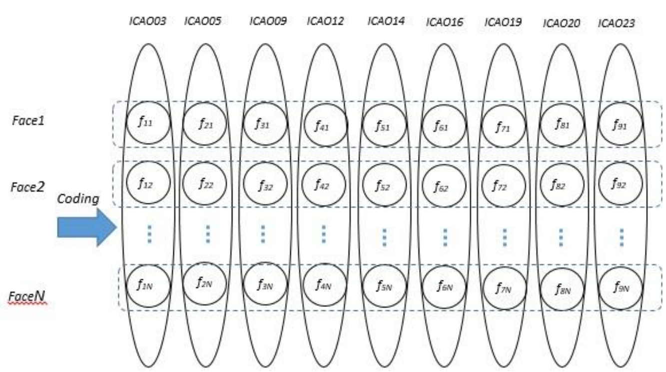


Fig. 7 Storage module structure.

5- Simulation Results

For simulating the proposed model, a system with 16GB of RAM, Core i7 processor, 2TB hard drive has been used.

In this study, 1741 images of the AR database with size of 576 * 768 [32] and 291 images of the PUT database with size of 1536 * 2048 [33] were used to train and test the proposed method, which include 310 fully compatible images (compatible with all requirement) and 1722 incompatible images (incomparable at least with one requirement). Database division to test and train set, has been done by K-fold cross-validation algorithm, where the best result was obtained with K=10. For objective performance evaluation in the database, ground-truth data has been employed. Some of these images are manually labeled, each image containing eye corners information and position and shooting features based on three logically compliance, noncompliance, and dummy modes. Dummy value is used for uncertainty situations (for example, when one uses sunglasses, it is almost impossible to detect open or closed eyes).

Two types of errors can occur during the compliancy assessment of face images:

1) Declaring compliancy for an image which is not compliant (False Match Rate) (FMR):

$$FMR = \frac{FP}{FP+FN} \tag{8}$$

2) Declaring incompliancy for an image which is compliant (False Non Match Rate) (FNMR).

$$FNMR = \frac{FN}{TP+TN} \tag{9}$$

A good biometric system should illustrate a small amount of FMR and FNMR. High FMR indicates high system error and low FNMR indicates low system functionality for all studied cases' acceptance. ROC, DET charts and EER are used to analyze a biometric system.

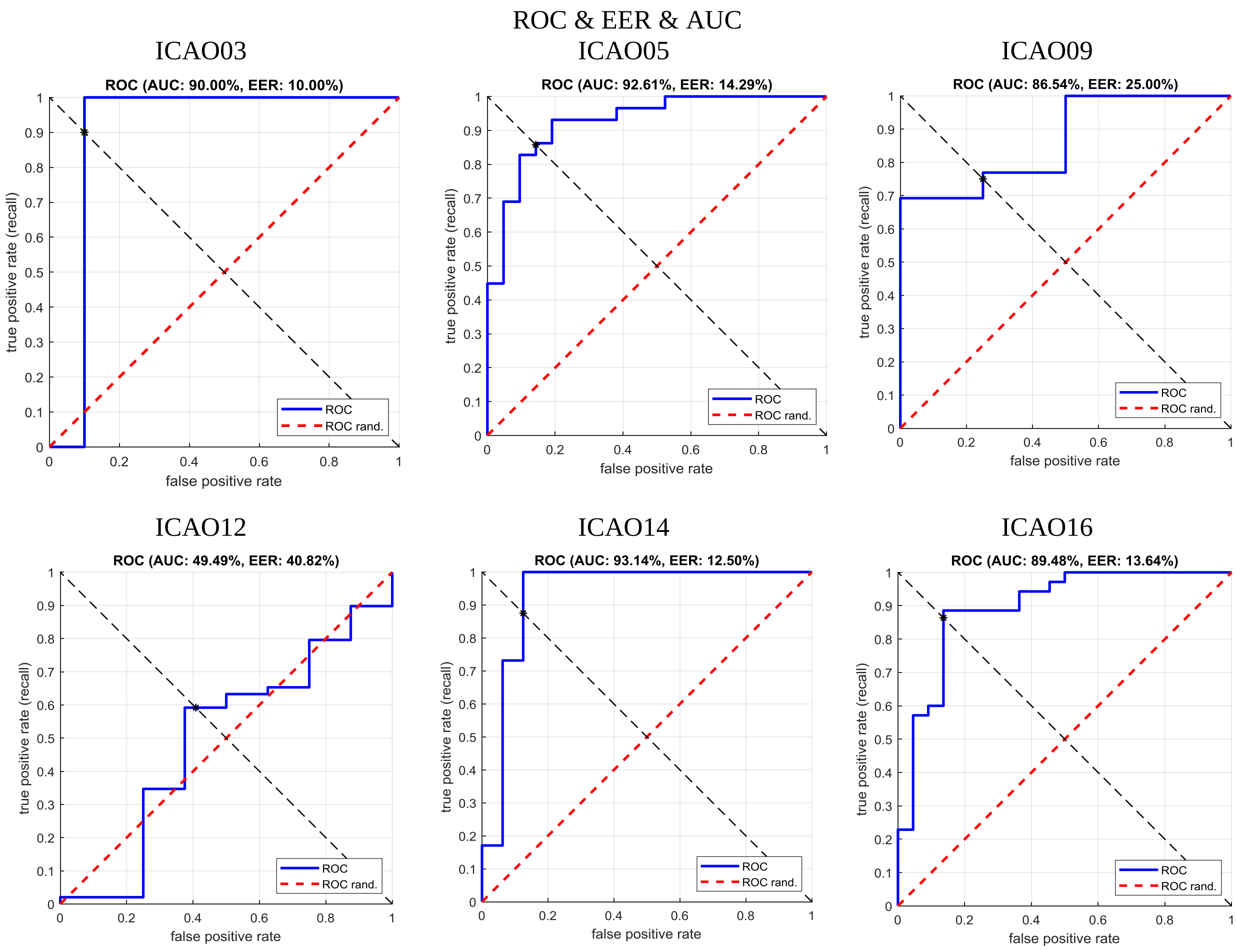
EER: Equal Error Rate (EER) is indicated by interaction between FMR and FNMR. EER represents the error rate at a *t* threshold such that the False match rate and the false non-match rate are equal (FMR(*t*) = FNMR(*t*)). EERs are calculated for compliancy rate checking and used for each feature performance evaluating.

Rej: Rejection Rate refers to the percentage of the face images, which cannot be processed by the proposed method. This rejection can be due to the pixellation, hair across eyes and shadow across the face. For comparing the results of simulation, three SDKs (two commercial SDKs and the BiolabSDK [9]) were used. The name of commercial SDKs cannot be disclosed, because of the specific agreement with their providers in [9]. Table 3 shows the EER and Rejection rates for comparing three SDKs results with the proposed method.

Table 3: EER and Rejection rates for the three SDKs and proposed method

Row	*Characteristic*	*SDK1*		*SDK2*		*BioLabSDK*		*HMAX Method*	
		EER	*Rej*	*EER*	*Rej*	*EER*	*Rej*	*EER*	*Rej*
1	ICAO03- Looking Away	*27.5%*	*7.1%*	-	-	*20.6%*	*0.0%*	*10.00%*	*0.16%*
2	ICAO05- Unnatural Skin Tone	*18.7%*	*4.8%*	*50.0%*	*0.8%*	*4.0%*	*0.2%*	*14.29%*	*0.3%*
3	ICAO09- Hair Across Eyes	*50.0%*	*81.9%*	-	-	*12.8%*	*0.0%*	*25.00%*	*0.0%*
4	ICAO12-Roll/Pitch/Yaw>8°	-	-	*26.0%*	*2.9%*	*12.7%*	*0.2%*	*40.82%*	*0.0%*
5	ICAO14- Red Eyes	*5.2%*	*4.5%*	*34.2%*	*0.0%*	*7.4%*	*0.0%*	*12.5%*	*0.2%*
6	ICAO16- Shadows Across Face	*36.4%*	*8.1%*	-	-	*13.1%*	*0.4%*	*13.64%*	*0.4%*
7	ICAO19- Frames Too Heavy	-	-	-	-	*5.8%*	*0.0%*	*0.0%*	*0.0%*
8	ICAO20– Frame Covering Eyes	*50.0%*	*62.3%*	-	-	*6.3%*	*0.0%*	*0.0%*	*0.1%*
9	ICAO23- Mouth Open	*3.3%*	*52.1%*	-	-	*6.2%*	*0.0%*	*10.71%*	*0.4%*
- *Shows that the SDK does not support the test for this Characteristic*									

ROC Curve: The System Performance Characteristic Curve or Receiver Operating Characteristic (ROC) is an objective evaluation method that is a two-dimensional diagram, where the x-axis corresponds to False Positive Rate or FMR and the y-axis corresponds to True Positive Rate or *1- FNMR* (often replaces by FNMR). The area under the ROC diagram represents the Area Under Curve (AUC). The high AUC value indicates higher accuracy of the model. For each of the ICAO requirements investigated in this research, the ROC diagrams are calculated and are shown in Figure 8.



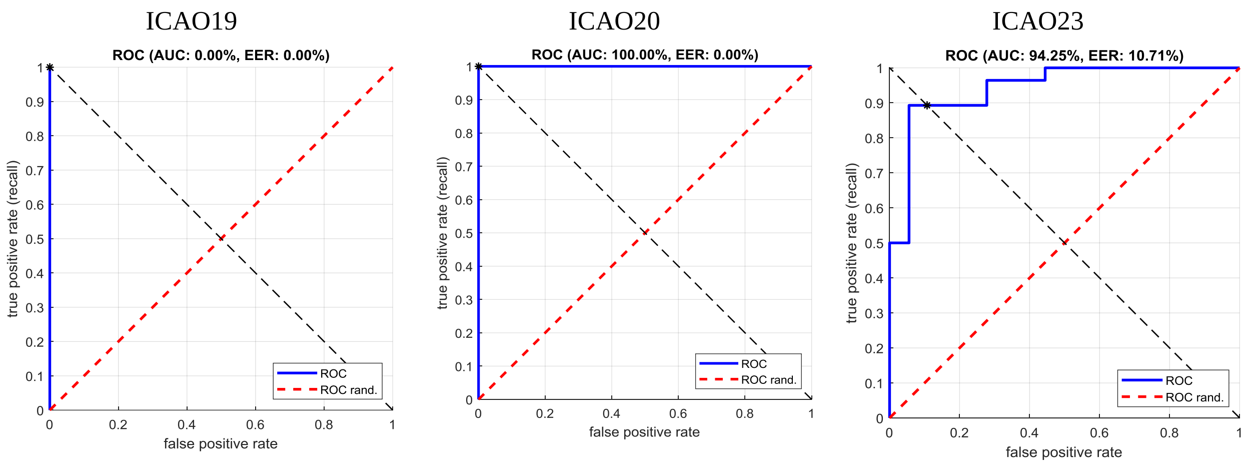


Fig. 8 ROC diagrams for each of the nine ICAO features investigated in this work.

In the ROC diagram, the greater the accuracy of the test, the closer the curve to the left boundary and then to the upper boundary of the ROC space. The closer the bend to the 45-degree diameter of the ROC space, the less accurate the test. The tangent slope at a cutting point shows the Likelihood Rate (LR) for the test.

Recall: A positive diagnosis probability being true when the actual results are positive. This parameter is also called the true positive rate.

$$Recall = TPR = \frac{TP}{TP+FN} \qquad (10)$$

Precision: A positive diagnosis probability being true when the experimental results are positive. This parameter is also called the false negative rate.

$$Precision = FNR = \frac{TP}{TP+FP} \qquad (11)$$

Average Precision: For each positive recall sample, the sum of precision values in positive recalls to all positive diagnoses, determines the average Precision parameter.

$$AP = \frac{\sum_{k=1}^{n}(P(k)\times rel(k))}{number\ of complient\ requirements} \qquad (12)$$

where *rel(k)* is an index that is equal to 1 if the requirement is compliant, otherwise it will be zero [34]. The average contains all the associated requirements and those associated requirements that have not been compliant, have a zero value for precision.

AP11: This is an index calculated by averaging the precision over a set of evenly spaced recall levels {0, 0.1, 0.2, ... 1.0}. This factor is used for reducing the impact of wiggles in the curve.

$$AP11 = {}^{1}/_{11} \sum_{r\in\{0,o.1,\ldots,1.0\}} p_{int}(r) \qquad (13)$$

where $p_{int}(r)$ is the interpolated precision, shows the maximum precision over all recalls greater than r (in 11 points) [35].

Figure 9 contains the accuracy-recall diagrams, the area under the curve (AUC), the average precision (AP) and AP11.

PRECISION & RECALL & AUC & AVERAGE PRECISION (AP)

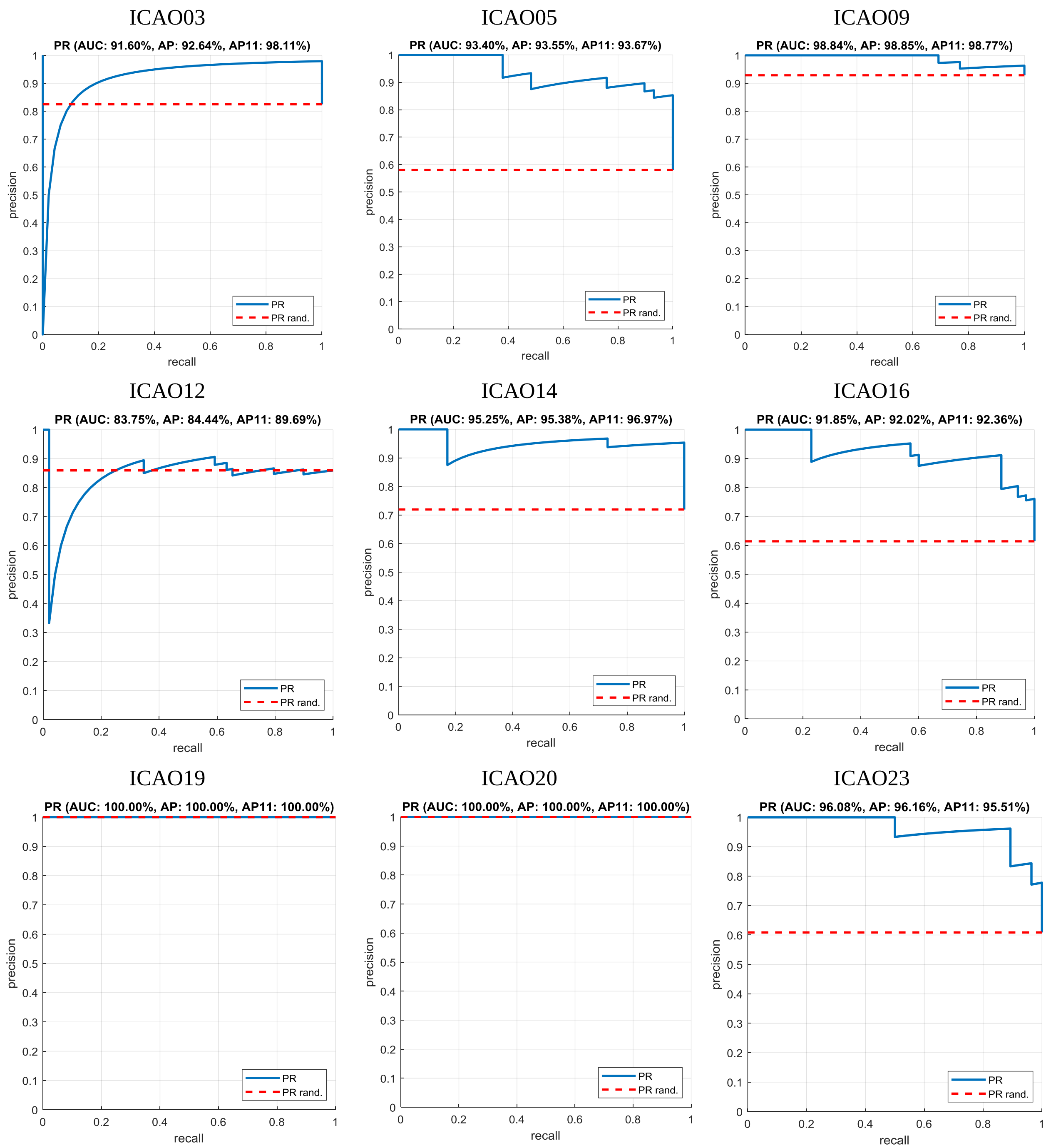


Fig.9 Accuracy-recall diagrams, Area under the Curve (AUC) and Average Precision (AP) for the selected ICAO requirements

6- Conclusions

One of the most important factors affecting the accuracy of automatic face recognition is the face images quality assessment. Accordingly, a benchmark was provided by the University of Bologna called BIOLAB-ICAO, which Facial image Compliancy part is called FICV. In this research we proposed a new approach for facial images quality assessment using HMAX model (as the perceptual brain modeling).

The way of information storing and fetching it for training, is like the way of storing information in the brain. Nine

ICAO requirements are used to assess quality. The AR and PUT databases were used to train and test the model. The assessment factors introduced in the FICV benchmark were used to evaluate the modeling results. The results showed improvement in the detection of some requirements, particularly Frame Too Heavy (ICAO19), Frame Across Eyes (ICAO20). So, it is recommended to use HMAX model for these requirements detecting in the SDKs. As a follow-up, a model based on brain decision-making paths approaches to assess the quality of facial images, can be suggested.

References

[1] ISO/IEC 19794-5, Information technology - Biometric data interchange formats - Part 5: Face image data, 2011.

[2] Y. Wong, Sh. Chen, S. Mau, C. Sanderson, B. C. Lovell ,"Patch-based Probabilistic Image Quality Assessment for Face Selection and Improved Video-based Face Recognition ", Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 ,IEEE.

[3] P. Ji, Y. Fang, Zh. Zhou, and J. Zhu, "Fusion of mSSIM and SVM for Reduced-Reference Facial Image Quality Assessment", (Eds.): CCBR 2012, LNCS 7701, pp. 75–82, Springer-Verlag Berlin Heidelberg.

[4] M. Ferrara, A. Franco, D. Maio, "BIOLAB-ICAO: A New Benchmark to Evaluate Applications Assessing Face Image Compliance to ISO/IEC 19794-5 Standard", 978-1-4244-5654-3/09/IEEE ICIP 2009, pp. 41-44.

[5] Th. Serre, L. Wolf, S. Bileschi, M. Riesenhuber, and T. Poggio, “Robust object recognition with cortex-like mechanisms,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 29, 2007, pp. 411–426.

[6] R. Hsu; M. Abdel-Mottaleb, A.K. Jain, "Face detection in color images," IEEE Transactions on Pattern Analysis and Machine Intelligence, May 2002, Vol.24, No.5, pp.696-706.

[7] C. Bouvier, A. Benoit, A. Caplier, P. Y. Coulon “Open Or Closed Mouth State Detection: Static Supervised Classification Based On Log-Polar Signature”, Springer Berlin Heidelberg on Advanced Concepts for Intelligent Vision Systems, 2008, pp. 1093-1102.

[8] W. Qi, Y. Sheng, L. Xianwei, “A Fast Mouth Detection Algorithm Based On Face Organs” IEEE International Conference on Power Electronics and Intelligent Transportation System, December 2009, pp. 250-252.

[9] M. Ferrara, A. Franco, D. Maio, D. Maltoni, "Face Image Conformance to ISO/ICAO Standards in Machine Readable Travel Documents", IEEE Transactions on Information Forensics and Security, AUGUST 2012,Vol 7, No. 4, pp. 1204-1213.

[10] T.H.B Nguyen, V.H. Nguyen, and H. Kim, "Automated conformance testing for ISO/IEC 19794-5 Standard on facial photo specifications", Int. J. Biometrics, Vol. 5, No. 1, 2013, pp.73–98.

[11] S. Coronel Castellanos, I. Solis Moreno, J. A. Cantoral Ceballos, R. Alvarez Vargas, P. L. Martinez Quintal , " An Approach to Improve Mouth-State Detection to Support the ICAO Biometric Standard for Face Image Validation", International Conference on Mechatronics, Electronics and Automotive Engineering, 978-1-4673-8329-5/15 ,2015 IEEE , DOI 10.1109/ICMEAE.2015.12

[12] R. L. Parente, L. V. Batista, Igor L. P. Andrezza, Erick V. C. L. Borges, Rajiv A. T. Mota, "Assessing Facial Image Accordance to ISO/ICAO Requirements", 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Sao Paulo, Brazil October 2016, IEEE, , DOI: 10.1109/SIBGRAPI.2016.033

[13] Y. Li, W. Wu, B. Zhang and F. Li, "Enhanced HMAX model with feedforward feature learning for multiclass categorization". Front. Comput. Neurosci, 2015, Vol.9, pp. 123. DOI: 10.3389/fncom.2015.00123

[14] HH. Gorji, S. Zabbah, R. Ebrahimpour, "A temporal neural network model for object recognition using a biologically plausible decision making layer", 2018, arXiv preprint arXiv:1806.09334.

[15] J. Filali, H. Zghal, J. Martinet. "Ontology and HMAX Features-based Image Classification using Merged Classifiers". International Conference on Computer Vision Theory and Applications2019 (VISAPP’19), Feb 2019, Prague, Czech Republic. hal-02057494.

[16] H. Yu, Z. Xu, Ch. Fu and Y. Wang, "Bottom-up attention based on C1 features of HMAX model", Proc. SPIE 8558, Optoelectronic Imaging and Multimedia Technology II, 85580W (21 November 2012), https://doi.org/10.1117/12.999263.

[17] S. Saraf Esmaili, K. Maghooli & A. Motie Nasrabadi , " A new model for face detection in cluttered backgrounds using saliency map and C2 texture features", International Journal of Computers and Applications, Vol. 40, No. 4, 2018, pp. 214-222.

[18] H.Zh. Zhang, Y.F. Lu, T. K. Kang and M.T. Lim, "B-HMAX: A Fast Binary Biologically Inspired Model for Object Recognition", Neurocomputing, Vol. 218, 19 December 2016, pp. 242-250. http://dx.doi.org/10.1016/j.neucom.2016.08.051

[19] C. Th´eriault, N. Thome, and M. Cord," Extended coding and pooling in the HMAX model", IEEE

Transactions on Image Processing , Vol. 22 , No. 2 , Feb. 2013 , pp. 764 – 777.

[20] E.T Rolls and G Deco, "Computational neuroscience of vision", Press: Oxford, 1st edition, 2006.

[21] J. Mutch and D.G. Lowe, “Object class recognition and localization using sparse features with limited receptive fields,” International Journal of Computer Vision, vol. 80, October 2008, pp. 45–57.

[22] I. Khan, H. Abdullah and M. Shamian Bin Zainal, "Efficient eyes and mouth detection algorithm using combination of viola jones and skin color pixel detection", International Journal of Engineering and Applied Sciences, Vol. 3, No. 4, June 2013, pp. 51-60.

[23] M. Jones and P. Viola., "Face Recognition Using Boosted Local Features". Mitsubishi Electric Research Laboratories Technical Report Number: TR2003-25. Date: April, 2003.

[24] NH. Barnouti, S. Sameer, "Face Detection and Recognition Using Viola-Jones with PCA-LDA and Square Euclidean Distance", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 7, No. 5, 2016, pp.371-377.

[25] S. Kaur, A. Chadha, "Supervised Descent Method Viola-Jones and Skin Color Based Face Detection and Tracking ", International Journal of Engineering Development and Research (IJEDR), Volume 5, Issue 2, ISSN: 2321-9939, 2017, pp. 1689-1695.

[26] I. Dagher and H. Al-Bazzaz, "Improving the Component-Based Face Recognition Using Enhanced Viola–Jones and Weighted Voting Technique", Hindawi Modelling and Simulation in Engineering, Volume 2019, Article ID 8234124, 2019, 9 pages. https://doi.org/10.1155/2019/8234124

[27] Q. Cheng, H. Zhou, and J. Cheng, "The Fisher-Markov Selector: Fast Selecting Maximally Separable Feature Subset for Multiclass Classification with Applications to High-Dimensional Data", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 33, No. 6, June 2011, pp. 1217-1233.

[28] V.N. Vapnik, "Statistical Learning Theory", Wiley, 1998.

[29] B. Scholkopf and A.J. Smola, "Learning with Kernels", MIT Press, 2002.

[30] M. Riesenhuber and T. Poggio, “Hierarchical models of object recognition in cortex,” Nature Neuroscience, vol. 2, 1999, pp. 1019–1025.

[31] A. Mittal, AK. Moorthy, AC. Bovik, "No-reference image quality assessment in the spatial domain". IEEE Trans. Image Process. Vol. 21, No. 12, 2012, pp. 4695–4708.

[32] A. M. Martinez and R. Benavente, "The AR Face Database", CVC Technical Report No.24, June 1998.

[33] A. Kasinski, A. Florek, A. Schmidt, "The PUT Face Database", Image Processing & Communication. Vol. 13. No. 3, 2008, pp. 59-64.

[34] A. Turpin, F. Scholer, "User performance versus precision measures for simple search tasks", Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (Seattle, WA, August 06–11, 2006). New York, NY: ACM. pp. 11–18.

[35] K.H. Brodersen, C.S. Ong, K.E. Stephan, J.M. Buhmann,
"The binormal assumption on precision-recall curves ", at the Wayback Machine. December 8, 2012, Proceedings of the 20th International Conference on Pattern Recognition, pp. 4263-4266.

Azamossadat Nourbakhsh received the B.S. degree in Computer Engineering from Azad University, Lahijan Branch, Iran in 1998, and M.S. degree in Artificial Intelligence & Robotics from Azad University, Science & Research Branch, Iran, in 2007. She is Ph.D. Candidate in Azad University, Science & Research Branch, Tehran, Iran. Her research interests include Image Processing, Machine Vision, Biometrics, Cognitive Science and Machine Intelligence.

Mohammad. Shahram Moin received his B.Sc. degree from Amir Kabir University of Technology, Tehran, Iran, in 1988; M.Sc. degree from University of Tehran, Iran, in 1991; and Ph.D. degree from École Polytechnique de Montréal, Montréal, Canada, in 2000, all degrees in electrical engineering. Dr. Moin is associate professor and head of IT Research Faculty in ICT Research Institute (ITRC). His research interests are Pattern Recognition, Image Processing, Biometrics, Data Mining and Big data Analytics.