

Técnicas de Clustering basada en Redes Neuronales para reconocer gestos estáticos del alfabeto dactilológico de la Lengua de Señas Peruana

Carlos Anibal Espinoza Hoyos, Ing¹, Pedro Huamaní-Navarrete, Dr²,
^{1,2}Ricardo Palma University, Perú, c.espinoza452@gmail.com, phuamani@urp.edu.pe

Abstract – *The Peruvian Sign Language (LSP) is the language used by people with hearing disabilities throughout the national territory [1]. At this point, in the social reality of our country there is still an isolation from the deaf community due to the lack of knowledge about this section of population and its needs, this situation has as result a social problem, turning them into an isolated and vulnerable community. Faced with this lack, the possibility of applying current technologies for the deaf community is envisaged using Artificial Neural Networks (RNA) in the field of pattern recognition and classification (clustering) as many authors have proven. The rationale behind this article was to recognize 24 static gestures of the LSP dactylological alphabet, acquired of an electronic glove and using two clustering techniques based on the neural networks of Generalize Regression and Learning by Vector Quantization with the support of the Toolbox Deep Learning del Matlab. Moreover, three different neural network topologies were applied for each technique: for the gestures, for spatial orientation sensors (MPU) a conjugate so-called for grouping results from the anterior topologies. The recognition of static gestures was assessed using the Contingency Table and Precision, Recall and Magnitude-F metrics. The results revealed the Neural Network of Generalize Regression gave better results in the recognition of gestures, obtaining precision and recall measurements equal to 100% and a Magnitude F in the same way, along with a noise resistance of 1%; by contrast, the Neural Network of Learning by Vector Quantization, reached just a Precision measurement equal to 81.25% and Recall to 87.5%, obtaining a Magnitude F equal to 84.26% and nearly not a thing of noise tolerance.*

Keywords—Clustering algorithm, Deep Learning Toolbox, Peruvian Sign Language, Precision, Recall.

I. INTRODUCCIÓN

Actualmente, existe un gran abismo lingüístico entre la comunidad practicante de la Lengua de Señas Peruana (LSP) y la no practicante, siendo las principales causas del problema social la continua carencia de políticas inclusivas para la comunidad sorda de nuestro país, la no profesionalización de la lengua oficial reconocida para la comunidad, y la insuficiente cantidad de profesionales capacitados para la interpretación de la LSP a lenguaje hablado. De lo anteriormente expuesto, se puede resumir que el común denominador entorno a la problemática, es la brecha lingüística existente entre la comunidad sorda practicante de la LSP y la comunidad oyente no practicante de esta lengua; este aislamiento lingüístico se debe a 3 factores importantes: en primer lugar la carencia de políticas inclusivas, si bien en el 2010 se promulgó la Ley N°29535 [1], se tuvieron que esperar unos cuantos años para su reglamentación y aplicación; en segundo lugar, la no profesionalización de la LSP en nuestro país, esto significa que no se puede tener un título profesional para ejercer de intérprete

para esta lengua en el territorio nacional; y, finalmente, la insuficiente cantidad de intérpretes (23) certificados en el extranjero y para una población creciente de aproximadamente 560 mil personas. Si bien la ley promulgada en el 2010 instaba a las instituciones estatales a contar con intérpretes para la comunidad sorda, esta necesidad no fue abastecida; adicionalmente, la carencia de promoción de la LSP implica ignorancia para la mayoría de personas acerca de esta comunidad y de sus necesidades, hasta en el ámbito político; arriesgando de esta manera a la población sorda, ya que no poseen un fácil acceso a establecimientos de servicios primordiales como lo son hospitales, comisarías, centros educativos, entre otros, debido a la imperativa necesidad de un intermediario para la comunicación de sus necesidades. Junto a ello, muchas personas son separadas de la comunidad al no poder comunicarse, propiciando escenarios de abusos que quedan impunes debido a que no pueden ser comunicados por los violentados.

Por lo tanto, la importancia de este proyecto radica en evaluar la suficiencia con la que las RNAs son capaces de reconocer gestos específicos, de manera que puedan suplir la necesidad de un intermediario para la comunicación de las necesidades de esta comunidad; evitando de esta forma, las dificultades existentes como lo son el no tener fácil acceso a servicios primordiales. Así como también, debido a la limitada bibliografía concerniente a la aplicación de la tecnología de los guantes electrónicos, para la solución de este problema en el país; es por ello que, la presente investigación toma una importancia académica. Igualmente, el interés central de esta investigación radica en el reconocimiento de los gestos estáticos, a través de dos topologías de redes neuronales artificiales basadas en la clasificación por clustering.

De esta manera, en el marco de la aplicación de las redes neuronales, esta investigación abarca la correlación entre las técnicas de clustering implementadas, y el distintivo reconocimiento y/o clasificación de los datos a los cuales la red neuronal fue sometida. Para lo cual, el instrumento de recolección de datos fue un “guante electrónico” implementado en un proyecto previo en la Universidad Ricardo Palma; el procedimiento radicó en la toma de 200 muestras a 3 personas distintas realizando las 24 señas estáticas pertenecientes al alfabeto dactilológico peruano. Asimismo, se eligió el Principio de Pareto para establecer la cantidad de datos necesarios para el entrenamiento, y para las pruebas de las redes neuronales artificiales.

Para un estudio sistematizado del problema, el presente artículo se ha estructurado de la siguiente manera. Una sección de Introducción que realiza un abordaje al tema de investigación, planteando el problema sobre el cual gira, la importancia de esta y los objetivos a los que se apunta para

Digital Object Identifier (DOI):
<http://dx.doi.org/10.18687/LACCEI2021.1.1.547>
ISBN: 978-958-52071-8-9 ISSN: 2414-6390

resolver la incógnita. En la sección Marco Teórico, se exponen los acercamientos entre la ciencia y la LSP a nivel mundial y local, a través de investigaciones previas con el fin de analizar procedimientos y resultados pasados, y algunas definiciones teóricas importantes. Por otro lado, en la sección Diseño de Ingeniería se planteó la manera como se llevó a cabo la experimentación, y se detallan las etapas necesarias para el cumplimiento de los objetivos. Luego, en la sección de Resultados se detallan los datos cuantitativos de la experimentación, junto con la comparación y análisis de los parámetros establecidos para la evaluación de las técnicas de clustering propuestas. Para finalizar la investigación se redactaron las conclusiones, recomendaciones y se describieron las referencias bibliográficas utilizadas.

II. MARCO TEÓRICO

En [2] se resume parte de la historia de las Redes Neuronales Artificiales (RNA), donde se puede encontrar que en 1969 Minsky y Papert describen las limitaciones computacionales de las redes neuronales de una sola capa (single layer); luego, en 1986, James Lloyd McClelland y David Everett Rumelhart publican acerca del algoritmo de aprendizaje de retro propagación que permite el entrenamiento de perceptrones multicapa, introduciendo por tanto el concepto de las capas ocultas de las RNA y con ello los perceptrones multicapa. De esta manera, en [3], se introduce el concepto de Red Neuronal poco profunda (Shallow Neural Network), conformada por unas pocas capas ocultas; por otro lado, una red neuronal profunda (Deep Neural Network) es aquella que tiene un elevado número de capas ocultas, y por ello se ha convertido en la herramienta estándar para resolver una variedad de problemas de visión computacional.

Existen algunos estudios previos que han sido planteados como soluciones ante la problemática existente, tal es el caso de [4] que utilizó una red neuronal Perceptron Muticapa de 5 entradas y 26 salidas, una para cada letra del abecedario, y compara una cierta cantidad de redes neuronales. Según los resultados obtenidos, las redes neuronales entrenadas demostraron un buen funcionamiento y reconocieron las señas de manera óptima; sin embargo, es el algoritmo de propagación rápida el cual demostró un mejor desempeño entre todas las demás; y, por ello, consideran implementarlo para un futuro prototipo final. Luego, en la Universidad de Cornell [5], se presentó un proyecto similar en el que se implementó un guante con sensores Flex y acelerómetro, con los que se captaron señales del Lenguaje de Señas Americano (ASL), y con apoyo del algoritmo Support Vector Machine (SVM), identificaron 28 grupos de gestos y lograron un proceso de clustering exitoso al 98%, debido a que necesitaban mejoras para evitar la ambigüedad entre las señas de las letras “u” y “v”. Asimismo, en la Universidad Politécnica de Hong Kong [6], se implementó una opción más, basada en un sistema de reconocimiento de lenguaje de señas usando sensores 3D, donde utilizaron Multiple Sensors Data Fusion (MSDF) y el procesamiento de los datos se llevó a cabo por medio de Modelos Ocultos de Markov (HMM). Los resultados presentados demuestran una alta tasa de reconocimiento para lo propuesto en comparación con los sistemas de un solo sensor. Por otro lado, una contribución nacional proviene de la tesis [7], donde se implementaron guantes con acelerómetro y sensores Foto-Flex para recabar

información de los dedos flexionados de cada seña realizada; además, la RNA utilizada en esta investigación fue Back Propagation y se evaluó entorno a su precisión, obteniendo una clasificación satisfactoria del 97%. Adicionalmente, en [8] se encuentra la propuesta para secuencias de videos que contienen letras del alfabeto ASL, donde se utilizó una Red Neuronal Convolutiva (CNN) para construir modelos de clasificación para cada una de las 26 letras del vocabulario. Los resultados muestran que la tasa de efectividad de este modelo fue del 98.66%, pero se debe tener en cuenta que solo reconoce las letras del alfabeto y ayuda cuando las personas deletrean determinadas palabras. Y en el artículo [9] se realizó la extracción de datos de secuencias de videos que contienen signos propios del ASL, a través de un algoritmo capaz de detectar y seguir el trayecto de una mano, identificando por cada “frame” si existe un signo deseable e importante. Por ello, los resultados se enfocan más en la capacidad de seguir la mano correctamente, lo cual se ve afectado por la calidad del video y/o distorsiones del mismo. La red fue evaluada conforme su precisión, recall y magnitud-F1, con los que sustentan numéricamente sus conclusiones.

Complementariamente a lo anterior, se indican algunas definiciones importantes y utilizadas en este artículo.

A. Red Neuronal de Regresión Generalizada (GRNN)

Según Donald F. Specht en [10], da el alcance de una red que usa un método que la libera de la necesidad de asumir una forma funcional específica. De manera que, le concede la forma apropiada para ser expresada como una función de densidad probabilística, la cual es determinada empíricamente de los datos observados utilizando una estimación de ventana Parzen. Es así que, en [11], la arquitectura multicapa de la red GRNN se define por la conformación de una capa de entrada, una capa de Base Radial (donde se encuentra específicamente la GRNN), y una capa lineal especial.

B. Red Neuronal de Aprendizaje por Cuantización Vectorial (LVQNN)

Según Bonifacio del Brío, en [12] se define como un modelo supervisado que realiza tareas de clasificación de patrones; y, en el software Matlab esta red neuronal está planteada con una capa de entrada, una competitiva (donde se encuentra específicamente la red LVQ), y una lineal.

C. Señas o Gestos Estáticos

Son gesticulaciones propias del Lenguaje de Señas, las cuales no implican variación en el tiempo o movimiento. Entre las señas propias del alfabeto dactilológico peruano se pueden contabilizar 24 señas estáticas, y 3 señas dinámicas (que implican movimientos) correspondientes a las letras J, Ñ y Z, tal como también fue utilizado en [13]. Ver la figura 1.

D. Métrica Precision (PR)

Se define matemáticamente como el número de Verdaderos Positivos (VP), dividido por el número total de Verdaderos Positivos (VP) y Falsos Positivos (FP). A continuación, se muestra la expresión correspondiente.

$$PR = \frac{VP}{VP + FP} \quad (1)$$



Fig. 1 Alfabeto dactilológico peruano [14].

E. Métrica Recall (RE)

Se define matemáticamente como el número de Verdaderos Positivos (VP) dividido por el número total de Verdaderos Positivos (VP) y Falsos Negativos (FN). A continuación, se muestra la expresión correspondiente.

$$RE = \frac{VP}{VP + FN} \quad (2)$$

F. Métrica Magnitud-F

Se utiliza para combinar las métricas Precision y Recall en un sólo valor; de esta manera, se facilita la comparación del rendimiento combinado de estas entre varias soluciones. A continuación, se muestra la expresión correspondiente conociendo que β es la importancia entre las métricas Precision y Recall, y siendo en este caso igual a la unidad.

$$Magnitud - F = \frac{(\beta^2 + 1) * PR * RE}{\beta^2 * PR + RE} \quad (3)$$

III. DISEÑO DE INGENIERÍA

El guante electrónico utilizado y mencionado en este artículo, pertenece a un proyecto previo que fue desarrollado en

el año 2019 en la asignatura Taller de Electrónica IV del ciclo 2019-II en la Carrera de Ingeniería Electrónica de la Universidad Ricardo Palma, Lima - Perú. En dicho proyecto el autor, Carlos Espinoza Hoyos, formó parte del grupo de trabajo, mientras que Pedro Huamaní Navarrete se desempeñó como docente de tal asignatura. De esta manera, estos guantes electrónicos fueron implementados con sensores Flex, ubicados en los dedos, para emular las falanges de las manos y captar el movimiento de flexión de estos; asimismo, se utilizó un módulo MPU6050 para capturar la afección de la gravedad sobre este sensor, así como un módulo Arduino NANO encargado de la lectura de los sensores anteriormente mencionados; por otro lado, la programación y el procesamiento de la señal se llevó a cabo en un computador utilizando el software LabVIEW. La Figura 2 muestra uno de los guantes electrónicos implementados.



Fig. 2 Prototipo de guante electrónico implementado en la asignatura Taller de Electrónica IV – Ciclo 2019 II - URP.

Igualmente, con la finalidad de apreciar mejor aún el proceso de reconocimiento de los gestos estáticos de la actual investigación, en la Figura 3 se muestra una representación de un diagrama de bloques general, donde se aprecia la etapa de adquisición a través del guante electrónico y el Arduino Nano, la etapa de preprocesamiento, establecimiento de la base de datos y entrenamiento con las RNA. A continuación, se detalla.

A. Adquisición de señales

A través del puerto serial del módulo Arduino NANO se recogen y se envían los niveles de voltaje proporcionados por los sensores Flex (a_1, a_2, a_3, a_4 y a_5), así como los valores de los ejes X e Y obtenidos de los MPU's (a_6 y a_7). Dichos niveles son enviados hacia el software Matlab, donde se ordenan en matrices de 7×1 . Asimismo, se utilizaron las librerías I2Cdev.h, MPU6050.h y Wire.h, con 57600 baudios para la comunicación.

B. Preprocesamiento de datos digitales

Consiste en el cálculo de un promedio y eliminación de los errores en la toma de datos, como el de los sobre picos producto de movimientos. De esta manera, se establece una base de datos conformada por 200 mediciones correspondientes a cada letra y por cada uno de los tres participantes. Posteriormente, basado en el Principio de Pareto, se realiza la separación del total de mediciones en dos grupos. El 80% para el grupo de entrenamiento, y el 20% para el grupo de prueba de la RNA.

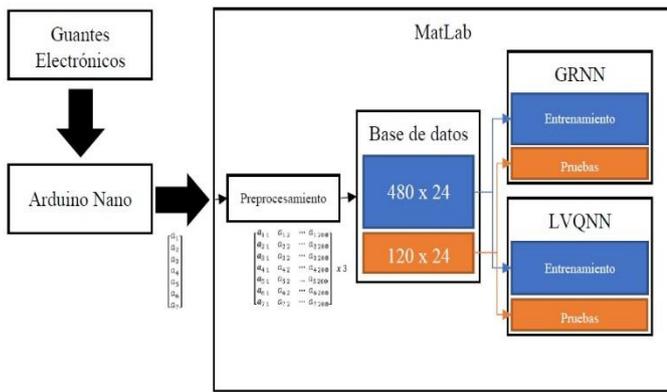


Fig. 3 Diagrama de bloques general del proyecto desarrollado.

C. Acondicionamiento de datos e implementación computacional

Para la correcta obtención de datos, los sensores Flex fueron colocados a lo largo de cada dedo de la mano cubriendo las falanges proximales, medias y parte de las distales. Esto fue realizado debido a que se necesitó tener un valor numérico preciso por cada señal; pues, en ciertos casos como el de la letra C, la flexión completa de los dedos no es tan requerida como en el caso de la letra S, donde los dedos meñique y pulgar si deben ser flexionados lo mejor posible. Además, debido a que los sensores Flex son comerciales, los valores de resistencia no son idénticos y mantienen una ligera diferencia que se hace notoria en los niveles de voltaje utilizados; por lo cual, se vio la necesidad de realizar un mínimo ajuste para aproximarlos a la unidad, cuando los dedos se encontraban totalmente estirados. Ver la Figura 4. Y, experimentalmente, se apreció que, al flexionar completamente los 5 dedos, los niveles de voltaje se encontraban por debajo de 0.8 voltios, tal es el caso de la representación de la letra S. Ver la Figura 5.

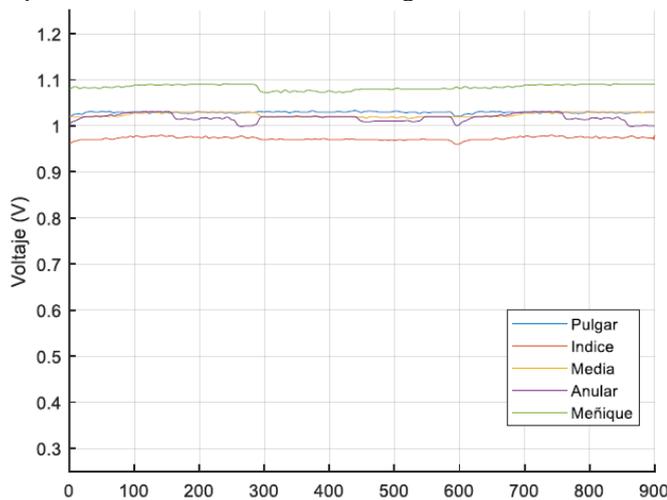


Fig. 4 Valores de voltaje cuando los dedos se encuentra totalmente estirados.

Por la cantidad de datos recopilados, a continuación, solamente se presentan los voltajes máximos y mínimos otorgados por los 5 sensores Flex para los gestos de las letras: A, B, C, D, E, F, G, H, I, K, L, X e Y. Ver la Tabla I.

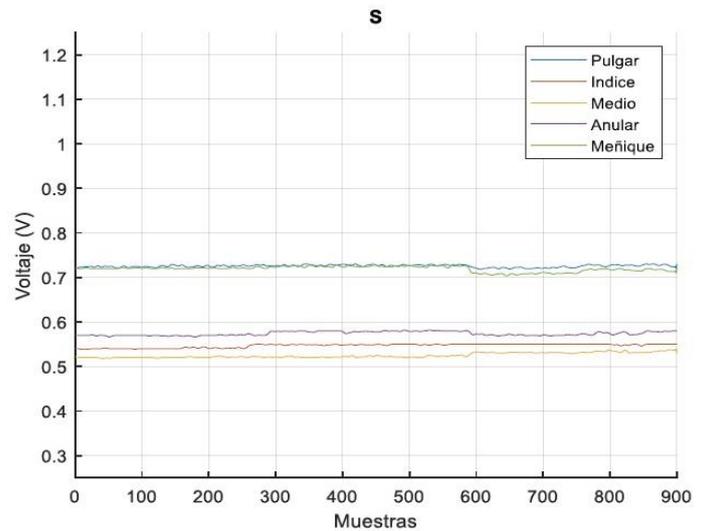


Fig. 5 Valores de voltaje cuando los dedos se encuentra totalmente contraídos.

TABLA I
RANGOS DE LECTURA DE VOLTAJES OBTENIDOS DE LOS SENSORES FLEX

GESTOS	PULGAR	ÍNDICE	MEDIO	ANULAR	MEÑIQUE
A	0.9671 ~ 1.0109	0.6900 ~ 0.7168	0.5185 ~ 0.5506	0.5385 ~ 0.5612	0.6817 ~ 0.7126
B	0.7997 ~ 0.8512	0.9900 ~ 1.0341	0.9864 ~ 1.0315	1.0902 ~ 1.1600	0.8900 ~ 0.9400
C	0.9834 ~ 1.0700	0.7581 ~ 0.8200	0.6200 ~ 0.6767	0.8063 ~ 0.8551	0.7323 ~ 0.8000
D	0.8506 ~ 0.9601	1.0622 ~ 1.1422	0.4676 ~ 0.5516	0.6504 ~ 0.7422	0.5018 ~ 0.6243
E	0.7500 ~ 0.7709	0.5700 ~ 0.6114	0.4300 ~ 0.4598	0.6300 ~ 0.6643	0.4965 ~ 0.5300
F	0.7744 ~ 0.8000	0.6016 ~ 0.6229	0.8400 ~ 0.8815	0.9600 ~ 1.0400	0.8372 ~ 0.8800
G	0.9300 ~ 0.9600	0.8885 ~ 0.9222	0.4592 ~ 0.5000	0.5625 ~ 0.5996	0.6447 ~ 0.6715
H	1.0486 ~ 1.0641	1.1210 ~ 1.1700	0.9327 ~ 0.9700	0.6963 ~ 0.7200	0.5271 ~ 0.5600
I	0.7406 ~ 0.7646	0.6400 ~ 0.6969	0.4986 ~ 0.5355	0.5925 ~ 0.6230	1.0385 ~ 1.0600
K	0.8482 ~ 0.8766	0.9209 ~ 0.9523	1.0100 ~ 1.0231	0.6013 ~ 0.6229	0.6802 ~ 0.7220
L	0.9783 ~ 1.0097	0.9283 ~ 0.9600	0.5002 ~ 0.5600	0.4872 ~ 0.5300	0.7085 ~ 0.7600
...					
X	0.7351 ~ 0.7700	0.7400 ~ 0.7952	0.5074 ~ 0.5179	0.5570 ~ 0.5862	0.6959 ~ 0.7100
Y	0.9785 ~ 1.0269	0.6985 ~ 0.7119	0.6151 ~ 0.6300	0.5768 ~ 0.6021	0.9943 ~ 1.0269

De esta manera, se logra apreciar que algunos gestos son críticos entorno a la similitud que poseen entre sí. Tal es el caso de la letra: D con X, C con O, E con M, N y S, K con H, L con G y Q, y R con V. De la misma forma, a continuación, se

presentan los valores obtenidos del sensor MPU para las distintas orientaciones descritas anteriormente, y que a su vez llegan a ser críticas para el proceso de decisión por parte de la red neuronal. Ver la Tabla II

TABLA II
RANGOS DE LECTURA DE VOLTAJES OBTENIDOS DEL SENSOR MPU

ORIENTACIÓN	EJE X	EJE Y
X	29.8495 ~ 31.2982	-47.7389 ~ -45.2999
Y ⁺	11.2813 ~ 17.0790	-76.9832 ~ -72.0087
Y ⁻	17.8459 ~ 21.4277	48.7415 ~ 50.7634

D. Implementación de la red neuronal GRNN

Para la implementación computacional de la GRNN se requirió del comando “newgrnn” del Toolbox Deep Learning del software Matlab. Es así que, dicho comando necesitó del vector de prueba, del vector de clases (definido por el usuario), y del parámetro “SPREAD” con su valor estándar debido a que todas las categorías tuvieron la misma importancia. Asimismo, se implementaron tres redes neuronales GRNN; la primera (GRNN-GESTO) para el reconocimiento de los 24 gestos, la segunda (GRNN-MPU) para el reconocimiento de 3 posiciones de la mano identificadas como orientaciones Y⁺, X e Y⁻ por el sensor MPU, y la tercera denominada conjugada (GRNN-CONJU) que combinó los resultados de las dos anteriores para ayudar a discernir la categoría correcta, en el caso de presentarse similitudes. Se comprobó también, de manera experimental, que el cálculo de los pesos sinápticos para dicha red neuronal se obtuvo desde ella misma y en función de los vectores de prueba y de categoría, previamente impuestos.

E. Implementación de la red neuronal LVQNN

Para la implementación computacional de la LVQNN se requirió del comando “lvqnet” del Toolbox Deep Learning del software Matlab. Y, para dicho comando, se necesitó de la consideración del valor de uno de sus parámetros de entrada LR, debido a que influye en el aprendizaje de la propia red neuronal. Para la presente investigación, con 24 categorías, el valor LR se mantuvo 10 veces por debajo del valor estándar. Por otro lado, también se requirió del comando “train” para el proceso de entrenamiento. Y como también, se implementaron tres LVQNN; la primera para el reconocimiento de los datos provenientes del sensor Flex (LVQNN-GESTO), la segunda para el reconocimiento de los datos provenientes del sensor MPU (LVQNN-MPU), y la tercera denominada conjugada (LVQNN-CONJU) que combinó los resultados de las dos anteriores para ayudar a discernir la categoría correcta, en el caso de presentarse similitudes. Y a diferencia de la anterior red neuronal, los pesos sinápticos se obtuvieron por un proceso de aprendizaje.

A continuación, en la Tabla III se muestran las principales características de las dos redes neuronales utilizadas GRNN y LVQNN.

IV. RESULTADOS

En esta sección se muestran los resultados de las Tablas de Contingencia, y de las métricas Precision, Recall y Magnitud-F, cuando se utilizaron las dos topologías de redes neuronales

GRNN y LVQNN. A continuación, se hace la descripción de la evaluación de cada tipo de RNA utilizada.

TABLA III
PRINCIPALES CARACTERÍSTICAS DE LAS REDES NEURONALES GRNN Y LVQNN

TIPOS DE REDES NEURONALES	VECTOR DE ENTRADA	CAPA OCULTA	CAPA DE SALIDA
RED NEURONAL GRNN-GESTO	5 x 1 (a ₁ , a ₂ , a ₃ , a ₄ , a ₅)	11520 neuronas	24 neuronas (A, B, C, ..., Y)
RED NEURONAL GRNN-MPU	2 x 1 (a ₆ , a ₇)	1440 neuronas	3 neuronas (Y ⁺ , X, Y ⁻)
RED NEURONAL GRNN-CONJU	2 x 1 (GESTO, MPU)	2880 neuronas	24 neuronas (A, B, C, ..., Y)
RED NEURONAL LVQNN-GESTO	5 x 1 (a ₁ , a ₂ , a ₃ , a ₄ , a ₅)	96 neuronas	24 neuronas (A, B, C, ..., Y)
RED NEURONAL LVQNN-MPU	2 x 1 (a ₆ , a ₇)	9 neuronas	3 neuronas (Y ⁺ , X, Y ⁻)
RED NEURONAL LVQNN-CONJU	2 x 1 (GESTO, MPU)	10 neuronas	24 neuronas (A, B, C, ..., Y)

A. Evaluación de la red neuronal GRNN

La evaluación de esta red neuronal se realizó de forma independiente para cada una de las tres topologías.

- En relación a la GRNN-GESTO, la Tabla de Contingencia con la clasificación de las 120 muestras por cada gesto utilizado, confirma una cantidad nula para los Falsos Positivos y Falsos Negativos; y con ello, las métricas Recall y Precision fueron iguales a 1.
- En relación a la GRNN-MPU, la Tabla de Contingencia correspondiente a la clasificación, no presentó errores en la clasificación de los datos MPU obtenidos de cada gesto. Por lo cual, se confirma una cantidad nula para los Falsos Positivos y Falsos Negativos; y con ello, las métricas Recall y Precision fueron iguales a 1.
- Y, en relación a la GRNN-CONJU, la Tabla de Contingencia correspondiente a la clasificación de los gestos en función a las dos redes GRNN-GESTOS y GRNN-MPU, también presentó una cantidad nula para los Falsos Positivos y Falsos Negativos; y con ello, las métricas Recall y Precision fueron iguales a 1.

B. Evaluación de la red neuronal LVQNN

De la misma manera que en la red neuronal anterior, la evaluación fue realizada de manera independiente para cada una de las tres topologías.

- En relación a la LVQNN-GESTO, la Figura 6 muestra la Tabla de Contingencia con errores en la clasificación de 3 gestos particulares. Es decir, 120 muestras de la letra G (clase 7) son consideradas como parte de la letra T (clase 19), 120 muestras de la letra Q (clase 16) son consideradas como parte de la letra X (clase 23), y 120 muestras de la letra S (clase 18) son consideradas como parte de la letra O (clase 14). Por lo cual, las categorías 14, 19 y 23 presentan Falsos Positivos iguales a 120, y las categorías 7, 16 y 18 presentan Falsos Negativos iguales a 120. Por lo tanto, las métricas Precision y Recall no son iguales a 1 para las 24 clases, solamente para 18.

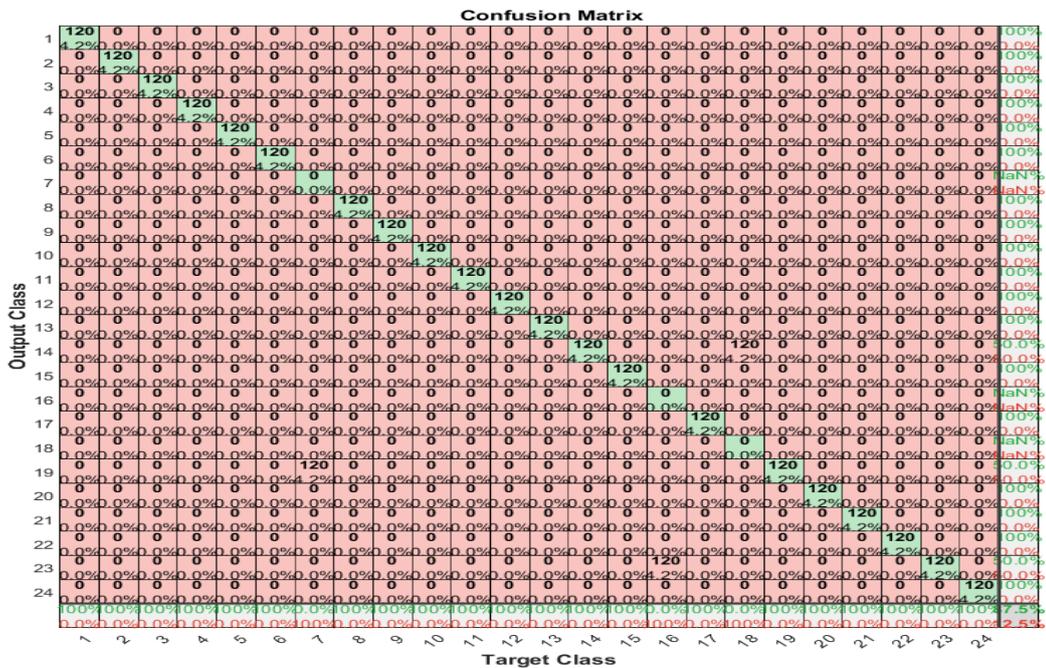


Fig. 6 Tabla de Contingencia para la red neuronal LVQNN-GESTO.

- En relación a la LVQNN-MPU, la Tabla de Contingencia correspondiente a la clasificación, no presentó errores en la clasificación de los datos MPU obtenidos de cada gesto. Por lo cual, se confirma una cantidad nula para los Falsos Positivos y Falsos Negativos; y con ello, las métricas Recall y Precision son iguales a 1.
- Y, en relación a la LVQNN-CONJU, la Figura 7 muestra la Tabla de Contingencia correspondiente a la clasificación de los gestos en función a las dos redes LVQNN-GESTO y LVQNN-MPU, donde se observa la presencia de un alto error de clasificación y por lo cual solamente la letra A fue reconocida; para esta clase la métrica Precision fue igual a 0.0417 y la de Recall igual a 1.

C. Análisis y Comparación de Resultados

Para realizar el análisis de los resultados obtenidos con las 3 topologías diferentes de redes neuronales planteadas, se optó por utilizar la Magnitud-F reemplazando $\beta = 1$ en (3). Esta magnitud permite determinar numéricamente la relación existente entre las métricas Precision y Recall de una red neuronal, de manera que se pueda definir numéricamente y con un solo parámetro el desempeño de esta para una función determinada. No obstante, para los cálculos se utilizaron los valores promedios obtenidos por las expresiones (1) y (2), al analizar las tres topologías por cada red neuronal propuesta.

Asimismo, debido al restringido nivel de clasificación de la red LVQNN-CONJU, se optó por reemplazarla por la red GRNN-CONJU. Esto dio origen a una mejora sustancial logrando la clasificación de aquellos gestos que no fueron clasificados anteriormente, y de esa manera se obtuvieron las nuevas métricas Precision y Recall iguales a 1. A continuación, en la Tabla IV, se muestran los valores promedio de las métricas Precision y Recall para cada topología de red neuronal propuesta, así como también la correspondiente Magnitud-F característica.

TABLA IV
VALORES PROMEDIO DE MÉTRICAS

RED NEURONAL GRNN						
	GRNN-GESTO		GRNN-MPU		GRNN-CONJU	
	PRECISION	RECALL	PRECISION	RECALL	PRECISION	RECALL
	1	1	1	1	1	1
Magnitud-F	1		1		1	
RED NEURONAL LVQNN						
	LVQNN-GESTO		LVQNN-MPU		LVQNN-CONJU	
	PRECISION	RECALL	PRECISION	RECALL	PRECISION	RECALL
	0.8125	0.8750	1	1	1.7375×10^{-3}	0.0417
Magnitud-F	0.8426		1		3.3359×10^{-3}	
RED NEURONAL LVQNN + GRNN						
	LVQNN-GESTO		LVQNN-MPU		GRNN-CONJU	
	PRECISION	RECALL	PRECISION	RECALL	PRECISION	RECALL
	0.8125	0.8750	1	1	1	1
Magnitud-F	0.8426		1		1	

D. Análisis de Resultados con Presencia de Ruido

Asimismo, para mostrar el desempeño de ambas redes neuronales propuestas, se agregó un ruido controlado a las señales existentes en proporciones de 0.1, 1, 5 y 10% desde el mismo software Matlab. Para ello, se empleó el comando RANDI para establecer números aleatorios en un rango establecido $[i_{min}, i_{max}]$, en un vector con dimensión $SZ_1 \times SZ_2$, y en proporción al porcentaje establecido según la siguiente expresión:

$$\text{Nivel de ruido \%} = \frac{i_{\max}}{\text{percentage}} * 100 \% \quad (4)$$

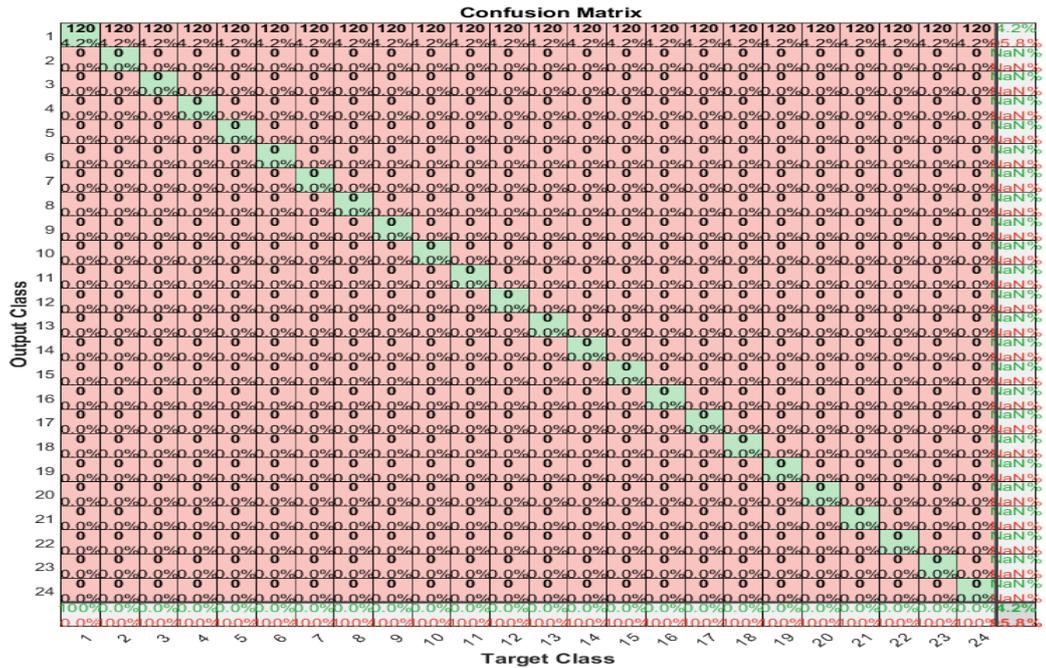


Fig. 7 Tabla de Contingencia para la red neuronal LVQNN-CONJU.

De esta manera, en la Tabla V se muestran los resultados alcanzados considerando una red GRNN-CONJU en vez de una red LVQNN-CONJU. Además, la comparación con presencia de ruido solamente se limitó a las redes neuronales de GESTO y CONJUGADA, porque la red correspondiente al movimiento de la mano, MPU, no presentó error alguno al ser sometida a los distintos niveles de ruido.

Igualmente, de la Tabla V, para el caso de 0.1% y 1% de ruido, se observan que las redes neuronales GR permanecen inalterables en el proceso de clasificación, en contraposición a los resultados arrojados por las redes LVQ, que si se ven afectadas en su tarea de clasificación. Luego, a partir de 5% de ruido, se llega a apreciar errores de clasificación en las redes neuronales GR, los cuales no pueden ser mejorados con los datos obtenidos de la red MPU según lo obtenido de su red Conjugada; de igual manera para la técnica LVQ, sus redes presentan un aumento en sus errores de clasificación, que obtienen una mejora con los datos MPU (un ejemplo particular de esto ocurre con las letras Q y S, que aumentan sus métricas de Precision y Recall).

V. CONCLUSIONES

Al implementar distintas redes neuronales GRNN, se comprobó de manera experimental que, el cálculo de sus pesos sinápticos se da por la propia red en función de los vectores de datos y de sus categorías previamente definidas, debido a que obedecen a una regla matemática. Por otro lado, en la implementación de las redes neuronales LVQNN se tuvo en consideración el valor del parámetro LR, porque influye en el aprendizaje de la misma; por lo cual, mientras más alto sea su valor el aprendizaje será más rápido, pero menos seguro; en el caso contrario, un valor LR más bajo hace un proceso de aprendizaje más lento, pero más fino y seguro.

TABLA V
PROMEDIO DE LAS MÉTRICAS PRECISION, RECALL Y MAGNITUD-F ANTE
DISTINTOS PORCENTAJES DE RUIDO

RUIDO DE 0.1%				
MÉTRICAS	GRNN-GESTO	LVQNN-GESTO	GRNN-CONJU	LVQNN-CONJU
PRECISION	1	0.8112	1	0.9814
RECALL	1	0.8750	1	0.9743
MAGNITUD-F	1	0.8419	1	0.9778
RUIDO DE 1 %				
MÉTRICAS	GRNN-GESTO	LVQNN-GESTO	GRNN-CONJU	LVQNN-CONJU
PRECISION	1	0.8245	1	0.9944
RECALL	1	0.8750	1	0.9941
MAGNITUD-F	1	0.8490	1	0.9942
RUIDO DE 5%				
MÉTRICAS	GRNN-GESTO	LVQNN-GESTO	GRNN-CONJU	LVQNN-CONJU
PRECISION	0.9823	0.8136	0.9823	0.9758
RECALL	0.9792	0.8677	0.9792	0.9778
MAGNITUD-F	0.9807	0.8398	0.9807	0.9768
RUIDO DE 10%				
MÉTRICAS	GRNN-GESTO	LVQNN-GESTO	GRNN-CONJU	LVQNN-CONJU
PRECISION	0.8897	0.7518	0.7965	0.9025
RECALL	0.8476	0.7767	0.8302	0.8805
MAGNITUD-F	0.8681	0.7640	0.8130	0.8914

REFERENCIAS

Del mismo modo, la evaluación de los algoritmos implementados para la conformación de una red neuronal mediante sus respectivas Tablas de Contingencia, brindó una idea clara y gráfica de la actuación de las distintas redes neuronales al ser sometidas tanto a los datos pre procesados, como también al ser sometidas a los datos afectados por ruido. El establecimiento de estas tablas con el comando “plotconfusion” del Matlab, permitió obtener los valores de las métricas Precision y Recall y en forma de porcentaje.

Respecto a la red neuronal GRNN-GESTO, con la aplicación del ruido, no se inmutó hasta la aplicación del 5% de este, en donde su precisión disminuye al 0.9823 con errores en 3 categorías que no pueden ser corregidas con los datos MPU, y manteniéndose este valor en su red neuronal conjugada GRNN-CONJU; luego, con la aplicación del 10% de ruido los elementos se distribuyeron a lo largo de 11 categorías distintas, siendo la categoría 11 (letra L) la que acumuló una mayor cantidad de elementos erróneos, obteniendo un valor de 0.5217, que tampoco obtiene mejoras con la aplicación de los datos de la red neuronal GRNN-MPU, y obteniendo como resultado final de la red neuronal conjugada algunos valores de clasificación correctamente nulos, y con ello la no existencia de la métrica precision en algunas categorías, y disminuyendo de esta manera el valor promedio de esta última métrica de un valor de 0.8897 en la red neuronal de GESTOS a 0.7965 en la red neuronal CONJUGADA.

Por otro lado, respecto a las redes neuronales LVQNN, son más susceptibles al ruido debido a la cantidad de errores arrojados desde la aplicación del 0.1% de ruido, y con mejoras sustanciales tras la adición de los datos clasificados por la red neuronal MPU en la estructura de la CONJUGADA que se le adicionó, logrando variar de 0.8112 en la red de GESTOS a 0.9814 en la red CONJUGADA; con la aplicación del 5% de ruido los errores se vuelven más notables en la red de GESTOS, manifestando una métrica de precision promedio igual a 0.8136, y siendo sustancialmente corregidos en la red CONJUGADA con valores de precisión muy cercanos a la unidad en todas las categorías, así como una métrica precision promedio de 0.9758; luego, al elevar el ruido al 10% se obtiene un mayor aumento de fallas en la clasificación en la red de GESTOS, con un promedio de 0.7518, pero con una mejora significativa en la red CONJUGADA con el aumento de los resultados de la red MPU, obteniendo un aumento de la métrica precision promedio a 0.9025.

De los valores promedio de las redes neuronales con datos pre-procesados, correspondientes a la Tabla IV, se observó que la Magnitud-F, en todas las redes neuronales GRNN fue igual a la unidad, por lo cual son más precisas en su clasificación al punto de no necesitar los datos de su red neuronal GRNN-MPU, para discriminar entre los gestos considerados como críticos; por el contrario, las redes neuronales LVQNN alcanzan similitudes entre los patrones, teniendo así un mayor nivel de confusión en los gestos considerados como críticos, lo cual se logró corregir con los datos obtenidos de la red neuronal LVQNN-MPU; esto, sumado a lo obtenido con la adición de ruido y los bajos niveles de tolerancia aceptados por ambas redes (1% para las GRNN y 0.1% para las LVQNN), evidencia la necesidad y la importancia del preprocesamiento de datos para una mayor precisión en la clasificación de los vectores, así como la importancia de una acertada y precisa toma de datos.

- [1] El Peruano. (2010, 21 de Mayo). El Peruano, Web. (Editora Perú) Recuperado el 27 de julio de 2020 de Diario El Peruano. [Online] Available: <https://busquedas.elperuano.pe/normaslegales/ley-queotorga-reconocimiento-oficial-a-la-lengua-de-senas-p-ley-n-29535-496850-3/>
- [2] Walczak, S. Encyclopedia of Physical Science and Technology: Artificial Neural Networks. 3rd ed., U.S.A.: Robert A. Meyers, 2001, pp. 631-645.
- [3] Kim, P., MATLAB Deep Learning with Machine Learning, Neural Networks and Artificial Intelligence. República de Corea: Apress. 2017.
- [4] J. López-Noriega, M. Fernández-Valladares, and V. Uc-Cetina, "Glove-Based Sign Language Recognition Solution to Assist Communication for Deaf Users", in 11th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE), Ciudad del Carmen, México, October 2014.
- [5] M. Lin and R. Villalba (2014). Sign Language Glove. (Cornell University) Recuperado el 15 de Julio de 2020 de Cornell University. [Online]. Available: https://people.ece.cornell.edu/land/courses/ece4760/FinalProjects/f2014/rdv28_mjl256/webpage/
- [6] F. Kai-Yin, G. Nuwan, Ch. Chi-Tsun and T. Chi, "A Real-Time ASL Recognition System Using Leap Motion Sensors", in International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery, Xi'an, China, September 2015.
- [7] K. Fernández and G. Sandoval, "Diseño y construcción de un prototipo de sistema electrónico para conversión," Tesis, Universidad Nacional Pedro Ruiz Gallo, Facultad de Ciencias Físicas y Matemáticas, Lambayeque, Perú, 2017.
- [8] S. Praatek, et al, "Dynamic Tool for American Sign Language Finger Spelling Interpreter," in International Conference on Advances in Computing, Communication Control and Networking, Greater Noida, India, October 2018.
- [9] F. Csóka, J. Polec and R. Vargic, "Key Frame extraction from video sequences containing ASL Sign with Concealed Transmission Errors." in International Conference on Control, Decision and Information Technologies (CoDIT), Barcelona, España, April 2017.
- [10] D. Specht, "A General Regression Neural Network" IEEE Transactions on Neural Networks, vol. 2, no. 6, pp. 568-576, november 1991.
- [11] MathWorks. (2021). MathWorks. (M. Works.Inc, Productor) Recuperado el 03 de 08 de 2020, [Online]. Available: <https://la.mathworks.com/products/deep-learning.html>
- [12] M. Del Brio. Redes Neuronales y Sistemas Difusos, 2 ed., Madrid, España: RA-MA. Recuperado el Julio de 2002.
- [13] J. Flores, "Clasificación y reconocimiento de gestos estáticos de la mano basado en el alfabeto dactilológico de la lengua de señas del Perú aplicando redes profundas bajo características invariantes," Tesis, Universidad Nacional San Antonio Abad del Cusco, Cusco, Perú, 2017.
- [14] DIGEBE. (2015). Lengua de Señas Peruana (2 ed.). Lima, Perú: Finishing S.A.C. Recuperado el 22 de Julio de 2020, [Online]. Available: <http://repositorio.minedu.gob.pe/handle/MINEDU/5545>