Practical emotional neural networks

Ehsan Lotfi a,∗, M.-R. Akbarzadeh-T. b

aDepartment of Computer Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran
bDepartment of Electrical and Computer Engineering, Center of Excellence on Soft Computing and Intelligent Information Processing, Ferdowsi University of Mashhad, Iran

ARTICLE INFO

Article history:
Received 12 July 2013
Received in revised form 25 May 2014
Accepted 30 June 2014
Available online 8 July 2014

Keywords:
Amygdala
BELBIC
Cognition
Emotional state
Learning
Emotion

ABSTRACT

In this paper, we propose a limbic-based artificial emotional neural network (LiAENN) for a pattern recognition problem. LiAENN is a novel computational neural model of the emotional brain that models emotional situations such as anxiety and confidence in the learning process, the short paths, the forgetting processes, and inhibitory mechanisms of the emotional brain. In the model, the learning weights are adjusted by the proposed anxious confident decayed brain emotional learning rules (ACDBEL). In engineering applications, LiAENN is utilized in facial detection, and emotion recognition. According to the comparative results on ORL and Yale datasets, LiAENN shows a higher accuracy than other applied emotional networks such as brain emotional learning (BEL) and emotional back propagation (EmBP) based networks.

© 2014 Elsevier Ltd. All rights reserved.

1. Introduction

In the literature, there are three paradigms regarding the concepts of neural networks and emotions. The first is related to emotion recognition and expression using artificial neural networks, known as affective computing (Caridakis, Karpouzis, & Kollias, 2008; Fragopanagos & Taylor, 2005; Ioannou et al., 2005; Mermillod, Bonin, Mondillon, Alleysson, & Vermeulen, 2010; Rao, Saroj, Maity, & Koolagudi, 2011). The second paradigm is the modeling of emotion as a biological process via connectionist approaches for neuropsychological issues (Frewen, Dozois, Joanisse, & Neufeld, 2008; Grossberg, 1975; Grossberg & Seidman, 2006; Levine, 2007). And the third which is less noted and will be addressed here is associated with motivation from emotion to improve or create artificial intelligence tools such as artificial emotional neural networks (Khashman, 2010). This paper aims to review and develop neural networks motivated from emotion and is concerned with the methods in which researchers have applied them successfully in various artificial intelligence-based application domains such as intelligent control, prediction and classification as well as pattern recognition. In this framework, we can explicitly address the brain emotional learning (BEL) based neural networks (Lotfi & Akbarzadeh-T, 2013a) and the emotional back propagation (EmBP) based neural networks (Khashman, 2010, 2012). Here we attempt to review them, introduce their abilities and drawbacks and to propose a powerful applied emotional neural network beneficial to engineering and real world problems.

These applied networks, each of which is associated with an emotional learning algorithm, have been produced through conceptual models called computational models of emotion (Marsella, Gratch, & Petta, 2010). BEL based networks have been created via anatomical computational models and EmBP based networks have been made via appraisal computational models of emotion. In the anatomical view, the focus is on the emotional brain. Emotional brain refers to the portions of the human brain that process external emotional stimuli such as reward and punishment received from the outside world. Emotional brain has a superior feature that is fast reacting. Researchers do not have an agreement on the source of this fast processing. Some researchers believe that this feature is obtained because of the existence of short paths in the emotional brain. Others like Pessoa (2008, 2009) argue that cortical transmission is fast enough that this short path is unnecessary and that emotional stimuli are still subject to intentional control. However, in this approach, what motivates employing models of emotional brain in engineering applications is the high speed of emotional processing possibly due to the inhibitory synapses and the short paths in the emotional brain (Lotfi & Akbarzadeh-T, 2014). It is likely that the most important characteristic of the practical models produced based on the emotional brain and especially the models including the short...
paths and the inhibitory connections is fast learning and quick reacting.

In contrast to the BEL based networks, EmBP based networks are more motivated by the appraisal approach of emotion. According to this approach, emotional states can be appraised as situational maps and can be elicited by appraisal variables. In the appraisal approach, the links between emotional states and situations is usually defined by IF-Then roles, actually the description levels higher than anatomical view are considered here. The innovative aspect of EmBP networks is applying the emotional states and situations in the learning process of artificial neural networks. Although the basic motivation behind the use of EmBP networks, like many other models in the appraisal approach, is the building of human-like agents that have emotions (Khashman, 2008). EmBP based networks have been successfully applied in various engineering applications. The present paper considers the neural networks which are motivated by emotion on one side, and those with successful engineering applications on the other side. In Sections 1.1 and 1.2 we review these networks in detail and in Section 2, present their features and introduce our aims. The proposed method is then presented in Section 3. Experimental results are evaluated through several simulations in Section 4. Finally, conclusions are drawn in Section 5.

1.1. BEL based neural networks

These networks are inspired by the anatomical findings of LeDoux (1991, 1996, 2000). The important findings of LeDoux include the characterizing of signal propagation paths in the emotional brain. He argues that due to the existence of shorter paths in the emotional brain, emotional stimuli are processed much faster than normal stimuli. This fast processing has motivated the researchers to model the emotional brain and employ the resulting models in various engineering applications. Studies of the neural basis of the emotionalbrain are described by the limbic system (LS) theory of emotion. As shown in Fig. 1, LS consists of a complex set of structures located in the cortical or subcortical areas such as (LeDoux, 1996) amygdala (AMYG), orbitofrontal cortex (OFC), thalamus, sensory cortex, hypothalamus and hippocampus. Among these structures, AMYG plays a critical role in emotional learning and reacting. AMYG stores the emotional memories and responds to each input stimulus retrieved by them. AMYG is a permanent memory (Fadok, Darvas, Dickerson, & Palmiter, 2010; Griggs, Young, Rumbaugh, & Miller, 2013; Lamprecht, Hazvi, & Dudai, 1997; Yeh, Lin, & Gean, 2004), which has a forgetting process (Hardt, Nader, & Nadel, 2013; Kim, Li, Hamlin, McNally, & Richardson, 2011) and is involved in the attention process (Bianchin, Mello e Souza, Medina, & Izquierdo, 1999; Rolls, 1992).

LeDoux argues that there are two different ways that external stimuli can reach the AMYG. One is short and fast but imprecise and comes directly from the thalamus. And the other is long and slow but precise and comes from the sensory cortex. These paths are presented in Fig. 2. Thus AMYG is properly situated to reach the stimulus extremely quickly and produce the required reaction. Thus emotional stimuli such as fear can bring about quick reactions, usually when there is no chance for the rational mind to process the danger. AMYG is the storage of emotional memories and responsible for emotional stimuli. AMYG receives reward signals in the learning process and interacts with the OFC. OFC receives connections from the sensory cortex and AMYG. AMYG responds to the emotional stimulus. OFC then evaluates AMYG’s response and tries to prevent inappropriate answers based on the context provided by the hippocampus (Balkenius & Morén, 2001; Morén, 2002; Morén & Balkenius, 2000). The AMYG–OFC model in Fig. 3 learns to react to the new stimulus based on the history of input rewards and punishment signals. Additionally, in the model, AMYG learns to associate with emotionally charged and neutral stimuli. The OFC prevents the formation of inappropriate experiences and learning connections. AMYG–OFC model consists of two subsystems which attempt to respond correctly to emotional stimuli. Each subsystem consists of a number of nodes which are related to the dimension of each stimulus. At first, the stimulus enters the thalamus part of the model to calculate the maximum input and submits it to AMYG as one of them. The OFC does not receive any input from thalamus. Instead, it receives AMYG’s output in order to update its learning weights (i.e. OFC weights $w_1, w_2, w_3$ in Fig. 3; Morén & Balkenius, 2000). Although Morén (2002) defined an internal reinforce $R_0$ to update the OFC’s weights as follows,

$$R_0 = \begin{cases} \left[ \sum A_i - R_E W \right]^+ - \sum O_i & \text{if } (R_E W \neq 0) \\ \left[ \sum A_i - \sum O_i \right]^+ & \text{Otherwise} \end{cases}$$

(1)

it is not clear how values are assigned to the REW signal while this signal plays a pivotal role in the AMYG learning process. Lucas, Shahmirzadi, and Sheikholeslami (2004) explicitly determined the reward signal REW and proposed the BEL base controller named BELBIC which has been successfully utilized in various control applications (Beheshti & Hashim, 2010; Chandra, 2005; Daryabeigi,
where $T$ is the target value, $P$ is the input pattern, $\alpha$ and $\beta$ are learning rates and $\gamma$ is decay rate in AMYG learning rule, where $T - E_\alpha$ is calculated error as an internal reinforcer and the $\max$ operator causes the monotonic learning. The added decay rate has in fact a neurobiological basis, simulating the forgetting role of AMYG and is discussed in Lotfi and Akbarzadeh-T (2014).

### 1.2. EmBP based neural networks

In contrast to the BEL based networks which are based on anatomical approaches, EmBP based neural networks emphasize the appraisal approach. Appraisal approach is the most fruitful source for designing symbolic AI systems (Marsella et al., 2010). It states that an emotion is a personal appraisal of person–environment relationship. Frijda and Swagerman (1987), Lazarus (1991), Ortony, Clore, and Collins (1988) and Scherer (2001) are among pioneers of appraisal approaches including the following models: ACRES (Frijda & Swagerman, 1987), AR (Elliot, 1992), EM (Reilly, 1996), WILL (Moffatt, Frijda, & Phaf, 1993), TABASCO (Staller & Petta, 2001), FLAME (El-Nasr, Yen, & Joerger, 2000), EMILE (Gratch, 2000), CBI (Marsella, Johnson, & LaBore, 2003), ACTAFFAC (Rank, 2004; Rank & Petta, 2005), PARLEE (Bui, 2004), EMA (Marsella & Grath, 2009), THESPIAN (Si, Marsella, & Pynadath, 2005), FEARNOT (Aylett, Louchart, Dias, Paiva, & Vala, 2005) as well as PACTIDM (Marinier & Laird, 2008). This approach emphasizes that an emotion must be appraised through situation maps. The maps have been frequently defined by If–Then roles. For example, in the FLAME model, fuzzy sets were applied to present the emotions and fuzzy rules were used to define the maps from events to emotions and emotions to actions. In this framework, some researchers investigated the impact of emotion upon learning. For example Poel, den Akker, Nijholt, and van Kesteren (2002) proposed the SHAME model to investigate the emotional states during learning, and Khashman (2008) investigated the effect of the added emotional factors on learning and decision making capabilities of the neural network. The innovative aspect of Khashman’s models (2008, 2009a, 2010) is applying the emotional states in the learning process of Multilayer Perceptron (MLP). He used anxiety and confidence as emotional states affecting the learning process and modified back the propagation (BP) learning algorithm, using them. The resulting algorithm was named the EmBP learning algorithm. EmBP has additional emotional weights that are updated using two emotional parameters: anxiety and confidence. He assumes that the anxiety level is high at the beginning of a learning task and the confidence level is low. After some time, practice and positive feedbacks decrease the anxiety level as the confidence level grows. He sets the initial confidence coefficient value to “0” and defines anxiety coefficient ($\mu$) and confidence coefficient ($k$) values:

$$\mu = Y_{AVPAT} + E$$

and,

$$k = \mu_0 - \mu_i$$

where $Y_{AVPAT}$ is the average value of all patterns presented to the neural network in each iteration, $E$ is the error feedback, $\mu_0$ is the anxiety coefficient value at the first iteration and $\mu_i$ is the anxiety coefficient value in the subsequent iteration. According to Eqs. (6) and (7), these variables (i.e. anxiety and confidence) are not independent and essentially opposite. At the beginning of a learning task, high anxiety causes more attention to be devoted to the current new learning samples. However, a greater reliance on the previously learned samples is resulted from a high level of confidence. In other words, anxiety and confidence can maintain a balance of attention between the new and previously learned data. The EmBP
Fig. 4. DouNN proposed by Khashman (2010).

Table 1
The summary of reviewed applied artificial emotional neural networks (+: feature presence, −: absence and ∗: not examined).

<table>
<thead>
<tr>
<th>Model</th>
<th>Architectures</th>
<th>Learning algorithm</th>
<th>AMYG–OFC</th>
<th>BELBIC</th>
<th>MLP</th>
<th>DouNN</th>
<th>BELPR</th>
<th>AMYG–OFC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reference</td>
<td></td>
<td></td>
<td>BEL</td>
<td>BEL</td>
<td>EmBP</td>
<td>EmBP</td>
<td>DBEL</td>
<td>(Lotfi &amp; Akbarzadeh-T, 2013b)</td>
</tr>
<tr>
<td>Neuro-psychological features</td>
<td>Long term memory</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Forgetting process</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>Inhibitory task</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Attention process</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Supervised learning</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>Emotional states</td>
<td>−</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Application</td>
<td>Simple pattern learning</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>Intelligent control</td>
<td>−</td>
<td>−</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>Classification</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>Prediction</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td></td>
<td>Facial recognition</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

profits from this mechanism and yields higher performance in the learning process.

Additionally, EmBP has been improved by adding cognitive and emotional neurons. In the resulting architecture named DuoNN (Khashman, 2010), a dorsal neuron for cognition and a ventral neuron for emotion are added to a hidden layer, as illustrated in Fig. 4. The DuoNN presented in Fig. 4 is motivated by the attention–emotion interaction model of Fragopanagos and Taylor (2006). The emotional processing occurs in a ventral network and cognitive processing is located in a dorsal network. Fragopanagos and Taylor (2006) argue that emotion can guide the individual’s attention which in turn controls cognition. In this model, a dorsal attention circuit of dorso-lateral prefrontal cortex and a ventral attention circuit are in interaction with OFC and in turn, all of them are in interaction with AMYG. In the DuoNN, dorsal and ventral circuits are modeled by two hidden neurons where the anxiety and the confidence coefficients affect the learning of the weights. Furthermore, DuoNN can successfully control attention in the learning process. In the learning algorithm, the high value of anxiety reduces the system’s attention to the derivative of errors in the output while an increase in the confidence level which is the result of a decrease in stress level means that more attention is attributed to previous adjustments.

Although the main motivation towards Khashman’s emotional modeling is to simulate the human emotions, these networks have been successfully utilized in various applications such as pattern recognition and classification (Khashman, 2009b, 2009c, 2012; Maglianesi & Stegmayer, 2012), optimization (Yang, Wang, Yuan, & Yin, 2012) and decision making (Khashman, 2011; Lim & Jain, 2010). In the next section we present the features of the networks and try to apply them in order to improve the emotional neural networks.

2. Features of the published models and the foundation of the proposed model

As it was reviewed, applied artificial emotional neural networks are inspired by emotions or based on anatomical studies or model the emotional states in the learning algorithms. Despite all the existing fundamental differences among these networks, they have two features in common: first, they have proved to be significantly effective in engineering applications and second, they have exhibited high speed and quick convergence in these applications. For example, by including emotional states, the EmBP algorithm has increased the learning speed compared to BP in facial detection, and BELPR algorithm has shown very low computational complexity. This distinguished feature is not accidental and it is also present in the biological processes relating to emotions. According to Goleman (2006), processing emotions in the brain is extremely fast which provides the capability of having quick reactions or responses.

Each of these methods holds a unique view regarding emotions and they have been successful in modeling the respective features based on it. The neuropsychological features and related applications are summarized in Table 1. For example, BELPR and ADBEL models include the features of memory decay and inhibitory connections but do not model attention and emotional states in the learning process, something that can be important from the appraisal view, and also something which is properly modeled in EmBP networks. However, these networks do not include the other neurological features of the emotional brain such as inhibitory connections which control emotional responses and the decaying process. It seems that the method employed in the EmBP algorithm for modeling emotional states in the learning process can be a turning point in including emotional descriptions in the neural processing of emotion. In other words, it can be a turning point in expressing emotional states in the form of anatomical models, instead of
If–Then rules, in the appraisal approach. Therefore, it can be employed in adding the effects of emotional states in BEL models. We have assumed the following in our ambitious view: the more a model can incorporate the behavioral characteristics of neurological emotions the more effective it can be in, firstly, explaining the biological considerations of emotions and, second, its engineering applications. The development of such methods has lead to the development of artificial intelligence instruments. On the other hand, since they are successful in solving real world problems, if they are based on neurological emotions, they are more reliable in explaining biological issues.

Here, we aim to propose a model which can simulate emotional states as well as EmBP networks on one hand, and model neurological emotional processes and BEL networks on the other hand. We expect the model to not only function as well as the mentioned networks, but to have a better application in artificial intelligence. Furthermore, we have identified areas of possible improvement within these networks and algorithms:

– Firstly, BEL models do not include the emotional states. Secondly, in BEL models, network nodes can be programmed identical to the artificial perceptron model, something which does not happen in such models. This is only possible by adding a bias to the model's nodes and using the activation function. The BEL models have properly modeled the important foundations of the emotional brain and what we do, equip them with more biological features of emotion. In fact, we add emotional states and have brought them to a lower level of processing (Kasabov, 2014) i.e. the level of neurons. Although at this level, applying a more biological type of neuron, e.g. spiking neurons (Kasabov, Dhoble, Nuntalid, & Indiveri, 2013), may be more interesting.

A simple model such as perceptron can properly introduce this issue.

– The EmBP networks are obviously different from the neurophysiological processes of the brain but it can be said that they are somehow close to the common multilayer artificial networks and they have been able to successfully model emotional behaviors by adding the related coefficients. Here, we attempt to study the way emotional states and behaviors are modeled in a model close to the brain physiology. We combine and develop these features in the form of a network model in order to obtain a more comprehensive simulation of emotions on one hand and improve the results in engineering applications on the other hand. The proposed model is discussed in Section 3.

3. The proposed limbic based artificial emotional neural networks

Here we propose a novel applied neural model of emotion to use the advantages of using emotional states in learning, like EmBP networks, and apply the inhibitory connections in the structure, like BEL networks. The proposed method named LiAENN can be used for multiple-inputs multiple-outputs pattern recognition, classification and prediction problems. The multiple-input single-output architecture of LiAENN is proposed in Fig. 5; and Fig. 6 shows the multiple-input multiple-output architecture. In the figures, the solid lines and the dashed lines present the data flow and the learning flow respectively. The input pattern is illustrated by vector \( p_{0<n+1} \) beginning the data flow and feed forward computing and \( t \) is the target value beginning the learning flow as well as backward learning computing.
3.1. Feed forward computing

The input signal \( p_{n+1} \) enters the thalamus and then goes to the sensory cortex. The AMYG receives the input pattern \( p_1, p_2, \ldots, p_n \) from sensory cortex, and receives \( p_{n+1} \) from the thalamus. \( p_{n+1} \) calculated by the following formula is the output of thalamus and one of AMYG inputs:

\[
p_{n+1} = \text{mean}_{j=1\ldots n}(p_j)
\]  

and as illustrated in Fig. 5, \( v_{n+1}s \) are related weights. Actually the \textit{mean} operator simulates the imprecise information coming from thalamus as discussed in Section 1.1. Furthermore the OFC receives the input pattern including \( p_1, p_2, \ldots, p_n \) from the sensory cortex. And according to the biological process there is not any connection between thalamus and OFC directly. AMYG and OFC are the two main subsystems. AMYG is modeled by a two layer perceptron with single output neuron and two hidden neurons, and also OFC is modeled by another two layer perceptron as presented in Fig. 5, to inhibit the AMYG responses. AMYG and OFC make two internal outputs. \( E_s \) is the internal output of AMYG and \( E_o \) is the output of OFC. They are calculated by the following formulas:

\[
E_a = f_a^2 \left( v_1, v_a \left( \sum_{j=1}^{n+1} (v_{j,1} p_j) + b_{a1} \right) + v_2, f_a^1 \left( \sum_{j=1}^{n+1} (v_{j,2} p_j) + b_{a2} \right) \right) + b_{a1}^2 \]  

\[
E_o = f_o^2 \left( w_1, f_o^1 \left( \sum_{j=1}^{n} (w_{j,1} p_j) + b_{o1} \right) + w_2, f_o^1 \left( \sum_{j=1}^{n} (w_{j,2} p_j) + b_{o2} \right) + b_{o1} \right)
\]  

where \( ba_1^1 \) is the bias of the single output neuron in AMYG, \( ba_2^1 \) is the bias of the second hidden neuron, \( ba_1^2 \) is related to the second hidden neuron, \( ps \) are the elements of the input pattern, \( vs \) are related learning weights where the superscript shows the number of layers and the subscript shows the related connection between two neurons. For example, \( v_{2,1}^2 \) is the AMYG weight in the output layer located between the second neuron in the hidden layer and the output neuron. In the equations, \( f_a^1 \) is the first layer’s activation function of AMYG and \( f_a^2 \) is the second layer’s activation function, \( b_o \) is OFC bias, \( ws \) are learning weights of OFC and \( f_o \) is the activation function of OFC. Actually Eq. (9) is the feed forward computation of AMYG as a two layer perceptron with two hidden nodes and single output node, and Eq. (10) is the feed forward computation of OFC. Finally, final output is simply calculated by the following formula:

\[
E = E_a - E_o.
\]  

Actually this subtraction implements the inhibitory task of OFC.

3.2. Learning backward computing

Firstly, the learning weights of AMYG must be adjusted. This adjustment is based on the error back propagation algorithm. Let \( T \) be the target value associated with the input pattern \( p \). Thus the error is

\[
\text{err} = T - E_a
\]  

and the sensitivity (Hagan, Demuth, & Beale, 1996) of the second layer is:

\[
S^2 = -2 \times (f_a^2)' \times \text{err}
\]  

where symbol \( \times \) denotes simple multiplication. Now, the second layer’s weights can be updated as follows:

\[
v_{1,1}^2 = v_{1,1}^2 - \alpha \times S^2 \times \left( v_{1,1}^2 \times f_o^1 \left( \sum_{j=1}^{n+1} (v_{j,1} \times p_j) + b_{a1}^1 \right) \right)
\]  

where symbol \( \times \) denotes simple multiplication. Now, the second layer’s weights can be updated as follows:

\[
v_{1,1}^2 = v_{1,1}^2 - \alpha \times S^2 \times \left( v_{1,1}^2 \times f_o^1 \left( \sum_{j=1}^{n+1} (v_{j,1} \times p_j) + b_{a1}^1 \right) \right)
\]
emotional states including anxiety and confidence states. The OFC weights are updated as follows:

\[ \text{err} = T - E_a + E_o \]  
(23)

and the OFC weights are updated as follows:

\[ v^{2}_{j,1} = v^{2}_{j,1} - \alpha \times S^2 \times \left( v^{2}_{j,1} \times f_a^1 \left( \sum_{j=1}^{n+1} (v^{1}_{j,2} \times p_j + ba^1_2) \right) \right) \]  
(15)

\[ ba^2_2 = ba^2_2 - \alpha \times S^2 \]  
(16)

where \( \alpha \) is the learning rate.

Then, the sensitivity on the hidden layer must be calculated and for the first hidden neuron is

\[ S^1 = f_a^{1'} v_{1,1}^1 \times S^2 \]  
(17)

that updates the learning weights of the first hidden neuron as follows:

\[ v^{1}_{j,1} = (1 - \gamma) \times v^{1}_{j,1} - \alpha \times S^1 \times p_j \quad \text{for} \quad j = 1 \ldots (n + 1) \]  
(18)

\[ ba^1_1 = (1 - \gamma) \times ba^1_1 - \alpha \times S^1 \]  
(19)

Similarly, for the second hidden neuron is as follows:

\[ S^1 = f_a^{1'} v_{2,1}^2 \times S^2 \]  
(20)

\[ v^{1}_{j,2} = (1 - \gamma) \times v^{1}_{j,2} - \alpha \times \mu \times S^1 \times p_j + k \times \Delta v^{1}_{j,2} \quad \text{for} \quad j = 1 \ldots (n + 1) \]  
(21)

\[ ba^1_2 = (1 - \gamma) \times ba^1_2 - \alpha \times \mu \times S^1 \times p_j + k \times \Delta ba^1_2 \]  
(22)

where \( \mu \) and \( k \) are updated by Eqs. (6) and (7) at each iteration and \( \gamma \) is decay rate in AMYG learning rule. In the proposed model, the emotional states including anxiety \( \mu \) and confidence \( k \) have been modeled only at the second hidden neuron of AMYG. We assume that all AMYG’s neurons are involved in the forgetting process and only one AMYG’s neuron includes anxiety and confidence states.

After updating the AMYG weights, OFC weights should be updated through similar steps. The OFC must be adjusted to correct the AMYG response. So the error for OFC backpropagation is

\[ \text{err} = T - E_a + E_o \]  
(23)

According to Eqs. (28)–(31), OFC does not include emotional states and decay mechanisms, because they have not been confirmed in the neuropsychological literature. Table 2 summarizes the agreements of the proposed model and the cognitive studies presented in Introduction section.

Now let us generalize the above model to a multi-input/output architecture \((n \times m)\) the number of inputs and \(m \) the number of outputs, Fig. 6 shows the result. In the proposed architecture, there are \( m \) OFC parts and \( m \) AMYG parts. The proposed model is a multi-input/output architecture that can be learned by following Anxious Confident Decayed Brain Emotional Learning Rules (ACDBEL).

In the algorithm, the inputs are the random learning weights of AMYG and OFC. And the notation \((\cdot \cdot \cdot)_i^j\) belongs to the \(j\)th AMYG and OFC parts.

**ACDBEL Algorithm**

**Step 1-AMYG Part:**

- Take \(k\)th pattern-target sample pair

\[- p_{i+1} = \text{mean}_{i=1..n}(p_i)\]

- For each output node \(i = 1 \ldots m\) do the following steps

\[ %\text{Using the following equations to calculate the output} \ E_o \]

\[ E_o = f_a^2 \left( (v^{2}_{1,1})_i \times f_a^1 \left( \sum_{j=1}^{n+1} ((v^{1}_{j,1})_i \times p_j) + (ba^1_1)_i \right) \right) \]

\[ + (v^{2}_{2,1})_i \times f_a^1 \left( \sum_{j=1}^{n+1} ((v^{1}_{j,2})_i \times p_j) + (ba^1_2)_i \right) + (ba^1_2)_i \]

Update the input weight of AMYG part \(i\), for \(j = 1 \ldots n + 1\)

\[ %\text{Calculate the error and sensitivity on second layer} \]

\[ \text{err} = T - E_a \]

\[ S^2 = -2 \times f_a^{2'} \times err \]

\[ %\text{Update the weights and bias of second layer} \]

\[ (v^{2}_{1,1})_i = (v^{2}_{1,1})_i - \alpha \times S^2 \times (v^{2}_{1,1})_i \]

\[ \times f_a^1 \left( \sum_{j=1}^{n+1} ((v^{1}_{j,1})_i \times p_j) + (ba^1_1)_i \right) \]

\[ (v^{2}_{2,1})_i = (v^{2}_{2,1})_i - \alpha \times S^2 \times (v^{2}_{2,1})_i \]

\[ \times f_a^1 \left( \sum_{j=1}^{n+1} ((v^{1}_{j,2})_i \times p_j) + (ba^1_2)_i \right) \]

\[ (ba^1_1)_i = (ba^1_1)_i \times \alpha \times S^2 \]
\( \text{calculate sensitivity on first neuron of first layer} \)
\[ S^1 = (f_a^1)' \times (v_{1,1}^1) \times S^2 \]

\( \text{Update the weights and bias connected to first neuron of first layer} \)
\[ (v_{1,1}^1)_i = (v_{1,1}^1)_i - \alpha \times \mu \times S^1 \times p_j + \beta \times (\Delta v_{1,1}^1)_i \]
\[ (ba1^1) = (ba1^1) - \alpha \times \mu \times S^1 + \beta \times (\Delta ba1^1)_i \]

\( \text{calculate sensitivity on the second neuron of first layer} \)
\[ S^1 = (f_a^1)' \times (v_{2,1}^1) \times S^2 \]

\( \text{Update the weights and bias of second neuron of first layer} \)
\[ (v_{1,2}^1)_i = (v_{1,2}^1)_i - \alpha \times S^1 \times p_j \]
\[ (ba2^1) = (ba2^1) - \alpha \times S^1 \]

- If \( k < \) number of training patterns then \( k = k + l \) and proceed to the first.
- Let epoch = epoch + 1 and \( k = 1 \).

\( \text{Update the anxiety and confidence coefficients} \)
\[ \mu = Y_{Aw/Pat} + err \]
\[ k = 1 - \mu. \]

\( \text{Update learning weight} \ \alpha \)
- if \( \text{current\_performance/previous\_perf} > 1.04 \)
  \[ \alpha = \alpha \times 0.7 \]
  else \( \alpha = \alpha \times 1.05 \) end.
- If the stop criterion has not satisfied proceed to the first.

**Step 2-OFc Part:**
- Take \( k \)th pattern-target sample pair
- For each output node \( i = 1..m \) do the following step

\( \text{Use the following equations to calculate the outputs} \)
\[ E_a = f_a^2 \left( (v_{1,1}^1)_i \times f_o^1 \left( \sum_{j=1}^{n+1} ((v_{1,2}^1)_j \times p_j) + (ba1^1)_i \right) \right) \]
\[ + (v_{2,2}^1)_i \times f_o^1 \left( \sum_{j=1}^{n} ((v_{1,2}^1)_j \times p_j) + (ba2^1)_i \right) + (ba2^1)_i \]
\[ E_o = f_o^2 \left( (v_{1,1}^1)_i \times f_o^1 \left( \sum_{j=1}^{n+1} ((w_{1,1}^1)_j \times p_j) + (bo1^1)_i \right) \right) \]
\[ + (v_{2,2}^1)_i \times f_o^1 \left( \sum_{j=1}^{n} ((w_{1,2}^1)_j \times p_j) + (bo1^1)_i \right) + (bo1^1)_i \]

- Update input weight \( j \) of OFC part \( i \), for \( j = 1...n \)

\( \text{calculate the error and sensitivity on second layer} \)
\[ err = T - E_a + E_o \]
\[ S^2 = -2 \times (f_o^1)' \times err. \]

\( \text{Update the weights and bias} \)
\[ (w_{1,1}^2)_i = (w_{1,1}^2)_i - \beta \times S^2 \times (w_{1,1}^2)_i \]
\[ \times f_o^1 \left( \sum_{j=1}^{n} ((w_{1,2}^1)_j \times p_j) + (bo1^1)_i \right) \]
\[ (w_{2,1}^2)_i = (w_{2,1}^2)_i - \beta \times S^2 \times (w_{2,1}^2)_i \]
\[ \times f_o^1 \left( \sum_{j=1}^{n} ((w_{1,2}^1)_j \times p_j) + (bo1^1)_i \right) \]
\[ (bo1^2)_i = (bo1^2)_i - \beta \times S^2 \]

\( \text{calculate sensitivity on the first layer and update the weights and bias of first layer} \)
\[ (w_{1,1}^1)_i = (w_{1,1}^1)_i - \beta \times (f_a^1)' \times (w_{1,1}^2)_i \times S^2 \times p_j \]
\[ (w_{1,2}^1)_i = (w_{1,2}^1)_i - \beta \times (f_a^1)' \times (w_{2,1}^2)_i \times S^2 \times p_j \]
\[ (bo1^2)_i = (bo1^2)_i - \beta \times (f_a^1)' \times (w_{2,1}^2)_i \times S^2. \]

- If \( k < \) number of training patterns then \( k = k + l \) and proceed to the first.
- Let epoch = epoch + 1 and \( k = 1 \)

\( \text{Update learning weight} \ \beta \)
- if \( \text{current\_performance/previous\_perf} > 1.04 \)
  \[ \beta = \beta \times 0.7 \]
  else \( \beta = \beta \times 1.05 \) end.
- If the stop criterion has not satisfied proceed to the start of Step 2.

In the ACDBEL algorithm, the learning weights \( \alpha \) and \( \beta \) are updated at the end of each iteration adaptively. See “Update learning weight \( \alpha \)” in the algorithm. According to this step, if performance is increased then the learning rate should be decreased. As mentioned in our previous work ([Lotfi & Akbarzadeh-T, 2014]), this adaptation may increase the general performance of the model. The LiAENN network is trained by the ACDBEL algorithm which is neuro-psychologically motivated and can be used in classification and prediction of problems. Although from a biological point of view, the AMYG is responsible for emotional stimuli, we can apply its artificial model LiAENN to make a response for any input patterns. The architecture presented in Fig. 5 can be used for single class classification and prediction, and the multi output architecture presented in Fig. 6 can be used for multi class classification problems.

### 4. Experimental studies

A toolbox of proposed algorithms has been prepared and is accessible at [http://bitools.ir/projects.html](http://bitools.ir/projects.html). This toolbox has been written and evaluated on Matlab2010b. Our aim in this section is to assess the application of the proposed method in facial detection and emotion recognition, and compare it with other applied emotional networks such as EmBP and BEL networks, respectively presented in Sections 4.1 and 4.2.

#### 4.1. Comparative studies with EmBP networks

The changes in the facial features complicate the face recognition task and researchers have tried to provide methods capable of recognizing human faces. DuoNN has been applied to recognize a person upon presenting his/her facial image ([Khashman, 2010]). According to the results, the emotional networks present better results than conventional networks. The adopted test bench was “ORL Database of Faces” that is accessible at [http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html](http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html). As illustrated in Fig. 7, the ORL database includes ten different images of 40 people with different genders, ethnicities, and ages. Here we compare our method with DuoNN and EmBP on ORL dataset. All testing conditions are the same as reported by [Khashman (2010)]. For example, the image size is 10 × 10 pixels, the patterns size is 100, the training samples number is 200, the testing samples number is 200 and the initial weights are randomly selected between \([-0.3, 0.3]\). Table 3 shows the parameters used in the learning phase.

The stopping criterion in learning process is to reach a certain error that is 0.007. Fig. 8 presents the error of the first 10 learning epochs for three methods. According to Fig. 8, the weights of the proposed method rapidly converge during the first 2 epochs.
Fig. 7. Examples of ORL images. Source: From Khashman (2010).

Table 3
The learning parameters used in ORL face recognition using DuoNN, EmBP and LiAENN.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DuoNN and EmBP</th>
<th>LiAENN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input neurons</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Hidden neurons</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td>Output neurons</td>
<td>40</td>
<td>40</td>
</tr>
<tr>
<td>Learning coefficients</td>
<td>0.0059</td>
<td>0.0059</td>
</tr>
<tr>
<td>Random initial weights range</td>
<td>-0.3 to +0.3</td>
<td>-0.3 to +0.3</td>
</tr>
<tr>
<td>Convergence error</td>
<td>0.007</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Confidence after learning, LiAENN presents a significant improvement in terms of accuracy and time complexity.

4.2. Comparative studies with BEL networks

In order to investigate the role of anxiety and confidence coefficients, the proposed ACDBEL can be compared with a BEL network such as BELPR. The structure of ACDBEL and BELPR are similar and what differentiates them is applying the emotional coefficients while the ACDBEL is closer to biological features. ACDBEL profits from both the emotional and decaying coefficients while BELPR just uses decaying coefficient. Here we utilize LiAENN to classify the Yale dataset and compare it with the model free BELPR. The dataset contains 165 grayscale images in GIF format of 15 individuals. There are 11 images per subject, one per different facial expression or configuration: center-light, w/glasses, happy, left-light, w/no glasses, normal, right-light, sad, sleepy, surprised, and winking. Here, the first 8 images of each class are used for training and the remaining images are used for testing. The examples of Yale are presented in Fig. 9. The dataset includes 15 classes. The targets should be encoded with binary numbers (i.e. 15 binary number for 15 classes) and the input pattern should be involved through one feature extraction step.

In the literature, a dimensionality reduction step (Hussein Al-Arashi, Ibrahim, & Azmin Suandi, 2014) has been applied to classify this dataset (Hussein Al-Arashi et al., 2014; Yin, Jiao, Shang, Xiong, & Wang, 2014). A well-known method for the dimensionality reduction is principal component analysis (PCA; Chen & Xu, 2014; Chen, Zheng, Xu, & Lai, 2013; Hussein Al-Arashi et al., 2014). PCA normally transform the original features space to a lower dimensional feature space and can be considered as a preprocessing step. The PCA calculates the data covariance matrix and then finds the eigenvalues and the eigenvectors of the matrix. According to the PCA algorithm, the only terms corresponding to the $K$ largest eigenvalues are kept. Here we used PCA as a preprocessing step. The resulting reduced feature vectors then can be considered for classification. Here we used the first 100 features for classification. So the number of input neurons is 100, the number of hidden neurons is 30 and the output neurons is 15. The learning and structural parameters of the methods are presented in Table 5. The number of input neurons depends on the number of attributes and the number of output neurons depends on the number of classes in each dataset. All learning and testing conditions of ACDBEL and BELPR are the same. The values $\alpha$ and $\beta$ are set at 0.01 and the values $\gamma = 0$. In the learning process, the stop criterion is the maximum epoch.
Table 4
The comparative results of ORL face detection using three methods with stop criterion error = 0.007.

<table>
<thead>
<tr>
<th>Model</th>
<th>EmBIa</th>
<th>DuoNNa</th>
<th>LiAENN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anxiety coef.</td>
<td>0.011425</td>
<td>0.011423</td>
<td>0</td>
</tr>
<tr>
<td>Conf. coef.</td>
<td>0.461276</td>
<td>0.516024</td>
<td>1</td>
</tr>
<tr>
<td>Iterations</td>
<td>16307</td>
<td>5030</td>
<td>1879</td>
</tr>
<tr>
<td>Correct...</td>
<td>75%</td>
<td>89.5%</td>
<td>100%</td>
</tr>
<tr>
<td>Correct...</td>
<td>70.5%</td>
<td>79.5%</td>
<td>99.5%</td>
</tr>
<tr>
<td>Correct...</td>
<td>72.75%</td>
<td>84.5%</td>
<td>99.5%</td>
</tr>
</tbody>
</table>

a From Khashman (2010).

Table 5
The learning parameters used in Yale dataset classification using BELPR and LiAENN.

<table>
<thead>
<tr>
<th>Model</th>
<th>BELPR</th>
<th>LiAENN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input neurons</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Hidden neurons</td>
<td>30</td>
<td>30</td>
</tr>
<tr>
<td>Output neurons</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td>Learning coeff.</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Random init.</td>
<td>−0.3 to +0.3</td>
<td>−0.3 to +0.3</td>
</tr>
</tbody>
</table>

(i.e. the maximum number of learning epochs has been reached). The maximum and minimum values of the inputs have been determined and the normalized data (between 0 and 1) have been used to adjust the weights. Here, the training and the testing approaches are repeated 5 times and the average accuracy is recorded.

Table 6 presents the average correct detection percentage obtained from BELPR and LiAENN during 5 executions. The average results indicated in Table 6 are based on the Student’s t-test with 95% confidence. It is obvious that ACDBEL with anxiety and confidence rates can significantly improve the accuracy during the learning epochs. By increasing the number of learning epochs, the accuracy of ACDBEL is increased significantly. The higher accuracy in 10,000 and 15,000 epochs is obtained from ACDBEL. According to Table 6, ACDBEL shows a higher accuracy than BELPR based on correct detection percentage.

The mean squared error (MSE) of the results during learning epochs of ACDBEL are presented in Fig. 10. The left side of Fig. 10 shows the performance measure in the AMYG learning step of ACDBEL. The right side of Fig. 10 shows the performance measure of OFC learning step of ACDBEL. As illustrated in the figures, MSE < 0.1 is obtained during the first 10 epochs of each two parts of the model and it is shown that these parts can effectively increase the learning accuracy.

Table 7 summarizes the percentage improvement of our method in two datasets ORL and Yale. Our method improves the previous emotional model results. The best detection accuracy of the ORL and Yale datasets are 79.50% (Table 4) and 63.94% (Table 6) respectively obtained from DuoNN and BELPR. ACDBEL improves the recognition accuracy about 24.5% and 25.2% respectively. The
percentage improvement of our method is summarized in Table 7 which is calculated through the following formulas:

\[
\text{Percentage improvement} = \frac{100 \times (\text{proposed method result} - \text{compared result})}{\text{compared result}}. \tag{32}
\]

The increased accuracy of the proposed model, compared to DuoNN, is due to the use of OFC–AMYG inhibitory structure along with emotional coefficients. The OFC output corrects the final response and thus increases the accuracy of the model. Also, the increased accuracy of the proposed model compared to BELPR is because of incorporating the emotional coefficients of anxiety, something which is neglected in BELPR. Although the emotional factors incorporated in Eqs. (21) and (22) are inspired by biological features but from the viewpoint of artificial learning algorithms, they are a kind of variable learning coefficient which performs dedicated learning speed adjustments. In other words, they increase the learning speed of the new samples at the beginning of the learning process and also draw the network’s attention to prior learnings and reduce changes in the weights.

5. Conclusions

A novel applied computational model of emotion named LiAENN is presented here. The learning weights of LiAENN are adjusted by the proposed ACDBEL algorithm. In contrast to BEL based networks, ACDBEL considers the emotional states and in contrast to EmBP based networks, it incorporates the anatomical bases of emotion. Actually, a common appraisal-anatomical modeling of emotion is applied here to produce LiAENN architecture with ACDBEL learning algorithm. The emotional states applied in the learning algorithm are anxiety and confidence and the anatomical features which have been utilized in the architecture are fast and imprecise paths in the emotional brain, inhibitory task of OFC and forgetting process of AMYG. The toolbox of the model has been written on Matlab2010b and is accessible at http://www.bitools.ir/projects.html. In numerical studies, LiAENN was utilized to recognize the facial datasets ORL and Yale. According to the results, the performance of proposed method is higher than that of the EmBP based emotional networks in facial detection. Furthermore, in the Yale classification problem, LiAENN with ACDBEL learning algorithm is more accurate than BELPR. From another point of view, the added anxiety and confidence states have in fact a neuropsychological basis and yield better learning and reacting as illustrated here. ACDBEL is based on neurophysiological aspect of the emotional brain and also the appraisal situation of emotion, is model free and can be used in classification, prediction and facial recognition problems. The proposed method can adjust its own weights in an online manner and correct the wrong answers simultaneously. Also, the learning rates \( \alpha \) and \( \beta \) which were constants in the previous BEL based algorithms are adaptively updated during the learning process. The proposed method is general which can accommodate various applications and it can be improved in many respects. In our future work, we concentrate on the operators in this model. The neurophysiological study of the brain indicates that there is a type of uncertainty in the behavioral and physiological functions of emotion. The operators of the neural networks are founded based on the lowest level of this uncertainty. Operators of addition, subtraction and multiplications do not exist in the physiological behaviors of the brain the way they are used in the models. We intend to incorporate this uncertainty in our proposed model by the placement of the fuzzy operators including t-norm and s-norm, so that the model acts more similarly to the brain’s behavior. We hope that the resulting model will be more successful in engineering and artificial intelligence applications.

Lastly, although the proposed model is inspired by the biological features of the emotional brain and facial datasets were the focus of the assessments, being model-free (a feature inherited from BELPR models) can greatly extend its applications in areas such as prediction functions. This is something which can be considered in future studies.

Acknowledgments

The authors thank the reviewers for their excellent feedback on the paper.

References


