Multilayer neural network recognition for sign language pattern motion detection with hand gesture control

M. Mauledoux[1], W. Amador[2], A. Vega[3]

***Abstract*—**This document shows the results of an interdisciplinary research design in the areas of electronical engineering and speech therapy/communication sciences, using a proposed classification method of electromyographic signals. Recognition technologies use signals from the forearm musculature and 27 signal points that coorspended to the alphabet, pronouns and some verbs high frequency of the Colombian Sign Language (CSL). A neural network of Pattern Recognition and comparative analysis of the number of hidden layers, plus performance was implemented. During the classification and recognition of these results high accuracy was obtained with only a 2.03% error in the classification of signals corresponding to the aforementioned areas. The data collection process included the participation of a natural linguistic model or deaf person/LSC user

Keywords—Emg, Neural network, Sign language Colombian.

I. Introduction

Historically, the deaf population has endeavored to participate in community equally. However, they often have few opportunities for participation in contexts such as employment and education [1] without communication hinderances. One of the factors limiting the participation of the deaf population is the low presence of interpreters to facilitate communication between the many different communication partners. When considering this barrier, the development of technologies that facilitate communication revealed a solution for the deaf when communicating with hearing people. This gives purpose to this research. A system of analysis and processing was developed using electromyography signals and movements from hand signals used by deaf persons in CSL. In this regard, it is important to mention that there are some technological developments that have sought to respond to the needs of the community, in terms of communication, particularly seeking to promote the communication of the deaf population (CSL user) within the hearing population.

Mauricio Felipe Mauledoux Monroy is with the Department of Mechatronics Engineering at the Nueva Granada Military University, Colombia as well as with the Manuela Beltran University, Bogota, Colombia.
William Amador and Ana Milena Rincon are with the Manuela Beltran University, Bogota, Colombia.

Example, [2] in 2013, the University Francisco José de Caldas developed, from an undergraduate study, an effort aimed to facilitate children learning Colombian Sign Language and a technological prototype. This being the first version, it met with some limitations in technology. A later, and second version, with fewer errors was proposed. However, it was geared towards perfecting the tool. In 2015, engineers develop a third version that would allow the deaf child population to interact with CSL [3] using a touch screen (about the size of a smart phone). Unfortunately, this was limiting the user to a one lexical repertoire and reduced that impact to short interactions. With advanced technological developments, that input system was patented by the Colombian Ministry of Trade and Industry.

There have been several engineers who have tried to translate the communication of the deaf. Nevertheless, among the most recent technological advances using an electronic system for language interpretation of Colombian sign language was made in 2015. This is the implementation of sensors in shoulder, wrist, elbow and attached to gloves. These in turn synchronized within a mobile device that transforms the signals into voice and text. It is worth noting that this technological input has been patented by the Colombian Ministry of trade and industry by order #52860 issued on 28 August 2015 [4].

This is one of the more relevant research studies because of their technique in the use of handles with EMG sensors and motion detecting patterns from the DTW (Dynamic Time Warping) algorithm. These movements, once performed are routed through a server and translated. [5]

However, within a therapeutic context, there is no evidence of proposals or technological developments to mediate the communication of the deaf population or users of CSL with hearing people. Not because of the ignorance of the linguistic features but for lack of interactions with engineering professionals. This posed new challenges for the scientific community, from an interdisciplinary perspective, and knowledge dialogue between professionals.

Moreover, these professionals are analyzing information for processing in the area of gesture recognition. Techniques exist in image processing electromyography, which uses signals and surface electromyography. This is highlighted in applying the research to the corresponding needs of the deaf community or in the therapeutic/treatment context. The surface electromyography signal has become one of the most commonly used, promptly responding to problems in different areas like prosthetic devices, rehabilitation programs, diagnostic evaluation and implementation in gadgets for everyday use.

Due to the growth in the use of these signals, classification studies have used different methods. Among which, we highlight studies in the area of rehabilitation and prosthesis for computational efficiency in the qualification of the movement and the application of prosthetic hands [6] [7] [8] [9]. One of the investigations in this case uses the evaluation of 4 movements associated with wrist muscles. From this, they were able to classify them using the adaptive neuro-fuzzy inference system (ANFIS) and the artificial neural network (ANN). Thanks to this classification and accuracy, results were achieved in 88.90% in ANFIS. Proving this to be a better technique than ANN in a matter of precision and speed during training and classification [6]. One of the forums found in the area of rehabilitation is seeking techniques for quick and precise strength in the upper limbs. They seek to implant a prosthesis by using techniques like Extreme learning machine (ELM), Support vector machine prediction (SVM), and Multiple nonlinear regression (MNR) while observing different characteristics in each technique. Throughout the trials, they demonstrate that the SVM is the best option concerning accuracy [10]. On the other hand, in the field of prostheses there are no comparative works. Techniques like Principal Component Analysis (PCA) are used with respect to Common spatial patterns (CSP) for the evaluation of 5 signals. In a population of 20 subjects validated and compared to an upper limb amputee, they obtained a results accuracy using PCA of 88.81 ± 6.58 and CSP of 89.35 ± 6.16%. This showed effectiveness in the results for the use of prosthesis in amputees [11]. Because of this, different ways to address efficiency when solving this problem are checked.

Within the optimal methods of electromyography, signals for characterization and classification using neural networks to determine hand force and movements are located[12] [13] [14]. Also, research and methods proposed use pattern recognition to minimize misclassification. For implementation in upper limb prosthesis, 8 kinds of movements compared with 6 signal processing techniques, based on Linear discriminant analysis [15], are used. Another study uses neural networks for the classification of hand movement with EMG signals from seven processing points of pre-processed EMG. This is used for detecting four movements of predefined hand gestures made from 10 neurons. Within a hidden layer, the study shows how neural networks, like Levenberg-Marquardt, can achieve an average rating of 88.4%. [16] Another technique to note is the presence of electrical activity in the musculoskeletal system, using wavelet maximum for EMG data analysis in conjunction with ANN for classification [17].

II. Problem Statement

However, based on such history, today it is important to develop technological tools not only with the aim of publicizing the Colombian Sign Language, or other that of other countries, and the characteristics of the deaf community, but also need to innovate based on artificial conductive pattern recognition techniques or intelligence technologies. The response times and signal processing thereof are characterized by better speed and also provide higher capacity in terms of recognition of sign, with the possibility of migrating data from different people, and of course with adaptive learning capability.

Therefore, this research aims to identify patterns and associated movements of hand signals used by deaf users of CSL. Using neural networks to thereby generate a system for interpretation of these signals.

To achieve this goal, it is necessary to make a list signs and gestures that correspond to alphabetic/linguistic units, personal pronouns and some high frequency verbs. For the measurement process and acquisition of data, signals are counted with a model of linguistic professionals in audiology, deaf persons using Colombian sign language, or an interpreter for te deaf and by placing a myo-gesture armband control around the proximal one-third of the forearm, making contact with the muscles, for movements in front of the face. Later, the same movements in the deaf person are captured and the process of signals is performed.

Subsequently, the selection of an analysis technique and neural network processing pattern recognition are used. The signal obtained matrices allow the construction of input and output for network training . Finally , we proceed to validate the results obtained by analyzing the number of hidden layers and neurons which yielded an error result on testing.

III. Design and Classification using Artificial Neural Networks

In this section, we propose the design of a neural network type pattern recognition for the classification 27 signs. The signs correspond to the alphabet, pronouns, and ten high frequency verbs, for a total of 46 signs. For each gesture, 15 samples were taken from 18 sensors measurements. The samples came from eight EMG sensors, 3 accelerometers sensors, 3 gyroscopes sensors and 4 that correspond to quaternion data. The figure 1 shows one of the captured gestures. The data sample for each movement was 70 recorded movements and these in turn are stored in a database for further processing.

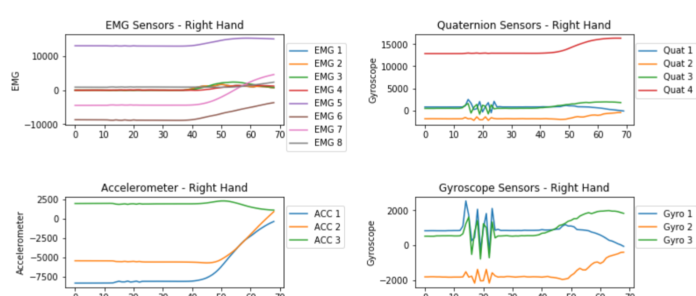

Fig. 1 Signals of Myo Gesture Control Armband

From the 15 samples, gave way to a selection process that allowed to take the 8 most significant gestures, thus discarding the remaining 7. As a result of this exercise, we were able to obtain an array of 368X1260 in system input data, which corresponds to the number of samples taken for gestures. We then analyzed all the data from the sensors and an array of 368X46. corresponding these to the outputs of the system associated with a specific gesture.

Based on these matrixes, the neural network was implemented. For the first stage of testing, a script that allowed stored information in a database, holding the records of hand signals taken from each of the persons, was sampled. A second stage, with the data collected, training the neural network was performed by a Matlab tool.

In order to calculate the mean square error of the ANN training process, to analyze the performance of the feedforward ANN [18] based in weights and biases is defined by:

$$F = mse = \frac{1}{N}\sum_{i=1}^{N}(e_i)^2 = \frac{1}{N}\sum_{i=1}^{N}(t_i - a_i)^2 \tag{1}$$

Where each individual mean square error is defined by:

$$F = mse = \frac{1}{N}\sum_{i=1}^{N} W_i^e (e_i)^2 = \frac{1}{N}\sum_{i=1}^{N} W_i^e (t_i - a_i)^2 \tag{2}$$

Which, through the matrix of collected data, proceeded to conduct training in a range of 20 to 80 neurons, with 50 repetitions per neuron, getting the error rate in testing average is 0.0244. This was achieved with 78 neurons. From these, a data classification accuracy of 97.56% was generated. This is seen more clearly in figure 2, which shows the mean and error trend neurons (i.e the maximum and minimum for each of the neurons).

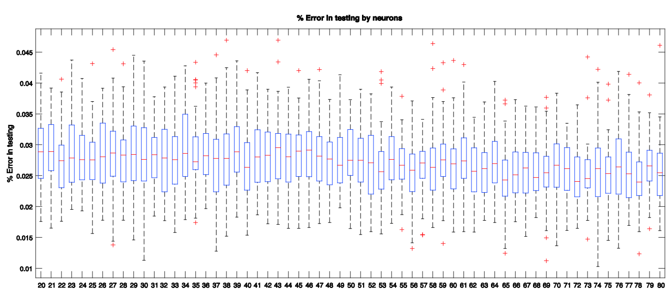

Fig. 2 Training with a layer

From the results, it was shown that the margin of error is very high for the performance requirements of the NNA. So, the need to reduce this error by reconfiguring the neural network was generated. This included a new layer and repetition of the training procedure. Retraining NNA was conducted with the same procedures as the previous network. Obtaining results, the lowest percentage of error in testing of 0.0203 (average), shown in figure 3. This was achieved with 43 neurons generating a classification accuracy of 97.97%.

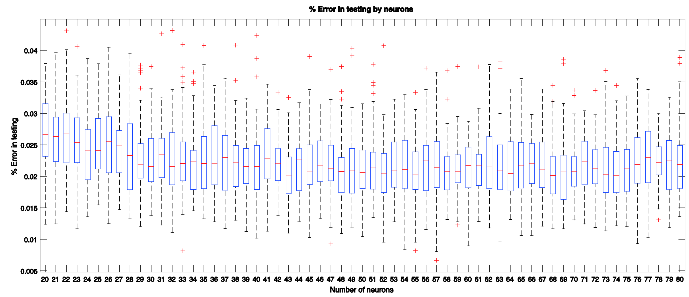

Fig. 3 Training with two layer

According to the results, a reduction of 28.73% in the percent accurate was identified with a layer of training in respect to the two layers, evaluating all training from 20 to 80 neurons. This percentage is not significant when evaluating the average training. However, minimal errors that are obtained by each neuron found generate a 43% reduction. In figure 4, the difference between the layers to their minimum values of each neuron is identified.

Fig. 4 Comparison of training by layers

IV. Analysis and Results

From the analysis of results obtained, the neural network had isolated cases with an error margin of 0.00663 with 57 neurons with two hidden layers. This was compared to the results that were acquired in a layer where a percentage of 0.0102 was obtained with 74 neurons. It is important to mention, in this case no online training is done, since these isolated cases can be obtained as a training base. However, it is evident that the least common denominator in a neural network training for the two-layer neurons is 43, to still retain the best option for retraining online. Research online for retraining may not be necessary. The best isolated case was obtained. This structure is shown in Figure 5.

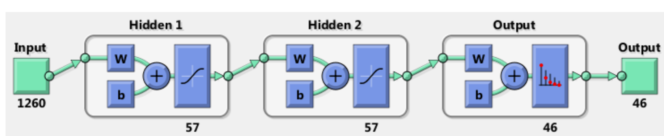

Fig. 5 Structure of the neural network

For evaluating the performance of the implemented neural network, it was based on the mean squared error (MSE). While the weights were adjusted using the back propagation learning rate to improve network performance. As a result, information obtained from performance analysis concerning different layers and number of neurons, validation network 0.01444 in time 38 epochs, seen in Figure 6 were achieved. This error is acceptable considering the error of validation and testing are even lower. This generated an appropriate response from the network to the training set.

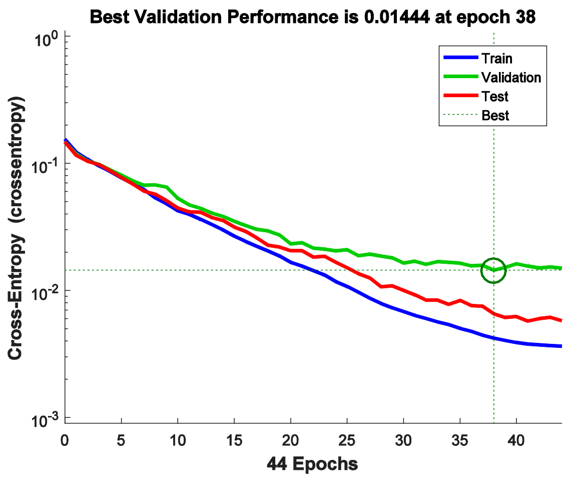

Fig. 6 Validation performance training

Using gradient data obtained, graphic Figure 7, it was observed, that in 44 times, it is possible to obtain a gradient value of 0.003997. This is consistent with the tolerance used.

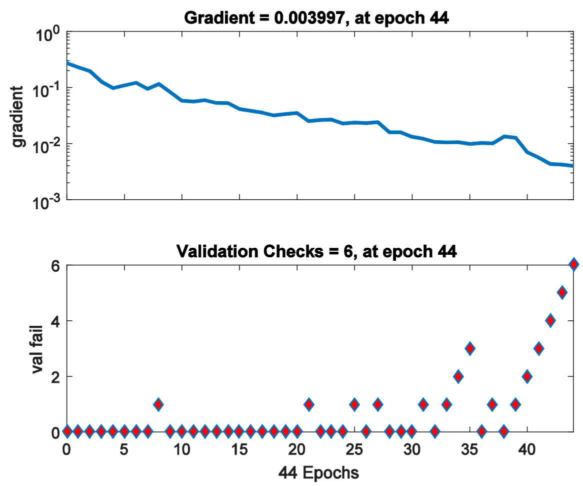

Fig. 7 Neural network training epochs

Culminating was evident, that being a multiplex type problem, classified gestures as being more dificult corresponded to "o and c" this is most thoroughly observed in the following Figure 8. This is due to the similarity of gestures and variations presented.

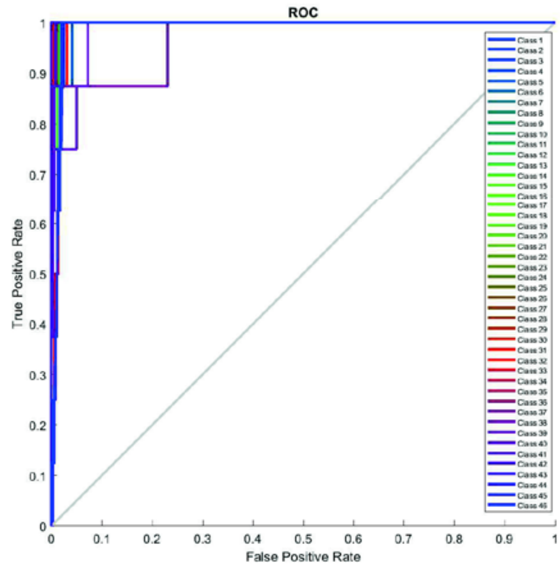

Fig. 8 ROC curves for artificial neural networks

V. Conclusion

Thanks to the results, it is concluded that the Pattern Recognition type of neural network generated good performance for pattern recognition electromyography associated with signs produced by a deaf person using CSL. Generating an average error with two hidden layers of 0.0203 evaluated on testing.

MSE results also identified in isolated cases up to 0.00663, generating error and checking accuracies of 99.34%, thus the effectiveness of the technique as a tool for pattern recognition in deaf individuals using CSL.

The analyzed results identified that this can improve a learning model for new classes or patterns of gestures. Taking average values that generate online training for constant learning of the neural network. This can improve learning within the neural network. In addition, the use of new learning algorithms suggests that, given the implications, Colombian sign language can become a problem of multiplex data. In other words, this language has a robust lexical repertoire. This will a develop software that each user learns online. It is suggested that these preloaded supported targets and a user interface are implemented.

Finally, the participation of relay center technologies for inclusion, funded by the Ministry of ICTs, is appreciated and from linking linguistic models in research, during the initial stages.

References

[1] C. X. Espinosa, V. G. Gómez, and C. M. Cañedo, “El Acceso y la Retención en la Educación Superior de Estudiantes con Discapacidad en Ecuador,” Form. Univ., vol. 5, no. 6, pp. 27–38, 2012.

[2] C. Hernández, J. L. Pulido, and J. E. Arias, “Las tecnologías de la información en el aprendizaje de la lengua de señas,” Rev. Salud Pública, vol. 17, no. 1, pp. 61–73, 2015.

[3] C. Hernández, H. Márquez, and F. Martínez, “Propuesta Tecnológica para el Mejoramiento de la Educación y la Inclusión Social en los Niños Sordos,” Form. Univ., vol. 8, no. 6, pp. 107–120, 2015.
[4] ANDRÉS MAURICIO CÁRDENAS TORRES, LEÓN MAURICIO RIVERA MUÑOS,BEATRIZ LILIANA GÓMEZ GÓMEZ ,LUZ MAGNOLIA TILANO VEGA, “Sistema electrónico para la interpretación de lengua de señas Colombiana,” 14-173060-00000–0000, de agosto-2015.
[5] P. Paudyal, A. Banerjee, and S. K. Gupta, “SCEPTRE: a Pervasive, Non-Invasive, and Programmable Gesture Recognition Technology,” presented at the Proceedings of the 21st International Conference on Intelligent User Interfaces, 2016, pp. 282–293.
[6] H. Jahani Fariman, S. A. Ahmad, M. Hamiruce Marhaban, M. Ali Jan Ghasab, and P. H. Chappell, “Simple and Computationally Efficient Movement Classification Approach for EMG-controlled Prosthetic Hand: ANFIS vs. Artificial Neural Network,” Intell. Autom. Soft Comput., vol. 21, no. 4, pp. 559–573, 2015.
[7] A. Suberbiola, E. Zulueta, J. M. Lopez-Guede, I. Etxeberria-Agiriano, and M. Graña, “Arm Orthosis/Prosthesis Movement Control Based on Surface EMG Signal Extraction,” Int. J. Neural Syst., vol. 25, no. 3, p. 1550009, 2015.
[8] D. Blana, T. Kyriacou, J. M. Lambrecht, and E. K. Chadwick, “Feasibility of using combined EMG and kinematic signals for prosthesis control: a simulation study using a virtual reality environment,” J. Electromyogr. Kinesiol., 2015.
[9] G. R. Naik and H. T. Nguyen, “Nonnegative matrix factorization for the identification of EMG finger movements: evaluation using matrix analysis,” IEEE J. Biomed. Health Inform., vol. 19, no. 2, pp. 478–485, 2015.
[10] V. Khoshdel and A. Akbarzadeh, “Application of statistical techniques and artificial neural network to estimate force from sEMG signals,” J. AI Data Min., 2016.
[11] F. Riillo et al., “Optimization of EMG-based hand gesture recognition: Supervised vs. unsupervised data preprocessing on healthy subjects and transradial amputees,” Biomed. Signal Process. Control, vol. 14, pp. 117–125, 2014.
[12] Z. Yang, Y. Chen, Z. Tang, and J. Wang, “Surface EMG based handgrip force predictions using gene expression programming,” Neurocomputing, 2016.
[13] S. P. ARJUNAN, D. K. KUMAR, and B. K. PANIGRAHI, “RECOGNITION OF FINGER/HAND GRIP MECHANISM BY COMPUTING S-TRANSFORM FEATURES OF SURFACE ELECTROMYOGRAM SIGNAL FROM HEALTHY AND AMPUTEE,” J. Mech. Med. Biol., p. 1650076, 2016.
[14] J. G. Ngeo, T. Tamei, and T. Shibata, “Continuous and simultaneous estimation of finger kinematics using inputs from an EMG-to-muscle activation model,” J. Neuroengineering Rehabil., vol. 11, no. 1, p. 1, 2014.
[15] S. Amsuss, P. M. Goebel, N. Jiang, B. Graimann, L. Paredes, and D. Farina, “Self-correcting pattern recognition system of surface EMG signals for upper limb prosthesis control,” Biomed. Eng. IEEE Trans. On, vol. 61, no. 4, pp. 1167–1176, 2014.
[16] M. Ibn Ibrahimy, R. Ahsan, and O. O. Khalifa, “Design and optimization of Levenberg-Marquardt based neural network classifier for EMG signals to identify hand motions,” Meas. Sci. Rev., vol. 13, no. 3, pp. 142–151, 2013.
[17] S. Mane, R. Kambli, F. Kazi, and N. Singh, “Hand Motion Recognition from Single Channel Surface EMG Using Wavelet & Artificial Neural Network,” Procedia Comput. Sci., vol. 49, pp. 58–65, 2015.

Mauricio Felipe Mauledoux Monroy was born in Bogotá, Colombia, in 1982. He received the B.S degree in Mechatronics engineering from the Nueva Granada Military University, Colombia, in 2005. In 2008 as a student of the Master in Information Technologies and Intelligent Systems in the St. Petersburg State Polytechnic University, Russia, at the automatic and intelligent distributed control department, he was promoted to a Ph.D. student. In 2011 He received the Ph.D. degree in Mathematical models, numerical methods and software systems (Red Diploma) from the St. Petersburg State Polytechnic University, Russia. In 2012, he joined the Department of Mechatronics Engineering at the Nueva Granada Military University, Colombia, as an Assistant professor and the electronic engineering program in Manuela Beltran University.

William Amador, was educated at Manuela Beltran University in Bogotá, Colombia receiving the BSc. Degree in Electronics Engineering in 2014. Since 2015 he has worked in the Engineering Faculty at Universidad Manuela Beltrán (Bogotá. Colombia) in Electronics Research Group (GIIB-UMB). Areas of interest: Robotics, Automation, Machine learning and imagen processing.

Ana Milena Rincon, Phonoaudiology, Magister Educational and social development, research coordinator Phonoaudiology program, Manuela Beltran University, research experience related technological development from interdisciplinary work, related to speech, language, and personal inclusion of persons with disabilities.