Research Article

Dongshu Wang, Heshan Wang*, Jiwen Sun, Jianbin Xin, and Yong Luo

Face Recognition in Complex Unconstrained Environment with An Enhanced WWN Algorithm

https://doi.org/10.1515/jisys-2019-0114
Received May 02, 2019; accepted Nov 21, 2019

Abstract: Face recognition is one of the core and challenging issues in computer vision field. Compared to computer vision, human visual system can identify a target from complex backgrounds quickly and accurately. This paper proposes a new network model deriving from Where-What Networks (WWNs), which can approximately simulate the information processing pathways (i.e., dorsal pathway and ventral pathway) of human visual cortex and recognize different types of faces with different locations and sizes in complex background. To enhance the recognition performance, synapse maintenance mechanism and neuron regenesis mechanism are both introduced. Synapse maintenance is used to reduce the background interference while neuron regenesis mechanism is introduced to regulate the neuron resource dynamically to improve the network usage efficiency. Experiments have been conducted on human face images of 5 types, 11 sizes, and 225 locations in complex backgrounds. Experiment results demonstrate that the proposed WWN model can basically learn three concepts (type, location and size) simultaneously. The experiment results also show the advantages of the enhanced WWN-7 model for face recognition in comparison with several existing methods.

Keywords: Face Recognition, Where-What Networks, Synapse Maintenance Mechanism, Neuron Regenesis Mechanism

1 Introduction

Biometric recognition can demonstrate an individual’s identity automatically based on one’s anatomical and behavioral characteristics [39]. With the development of artificial intelligence, computer vision, cognitive science and psychology, biometric recognition method has become an important technique in national safety and public security, due to its convenience and non-intrusiveness. It has been widely used in public security, financial system, intelligent surveillance, information security, civil aviation and military security.

Compared with other biometric features, such as fingerprints, palms, irises, voice, gait, and ears, human face has many distinct advantages [22]. Even at a long distance, face characteristics can be extracted from camera images, which provide a convenient and non-intrusive way to monitor remotely. Moreover, the face also has richer structure and larger area than other biometric features so that the face region is not easily occluded. Hence, face recognition has become an indispensable biological authentication method and has drawn much attention of scholars over the past decades in various domains [17].

Performance of the face recognition is affected by many factors, such as lighting condition, pose variation, occlusion, images with low resolution and complex background [6]. So the face recognition under un-
controlled conditions, e.g., complex backgrounds and variable resolutions, is still greatly challenging in the field of image processing and computer vision.

A typical way for face recognition is to use handcrafted-descriptor approaches which can be categorized into three groups: feature-based, holistic based and hybrid-based approaches [21]. In the feature-based method, a geometric vector representing the facial features is extracted by measuring and computing the locations and geometric relationships among facial features, such as the mouth, ears, eyes and nose, and input the geometric vector to a structural classifier [11]. Local geometric feature extraction approaches mainly include local binary pattern (LBP) [17, 24], Gabor features [16, 36], SIFT methods [10, 29] and their modified approaches [2, 42]. Compared with the feature-based methods, the holistic approaches generally extract the feature vector through operating on the whole face image rather than measuring the local geometric characteristics. The eigenface method [8, 28] is a well-known example that can be analyzed and extracted by several typical approaches, for instance Principle Component Analysis (PCA) [4, 20], Linear Discriminant Analysis (LDA) [1, 37] and Independent Component Analysis (ICA) [4]. Finally, the hybrid approaches are both holistic and feature extraction methods, and examples of this category can be found in [5, 40, 41, 43].

Unfortunately, recognition performance using handcrafted-descriptors approaches declines dramatically in unconstrained conditions, resulting from the fact that the constructed face representations are very sensitive to the highly nonlinear intra-personal variations, such as expression, illumination, pose, and occlusion. To address these drawbacks, considerable attention has been paid to the learning-based approaches that learn features from labeled training samples using machine learning techniques [3, 27] (e.g., deep learning neural networks [15, 19]). This approach learns automatically a set of effective feature representations through hierarchical nonlinear mappings, which can handle the nonlinear variations (intra- and inter-personal variations) of face images. Deep belief network (DBN) is one of the most popular unsupervised deep learning methods, which has been successfully applied to learn hierarchical representation from unlabeled data in a wide range of fields, including face recognition [18]. Nevertheless, a key limitation of the DBN when the pixel intensity values are assigned directly to the visible units is that the feature representations of the DBN are sensitive to the local translations of the input image. This can lead to disregarding local features of the input image known to be important for face recognition.

Despite the accumulation of literature on face recognition, these learning-based approaches above use symbolic representation and their results are confined to particular tasks [32]. For this, a type of emergent models referred to as Where-What Network (WWN) model is proposed for face recognition. The WWN model is a particular type of network models simulating human visual system which outperforms computer visual systems as the human visual system recognizes the generic objects more easily [23]. WWN models simulate the dorsal stream (where pathway) and the ventral stream pathway (what pathway) [14], and the objects type, location and size can be recognized simultaneously in constrained environments [25]. However, when it comes to face recognition in unconstrained environments, the latest WWN model (WWN-7) cannot achieve satisfactorily high recognition rates and its location errors and size errors are still relatively high.

The contribution of this paper is an improved WWN model for face recognition in unconstrained environments where the objects type, the location and the size can be recognized simultaneously. This new model, which is based on the WWN-7 model, incorporates synapse maintenance mechanism and neuron regenesis mechanism to improve the recognition performance. We deal with emergent representation of the face images in unsupervised learning in the internal neurons of the WWN-7 model. Synapse maintenance is used to reduce the background interference while neuron regenesis mechanism is introduced to regulate the neuron resource dynamically for enhancing the network usage efficiency. The proposed model is tested to identify human faces in the complex natural backgrounds. The experiment results show the advantages of the improved WWN-7 model for face recognition in comparison with several existing methods.

The remainder of the paper is organized as follows. Section II describes the related concepts and algorithm behind our model, e.g., synapse maintenance, neuron regenesis mechanism. Section III designs the experiment procedure, and analyzes the experiment results, while the conclusions and future studies are given in the last section.
2 Modeling

This section presents the improved WWN model for face recognition in unconstrained environment. The WWN-7 model, which is the basis of the proposed model, is described in the beginning and then its associated algorithm is given. Subsequently, the synapse maintenance and the neuron regenesis are introduced for improving the recognition performance.

2.1 The WWN-7 model

Where-What Networks model [14] is an embodiment of the developmental network model [31, 34], whose structure is shown in Figure 1. It is composed of three areas: X area, Y area and Z area.

As the sensor, X area is the retina and the perception part of the agent, responsible for external information input. The entire Y area, as the “brain” of the agent, is enclosed in the skull. Y area is not under the direct supervision of the outside world and it realizes the information exchange between X area and Z area, just like a “bridge”. It is divided into four parts: Y1, Y2, Y3 and Y4, with different neuron numbers. The preprocessing area is responsible for calculating the pre-action energy of neurons in the learning process to determine which neuron can fire (be activated). Z area is the motor terminal of the agent and it is divided into three parts: LM (location-motor), SM (size-motor) and TM (type-motor). LM and TM simulate the dorsal pathway and ventral pathway in human visual system, respectively.

In the WWN-7 model, X area unidirectionally inputs the images to Y area, and this implies a bottom-up connection, while the connections between Y area and Z area are bidirectional. The top-down connections indicate that the motor concepts of Z area can guide the brain to learn. In Y area, its bottom-up connections feed forward the image understanding of the network (agent) to the motor terminal.
(1) Connecting modes among neurons
Three connecting modes exist in the WWN-7 model, i.e., the bottom-up connection, the top-down connection and lateral connection. The information transmissions from X area to Y area and from Y area to Z area belong to the bottom-up connections, while the transmission from Z area to Y area means the top-down connection. There exist two modes of the lateral connection among neurons in the same area, i.e., activation and inhibition. The lateral connections exist in Y and Z areas, as shown in Figure 2, which is a simplified connection diagram of the WWN-7.

Figure 2: Three Connection Modes in WWNs.

(2) Receptive fields
The corresponding neurons in X area and Y area are locally connected. Y area receives the top-down input and bottom-up input simultaneously. Only the connections between X area and Y area, i.e., the bottom-up input, are discussed as an example in this section. The WWN-7 model puts the images apperceived by X area into Y area to operate directly. According to the location, each neuron in Y area is connected to the specific neuron in X area. In human visual system, inputs in early processing area can decide the receptive fields in the visual cortex region. Neuro-anatomical studies [13, 26] also confirmed that some neurons in the anterior part of the cerebrum had small receptive fields while some neurons in the later processing area had large receptive fields. From Figure 1, we can see that the four parts in Y area have different neurons, and sizes of the receptive fields increase gradually from $Y_1$ to $Y_4$, i.e., 7×7 pixels, 11×11 pixels, 15×15 pixels and 19×19 pixels, respectively. Shapes of the receptive fields are all square. Each receptive field starts from the upper left corner of the image and turns to the right and downward by a pixel, respectively. Information obtained from the receptive fields by each shift will be input to a corresponding neuron in Y area. This process continues until the total image has been input.

(3) Effective fields
WWN-7 has three motor areas: TM, LM and SM, which characterize the behaviors of the network. In the network learning process, the motor areas provide the agent with the location, type and size as guidance. According to the neuron activation, these concepts are related to the characteristics of the foreground input by X area to the agent. If the bottom-up inputs and the top-down inputs match well, the neurons in Y area that associate the characteristics with the concepts are likely to win in the top-k competition. Based on this, the teacher sets up the inputs of the corresponding motor concepts, i.e., the top-down inputs, according to the foreground information. Such inputs will enable the neurons in Y area to pay more attention to the foreground, thus the neurons in Y area which associate the characteristics with the concepts are more likely to fire constantly. Since the background is not related to the concepts of TM, LM or SM, only a small number
of neurons in Y area will learn the background, which is beneficial to the foreground identification from the complex background.

(4) Match of two inputs for the neurons in Y area

Each neuron in Y area has a weight vector $v=\left(v_b, v_t\right)$, corresponding to the area inputs ($b$, $t$). $b$ and $t$ denote the bottom-up and top-down inputs, respectively. Calculation of the pre-action energy of a neuron in Y area is depicted as follows:

$$r(v_b, b, v_t, t) = \frac{v_b}{||v||} \cdot \frac{b}{||b||} + \frac{v_t}{||v||} \cdot \frac{t}{||t||} = \hat{v} \cdot \hat{p}$$  \hspace{1cm} (1)

where $\hat{v}$ is the unit vector of the normalized synaptic vector $v=(\hat{v}_b, \hat{v}_t)$, and $\hat{p}$ is the unit vector of the normalized input vector $p=(\hat{b}, \hat{t})$. The inner product in the formula (1) evaluates the match degree between $\hat{v}$ and $\hat{p}$, because $r(v_b, b, v_t, t) = \cos(\theta)$, where $\theta$ represents the angle between the two unit vectors $\hat{v}$ and $\hat{p}$.

When one neuron in Y area can find a good match vector from the input vector $\hat{p}$, e.g., match from bottom to top, i.e., $v_b \cdot \hat{b}$, and the match from top to bottom, i.e., $v_t \cdot \hat{t}$, the corresponding neuron will fire and be updated. It means that the neurons learned the background have a small chance to fire because their top-down matches cannot be the best.

All the neurons in Y area compete to fire through the top-$k$ competition mechanism, which can be described as follows:

$$r'_q = \begin{cases} \frac{r_{q+1}-r_{q+1-1}}{r_{q+1}-r_{q+1-1}} & 1 \leq q \leq k \\ 0, & \text{otherwise} \end{cases}$$  \hspace{1cm} (2)

where the subscripts 1, $q$ and $k$+1 indicate the position ranking list of the pre-action energy in descending order.

In the paper, only $k=1$ is considered, so the winner neuron $j$ can be identified by the following calculation:

$$j = \arg \max_{1 \leq i \leq c} r(v_{bi}, b, v_{ti}, t)$$  \hspace{1cm} (3)

where $c$ is the neuron number in the corresponding area.

2.2 Learning Algorithm: Lobe Component Analysis

Lobe component analysis (LCA) [13] is used as the learning algorithm in the WWNs model. In terms of the individual development, individual difference will lead to different learning rates. Because of the different learning rates, just as we teach students in accordance with their aptitudes, appropriate mechanism of the network should develop according to different individuals and each neuron can achieve the best learning rate. So it is very important to design a reasonable mechanism of the learning rate on the basis of temporal optimality. WWNs model is a neural network to roughly simulate the visual processing pathway of the human being in this work. The learning process is not once and for all, but a cycle of progressive process. In the learning process, the new knowledge will be learned, and the knowledge that is not used will be forgotten gradually. The repetitive learning process can enhance the memory of the agent. Similarly, the learning rate of WWNs is designed to reflect these ideas which are embodied in the following weight updating formulas:

$$v_j \leftarrow \omega_1(n_j)v_j + \omega_2(n_j)y_j\hat{p}$$  \hspace{1cm} (4)

$$\begin{cases} \omega_2(n) = \frac{1+\mu(n)}{n} \\ \omega_1(n) + \omega_2(n) \equiv 1 \end{cases}$$  \hspace{1cm} (5)
where \( \omega_1(n_j) \) and \( \omega_2(n_j) \) are the learning rate and retention rate, respectively, both depending on the firing age \( n_j \) of neuron \( j \). The retention rate \( \omega_2(n_j) \) consolidates the knowledge that the neurons have learned. In the learning rate \( \omega_1(n_j) \), the amnesic mechanism \( \mu(n) \) is introduced, where \( m=2, r=10000, t_1=10, t_2=30 \) (based on the literature [13, 33]). In this way, we can dynamically adjust the learning rate and the retention rate in the weight updating process. An appropriate retention factor is set to make the network learn fast in early development stage. When it comes to a certain degree and the neuron age is large enough, the learning rate tends to be a constant gradually. So this mechanism can keep learning new knowledge persistently, and forgetting the knowledge less used.

### 2.3 Synapse Maintenance Mechanism

In this paper, the faces in the complex backgrounds have different types, locations and sizes. Especially in the case of different receptive fields, the backgrounds have great impacts on the detection of the face features. Therefore, it is able to estimate the standard deviation of each pixel in the input from X area to Y area by using the synapse maintenance mechanism (SMM) [31]. In training process, these pixels whose standard deviations vary greatly (corresponding to the background) will be suppressed, which will reduce the influence on the calculation of the pre-action energy of the neuron in Y area.

Receptive fields in computer vision are generally regular geometry, while the object outlines in real world are arbitrary. Therefore, the receptive field often contains partial background. In WWN-7, receptive field of the neuron in Y area is also a regular square. Obviously, the background will interfere the recognition result of the foreground. So the SMM is introduced into the WWN model to solve this problem. The SMM is expected to recognize the foreground and background in the receptive field. After the foreground outline is sought out, the background can be easily suppressed. It works to estimate the standard deviation of each pixel in the receptive field, through restraining the pixel with large standard deviation, to decrease the influence of the unstable pixel on the pre-action energy of the neuron. So this mechanism reduces the influence of the background on the neuron activation.

Supposing the input of Y neuron is \( p = (p_1, p_2, \ldots, p_d) \) and the value of a neuron in Y area is \( v = (v_1, v_2, \ldots, v_d) \). For the neuron \( i \) \((i=1, 2, \ldots , d)\), the matching degree between \( v_i \) and \( p_i \) can be calculated as follows:

\[
\sigma_i = E[|v_i - p_i| \mid \text{the neuron fires}] \tag{7}
\]

where \( \sigma_i \) is used to evaluate the matching degree of the input and the value of the neuron in Y area.

According to the matching degree \( \sigma_i \), each neuron decides to extract or retract the synapse dynamically. In Hebbian algorithm, \( v_i \) is the good estimate of the amount of \( p_i \) and the expected value of \( p_i \). If \( v_i \) is estimated well, \( \sigma_i \) can be seen as the standard deviation of the input pixel. So the brightness dispersion of the input pixel in the same position is estimated by \( \sigma_i \). A higher \( \sigma_i \) indicates a higher dispersion degree, meanwhile, the input pixel is more unstable and the synapse should be retracted. Now a given threshold is set to determine the synapse state. But the synapse will extend or retract repeatedly when the \( \sigma_i \) approaches the threshold we set. So the smooth synaptogenetic factor \( f_i \) is added in the following formula:

\[
f_i = \frac{1}{\eta(\sigma + \varepsilon)} \tag{8}
\]

where \( \varepsilon \) is an infinitesimal quantity, used to avoid the zero denominator. \( f_i \) will be normalized to prevent itself from being too large.

\[
\begin{align*}
\eta &= \sum_{i=1}^{d} (\sigma_i + \varepsilon)^{-1} \\
\sum_{i=1}^{d} f_i &= 1
\end{align*}
\tag{9}
\]
The synaptogenetic factor $f_i$ reflects the connecting strength between the pixel and the corresponding neuron. Range of the synaptogenetic factor $f_i$ is between 0 and 1. If $f_i$ is equal to 1, the corresponding pixel is completely connected, while 0 means completely suppressed.

Synapse maintenance requires pruning the input $p$ and the value $v$ as follows:

$$p'_i \leftarrow f_i \times p_i, \quad v'_i \leftarrow f_i \times v_i$$  \hspace{1cm} (10)$$

Then $p'_i$ and value $v'_i$ are used to calculate the pre-action energy of the neuron in Y area. Interested readers are suggested to refer the [30] for more concrete description of the SMM.

### 2.4 Neuron Regenesis Mechanism

Hebbian learning changes the synaptic weights between neurons, while neuron regenesis mechanism makes the inactive neurons have the chance to fire (meaning “new birth, renewal”) and make new connections with other neurons. Firing frequency is used to evaluate the neuron activation degree. If the firing frequency is low, the corresponding neurons are regarded not active, i.e., few winning times in competition. Therefore, they are possibly suppressed by their surrounding active neurons, and can’t learn new knowledge about the foreground. If these low level neurons don’t have a chance to regenerate and learn new features, it is a great waste of the neuron resources. So the neuron regenesis mechanism is designed to regulate the neuron resources, hence promote the WWNs performance.

At time $t$, the firing frequency is defined as follows:

$$f(t) = \frac{n(t)}{N(t)}$$  \hspace{1cm} (11)$$

where $f(t)$ is the firing frequency of the neuron, $n(t)$ is the neuron age (fire times) and $N(t)$ is the current running time of the network. Accordingly, at time $t+1$, the firing frequency can be defined:

$$f(t + 1) = \frac{n(t + 1)}{N(t + 1)}$$  \hspace{1cm} (12)$$

And the current running time of the network increases one unit:

$$N(t + 1) = N(t) + 1$$  \hspace{1cm} (13)$$

To sum up, the firing frequency of the neuron can be calculated in the recursive manner:

$$f(t + 1) = \begin{cases} 
1, & \text{if } t = 0 \\
\frac{n(t + 1) - f(t)}{n(t)} - f(t), & \text{if } t > 0
\end{cases}$$  \hspace{1cm} (14)$$

Then, the neuron firing frequency of one neuron should be compared with its neighbors to determine which neuron should be regenerated. In 3D space, one neuron has a total of six nearby neurons with the distance equalling to 1. If a neuron competes to fire after winning, it should be compared with other six neurons. Suppose the neuron A is the winner and the neuron B is one of the six nearby neurons. The following criterion is adopted to compare the firing frequencies of neurons.

$$\begin{align*}
&n_A(t) > n_0 \\
&f_A(t) > 4f_B(t)
\end{align*}$$  \hspace{1cm} (15)$$

Only the active neurons at high active level have the ability to suppress other neurons, which means that “age” (denoting its firing times) of the neuron A is already very high. In the expression (15), $n_0$ is the “age” threshold, and in this work, it is set to 40, i.e., $n_0 = 40$. In addition, the firing frequencies of the nearby neurons are also relatively low. In the paper, the firing frequencies of the nearby neurons should be less than $1/4$ of that of the neuron A. If the criterion is met, the neuron B should die and then regenerate to learn new knowledge.
2.5 How the Y Neurons Learn the Object Features

Firings of all Y neurons go through top-k competition, and only the winners can fire and update their weights, i.e., object features they have learned. In the top-k competition, neuron active degree is evaluated by its pre-action energy. Only the former k neurons with the largest pre-action energy can fire. The pre-action energy is composed of four parts: action value corresponding to the X input (bottom-up input), and action values corresponding to LM, SM and TM inputs (top-down inputs). Only when the four action values are large (match well) can the neuron fire. The firing neuron can learn the corresponding feature of the input image, and connect with the corresponding neurons in LM, SM and TM (supervised by the external environment), in other words, connecting the features with the concepts. After that, if this neuron wants to fire frequently (age increasing, more mature, learning more stable feature), it means that the feature in X layer and the corresponding concepts in LM, SM and TM must happen simultaneously, i.e., the feature in X layer is relevant with these concepts. In other words, if some feature in X layer has almost no relevance with the concepts in LM, SM and TM, this feature is difficult to be learned. Even the neuron corresponding to this feature fires a few times occasionally, its firing frequency will be not high. Since the concept is irrelevant to the feature, it cannot ensure that the four pre-action values are high. These neurons should die and regenerate through the neuron regenerate mechanism, to learn new features. Therefore, most neurons in WWNs can learn the foreground features, while few neurons can learn the background. Since compared with the foreground, the background has almost no relevance with the concepts denoted by LM, SM and TM which are provided by the external teacher.

3 Experiment Design and Analysis

As introduced in the previous section, the improved WWN-7 model will be designed to identify human faces in the complex natural backgrounds. Effects of synapse maintenance and neuron regenesis mechanism will be shown in the experiment. The internal weights will be analyzed to demonstrate the accuracy of face recognition with the proposed WWN-7 model. Comparative experiments will show the advantages of the proposed WWN-7 model for face recognition in comparison with several existing methods.

3.1 Map Library

Face pictures used in this paper are LFW Face Database, and the complex natural backgrounds used are selected randomly from the natural images whose dimension is 38×38 pixels, as shown in Figure 3. SIZES of the foreground objects range from 24×24 pixels to 34×34 pixels, with a pixel increase each time, so there are 11 kinds of sizes. Due to the different sizes, the foreground objects have different locations, so the foreground objects with smaller sizes have more locations. The location number ranges from 225 ((38-24+1)×(38-24+1)) to 25 ((38-34+1)×(38-34+1)).

Objects to be recognized are 5 different human faces and each face has 11 different sizes. So there are 55 types in total, and the face location numbers in the complex backgrounds change from 5×5 to 15×15, as shown in the Table 1. For the same face type, images with even sizes (24×24 pixels, 26×26 pixels, etc.) are used for training, while those with odd sizes (25×25 pixels, 27×27 pixels, etc.) are used for testing. There are 25 face images to be tested in experiment. For each face, it can appear all possible locations in the backgrounds. Thus the recognition rate can be calculated.

The network will be randomly assigned to a complex natural background first, then the foreground object is embedded in the background in the training phase. After the foreground objects at all locations have been trained, the WWN performs a whole train procedure.
3.2 Inputs and Outputs

Inputs from X to Y layer are introduced first. The connections between X and Y layer are local and the different receptive fields mean the different characteristic information. The network involves the receptive fields of 7×7 pixels, 11×11 pixels, 15×15 pixels and 19×19 pixels, corresponding to the four parts of Y layer, respectively. Taking the receptive field of 7×7 pixels as an example. The retina can only capture the square range of 7×7 pixels each time. Each receptive field starts from the top left corner of the image and moves down by a pixel each time. When it moves to the bottom of the image, the receptive field moves to right side by one pixel and starts to scan from the second column. This process will be carried out until the entire image has been input. The scanned data are stored in a sequential order, so that the inputs of the first part of the Y layer are obtained. Other receptive fields can replace the former one to repeat this process and the inputs of the remaining three parts of Y layer will be obtained according to the same steps.

There are some independent inputs from the three parts of Z layer to Y layer. Because of the five different human faces, input dimension of TM is 5×1. Due to the 11 different sizes of the face images, input dimension of SM is 11×1. Location number must contain all the possible locations, so the input dimension of LM is 225×1, as shown in Table 2.

Above sections illustrate that different receptive fields correspond to the different parts of Y layer. Taking the receptive field of 7×7 pixels as an example. There are 9 sub-layers in Y1 and the input image dimension is 38×38 pixels. Each small scanning frame corresponds to a neuron, so there are (38-7+1)×(38-7+1)×9=32×32×9 neurons in Y1 sub-layer. In the similar way, the neuron numbers in Y2, Y3 and Y4 sub-layers can be obtained and showed in Table 1.

| Y layer | Receptive field | Number of sub-layer | Neuron number of sub-layer | Total neuron number |
|---------|-----------------|---------------------|---------------------------|--------------------|
| Y1      | 7×7             | 9                   | 32×32                     | 32×32×9            |
| Y2      | 11×11           | 9                   | 28×28                     | 28×28×9            |
| Y3      | 15×15           | 6                   | 24×24                     | 24×24×6            |
| Y4      | 19×19           | 6                   | 20×20                     | 20×20×6            |
Y layer receives the inputs from X layer and the three kinds of inputs (type, size and location) from Z layer simultaneously. The pre-action energy of Y neuron indicates the matching degree between the input and its corresponding weight, which represents the memory of the neuron. The greater the pre-action energy is, the more consistent it is between the input and the memory information, and the easier it is for the neuron to fire. After the pre-action energy are obtained, neurons in the Y layer compete to get the output according to the top-k competition mechanism.

### 3.3 Calculation and Distribution of Neurons

Before calculating the pre-action energy, neuron input should be normalized. Variable \textit{input} and \textit{weight} represent the input and connection weight between layers, respectively, while \textit{epsilon} represents a small positive number. The normalization process can be depicted as follows:

1. Search the minimum and maximum of each row to form a vector \textit{min} and \textit{max}, respectively;
2. Calculate the difference between the maximum and the minimum. To avoid the zero denominator, a small positive number \textit{epsilon} is introduced: \textit{diff} = \textit{max} - \textit{min} + \textit{epsilon}.
3. Execute the normalization process: \textit{input} = (\textit{input} - \textit{min}) / \textit{diff}; \textit{weight} = (\textit{weight} - \textit{min}) / \textit{diff}.

After normalizing the inputs and weights, they are used to calculate the pre-action energy of the Y neurons in the next step.

Pre-action energy determines whether the neurons can fire. If it is high, the neuron will become the winner in the top-k competition and will fire, and updates its weight and age through the Hebbian learning mechanism.

Split the formula 1 into three parts and calculation of the pre-action energy of the Y neuron can be depicted as follows:

\[
\begin{align*}
      r_b(v_b, b) &= \dot{v}_b \cdot \dot{b} \\ 
      r_t(v_t, t) &= \dot{v}_t \cdot \dot{t} \\ 
      r &= (1 - a)r_b + ar_t
\end{align*}
\]

where \( r_b \) and \( r_t \) are the pre-action value of the bottom-up connection and top-down connection, respectively. Dot above the vector represents the normalization of the vector. Subscript \( b \) and \( t \) represent the bottom-up input and top-down input, respectively. \( a \) refers to the ratio of \( r_t \) to \( r \). When the network is in training state, \( a = 0.5 \). When the network is in testing state, \( a = 0 \) and the pre-action energy of Y neuron is only determined by the bottom-up input. \( r_t \) is made up of three parts: \( Z_{TM}, Z_{LM} \) and \( Z_{SM} \), and each of them accounts for 1/3.

Then the achieved pre-action energy values are ranked and the neurons ranked in the former \( k \) (\( k \) can be selected according to the network situation, in this work, \( k = 1 \)) are activated. Energy value of the activated neuron is set to 1 and those of the rest neurons are all set to zero.

### 3.4 Visualization of the Internal Weights

After the network is trained 20 times, Figure 4 visualizes the bottom-up weights of partial neurons in the four sub-layers of Y layer. Each small box in Figure 4 represents the bottom-up weight of a Y neuron and the feature that the neuron has learned. Each box is separated by a white line and the dimension represents the size of the corresponding receptive field. The black box represents the dormant neuron which is not activated and is in the initial state. In the small receptive field (i.e., \( 7 \times 7 \)) of \( Y_1 \), the faces with different sizes have more similar characteristics, so more characteristics are cumulative. In this receptive field, most features that the neurons have learned are about the object margin orientation. While in the largest receptive field (i.e., \( 19 \times 19 \)) of \( Y_4 \),
Figure 4: Visualization of the Weight Value from X Layer to Y Layer.
local or global features learned can be mainly identified. Figure 4 indicates that most neurons have learned the face characteristics.

Furthermore, the receptive field of $Y_4$ is visualized in Figure 5. Sub-figures a, b, and c represent the top-down weights of the type, location and size, respectively. The weights are normalized and they are separated by white lines. Weight dimension of single neuron in sub-figure a and b is $5\times1$ and $11\times1$, respectively, which cannot deform into a square. So the box with full zeros is set up and the weight of a single neuron is placed into it to display. Weight dimension of a single neuron in sub-figure c is $225\times1$, which can deform into a square with $15\times15$. The white spots in different positions represent the different activated neurons.

![Visualizations of the Weight Value from Z Layer to Y Layer.](image)

**Figure 5:** Visualization of the Weight Value from Z Layer to Y Layer.

Visualizations of the synaptogenic factors, and standard deviations of the synapse matching are shown in Figure 6 and each box in the figure represents a neuron. In sub-figure a, foreground contour has been automatically formed by the synapse maintenance mechanism. In sub-figure b, standard deviations of the synapse matching are constantly updated in accordance with the Hebbian learning. Black box represents the dormant neuron which is not activated.
3.5 Effect Demonstration of Neuron Regenesis

WWN-7 is “skull-closed”, with the autonomous developmental ability, so the network internal resource can be regulated dynamically in the network learning. Here, resource dynamic regulation denotes that the connection among neurons can be built up or cut off dynamically. WWN-7 has the autonomous developmental ability to strengthen the right knowledge learned and weaken the false cognition, through continuous learning.

Figure 7 visualizes the bottom-up weights of Y neurons corresponding to the receptive field with 19×19 pixels, in different training epochs. One training epoch means that WWN has learned all training foreground images in all possible locations in the complex background. The top left sub-image (epoch=0) shows the initial state of the network, and there are no connections among the neurons, which means that the current WWN has strong ability to learn new knowledge. Just like a newborn baby, it has strong plasticity. The top right sub-image (epoch=1) displays the object characteristics that Y neurons have learned after the first network training epoch. We can see that some of the current characteristics are obscure, and it is difficult to distinguish them. In other words, the current WWN is not sure whether the characteristics learned is useful and should be kept. The bottom right sub-image (epoch=5) illustrates the weights after 5 training epochs, it is clear that partial characteristics are very obvious and a few neurons have regenerated to learn new features. Through the learning, the WWN has retained partial useful features, and the neuron regenesis mechanism has begun to regulate the neuron resource. With the network learning (epoch becomes larger), more and more neurons enter learning state and learn the object features. The bottom left sub-image (epoch=10) shows the weights after 10 network training epochs, we can see that most features are very obvious, and the neurons have been activated efficiently to learn knowledge, and the suppressed neurons have been reactivated.
to learn new features. Just like the human brain can produce new neurons to learn new knowledge, through the neuron regenesis mechanism, the WWN can reactivate the suppressed neurons to regenerate and learn new knowledge. So in the limited resource space (i.e., fixed neuron number), WWN can carry out efficient and accurate learning.

Furthermore, to display the effect of the neuron regenesis, the content in red box in Figure 8 is amplified. The sub-figure (a) shows that, after one training epoch (i.e., epoch=1), the features that the neurons in the right top corner learned contain parts of foreground features and partial backgrounds. Since the background has little relation with the TM, LM and SM in Z layer, the fire frequencies of these neurons are not high. Through the neuron regenesis mechanism, these neurons will die and regenerate to learn the foreground. Sub-figure (b) displays that after 5 training epochs (i.e., epoch=5), these regenerating neurons have learned new face characteristics which contain most part of the foreground and little background.
Figure 8: The first neuron regenesis case.

Figure 9: The second neuron regenesis case.
Sub-figures (a) and (b) in Figure 9 illustrate that, after 1 training epoch and 5 training epochs, the neurons in the corresponding red box are in the learning state, but the foreground features they learned have not changed, and several neighbouring neurons learn the same or similar foreground characteristics. Sub-figure (c) shows that after 10 epochs, face features that the neurons in corresponding location learned have changed, and the neighbouring neurons have begun to learn different face features which means these neurons have experienced the procedure from dying to regenerating, and learned new foreground features.

Figures 8 and 9 illustrate the dynamic regulation of the neuron regenesis mechanism for the network internal resource: it not only makes the neurons that have learned the useless background information to regenerate and learn new foreground features (denotes in Figure 8), but also let the neurons with low fire frequencies in the neighbouring neurons learned the same or similar foreground features to regenerate and learn other foreground features, thus more foreground features can be learned (denotes in Figure 9). Therefore, the neuron regenesis mechanism is a good supplement to the Hebbian learning.

3.6 Performance Evaluation

Since the WWN-7 can learn three types of features: type, location and size, therefore, performance of WWN-7 can be evaluated from the recognition results of these three aspects. The recognition rate of the type is the proportion of the test images which are correctly identified. The recognition error of the location is the Euclidean distance between the recognition location and the real location. Recognition error of the size is the Euclidean distance between the recognition dimension and the real dimension. Recognition error of the location and size is the average of all the recognition errors of the test images. The performance of WWN-7 can be discussed from three situations: WWN-7 without the synapse maintenance and neuron regenesis, WWN-7 with the synapse maintenance only, and WWN-7 with both synapse maintenance and neuron regenesis. The diagrams of the recognition effect are shown in Figure 10.

Figure 10(a) shows that the recognition rate of the WWN-7 without the synapse maintenance and neuron regenesis mechanism can be 98.1%, while that of the WWN-7 with these two biological mechanisms can reach 99.6%. Performance of the WWN-7 with these two mechanisms is better than that of only with the synapse maintenance. But during epoch=4 and epoch=9, performance of the former is worse than that of the latter, resulting from the fact that during the early time of the neuron regenesis, the neurons that have learned part of foreground characteristics and partial backgrounds will die. And the connections between these neurons are cut off totally. Their ages are set to 1 again, and their weights return to initial states. Before these neurons learn the new features and set up stable connections, they are not conducive to the performance of WWN-7. But with the training progress, these neurons learn more stable foreground features quickly, so its total performance is better.

Figure 10(b) displays that location error of the WWN-7 with these two mechanisms is the smallest among the three cases, i.e., 1 pixel. In face images, the foreground objects (faces) have little difference, while the backgrounds are randomly added from natural images which bring great difference. With the synapse maintenance mechanism, the background interference can be suppressed effectively, so it is a great help to identify the types and detect the locations. Similarly, Figure 10(c) illustrates that the size (or dimension) error of the WWN-7 with these two mechanisms is the smallest among the three cases which is close to 1 pixel. As previously mentioned, in training phase, face images with even size are trained, while in testing phase, odd size are adopted. There exists one pixel difference between the even size (e.g., 24×24 pixels) and the corresponding odd size (e.g., 25×25 pixels), so the best result of the dimension error is close to 1 pixel.

To sum up, WWN-7 with the synapse maintenance significantly improves the face recognition effect, compared with that of without the synapse maintenance mechanism. If it is added with the neuron regenesis mechanism, its recognition performance can be further promoted but the promotion is not very obvious.
Figure 10: Face recognition rates with WWN-7 in three cases. (a) Total recognition rate, (b) location error (unit: pixel), (c) dimension error (unit: pixel).
3.7 Performance Comparison

To demonstrate the performance of WWN-7 on face recognition further, 3 classical methods: PCA+third-order nearest neighbor algorithm [12], the BP neural network [38], robust sparse representation algorithm [7] are employed to compare the effect with the WWN-7 model. In above experiment, performance of WWN-7 is evaluated from three indexes: the recognition rate of types, the identification error of locations and the sizes. However, these classical methods can only classify the face images, while the location and size of the face can’t be identified. Therefore, in this comparative experiment, only the recognition rates of the type are evaluated. The recognition results are provided in Table 2 to Table 6. The PCA dimension and the feature dimension of sparse representation in Table 2 and Table 4 are the dimensions of feature vector after extracting features.

Table 2: Relationship between PCA Dimension and Recognition Rate

| PCA dimension | Recognition rate |
|---------------|------------------|
| 50            | 0.5731           |
| 90            | 0.6752           |
| 120           | 0.7136           |
| 140           | 0.7703           |
| 160           | 0.7523           |
| 170           | 0.7344           |

Table 3: Relationship between BP Training Epoch and Recognition Rate

| Training epoch | Recognition rate |
|----------------|------------------|
| 500            | 0.6038           |
| 1000           | 0.7353           |
| 2000           | 0.7683           |
| 3000           | 0.8157           |
| 3500           | 0.8565           |
| 4000           | 0.8567           |
| 4500           | 0.8567           |
| 5000           | 0.8567           |

Table 4: Relationship between Feature Dimension of Sparse Representation and Recognition Rate

| Feature dimension | Recognition rate |
|-------------------|------------------|
| 75                | 0.6747           |
| 100               | 0.7563           |
| 125               | 0.8116           |
| 150               | 0.8327           |
| 175               | 0.8259           |
| 200               | 0.8221           |
Table 5: Relationship between training times and recognition rate of the original WWN-7

| Training times | Recognition rate |
|----------------|------------------|
| 2              | 0.9346           |
| 4              | 0.9620           |
| 6              | 0.9712           |
| 8              | 0.9728           |
| 10             | 0.9808           |
| 12             | 0.9815           |
| 14             | 0.9811           |
| 16             | 0.9813           |

Table 6: Comparison of recognition rate for the 6 methods

| Approaches                                    | Recognition rate |
|----------------------------------------------|------------------|
|                                              | Average | Standard deviation | Minimum | Maximum |
| PCA+third-order nearest neighbor algorithm   | 0.7721  | 0.0253              | 0.6743  | 0.7987  |
| BP neural network                            | 0.8457  | 0.0471              | 0.6564  | 0.8743  |
| Robust sparse representation algorithm       | 0.8385  | 0.0334              | 0.7418  | 0.8833  |
| WWN-7 without the two mechanisms             | 0.9807  | 0.0034              | 0.9741  | 0.9832  |
| WWN-7 with synapse maintenance only          | 0.9924  | 0.0068              | 0.9831  | 0.9947  |
| WWN-7 with the two mechanisms                | 0.9960  | 0.0061              | 0.9823  | 0.9977  |

In the comparative experiments, for the BP network, the neuron numbers of input layer, hidden layer and output layer are 5, 10 and 5, respectively. Each person has 11 face images, 6 images are used to train the network and 5 images are used to test (the training set and test set are the same as those used in the WWN-7, respectively). Hidden layer neurons use ‘tansig’ transfer function, and output layer neurons adopt ‘purelin’ transfer function. Network train function employs the ‘traingdx’, and its weights and thresholds are automatically initialized. Network training performance function adopts mean squared error (mse), when the mse=0.001, the network stops. According to the Table 2 and 4, PCA dimension adopts 140, and the feature dimension of sparse representation uses 150. To avoid the random error, the same recognition procedure is done 30 times with the 6 methods, thus to calculate the average, standard deviation, minimal and maximal value of the recognition rate. The final results are shown in Table 6.

Table 6 displays that the recognition rates of the former three classical methods are lower than those of the WWN-7, which result from two reasons: the first reason is the interference of the different backgrounds, and the second one is the location and size variation of the foreground. Large standard deviations imply that the former three methods cannot well adapt the changes of location and size of the foreground in complex background. Due to its biomimetic mechanism, WWN-7 with the synapse maintenance mechanism can seek out the foreground outline to recognize the foreground and the parts of the background in the receptive field. Moreover, the neuron regenesis mechanism can activate the suppressed neurons in the competition to regenerate and learn new foreground features, so the recognition rate of WWN-7 with the two mechanisms is higher than other approaches. These comparative experiments indicate that WWN-7 can well realize the recognition and detection of the types of human faces in complex backgrounds.

Finally, we compared the location error and size error between our work and the standard WWN-7 [35]. The results are provided in Table 7 and 8. From the Table 7 and 8, we can see that the improved WWN-7 model including the synapse maintenance mechanism and the neuron regenesis mechanism can effectively promote the recognition rate of the location and size of the foreground in complex backgrounds. For these two mechanisms, the synapse maintenance mechanism contributes to the performance improvements more than the neuron regenesis mechanism.
### 3.8 Computational Complexity

The previous part compares the performance of different methods on the face recognition, in this part, we theoretically compare the computational complexity of these methods.

Since the WWN is a typical feed-forward network, and the computational complexity of common feed-forward network is $O(n^4)$ [9], where $n$ denotes the number of neurons in each layer. Considering the number of layers in WWN is very small, so we can neglect its influence on the computation complexity, thus we can achieve the computational complexity of WWN, i.e., $O(n^3)$. Similarly, we can get the computational complexity of BP neural network, i.e., $O(n^3)$. While that of the PCA and robust sparse representation algorithm is $O(n^2)$, where $n$ represents the data dimension.

Based on the above discussion, we can approximately assess the computational complexity of the three approaches. For the same learning task, WWN and BP neural network have the highest computation complexity, which means a longer operation time. While PCA and robust sparse representation algorithm has the lower one, so their computation time is relatively little. Although WWN has the highest recognition rate among the three methods, how to design better network architecture to enhance its computation speed is an important research direction.

### 4 Conclusions and Future Work

This paper proposes an improved Where-What Networks (WWN) model that simulates the working mechanism of the human brain visual system for identifying human faces in complex background. The main internal mechanisms of Where-What Networks, e.g., Hebbian learning rule, receptive fields, top-$k$ competition, update rules, and so on, are explained detailedly in the work. To study the internal mechanism of the WWN-7 further, the bottom-up and top-down weights are visualized in the experiment. Based on the WWN-7 model, synapse maintenance is used for reducing the background interference while neuron regenesis mechanism is designed to regulate the neuron resource dynamically for enhancing the network usage efficiency. The performance of the WWN-7 with (without) the synapse maintenance and neuron regenesis mechanism for the face recognition in complex backgrounds are studied, and influences of synapse maintenance and neuron regenesis mechanism are analyzed. Experiment results indicate that the average recognition rate of the improve model reaches 99.6%, the location errors and the size errors are reduced by 18% and 24% when these two mechanisms are both considered. Although this model has achieved good performance, it is only for recognizing different types of faces with different locations and sizes in complex background, without considering
the angle of the faces, occlusion, which will affect the final recognition rate. Moreover, its high computation complexity is also a problem to be solved.

Future research will consider a more detailed model to simulate the human brain visual system which is extremely complex, and it will extend the current work which considers the dorsal pathway and ventral pathway only. The objects to be recognized should not be limited to the static images but extended to the moving objects in video.

Acknowledgement: This research is supported by the National Natural Science Foundation of China under Grant 61603343 and 61703372 and the Science & Technology Innovation Team Project of Henan Province under Grant 17IRTSTHN013. Scientific Problem Tackling of Henan Province under Grant 192102210256.

References

[1] K. Mehta A. Bansal and S. Arora. Face recognition using pca and lda algorithms. In Proceedings of the 2012 Second International Conference on Advanced Computing and Communication Technologies, pages 251–254, Rohtak, Haryana, India, May 14-16, 2012.

[2] N. V. Manoj Ashwin A. Lawrence and K. Manikantan. Face recognition using background removal based on eccentricity and area using ycbcr and hsv color models. In Proceedings of the International Conference on Signal, Networks, Computing, and Systems, pages 33–43, Springer, India, 2017.

[3] S. Ipson A. S. Al-Waisy, R. Qahwaji and S. Al-Fahdawi. A multimodal deep learning framework using local feature representations for face recognition. Machine Vision and Applications, (1):1–20, 2017.

[4] A. Aldhahab and W. B. Mikhail. Face recognition employing dwt followed by fastica. Circuits, Systems, and Signal Processing, DOI: 10.1007/s00034-017-0653-z, 2017.

[5] M. S. Al Ani and A. S. Al-Waisy. Face recognition approach based on wavelet-curvelet technique. International Journal of Signal Image Processing, 3(2):21–31, 2012.

[6] K. Huang C. Ren, D. Dai and Z. Lai. Transfer learning of structured representation for face recognition. IEEE Transactions on image processing, 23(12):5440–5454, 2014.

[7] Y. Hou C. Zheng and J. Zhang. Improved sparse representation with low-rank representation for robust face recognition. Neurocomputing, 198:114–124, 2016.

[8] T. Dey and D. Ghoshal. Pose invariant face recognition technique based on eigen space approach using dual registration techniques after masking. Advances in Optical Science and Engineering, 194:335–343, 2017.

[9] K. Fredenslund. Computational complexity of neural networks. https://kasperfred.com/posts/computational-complexity-of-neural-networks.

[10] Y. Gao and H. J. Lee. Viewpoint unconstrained face recognition based on affine local descriptors and probabilistic similarity. Journal of Information Processing Systems, 11(4):6–43, 2015.

[11] H. Imtiaz and S. A. Fattah. A curvelet domain face recognition scheme based on local dominant feature extraction. ISRN Signal Processing, 2012(1):4615–4621, 2012.

[13] Z. Ji and J. Weng. A developmental where-what network for concurrent and interactive visual attention and recognition. Robotics and Autonomous Systems, 71:35-48, 2015

[14] Z. Ji, J. Weng, and D. Prokhorov. Where-what network 1: “where” and “what” assist each other through top-down connection. In IEEE International Conference on Development and Learning, pages 61–66, Montreal, Canada, August 9-12, 2008.

[15] S. Ren K. He, X. Zhang and J. Sun. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), pages 770–778, Las Vegas, NV, United States, June 27-30, 2016.

[16] F. J. Galdames L. A. Cament and K. W. Bowyer. Face recognition under pose variation with local gabor features enhanced by active shape and statistical models. Pattern Recognition, 48(11):3371–3384, 2015.

[17] C. Li, W. Wei, J. Li, and W. Song. A cloud-based monitoring system via face recognition using gabor and cs-lbp features. The Journal of Supercomputing, 73:1532–1546, 2017.

[18] C. Li, W. Wei, J. Wang, W. Tang, and S. Zhao. Face recognition based on deep belief network combined with center-symmetric local binary pattern. International Journal of Multimed Ubiquitous Engineering, 354:277–283, 2016.

[19] J. Liu, C. Fang, and C. Wu. A fusion face recognition approach based on 7-layer deep learning neural network. Journal of Electrical and Computer Engineering, 2016(5786):1–7, 2016.

[20] S. Mondal and S. Bag. Face recognition using pca and minimum distance classifier. In Proceedings of the 5th International Conference on Frontiers in Intelligent Computing: Theory and Applications, pages 397–405, Springer Singapore, 2017.
[21] D. N. Parmar and B. B. Mehta. Face recognition methods and applications. *International Journal of Computer Technology and Applications*, 4(1):84–86, 2014.

[22] N. Poh R. Blanco-Gonzalo and R. Wong. Time evolution of face recognition in accessible scenarios. *Human-centric Computer Information Science*, 5(1):1–11, 2015.

[23] M. Riesenhuber and T. Poggio. Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2(11):1019–1025, 1999.

[24] A. Munir A. Nawaz S. Arshid, A. Hussain and S. Aziz. Multi-stage binary patterns for facial expression recognition in real world. *Cluster Computing*, DOI: 10.1007/s10586-017-0832-5, 2017.

[25] X. Song, Z. Zhang, and J. Weng. Where-what network 5: Dealing with scales for objects in complex backgrounds. In *Proceedings of International Joint Conference on Neural Networks*, pages 2795–2802, San Jones, California, USA, July 31-August 5, 2011.

[26] M. Sur and J. L. R. Rubenstein. Patterning and plasticity of the cerebral cortex. *Science*, 310:805–810, 2005.

[27] B. K. Tripathi. On the complex domain deep machine learning for face recognition. *Applied Intelligence*, 47(3):382–396, 2017.

[28] M. Turk and A. Pentland. Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1):71–86, 1991.

[29] J. Weng. Symbolic models and emergent models: A review. *IEEE Transactions on Autonomous Mental Development*, 4(1):29–53, 2012.

[30] J. Weng and M. Luciw. Dually optimal neuronal layers: Lobe component analysis. *IEEE Transactions on Autonomous Mental Development*, 1(1):68-85, 2009

[31] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, and E. Thelen. Autonomous mental development by robots and animals. *Science*, 291(5504):599–600, 2001.

[32] X. Wu, G. Guo, and J. Weng. Skull-closed autonomous development: WWN-7 dealing with scales. In *Proceedings of the International Conference on Brain-Mind*, pages 10–18, East Lansing, MI, USA, July 27-28, 2013.

[33] Yong Xu, Zhengming Li, Jeng-Shyang Pan, and Jing-Yu Yang. Face recognition based on fusion of multi-resolution gabor features. *Neural Computing and Applications*, 23(5):1251–1256, 2013.

[34] Mingliang Xue, Wanquan Liu, and Xiaodong Liu. A novel weighted fuzzy lda for face recognition using the genetic algorithm. *Neural Computing and Applications*, 22(7-8):1531–1541, 2013.

[35] Q. Lv Y. He, B. Jin and Y. Yang. Improving bp neural network for the recognition of face direction. In *2011 International Symposium on Computer Science and Society (ISSCCS)*, pages 79–82, Washington, DC, USA, July 16-17, 2011.

[36] J. Yin, W. Zeng, and L. Wei. Optimal feature extraction methods for classification methods and their applications to biometric recognition. *Knowledge-Based Systems*, 99:112–122, 2016.

[37] H. Mendez-Vazquez R. He Z. Chai, Z. Sun and T. Tan. Gabor ordinal measures for face recognition. *IEEE Transactions on Information Forensics and Security*, 9(1):14–26, 2014.

[38] X. Li Z. Li, D. Gong and D. Tao. Learning compact feature descriptor and adaptive matching framework for face recognition. *IEEE Transactions on Image Processing*, 24(9):2736–2745, 2015.

[39] B. Zhang and Z. Mu. Robust classification for occluded ear via gabor scale feature-based nonnegative sparse representation. *Optical Engineering*, 53(6):667–677, 2013.

[40] B. Zhang and Y. Qiao. Face recognition based on gradient gabor feature and efficient kernel fisher analysis. *Neural Computing and Application*, 19(4):617–623, 2010.