**Phases recongnition with Machine Learning using MYO device based on FPGA**

Escuela superior Politécnica del Litoral ESPOL

Albuja Abraham, Joseph González

Diseño de Sistemas Digitales

Guayaquil, Ecuador

**Abstract.**

The project described is the implementation of an embedded system based on FPGA which performs the reading and classification of temporary acceleration signals measured by the MYO electronic device; These signals have been sorted in files that will be processed for subsequent classification of words according to the red neural that has been chosen based on the RMS values ​​of the file sets that served as the basis for the execution of the same. These contain information on the position of the device in the time domain and frequency of the motor activity of the subject carrying the device which performs the movement that is within the sign language represents words that were raised from the beginning. Finally, the system shows the result of the classification through an HDMI screen with 96% accuracy when classifying the files

1. **Introducción**

Las personas con discapacidad auditiva utilizan la lengua de señas para comunicarse, lengua que solo ellos y un grupo reducido de personas dominan. Por esta razón, hasta el día de hoy estas personas no pueden comunicarse con otros individuos, teniendo en cuenta que son un porcentaje menor en la población, por lo tanto, se puede entender que estas personas deben sentirse excluidas, por gran parte de la población. La única forma de poder interactuar con este grupo no oyente, es aprendiendo la lengua de señas a través de cursos presenciales y/o cursos E-Learning.

La motivación que nos lleva a la realización de este proyecto es, principalmente, el hecho de poder ayudar a personas que tienen problemas para comunicarse debido a que la estadística dada por el Consejo Nacional de igualdad de Discapacidades determina que el 14.03% de las personas con discapacidades en el Ecuador son sordomudas. Tomando en cuenta que la lengua de señas se basa en la utilización de gestos como comunicación, y que las tecnologías de reconocimiento de gestos pertenecientes a la lengua de señas se encuentran aún en investigación y desarrollo, se encontró una nueva motivación para contribuir y constituir una base sólida de conocimientos para el posterior desarrollo de tecnologías que permitan utilizar el reconocimiento gestual en diferentes ámbitos de progreso.

Como método de solución para el presente proyecto se realizan las mediciones de las señales de aceleración de los movimientos, entregando valores de tres ejes x, y, z. La frecuencia de muestreo es de 1503,759hz debido a que se realiza la toma de 4000 datos en un lapso de 2.66s. Existen 200 archivos por cada seña es decir que en total son 1000 archivos los cuales servirán para la creación de la red neuronal que se utilizara para la clasificación de archivos que representan movimientos de 5 señas establecidas.

1. **Arquitectura del Hardware**

El proyecto tiene 2 procesadores, uno contiene la red neuronal y recibe los archivos que son datos de los movimientos que serán clasificados y posteriormente el otro procesador se encargara de la comunicación para la pantalla HDMI.

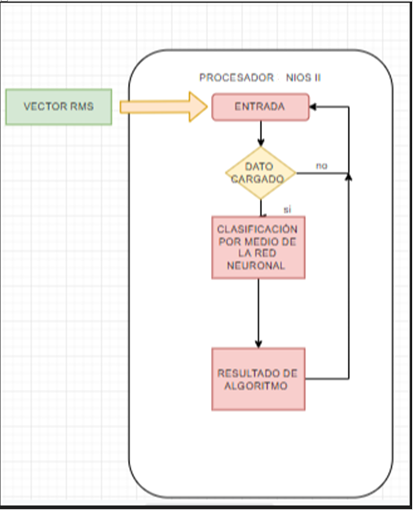


Figura 1 Pseudocodigo del procesador 1

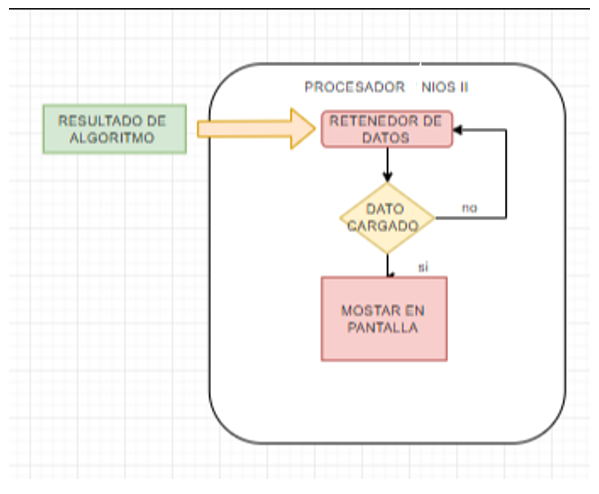
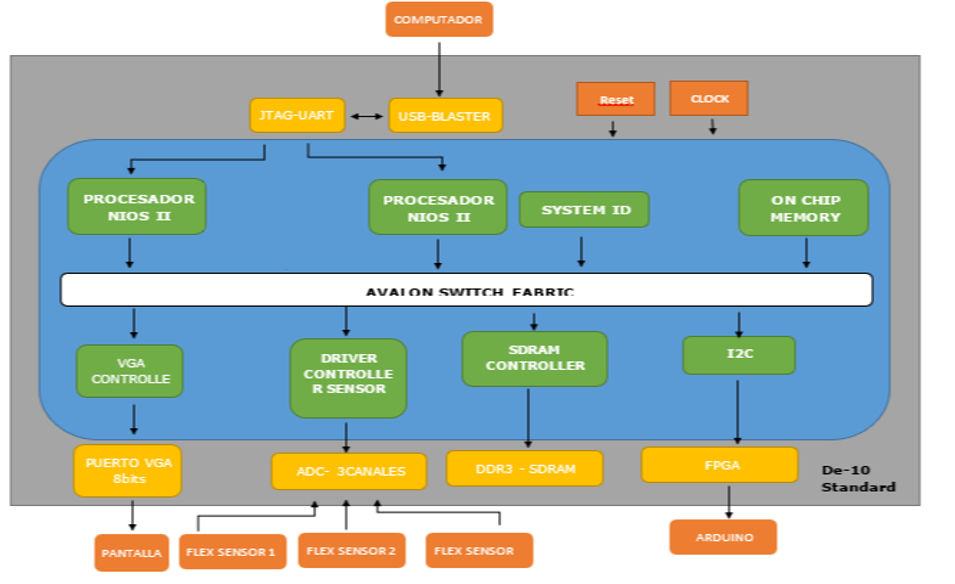


Figura 3 Arquitectura del sistema

Figura 2 Pseudocodigo del procesador 2

El primer procesador convierte una matriz de 4000x3 a un vector de 1x3 calculando el valor rms de cada columa. Este procesamiento de datos lo realiza de la siguiente manera, primero pregunta si existe o no un archivo en . Mat si es así entonces procede a realizar la lectura de todos los datos por fila, preguntando siempre si se llegó a la fila 4000, cuando llega a la fila 4000 inmediatamente el procesador procede a realizar el cálculo de los rms de las 4000 filas y nos entrega un vector de 1 fila y 3 columnas. Este procesador tambien posee el entrenamiento de la red neuronal,a este le llega el vector de 1 x3 y le aplica el algoritmo para la identificación de la palabra, al finalizar esto este procesador nos entrega una palabra a partir de los datos rms.

El segundo procesador está encargado de mostrar la palabra por pantalla, a su entrada recibe la palabra de manera codificada, por tal motivo este la procesa y establece la conexión vga con el monitor.

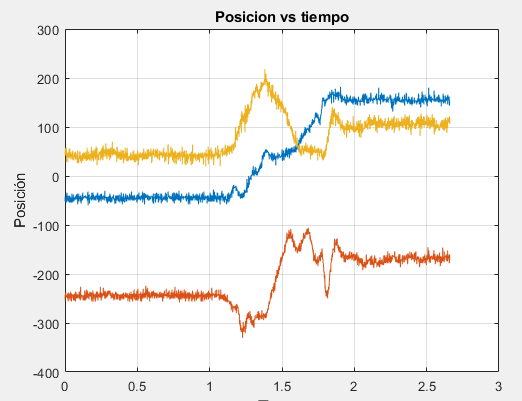


Figura 4 Gráfica de posición vs Tiempo

**Dataset**

Cuando el usuario realice la ejecución de las frases mediante el lenguaje de señas los dispositivos comerciales MYO, darán los datos de posición mientras realiza el movimiento y serán registrados en archivos tal cual se ha mostrado en la figura 4. Se ha creado en Matlab un script que permite cargar el archivo que contiene las 4000 filas x 3 columnas que serán enviadas por comunicación RS-232 al arduino y este convierte los datos y comunica por I2C a la FPGA la cual con la programación adecuada de los bloques que se indicara a continuación mostrará por pantalla las palabras que representen el movimiento en el cual ha sido clasificado el archivo según la red neuronal creada a partir de los datos RMS de cada archivo.

Se procesan los datos y se tiene los siguientes RMS entre los archivos que están almacenados para la red neuronal.

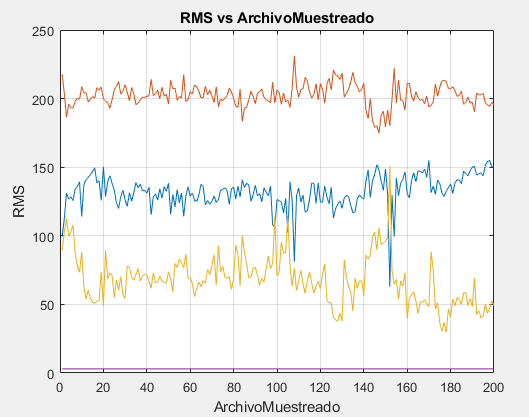


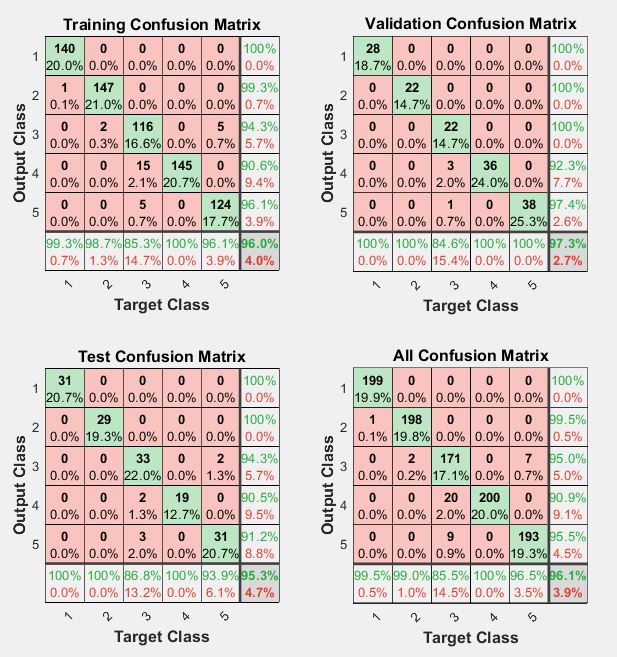
Figura 5 Valores RMS

**Red neuronal**

Se ha creado en Matlab un script que permite cargar el archivo que contiene las 4000 filas x 3 columnas que serán enviadas por comunicación RS-232 al arduino y este convierte los datos y comunica por I2C a la FPGA la cual con la programación adecuada de los bloques que se indicara a continuación mostrará por pantalla las palabras que representen el movimiento en el cual ha sido clasificado el archivo según la red neuronal creada

La creación de la red neuronal para la cual se toma el valor RMS de cada uno de los archivos tiene un accuracy del 96.1% de precision de la clasificacion de las palabras tal como lo muestra la gráfica 6 donde muestra la dificultad que tuvo para reconocer los archivos que representan movimientos que serán reconocidos luego en consola como palabras. Las palabras que son parte de la red neuronal son : hello, deaf,me, you, sign.

Siendo la palabra “me” un poco más difícil de reconocer para la red neuronal.

****

***Resultados***

Una vez realizado el envió de datos por medio de Matlab al Arduino, este convierte los datos y envía por medio de comunicación I2C a la FPGA este envío de datos se muestra por consola en la figura 7 y una vez realizada el envío de datos del archivo que contiene el movimiento realizado, la FPGA realiza la clasificación de la palabra según la red neuronal prestablecida, determinando y mostrando por consola la palabra realizada por medio de lenguaje de señas tal como lo muestra en la figura 8.

Posteriormente la FPGA permite que realizando las conexiones y programación necesaria se muestre esta clasificación por medio de salida HDMI de la tarjeta DE10-NANO, la cual usando un convertidor a VGA podrá ser mostrada por pantalla.



Figura 7 Envío de datos por medio de comunicación I2C

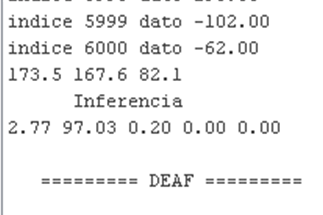


Figura 8. Resultados mostrados por consola

Figura 6 Matriz de confusión

**Referencias**

La investigación para el reconocimiento de gestos hechos con las manos empezó a realizarse hace más de 40 años [1]. En 1977 se realizó un sistema de identificación de movimiento de doblamiento de dedos mediante un guante en mano propuesto por Zimmerman. [2]

Hoy en día existen varias aplicaciones entre esta, está la detección de frases claves usando acelerador compacto CNN diseñado por Lattice semiconductor, el cual consta de un sistema embebido con red neuronal basado en FPGA. Se toma los datos mediante el acelerador NN IP, el cual se conecta a una FPGA tolos la cual hace la transferencia de datos a la FPGA por medio de un FPGA bitsream donde se los procesa y se clasifica para la toma de la respectiva acción [3]

En otra aplicación se tiene un analizador sintáctico basado en FPGA para gramática libre de contexto el cual utiliza el FPGA CHIP para la toma de decisiones y para las memorias emplea un Chip SDRAM donde se encuentra la memoria de análisis y de memoria de gramática GMI [4]

# **Bibliografía**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| [1] |  |  | P. Premaratne, "human computer interaction using hand gestures," Springer, 2014. | | | |
| [2] |  |  | P. Sharma and S. Sharma, "Evolution of Hand Gesture Recognition : A Review," Pilani. | | | |
| [3] |  |  | L. Semiconductor, "Key Phrase Detection Using Compact CNN Accelerator IP," Reference Design, 2019. | |
| [4] |  |  | C. Ciressan, E. Sanchez, M. Rajman and J. C. Cha, "An FPGA-based syntactic parser for real-life unrestricted context-free grammars". |
|  |  |  |  | | |