

Extreme Learning Machine-Based Age-Invariant Face Recognition With Deep Convolutional Descriptors

Leila Boussaad, Computer Science Department, Batna 2 University, Algeria

Aldjia Boucetta, Computer Science Department, Batna 2 University, Algeria

ABSTRACT

The principal intention of this paper is to study face recognition across age progression at two levels: feature extraction and classification. In other words, this work aims to prove the benefit of replacing the Softmax layer of the deep-convolutional neural networks (CNN) by extreme learning machine (ELM) classifier based on deep features computed from fully-connected layer of pre-trained AlexNet CNN model in a context of age-invariant face recognition. Experimental results indicate that the ELM classifier combined with feature extracted by the pre-trained AlexNet CNN model worked effectively for face recognition across age progression. As significant highest mean accuracy rates are always obtained using ELM classifier, these results are more significant, following a 95% confidence level hypothesis test.

KEYWORDS

Age progression, Automatic Feature Extraction, Classification, Convolutional Neural Network (CNN), Face Identification

INTRODUCTION

Age-invariant face recognition, as a focus topic of face recognition in uncontrolled environment, is a very useful technology which may be applied in large real-world applications in which age compensation is required, like criminal and missing children identification, and biometric security systems.

Dealing with aging related variations is a challenging task because age related effects differ for different individuals and it is in combination with external factors, like health conditions and lifestyle, which have been shown to contribute to facial aging effects (Lanitis et al. (2009)), thus, elaborating age-invariant face recognition systems becomes a major necessity.

Feature extraction is a key step in face recognition system. It involves extracting information which best represent the image and which are invariant in context of face recognition across age progression.

Prior to developing Deep-CNN, multiple manual age-invariant methods have been proposed for features extraction (Sungatullina et al. (2013), (Bereta et al. (2013)), and (Ling et al. (2009)), that are computed from low level characteristics and statistical representation. Recently, Deep-CNN has become the most common method in use for automatic feature extraction (Agrawal et al. (2019)), (Shakeel et al. (2019)), and (Moustafa et al. (2020)). Also, Deep-CNN can be used in transfer learning with features from the pre-trained CNN model, in a classification task.

DOI: 10.4018/IJAMC.290540

This article published as an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>) which permits unrestricted use, distribution, and production in any medium, provided the author of the original work and original publication source are properly credited.

The success of CNN in the past few years is justified on the following grounds: the accessibility of great, labeled databases for training and validation such as the Large Scale Visual Recognition Challenge (ILSVRC) (Russakovsky et al. (2015)), also the availability of parallel graphics units (GPU) that allows speeding up the learning process.

Furthermore, before the CNNs started to take over, other classifiers like Support Vector Machines (SVM) and Extreme learning Machine (ELM) have proved already a high degree of accuracy in many applications. Therefore, it is a great deal of logic to ask the question:

Which is highly advisable, Deep-CNN classifier or Extreme Learning Machine (ELM) classifier for age-invariant face recognition?

Motivated by few partial studies that have already reported on the topic, such as (Tang et al. (2013)), (Sakr et al. (2016)), and (Zhang et al. (2016)), this study is an attempt to answer this question, and particularly, in a context of age-invariant face recognition.

The main purpose of this work is to use a pre-trained Deep-CNN for automatic feature extraction, and to use Extreme Learning Machine (ELM) classifier instead of using the Softmax classifier of the CNN, at the recognition level. To reach this objective, we perform an extensive comparative study in which we consider Softmax classifier, Support Vector Machine (SVM) classifier, Extreme Learning machine (ELM) classifier, and Kernel Extreme Learning Machine (KELM) classifier, based on deep features computed from fully-connected layers 'Fc6', 'Fc7' and 'Fc8' of a pre-trained AlexNet CNN model. It is worth noting that the considered pre-trained deep-CNN model is among the largest applied in many research areas, and it is seen to be among the most successful in use for recognition tasks.

The rest of the paper is organized as follows: in section 2, we review some related works on existing methods for face recognition across age progression. In section 3, we present our proposed approach, the aging database, the pre-processing stage, the deep-learning based feature extraction method and the considered classifiers. Section 4 reports results and discussions and section 5 concludes the paper.

RELATED WORKS

The methods that have been proposed in regards to the aging effects on face recognition can be categorized in two main classes (Ramanathan et al. (2009a)): “generative”, and “discriminative” methods.

Generative methods (Ramanathan et al. (2009b)), (Ramanathan et al. (2008)), (Park et al. (2010)), and (Park et al. (2008)) are based directly on age estimation and age transformation to convert the query image to the convenient age. Then, any standard recognition algorithm can be used to obtain the query identity.

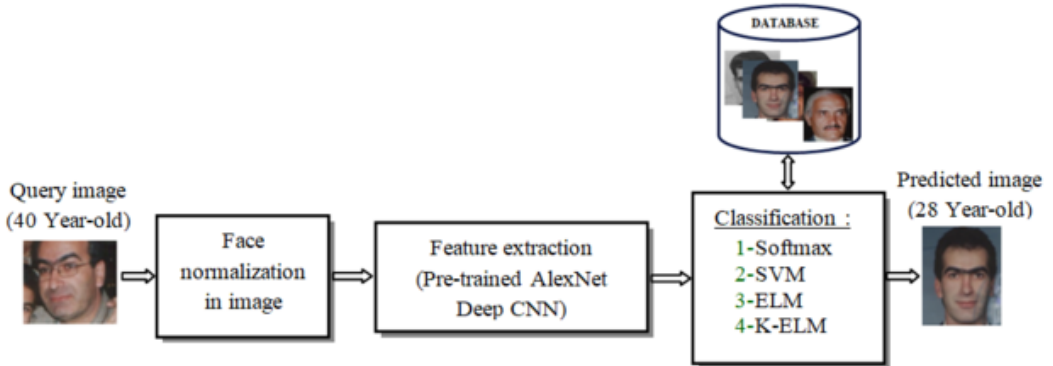
However, discriminative methods (Li et al. (2017)), (Boussaad et al. (2016)), (Sajid et al. (2018a)), and (Sajid et al. (2018b)) focus particularly on the choice of discriminatory features and metric learning that are invariant over time. Lately, another category can be added, it includes deep-learning based methods (Wang et al. (2018)), (El Khiyari et al. (2016)), (El Khiyari et al. (2017)), (Sajid et al. (2018c)), (Zhao et al. (2019)), and (Ni et al. (2019)).

A survey on age-invariant face recognition systems, effect of aging on the accuracy of age-invariant face recognition systems as well as some facial databases may be found in (Sawant et al. (2019)).

In this section, we briefly list some of the recent researches related to age-invariant face recognition. Duong et al. (Duong et al. (2017)) presented a probabilistic model, called Temporal Non-Volume Preserving (TNVP) transformation which can model the long-term aging process by separating it into various short-term stages. The structure may be processed into a deep convolutional network while ensuring the benefits of probabilistic models with tractable log-likelihood density estimation.

Also, Agrawal et al. (Agrawal et al. (2019)) proposed an age-invariant face recognition method based on a Galaric Swarm Optimization (GSO) algorithm for optimal selection of local texture features, and Extreme Learning Machine (ELM) classifier for modeling the aging function.

Figure 1. Different modules involved in the proposed age-invariant face recognition approach



Similarly, Shakeel et al. (Shakeel et al. (2019)) presented a discriminative model for age-invariant face recognition that use Euclidean locality adaptor for encoding Deep features obtained using the pre-trained AlexNet CNN model and Canonical Correlation Analysis (CCA) for feature fusion.

In the same context, Moustafa et al. (Moustafa et al. (2020)) introduced an age-invariant method, where features are extracted by activation of different layers of a pre-trained VGG-face CNN model, and selected by Multi-Discriminant Correlation Analysis (MDCA).

Furthermore, based on the fact that aging influences the facial parts (eyes, nose, and mouth) differently, Nimbarthe et al. (Nimbarthe et al. (2020)) suggest to use a weighted sum as a score-level fusion algorithm to combine scores given by a Support Vector Machines (SVM) classifier based on deep features computed from vertical and horizontal partitions of a face. However, Boussaad et al. (Boussaad et al. (2020a)) propose to use a Discriminant Correlation Analysis (DCA) as a feature-level algorithm to combine deep-based features computed from separated facial components (eyes, mouth and nose) and a Support Vector Machines (SVM) as a classifier.

Proposed Approach Description

The proposed age-invariant face recognition approach includes three modules, as shown in figure 1. After a normalization stage, where images are rotated, cropped and resized, feature extraction is performed using the pre-trained AlexNet CNN and classification is completed using four classifiers (Softmax classifier, Support Vector Machine (SVM) classifier, Extreme Learning machine (ELM) classifier, and Kernel Extreme Learning Machine (KELM) classifier). The details of each module as well as the used aging databases are described in the following subsections:

Aging Databases

For the experiment described in this paper, we used two well known and publicly available databases that are: FG-NET (Face and Gesture Recognition Research Network) aging database (G. Face (2020)) and MORPH aging database (Ricanek et al. (2006)).

FG-NET database contains 1002 face images of 82 subjects at different ages ranging in age from newborns to 69, for which there were a different few number of images per subject, about 4 to 12 images per subject. Typical images from the database are shown in Figure 2.

MORPH database includes two parts: MORPH-1 contains 1690 images from different subjects, for which there were about 2 to 4 images, and MORPH-2 contains 15024 images from different subjects, with about 3 to 8 images per person.

The age difference in MORPH-1 is from 0 to 30 but that in MORPH-2 is less than 5. This made the MORPH-1 more advisable in studying the aging effects. Typical images from the database are shown in Figure 3.

Figure 2. Examples of images from the FG-NET aging database (The number under each face image represents the age of the image).



The databases include old photos with strong variations such as illumination, pose, occlusion, and expression. Given these characteristics, the databases are considered as challenging databases.

Face Images Normalization

In experiments performed in this study, the images are preprocessed using the steps that are described in the following:

- Original FG-NET and MORPH images were first rotated (to get eyes at fixed points in image so that the inter-ocular segment is horizontal) based on eye coordinates in files provided with the original FG-NET and MORPH databases.
- Images were cropped (using the eyes coordinates) to remove the background.
- Images were resized to size required by the pre-trained AlexNet CNN model and each gray-scale image was converted to RGB image. Figure 4 shows the steps of normalization process.

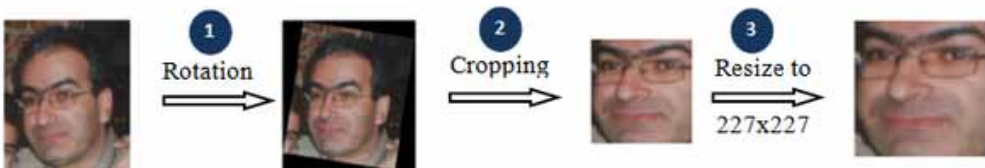
Features Extraction

As we have already indicated, we use a pre-trained Deep-CNN model for feature extraction; the considered pre-trained model is AlexNet. Our choice of AlexNet CNN model for feature extraction is based on a recent work (Boussaad et al. (2020b)), which evaluates five popular pre-trained CNN

Figure 3. Examples of images from the MORPH aging database (The number under each face image represents the age of the image).



Figure 4. Different steps of normalization process.



(AlexNet, GoogleNet, Inception V3, ResNet50 and SqueezeNet) and demonstrates that AlexNet model appears to be an excellent tool for feature extraction for age-invariant face recognition systems.

For the pre-trained AlexNet Deep-CNN model, feature extraction is computed from the fully-connected layers 'Fc6', 'Fc7', and 'Fc8'.

Pre-Trained AlexNet Deep CNN Model:

AlexNet (Krizhevsky et al. (2012)) was the winner of LSVRC-2012. It is simple, and powerful, network architecture with convolutional and pooling layers one on top of the other, followed by fully connected layers at the top. This architecture is commonly used as a starting point when applying a deep learning approach to computer vision tasks.

Specifically, AlexNet includes five convolution layers and three fully connected layers, which are all followed by a ReLu layer except for the last fully connected layer where a Softmax function is used. The architecture also consists of three pooling layers, two normalization layers and one dropout

layer. This architecture is clearly shown in Figure 5; further table 1 shows AlexNet implementation in Matlab, its layers, and required parameters.

CLASSIFICATION

Softmax Classifier

In a classification role, Deep-CNN commonly uses a Softmax layer. It is usually the last layer in the architecture, it has K inputs nodes, and it provides K outputs interpreted as a discrete probability distribution $p_i, (i=1, \dots, K)$ of K labeled classes, thus:

$$\sum_{i=1}^K p_i = 1$$

So, the total input into a Softmax layer, given by a_i , is given by:

$$a_i = \sum_k h_k W_{ki} \quad (1)$$

Where h is the activation of the next to last layer nodes, W is the weight connecting the next to last layer to the Softmax layer, and the output of a_i is given by:

$$p_i = \frac{e^{a_i}}{\sum_j^K e^{a_j}} \quad (2)$$

The predicted class \hat{i} is given by:

$$\hat{i} = \arg \max_i p_i = \arg \max_i a_i \quad (3)$$

Support Vector Machine (SVM)

Support Vector Machines (SVMs) were originally developed for binary classification (Cortes et al. (1995)). They are the most favorite for many, since they may provide meaningful accuracy with a computational efficiency. In this subsection, we briefly summarize the principle of SVM. We keep the same notations used in (Zong et al. (2011)):

Given a set of labeled training data (x_i, t_i) , where the point $x_i \in \mathbb{R}^d$, the label $t_i \in \{-1, 1\}$, and $i=1, \dots, N$. SVM seeks for a hyper plane that distinctly classifies the data into two classes with maximum separating margin in the feature space provided by:

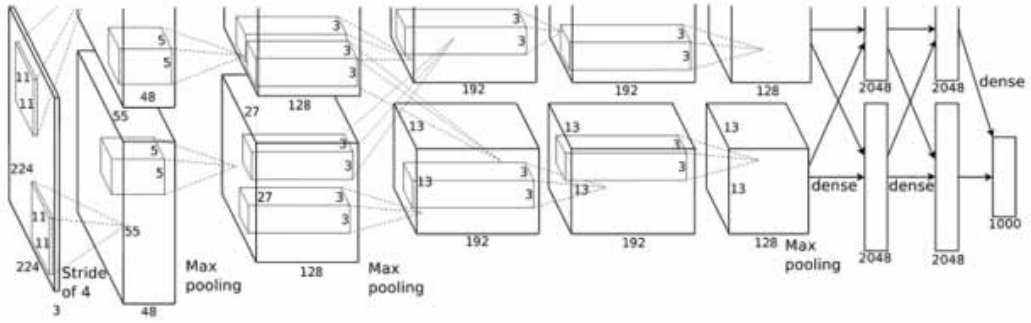
$$\phi(x) : x_i \rightarrow \phi(x_i).$$

To maximize the separating margin distance, the optimization function is given by the following equation:

Table 1. AlexNet implementation in Matlab, different layers, and required parameters

Layer number	Layer name	Layer type	Description
1	'data'	Image Input	227x227x3 images with 'zerocenter' normalization
2	'conv1'	Convolution	96 11x11x3 convolution with stride [4 4] and padding [0 0 0]
3	'relu1'	Relu	Relu
4	'norm1'	Cross channel Normalization	Cross channel normalization with 5 channels per element
5	'pool1'	Max Pooling	3x3 max pooling with stride [2 2] and padding [0 0 0]
6	'conv2'	Grouped Convolution	2 groups of 128 5x5x48 convolutions with stride [11] and padding [2 2 2]
7	'relu2'	Relu	Relu
8	'norm2'	Cross channel Normalization	Cross channel normalization with 5 channels per element
9	'pool2'	Max Pooling	3x3 max pooling with stride [2 2] and padding [0 0 0]
10	'conv3'	Convolution	384 3x3x256 max pooling with stride [2 2] and padding [0 0 0]
11	'relu3'	Relu	Relu
12	'conv4'	Grouped Convolution	2 groups of 192 3x3x192 convolutions with stride [11] and padding [1 1 1]
13	'relu4'	Relu	Relu
14	'conv5'	Grouped Convolution	2 groups of 128 3x3x192 convolutions with stride [11] and padding [1 1 1]
15	'relu5'	Relu	Relu
16	'pool5'	Max Pooling	3x3 max pooling with stride [2 2] and padding [0 0 0]
17	'fc6'	Fully Connected	4096 fully connected layer
18	'relu6'	Relu	Relu
19	'drop6'	Dropout	50% dropout
20	'fc7'	Fully Connected	4096 fully connected layer
21	'relu7'	Relu	Relu
22	'drop7'	Dropout	50% dropout
23	'fc8'	Fully Connected	1000 fully connected layer
24	'prob'	Softmax	softmax
25	'output'	Classification Output	Cross entropy ex with 'tench' and 999 other classes

Figure 5. Pre-trained AlexNet CNN model architecture



$$\text{Minimize} : L_p = \frac{1}{2} W^2 + C \sum_{i=1}^N \xi_i$$

$$\text{Subjectto} : t_i (W \cdot \phi(x_i) + b) \geq 1 - \xi_i, \quad i = 1, \dots, N$$

$$\xi_i \geq 0, \quad i = 1, \dots, N \quad (4)$$

Where C is a parameter stated by the user.

This is equivalent to minimize the training error, which is given by the following equation:

$$\text{Minimize} : L_D = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N t_i t_j \alpha_i \alpha_j \phi(x_i) \phi(x_j) - \sum_{i=1}^N \alpha_i$$

$$\text{subjectto} : \sum_{i=1}^N t_i \alpha_i = 0$$

$$0 \leq \alpha_i \leq C, \quad i = 1, \dots, N \quad (5)$$

Where for each training example (x_i, t_i) corresponds a Lagrange multiplier α_i

If we consider the kernel function $K(u, v) = \phi(u) \cdot \phi(v)$, the equation (5) can be written as:

$$\text{Minimize} : L_D = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N t_i t_j K(x_i, x_j) \alpha_i \alpha_j - \sum_{i=1}^N \alpha_i$$

$$\text{subjectto} : \sum_{i=1}^N t_i \alpha_i = 0$$

$$0 \leq \alpha_i \leq C, \quad i = 1, \dots, N \quad (6)$$

Extreme Learning Machine (ELM)

Extreme Learning Machines (Huang et al. (2006)) are simple learning methods that are based on single hidden layer feed-forward networks. Compared to classical feed-forward network, the ELM does not need a gradient-based back-propagation to determine its weights.

The ELM is distinguished by higher generalization performance and a very fast learning process (Huang et al. (2006)). In this subsection, we briefly introduce the concept of ELM as a classifier (Huang et al. (2006)):

Given a set of N labeled training data (x_i, t_i) , where $x_i \in \mathbb{R}^n$, the label $t_i \in \mathbb{R}^m$. ELM seeks for determining the output weights of a single layer feed-forward network with minimum norm and minimum training error. The output of ELM with L hidden nodes is given by:

$$o_j = \sum_{i=1}^L \beta_i g(x_i) = \sum_{i=1}^L \beta_i g(w_i \cdot x_i + b_i), j = 1, \dots, N \quad (7)$$

Where w_i and β_i are, respectively, the weights connecting the i^{th} hidden node with input and output nodes. b_i is the threshold of the i^{th} hidden node. The optimization function is given by the following equation:

$$\text{Minimize} : \sum_{j=1}^L o_j - t_j, j = 1, \dots, N \quad (8)$$

In other words, there exist w_i , β_i , and b_i such as:

$$\sum_{i=1}^L \beta_i g(w_i \cdot x_i + b_i) = t_i, j = 1, \dots, N \quad (9)$$

EXPERIMENTATION

Experimentation Methodology Description

The whole algorithm was evaluated using Matlab (R2018b) environment. All the images in the FG-Net and the MORPH-1 databases are used for training and test, and as we previously noted, all images are resized to the size required by the pre-trained AlexNet CNN that is 227x227, and any gray-scale image was converted to RGB image.

The study considers four classifiers, that are Softmax, SVM, ELM, and KELM on deep features computed from fully-connected layers ‘Fc6’, ‘Fc7’, and ‘Fc8’ of the pre-trained AlexNet CNN model. The overall process is given in table 2.

Table 2. Selecting the greatest classifier for age-invariant face recognition based on higher 10-folds cross validation (c.v.) average recognition accuracy rates.

Pre-trained AlexNet CNN layer	Classifiers			
	Softmax	SVM	ELM	KELM
‘Fc6’	c.v.	c.v.	c.v.	c.v.
‘Fc7’	c.v.	c.v.	c.v.	c.v.
‘Fc8’	c.v.	c.v.	c.v.	c.v.

The same selection process is repeated three times, using FG-NET database, MORPH-1 database, and fused FG-NET and MORPH-1 databases.

Experimental Results

Experimental results are reported in terms of average and standard deviation of the recognition accuracy rates following a 10-fold cross validation scheme. Recognition accuracy represents the percentage of samples that are correctly predicted.

Table 3 and table 4 present the mean recognition accuracy rates and the standard deviations for the FG-NET database (table 3) and MORPH-1 database (table 4).

With regard to these results, we can note that always the accuracy rates provided with Deep-features obtained from the fully-connected ‘Fc7’ layer outperform the results given by the other layers (‘Fc6’ and ‘Fc8’), that clearly proves the powerful of the ‘Fc7’ layer as age-invariant face descriptor.

Also, from the comparison of the achievements by the four classifiers, we can note that SVM and ELMs still give the best results than a Softmax classifier, particularly when we compare the Softmax classifier with KELM, the improvement in accuracy rate is 9.08% (‘Fc7’ layer), 8.11% (‘Fc6’ layer), and 6.91% (‘Fc8’ layer) in the case of FG-NET database, and 7.56% (‘Fc7’ layer), 8.99% (‘Fc6’ layer), and 7.03% (‘Fc8’ layer) in the case of MORPH-1 database. Note that these Softmax accuracies were reached with an epoch number equal to 28.

Moreover, we can note that results given by ELMs classifiers even exceed those provided by SVM classifier, where we can find an increase of about 1% to 2% when using ELMs classifiers.

Furthermore, accuracy rates provided by KELM classifier are a little greater than those given by ELM, as we can see, that the differences are almost insignificant.

Table 3. Recognition accuracy rates, and standard deviations for different classifiers (FG-NET database).

Pre-trained AlexNet CNN layer	Softmax		SVM		ELM		KELM	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD
‘Fc6’	82.4%	2.1%	87.9%	1.35%	90.01%	1.8%	90.51%	1.17%
‘Fc7’	85.56%	1.78%	91.52%	0.88%	92.71%	2.21%	94.64%	1.67%
‘Fc8’	72.1%	1.7%	75.16%	1.54%	78.51%	1.65%	79.01%	0.9%

Table 4. Recognition accuracy rates and standard deviations for different classifiers (MORPH-1 database).

Pre-trained AlexNet CNN layer	Softmax		SVM		ELM		KELM	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD
‘Fc6’	84.01%	1.65%	85.1%	0.91%	92.15%	1.24%	93.0%	1.12%
‘Fc7’	88.12%	2.44%	93.91%	1.49%	95.57%	1.06%	95.68%	0.80%
‘Fc8’	73.2%	1.82%	79.01%	1.51%	80.1%	1.6%	80.23%	0.83%

Figure 6. Recognition accuracy rates (FG-NET database).

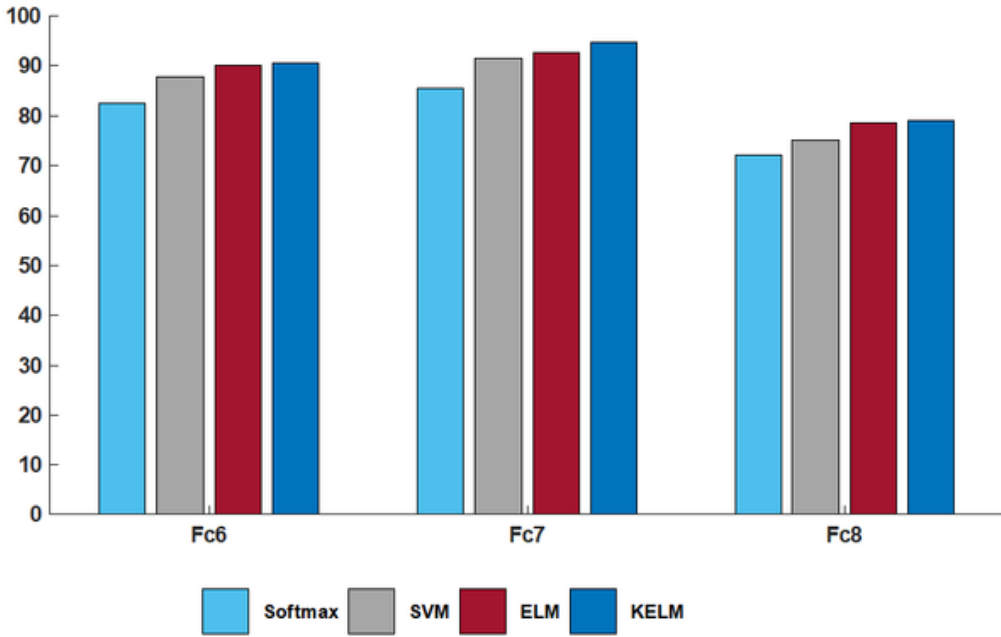
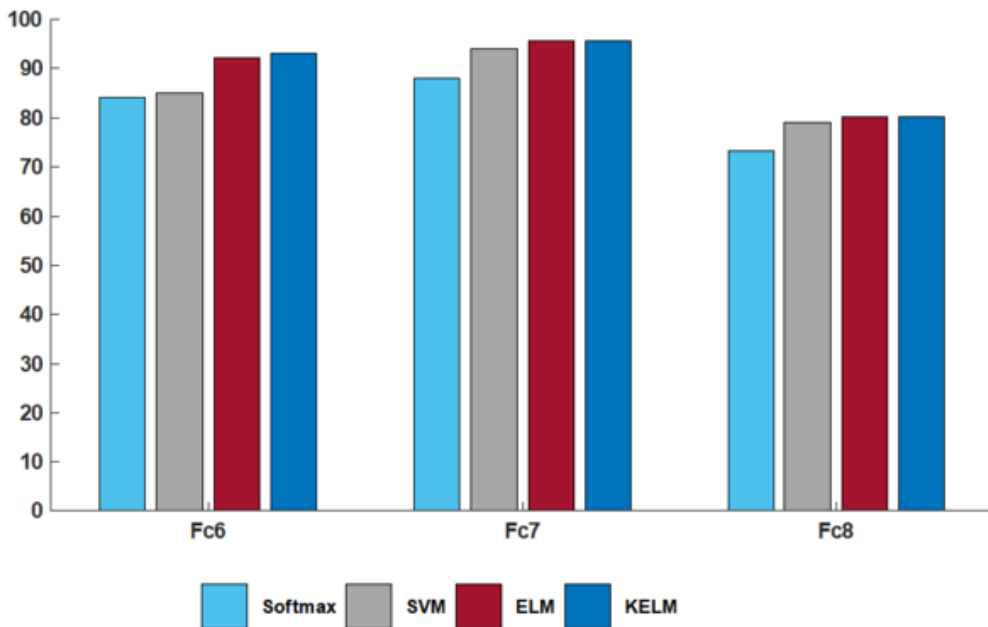


Figure 7. Recognition accuracy rates (MORPH-1 database).



For easier viewing, results were presented in figure 6 for the FG-NET database and figure 7 for MORPH-1 database. It is even more clear, that ELMs mainly KELM prove their power as classifiers in a context of age-invariant face recognition, and they may guarantee improved accuracy results most importantly when they used in conjunction with deep-CNN based descriptor.

Similarly, from the showed results in table 5 and figure 8, which represent accuracy rates using combination of the two databases (FG-NET and MORPH databases), we can specifically conclude that ELMs preserve their advantages, where the great accuracy rate (94.31%) was provided by KELM.

Table 5. Recognition accuracy rates and standard deviations for different classifiers (Fused FG-NET and MORPH-1 database).

Pre-trained AlexNet CNN layer	Softmax		SVM		ELM		KELM	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD
'Fc6'	80.91%	1.78%	85.24%	1.4%	91.92%	1.2%	91.51%	0.98%
'Fc7'	84.26%	1.62%	91.84%	2.34%	93.60%	0.94%	94.31%	1.7%
'Fc8'	70.0%	2.01%	77.13%	1.9%	81.3%	0.89%	81.2%	1.4%

Also, the figure 9 gives a clear picture of the relative stability of the different classifiers accuracy rates over the different databases.

Finally, we verify if the obtained results are significant, according to a 95% confidence level hypothesis test. A confidence interval approach can be used (Ferreira et al. (2017)). It consists of calculating the confidence interval at 95% of the difference of the average accuracy rates of two different classifiers:

Figure 8. Recognition accuracy rates (Fused FG-NET-MORPH databases).

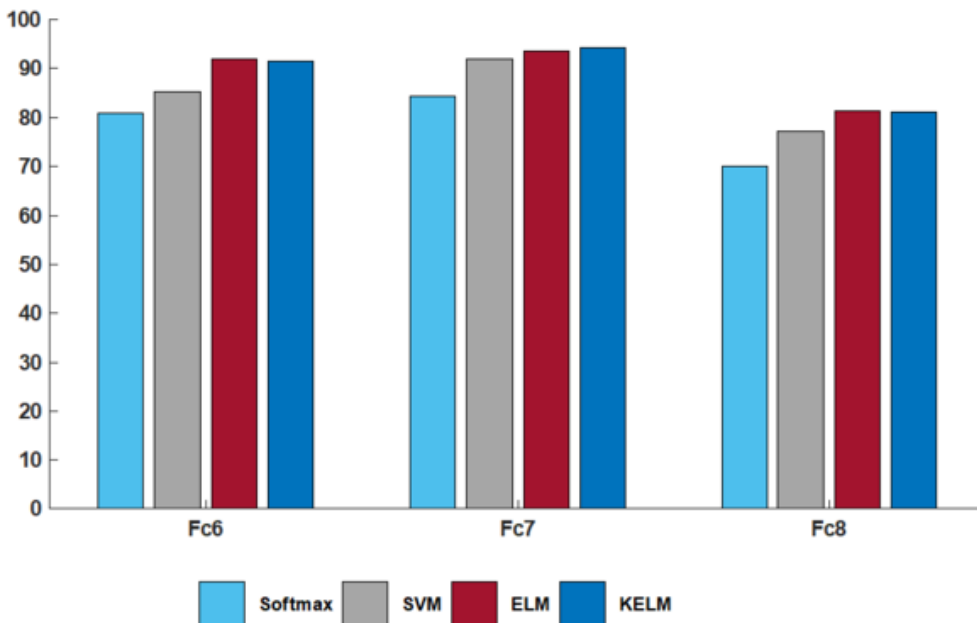
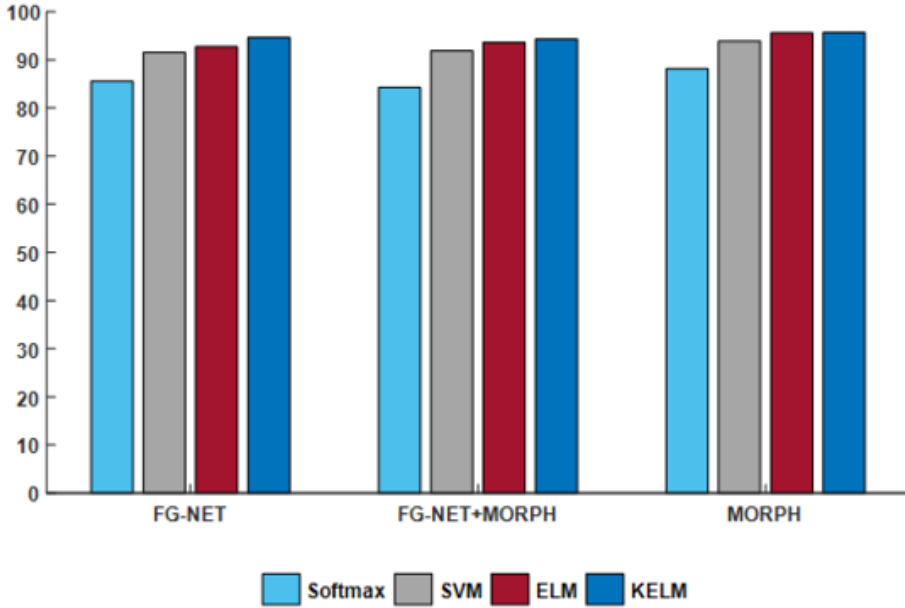


Figure 9. Recognition accuracy rates over different databases.



- if this 95% confidence interval does not contain 0 as a point, then we will conclude that the two average accuracy rates are different,
- if this confidence interval contains 0, then we will conclude that there is nothing to confirm that the two means are different.

The confidence interval IC_{ij} of difference mean accuracy rates at 95% confidence level of two classifiers i and j , is computed using the following equation:

$$IC_{ij} = (\mu_i - \mu_j) \pm 1.96\sqrt{(\delta_i^2 + \delta_j^2)} \quad (10)$$

Where μ_i and δ_j are, respectively, the mean accuracy and the standard deviation of the classifier i , and μ_j and δ_j are the mean accuracy and the standard deviation of the classifier j .

We compare four classifier accuracy rates, which represents six confidence intervals (see table 6). Each number in table 6 (1 to 6) represents two classifier combination for which a confidence interval at 95% of the difference of the average accuracy rates is computed.

From the hypothesis test result, presented by confidence intervals at 95% of means accuracy rates difference of two classifiers, computed by the equation 10, and showed in figure 10, figure 11, and figure 12, we see that for FG-NET database, all confidence intervals do not include the point zero, which means that the rates achieved by ELM classifier are significantly better than the other models (figure 10).

However, in the case of MORPH-1 database and fused-databases (figure 11 and figure 12), the interval 1, 2, 3, 4, and 5 do not contain the point zero, which always proves that results achieved by ELM classifiers still remain significantly better than other classifiers, but the interval 6 includes the point zero which still shows that there is no difference between accuracy rates provide by ELM, and KELM.

Table 6. Classifier comparisons considered in the hypothesis test.

	Softmax	SVM	ELM	KELM
Softmax	-	1	2	3
SVM	-	-	4	5
ELM	-	-	-	6
KELM	-	-	-	-

Figure 10. Hypothesis test confidence intervals for the mean accuracy differences (FG-NET_{database}).

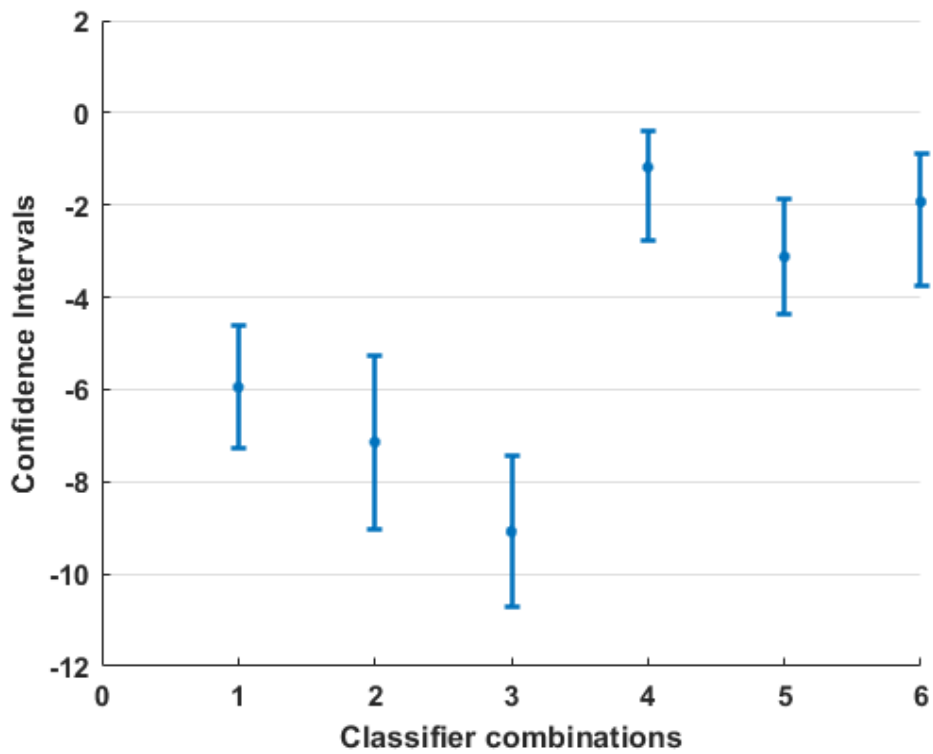


Figure 11. Hypothesis test confidence intervals for the mean accuracy differences (MORPH-1 database).

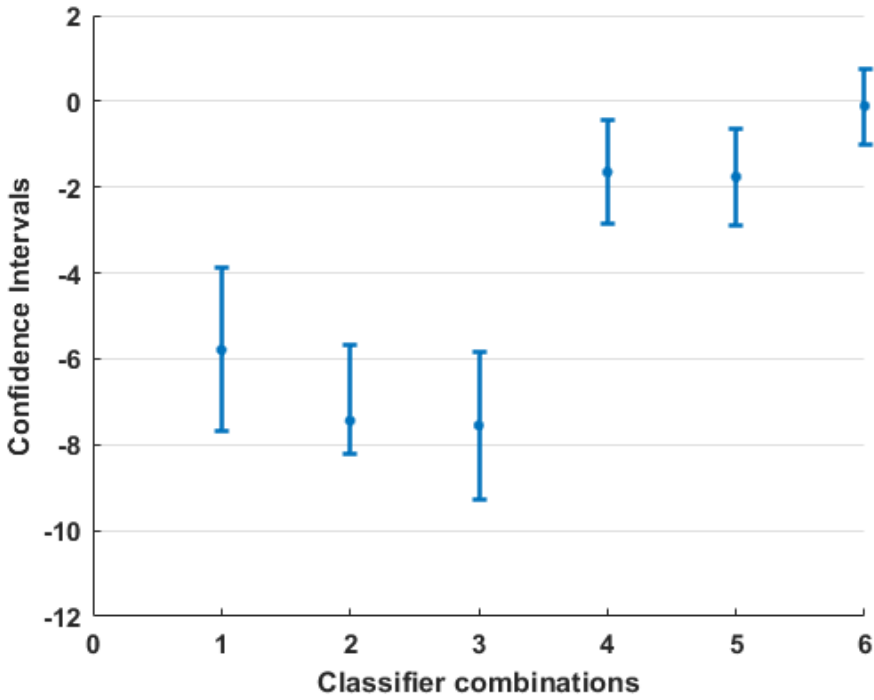
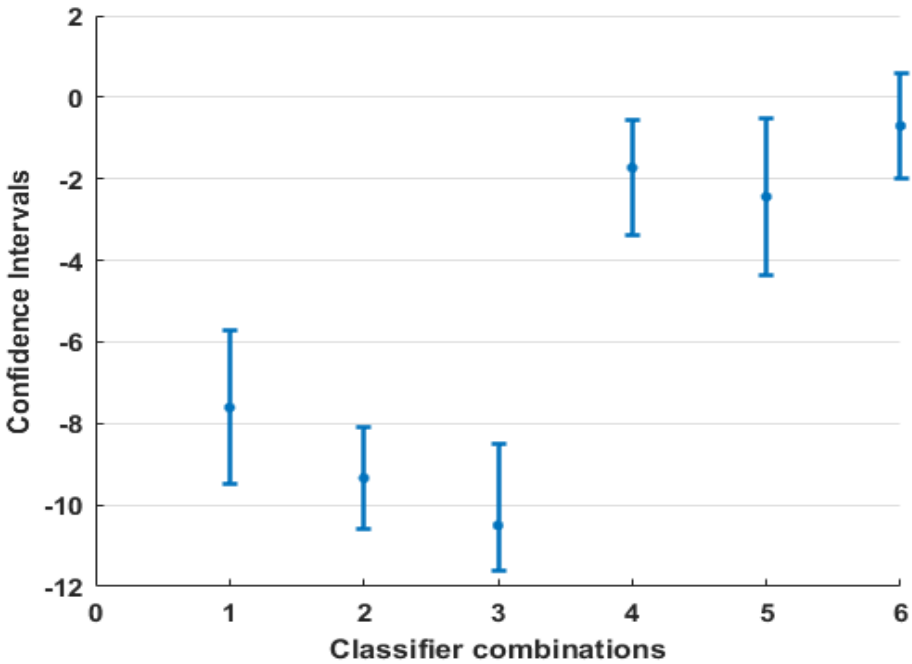


Figure 12. Hypothesis test confidence intervals for the mean accuracy differences (FusedFG-NET-MORPH-1 database).



CONCLUSION

This paper addresses face recognition across age progression. It aims to prove the benefit of replacing the Softmax layer with Extreme Learning Machine (ELM) classifier based on deep features computed from fully-connected layers 'Fc6', 'Fc7' and 'Fc8' of a pre-trained AlexNet CNN model, in a context of age-invariant face recognition.

To reach this purpose, a comparative study was performed, in which Softmax classifier; Support Vector Machine (SVM) classifier, Extreme Learning Machine (ELM) classifier, and Kernel Extreme Learning Machine (KELM) classifier were considered.

By examining and comparing the performances of the four classifiers, the question asked as problem can be answered: Replacing a Softmax layer by an ELM is very simple and it seems to be powerful for classification in context of age-invariant face recognition based on deep features computed from fully-connected layer 'Fc7' of a pre-trained AlexNet CNN model, as we have demonstrated that ELM classifier works better than Softmax on two standard databases.

As future work, it is important to test the proposed method by other effects, like illumination, expression, occlusion and pose by using dedicated face databases.

REFERENCES

- Agrawal, S., Kumar, S., Kumar, S., & Thomas, A. (2019). A novel robust feature extraction with GSO-optimized extreme learning for age-invariant face recognition. *Imaging Science Journal*, 67(6), 319–329. doi:10.1080/13682199.2019.1658914
- Bereta, M., Karczmarek, P., Pedrycz, W., & Reformat, M. (2013). Local descriptors in application to the aging problem in face recognition. *Pattern Recognition*, 46(10), 2634–2646. doi:10.1016/j.patcog.2013.03.010
- Boussaad, L., Benmohammed, M., & Benzid, R. (2016). Age Invariant Face Recognition Based on DCT Feature Extraction and Kernel Fisher Analysis. *Journal of Information Processing Systems*, 12(3).
- Boussaad, L., & Boucetta, A. (2020a). An effective component-based age-invariant face recognition using Discriminant Correlation Analysis. *Journal of King Saud University-Computer and Information Sciences*. 10.1016/j.jksuci.2020.08.009
- Boussaad, L., & Boucetta, A. (2020b). Deep-learning based descriptors in application to aging problem in face recognition. *Journal of King Saud University-Computer and Information Sciences*. 10.1016/j.jksuci.2020.10.002
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297.
- El Khiyari, H., & Wechsler, H. (2016). Face recognition across time lapse using convolutional neural networks. *Journal of Information Security*, 7(3), 141–151.
- El Khiyari, H., & Wechsler, H. (2017). Age invariant face recognition using convolutional neural networks and set distances. *Journal of Information Security*, 8(03), 174.
- Ferreira, F. P., & de Carvalho, T. B. A. (2017). Analysis of wavelet families for face recognition. In 2017 Workshop of Computer Vision (WVC) (pp. 90-95). IEEE.
- Group Face. (2000). *Fg-net aging database*. <https://web.archive.org/web/20070217193535/http://www.fgnet.rsunit.com/>
- Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing*, 70(1-3), 489–501.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105). Academic Press.
- Lanitis, A. (2009). Facial Biometric Templates and Aging: Problems and Challenges for Artificial Intelligence. In *AIAI Workshops* (pp. 142-149). Academic Press.
- Li, H., Zou, H., & Hu, H. (2017). Modified hidden factor analysis for cross-age face recognition. *IEEE Signal Processing Letters*, 24(4), 465–469. doi:10.1109/LSP.2017.2661983
- Ling, H., Soatto, S., Ramanathan, N., & Jacobs, D. W. (2009). Face verification across age progression using discriminative methods. *IEEE Transactions on Information Forensics and Security*, 5(1), 82–91. doi:10.1109/TIFS.2009.2038751
- Moustafa, A. A., Elnakib, A., & Areed, N. F. (2020). Age-invariant face recognition based on deep features analysis. *Signal, Image and Video Processing*, 14(5), 1–8. doi:10.1007/s11760-020-01635-1
- Nhan Duong, C., Gia Quach, K., Luu, K., Le, N., & Savvides, M. (2017). Temporal non-volume preserving approach to facial age-progression and age-invariant face recognition. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 3735-3743). Academic Press.
- Ni, T., Gu, X., Zhang, C., Wang, W., & Fan, Y. (2019). Multi-Task Deep Metric Learning with Boundary Discriminative Information for Cross-Age Face Verification. *Journal of Grid Computing*, 1–14.
- Nimbarte, M., & Bhojar, K. K. (2020). Biased face patching approach for age invariant face recognition using convolutional neural network. *International Journal of Intelligent Systems Technologies and Applications*, 19(2), 103–124.
- Park, U., Tong, Y., & Jain, A. K. (2008, September). Face recognition with temporal invariance: A 3d aging model. In *2008 8th IEEE International Conference on Automatic Face & Gesture Recognition* (pp. 1-7). IEEE.

- Park, U., Tong, Y., & Jain, A. K. (2010). Age-invariant face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(5), 947–954. doi:10.1109/TPAMI.2010.14 PMID:20299717
- Ramanathan, N., & Chellappa, R. (2008, September). Modeling shape and textural variations in aging faces. In *2008 8th IEEE International Conference on Automatic Face & Gesture Recognition* (pp. 1-8). IEEE. doi:10.1109/AFGR.2008.4813337
- Ramanathan, N., Chellappa, R., & Biswas, S. (2009). Age progression in human faces: A survey. *Journal of Visual Languages and Computing*, 15, 3349–3361.
- Ricanek, K., & Tesafaye, T. (2006, April). Morph: A longitudinal image database of normal adult age-progression. In *7th International Conference on Automatic Face and Gesture Recognition (FGR06)* (pp. 341-345). IEEE.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., & Berg, A. C. et al. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3), 211–252. doi:10.1007/s11263-015-0816-y
- Sajid, M., & Shafique, T. (2018c). Hybrid generative–discriminative approach to age-invariant face recognition. *Journal of Electronic Imaging*, 27(2), 023029.
- Sajid, M., Shafique, T., Manzoor, S., Iqbal, F., Talal, H., Samad Qureshi, U., & Riaz, I. (2018b). Demographic-assisted age-invariant face recognition and retrieval. *Symmetry*, 10(5), 148.
- Sajid, M., Taj, I. A., Bajwa, U. I., & Ratyal, N. I. (2018a). Facial Asymmetry-Based Age Group Estimation: Role in Recognizing Age-Separated Face Images. *Journal of Forensic Sciences*, 63(6), 1727–1749.
- Sakr, G. E., Mokbel, M., Darwich, A., Khneisser, M. N., & Hadi, A. (2016, November). Comparing deep learning and support vector machines for autonomous waste sorting. In *2016 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET)* (pp. 207-212). IEEE. doi:10.1109/IMCET.2016.7777453
- Sawant, M. M., & Bhurchandi, K. M. (2019). Age invariant face recognition: A survey on facial aging databases, techniques and effect of aging. *Artificial Intelligence Review*, 52(2), 981–1008.
- Shakeel, M. S., & Lam, K. M. (2019). Deep-feature encoding-based discriminative model for age-invariant face recognition. *Pattern Recognition*, 93, 442–457. doi:10.1016/j.patcog.2019.04.028
- Sungatullina, D., Lu, J., Wang, G., & Moulin, P. (2013, April). Multiview discriminative learning for age-invariant face recognition. In *2013 10th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG)* (pp. 1-6). IEEE. doi:10.1109/FG.2013.6553724
- Tang, Y. (2013). *Deep learning using linear support vector machines*. arXiv preprint arXiv:1306.0239.
- Wang, Y., Gong, D., Zhou, Z., Ji, X., Wang, H., Li, Z., . . . Zhang, T. (2018). Orthogonal deep features decomposition for age-invariant face recognition. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 738-753). Academic Press.
- Zhang, L., Zhang, D., & Tian, F. (2016). SVM and ELM: Who Wins? Object recognition with deep convolutional features from ImageNet. In *Proceedings of ELM-2015 Volume 1* (pp. 249-263). Springer. doi:10.1007/978-3-319-28397-5_20
- Zhao, S., Li, J., & Wang, J. (2019). Disentangled representation learning and residual GAN for age-invariant face verification. *Pattern Recognition*.
- Zong, W., & Huang, G. B. (2011). Face recognition based on extreme learning machine. *Neurocomputing*, 74(16), 2541–2551.