International Journal of Scientific & Engineering Research (IJSER), January 2013
ISSN 2229-5518

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO
EN INGENIERÍA

FACULTAD DE INGENIERÍA

NUMBER OF HIDDEN NEURONS IN A RECURRENT NEURAL NETWORK ESTIMATION

T E S I S

QUE PARA OPTAR POR EL GRADO DE:

DOCTOR EN INGENIERÍA

SISTEMAS – INVESTIGACIÓN DE OPERACIONES

PRESENTA:

PAOLA LUCÍA TELLEZ BALLESTEROS

TUTORES:
MANUEL ORDORICA MELLADO
ANGEL KURI MORALES
JOSE DE JESUS ACOSTA FLORES

1 CONTENTS

2 TABLES

ISSN 2229-5518

3 PICTURES

4 FIGURES

5 ABSTRACT

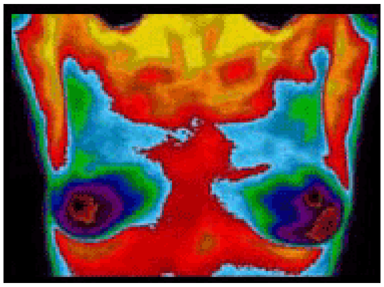
Genetic algorithms are an alternative for optimizing the performance of recurrent neural networks by searching for the best number of hidden neurons. Genetic algorithms have been used for heuristics, but never to find the best number of hidden neurons in a recurrent neural network. Three architectures of recurrent neural networks were used to measure the performance with spoken Spanish digits. 13 was the number of hidden neurons used for a Jordan network to give the best performance. This number lets to optimize resources in hardware implantation.

ISSN 2229-5518

6 INTRODUCTION

This thesis is the result of painful months for my family, caused by my mother cancer. She used to sleep with the cell during the last 12 years, to be able to wake up on time in the morning. The cancer was detected when nothing could be done. One month was the remaining time of her life.

It takes 8-10 years for a dime-size tumor to grow (1). Statistics indicate that 15% of all breast cancers occur between the ages of 20 and 44, and the risk increases with age. Current research indicates that 1 of 8 women in the U.S. will get breast cancer in their lifetime. Women without children and also who have had their first child after age 30 seem to be at higher risk (2).



Thermograph 1. Breast warming

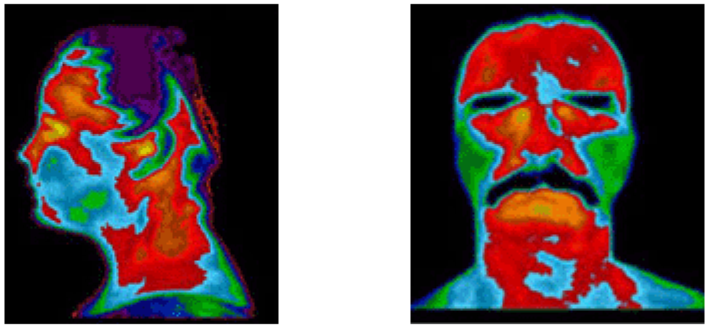
Electromagnetic fields caused by the cell and another electromagnetic appliances are a main research problem. The body warming provoked by electromagnetic forces has fast solutions, but they are expensive in any case. It is compulsory to invest in new alternatives research for the growing problems in public health. We must optimize our electrical appliances.

Neural networks can be implanted electronically, but they must be optimized first. There are a lot of ways of doing it. It is only one example that the statistics and operations research tools let us to improve our electronic appliances. The problem is not new, and it must be attacked from the root. Developing countries have invested great incomes for making it a reality.

6.1 Problem exposition and its importance

Sophisticated cameras and computer systems can be used to measure heat from the surface of the body and produce an image that can be evaluated for abnormalities. It can be showed in the Thermograph 2. Alterations in these images are caused when cellular changes increase blood flow, by warming the body. These changes may be among the earliest signs of cancerous formation.

Body thermograph has been used to find signs of pre-cancerous tissue, or early stage cancer; that are difficult to find by physical exam or mammography, with no use of radiation so patients can be monitored as often as necessary. Hotter temperatures show up as reds, oranges and yellows and those would be a cause for concern. Thermal Imaging is able to pin-point precisely where a mass is forming for additional testing. Looking abnormal heat changes produced by a diseased body allows for extremely early cancer detection.



Thermograph 2. Heating in body surface

A positive infrared image may indicate the presence of many different breast diseases such as mastitis, benign tumors, fibrocystic breast disease, cancer and others. In patients without cancer, the examination results are used to indicate the level of possible future cancer risk. This gives a woman time to take a pro-active approach to her breast health by initiating anti-carcinogenic lifestyle modifications and decreasing as many known risk factors as possible. If cancer is suspected, this information is used to direct further examinations and tests to insure prompt treatment.

6.2 Proof

This kind of technology lets us to watch the warming caused by a mobile telephone in the head. The Thermograph 3 shows the human head at the beginning of the call, after one minute and half, after three minutes and after four minutes. Electromagnetic fields are generated also from microwave ovens, hairdryers, the electric wiring in the house, and remote control devices, by example. They are also generated by computer screens, industrial electric furnaces, electric motors in the workplace (3).

Mobile telephone technology produces microwaves in two ways: from the antennas that are placed around our cities, towns, and motorways; and from the telephones themselves which are also antennas. Human bodies are transparent to some frequencies and not to others. Energy is also important, because we use sunscreen creams to protect us from sunlight. For similar reasons there are recommended exposure limits for certain electromagnetic fields.

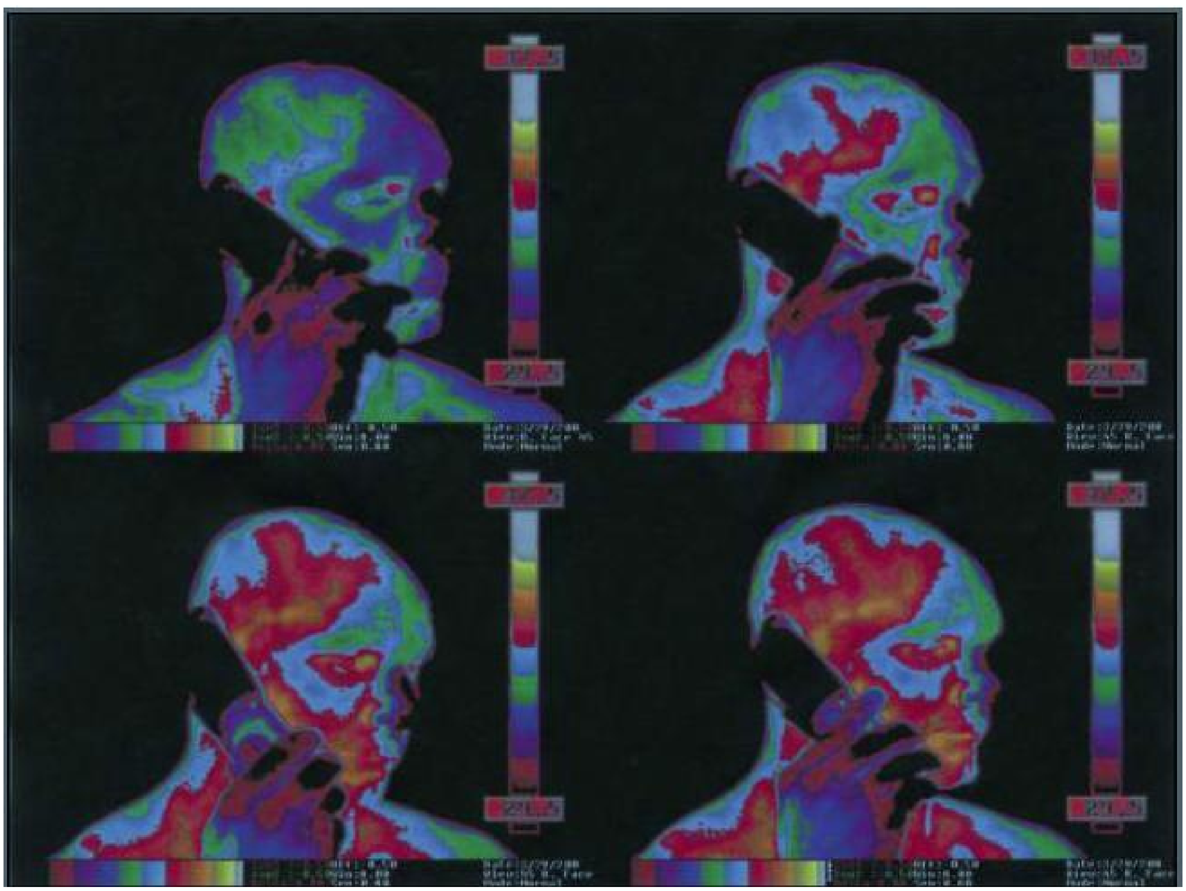
The International Commission on Non-Ionizing Radiation Protection (ICNIRP) gives the guidelines for current European Union exposure limits. Mexico does not have any commission to regulate these limits, by this reason this country receives all electromagnetic appliances that are not accepted inside U.S.A. by surpassing these exposure limits. The Table 1 and the Table 2 show the typical field strengths from European household appliances compared to ICNIRP recommended limits.

Table 1. Electric fields of European appliances

Electrical appliance	Magnetic field strength in volts/meter At 30 cm	ICNRP recommended exposure limit in volts/metre
Stereo receiver	180	5000
Electric iron	120	5000
Toaster	80	5000
Electric oven	8	5000

Table 2. Magnetic fields for European appliances

Electrical appliance	Magnetic field strength in volts/meter At 30 cm	ICNRP recommended exposure limit in volts/meter
Stereo receiver	180	5000
Electric iron	120	5000
Toaster	80	5000
Electric oven	8	5000



Thermograph 3. Head with mobile telephone

6.3 Objective

There are neural networks books that show how to use statistics to improve the neural networks performance (4). Watching in future, we must optimize all kind of neural networks because they can be implanted in electronic appliances for handicapped people. We look with this thesis to show one way to improve the recurrent neural networks performance. The application can be any for disabled people, but word recognition was used.

6.4 Next chapters explanation

The next chapter shows what the recurrent neural networks are. There are different kinds of them, and also there are some more efficient than others. The next chapter gives a way to improve them by using genetic algorithms. Conclusions are also exposed at the end.

7 RECURRENT NEURAL NETWORKS

7.1 Recurrent Neural Prediction Model

Recurrent neural prediction model was created by one master student of Electro-Communications University (5). It has been used for classification of patterns, so it can be used for any problem that involves pattern recognition in time. There is one recurrent neural network to each pattern to recognize in the time. The input of each network is one frame of a spoken word, and the output is the next frame of this word in the time.

The recurrent neural network is a novel model because has recurrent connections in different layers of the neural network. This lets that the recurrent neural network becomes a dynamical system with a big computational capability (6). The recurrence lets a temporal evaluation of states. There are several architectures to evaluate the time signal: Jordan, Elman, VSRN+ recurrent neural networks.

7.1.1 Elman Model

The Elman network was developed by him in 1990 (7). He has always interested in the brain study. He thought how to simulate it with the computer. He thought that there are connections from the hidden layer to a context layer, which also has connections as input to the hidden layer. This model is shown in the Figure 1.

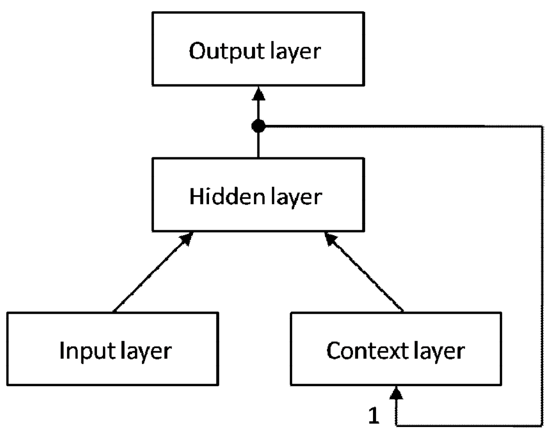


Figure 1. Elman Recurrent Network

7.1.2 Jordan Model

Jordan thought that it is very time consuming for the computer to keep registers to manipulate time signals. He developed a network model capable of computing time signals with a new architecture (8). Jordan retrieved the output from the output layer to the context layer, which is the new input to the hidden layer in the next training time, as it is showed in the Figure 2.

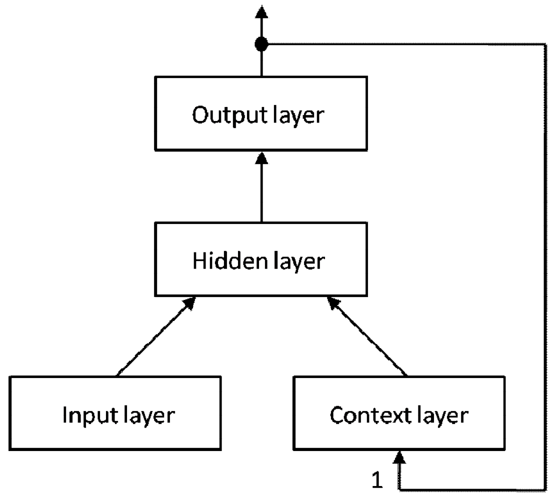


Figure 2. Jordan Recurrent Network

7.1.3 VSRN+ Model

Haruhisha Takahashi also created his own recurrent architecture (9). He studied in the Osaka University and he thought that each layer must have its own context layer to receive the output of this layer, which must keep the input for the same layer for the next training time. This model can be seen in the Figure 3.

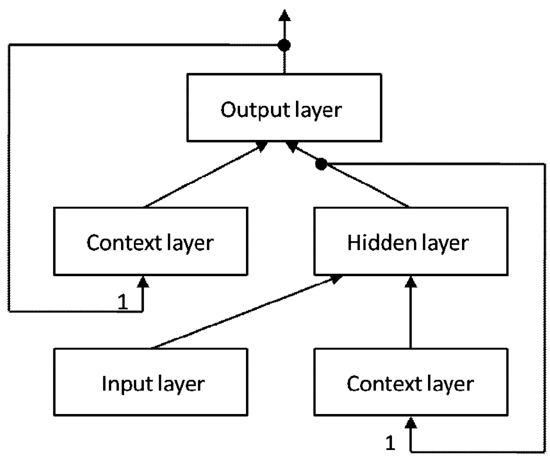


Figure 3. VSRN+ Recurrent Network

7.2 Training method

The multilayer perceptron has traditionally been used in the literature (10). The training method is the same for recurrent networks. It is the back-propagation rule, but now we must take also the input of the context layer in the rule. The training method will train the input signal less the output signal to minimize the prediction error of output layer. The output signal is the predicted signal which really comes as the input signal in the next training time. This rule involves two steps (11):

1. During the first phase the input is propagated forward through the network to compute the output values for each output layer. It is showed in the left side of the Figure 4. This output is compared with its desired value, resulting in an error signal for each output layer.
2. The second phase is a backward pass through the network during which the error signal is passed to each layer in the network and appropriate weight changes are calculated. It is showed in the right side of the Figure 4.

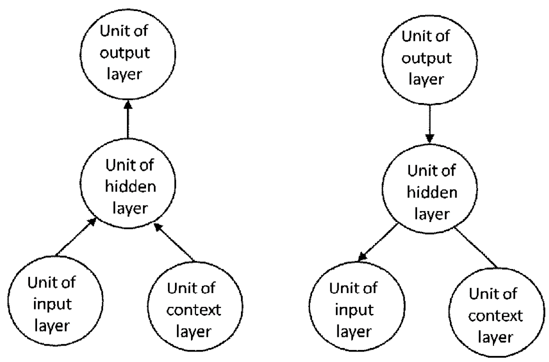


Figure 4. RNN Training

7.3 Evaluation method.

The pattern recognition is developed with the next steps (5):

1. Each recurrent neural network was trained to recognize a time signal. *N* time sequenced feature vectors of length *T*: $a_1, a_2, \ldots, a_T$ are the input of each one of he recurrent neural networks,
2. Each recurrent neural network gives a prediction value for each one of the *N* input vectors: $\hat{a}_1, \hat{a}2, \ldots, \hat{a}_T$. The square difference between the prediction value and the real value gives un prediction error E_c in the equation 1 for each recurrent neural network.

$$\boldsymbol{E_c = \frac{1}{2}\sum_{t=1}^{T}\sum_{j}^{n}(a_{i+1,j} - \hat{a}_{i+1,j})^2} \quad (1)$$

3. The recurrent neural network that has the smallest prediction error is used to classify the category of the input time signal. It is showed in the Figure 5.

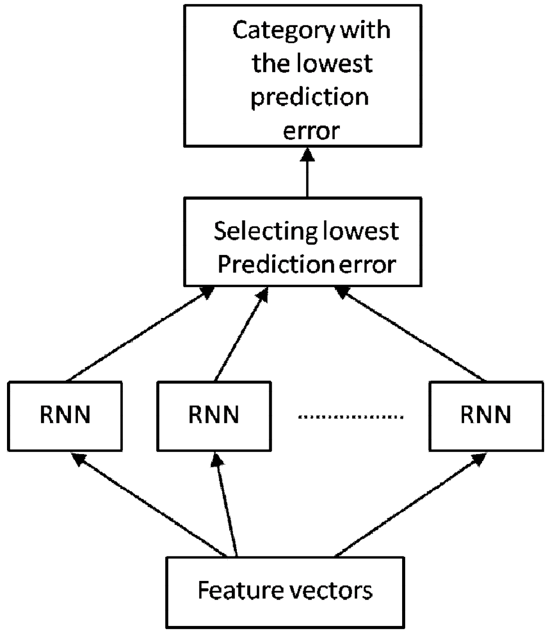


Figure 5. Evaluation Method

8 OPTIMIZING THE NUMBER OF HIDDEN NEURONS

The input time signals were the spoken digits. Each spoken digit was used to train a recurrent neural network. Several architectures were used: Elman or SRN, Jordan and VSRN+. Each one used the Recurrent Neural Prediction Model. The results can be seen in the Table 3 (11). The best performance is given with the Jordan architecture for spoken Spanish digits. Improving these results with genetic algorithms is the objective for this chapter.

Table 3. Spoken Digits Recognition

Seed	Jordan %	SRN %	VSRN+ %
30	96.25	90.00	95.00
31	96.25	93.75	92.50
32	95.00	88.75	93.75
33	96.25	92.50	95.00
34	97.50	91.25	96.25
35	95.00	90.00	93.75
36	93.75	92.50	93.75
37	97.50	90.00	93.75
38	97.50	93.75	95.00

8.1 Genetic algorithms

The number of neurons in the input layer and in the output layer is the same for this kind of neural network. This is determined by the length *T*, in which the signal is divided. So the number of neurons in the hidden layer is a variable. It has been written that the number of neurons can be between 2 and any number. But there is no human being with an endless number of neurons in the brain. Optimizing neurons means using less number of neurons to get better or best performance.

The number of hidden neurons is a variable that must be coded with a binary representation. Each neural network has a binary representation that is the number of neurons in the hidden layer. This variable was watched between its lower values that are between 2 just 17. The steps are the next:

8.1.1 Initial population

N recurrent neural networks are created. Each member with its binary representation that is random. This representation is called genotype. The performance of each member is got with the recurrent neural prediction model. It is called the fitness of each recurrent neural network. All genotypes are ordered according with their fitness value from the worst to the best.

8.1.2 Deterministic Crossover

Crossing the genotype *i* with the genotype *N-i+1*, with *i* since 1 to *N* is done in this step (12). This kind of crossover is called Vasconcelos. It is used to explore the hole Boolean search space. Vasconcelos (13) was a Mexican that improved the Mexican public education system during the Eulalio Gutierrez government since 1915. The result of this step is called offspring.

8.1.3 Mutation

The features of each offspring genotype can be changed randomly. These features are called genes. Each gene can be mutated with a mutation probability. Doing it means to change the genes inherited from their parents genotypes.

8.1.4 Selection

Ordering the offspring and the other genotypes is the next step. The best *N* resulting genotypes are chosen if the objective has not been reached to repeat the process. Otherwise the algorithm ends with the best genotype.

8.2 Results

The Table 4 (14) shows the performance with different recurrent neural network models. The Elman model does not offer a best performance with less neurons. The other two models offer an alternative for optimizing neurons. The Jordan model gives the best recognition performance but uses more neurons in the hidden layer than the VSRN+ model

Table 4. Optimizing with genetic algorithms

	Hidden neurons	Recognition
VSRN+	6	97.50 %
Jordan	13	98.75 %
Elman	17	93.75 %

ISSN 2229-5518

9 CONCLUSIONS

The time will be over in the future. But for now, we must optimize it. There are few people that can attend disable people that are more sensitive to electromagnetic fields. If electronic appliances are improved with word recognition to be used for handicapped people, not only monetary savings can be got, higher life quality would be a plus.

10 REFERENCES

1. International Academy of Clinical Thermology. [Online] [Cited: October 17, 2012.] http://www.iact-org.org.
2. Naturelle Health Systems. [En línea] [Citado el: 17 de October de 2012.] http://www.naturelle.com.mx.
3. **European Comission - Research Directorate - General- European Communities.** *Health and electromagnetic fields.* 2005. SSPE-CT-2004-502173.
4. **Haykin, Simon.** *Neural Networks.* s.l. : Prentice Hall, 1999.
5. **Uchiyama Toru, Takahashi Haruhisha.** *Japanese digits recognition using recurrent neural prediction model.* Tokyo : Electro communications University, 1999. J82-D-11.
6. **Watanabe, Ken-ichi, Iso & Takao.** *Speaker independent word recognition using a neural prediction model.* s.l. : IEEE Proc ICASSP, 1990. págs. 441-444. 8.
7. *Finding structure in time.* **Elman.** 1990, Cognitive Science, Vol. 14, págs. 179-211.
8. **Jordan, Michael.** *Serial order: a parallel distributed processing approach.* San Diego : Institute of Cognitive Science. 8604.
9. **Takahashi, Haruhisha.** *Recurrent Neural Networks.* Osaka : Osaka University, 1993. SP3-111.
10. *Phoneme Recognition using Time-Delay Neural Networks.* **Waibel, Hanazawa, Hinton, Shkano.** s.l. : IEEE Transactons Acoustics Speech Signal Processing, 1989, Vol. 37, págs. 328-339.
11. *Recurrent Neural Prediction Model for Digits Recognition.* **Tellez, Paola.** 5, 2011, International Journal of Scientific & Engineering Research, Vol. 2. ISSN 2229-5518.
12. **Kuri, Galaviz.** *Genetic Algorithms.* Mexico : Fondo de Cultura Económica, 2002.
13. *Grandes personajes universales y de México.* Barcelona : Oceano, 1997.
14. *Number of hidden neurons in a recurrent neural networks estimation.* **Ordorica, Tellez.** 2012, International Journal of Scientific & Engineering Research.