

Strategy for the detection of types of facial gestures using SOM neural network.

Estrategia para la detección de tipos de gestos faciales usando red neuronal SOM.

Huarote Zegarra Raúl Eduardo¹, Vega Luján Yensi², Flores Masías Edward José³, Katherine Susan Llanos Chacaltana⁴, Díaz Reátegui Mónica⁵ and Lévano Stella Miguel Alfredo⁴

¹Universidad Nacional Tecnológica de Lima Sur, Perú, rhuarote@untels.edu.pe

²Universidad Nacional de Trujillo, Perú, yensi.vega@gmail.com

³Universidad Nacional Federico Villarreal, Perú, eflores@unfv.edu.pe

⁴Universidad Privada del Norte, Perú, susan.llanos@upn.edu.pe, miguel.levano@upn.edu.pe

⁵Universidad Privada Norbert Wiener, Perú, monica.diaz@uwiener.edu.

Abstract– This research aims to cover a need to be able to classify gestures, specifically the gestures of people's faces, which reflects the emotions of each person such as anger, fear, happiness and sadness. To be able to identify these gestures, it is necessary to apply a strategy, which is to prepare the digital image matrices in a sequence, such as converting to gray tone, finding the orientation, applying the sobel and medfilt2 algorithm, so that this result can enter to a SOM neural network and be able to be classified according to the gestures. Labeling as 0 to anger, 1 to fear, 2 to happiness and 3 to sadness. To corroborate this strategy, a public database of faces has been taken, being 160 images of faces for the training and for the tests 15 images were used that were not part of the training and each image obtained in .jpg format in different dimensions, achieving demonstrate with this strategy an affectivity of 96.0% certainty in the identification of gestures.

Keywords– Strategy, SOM neural network, gestures, emotions.

I. INTRODUCCIÓN

Según [1] considera que en el mundo, uno de cada siete jóvenes de 10 a 19 años padece algún trastorno mental, un tipo de trastorno que supone el 13% de la carga mundial de morbilidad en ese grupo etario, esta información lleva a una expresión gestual que se debe de evidenciar cuando se le realiza una determinada evaluación, donde estos cambios emocionales deben ser considerados o por lo menos los cambios abruptos. Según [2] considera que los trastornos afectivos, de ansiedad y por consumo de sustancias en adultos, graves y moderados, tienen una brecha de tratamiento de 73,5% en la Región de las Américas, 47,2% en América del Norte y 77,9% en América Latina y el Caribe (ALC), estos datos deben ser considerados e identificados de manera gestual. Así también [3] considera que las emociones se centra en los rasgos gestuales avizorando el comportamiento de la

personalidad siendo estable y variable, de las cuales son la empatía, asertividad, optimismo, impulsividad, extroversión, motivación entre otros. Para [4] considera mediante interacciones que son emocionalmente propicias y receptivas se refleja en la expresión gestual. Así también [5] hace referencia a poder regular el estado emocional, contribuyendo a que el bebé desarrolle estrategias temporales para tranquilizarse, expresados en sus gestos. Así [6] considera que la expresión facial representado en las diferentes emociones, tiene una función social muy importante, y que a su vez facilita la interacción entre las personas. Este suceso de gestualizar en la iteración con otras personas tiene una base neurológica, que no se aísla del contexto en el cual quiera expresar sus ideas, tampoco de la experiencia acumulada por la interacción entre las personas en dicho contexto. [7] menciona que la neurociencia ayuda a entender como ocurren los procesos emocionales y de la expresiones gestuales (lenguaje no verbal).

En este campo se han realizado múltiples y relevantes investigaciones enmarcadas en este escenario, tales como [8] logra analizar la presencia/ausencia de habilidades preverbiales (gestuales) en niños con ausencia del lenguaje no autistas.

Por tanto la presente investigación lo que pretende es identificar en que grupo de los diferentes gestos pertenecen en base a lo aprendido por la red neuronal artificial SOM.

A. Inteligencia Artificial

Para [9] la inteligencia artificial lo define como “el medio por el cual las computadoras, los robots y otros dispositivos realizan tareas que normalmente requieren de inteligencia humana”, así también [10] considera a la IA que “es el estudio de cómo lograr que las computadoras realicen tareas que, por el momento, los humanos hacen mejor”, también para [11] considera a la inteligencia artificial como “la construcción de

Digital Object Identifier (DOI):

<http://dx.doi.org/10.18687/LEIRD2022.1.1.80>

ISBN: 978-628-95207-3-6 ISSN: 2414-6390

algoritmos ha logrado que las computadoras y las máquinas aprendan a resolver problemas” en los diferentes campos como la medicina, entre otros campos.

Por tanto la inteligencia artificial emula en lo posible todas las capacidades propias del humano plasmadas en algoritmos, agenciándose de tecnologías de información (software y hardware) para su implementación.

B. Red neuronal Artificial

Para [12] las redes neuronales artificiales “están formadas por capas de variables interconectadas entre sí con unos pesos que el algoritmo va calculando y ajustando mediante iteraciones (repeticiones) de un proceso que se conoce como «descenso del gradiente»”. Para [13] atribuye el éxito de las redes neuronales artificiales en usar varios algoritmos de optimización y varias estructuras de la red, como la normalización por lotes.

Por tanto la red neuronal artificial emula en algoritmos la manera de aprender basándose del modelo biológico del humano, a pesar que existen otros factores o niveles que aún no están contempladas.

C. Red neuronal Artificial

A partir de la publicación de [14] ha mostrado el modelo de red neuronal autoorganizativo, y se basó para su investigación en la manera como la información es captada a través de los órganos sensoriales de los animales (específicamente en el córtex superior), donde las zonas detectoras se ordenan y representan internamente formas de mapas bidimensionales, en este mapa las neuronas compiten y se agrupan según las características ingresadas. Este tipo de red neuronal permite clasificar de acuerdo a los patrones característicos, este tipo de red neuronal tiene un aprendizaje basado en competencia, por tanto no se considera un tutor para que maneje el error en el proceso de aprendizaje, ya que no se considera un valor deseado en el cual tienda los resultados. Formalizado en la ecuación (1) para encontrar la neurona vencedora ya se realiza por un método competitivo, la ecuación (2) refleja la actualización de pesos, la ecuación 3 para encontrar el nivel de afectación de las neuronas vecinas, considerando la matriz gaussiana y la ecuación (4) para encontrar el radio. En la Fig. 1 un ejemplo de agrupación de datos mediante colores. Este modelo de clasificación por medio de la red SOM se puede aplicar en la clasificación enfermedades oculares a partir de las imágenes que son obtenidas de una base de datos públicos, considerando los parámetros de entrada para poder utilizar la red neuronal.

$$\min \|X_p - W_j\| = \min \sum_{i=1}^N (X_{pi} - W_{ji})^2 \quad (1)$$

Dónde: $X_p: X_{p1}, X_{p2}, \dots, X_{pN}$, son los patrones de entrada y el vector de pesos $W_j: W_{j1}, W_{j2}, \dots, W_{jN}$

$$\Delta W_{ji}(n+1) = \alpha(n)(X_{pi} - W_{ji}(n)) \quad (2)$$

Donde n : es el número de ciclos, α es la tasa de aprendizaje

$$h(|i-g|, t) = \begin{cases} 0, & \text{si } |i-g| > R(t) \\ 1, & \text{si } |i-g| \leq R(t) \end{cases} \quad (3)$$

Donde $|i-g|$ hace es medir el nivel de afectación de las neuronas vecinas en la iteración t .

$$r(t) = r_0 e^{-\frac{t}{\tau}} \quad (4)$$

Dónde: r_0 es el radio inicial para cubrir todos los nodos.

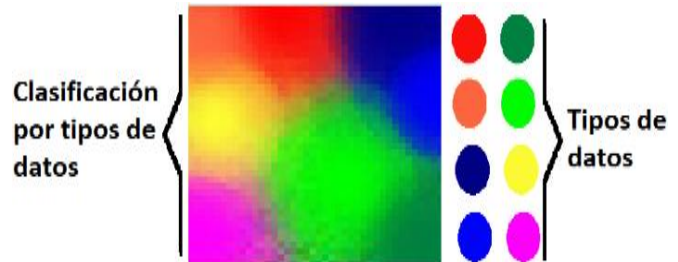


Fig. 1. Creación de figura en ejemplo del mapa autoorganizado de kohonen

En la Fig. 2 muestra la arquitectura competitiva de la red neuronal SOM, en la capa de entrada que tiene N neuronas de manera lineal donde representan la información inicial y a su vez estos se transmite los datos a la capa competitiva, formada por M neuronas para realizar el proceso de aprendizaje. En cada neurona i está conectada con cada una de las neuronas de la capa de competitiva j mediante un W_{ji} que tiene asociado, considerando que una neurona elegida pertenece a una vecindad N_j de la neurona j , y la cantidad de sus vecinas es fijo durante el proceso de entrenamiento, por tanto la elegida es resaltada y los que están a su costado tienen cierto grado de afectación en función de la distancia del elegido, donde por lo general se establece con una matriz llamada gaussiana.

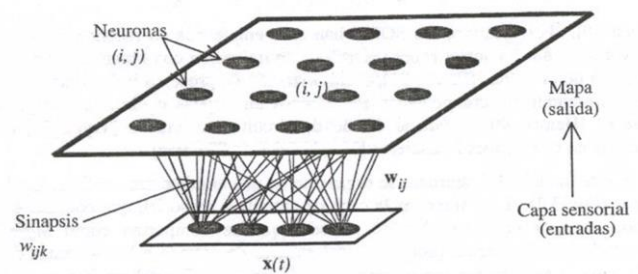


Fig. 2. Representación de la clasificación de tipo de gestos usando la red SOM

D. Estrategias aplicadas

Existe diferentes puntos de vista respecto a las estrategias, en esta investigación se va a considerar a las vinculadas con las redes neuronales artificiales, como [15] usa una estrategia previa para convertirlo a patrones de entrada a las redes neuronales artificiales, que consta de una serie de pasos que restringe los contactos neuronales durante el cableado cerebral siendo procesos dinámicos en tiempo y espacio. Otra estrategia aplicada en este sector es de [16] donde clasifica los tipos de células, basándose en el análisis de la vinculación morfológica de la proyección axonal, ya que cambian en

función del tiempo, buscando una medida adecuada, ya que son diversos. También [17] aplica estrategia sensoriomotora a las neuronas, para la discriminación de formas usando modelo lineal generalizado. La estrategia aplicada de [18] para la detección de Covid-19 en imágenes de torax, fue aplicando una eliminación de ruido previa al ingreso de la red, mejorando la sensibilidad y especificidad. Otra estrategia aplicada por [19] consiste en aplicar previamente al ingreso de datos a la red neuronal artificial, a cada imagen de muestra granular extrae información del PSD y morfología, a partir de ello se aplica la red neuronal BiLSTM. En el ámbito de la energía [20] aplica una estrategia basado en CEEMDAN para posteriormente aplicar a las redes neuronales artificiales, para poder pronosticar la radiación solar con mayor precisión. También en la investigación de [21] aplica la estrategia a las imágenes, que es una combinación de contornos CT y GTV, para tener el insumo para la red neuronal artificial y así poder predecir los diferentes resultados clínicos en los datos de cáncer de H&N.

Tal como se muestra en la Fig. 3. Realiza un gráfico genérico de este punto acerca de la estrategia previa para la aplicación de redes neuronales artificiales a las imágenes de los rostros (gestos).



Fig. 3. Estrategia previa a la aplicación de la IA

II. ESTRATEGIA

A. Estrategia y preparación de datos

Se realiza la estrategia previa para la aplicación de redes neuronales artificiales a las imágenes de los rostros (gestos), lo cual se basa en preparar las imágenes a nivel pixel desde el punto de vista Visión artificial.

Queda plasmada este código fuente de la estrategia específicamente aplicando el procesamiento digital a las imágenes, en la implementación en el código fuente python 3.x, tal como se muestra en la Fig. 4.

```
def sobel(i):
    m,n = i.shape
    Gx = np.zeros([m-2,n-2],np.float32)
    Gy = np.zeros([m-2,n-2],np.float32)
    gx = [[-1,0,1],[-1,0,1],[-1,0,1]]
    gy = [[1,-1,1],[0,0,0],[1,-1,1]]
    for j in range(m-2):
        for i in range(n-2):
            Gx[j,i] = sum(sum([gx[i1,j1]*i[i1+j1,i2+j1]]))
            Gy[j,i] = sum(sum([gy[i1,j1]*i[i1+j1,i2+j1]]))
    return Gx,Gy

def medfil2(Gx,Gy):
    m,n = Gx.shape
    temp = np.zeros([m*2,d/2],n*2,d/2),np.float32
    salida = np.zeros([m,n],np.float32)
    temp[:,0:m] = Gx
    temp[m:,0:m] = Gy
    for i in range(1,m):
        for j in range(1,n):
            A = np.asarray(temp[i-1:i+2,j-1:j+2]).reshape(-1)
            salida[i,j-1] = np.sort(A)[4]
    return salida

def orientacion(patron,p):
    Gx,Gy = sobel(patron)
    Gx = medfil2(Gx)
    Gy = medfil2(Gy)
    m,n = Gx.shape
    mOrientaciones = np.zeros([m/n,n/n],np.float32)
    for i in range(m/n):
        for j in range(n/n):
            YY = sum(sum([Gx[i1*j1,j1]*Gy[i1*j1,j1]]))
            XX = sum(sum([Gx[i1*j1,j1]**2+Gy[i1*j1,j1]**2]))
            mOrientaciones[i,j] = (0.5*math.atan2(YY,XX) + math.pi/2.0)/(180.0/math.pi)
    return mOrientaciones

def normalizar(archivo):
    data = loaddata(archivo+'.tif')
    r = resize(data, (256, 256))
    resized_data = gabor(r, frequency=0.5)
    m = len(resized_data[0])
    n = len(resized_data[0][0])
    larray = np.array(resized_data[0],np.float32)
    larray = larray[1:n-1,1:n-1]
    patron = larray[1:n-1,1:n-1]
    EE = orientation(patron,14)
    return np.asarray(EE).reshape(-1)
```

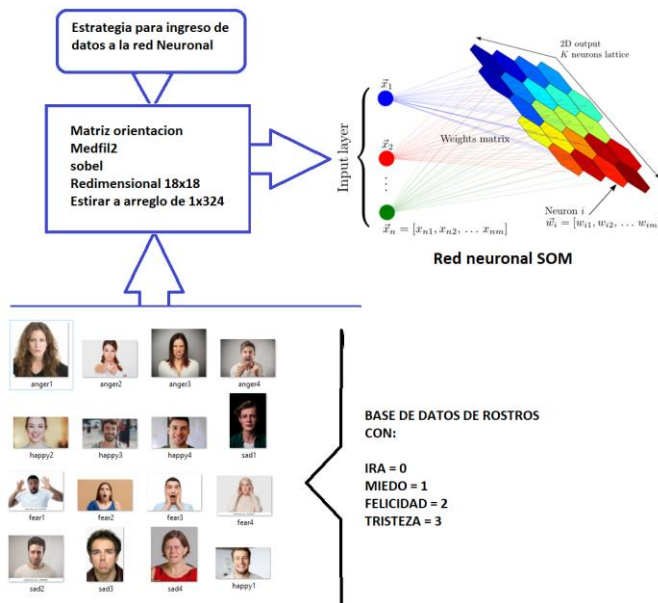
Fig. 4. Código fuente de la estrategia previa a la aplicación de la IA

En la Fig. 5, se evidencia en detalle de que cada uno de los pasos para aprender los gestos a querer agrupar se encuentra en formato .jpg y tienen una dimensión variada en pixeles, esto antes de aplicar la estrategia, ya que después de aplicar la estrategia se obtiene su representativo en dimensión 18x18 de cada imagen. Si este último se le estira a se obtendrá un arreglo de 1x784, con valores de [0 a 255], pero este arreglo contiene valores enteros, lo que implica de manera necesaria pasarlo a reales en un rango de [0.0 a 1.0], para ello se usaría el método de escalación, tal como se muestra en la Table I, cabe resaltar que el proceso inverso es de regresar al tono de gris en un valor de tipo pixel, respetando su valuación entera.

TABLE I
TABLA DE ESCALADO Y EL PROCESO INVERSO

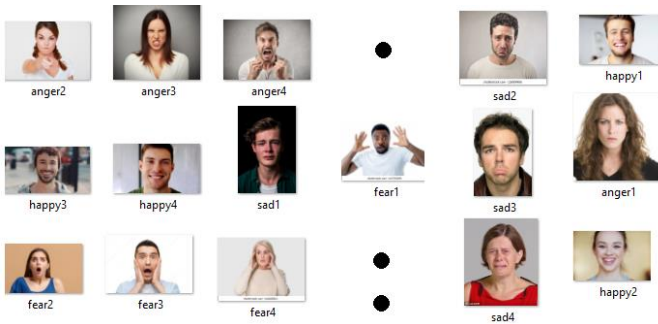
Ecuación	Operaciones	
	Observación	Detalle
$r = \frac{Y - Min}{Max - Min}$	Necesario para escalar en el rango de 0 a 1	r: valor escalado Y: valor a escalar,
$Y = r(Max - Min) + Min$	Necesario para regresar la proporción original en función de r	Min: Valor mínimo del pixel ósea 0 Max: máximo del pixel ósea 255

El conjunto de datos de imágenes de rostros con los gestos establecidos, tienen la particularidad expresar los gestos IRA, MIEDO, FELICIDAD, TRISTEZA, que consta de un conjunto de 160 ejemplares de cada uno, así como para el testeo se tiene un conjunto de prueba de 15 imágenes ejemplares. Cada ejemplar es una imagen en formato .jpg de dimensión variable, asociada con una etiqueta de cada gesto. Por tanto son 4 clases las que se van a considerar para la presente investigación. Las imágenes de esta fuente de datos de rostros sirven para poder comprobar la investigación de poder agruparlas en las 4 clases gestuales establecidas y para la evaluación comparativa de los algoritmos de aprendizaje automático basado en red SOM. Teniendo en cuenta que para la evaluación o mapeo se debe realizar el mismo procedimiento a las imágenes para el entrenamiento (ósea hacer su representativo).



III. PROCEDIMIENTO

En el proceso de aprendizaje de la red neuronal SOM al aplicar la estrategia establecida, genera un gráfico de índices, propio de la red SOM, donde se ve agrupada por las 4 clases a las 15 imágenes gestuales, tal como se ve en la Fig. 6. Estas clases se pueden representar como clase 0 a IRA, 1 a MIEDO, 2 a FELICIDAD y 3 a TRISTRZA.



La previa de pasos para aplicar la Fig. 6, como resultado obtenemos en, necesario para que pueda ingresar a la red neuronal artificial SOM, considerando que estos dato (finales) es necesario estirarlo para que ingrese como patrones de entrada. Cabe resaltar que la aplicación de la estrategia es por cada gesto en el formato de la Fig. 7, mas no al conjunto.

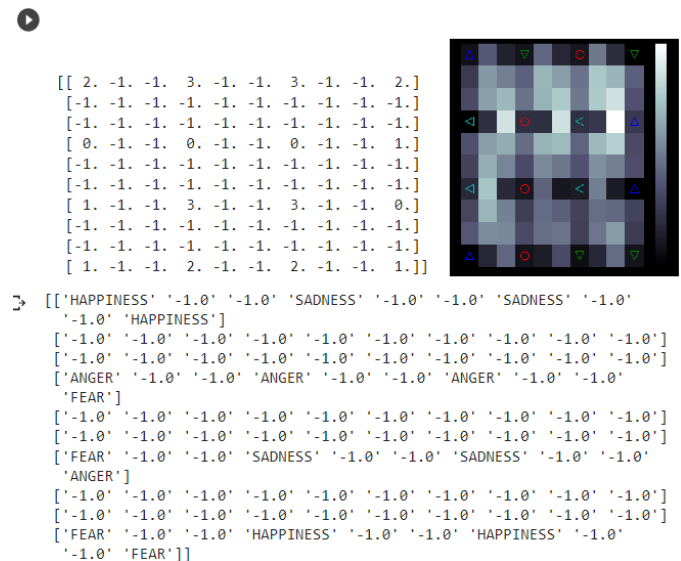


En la Fig 8 que es su representación de etiquetas de la matriz index, agrupada después del aprendizaje no supervisado, donde la parte circular rojo representa a la clase IRA, la parte triangular verde representa a la clase de MIEDO, la parte triangular verde representa a la clase de FELICIDAD y la parte triangular celeste representa a la clase de TRISTEZA. A partir de los datos entrenados, representado en la matriz index, se ha realizado las pruebas, donde a las pruebas de las mismas imágenes de entrenamiento (160 imágenes) se ha obtenido un 100% de certeza, frente a las imágenes que no han sido parte del entrenamiento (15 imágenes), obtuvo un resultado de 96.0% en promedio, tal como se muestra en la Table II, y una de las pruebas se evidencia en la Fig. 9.

TABLE II
TABLA DE PORCENTAJE DE CERTAZA EN MAPEO SEGUN LAS
CLASE DE GESTOS

Clase	Etiqueta	% de certeza con imágenes de entrenamiento (160)	% de certeza con imágenes que no son parte del entrenamiento (15)
0	IRA	100 %	100 %
1	MIEDO	100 %	93 %
2	FELICIDAD	100 %	100 %
3	TRISTEZA	100 %	91 %

En la Fig. 9 parte derecha se visualiza que en base a la variable de la cantidad de neuronas y ver la variación según la cuantización del error, se observa que van desde 10 hasta el valor 40 como límite para el valor las cantidades de neuronas, según los resultados reflejados en la gráfica, con valores 10 el aprendizaje toma un tuna cuantización del error más visible de 24.5%, mientras más neurona utilicemos, hasta llegar a un valor ideal para las neuronas 15, generando una cuantización de error de 18,11%, y decreciendo conforme incrementar la cantidad de neuronas.



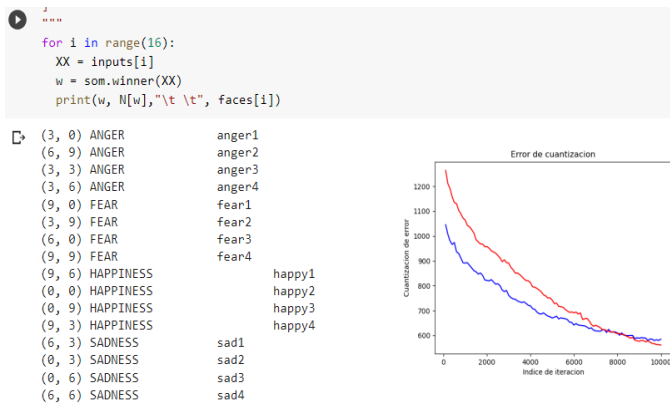


Fig. 9. Porción de resultados de imágenes gestuales

Un análisis de la variación de la cuantización de error con respecto a la cantidad de neuronas en la capa competitiva, se demuestra que a partir de la dimensión de 15x15 neuronas en adelante el error se mantiene constante, por tanto bajo este caso en particular, se recomienda tener esta cantidad de dimensión, ya que neuronas adicionales a 15x15 es innecesarios por que se obtiene el mismo error, tal como se muestra en la Fig. 10.

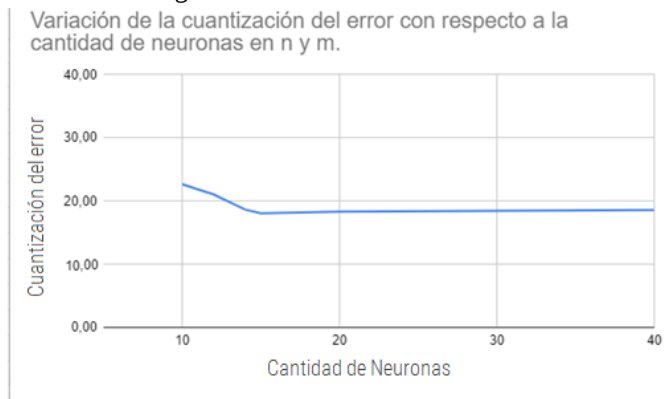


Fig. 10. Cuantización de error con respecto a la cantidad de neuronas cuadradas

IV. CONCLUSIONES

Al aplicar la estrategia propuesta para esta investigación y aplicar la red SOM, se obtuvo como resultado de efectividad de 93.7% en la clasificación de gestos faciales (ira, miedo, felicidad y tristeza), para lograr ello se usó 160 imágenes de la base de datos pública de Fundus-images para el entrenamiento de la red, representando un valor alentador, considerando que las pruebas para el testeo se ha realizado las 15 imágenes clasificadas en cada una de las clases establecidas y que no fueron parte del entrenamiento.

Cada vez que la matriz index genere los datos de las clases, se agruparán las clases de manera diferente, los valores de las etiquetas que representa cada estado de ánimo estarán en posiciones diferentes. En la Fig. 9, se muestra la gráfica del resultado de la aplicación del modelo de la red SOM

propuesto por [18] y la aplicación de la estrategia en la Fig. 5. Al aplicar la red neuronal SOM para clasificar tipos de tipos de gestos faciales aporta a nuevas investigaciones, ya que se puede extender a otros gestos no considerados en la presente investigación. Se ha utilizado la librería Minisom en un entorno Colab, donde nos proporcionan datos suficientes para realizar un buen análisis de nuestros resultados y obtener más información. Tal como se muestra en la Fig. 11, con los parámetros para obtener los resultados esperados, como la dimensión de la capa competitiva de 10x10, la dimensión de la capa de entrada de 1x324, el valor sigma = 1.0, el valor alpha inicial máximo es de 0.5, la topología para el nivel de afectación de los vecinos es de tipo hexagonal, la cantidad de iteraciones es de 50000 y la matriz de afectación a los vecinos es la gaussiana.

```
[13] max_iter = 50000
    alpha_max = 0.5

    som = MiniSom(x=10, y=10, input_len=324, sigma=1.0, learning_rate=alpha_max, activation_distance='euclidean',
                  topology='hexagonal', neighborhood_function='gaussian')
    som.random_weights_init(inputs)
    som.train_random(data=inputs, num_iteration=max_iter)
```

Fig. 11. Parte del código Fuente para realizar el aprendizaje

V. DISCUSIÓN

Si hacemos la comparativa de los resultados obtenido por [26] indica que usando árbol de decisión tiene una precisión de 87.5% frente a la presente investigación usando una estrategia de datos aplicando el algoritmo Sobel, orientación, medfilt2 y a esto redimensionarlo para obtener su representativo en una dimensión pequeña tal como se aprecia en la Fig 3. A estos nuevos datos vienen a ser los patrones de entrada a la red neuronal SOM obteniendo un 93.7% de precisión. Teniendo en cuenta que los casos usan imágenes que tienen retinopatía diabética, tal como la investigación de [23] que tuvo un 96.11% de certeza implicando una diferencia mayor de 2.41% de precisión frente a la presente investigación, por tanto se requiere unos ajustes para alcanzar o superar esta precisión.

VI. RECOMENDACIONES

Para tener un criterio más cercano en función del análisis psicológico de la persona usando la red neuronal SOM, no necesariamente en un área de ciencia, si no en un área donde necesariamente se requiera saber que tan cambiante es el estado de ánimo de la persona como en colegios, institutos, universidades, centros laborales, etc.

REFERENCES

- [1] Organización Mundial de la Salud, "Salud mental del adolescente", 2021. [Online]. Available: <https://www.who.int/es/news-room/fact-sheets/detail/adolescent-mental-health> [Accessed: Mar. 01, 2022].
- [2] Organización Panamericano de la Salud, "Salud Mental", 2021. [Online]. Available: <https://www3.paho.org> [Accessed: Mar. 01, 2022].

- [3] F. Ahumada, "La relación entre inteligencia emocional y salud mental" Buenos Aires: MERCOSUR, 2011
- [4] Organización Panamericano de la Salud, "El cuidado cariñoso y sensible para el desarrollo en la primera infancia", 2018. [Online]. Available: <https://www.who.int/> [Accessed: Mar. 01, 2022].
- [5] Unicef, "Orientación para familias sobre Autorregulación emocional y alternativas para eliminar la violencia en la crianza", 2020. [Online]. Available: <https://www.unicef.org/lac/media/13906/file/Autorreguacion-Emocional-Famiias-Prevencion-Violencia.pdf> [Accessed: Mar. 01, 2022].
- [6] Gordillo F, Pérez MA, Arana JM, Mestas L, López RM, "Importancia de la expresión facial de las emociones", 2015
- [7] No Verbal, "Las emociones y la expresión en el lenguaje corporal", 2019. [Online]. Available: <https://www.no-verbal.com/emociones-y-expresion-lenguaje-corporal/en> [Accessed: Mar. 01, 2022].
- [8] V. Maggio, C. Abadia, A. Jacob, D. Vázquez Durrieu and M. Wolman, "La comunicación preverbal en niños con ausencia del lenguaje", Revista de Logopedia, Foniatría y Audiología, 2021.
- [9] Gómez, R., La inteligencia Artificial ¿Hacia dónde nos lleva? ¿Cómo lo vez?, México, UNAM, 2013.
- [10] Elaine Rich y Knight Kevin. Inteligencia Artificial. Segunda Edición. McGraw Hill: México, 1994
- [11] M. Regalado, A. Medina and C. Pizarro, "Artificial intelligence: phones smart phones and advanced technology stethoscopes", Atención Primaria, Volume 53, Issue 10, December 2021.
- [12] A. Martorell, A.Martin-Gorgojo, E.Ríos-Viñuela, J.M.Rueda-Carnero, F.Alfageme and R.Taberner, "Artificial Intelligence in Dermatology: A Threat or an Opportunity?", Actas Dermo-Sifiliográficas, Volume 113, Issue 1, January 2022, Pages 30-46.
- [13] H. Zhang, L. Feng, X. Zhang, Y. Yang and J. Li, "Necessary conditions for convergence of CNNs and initialization of convolution kernels", Digital Signal Processing, Volume 123, 30 April 2022.
- [14] K. Teuvo, "Self-organizing maps", Berlin: Springer, 1989
- [15] E. Agi1, A. Kulkarni and P. Robin Hiesinger, "Neuronal strategies for meeting the right partner during brain wiring", Current Opinion in Neurobiology, Volume 63, August 2020, Pages 1-8.
- [16] J. Winnubst, N. Spruston and J. AHarris, "Linking axon morphology to gene expression: a strategy for neuronal cell-type classification", Current Opinion in Neurobiology, Volume 65, December 2020, Pages 70-76.
- [17] C.Rodgers, R. Nogueira, B. Pil, E. Greeman, J. Park, Y. Hong, S. Fusi and R. Bruno, "Sensorimotor strategies and neuronal representations for shape discrimination", Neuron, Volume 109, Issue 14, 21 July 2021, Pages 2308-2325.e10.
- [18] M. Momeny, A. Neshat, M. Hussain, S. Kia, M. Marhamati, A. Jahanbakhshi and G. Hamarneh, "Learning-to-augment strategy using noisy and denoised data: Improving generalizability of deep CNN for the detection of COVID-19 in X-ray images", Computers in Biology and Medicine, Volume 136, September 2021.
- [19] P. Zhang and Y. Zhen-Yu, "A novel deep learning-based modelling strategy from image of particles to mechanical properties for granular materials with CNN and BiLSTM", Computer Methods in Applied Mechanics and Engineering, Volume 382, 15 August 2021
- [20] B. Gao, X. Huang, J. Shia, Y. Taia and J. Zhang, "Hourly forecasting of solar irradiance based on CEEMDAN and multi-strategy CNN-LSTM neural networks", Renewable Energy, Volume 162, December 2020, Pages 1665-1683.
- [21] S. Pang, M. Field, J. Dowling, S. Vinod, L. Holloway and A. Sowmy, "Training radiomics-based CNNs for clinical outcome prediction: Challenges, strategies and findings", Artificial Intelligence in Medicine, Volume 123, January 2022.
- [22] S. Wy, S. Choe, Y. Jeong Lee, E. Bak, M. Jang, S. Chan Lee, A. Ha, J. Wook Jeong, K. Ho Park and Y. Kook Kim, "Decision Tree Algorithm-Based Prediction of Vulnerability to Depressive and Anxiety Symptoms in Caregivers of Children with Glaucoma", American Journal of Ophthalmology, 2022.
- [23] G.Sivapriya, V.Praveen, P.Gowri, S.Saranya, S.Sweetha, Kukunoor Shekar, "Segmentation of Hard exudates for the detection of Diabetic Retinopathy with RNN based semantic features using fundus images", 2022, MaterialsToday proceedings