Study of Robust Arabic handwriting recognition by Neural Network Based on Single Electron Transistors

Abstract Thisstudy proposes a new technique of Arabic Handwriting Recognition using multilayer perceptron neural network circuit based on single electron transistor. The basic idea of the present study is how to integrate a single electron transistor in neural network instead of conventional electronics for better robust recognition of Arabic letters and words using a new architecture based on single electron transistors. The simulations results are analysed in the form of the average recognition rates and variances of handwritten Arabic letters. A study of robustness against different noisy images willbepresented with recognition rates specific to each letter, group of letters and Arabic words.

In recent years, some research has been done on the problem of offline Arabic handwriting recognition [1-2]. The lack of Arabic handwritten systems and the absence of nano-electronics recognition systems in the future of nanocircuits based on single electron transistors (SET) are some of the reasons for the scarcity of research on Robust Arabic handwriting recognition. A winner-take-all (WTA) network is considered as competitive neural network for implementing a nanoelectronicanalog neural processor. In the recent years, The SET–WTA neuron circuit for a lateral inhibition configuration has attracted researchers[3].The combining neural network completely designed using SET devices on Arabic recognition was simulated using building blocks by SIMON software [4].

In this study, we combined the advantages of SET and the neural network and more specifically multilayer perceptrons paradigms in order to take the advantage from the properties of both nanoelectronic devices and WTA network (flexibility, learning capabilities and recognition performance).

The SET is a solid-state electronic device which controls the transport of unique or few electrons[5].We propose, for the first time, a new neural network based on 9 SET-WTA neuron circuits developing an appropriate activation function[6]. A lateral inhibition configuration is simulated with SIMON for offline Arabic handwriting recognitions[7].
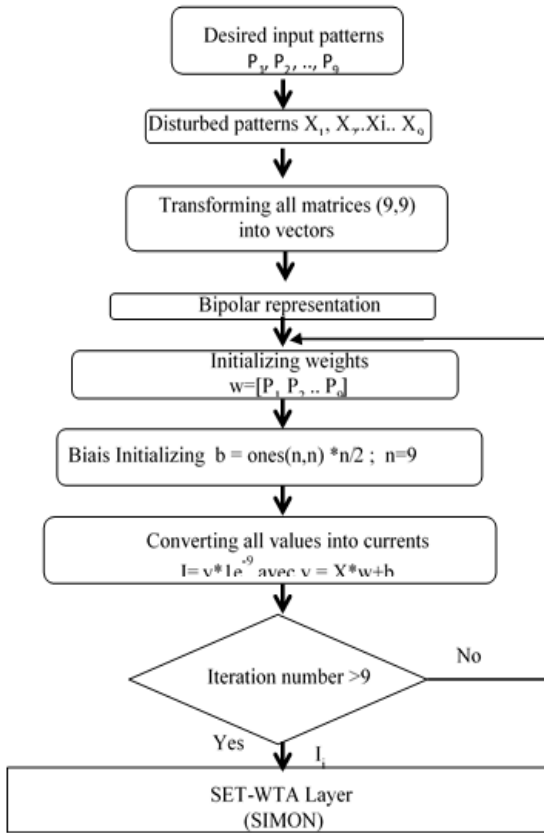
The neuron's output is a function of its input current $I_i$and of the output voltages from the nearest neurons in the network. Currents I_i traverse neuron from the current source and bring the data information to the network, and each output voltage spreads to the following neighbour's SET-WTA[8]..This choice is justified by the little progress achieved in offline recognition of Arabic letters and words.

Each pattern represents an Arabic handwritten letter or word in 9*9 matrix. The next step, in pre-processing layer, is transforming each exemplar and disturbing letter/word matrix.The pre-processing layer processes different samples of letters or words and makes them ready to be exploited in the neural network training. Fig.1 describes our implemented MATLAB code represented in block diagram modelling the pre-processing layer .

The outputs provided from the pre-processing layer, shown in Fig 1, will be processed SIMON in order to know which one is the winner input and, after that, we calculate a recognition rate using MATLAB. In fact, the most practical handwriting recognition systems consist of two modules.

The first is the front-end feature module and the second is the back end classifier using a multilayer perceptron neural network.

Desired input patterns
$P_1, P_2, .., P_9$

↓

Disturbed patterns $X_1, X_2.Xi.. X_9$

↓

Transforming all matrices (9,9) into vectors

↓

Bipolar representation

↓

Initializing weights
$w=[P_1, P_2, .. P_9]$

↓

Biais Initializing $b = ones(n,n) *n/2$; $n=9$

↓

Converting all values into currents
$I= v*1e^{-9}$ avec $v = X*w+b$

↓

Iteration number >9 — No

Yes ↓ I_i

SET-WTA Layer (SIMON)

The recognizer system was trained and tested with our Arabic database[9]. It consists of 28 letters at different handwritten forms and perturbations (1028 samples) and 96 words. All trainings were carried out using 100 letters of clean data. A set of 50 letterswas used for testing. We note that the recognition rate of a group of letters in a network with 3 neurons is lower than that of letters distributed letter by letter in a network with 4 neurons. The letter complexity influences the recognition rate of a neural network. We conclude that the smaller the number of neurons in a network, smaller the number of erroneous letters and smaller the bigger the recognition rate of neural network is. The complexity of the neural networks influences clearly the recognition rate. In order to observe more the complexity of the letter shape, we have found variance and gap type of each letter which increase with the complexity of group-letter. Finally, the smaller the letter shape, the more reliable the recognition will be.

Fig 1 Flow diagram of pre-processing layer

Table 1 Performance of Arabic alphabet Recognition System

RR(%) / Arabic alphabet	Group of letters in a network with 3 neurons	Recognition letter one by one in a network with 3 neurons	4 erroneous letters in a network with 4 neurons	9 erroneous letters in a network with 9 neurons
أ	71.216	72.850	70.000	60.107
ث ت ب	90.278	95.000	81.427	66.800
خ ح ج	79.000	85.900	97.520	57.348
ذ د	**98.363**	**98.363**	**89.000**	**84.387**
ز ر	**98.363**	**98.363**	**76.500**	**54.580**
ش س	73.118	73.500	96.952	56.728
ض ص	65.197	66.047	95.676	60.426
ظ ط	**94.030**	**97.255**	**94.000**	**93.166**
غ ع	51.850	73.656	95.670	65.287
ق ف	55.060	50.000	96.500	65.294
ك	**42.373**	**68.764**	**37.064**	**45.650**
ل	72.631	84.382	96.227	44.814
م	90.435	92.823	92.251	86.7693
ن	41.873	91.619	91.619	39.700
ه	61.270	79.536	77.561	71.225
و	**96.645**	**97.103**	**91.136**	**96.430**
ي	48.700	66.980	59.173	58.345
Average of RR (%)	**72,730**	**81.710**	**84.330**	**65.120**

We have developed our Arabic database which consists of 24 Arabic words handwritten by 4 different writers. Our aim was to collect an Arabic database of handwritten town names written in a similar quality for postal sorting. After the smoothing and detection of the reference line, the off-line handwriting is pre-segmented into strokes and sub-strokes. A MATLAB code can segment the image horizontally by extracting the sub images that are separated from the adjacent letters.After that, we separate the modifiers and segment image to get feature extraction.

Patterns are disturbed in recognition system based on the hybrid configuration based on neural network and SET.

We note that the RR of Arabic words is great for non-complex words. For town names with complex letters which have low RR separately (such as the letter ي or ك), there is a low RR (about 50% for 4 perturbed patterns). For each word, representingof 24 Tunisian city names with four handwritten writers, there were 28 training and 4 test instances. We can conclude that a correct RR of over 98.7 % can considered quite satisfactory in the context of clearly words.

The RR of our Arabic words handwritten systemhas a higher value than only Neural network techniques which used for Arabic handwriting recognition. [10]

We applied some noises and used three filter types. The first is the linear Middle filter. The second is the Gaussian filter. The third is the non-linear median filter. The peak signal-to-noise ratio (PSNR) criterion is adopted to measure the restoration performance quantitatively. Fig.2shows the comparative PSNR results of the restored Arabic letters with different filter types.
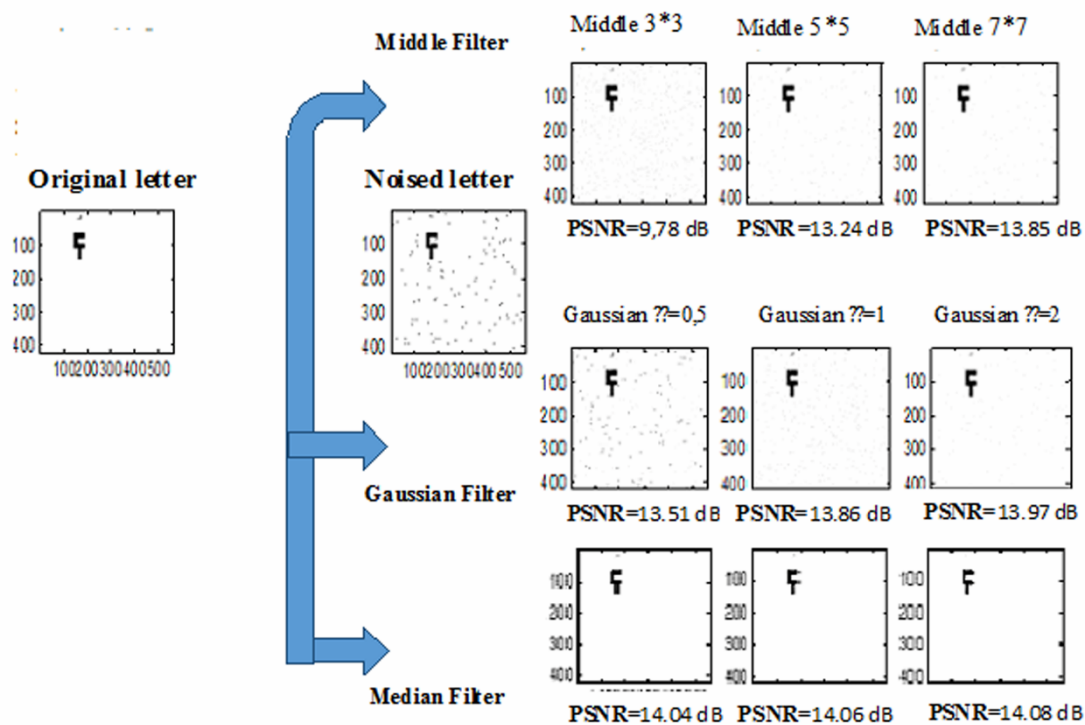
Fig 2 Effect of filtering noised Arabic letters and comparison of PSNR criterion

We used a dictionary consisting of 24* 4 writers Tunisian city names with four writers of each town name is considered in our experiments. After training with the SET-WTA circuit and tested our recognition system. The recognition rate deteriorates with additional written words if they are already noisy, the recognition decreases with Gaussian noises. For this reason, a word filtering system seems necessary in order to eliminate noise to improve RR. From the obtained results shown in Fig.3, we noticed that the middle filter keeps the sharpness of the image for large dimensions.
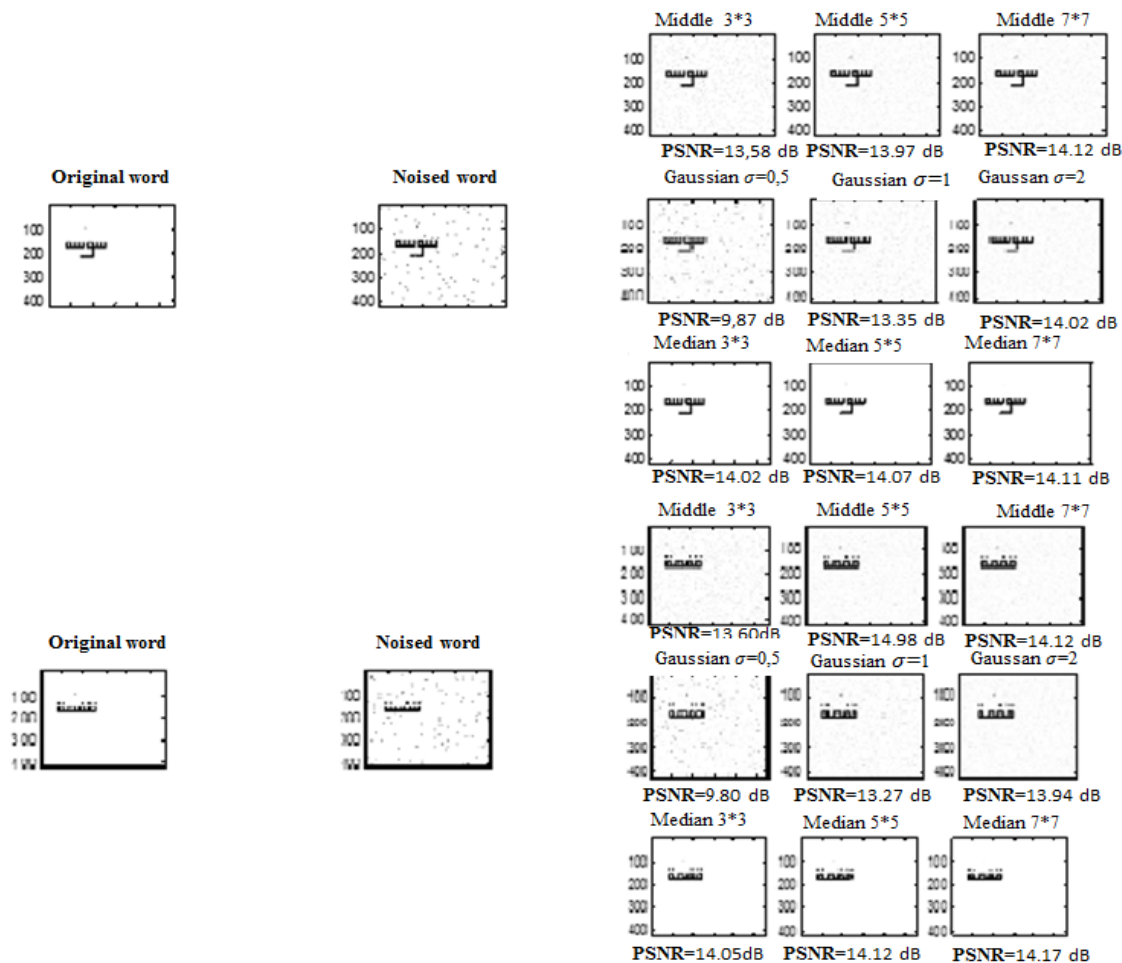
Fig. 3 Effect of filtering noised Arabic town names and comparison of PSNR criterion

The performances are detailed according to the different forms of the handwritten letters or words.Results showed an important improvement in term of RR. We investigated the robustness of the conceived neural network based on SET-WTA architecture when tested data were contaminated by specific noises.

A robustness study has been detailed and numerous filter models are compared to ensure a better quality denoising.

Conflict of interest

The author confirms that this article content has no conflict of interest.

Acknowledgements

Declared none.

References

1. Benouareth A, Ennaji A, Sellami M. (2006) HMMs with explicit state duration applied to handwritten Arabic word recognition. ICPR Hong Kong: 897-900.
2. *H*ajj,R. Al-Hajj,C. Mokbel,C. Mokbel,L. Likforman-Sulem,L. Likforman-Sulem.(2007) Combination of HMM-based classifiers for the recognition of Arabic handwritten words. ICDAR(02) : 959-963
3. Guimarães JG, do Carmo H.C, da CostaJC(2004)Single-electron winner-take-all network, Microelectronics Journal 35(2):173–178.
4. WasshuberC,KosinaH, SelberherrS (1997) Simon a Simulator for single electron tunnel devices and circuits. IEEE Trans. ComputerAided Design Integr. Circuits Syst(16):937 – 944.
5. Pott V (2007) Gate-all-around silicon nanowires for hybrid single electron

transistor/CMOSapplications. Thesis EPFL, Lausanne.

6. Guimarães JG, do Carmo H.C, da CostaJC(2004)Single-electron winner-take-all network, Microelectronics Journal 35(2):173–178.

7. Abdel AzeemS, AhmedH (2013) Effective technique for the recognition of offline Arabic handwritten words using hidden Markov models.IJDAR, 16(4), pp. 399-412.

8. Lientschnig G, Weymann I, Hadley P (2003)Simulating Hybrid Circuits of Single-Electron Transistors and Field-Effect Transistors.Jpn. J. Appl. Phys 42(10): 6467–6472

9. AbedH E, KherallahM, MârgnerV,Alimi A. M (2010) On-line Arabic handwriting recognition competition ADAB database and participating systems.IJDAR 14(1): 15-23.