Pattern Separation Network Based on the Hippocampus Activity for Handwritten Recognition

NAZAL MODHEJ¹, AZAM BASTANFARD², MOHAMMAD TESHNEHLAB³, (Member, IEEE), AND SOMAYEH RAIKESDA³

¹Department of Computer Engineering, Karaj Branch, Islamic Azad University, Karaj 31485-313, Iran
²Department of Control, Electrical and Computer Engineering, K. N. Toosi University of Technology, Tehran 16468-74415, Iran
³Department of Biomedical Engineering, Qazvin Branch, Islamic Azad University, Qazvin 15195-34199, Iran

Corresponding author: Azam Bastanfard (bastanfard@kiau.ac.ir)

ABSTRACT

Reaching high accuracy in handwritten character recognition is an essential challenge since it is widely used in many fields such as signature analysis and forgery detection. Recently, deep learning has demonstrated efficiency in this field. The problem with deep learning is that it uses a vast number of parameters that require a large dataset for training. To overcome this problem, an intelligent network is proposed in this study, based on the computational function of the dentate gyrus of the brain’s hippocampus. The ability to separate patterns with high overlapping is a task that is referred to as the dentate gyrus. Handwritten character images have high overlapping due to various writers’ styles, or even one writer’s style under different conditions. Therefore, proposing a network based on the dentate gyrus’ functional computation can be useful in this field. One of the prominent features of the proposed network is employing two excitation steps and two inhibition steps, augmenting the accuracy of recognizing handwritten characters. The proposed network was evaluated with six datasets of digits and characters from five languages. Experiments on all of the used datasets showed promising results. Moreover, a comparative and detailed analysis of the proposed network with other SOM-based and deep learning methods is provided. Experimental results show a significant boost in accuracy. While the character error rate (CER) was smaller than 1.85% for all the experiments, the smallest CER of 0.6% was achieved by the MNIST dataset. Moreover, in recognizing patterns with high noise, the proposed network showed satisfactory results.

INDEX TERMS

Hippocampus, excitation, inhibition, intelligent network, unsupervised learning.

I. INTRODUCTION

The increasing availability of data is changing the way decisions are made in the industry [1]. Recognition of patterns has always been of interest to researchers due to its wide range of applications. Psychological science [2] showed the significance of using handwriting improving comprehension and personal performance.

Handwritten texts comprise various trajectory shapes forming the letters depending on their lexicon and writing script [3]. In recent years, handwritten digit recognition has become an active research topic. High variation in the handwriting of different individuals is commonplace. Recognition of unconstrained handwritten digits has become challenging. It is still considered an open research topic for the community of document analysts [4] because a variation of handwritten has been observed. There are variant writers; everyone has his handwriting style; even this variation is observed for one writer at different times or in certain mental conditions. Automated handwriting recognition aims to reserve old manuscripts and create electronic libraries of digitized handwritten documents [5]. It is useful in many other areas like forgery detection and signature analysis [5].

Several solutions have been proposed for the problem of handwritten character recognition. Finding how a computer can imitate human functions, such as reading, writing, and viewing objects, has always been impressive. Many approaches based on human activity have been proposed,
such as deep learning-based classification algorithms, artificial neural networks, support vector machine classifiers, and Self-Organization Map (SOM). These methods have been applied to recognizing handwritten characters as well. Deep learning methods have attracted a lot of attention [6] and have brought great success in diverse application domains [7]. There are many branches of deep learning [8]. In many applications, deep learning can achieve a near or even better performance than the human level [5]. However, training a deep learning model is an extremely time-consuming task since it often involves many parameters [9]. The critical point is that deep learning algorithms, such as deep Convolutional Neural Networks (CNN), depend on the availability of large labeled datasets [10]. It is difficult to obtain sufficient labeled data to build an efficient classifier because manual data labeling is highly time-consuming [11].

To overcome these methodological problems, a network based on the functionality of Dentate Gyrus (DG) of the hippocampus is proposed in this study. The mammalian hippocampus is an essential region of the brain, which is necessary for the encoding of spatial and episodic memories [12]. Pattern separation is a component process of episodic memory which touches on, encodes, and resolves inference between memories [13] and is supported by the DG of the hippocampus [14], [15]. So, modeling the functionality of pattern separation in this part of the brain can help recognize handwritten images with high overlap. This study proposed a model based on the computational functionality of the DG as a network. As will be explained in the following sections, learning in the proposed network follows the policy of “winner take all.” For this reason, competitive learning is used for training in the proposed network. Since SOM uses the same learning algorithm, the major focus in this paper is comparing the proposed method with other SOM-based methods in recognizing handwritten digits. Also, the proposed network was compared with deep learning algorithms due to being widely used in character recognition. The proposed network has the responsibility of recognizing the handwritten digit and character images with high overlapping. The results of the simulation show this expected higher accuracy in comparison with other SOM-based and deep learning methods. The proposed network is evaluated using the six datasets of five languages. The results obtained from the experiments are compared with SOM, SOM-based methods such as deep SOM and deep learning methods. The results show the superiority of the proposed network over the compared methods both in the accuracy and the number of training of the network. The remainder of this paper is organized as follows: Section II summarizes previous related work. The structure of the brain’s hippocampus and the functionality of the DG of the hippocampus are indicated in Section III. The full description of the proposed network is provided in Section IV. The experimental results are represented in Section V. Section VI is dedicated to the discussion. Finally, Section VII contains the conclusion and suggestions for future work in this field.

II. RELATED WORK
Since the proposed network is evaluated in comparison with SOM-based and deep learning methods in handwritten recognition, this section focuses on these methods, which have been used for handwritten recognition.

A. SOM-BASED METHODS
Self-Organization Map is a useful tool for exploring data [15]. It is a popular data mining tool aimed at mapping high dimensional data into a lower-dimensional representation space [16], usually a 2-dimensional grid [17]. Over the years, SOM has proved to be a powerful and convenient tool for clustering and visualization data [18] like that of handwritten character recognition. The SOM algorithm is classified as an artificial neural network algorithm with a visual [19] and an unsupervised vector quantization algorithm [19], [20]. The unsupervised feature of SOM is its superiority to other supervised methods such as deep learning algorithms, which have shown acceptable results. The critical point is that deep learning algorithms such as deep Convolutional Neural Networks (CNN) are dependent on the availability of large labeled datasets [21]. It is difficult to obtain sufficient labeled data to build an efficient classifier because of the lack of technical support from experts and the time-consuming process of manual data labeling [11].

Different versions of SOM have been employed in handwritten recognition. In [22], a hybrid method, combining the convolutional neural network and SOM (CNN-SOM), was introduced for the classification of data. In this research, a rejection strategy was also presented to change the topology of the map during the training phase for improving the reject quality. In [23], an improved version of the adaptive subspace self-organizing map (ASSOM) [24] for adaptive offset subspace SOM is proposed, which is named AOSOM [23]. AOSOM can learn a set of ordered subspaces. The linear manifold SOM (LMSOM) was proposed in [25]. It considers offsets of linear manifolds from the origin and aims to learn linear manifolds by minimizing a projection error function in a gradient-descent fashion [25]. In reference [26], a mixture model under the self-organizing framework was proposed, applied to the recognition of handwritten digit images based on mixtures of linear manifolds [26]. The major drawback of SOM is its limited capability of high-level feature abstraction due to the shallow structure [27]. To overcome SOM’s drawback, the deep SOM (DSOM) algorithm [28] was explored. DSOM consists of multiple layers alternating the self-organizing and the sampling layer [28]. In DSOM, the self-organizing layer consists of certain numbers of maps, each of which receives input from a local region on its input map.

In contrast, the sample layer is organized by the winning neuron from each SOM in a self-organizing layer. The sample layer prepares data for a second self-organizing layer [28]. DSOM has many advantages, such as the ability to visualize [29], optimization [30], understandability...
and interpretability [31]. The whole DSOM is trained with a supervised self-organizing learning algorithm [28] that depends on the availability of labeled data [21]. DSOM is computationally expensive, and like traditional SOMs, it has the problem of static map size [32]. Reference [32] presents an easily parallelizable Deep Self-Organizing Maps architecture (PP-SOM). However, with the main intention of improving the computational time of DSOM while retaining the performance. Also, PP-SOM partially overcomes the static map size problems in traditional SOM and DSOM [32]. An extended architecture of initially DSOM called E-DSOM is presented in [21]. E-DSOM enhances the DSOM in two ways: the learning algorithm is completely unsupervised, and the architecture learns features of different resolutions in parallel in a single hidden layer [21]. Although the above deep algorithms show promising results, the time of recognition is still increased due to their deep structure. The notion of similarity between objects or examples plays a critical role in several machine learning tasks [33]. In SOM, the similarity of patterns is calculated in terms of Euclidean distance to perform the Best Matching Unit (BMU). In DSOM, each input image is segmented into smaller local regions that are sent to its SOM unit in that layer [28]. The entire hidden layer needs to do this process, which results in reducing the speed.

In E-DSOM, hidden layers contain several parallel layers and each of which has its SOM phase. For each parallel layer, a feature map is created and then feature maps are combined to generate a feature map [21]. However, this architecture decreases the speed of recognition because of the hidden layers. In other words, the hidden layer is still a challenge.

A new CNN architecture based on morphological filters was proposed in [43]. For the sake of convenience, the new presented architecture is called morphological CNN (Morph-CNN). More specifically, an interpretable architecture of a CNN, which uses morphological filters in the convolutional layer, was investigated [43]. Enhancing the convolutional neural network was also considered in the studies. For example, Drop Sample, as a new training method, was introduced to enhance deep convolutional neural networks for large-scale unconstrained handwritten character recognition [44]. A deep CNN architecture, DIGInet, was proposed for recognizing the above formats of digits. Moreover, the capability of deep CNN is explored in [2].

Convolutional neural network has been used in combination with other methods too. A learning method based on convolutional neural network is presented in [45]. This method, as a powerful feature extraction with support vector machines (SVM) as a classifier, is used for recognizing handwritten characters in documents.

As discussed above, the problem of deep learning-based methods is that they depend on the availability of large labeled datasets [10].

To overcome the problems of SOM-based and deep learning-based methods, this paper proposes a novel intelligent algorithm for handwritten recognition based on the DG learning model of the brain’s hippocampus.

Pattern separation in a given set of patterns is one of the issues in pattern recognition. A fundamental problem in the pattern recognition is the matching of different patterns with patterns previously represented in a system. In many methods, this process is performed by measuring the similarity of patterns. The network proposed in the current study overcomes the problem of SOM-based and deep-based methods by modeling pattern separation of the DG of the hippocampus of the brain.

III. PATTERN SEPARATION IN THE HIPPOCAMPUS

The pattern separation in the brain is done by a part of the hippocampus called the DG [46], [47]. The hippocampus has an essential role in memory [48], [49]. This area of the brain is responsible for certain functions, including episodic memory activity [50], [51], the ability to collect past experiences and respond to them, and the ability to display events based on time [48]. The hippocampus consists of three major parts: DG, CA1, and CA2 [52]. The entorhinal cortex (EC) is the entry to the hippocampus [53], [54], and DG is the first part of the Hippocampus receiving input from the other parts of the brain, mainly from the EC [55].

As explained previously, the hippocampus plays an essential role in episodic memory. Humans have similar experiences with general features, and by having the episodic memory, they are capable of delivering a unique personal presentation of events that have occurred with details [56].

1 Cornu ammonis
The ability to distinguish between similar episodes is the main characteristic of episodic memory. DG in the hippocampus performs this process through certain operations called pattern separation [57]. Different classifications of the cells and various parts of DG have been presented so far [58]. Generally, the DG’s most important cells are mossy cells, inhibitor cells, and hilar perforant path (HIPP) associated cells [52].

Granule cells (GC) and mossy cells have an excitatory effect, and DG is the only part of the brain that includes two excitatory cells [53]. Figure 1 demonstrates the structure of the relationship between the DG components. As depicted in Figure 1, granule cells receive inputs from the perforant path (PP). Mossy cells have an excitatory effect on granule cells, while the HIPP cells have an inhibition effect on them. Several studies have been carried out in the field of pattern separation of the brain, and various computational models have been proposed. The network structure proposed in the current study is modeled based on the behavioral computational model presented in [52], [59]. The basic computational model of DG was presented in [52] to evaluate the hypothesis of pattern separation. In [59], a basic circuitry of DG and CA3 based on the standard model of [52] was presented. The granule and mossy cells are excitatory cells and, the internal inhibitory cells and HIPP cells act as inhibitors [52]. Each trial begins when a new entry is activated in the PP, and this pattern is displayed for the DG. This executive sequence ends with a reading of the granule cells’ response to that input pattern [59]. Each granule cell j that remains active after all the excitations and inhibitions produce the output \( V_j \) (suprathreshold). For all the remaining granule cells, \( V_j = 0 \) (subthreshold). By strengthening the pattern separation, the hippocampus will store new information under new or unpredictable conditions [60]. Based on the process described above, each granule cells’ activity can be summarized as follows [52]:

\[
V_j = V_{rest}(j) + g_{e-pp}(j) + g_{int}(j) + g_{e-mc}(j) - g_{e-HIPP}(j)
\]

(1)

where, \( V_j \) represented the potential of granule cell \( j \). \( V_{rest}(j) \) is the remaining potential granule cell, \( g_{e-pp}(j) \) is the excitation from PP and defined as follows [53]:

\[
g_{e-pp}(j) = \sum_i y_i w_{ij}
\]

(2)

in relation 2, the activity of the \( i \)th afferent represented as \( y_i \) and, the strength of the synaptic connections between \( i \)th afferent and the granule cell \( j \) served as \( w_{ij} \). \( g_{int}(j) \) is the inhibitory conductance to granule cell \( j \) provided by the interneurons, calculated by (3) [52]:

\[
g_{int}(j) = \beta_{INT} \max_k \left( V_i y_i \right) 
\]

(3)

where \( j \) is the granule cell in the \( x \)th layer, and \( \max_k \) returns the \( k \)th maximum value. In the current simulations, \( k \) is considered one. Each interneuron is activated by granule cells in the layer. In turn, it is projected back to inhibit granule cells in that layer, implementing a form of “winner-take-all” competition. \( \beta_{INT} \) is a constant value governing the strength of lateral surround inhibition. \( g_{e-mc}(j) \) indicates the excitation from the mossy cell. Equation (4) represents the calculations performed by this section [52]:

\[
g_{e-mc}(j) = \beta_{Mc} \sum y_m w_{mj}
\]

(4)

where \( y_m \) represents the activity of the \( m \)th mossy cell. The weight of the connections between the \( m \)th mossy cell and the \( j \)th granule cell is considered \( w_{mj} \). This weight is 1 for the cells that receive input. Otherwise, it is 0. \( \beta_{Mc} \) is a constant value. \( y_m \), or the activity of any mossy cell, is calculated as follows [52]:

\[
y_m = \sum_j y_j
\]

(5)

where the activity of the \( j \)th granule cell is represented as \( y_j \). The inhibitory input of the hilar interneurons is served as \( g_{e-HIPP}(j) \) and is calculated as follows [52]:

\[
g_{e-HIPP}(j) = \beta_{HIPP} \sum_i y_h w_{ij}
\]

(6)

where \( y_h \) indicates the activity of the HIPP cell. \( w_{ij} \) shows the weight between this cell and the \( j \)th granule cell. Moreover, \( w_{ij} \) is considered one for all the granule \( j \) cells that receive input from \( h \), and zero for the rest of the cells. \( \beta_{HIPP} \) is a constant value; the high amount of this factor silences most of the granule cells. Also, \( y_h \) shows the activity of each HIPP cell [52]:

\[
y_h = \sum w_{ih} y_i
\]

(7)

where \( y_i \) is the \( i \)th PP afferent activity, and \( w_{ih} \) is the strength of the synaptic connections between \( i \)th afferent and the HIPP cell \( h \).
IV. PROPOSED NETWORK STRUCTURE

The proposed network separated input digit handwritten images based on the computation performed in the DG. The human brain exploits a computational process to accomplish this separation of patterns with high overlapping, as explained in the previous section. The proposed network is capable of performing sparsity of various handwritten digit images with a high level of overlapping using several steps. In the proposed network, two excitation steps and two inhibition steps occur. Granule cells are an essential part of the proposed network, directly connected with the input (handwritten images), mossy cells, and HIPP cells. Granule cells are arranged in separate layers [61]. The final value of these granule cells determines the network output.

The cells involved in this network are the granule cells, interneurons, mossy cells, and HIPP inhibitory cells. For two input digit images with a high level of overlap, the corresponding activity of granule cells is similar, and the same granule cell is silenced. For digit images with low overlapping, the opposite will occur.

The proposed network includes six main steps. The structure of the proposed network is depicted in Figure 2. As shown in Figure 2, granule cells are the essential cells in this network affected by the input cells, internal inhibitor cells, mossy cells, and HIPP cells, and ultimately determine the network output. These cells are excited by the input and mossy cells and inhibited by the internal inhibitor cells and HIPP cells. Therefore, the proposed network employs two excitation steps \((g_{\text{e-pp}})\) and \((g_{\text{e-mc}})\) and two inhibition steps \((g_{\text{int}})\) and \((g_{\text{i-HIPP}})\) that are a prominent feature. These two levels of excitation and two levels of inhibition increase the accuracy of recognition.

In the following sections, various stages of the proposed network in Figure 2 with the feedforward equations are entirely described in steps one to five. These equations are obtained based on the computational equations of DG described in section III. The feedback equation of the proposed network is defined in the sixth step.

A. STEP 1: EXCITATORY FROM THE INPUTS

Initially, inputs (handwritten images) are given to the granule cells. These inputs will have an excitatory effect on the granule cells in the initial step. Figure 3 illustrates the details of this section of the network’s architecture. As demonstrated in Figure 3, granule cells are arranged in several layers. There is a weight between each of the granule cells in each layer and each of the inputs. This weight indicates the strength between their connections. The output of this part is calculated based on relation 2 as follows:

\[
g_{\text{e-pp}}(j) = \sum_{in} S_{in} W_{S_{in}ij} \tag{8}\]

where \(S_{in}\) is the two-dimensional network handwritten images as inputs of the proposed network, and \(W_{S_{in}ij}\) represents the weight between the input and the \(j\)th cell of the \(i\)th layer. After applying this excitation, an activation function is applied to the output of this step. Since the proposed network is used for pattern detection, the Gaussian function is the most suitable option. Equation (9) indicates the Gaussian function.

\[
f(\text{net}) = e^{-\frac{(\text{net} - c)^2}{2\sigma^2}} \tag{9}\]

by applying equation (9) to \(g_{\text{e-pp}}\), the following equation, i.e., the activation of granule cells, is achieved:

\[
f\left(g_{\text{e-pp}}(j)\right) = e^{-\frac{1}{2}\left(\frac{g_{\text{e-pp}}(j) - c}{\sigma}\right)^2} \tag{10}\]

where \(c\), and \(\sigma\) are the center of the function and the standard deviation, respectively.

B. STEP 2: INHIBITION FROM INTERNEURONS

The second step of the proposed network is to turn off some of the granule cells in each layer using the interneurons. This step is the first inhibition in the network. The purpose of this stage is to transfer only some of the granule cells, rather than all of them, to the next step. Figure 4 illustrates step 2 of the proposed network. In Figure 4, the granule cells that have
been silenced are shown in dark colors. In each layer, some of the granule cells are silenced. The rest of the operations will only be performed on the cells that have survived from local inhibition. Removing several cells from the recognition process will enhance the operation of the next steps, and, overall, increase the speed of recognition. The functioning of this step of the proposed network is defined as:

\[ g_{\text{int}}^i(j) = \beta_{\text{INT}} \max_k (\forall_j Y_j) \]  

where \( g_{\text{int}}^i(j) \) demonstrates the \( j \)th granule cell’s excitation in the \( i \)th layer. \( \max_k (\forall_j Y_j) \) returns the value of the \( k \)th maximum value in granule cells in layer \( i \).

### C. STEP 3: EXCITATION FROM THE MOSSY CELLS

This step reflects the effect of granule cells on mossy cells, and in contrast, the excitation of mossy cells against the same granule cells. Figure 5 depicts the hierarchy of operations performed in this step. As illustrated in Figure 5, mossy cells excite the granule cells that have survived from the previous inhibition, shown in white color in Figure 5. Calculations of the excitation of granule cells by mossy cells are straightforward and are performed as follows:

\[ g_{\text{e-mc}}^i(j) = \beta_{\text{Mc}} \sum Y_m W_{ij} \]  

where \( g_{\text{e-mc}}^i(j) \) demonstrates the \( j \)th granule cell’s excitation in the \( i \)th layer. \( Y_m \) is the activity of the \( m \)th mossy cells, which is calculated as follows:

\[ Y_m = \sum_j y_j \]  

where \( y_j \) is the activity of the \( j \)th granule cell. The movement of this step will make the remaining active granule cells to be stronger than before. The critical point at this step is that some granule cells release their effects on the mossy cells, and, in response, the mossy cells will excite only the same granule cells. The question of which cells to choose for this step is fundamental. The Poisson distribution is employed to select these cells in the proposed network.

### D. STEP 4: INHIBITION FROM THE HIPP CELLS

At this step, the second and final steps of inhibition are applied to granule cells. At this step, initially, HIPP cells calculate their activity by using (7). Figure 6 depicts this step. As illustrated in Figure 6, there is a weight between every single input and each HIPP cell. The activity of each HIPP cell is calculated as:

\[ Y_h = \sum_{kp} W_{hn}^i S_n \]  

for input (\( S_n \)), the weight of the connections between the input and the \( p \)th HIPP cell (\( h_p \)) is presented as \( W_{hn}^i \). Step 1 and this part of step 4 can be performed parallel, increasing the speed of the operations in the proposed network.

After calculating \( Y_h \), the activation function is applied to the output of the HIPP cells by using (9). In the next step, these HIPP cells affect the remaining granule cells that survived from the previous steps. As demonstrated in Figure 7, there are weight connections between HIPP and granule cells named \( W_{hp}^{ij} \). \( W_{hp}^{ij} \) is one for the granule cells \( j \), which receive
input from h and is zero otherwise. So all HIPP cells will not take part in the inhibition operation. Therefore, the selection of appropriate cells will be an essential issue. The selection of the most suitable cells will be along with a significant effect in improving the recognition operations. Poisson distribution is a proper choice in selecting the most suitable cells for this step, as described in relations 9. By calculating all the HIPP cells’ activity, the granule cells’ inhibition is applied as follows:

\[ g^{ie}_{i-HIPP}(j) = \beta_{HIPP} \sum Y_h W_{hp}^{ij} \]  

(15)

the weight between the pth HPIP cell and ith granule cells from the jth layer is represented as \( W_{hp}^{ij} \). Also, \( Y_h \) indicates the rate of activity of the pth HPIP cell. \( g^{ie}_{i-HIPP} \) increases linearly with the input activity density and provides the effect of normalization on the output of granule cells for the dense active inputs [52].

E. STEP 5: OUTPUT CALCULATION

After performing the above four steps, the remaining active granule cells will form the system’s output. Figure 8 shows the final granule cell layers. In Figure 8, the silenced granule cells are shown in black, and the output of the proposed network (remaining active granule) is shown in white. The output of each granule cell in each layer will be determined based on the following (1):

\[ V^i_j = V^i_{rest}(j) + g^{ie-pp}(j) - g^{ie-int}(j) + g^{ie-mc}(j) - g^{ie-HIPP}(j) \]  

(16)

where, \( V^i_j \) represents the potential of ith granule cell from the jth layer. Moreover, the resting potential in this cell is represented as \( V^i_{rest}(j) \), which is considered a constant value. The excitation from the input, the inhibition from the interneurons, the excitation from the mossy cells, and ultimately, the inhibition from the HIPP cells into the granule cells are represented as \( g^{ie-pp}(j) \), \( g^{ie-int}(j) \), \( g^{ie-mc}(j) \), and \( g^{ie-HIPP}(j) \), respectively.

F. STEP 6: LEARNING IN THE PROPOSED NETWORK

Naturally, at the end of each trial of the network, and after calculating the network output, it is necessary to train network weights. In the proposed network, two weight sets need to be updated: 1) Weight of connections between the granule cells and input cells, and 2) Weights of connections between HIPP cells and input cells. Granule cells in the proposed network use the “winner take all” structure [52], meaning that the more robust cells silence the rest of the cells. The two inhibition steps in the proposed network structure also confirm this issue. Competitive learning [62], [63] is the best way to train the proposed network with this feature of granule cells. Both vectors of the preceding weights are trained by competitive learning at the end of each trial of the proposed network.

V. SIMULATION OF THE PROPOSED NETWORK

In this study, handwritten digit and character images with high overlaps are utilized to evaluate the proposed method. Recognition of such images requires high accuracy and speed. The experimentation was carried out on the system with Intel Core i5 6200 CPU clocked at 2.30 GHz and with a system memory of 8 GB RAM. The number of layers of the granule cells in separate experiments was separately determined to be 3 and 5. The number of granule cells in each layer was 20. The number of mossy and HIPP cells in various experiments was 10 and 7, respectively. Table 1 shows the initial values of multiple parameters of the proposed network in different experiments. Preprocesses such as noise removal and normalization were performed on the handwritten images initially.

The experiments were carried out using equal but different datasets from five script languages: MNIST dataset for
English digit recognition, Chars74K dataset for English digits and characters recognition, CMATERdb3.1.2 for Devanagari digits, CMATERdb3.4.1 for Telugu digit, HODA dataset for Persian/Arabic digits, and SVHN dataset for non-character English digit recognition. Image resizing was performed differently for each dataset, which will be explained in the relevant section.

Feature extraction was performed by the PCA method, and then the images were trained and tested by the proposed network.

A. DATASETS

1) MNIST DATASET

MNIST contains samples of images of the numbers 0 to 9 in two categories of training and test. The training and test sets in the MNIST dataset include 60,000 and 10,000 pattern samples respectively, which were collected from 250 different authors [64].

2) Chars74K DATASET

This dataset consists of 62 classes (0-9, a-z, A-Z) comprising of 7705 characters obtained from natural images, 3410 hand-drawn characters using a tablet PC, and 62992 synthesized characters from computer font [65].

3) CMATERDB

CMATER is a pattern recognition database repository created at the CMATER laboratory in Jadavpur University using special pre-formatted data sheets, which are filed by people from different age groups and education levels. For various scripts, individual digits and characters were extracted from the sheets [66]. Two of these CMATER datasets for Devanagari numerals and Telugu numerals were used in this study to test the proposed network.

- **CMATERdb3.1.1 dataset** The CMATERdb3.1.1 was performed for Devanagari handwritten digit script [66]. For this dataset, 3000 samples were collected. For training, 2000 samples were used and the rest of the samples were used for the test.

- **CMATERdb3.4.1 dataset** This dataset contains samples for the Telugu digit script. Similar to CMATER3.4.1, this dataset includes 4000 training digit images and 2000 test digit images [66].

4) HODA DATASET

This dataset was collected from approximately 12000 entrance exam registration forms. This dataset contains 80000 handwritten digit images with a resolution of around 200 dpi (dots per inch). HODA dataset contains 60000 samples for training and 20000 samples for testing [67].

5) SVHN DATASET

SVHN is a real-world dataset obtained from house numbers in Google Street View Images [68]. There were 73257 images for training and 26032 for testing. These images were 32 × 32 pixels.

B. EXPERIMENTS ON MNIST DATASET

1) NETWORK TRAINING AND TESTING

Two experiments were performed for evaluating the proposed network using MNIST dataset. For all experiments, the images were changed to 28 * 28 gray scales, similar to the study by [17].

**Experiment 1:** in this experiment, the accuracy of the proposed network was evaluated through training 2,000 samples out of 60,000 samples and applying 10-fold cross-validation. Results of this experiment were compared with two methods: CR-SOM and CR-MSOM. The results for the recognition accuracy of this experiment are presented in Table 2. While the CR-MSOM algorithm is more accurate than the CR-SOM in 9 steps of the 10-cross validation, the proposed network has higher accuracy in validation. However, the proposed network has higher accuracy in all ten levels than the CR-SOM and CR-MSOM methods.

| 10-fold | CR-SOM | CR-MSOM | Proposed Network |
|---------|--------|---------|-----------------|
| 1       | 95.7   | 97.6    | 98.01           |
| 2       | 97.4   | 99.3    | 99.5            |
| 3       | 95.6   | 98.0    | 99.0            |
| 4       | 93.6   | 96.3    | 97.7            |
| 5       | 93.3   | 95.0    | 97.0            |
| 6       | 96.0   | 98.0    | 99.01           |
| 7       | 95.3   | 99.0    | 99.5            |
| 8       | 99.0   | 98.6    | 99.3            |
| 9       | 96.0   | 97.6    | 98.3            |
| 10      | 94.3   | 98.3    | 99.03           |

Test data were recognized with the proposed network by a significant speed in the early steps of network training. This significant speed is due to the two stages of inhibition which exist in the proposed network. The silenced weak granule cells reduced the number of subsequent operations of the proposed network in the later steps. Only the weights of granule cells that survived from inhibition were updated, so the network’s training was accelerated. Table 3 indicates the number of granule cells before and after each inhibition step in this experiment. The proposed network in this experiment trained for 20 epochs with three layers of granule cells, with each layer containing 20 granule cells. That is, there is a total of 60 granule cells in this network. Table 3 demonstrates the number of these cells after each inhibition. The first inhibition was performed by interneuron cells in step 2 of the proposed network, and the second inhibition was performed by HIPP cells in step 4 of the proposed network. As it is clear, the number of granule cells has been reduced drastically after the first inhibition by the interneurons and after the second inhibition by the HIPP cells. As depicted in Table 3, the number of granule cells was reduced after two inhibitions and became nearly half of the granule cells’ initial amount.
TABLE 3. The number of granule cells in different stages of the proposed network.

| Layers | number of GCs | number of GCs after first inhibition | number of GCs after first inhibition |
|--------|---------------|--------------------------------------|--------------------------------------|
| layer1 | 20            | 14                                   | 10                                   |
| layer2 | 20            | 15                                   | 10                                   |
| layer3 | 20            | 13                                   | 11                                   |
| Total  | 60            | 42                                   | 31                                   |

TABLE 4. Proposed network testing by samples with and without noise.

| Experiment         | Number of the training set | Number of the test set | Accuracy % |
|--------------------|-----------------------------|-------------------------|------------|
| with noise         | 2000                       | 200                     | 91.5       |
| Without noise      | 2000                       | 200                     | 99.5       |

TABLE 5. Confusion matrix for testing of the noisy images.

| Truth | Prediction |
|-------|------------|
| 0     | 0          |
| 1     | 1          |
| 2     | 2          |
| 3     | 2          |
| 4     | 2          |
| 5     | 2          |
| 6     | 2          |
| 7     | 1          |
| 8     | 1          |
| 9     | 1          |

TABLE 6. Test accuracy comparison between proposed network, DSOM, and E-DSOM.

| Model       | Test Accuracy for different noise level(%) |
|-------------|---------------------------------------------|
|             | 2   | 5   | 10  | 20  | 40  | 50  | 60  |
| DSOM [28]   | 83.37| 83.14| 83.14| 82.39| 74.46| 62.00| 20.37|
| E-DSOM [21] | 87.12| 87.15| 86.88| 86.51| 79.91| 69.34| 23.64|
| Proposed Network | 91.5 | 91.47 | 87.75 | 85.71 | 79.59 | 77.55 | 71.42 |

FIGURE 9. Change in accuracy with different noise levels for DSOM, E-DSOM, and proposed network.

Experiment 2: In this experiment, 400 images of the 60,000 training samples were used in the training phase. The network was tested with a total of 10,000 test samples from the database. The proposed network in this experiment yielded an accuracy of 99.7%. The results of this experiment are compared with the latest SOM-based and deep-based method in section G. COMPARATIVE ANALYSIS.

2) THE GENERALIZED CAPABILITY OF THE PROPOSED METHOD

Noise is one of the most critical issues in the handwritten recognition. In this study, generalization capability was analyzed using noisy test data. An experiment was carried out with a total of 2,000 samples for training. In turn, 200 test samples were used for the test of the proposed network. An accuracy of 99.5% was achieved. Only one incorrect recognition was obtained, which is related to class “7”. Then, these 200 images were tested by adding noise level of 2. The accuracy was decreased from 99.5% to 91.5% and the number of incorrect recognition cases was increased to 17. Table 4 illustrates the results of this assessment. Confusion matrix of incorrect recognition noisy images is presented in table 5. As shown in the confusion matrix, class “7” has the most incorrect recognition cases. Four cases of all incorrect cases occurred in class “7”. Class “1”, “3”, and “9” are at the next level, with each class having 3 incorrect recognition images. Images related to class numbers “2”, “4”, and “5” were recognized correctly.

Also, the proposed network was tested against the various noise levels of 5, 10, 20, 40, 50, and 60.

When the proposed network’s generalization capability, DSOM, and the E-DSOM are compared, the proposed network consistently outperforms the DSOM and the E-SOM. These results are presented in Table 6 and Figure 9. As shown in Table 6 and Figure 9, for all noise levels, the proposed network showed better accuracy in comparison with DSOM and E-DSOM. Table 6 indicates that the accuracy of DSOM and E-DSOM was highly decreased when the level of noise increased. However, the proposed network showed a slight decrease. Notably, when the noise increased from 50 to 60, the accuracy of E-DSOM showed a considerable reduction from 69.34% to 23.64%. This reduction is much more than half.

Contrary to the proposed network, the DSOM showed similar behavior. This result means that even by increasing noise levels, the proposed method is much more capable than other comparable methods.

3) TESTING OF THE PROPOSED NETWORK WITH HIGHLY CHALLENGED EXAMPLES

The MNIST database includes 25 test cases with the highest challenge [69]. In reference [69], a CNN-SVM-based approach suggested that only six samples could be identified from among these 25 samples. Reference [17] proposed a SOM-based approach that successfully identified 24 samples of these challenging samples. In the study mentioned above, only one sample was identified wrongly. However, the
The proposed method was successful in the correct recognition of all the 25 samples with reasonable accuracy.

**C. EXPERIMENTS ON Chars74K DATASET**

Chars74K dataset was used for testing of the proposed network in order to evaluate the behavior of the proposed network against a dataset with more than 10 classes.

To study the potential competence of the proposed network for recognizing all of the 62 classes, the characters and digits in Chars74K dataset were considered. As discussed in above, Chars74K has a combination of handwritten characters, natural image characters and characters extracted from computer fonts. In this study, handwritten images were used.

In the experiments, 3410 handwritten character images were used for the training. Testing the proposed network was performed using 30% of these images. Images of Chars74K have different dimensions. So, they were cropped according to the location of the bounding boxes on individual characters to obtain images of the same size. They were also converted from color images to grayscale images. The accuracy of 98.75% was achieved by the proposed network.

**D. EXPERIMENTS ON CMATER DATASETS**

In order to demonstrate the capability of the proposed network in recognizing the various character scripts, CMATER datasets from two various types of scripts were used in this study. Devanagari and Telugu digit datasets used in this study belong to popular Indic scripts and are genealogically completely different from each other [70]. CMATER3.2.1 and CEMATER3.4.1 were used for Devanagari and Telugu digit recognition, respectively.

1) EXPERIMENTS ON CMATERDB 3.2.1

In this experiment, 2000 Devanagari digit images of CMATERdb3.2.1 were used for training. Testing the proposed network was performed using 1000 Devanagari digit images. The Results of the experiment show a good accuracy recognition of 99.6% for Devanagari digit recognition.

2) EXPERIMENTS ON CMATERDB 3.4.1

For the recognition of Telugu digits, 4000 images from CMATERdb3.4.1 were used to train the proposed network. Testing was performed using 2000 digit images. The accuracy of 99.65% was achieved by the proposed network for this experiment.

**E. EXPERIMENTS ON HODA DATASET**

The same numeral script was used for Arabic and Persian script. Persian/Arabic handwritten digits have a high overlapping. So, the evaluation of the proposed network for recognizing Persian/Arabic handwritten digits can demonstrate the efficiency of the proposed network. The experiment was carried out using all 60000 samples of the training. The proposed network was tested using 20000 images. An accuracy of 99.55% was obtained.

**F. EXPERIMENTS ON SVHN DATASET**

In order to prove the effectiveness of the proposed network, an experiment was carried out using the non-character SVHN dataset. Training of the proposed network was performed using all 73257 training samples of the dataset. All 26032 test samples of VSHN were used for evaluating the proposed network. Images had different dimensions. For this reason, they were cropped according to the location of the bounding boxes on individual digits to obtain images of the same size. They were also converted from color images to grayscale images. The accuracy recognition of 98.5% was achieved by the proposed network using SVHN dataset.

**G. COMPARATIVE ANALYSIS**

The proposed network was compared with some of the most popular SOM-based methods, including deep SOM. As mentioned in the previous sections, the proposed network was compared with SOM-based methods because of the similarity in their learning algorithm. Since the deep learning methods have attracted much attention in recent years due to great success in different domains such as the recognition field, the proposed network was compared with some of the most popular deep learning methods as well.

1) COMPARISON ON USED DATASETS

The proposed network was compared with SOM-based and deep learning methods. Results of accuracy and error rate comparison on MNIST dataset are presented in Tables 7 and 8, respectively.

As depicted in Table 8, the proposed network resulted in the least error rate as compared with other deep-based methods.

Tables 9 to 13 present the comparison of the experiments of the proposed network on the other used datasets.

The results of Table 7 to Table 13 show that the proposed network outperforms the other methods using different datasets of various languages.

2) STATISTICAL TESTS

For the evaluation of the results, error measurement is useful for a better evaluation. Character error rate (CER) was calculated for all the experimental results. CER is the percentage of erroneous characters in the system output and is a
TABLE 8. Comparative results (error rate) on MNIST dataset.

| Reference          | Method            | Error rate % |
|--------------------|-------------------|--------------|
| Madlakannu [72]    | DIGI-Net CNN      | 0.89         |
| Dorra Melloul [44] | Deep Morph-CNN    | 0.34         |
| Sergey, Loffe [73] | CNN               | 0.53         |
| Proposed network   | DG OF HIPPOCAMPUS | 0.3          |

TABLE 9. Comparative analysis on Char574K dataset.

| Reference          | Method            | Accuracy % |
|--------------------|-------------------|------------|
| Moazam Soomro [74] | CNN (GoogleNet)   | 88.89      |
| Moazam Soomro [74] | CNN (AlexNet)     | 77.77      |
| Proposed network   | DG OF HIPPOCAMPUS | 98.75      |

TABLE 10. Comparative analysis on CMATERdb3.2.1 dataset.

| Reference          | Method                          | Accuracy % |
|--------------------|---------------------------------|------------|
| Pramanic [75]      | CNN-based transfer learning     | 94.75      |
| Acharya [76]       | CNN                             | 98.47      |
| Sarkhel [71]       | CNN-based                       | 99.5       |
| Proposed network   | DG OF HIPPOCAMPUS               | 99.6       |

TABLE 11. Comparative analysis on CMATERdb3.4.1 dataset.

| Reference         | Method              | Accuracy % |
|-------------------|---------------------|------------|
| Nanehkaran [36]   | CNN                 | 99.45      |
| Proposed network  | DG OF HIPPOCAMPUS   | 99.55      |

TABLE 12. Comparative analysis on HODA dataset.

| Reference         | Method              | Accuracy % |
|-------------------|---------------------|------------|
| Dorra Melloul [43]| Deep Morph-CNN      | 97.33      |
| Proposed network  | DG OF HIPPOCAMPUS   | 98.5       |

TABLE 13. Comparative analysis on SVHN dataset.

| Reference         | Method             | Accuracy % |
|-------------------|--------------------|------------|
| Goodfellow [76]   | CNN                | 96.04      |
| Dorra Melloul [43]| Deep Morph-CNN     | 97.33      |
| Proposed network  | DG OF HIPPOCAMPUS  | 98.5       |

TABLE 14. CER and Confidence Interval of the experimental results for various datasets.

| Dataset  | CER on validation (%) | Confidence Interval |
|----------|-----------------------|---------------------|
| MNIST    | 0.6 ± 0.1             | (0.5, 0.7)          |
| Chars74K | 1.6 ± 0.14            | (1.25, 1.96)        |
| CMATERdb3.2.1 | 1.05 ± 0.15 | (0.9, 1.2)          |
| CMATERdb3.4.1 | 1.05 ± 0.075 | (0.98, 1.13)        |
| HODA     | 0.65 ± 0.15           | (0.5, 0.8)          |
| SVHN     | 1.73 ± 0.23           | (1.25, 1.55)        |

TABLE 15. Pair t-test and standard deviation on the experimental results of used datasets.

| Dataset  | p-value  | Null hypothesis | Standard Deviation |
|----------|----------|-----------------|-------------------|
| MNIST    | 0.00003  | Rejected        | 0.8445            |
| Chars74K | 0.00482  | Rejected        | 0.3747            |
| CMATERdb3.2.1 | 0.00008 | Rejected        | 0.3151            |
| CMATERdb3.4.1 | 0.00001 | Rejected        | 0.3389            |
| HODA     | 0.00020  | Rejected        | 0.2651            |
| SVHN     | 0.00150  | Rejected        | 0.3323            |

was calculated for all experimental results with a confidence interval of 95% (statistical significance level of 5%), as represented in Table 15.

VI. DISCUSSION

The DG of the hippocampus of the brain is responsible for recognizing the patterns with high overlapping. In this study, a network is proposed based on the computational function of the DG. In the proposed network, granule cells, which are the essential cells, are arranged in several layers. The network output will be the activity of these cells following the general steps of the network. During the operations of the network, two excitation operations and two inhibitory operations on the granule cells will enhance recognition accuracy, one of the unique features in the structure of the proposed network. The speed of operation will be increased due to the removal of some granule cells in the two inhibition stages. As in the next steps, only the remaining active cells are considered. The policy used in the proposed network structure is that the stronger cells will become stronger. In other words, there is a kind of competition among granule cells to remain active. Since the stronger cells get much stronger in the proposed network at different steps of excitation of the granule cells, the recognition of samples is performed with higher accuracy. The same corresponding granule cells are activated by activating two similar inputs, and the same corresponding cells are silenced.

The accuracy of recognizing and clustering 200 test images using MNIST dataset is 99.5%. The good speed accompanies high accuracy in the proposed network in the early steps of training the proposed network. This excellent efficiency is due to the two steps of inhibition and two steps of excitation of the proposed network, as mentioned above. With inhibition steps, some of the weak granule cells are silenced. With two excitation steps, the strong granule cells become stronger. For instance, by having 20 granule cells in each of the three layers in the network, the number of initially inhibited granule cells was reduced to 14, 12, 13 in the three layers. The number of
Deep learning was considered because of the importance of using the same learning algorithms, i.e., competitive learning. SOM-based methods were considered for comparing various languages. Experimental results were compared with evaluated using six digit and character datasets from five datasets from various languages. The proposed network was proposed network could recognize all of these challenging images, the reference [17] proposed a SOM-based method that successfully recognized 24 samples of these challenging images. While the MNIST database includes 25 test samples with the best previous accuracy of 99.7% outperforms that of the-state-of-the-art SOM-based methods, namely the DSOM and CR-MSOM, which are 99.03% and 96.17%, respectively. The results show that the proposed network outperforms the best compared deep learning method by an error rate of 0.3% instead of the Deep Morph-CNN with an error rate of 0.34%. CMATERdb3.2.1 and CMATERdb3.4.1 for Devanagari and Telugu digit scripts, which are popular scripts in India, were tested with the proposed network. The accuracy recognitions of 99.6% and 99.65% were achieved on these two datasets, respectively. The fourth digit dataset, HODA, was selected from Persian/Arabic languages. The same numeral script was used for these two languages. The good accuracy of 99.55% was obtained by experimenting on HODA dataset. For better evaluation, the proposed network was tested on a non-character digit dataset, SVHN. The accuracy recognitions of 98.5% was achieved by experiment on SVHN. Experimental results showed better accuracy recognition achieved for handwritten digit samples than non-character digit samples. Moreover, a comparison of the experimental results of four handwritten digit datasets demonstrated that a better accuracy recognition obtained on MNIST dataset. To evaluate the proposed network against datasets with more than 10 classes, Chars74K dataset with 62 classes was used. The notable point is that for digit recognition of various datasets and different languages, the training of the proposed network was performed in 20 training epochs, while for character datasets of Chars74K, the training was performed in 50 training epochs. However, the training epochs of the proposed network for digit recognition experiments are fewer than other compared methods such as [71] which reported their number of training epochs.

Results on all digit, character and non-character datasets depicted good standard deviation as presented in Table 15. As demonstrated in Table 15, the null hypothesis is rejected for all datasets by a pair t-test for a confidence coefficient of 0.05. In order to complete the statistical analysis, CER for all the experiments was computed. The results showed a significant boost in accuracy, resulting in CER less than 1.85% for all experiments on various datasets. The smallest CER was achieved for experiments carried on MNIST digit dataset. Experimental results showed that an almost uniform confidence interval was achieved for all experiments with the used datasets. All these results show the superiority of the proposed network.

Another advantage of the proposed network is its capability of recognizing noisy patterns. Experiments were performed for 200 test images with a noise level of 2, 5, 10, 20, 40, 50, and 60. Results indicated that recognition of these noisy images was accomplished with good accuracy. Since the learning algorithm of the proposed network is implemented using competitive learning like the SOM-based methods, the proposed network is compared with SOM-based methods. Specifically, the proposed network is compared with two previous successful SOM-based methods, namely the DSOM and the E-DSOM. Noisy digit images slightly reduce the accuracy of the proposed network. These good results are especially noticeable for noisy images with a noise level of 60. In DSOM, the accuracy of these samples was reduced from 83.37% to 20.37%. In E-DSOM, the reduction of the accuracy is from 87.12% to 23.64%. The proposed network shows superiority in generalization capability; hence its accuracy is slightly reduced from 91.5% to 71.42% for the worst noise level.

The proposed network was compared with the other two SOM-based methods, CR-SOM and CR-MSOM [17], for 10-fold cross-validation. Results showed that the proposed network was more accurate in comparison with CR-SOM and CR-MSOM. The proposed network achieved higher accuracy in all ten steps compared to CR-SOM and CR-MSOM. The network presented in this study achieved an accuracy of 99.7%. The best previous accuracy of 99.03% for SOM-based methods was reported in [17]. As reported in [69], the MNIST database includes 25 test samples with the highest challenge due to rotation, fracture, or misalignment. The CNN-SVM-based approach [69] could recognize only six images from among these challenging samples. While reference [17] proposed a SOM-based method that successfully recognized 24 samples of these challenging images, the proposed network could recognize all of these challenging samples.

In order to prove the effectiveness of the proposed network, other experiments were carried out using character and digit datasets from various languages. The proposed network was evaluated using six digit and character datasets from five various languages. Experimental results were compared with two categories of methods: SOM-based methods and deep learning. SOM-based methods were considered for comparison because the proposed network, like the SOM-methods, used the same learning algorithms, i.e., competitive learning. Deep learning was considered because of the importance of these algorithms in recent character recognition researches.

The performed experiments can be divided into three sets. Experiments for handwritten digit recognition, experiment for non-character images of digit recognition, and experiment for handwritten character recognition. Experiments on handwritten digit recognition were performed on four datasets of five different languages. Results of the experiment on MNIST dataset outperformed both the SOM-based and the deep learning methods. The achieved recognition accuracy of 99.7% outperforms that of the-state-of-the-art SOM-based methods, i.e., deep SOM and CR-MSOM, which are 99.03% and 96.17%, respectively. The results show that the proposed network outperforms the best compared deep learning method by an error rate of 0.3% instead of the Deep Morph-CNN with an error rate of 0.34%. CMATERdb3.2.1 and CMATERdb3.4.1 for Devanagari and Telugu digit scripts, which are popular scripts in India, were tested with the proposed network. The accuracy recognitions of 99.6% and 99.65% were achieved on these two datasets, respectively. The fourth digit dataset, HODA, was selected from Persian/Arabic languages. The same numeral script was used for these two languages. The good accuracy of 99.55% was obtained by experimenting on HODA dataset.

For better evaluation, the proposed network was tested on a non-character digit dataset, SVHN. The accuracy recognitions of 98.5% was achieved by experiment on SVHN. Experimental results showed better accuracy recognition achieved for handwritten digit samples than non-character digit samples. Moreover, a comparison of the experimental results of four handwritten digit datasets demonstrated that a better accuracy recognition obtained on MNIST dataset.

To evaluate the proposed network against datasets with more than 10 classes, Chars74K dataset with 62 classes was used.

The notable point is that for digit recognition of various datasets and different languages, the training of the proposed network was performed in 20 training epochs, while for character datasets of Chars74K, the training was performed in 50 training epochs. However, the training epochs of the proposed network for digit recognition experiments are fewer than other compared methods such as [71] which reported their number of training epochs.

Results on all digit, character and non-character datasets depicted good standard deviation as presented in Table 15. As demonstrated in Table 15, the null hypothesis is rejected for all datasets by a pair t-test for a confidence coefficient of 0.05. In order to complete the statistical analysis, CER for all the experiments was computed. The results showed a significant boost in accuracy, resulting in CER less than 1.85% for all experiments on various datasets. The smallest CER was achieved for experiments carried on MNIST digit dataset.

Experimental results showed that an almost uniform confidence interval was achieved for all experiments with the used datasets. All these results show the superiority of the proposed network.
VII. CONCLUSION

In humans, pattern separation is performed by a part of the brain called the dentate gyrus of the hippocampus. In this study, a novel network is presented for automatic pattern separation based on the DG activity of the hippocampus. Mainly, the computational function of the DG of the hippocampus is considered for intelligent pattern recognition of handwritten images. The proposed network employs two excitation steps and two inhibition steps, which constitute a prominent feature.

Several experiments were carried out on the proposed network using six datasets of five various languages, some of which are presented in this study. Simulation results are compared with the SOM-based methods and deep learning methods, indicating that the proposed method is more accurate in recognizing numbers and characters than other compared methods. Moreover, numerical simulation results indicate that the proposed network requires less training iterations to achieve a high accuracy, because weak granule cells are silenced in different steps of the proposed network. This recognition accuracy is prominently manifested in images that have certain levels of noise. Promising results achieved by the proposed network for datasets from various languages including, English, Devanagari, Telugu and, Persian/Arabic handwritten digit datasets. The proposed network was also evaluated using a non-character dataset named SVHN. The evaluation of the proposed network using Chars74K dataset with 62 classes demonstrated that the proposed network is capable of recognizing datasets with various numbers of classes. Experimental results show a significant boost in accuracy, resulting in a CER of smaller than 1.85% at uniform confidence intervals.

The notable point is that for digit recognition of various datasets and different languages, the training of the proposed network was performed in 20 training epochs. In contrast for character datasets of Chars74K, the training was increased to 50 training epochs.

For future work, the functionality of DG in pattern separation can be combined with the functionality of CA3 in the hippocampus. CA3 is a region of the hippocampus that plays an important role in pattern completion [79], [80]. Pattern completion refers to a process in which the entire pattern stored in the network can be retrieved and completed from just a small part of the pattern [81]. So, the combination of pattern separation and pattern completion will help recognize damaged handwritten digit and character images. Many neural networks only are trained with the excitation of neurons. Cerebellar Model Articulation Controller (CMAC) [82], [83] has one step of inhibition. In future work, the proposed network’s inhibition steps can be combined with other neural networks to deal with their performance with and without inhibition steps. Utilizing the proposed network for other fields with highly overlapping patterns is another idea for future work.

REFERENCES

[1] G. A. Susto, A. Schirru, S. Pampuri, S. McLoone, and A. Beghi, “Machine learning for predictive maintenance: A multiple classifier approach,” IEEE Trans. Ind. Informat., vol. 11, no. 3, pp. 812–820, Jun. 2015.

[2] P. A. Mueller and D. M. Oppenheimer, “The pen is mightier than the keyboard: Advantages of longhand over laptop note taking,” Psychol. Sci., vol. 25, no. 6, pp. 1159–1168, 2014.

[3] H. Chaudhury and S. R. M. Prasanna, “Handwriting recognition using sinusoidal model parameters,” Pattern Recognit. Lett., vol. 121, pp. 87–96, Apr. 2019.

[4] S. Sarkar, M. Ghosh, A. Chatterjee, S. Malakar, and R. Sarkar, “An advanced particle swarm optimization based feature selection method for trascript handwritten digit recognition,” in Proc. 2nd Int. Conf. Comput. Intell., Commun., Bus. Anal. (CICBA). New York, NY, USA: Springer-Verlag, 2018, pp. 82–94.

[5] T. M. Ghanim, M. I. Khalil, and H. M. Abbas, “Comparative study on deep convolution neural networks DCNN-based offline Arabic handwriting recognition,” IEEE Access, vol. 8, pp. 95465–95482, 2020.

[6] H. Sadar, M. M. Pedram, and M. Teshnehlab, “Multi-view deep network: A deep model based on learning features from heterogeneous neural networks for sentiment analysis,” IEEE Access, vol. 8, pp. 86984–86997, 2020.

[7] W. Choi, K. Duraisamy, R. G. Kim, J. R. Doppa, P. P. Pande, D. Marculescu, and R. Marculescu, “On-chip communication network for efficient training of deep convolutional networks on heterogeneous manycore systems,” IEEE Trans. Comput., vol. 67, no. 5, pp. 672–686, May 2018.

[8] X. Jiang, J. Sun, C. Li, and H. Ding, “Video image defogging recognition based on recurrent neural network,” IEEE Trans. Ind. Informat., vol. 14, no. 7, pp. 3281–3288, Jul. 2018.

[9] Q. Zhang, L. T. Yang, Z. Yan, Z. Chen, and P. Li, “An efficient deep learning model to predict cloud workload for industry informatics,” IEEE Trans. Ind. Informat., vol. 14, no. 7, pp. 3170–3178, Jul. 2018.

[10] C. Boufenar, A. Kerboua, and M. Batouche, “Investigation on deep learning for off-line handwritten Arabic character recognition,” Cognit. Syst. Res., vol. 50, pp. 180–195, Aug. 2018.

[11] D. Wu, X. Luo, G. Wang, M. Shang, Y. Yuan, and H. Yan, “A highly accurate framework for self-labeled semisupervised classification in industrial applications,” IEEE Trans. Ind. Informat., vol. 14, no. 3, pp. 909–920, Mar. 2018.

[12] S. N. Tuncdemir, C. O. Lacefield, and R. Hen, “Contributions of adult neurogenesis to dentate gyrus network activity and computations,” Behav. Brain Res., vol. 374, Nov. 2019, Art. no. 112112, doi: 10.1016/j.bbr.2019.112112.

[13] E. E. Bernstein and R. J. McNally, “Examining the effects of exercise on pattern separation and the moderating effects of mood symptoms,” Behav. Therapy, vol. 50, no. 3, pp. 582–593, May 2019.

[14] S. M. Stark, M. A. Yassa, J. W. Lacy, and C. E. L. Stark, “A task to assess behavioral pattern separation (BPS) in humans: Data from healthy aging and mild cognitive impairment,” Neuropsychologia, vol. 51, no. 12, pp. 2442–2449, Oct. 2013.

[15] T. Kohonen, Self-Organizing Maps, vol. 30, 3rd ed. Berlin, Germany: Springer, 2001.

[16] A. Diaz Ramos, E. Lopez-Rubio, and E. J. Palomo, “The forbidden region A. W-SOM, an algorithm for high-speed learning in hardware self-organizing maps for unsupervised image classification,” IEEE Trans. Circuits Syst. II, Exp. Briefs, vol. 67, no. 2, pp. 104–118, Dec. 2014.

[17] J. Mariette, M. Olteanu, and N. Villa-Vialaneix, “Efficient interpretable variants of online SOM for large dissimilarity data,” Neurocomputing, vol. 225, pp. 31–48, Feb. 2017.

[18] S. Zhang and S. Ge, “User power interaction behavior clustering analysis that is based on the self-organizing-center K-means algorithm,” IEEE Access, vol. 7, pp. 175879–175888, 2019.

[19] G. C. Cardarilli, L. D. Nunzio, R. Fazzolari, M. Re, and S. Spanò, “AW-SOM, an algorithm for high-speed learning in hardware self-organizing maps,” IEEE Trans. Circuits Syst. I, Exp. Briefs, vol. 67, no. 2, pp. 380–384, Feb. 2020.

[20] C. S. Wickramasinghe, K. Amarasinghe, and M. Manic, “Deep self-organizing maps for unsupervised image classification,” IEEE Trans. Ind. Informat., vol. 15, no. 11, pp. 5837–5845, Nov. 2019.
[69] X.-X. Niu and C. Y. Suen, “A novel hybrid CNN–SVM classifier for recognizing handwritten digits,” Pattern Recognit., vol. 45, no. 4, pp. 1318–1325, Apr. 2012.

[70] R. Sarkhel, N. Das, A. Das, M. Kundu, and M. Nasipuri, “A multi-scale deep quad tree based feature extraction method for the recognition of isolated handwritten characters of popular indic scripts,” Pattern Recognit., vol. 71, pp. 78–93, Nov. 2017.

[71] A. Madakannu and A. Selvaraj, “DIGI-Net: A deep convolutional neural network for multi-format digit recognition,” Neural Comput. Appl., vol. 32, no. 15, pp. 11373–11383, Aug. 2020.

[72] S. Ioffe and C. Szegedy, “Batch normalization: Accelerating deep network training by reducing internal covariate shift,” 2015, arXiv:1502.03167. [Online]. Available: http://arxiv.org/abs/1502.03167

[73] M. Soomro, M. A. Farooq, and R. H. Raza, “Performance evaluation of advanced deep learning architectures for offline handwritten character recognition,” in Proc. Int. Conf. Frontiers Inf. Technol. (FIT), Islamabad, Pakistan, Dec. 2017, pp. 362–367.

[74] R. Pramanik, P. Dansena, and S. Bag, “A study on the effect of CNN-based transfer learning on handwritten Indic and mixed numeral recognition,” in Document Analysis and Recognition, vol. 1020, S. Sundaram and G. Harit, Eds. Singapore: Springer, 2019. [Online]. Available: http://doi.org/443.webvpn.fjmu.edu.cn/10.1007/978-981-3-93617-7-4

[75] S. Acharya, A. K. Pant, and P. K. Gyawali, “Deep learning based large scale handwritten devanagari character recognition,” in Proc. 9th Int. Conf. Softw., Knowl., Inf. Manage. Appl. (SKIMA), Dec. 2015, pp. 1–5.

[76] I. J. Goodfellow, Y. Bulatov, J. Ibarz, S. Arnoud, and V. Shet, “Multi-digit number recognition from street view imagery using deep convolutional neural networks,” 2013, arXiv:1312.6082. [Online]. Available: http://arxiv.org/abs/1312.6082

[77] S. Drobac and K. Lindén, “Optical character recognition with neural networks and post-correction with finite state methods,” Int. J. Document Anal. Recognit. (IJDAR), vol. 23, no. 4, pp. 279–295, Dec. 2020, doi: 10.1007/s10032-020-00359-9.

[78] S. España-Boquera, M. J. Castro-Bleda, J. Gorbe-Moya, and L. R. Spencer and T. S. Bland, “Hippocampus and hippocampal neurons,” in Stress: Physiology, Biochemistry, and Pathology (Handbook of Life Sciences Series), vol. 3. Amsterdam, The Netherlands: Elsevier, 2019, pp. 57–68. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780128131466000059, doi: 10.1016/B978-0-12-813146-6.00005-9.

[79] J. J. Knierim and J. P. Neunuebel, “Tracking the flow of hippocampal computation: Pattern separation, pattern completion, and attractor dynamics,” Neurobiol. Learn. Memory, vol. 129, pp. 38–49, Mar. 2016.

[80] Y. Senzai, “Function of local circuits in the hippocampal dentate gyrus-CA3 system,” Neurosci. Res., vol. 140, pp. 43–52, Mar. 2019.

[81] W. Fang, F. Chao, L. Yang, C.-M. Lin, C. Shang, C. Zhou, and Q. Shen, “A recurrent emotional CMAC neural network controller for vision-based mobile robots.” Neurocomputing, vol. 334, pp. 227–238, Mar. 2019.

[82] N. Modhej and M. Teshnehlab, “Two novel learning algorithms for CMAC neural network based on changeable learning rate,” IJICTR, vol. 1, no. 1, pp. 35–40, 2012.

AZAM BASTANFARD received the bachelor’s degree in mathematics (applied in the computer) from Islamic Azad University, Karaj Branch, in 1997, and the M.S. and Ph.D. degrees in computer engineering from the Tokyo Institute of Technology, in 2001 and 2004, respectively.

She held a postdoctoral position with MIRALab, University of Geneva. Since 2006, she has been an Assistant Professor with Islamic Azad University, Karaj Branch. Her research areas include artificial intelligence, computer games, image processing, and multimedia. She was awarded “The National Elite Woman.”

MOHAMMAD TESHNEHLAB (Member, IEEE) received the B.S. degree from Stony Brook University, USA, in 1980, the M.S. degree from Oita University, Japan, in 1990, and the Ph.D. degree from Saga University, Japan, in 1994.

He is currently a Faculty Member of the Electrical Engineering Department, K. N. Toosi University of Technology. He is a member of the Industrial Control Center of Excellence and the Founder of the Intelligent Systems Laboratory (ISLab). He is also a Co-Founder and a member of the Intelligent Systems Scientific Society of Iran (ISSSI). His main research interests include artificial rough and deep neural networks, fuzzy systems and neural nets, optimization and its applications in identification, prediction, classification, and control. He is a member of the International Journal of Information and Communication Technology Research (IJICTR).

AZAM BASTANFARD received the bachelor’s degree in mathematics (applied in the computer) from Islamic Azad University, Karaj Branch, in 1997, and the M.S. and Ph.D. degrees in computer engineering from the Tokyo Institute of Technology, in 2001 and 2004, respectively.

She held a postdoctoral position with MIRALab, University of Geneva. Since 2006, she has been an Assistant Professor with Islamic Azad University, Karaj Branch. Her research areas include artificial intelligence, computer games, image processing, and multimedia. She was awarded “The National Elite Woman.”

NAZAL MODHEJ received the B.S. degree in computer engineering from Shahid Chamran University of Ahvaz, Ahvaz, Iran, in 2005, and the M.S. degree in computer engineering from the Science and Research Branch, Islamic Azad University, Ahvaz, in 2010. She is currently pursuing the Ph.D. degree in computer engineering with Islamic Azad University, Karaj Branch, Karaj, Iran.

She has also been a Faculty at Islamic Azad University, Susangerd Branch, since 2005. Her research interests include intelligent systems, neuroscience, and learning.

SOMAYEH RAIESDANA is currently an Assistant Professor of Biomedical Engineering with Qazvin Islamic Azad University, Iran. She does researches in the fields of biomedical engineering and computational neuroscience. Her research topics include epilepsy detection and modeling, sleep analysis, and cognitive processing analyses, including learning and memory, neuromarketing, and neuromuscular control. She is also involved in clinical breast cancer detection and modeling.

***

N. Modhej et al.: Pattern Separation Network Based on the Hippocampus Activity for Handwritten Recognition