Word Recognition in Continuous Speech and Speaker Independent by Means of Recurrent Self-organizing Spiking Neurons

Tarek Behi tarekbehi@gmail.com
Ecole Nationale d'Ingénieurs de Tunis
Electrical Engineering Department
Signal, Image and Pattern Recognition Research Unit
University Tunis El Manar
Tunis, 1002, Tunisia

Najet Arous Najet.Arous@enit.rnu.tn
Ecole Nationale d'Ingénieurs de Tunis
Electrical Engineering Department
Signal, Image and Pattern Recognition Research Unit
University Tunis El Manar
Tunis, 1002, Tunisia

Noureddine Ellouze N.Ellouze@enit.rnu.tn
Ecole Nationale d'Ingénieurs de Tunis
Electrical Engineering Department
Signal, Image and Pattern Recognition Research Unit
University Tunis El Manar
Tunis, 1002, Tunisia

Abstract

Artificial neural networks have been applied successfully in many static systems but present some weaknesses if patterns involve a temporal component. Let's note for example in speech recognition or contextual information, where different of the time interval, is crucial for comprehension. Speech, being a temporal form of sensory input, is a natural candidate for investigating temporal coding in neural networks. It is only through comprehension of the temporal relationship between different sounds which make up a spoken word or sentence that speech becomes intelligible. In fact we present in this paper presents three variants of self-organizing maps (SOM), the Leaky Integrators Neurons (LIN), the Spiking_SOM (SSOM) and the recurrent Spiking_SOM (RSSOM) models. The proposed variants is like the basic SOM, however it represents the characteristic to modify the learning function and the choice of the best matching unit (BMU). The case study of the proposed SOM variants is word recognition in continuous speech and speaker independent. The proposed SOM variants show good robustness and high word recognition rates.

Keywords: Word Recognition, Kohonen Map, Spiking Neural Networks, Leaky Integrators Neurons, Spiking SOM, Recurrent Spiking SOM.

1. INTRODUCTION

The majority of artificial neural network models were based on a computational paradigm involving the propagation of continuous variables from one unit to another. A new generation of pulsed neural networks has emerged, which focuses upon the mathematical formalization of the computational properties of biological neurons [1], [2]. The models which communicate through spikes use the timing of these spikes to encode and compute information. In Spiking Neuron Networks (SNNs), the presence and timing of individual spikes is considered as the means of communication and neural computation. This compares with traditional neuron models where analog values are considered, representing the rate at which spikes are fired. A simple spiking neural model can carry out computations over the input spike trains under several different modes [3]. Thus, spiking neurons compute when the input is encoded in temporal patterns, firing rates, firing rates and temporal correlations, and space–rate codes. An essential feature of the spiking neurons is that they can act as coincidence detectors for the incoming pulses, by detecting if they arrive in almost the same time [4, 5].

Spontaneous speech production is a continuous and dynamic process. This continuity is reflected in the acoustics of speech sounds and, in particular, in the transitions from one speech sound to another [6]. To take account of time in a system of data processing poses two great constraints. First, this system must be able to manage the succession of the various events which must be to treat in a sequential way, it is then a question of sequential treatment. Thus, if the duration of the events is relevant for the task to carry out, the system must be able to treat the temporal structure. However, in the context of the speech recognition, the use of the static networks of neurons is difficult sight the absence of the parameter time in their structure.
In order to classify temporal sequences many technique have been used to model temporal relation in connectionist model [7], [8] like the networks of recurring neurons [9], the temporal self-organizing map [10], [11], [12] and networks of impulse neurons [13], [8] which prove the existence of robust techniques of recognition and classification.

In our model the temporal information is taken into account by using spiking neurons. Spiking neural networks (SNN) have become quite popular recently, due to their biological plausibility. Using spiking neuron models, SNN are able to encode temporal information into both spike timing and spiking rates. The model which realizes the spiking neurons as coincidence detectors encodes the training input information in the connection delays.

The structure of the rest of this paper is as follows: In section 2, we present the self-organizing map and spiking self organizing map. In section 3, we propose the new variant, recurrent spiking self organizing map. In Section 4 we explain the principles of the new variant leaky integrators neurons model. In section 5, we illustrate experimental results of the application of SOM, RSOM, LIN, SSOM and RSSOM models in word recognition of TIMIT speech corpus.

2. SPIKING SELF ORGANIZING MAP

Self-organizing in networks is one of the most popular neural network fields [14], [15]. Such networks can learn to detect regularities and correlations in their input and adapt their future responses to that input accordingly [16]. The neurons of competitive networks learn to recognize groups of similar input vectors. Self-organizing maps learn to recognize groups of similar input vectors in such a way that neurons physically near each other in the neuron layer respond to similar input vectors.

A self-organizing map learns to categorize input vectors. It also learns the distribution of input vectors. Feature maps allocate more neurons to recognize parts of the input space where many input vectors occur and allocate fewer neurons to parts of the input space where few input vectors occur. Self-organizing maps also learn the topology of their input vectors.

The self-organizing map output represents the result of a vector quantization algorithm that gives a fixed number of references or prototype vectors onto high dimensional data sets in an ordered fashion. A mapping from a high dimensional data space ($\Re^n$) onto a two dimensional lattice of units is thereby defined.

An input vector $x \in \Re^n$ is compared with all mi, in any metric; in practical applications, the smallest of the Euclidian distances is usually used to define the best matching unit (BMU). The BMU is the neuron whose weight vector mi is closest to the input vector x determined by:

$$\|x - m_c\| = \min\{\|x - m_i\|\}, \forall\ i \in [1..n] \quad (1)$$

Where n is the number of map units and $||x – m_i||$ is a distance measure between x and m_i.
After finding the BMU, his weight vector is updated so that the BMU is moved closer to the current input vector. The topological neighbors of the BMU are also updated. This adaptation procedure stretches the BMU and its topological neighbors towards the sample vector. Kohonen update rule for weight vector of the unit i in the BMU neighborhood is:

$$m_i(t+1) = m_i(t) + \alpha(t)\, h_{ci}(t)\, [x(t) - m_i(t)],\ \forall\, i \in [1..n] \quad (2)$$

x(t) is the input vector randomly drawn from the input data set at time t, $h_{ci}(t)$ the neighborhood kernel around the winner unit c and α(t) the learning rate at time t [17].

In the context of spiking neuron networks we given a set S of m-dimensional input vectors s = (s_1, ..., s_m) and a spiking neuron network with m input neurons and n output neurons, where each output neuron v_j receives synaptic feedforward input from each input neuron u_i with weight w_{ij} and lateral synaptic input from each output neuron v_k, with weight w_{kj}. At every epoch of the learning procedure one sample is chosen and the input neurons are made fire such that they temporally encode input vectors [18], [19].

In the algorithm proposed here, the winner is selected from the subpopulation of units that fire the quickest in one simulation step. After choosing a winner, learning is applied as follows. The afferent weights of a competitive neuron i are adapted in such a way as to maximize their similarity with the current input pattern j. A measure of the similarity is the difference between the postsynaptic potential s_i that encodes the input stimulus and the connection weight w_{ij}. Furthermore, a spatial and a temporal neighborhood of the winner are created, such that only the neurons inside the S area and which have fired up until a reference time T_{out} are subject to learning. The learning rule is adapted from [29] and is given by:

$$\Delta w_{ij} = \eta \frac{T_{out} - t_j}{T_{out}} (s_i - w_{ij}) \quad (3)$$

Where $\mathbf{T_j}$ is the firing time of the **j** neuron, $\mathbf{T_{out}}$ is a time out limit, and **η** is the learning rate.

3. RECURRENT SPIKING SELF ORGANIZING MAP

The recurrent Self-Organizing Map (RSOM) [20], [21] as an extension to the Self-Organizing Map (SOM) that allows storing certain information from the past input vectors. The information is stored in the form of difference vectors in the map units. The mapping that is formed during training has the topology preservation characteristic of the SOM. Recurrent SOM differs from the SOM only in its outputs. The outputs of the normal SOM are reseated to zero after presenting each input pattern and selecting best matching unit with the typical winner takes all strategy. Hence the map is sensitive only to the last input pattern. In the RSOM the sharp outputs are replaced with leaky integrator outputs, which once activated gradually lose their activity. The modeling of the outputs in RSOM is close to the behavior of natural neurons, which retain an electrical potential on their membranes with decay.
The use of impulsionnel neuron makes it possible to improve the taking into account of the temporal parameters.

The model presented in this part has the same principle that spiking self organizing map except that the choice of the BMU is defined by a difference vector in each unit of the map. The difference vector is included in the recurring bond. Thus, the memory stores a linear sum of the preceding vectors.

The difference vector $I_{(n)}$ in each unit of the map is defined there according to this equation:

$$y_i(t) = (1-\alpha) y_i(t-1) + \alpha(x(t) - m_i(t)) \quad (4)$$

Where $y_i(n)$ is the leaked difference vector in unit i, $0 < \alpha \leq 1$ is the leaking coefficient. x(t) is the input vector and $m_i(t)$ is the weight vector of the unit i.

4. LEAKY INTEGRATOR NEURONS

We present neural networks consisting of leaky integrator units as a universal paradigm for neural and cognitive modeling. In contrast to standard recurrent neural networks, leak integrator units are described by ordinary differential equations living in continuous time. We present an algorithm to train the temporal behavior of leaky integrator networks.

In order to use leaky integrator units to create network models for simulation experiments, a learning rule that works in continuous time is needed. The following formulation is motivated by [22] and describes how a backpropagation algorithm for leaky integrator units can be derived.

In this approach, the state of each neuron *(i)* is represented by a membrane potential $P_i(T)$, which, is a function of the input I(t) which measures the degree of matching between the neuron's weight vector and the current input vector.

The differential equation of a membrane potential is:

$$\frac{dP_i}{dt} = h P_i(t) + I_i(t) \tag{5}$$

Where η < 0.

Particularly, the discrete version of the equation (5), often written as:

$$P_i(t) = l\, P_i(t - 1) + I_i(t) \tag{6}$$

LIN memorise the last activation of each neuron i by means of a Leaky Integrators potential noted $a_i(t)$ [23], [24], [25]:

$$a_i(t) = l\ a_i(t - 1) - \frac{1}{2} \mathrm{P} x(t) - w_i(t) \mathrm{P}^2 \tag{7}$$

Where λ is a depth memory constant($0 \le \lambda \le 1$), x(t) is the input vector, and $w_i(t)$ is the weight vector of neuron *i.* Comparing equations (6) and (7), we find that $I_i(t) = -(½) \; || x(t) – w_i(t)||^2$.

5. RESULTS AND DISCUSSIONS

We have implemented a variant of Kohonen network for continuous speech recognition. The realized system is composed of three main components [26], [27]: a pre-processor sounds and producing mel cepstrum vectors. The sound input space is composed by 12 mel cepstrum coefficients each 16 ms frame. 9 frames are selected at the middle of each word to generate data vectors. The second component is a competitive learning module. The third component is a word recognition module.

We used the DARPA TIMIT speech corpus for all experiments.

The TIMIT corpus is considered as a reference database [28]. Its broad diffusion in the international community allows an objective evaluation and shares performances of the developed systems.
The TIMIT database contains the recordings of 630 American speakers, divided into 8 regional dialects of the American English (' dr1 ' to ' dr8 ') and pronouncing each one 10 sentences. These sentences come from 3 corpus:

- Two sentences of calibration, pronounced by all the speakers are used to illustrate the variations of the dialects ('sa1 ' and ' sa2 ').

- Five sentences are taken randomly among 450 phonetically balanced and compact sentences conceived with MIT (identified ' sx3 ' with ' sx452 ') [29]. Each sentence is pronounced by 7 different speakers.

- Three sentences are selected to maximize the acoustic contexts, each sentence is marked only one time, On the hole 1890 different sentences for the 630 speakers (identified ' si453 ' to ' si2343 ') phonetically various selected with TI.

The total vocabulary of the base is 6300 sentences. The 630 speakers of the base (438 men and 192 women) are divided between the whole of training (462 speakers including 326 men and 136 women) and the whole of test (168 speakers including 112 men and 56 women). For each sentence, we have the English text, the sampled signal with 16 KHz with a resolution of 16 bits, segmentation in words and phonemic classification in 61 classes.

In our experiments, we have used the New England dialect region (DR1) composed of 24 male and 14 female. The corpus contains 7380 word units for training. Each word unit is represented by 9 frames selected at the middle of each word to generate data vectors. Training has been made on words for ten sentences of TIMIT database. The sentences and words can be found in table 1.

Sentences	Words
SA1	{'she'} {'had'} {'your'} {'dark'} {'suit'} {'in'} {'greasy'} {'wash'} {'water'} {'all'} {'year'}
SA2	{'don't'} {'ask'} {'me'} {'to'} {'carry'} {'an'} {'oily'} {'rag'} {'like'} {'that'}
SX56	{'academic'} {'aptitude'} {'guarantees'} {'your'} {'diploma'}
SI1377	{'as'} {'these'} {'maladies'} {'overlap'} {'so'} {'must'} {'the'} {'cure'}
SX395	{'i'} {'took'} {'her'} {'word'} {'for'} {'it'} {'but'} {'is'} {'she'} {'really'} {'going'} {'with'} {'you'}
SI921	{'differences'} {'were'} {'related'} {'to'} {'social'} {'economic'} {'and'} {'educational'} {'backgrounds'}
SI1027	{'even'} {'then'} {'if'} {'she'} {'took'} {'one'} {'step'} {'forward'} {'he'} {'could'} {'catch'} {'her'}
SX159	{'the'} {'government'} {'sought'} {'authorization'} {'of'} {'his'} {'citizenship'}
SX117	{'the'} {'mango'} {'and'} {'the'} {'papaya'} {'are'} {'in'} {'a'} {'bowl'}
SI1244	{'the'} {'sculptor'} {'looked'} {'at'} {'him'} {'bugeyed'} {'and'} {'amazed'} {'angry'}

TABLE 1: List of words of each sentence

Table 2 shows the number of samples of training data set of TIMIT speech corpus and the size of map for each sentence.

Sentences	Number of samples of training data set	Size of map
SA1	3735	20*15
SA2	3028	19*15
SX56	42	7*5
SI1377	72	7*6
SX395	109	9*6
SI921	81	8*6
SI1027	108	9*6
SX159	59	8*5
SX117	66	8*5
SI1244	80	9*5

TABLE 2: Number of samples of training data set of TIMIT speech corpus and the size of map

According to table 3, RSSOM provides the best classification rate 73.70%. With RSSOM we obtained an improvement of the classification rate in order to 16 % in comparison with SOM.

It is also noticed that for the sentence 'SA1', the three variants (SSOM, RSSOM and LIN) have capacities of roughly similar recognition.

Word	SOM	RSOM	SSOM	RSSOM	LIN
she	80.11	82.16	84.21	84.21	76.02
had	73.97	83.62	87.13	88.30	86.25
your	36.47	42.85	52.88	54.71	58.53
dark	65.78	82.74	84.50	89.47	90.35
suit	60.52	75.73	80.40	79.82	71.34
in	74.08	76.52	85.06	81.09	74.08
greasy	31.28	51.16	54.97	50.87	50.87
wash	46.78	65.49	55.55	61.98	71.05
water	37.71	45.02	48.24	60.52	67.54
all	74.26	73.68	80.11	79.23	78.94
year	55.26	77.77	77.77	80.11	76.31
Average	**57.85**	**68.86**	**71.91**	**73.70**	**72.87**

TABLE 3: Sentence SA1 recognition rates

Table 4 shows that RSSOM provide the best recognition accuracy in order to 69.11%. LIN provides best rate for the word /rag/ in order to 83.49%

Word	SOM	RSOM	SSOM	RSSOM	LIN
don't	67.30	68.88	80.63	75.87	80.63
ask	46.98	51.11	57.77	60.00	56.82
me	61.05	56.84	74.03	62.10	63.85
to	44.63	57.04	62.41	76.84	65.77
carry	57.14	54.28	65.07	69.20	66.98
an	47.95	61.56	64.28	68.02	75.17
oily	58.41	67.93	68.88	76.19	75.55
rag	62.22	73.96	80.00	82.22	83.49
like	19.36	33.33	50.79	38.73	33.33
that	72.06	71.42	78.41	77.46	77.14
Average	**52.30**	**58.79**	**68.20**	**69.11**	**67.71**

TABLE 4: Sentence SA2 recognition rates

From table 5 RSSOM, LIN and SSOM provide best classification accuracy in comparison with SOM. The variant LIN reaches good classification rates in order to 97.22%. However, this model reaches good recognition rates (in the range of 90 and 100%).

We also note that for the words /step/ and /her/ the SOM provide recognition rates in the range of 10% and 30%, on the other hand the models LIN and RSSOM provide a higher value of recognition rate in order to 100%.

TABLE 5: Sentence SI1027 recognition rates

Word	SOM	RSOM	SSOM	RSSOM	LIN
even	100	100	100	100	100
then	100	100	100	100	100
if	88.88	77.77	100	100	100
she	66.66	88.88	77.77	77.77	88.88
took	88.88	100	100	88.88	88.88
one	88.88	100	100	100	88.88
step	33.33	100	88.88	100	100
forward	100	100	100	100	100
he	88.88	88.88	100	88.88	100
could	100	88.88	100	100	100
catch	100	100	100	100	100
her	11.11	77.7	88.88	100	100
Average	**80.55**	**93.51**	**96.29**	**96.29**	**97.22**

From table 6 with RSSOM we obtained an improvement of the classification rate of 14 % in comparison with SOM and 21% in comparison with RSOM. RSSOM reaches good classification rates in order to 92.85% in training set.

For the word /guarantees/ RSOM has the low recognition rate in order to 10%, but with the recurrent spiking SOM (RSSOM) model we obtained a recognition rate in order to 70%, this result prove the stability and performance of the variant RSOM.

Word	SOM	RSOM	SSOM	RSSOM	LIN
academic	66.66	77.77	100	100	100
aptitude	100	100	88.88	100	100
guarantees	33.33	11.11	44.44	66.66	55.55
your	100	66.66	100	100	100
diploma	100	100	100	100	100
Average	**78.57**	**71.42**	**85.71**	**92.85**	**90.47**

TABLE 6: Sentence SX56 recognition rates

According to table 7, Leaky integrators neurons (LIN) and recurrent spiking SOM (RSSOM) provide best classification rate in order to 100% in training set.

SOM model gives the weak result for the word /so/ in order to 44%, on the other hand, all the other models provide a higher value of recognition rate in order to 100%.

Word	SOM	RSOM	SSOM	RSSOM	LIN
as	100	100	100	100	100
these	100	100	100	100	100
maladies	88.88	77.77	100	100	100
overlap	100	100	100	100	100
so	44.44	100	88.88	100	100
must	100	88.88	77.77	100	100
the	100	88.88	100	100	100
cure	100	100	100	100	100
Average	**91.66**	**94.44**	**95.83**	**100**	**100**

TABLE 7: Sentence SI1377 recognition rates

According to table 8, the variants LIN and RSSOM provide best classification rate in order to 100% in training set.

Word	SOM	RSOM	SSOM	RSSOM	LIN
the	100	62.50	100	100	100
mango	100	100	100	100	100
and	66.66	44.44	100	100	100
papaya	100	88.88	100	100	100
are	100	77.77	100	100	100
in	88.88	100	100	100	100
a	100	100	75.00	100	100
bowl	100	100	100	100	100
Average	**93.93**	**84.84**	**98.48**	**100**	**100**

TABLE 8: Sentence SX117 recognition rates

According to table 9, Leaky Integrators Neurons (LIN) provides the best classification rate in order to 96.25%. With LIN we obtained an improvement of the classification rate in order to 19 % in comparison with SOM and RSOM in training set. A higher value of recognition rate for the words /looked/, / bugeyed/, /and/ and /him/ means a better performance of variant SOM.

Word	SOM	RSOM	SSOM	RSSOM	LIN
the	87.50	87.50	87.50	100	100
sculptor	100	55.55	100	100	100
looked	66.66	77.77	88.88	88.88	100
at	88.88	88.88	100	88.88	100
him	77.77	55.55	100	100	100
bugeyed	44.44	77.77	100	100	100
and	44.44	66.66	66.66	88.88	66.66
amazed	88.88	100	88.88	88.88	100
angry	100	100	100	100	100
Average	**77.50**	**78.75**	**92.50**	**95.00**	**96.25**

TABLE 9: Sentence SI1244 recognition rates

According to table 10, LIN and RSSOM models provide best classification rate in order to 97.24% in training set. With RSSOM and LIN we obtained an improvement of the classification rate in order to 30 % in comparison with SOM.

With SOM model we can’t recognize some words like /for/ and /is/, recognition rates (in the range of 0 and 20%), on the other hand the models LIN and RSSOM provide a higher value of recognition rate (in the range of 90 and 100%)

Word	SOM	RSOM	SSOM	RSSOM	LIN
i	100	77.77	100	100	100
took	88.88	77.77	100	100	100
her	77.77	88.88	100	100	100
word	77.77	77.77	100	100	100
for	0.00	77.77	55.55	100	88.88
it	60.00	100	80.00	100	100
but	77.77	77.77	88.88	100	100
is	20.00	100	100	80.00	80.00
she	100	100	100	100	100
really	44.44	88.88	77.77	88.88	100
going	55.55	100	100	100	100
with	55.55	88.88	88.88	88.88	88.88
you	77.77	100	88.88	100	100
Average	**66.05**	**87.15**	**90.82**	**97.24**	**97.24**

TABLE 10 : Sentence SX395 recognition rates

According to table 11, the variant LIN provide best classification rate in order to 100% in training set. Table 11 shows that for words / related / and / backgrounds / classic SOM and RSOM are not able to recognize these classes (small recognition rates). However, SSOM, RSSOM and LIN reach good recognition rates (in the range of 90 and 100%).

Word	SOM	RSOM	SSOM	RSSOM	LIN
differences	100	100	100	100	100
were	100	100	100	100	100
related	55.55	44.44	88.88	100	100
to	100	100	100	100	100
social	100	100	100	100	100
economic	88.88	88.88	88.88	100	100
and	100	100	88.88	77.77	100
educational	100	100	88.88	100	100
backgrounds	66.66	100	100	100	100
Average	**90.12**	**92.59**	**95.06**	**97.53**	**100**

TABLE 11 : Sentence SI921 recognition rates

Table 12 shows that LIN provides the best recognition accuracy in order to 98.30%. With LIN we obtained an improvement of the classification rate in order to 14 % in comparison with SOM.
It is also noticed that for the sentence 'SX159', the three variants (SSOM, RSSOM and LIN) have capacities of roughly similar recognition.

Word	SOM	RSOM	SSOM	RSSOM	LIN
the	80.00	80.00	100	86.92	100
government	100	100	100	100	100
sought	100	100	88.88	100	100
authorization	66.66	88.88	100	100	100
of	88.88	100	100	100	100
his	55.55	44.44	66.66	77.77	88.88
citizenship	100	100	100	100	100
Average	**84.74**	**88.13**	**93.22**	**96.61**	**98.30**

TABLE 12 : Sentence SX159 recognition rates

6. CONCLUSION

In this paper, we have proposed new variants of self organizing neural network algorithm in the unsupervised learning category, and we are interested in word recognition using sentences from TIMIT databases by means of new SOM variants with impulse neurons.

The use of impulse neurons in the SOM makes it possible to establish temporal associations between the consecutive models in a temporal order through the impulses produced according to the entry and makes it possible to improve the taking into account of the temporal parameters in the recurring SOM.

In this paper, we have presented three SOM variants, Leaky Integrators Neurons model (LIN) which consider the temporal order between the successive samples by using a mechanism called Leaky Integrators, In this approach, the state of each neuron is performed by a membrane potential which is function of the input, this potential measure the adaptation degree between the neuron weight vector and the current input vector. Then the Spiking_SOM model (SSOM), based on the timing or the order of single spike events, the coding which represents information is through the differences in the firing times of different neurons and the recurrent Spiking_SOM (RSSOM) has the same principle that spiking self organizing map except that the choice of the BMU is defined by a difference vector in each unit of the map.

The proposed SOM variants provide best classification rates in comparison with the basic SOM and RSOM models. The RSSOM and LIN provide best classification rates in order to 100%.

As a future work, we suggest proposing other SOM variant, the Growing Hierarchical Self-Organizing map (GHSOM) using spiking neurons. GHSOM is a network of neurons whose architecture combines two principal extensions of SOM model, the dynamic growth and the tree structure in order to reduce the complexity of the task of classification and improve classification rates.

7. REFERENCES

[1] W. Maass, M. Schmitt. “On the complexity of learning for a spiking neuron”. In COLT’97, Conf. on Computational Learning Theory, ACM Press, 1997. pp. 54–61.

[2] W. Gerstner. “What´s different with spiking neurons”. In Henk Mastebroek and Hans Vos, editors, Plausible Neural Networks for Biological Modelling, Kluwer Academic Publishers;2001. pp. 23–48.

[3] W. Maass, C. M Bishop. “Pulsed Neural Networks”. MIT Press, 1999.

[4] R. Kempter, W. Gerstner. J. L VanHemmen and H. Wagner. “Extracting oscillations: Neuronal coincidence detection with noisy periodic spike input”. Neural Comput., vol. 10, pp. 1987-2017, 1998.

[5] W. R Softky, C. Koch. “The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs”. J. Neurosci. Vol. 13, pp. 334–350, 1993.

[6] F. Santiage , G. Alex, S. Jurgen. “Phoneme recognition with BLSIM-CIC”. 2008.

[7] S. Durand. “Réseaux neuromimétiques spatio-temporels pour l’organisation des sens. Application à la parole. Dans Actes Rencontres Nationales des Jeunes chercheurs en Intelligence Artificielle. Marseille, 1994.

[8] S. Durand. “TOM, une architecture connexionniste de traitement de séquence. Application à la reconnaissance de la parole”. PhD thesis Université Henri Poincaré, Nancy I ; 1995.

[9] P. Danilo, Mandic, A. Jonathon. “Chambers, Recurrent Neural Networks for Prediction”, John Wiley and Sons Ltd, 2001.

[10] G Vaucher. “Un modèle de neurone artificiel conçu pour l’apprentissage non supervise de séquences d’événements asynchrones”. In Revue VALGO, ISSN 1243-4825, Vol 1, pp. 66–107, 1993.

[11] T. Behi, N. Arous. “Modèle auto-organisateur à composante temporelle pour la reconnaissance de la parole continue”. Huitième journée scientifiques des jeunes chercheurs en génie électrique et informatique, GEI2008, Sousse-Tunisie, 2008.

[12] T. Behi, N. Arous. “Modèles auto-organisateur à apprentissage spatio-temporels Evaluation dans le domaine de la classification phonémique”. Cinquième conférence internationale JTEA2008, Hammamet-Tunisie , 2008.

[13] R. Brette. “Modèles Impulsionnels de Réseaux de Neurones Biologiques”. PhD Thesis, University of Cerveau-Cognition – Comporteme, 2003.

[14] 14 T, Kohonen. “Self_Organized Formation of Topologically Correct Feature Maps”. Biological Cybernetics. Vol.43, pp. 59-69,1982.

[15] N. Arous. “Hybridation des Cartes de Kohonen par les algorithmes génétiques pour la classification phonémique. PhD Thesis Ecole Nationale d’ingénieurs de Tunis , 2003.

[16] S. Haykin. “Neural Network A Comprehensive Foundation”, Prentice Hall Upper Saddle River, New Jersey, 1999.

[17] T. Kohonen. “Self-organizing map”, third edition, Springer, 2003.

[18] W. Maass. “Computing with spiking neurons”.In Maass, W. and Bishop, C. M., editors, Pulsed Neural Networks, chapter 2, MIT-Press., pp. 55-85 (1998)

[19] W. Maass, C. M. Bishop. “Pulsed Neural Networks”. The MIT Press, 1st edition, Cambridge, 1998.

[20] M. Varsta, J. Heikkonen and R; Milan. “A recurrent self-organizing map for temporal sequence processing”. Proc. Int. Conf. on Artificial Neural Networks (ICANNP'P97), Lausanne, Switzerland, 1997.

[21] T. Koskela, M. Varsta, J. Heikkonen, K. Kaski. “Time Series prediction using recurrent SOM with local linear models”. International Journal of Knowledge-based Intelligent Engineering Systems,2(1), : 60-68, 1998.

[22] J.G. Taylor. “Temporal patterns and leaky integrator neurons”. Proc. Int. Conference. Neural Networks (ICNN'90), Paris, 1990.

[23] T. koskela, M. varsta. “Recurrent SOM with local linear models in time series prediction”, PhD Thesis, University of Helsinki university of technologie-labo of computational engineering-Finland, April 1998.

[24] M. Varsta. “Temporal sequence processing using recurrent SOM”, PhD Thesis, University of Helsinki university of technologie-labo of computational engineering-Finland, 1998.

[25] T. Voegtlin. “Réseaux de neurones et autoréférence”, PhD Thesis, University of lumière lyon II, 2004.

[26] N. Arous, N. Ellouze. “Phoneme classification accuracy improvements by means of new variants of unsupervised learning neural networks”, 6th World Multiconference on Systematics, Cybernetics and Informatics, Floride, USA, 2002.

[27] N. Arous, N. Ellouze. “Cooperative supervised and unsupervised learning algorithm for phoneme recognition in continuous speech and speaker-independent context”, Elsevier Science, Neurocomputing, Special Issue on Neural Pattern Recognition, vol. 51, pp. 225 – 235, 2003.

[28] J. Garofalo, L. Lamel, W. Fisher, J. Fiscus, D. Pallett, N. Dahlgren and V. Zue. “TIMIT acoustic-phonetic continuous speech corpus”, Linguistic Data Consort, 2005.

[29] V. Zue, S. Seneff, J. Glass. “Speech database development at MIT, TIMIT, and beyond”, Speech Commun, vol. 9, pp. 351–356, 1990.