

Reconocimiento de figuras geométricas por medio de Redes Neuronales empleando Momentos Invariantes de Hu.

Saúl Renato Aragón Moreyra, Víctor Gudberto Jimarez Rojas

Departamento de Sistemas y Computación, Instituto Tecnológico de Oaxaca
Oaxaca, México

drone_cam@outlook.es
vicgtor2004@gmail.com

Abstract - En este trabajo se presenta el uso de las redes neuronales para el reconocimiento de figuras geométricas a través de imágenes tomadas en tiempo real. Se explora la estructura fundamental de las RNA (Las Neuronas y las capas que la componen), así como el perceptrón y uno de los principales algoritmos de entrenamiento llamado Backpropagation.

Se realizan pruebas con el entrenamiento, que es parte fundamental de una RNA, para que esta sea capaz de arrojar resultados por medio de una serie de datos de entrada.

Este trabajo emplea procesamiento de imágenes para así poder extraer los datos necesarios para posteriormente ser evaluados por la red.

Gracias a la librería OpenIMAJ se facilita el procesamiento de la imagen, como la escala de grises, negativos, método OTSU hasta la extracción de los componentes conectados, que posteriormente pasan a otro proceso donde por medio del cálculo del factor de Compacidad y los cuatro primeros Momentos Invariantes de Hu se determina si alguno de los componentes de la imagen representa alguna de las figuras geométricas.

Con el correcto entrenamiento de la red, se denota una efectividad suficiente para que esta sea capaz de determinar las figuras que procesa, aunque no se puede decir que la red es 100% confiable sin embargo es suficiente para este trabajo.

I. INTRODUCCIÓN

Desde que la primera computadora fue construida, esta ha sido ocupada para la resolución de problemas cotidianos que el humano experimentaba día a día, tareas que en algunas ocasiones resultaban monótonas y bastante tediosas, gracias al uso de algoritmos básicos los problemas que resultaban más engorrosos se resolverían más fácil, pero con el paso del tiempo y el incremento de la complejidad de los problemas, se dieron cuenta que no todos los problemas se resuelven a través de uno o varios algoritmos.

Dichos problemas se presentaban principalmente a tareas básicas como la clasificación objetos o demás tareas que se necesite de la experiencia humana. Fue por ello que se fue desarrollando más y más la inteligencia artificial y con ello las redes neuronales artificiales, con dichas redes se podrían emular ciertas características propias de los humanos como la capacidad de memorizar y asociar hechos.

En sí las redes neuronales artificiales son un modelo artificial y simplificado de un cerebro humano que es capaz de

obtener experiencia a través del conocimiento. Su principal composición es de neuronas las cuales están conectadas unas a otra, así formando capas de las cuales contamos la capa de entrada, las ocultas y por último la capa de salida, estas a su vez conforman la red neuronal misma.

Estas redes neuronales artificiales se caracterizan por aprender de la experiencia y de tomar en cuenta de características principales de las secuencias de datos. Al aprender por experiencia, estas redes generalizan automáticamente por la estructura que las componen, la ventaja de dichas redes es la abstracción de características esenciales que presentan sus respectivas entradas al no tener parecido alguno.

II. FUNDAMENTOS TEÓRICOS

A. Modelo Biológico

La neurona es la unidad fundamental del sistema nervioso y en particular del cerebro. Cada neurona es una simple unidad procesadora que recibe y combina señales desde y hacia otras neuronas. Si la combinación de entradas es suficientemente fuerte la salida de la neurona se activa [1].

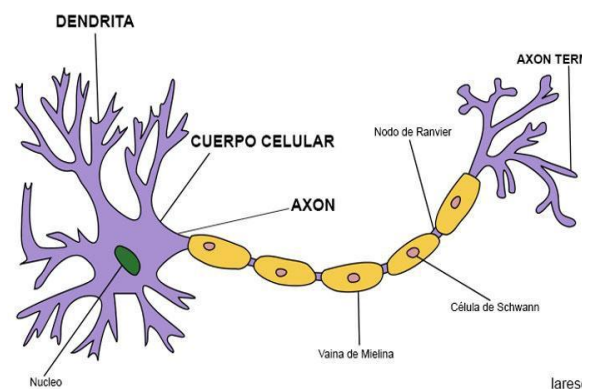


Fig. 1. Estructura de una Neurona Biológica.

El cerebro consiste en uno o varios billones de neuronas densamente interconectadas. El axón (salida) de la neurona se ramifica y está conectada a las dendritas (entradas) de otras neuronas a través de uniones llamadas sinapsis. La eficacia de la sinapsis es modificable durante el proceso de aprendizaje de la red [1].

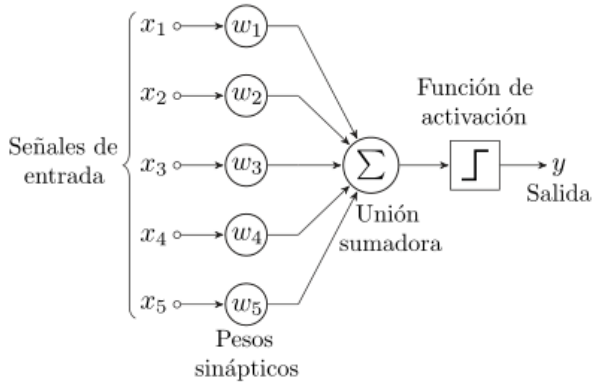


Fig. 2. Estructura de una Neurona Artificial.

B. Red Neuronal Artificial

La red neuronal intenta imitar los procesos del cerebro humano. Esta red consta de elementos primordiales para su funcionamiento que son las neuronas artificiales. A partir de estas, se pueden generar representaciones específicas, en base a un estado conjunto de ellas, se puede dar significado, a un objeto, una letra, una forma entre otras cosas.

Dentro de la red neuronal podemos encontrar tres tipos de neuronas, aquellas llamadas capa de entrada, que reciben estímulos externos y captan la información, dichas entradas son procesadas y enviada al siguiente tipo de neurona, la capa oculta, se denominan así ya que no tienen relación directa con la información de entrada ni de salida, una vez terminado la fase de procesamiento la información viaja a la etapa final, conocida como capa de salida, cuya función es dar una respuesta al sistema.

Las entradas a una neurona pueden provenir de fuentes externas o de otras neuronas en la red. Asimismo, la salida de una neurona es enviada a otras neuronas o al entorno. El conocimiento en una red neuronal está distribuido a lo largo de todo el sistema, debido a esto, se utilizan muchas interconexiones para obtener la solución de un problema en particular. [1]

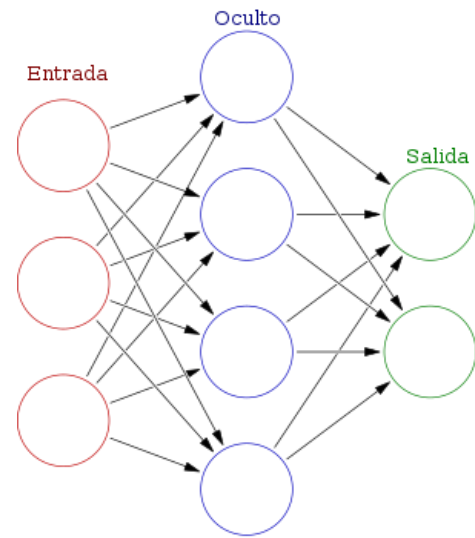


Fig. 3. Arquitectura de una Red Neuronal Simple.

C. Características de una RNA

La similitud con el cerebro humano les otorga una serie de características propias de este. Por ejemplo las RNA aprenden de la experiencia, generalizan de ejemplos previos a ejemplos nuevos y abstraen las características principales de una serie de datos.

Aprender: Adquirir el conocimiento de una cosa por medio del estudio, ejercicio o experiencia. Las RNA pueden cambiar su comportamiento en función del entorno. Se les muestra un conjunto de entradas y ellas mismas se ajustan para producir unas salidas consistentes.

Generalizar: extender o ampliar una cosa. Las RNA generalizan automáticamente debido a su propia estructura y naturaleza. Estas redes pueden ofrecer, dentro de un margen, respuestas correctas a entradas que presentan pequeñas variaciones debido a los efectos de ruido o distorsión [3].

Abstraer: aislar mentalmente o considerar por separado las cualidades de un objeto. Algunas RNA son capaces de abstraer la esencia de un conjunto de entradas que aparentemente no presentan aspectos comunes o relativos.

D. Ventajas de una RNA

Debido a su constitución y a sus fundamentos, las redes neuronales artificiales presentan un gran número de características semejantes a las del cerebro. Por ejemplo, son capaces de aprender de la experiencia, de generalizar de casos anteriores a nuevos casos, de abstraer características esenciales a partir de entradas que representan información irrelevante. Entre las ventajas que poseen se encuentra:[3]

- Aprendizaje Adaptativo
- Auto-organización
- Tolerancia a fallos
- Operación en tiempo real
- Fácil inserción dentro de la tecnología existente

E. Perceptrón

McCulloch – Pitts (1943) proponen la Unidad Lógica de Umbral, que demostraba que sistemas formados por varias neuronas artificiales eran capaces de procesar funciones lógicas y aritméticas [1].

Rosenblatt (1957) publicó el mayor trabajo de investigación en computación neuronal realizado hasta esas fechas. Su trabajo consistía en el desarrollo de un elemento llamado “Perceptrón” basado en la obra de McCulloch – Pitts.

Un sistema clasificador de patrones que puede identificar patrones geométricos y abstractos. El primero era capaz de aprender algo y era robusto, de forma que su comportamiento variaba sólo si resultaban dañados los componentes del sistema. Este modelo presentaba algunas limitaciones debido a que se trataba de un dispositivo en desarrollo. La mayor limitación la reflejan Minsky y Papert años más tarde, y ponían de manifiesto la incapacidad del perceptrón en resolver algunas tareas o problemas sencillos como por ejemplo la función lógica OR exclusivo.

F. Backpropagation

Rumelhart (1986) formalizó un método para que una red neuronal aprendiera una asociación que existe entre sus patrones de entrada y las clases correspondientes. Este método es conocido como backpropagation, propagación del error hacia atrás o retropropagación, y está basado en la regla de aprendizaje que es posible aplicar solo a modelos de redes multicapa.

Una característica importante de este algoritmo es la representación interna del conocimiento que es capaz de organizar en la capa o capas intermedias, para conseguir cualquier correspondencia entre la entrada y la salida de la red. Este algoritmo tiene la capacidad de autoadaptar los pesos de las neuronas de las capas intermedias para aprender la relación que existe entre un conjunto de patrones dados como ejemplo y sus salidas correspondientes [1].

G. Reconocimiento de Patrones

Los patrones se obtienen a partir de los procesos de segmentación, extracción de características y descripción donde cada objeto queda representado por una colección de descriptores. El sistema de reconocimiento debe asignar a cada objeto su categoría o clase (conjunto de entidades que comparten alguna característica que las diferencia del resto). Para poder reconocer los patrones se siguen los siguientes procesos:

- Adquisición de datos
- Extracción de características
- Toma de decisiones

El punto esencial del reconocimiento de patrones es la clasificación: se quiere clasificar una señal dependiendo de sus características. Señales, características y clases pueden ser de cualquiera forma, por ejemplo se puede clasificar imágenes digitales de letras en las clases «A» a «Z» dependiendo de sus píxeles o se puede clasificar ruidos de cantos de los pájaros en clases de órdenes aviares dependiendo de las frecuencias.

El reconocimiento de patrones se define como una ciencia que se encarga de los procesos propios de la ingeniería, computación y matemáticas los cuales están relacionados con objetos físicos o abstractos, con el propósito de extraer información que permita establecer propiedades de o entre conjuntos de dichos objetos [3].

Ya que la parte esencial del reconocimiento de patrones es la clasificación hay diversos enfoques o procedimientos que mejor se adaptan a la naturaleza del problema del reconocimiento de patrones, entre los cuales se encuentran:

- Geométrico: [3] En este enfoque los patrones deben ser graficables, se emplea el cálculo de distancias, geometría de formas, vectores números.
- Estadístico: [3] Este enfoque se basa en la teoría de la probabilidad y estadística, en el cual se tiene un conjunto de medidas numéricas con distribuciones de probabilidad conocidas o estimables, y a partir de ellas se realiza el reconocimiento.
- Sintáctico-estructural: [4] Se basa en encontrar relaciones estructurales que guardan los objetos de estudio, el objetivo es construir una gramática que describa la estructura del universo de los objetos de estudio, ha sido desarrollado utilizando la teoría de lenguajes formales, teoría de autómatas entre otras, para el estudio de objetos que no están descritos como vectores de atributos, por ejemplo: cadenas de símbolos o grafos.
- Lógico-combinatorio: [5] Se basa en la idea de que el modelado del problema debe ser lo más cercano posible a la realidad del mismo, sin realizar suposiciones que no estén fundamentadas, los atributos utilizados para describir a los objetos de estudio deben ser tratados cuidadosamente para no realizar operaciones que resulten antinaturales respecto al problema que se representa, comúnmente este enfoque se utiliza para conjuntos difusos y utiliza lógica simbólica, circuitos combinacionales y secuenciales.
- Redes neuronales: [4] Busca resolver problemas de reconocimientos de patrones mediante redes neuronales artificiales, las cuales son entrenadas para dar una respuesta ante determinados valores. De manera que una red neuronal puede dar una respuesta similar cuando se presenta un valor parecido a los que se utilizaron para entrenarla.

H. Librería OpenIMAJ

Es un conjunto de bibliotecas y herramientas para el análisis y la generación de contenido multimedia desarrollado por la Universidad de Southampton.

Contiene desde visión por computadora de vanguardia (por ejemplo, descriptores SIFT, detección de regiones sobresalientes, detección de rostros, etc.) y agrupamiento de datos avanzado, hasta software que realiza análisis sobre el contenido, el diseño y la estructura de páginas web [5].

I. Etiquetado de Componentes Conectados

De una imagen se agrupan sus píxeles en componentes basados en la conectividad, esto quiere decir que todos los píxeles en un componente conectado comparten valores de intensidad de píxeles similares y de alguna manera están conectados entre sí. Una vez que se han determinado todos los grupos, cada píxel se etiqueta con un nivel de gris o un color según el componente al que se le asignó.[6]

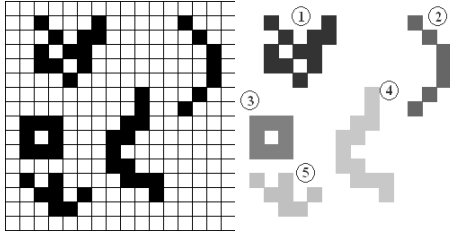


Fig. 4. Ejemplo de Etiquetado de componentes conectados.

J. Momentos Invariantes de HU

Los momentos de Hu son siete descriptores invariantes que cuantifican la forma de un objeto.

Pretenden extraer características para poder reconocer objetos que no se encuentren siempre en la misma posición, estén girados o su tamaño sea distinto [6].

Se ha demostrado que los primeros 6 momentos son invariables para la traducción, la escala, la rotación y la reflexión. Mientras que el signo del séptimo momento cambia para la reflexión de la imagen.

III. DESARROLLO

A. Delimitación de los datos

Para que el reconocimiento de patrones funcione de una manera óptima es necesario delimitar las clases que se emplearan para el entrenamiento de la red, por lo cual se opta por usar círculos, cuadros, triángulos y pentágonos debido a que existe un cierto margen de diferencia entre estas figuras.

Cabe mencionar que las figuras deben estar completamente rellenas para evitar el ruido en la imagen.

B. Obtención de los datos

Los datos se extraen a partir de un conjunto de imágenes con las figuras trazadas y rellenas a mano plasmadas en una hoja de papel blanco.

Se tomaron en cuenta 5 clases, círculos, cuadros, triángulos, pentágonos y basura, siendo esta ultima un factor importante en la correcta detección de las figuras.

Con un total parcial de 30 figuras por clase y un total general de 150 formas incluyendo ruido de imágenes.

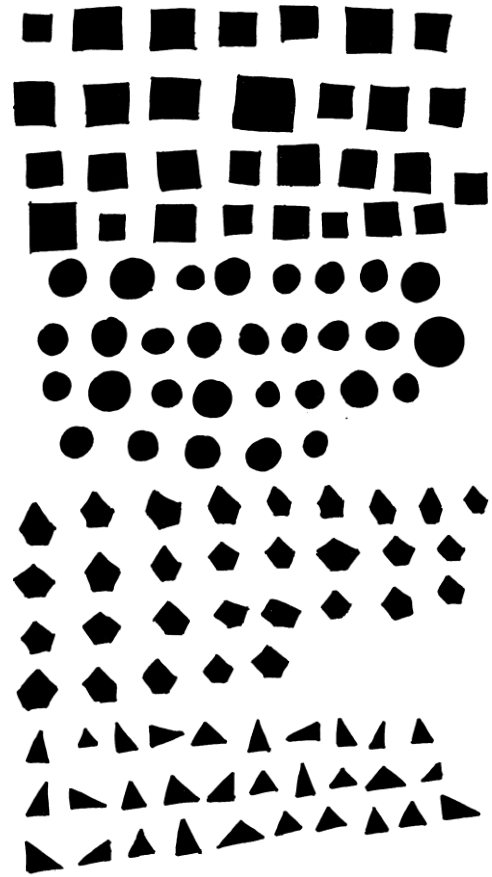


Fig. 5. Figuras geométricas dibujadas a mano.

C. Imagen a escala de grises

Para tener una mejor escena de la imagen a procesar es necesario aplicar una escala de grises, en la figura 3 se muestra un conjunto de imágenes completamente en blanco y negro, posteriormente las imágenes que ingresen de la webcam será necesario aplicar esta transformación para obtener un mejor resultado sobre la imagen esperada.[7]

Formula de la escala de grises.

$$P_{xy} = (R_{xy} + 0.3) + (G_{xy} + 0.59) + (B_{xy} + 0.11) \quad (1)$$

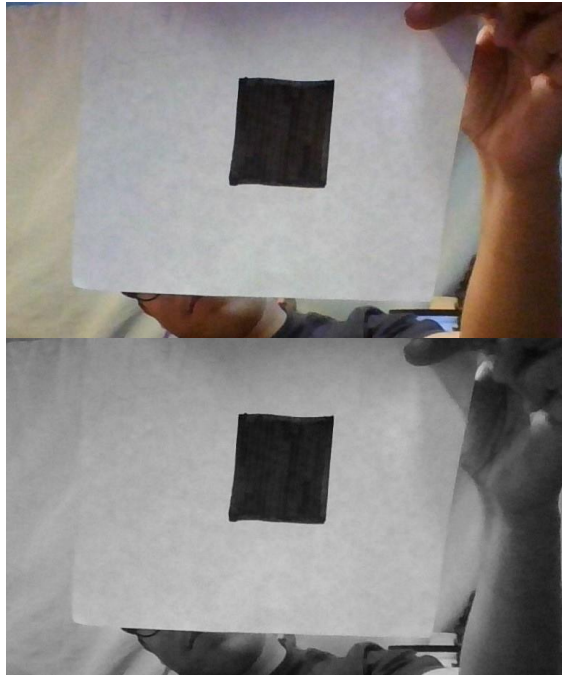


Fig. 6. Imagen a color y en escala de grises.

D. Imagen a color negativo

Posterior a la escala de grises se hace un proceso a escala negativa, ya que el color negro representa huecos en la imagen y las figuras hechas a mano son pintadas con pulmones oscuros.

Ya que se trabaja con imágenes de 8 bits por pixel la representación del color máximo de cada pixel es 255.

Formula del negativo

$$P_{xy} = 255 - P_{xy} \quad (2)$$



Fig. 7. Imagen en negativo.

E. Imagen con Método OTSU

Este método consiste devolver un umbral de intensidad único que separa los píxeles en dos clases, primer plano y fondo. Este umbral se determina minimizando la varianza de intensidad intercalase, o de manera equivalente, maximizando la varianza entre clases.[7]

En otras palabras separa la imagen principal del fondo, generando un umbral de color.

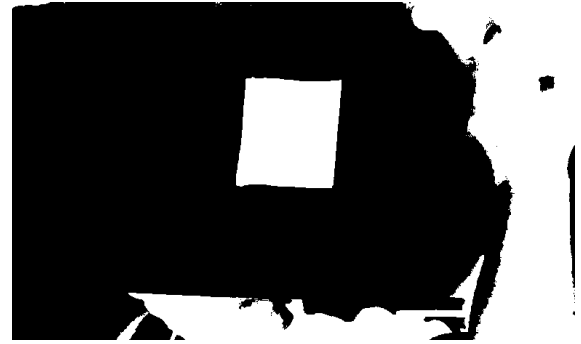


Fig. 8. Imagen aplicando método OTSU.

F. Extracción de los Componentes Conectados

Una vez realizado el etiquetado se procede a extraer los componentes por si solos de la imagen original, para así probar si alguno de estos coincide con alguna figura geométrica esperada o solo son ruido en la imagen.[6]

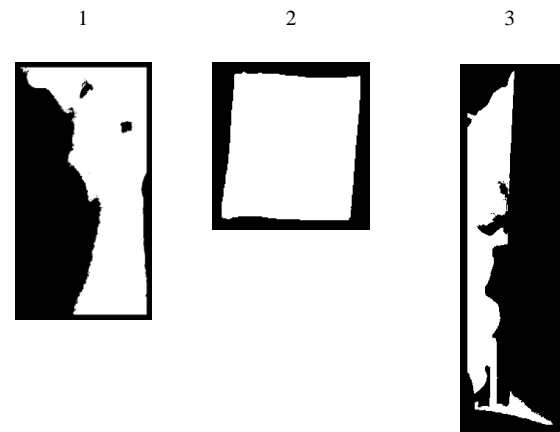


Fig. 9. Componentes conectados de la figura 6.

G. Calculo del Factor de Compacidad

El factor de compacidad está definido como la relación entre el área y el perímetro del componente[7].

Está representado por la siguiente formula.

$$f = \frac{Area}{Perimetro^2} \quad (3)$$

Tabla I
FACTORES DE COMPACIDAD

Figura	Área	Perímetro	Compacidad
1	58691	2178	0.012372451453
2	19390	601	0.053682021921
3	19356	1777	0.006129721708

Las los cálculos anteriores muestran las áreas, perímetros y factores de compacidad de los componentes conectados en la figura 8, donde el numero 2 representa los cálculos del cuadrado.

H. Calculo de Momentos Invariantes de Hu.

Una vez que el componente está separado se extraen los primeros cuatro momentos de Hu.[8]

Tabla II
PRIMEROS CUATRO MOMENTOS DE HU

Figura	M1	M2	M3	M4
1	0.3832	0.1006	473.17	127.34
2	0.1686	0.0004	0.0131	0.0009
3	0.9073	0.7529	1302.6	1034.7

IV. ENTRENAMIENTO

A. Entradas y Salidas

Como se mencionó anteriormente cada entrada está compuesta del factor de compacidad y los primeros cuatro momentos invariantes de Hu.

El set de entrada se compone de 150 patrones utilizados para el entrenamiento incluyendo basura o ruido de fondo.

Los pesos son establecidos en random de 0.0 a 1.0 dependiendo de la cantidad de entradas, posteriormente estos se van ajustando por cada época.

A continuación se muestran las salidas por clase

Tabla III
CLASE RESPECTO A SALIDA ESPERADA

Clase	Salida
Círculo	0,0,1
Cuadro	0,1,0
Pentágono	1,0,0
Triangulo	0,1,1
Basura	1,1,1

B. Proceso de Entrenamiento

Las siguientes pruebas de entrenamiento fueron realizadas de la siguiente forma.

Tabla IV
ENTRENAMIENTO DE PATRONES POR CLASE

Patrones	Validación	Error Mínimo
10	20	0.0047891954
20	10	0.0049998932
30	30	0.0049897947

C. Pruebas con 10 formas por Clase

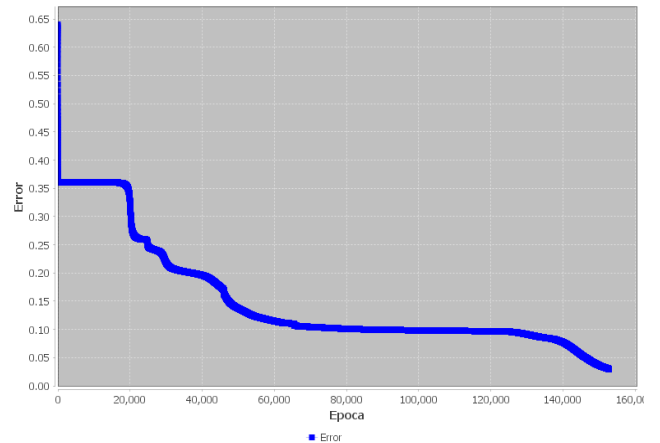


Fig. 10. Grafica de error en prueba 1.

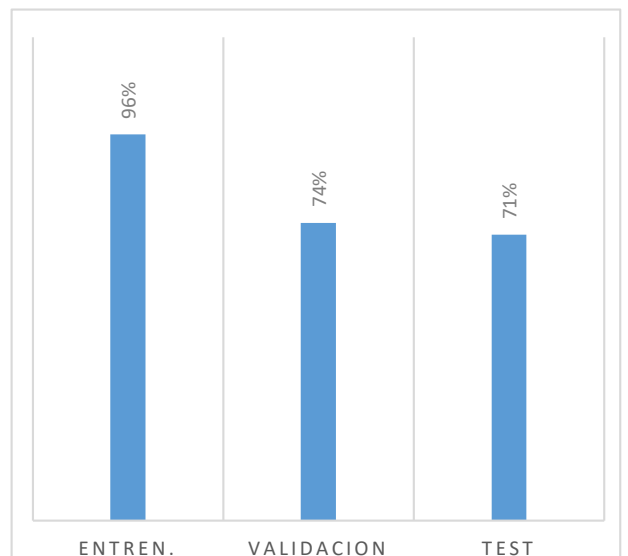


Fig. 11. Grafica de prueba 1 de 10 por clase.

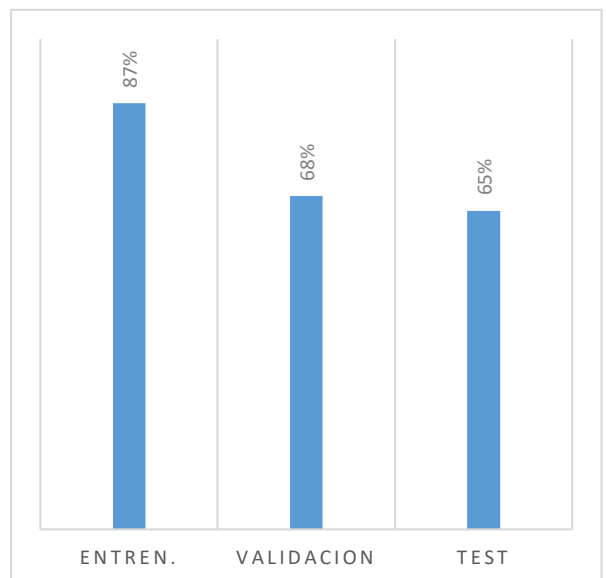


Fig. 12. Grafica de prueba 2 de 10 por clase.

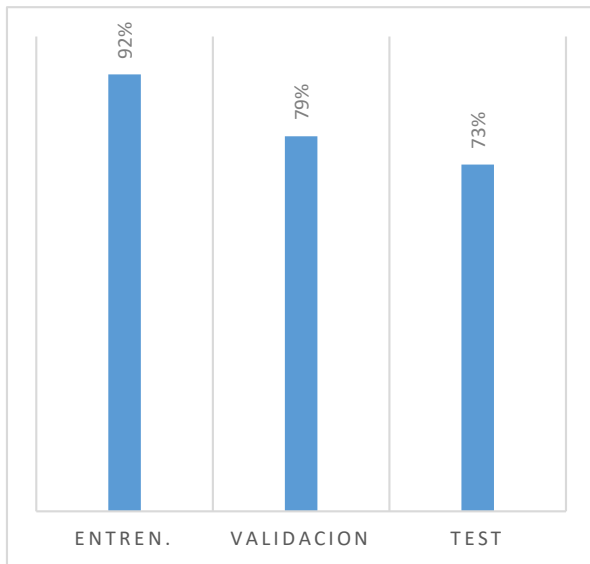


Fig. 13. Grafica de prueba 3 de 10 por clase.

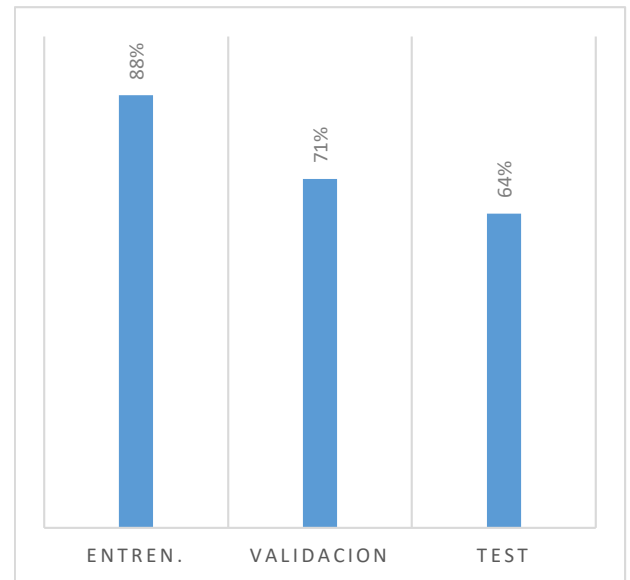


Fig. 16. Grafica de prueba 6 de 10 por clase.

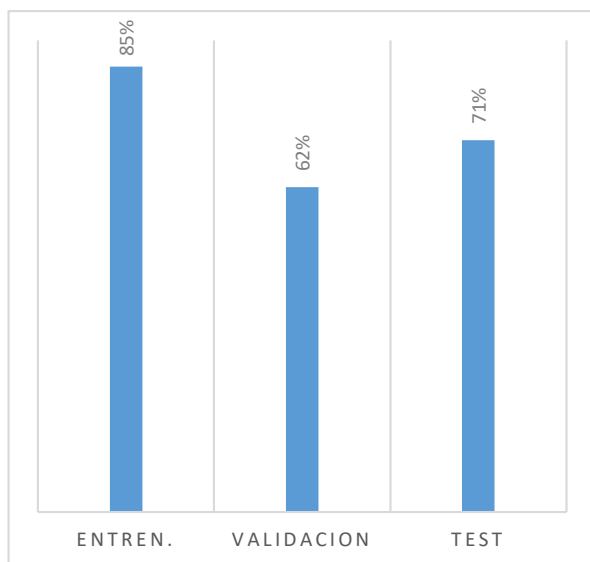


Fig. 14. Grafica de prueba 4 de 10 por clase.

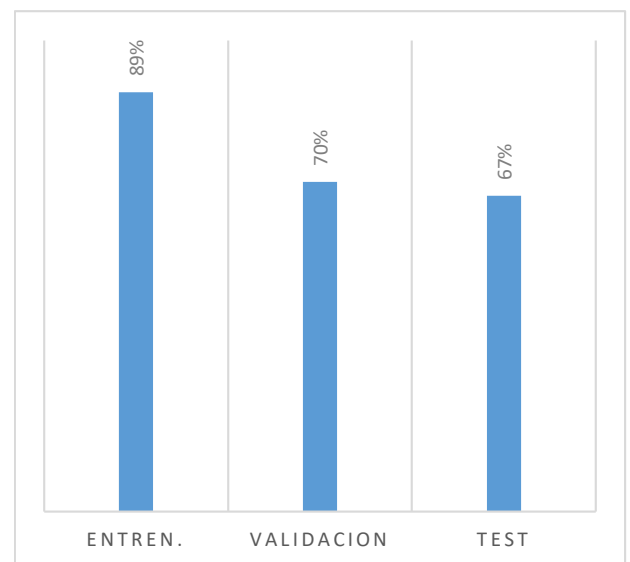


Fig. 17. Grafica de prueba 7 de 10 por clase.

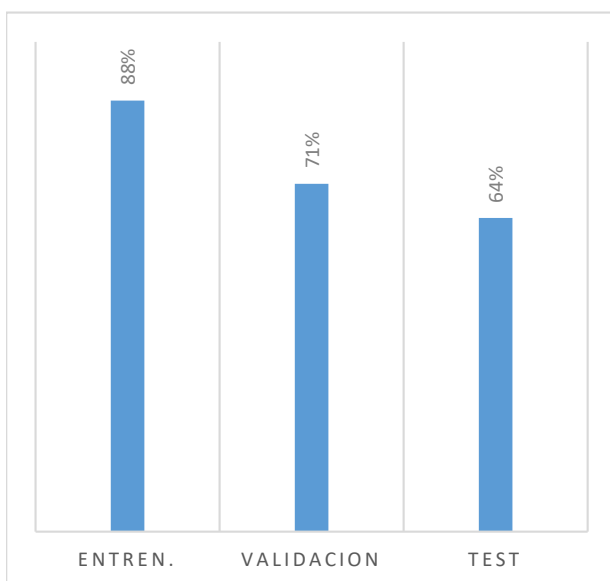


Fig. 15. Grafica de prueba 5 de 10 por clase.

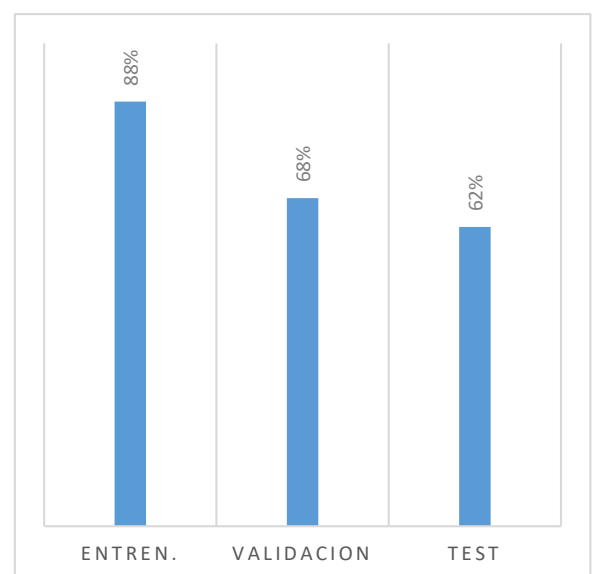


Fig. 18. Grafica de prueba 8 de 10 por clase.

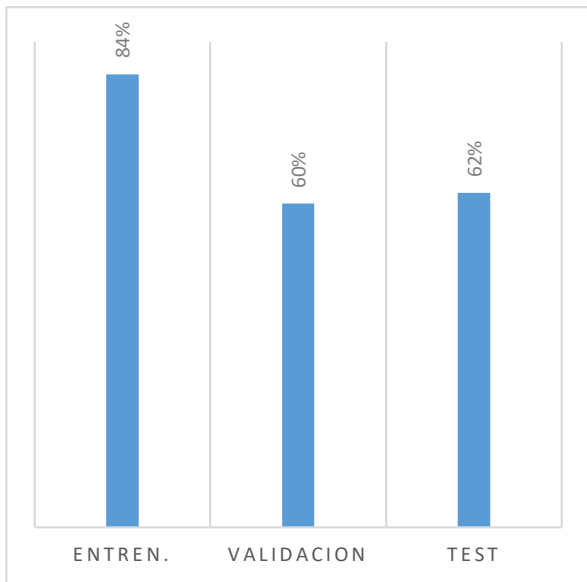


Fig. 19. Grafica de prueba 9 de 10 por clase.

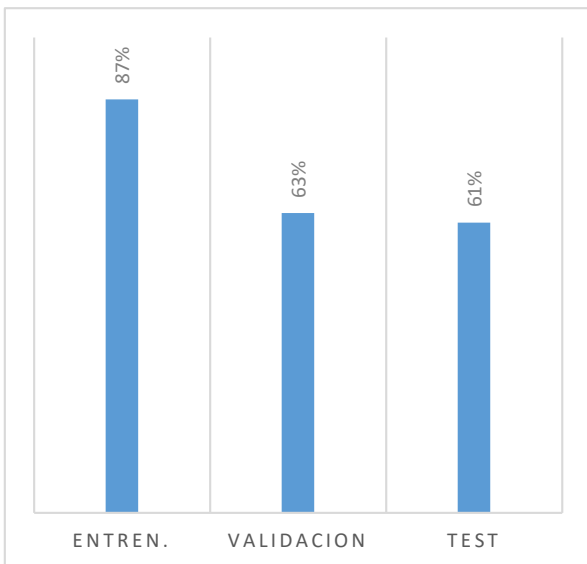


Fig. 20. Grafica de prueba 10 de 10 por clase.

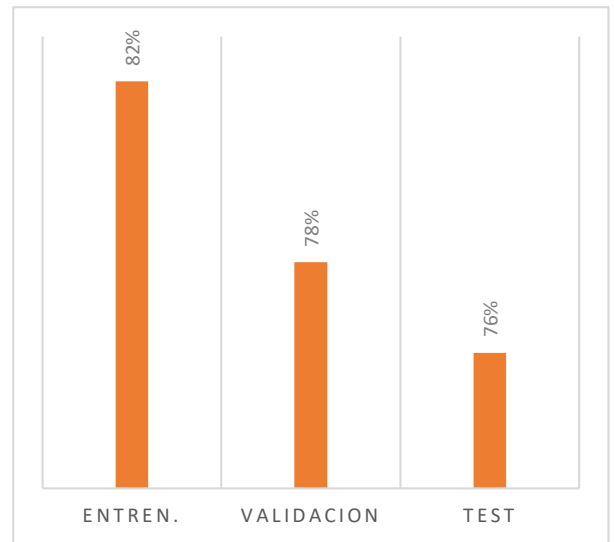


Fig. 22. Grafica de prueba 1 con 20 por clase.



Fig. 23. Grafica de prueba 2 con 20 por clase.

D. Pruebas con 20 formas por Clase

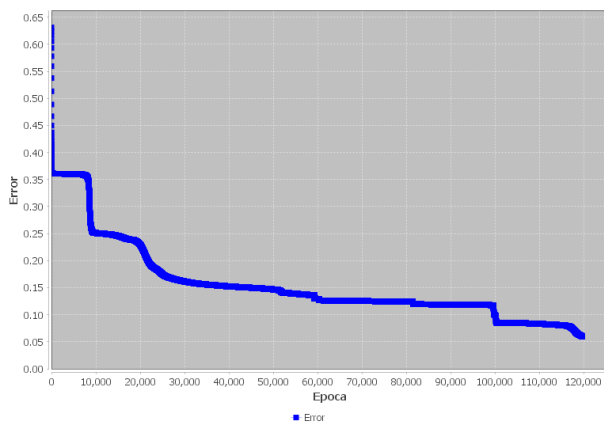


Fig. 21. Grafica de prueba 1 con 20 por clase.



Fig. 24. Grafica de prueba 3 con 20 por clase.



Fig. 25. Grafica de prueba 4 con 20 por clase.

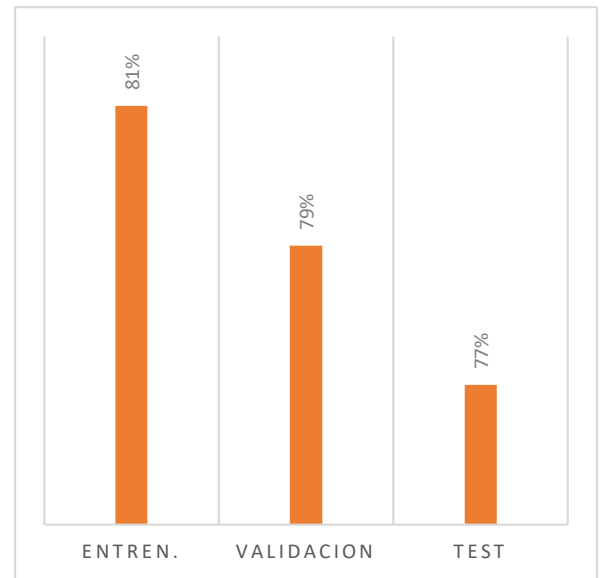


Fig. 28. Grafica de prueba 7 con 20 por clase.



Fig. 26. Grafica de prueba 5 con 20 por clase.

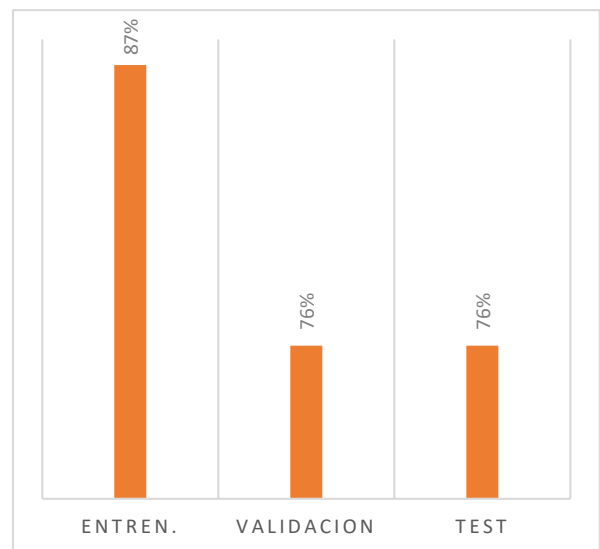


Fig. 29. Grafica de prueba 8 con 20 por clase.



Fig. 27. Grafica de prueba 6 con 20 por clase.

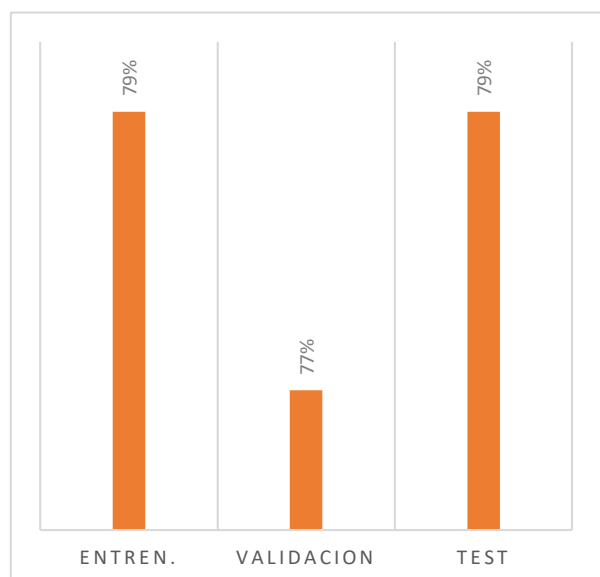


Fig. 30. Grafica de prueba 9 con 20 por clase.

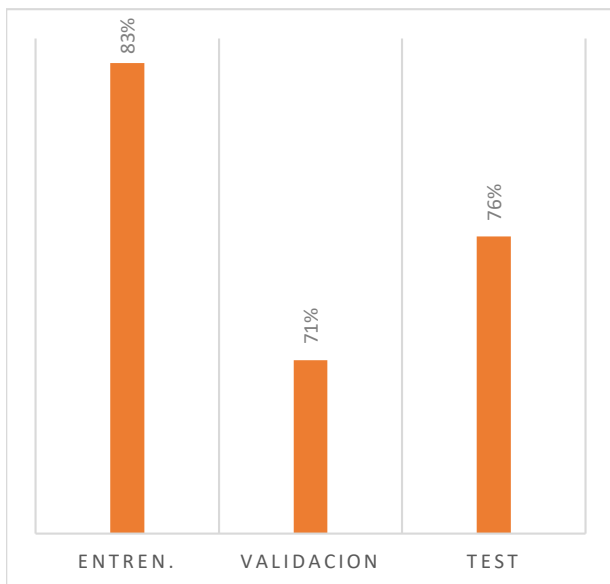


Fig. 31. Grafica de prueba 10 con 20 por clase.

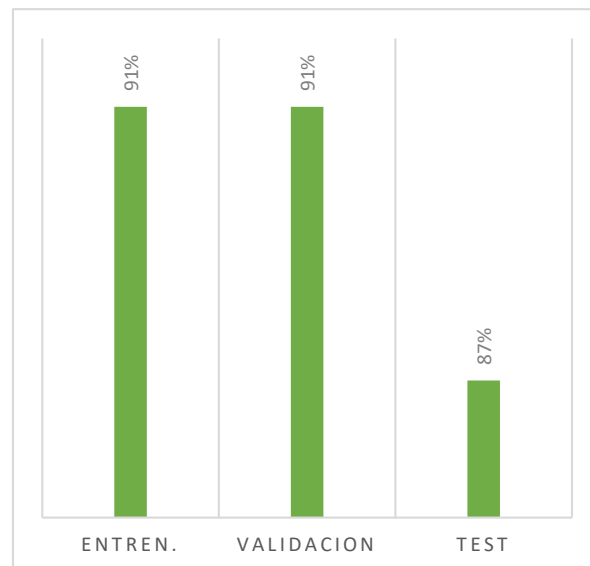


Fig. 34. Grafica de prueba 2 con 30 por clase.

E. Pruebas con 30 formas por Clase

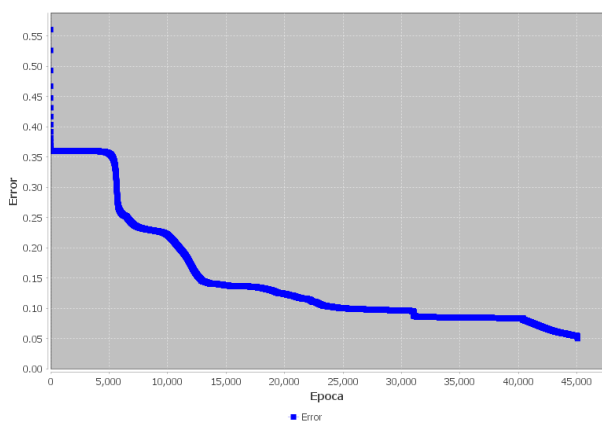


Fig. 32. Grafica de prueba 10 con 30 por clase.

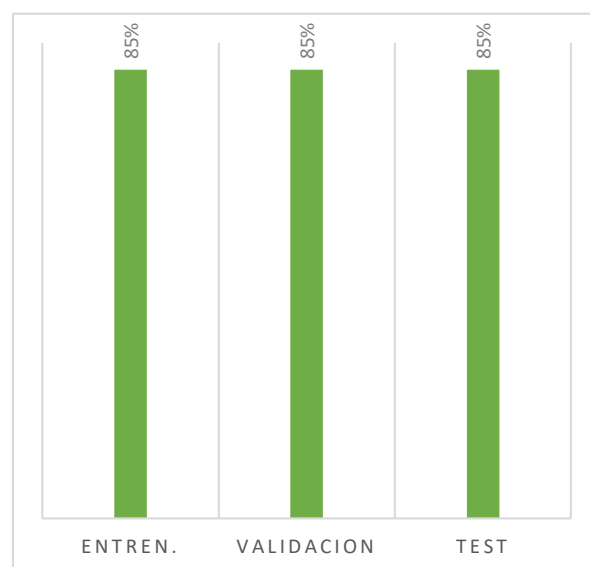


Fig. 35. Grafica de prueba 3 con 30 por clase.

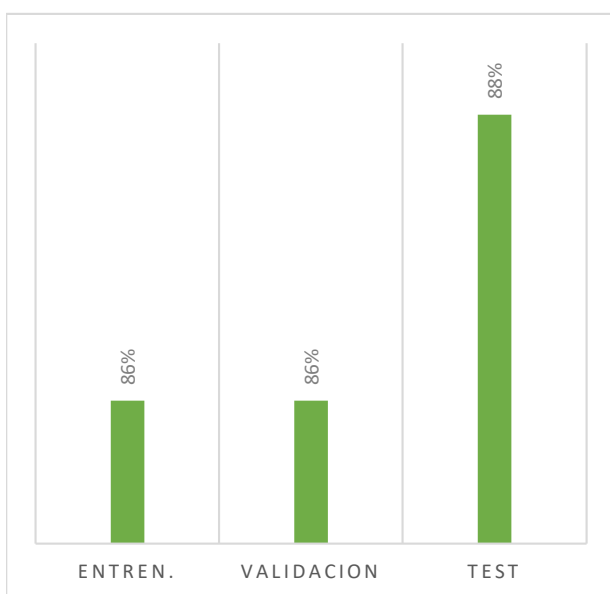


Fig. 33. Grafica de prueba 1 con 30 por clase.

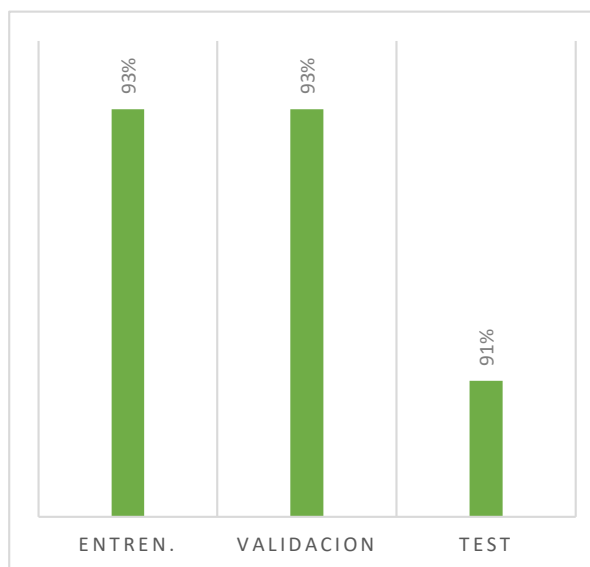


Fig. 36. Grafica de prueba 4 con 30 por clase.



Fig. 37. Grafica de prueba 5 con 30 por clase..



Fig. 40. Grafica de prueba 8 con 30 por clase.

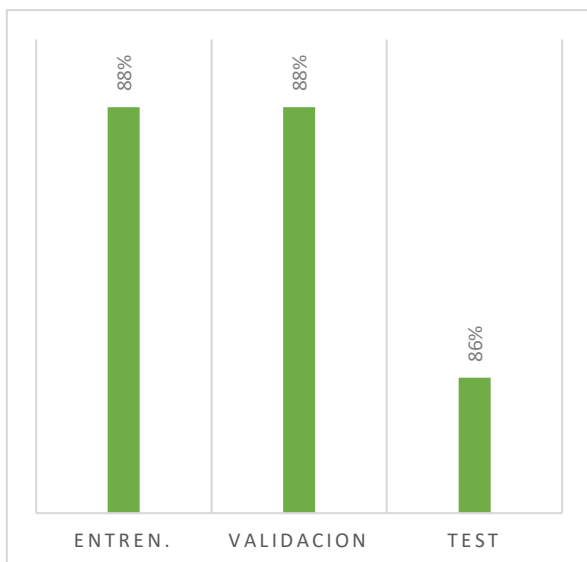


Fig. 38. Grafica de prueba 6 con 30 por clase.



Fig. 41. Grafica de prueba 9 con 30 por clase.



Fig. 39. Grafica de prueba 7 con 30 por clase.

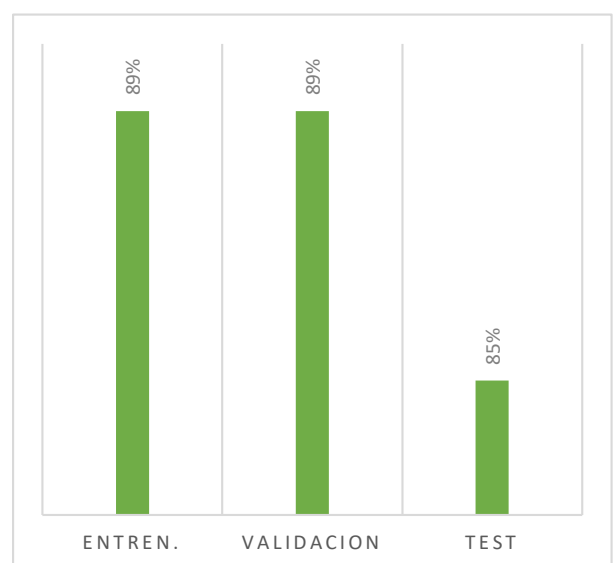


Fig. 42. Grafica de prueba 10 con 30 por clase.

V. CONCLUSIONES

Como resultado de este trabajo, podemos concluir que las Redes Neuronales Artificiales son un camino certero para el reconocimiento de figuras geométricas a través de los momentos de Hu.

Cabe mencionar que no es el único o el más efectivo, pero con ayuda de las librerías que ofrece OpenIMAJ es uno de los más sencillos de realizar en cuanto al procesamiento de imágenes.

Asimismo se hace hincapié en el entrenamiento, este debe ser cuidadosamente preparado y ejecutado, pues en las pruebas de entrenamiento arrojaban 93% de efectividad en el mejor de los casos pero al comprobar este porcentaje con figuras en tiempo real arrojó un 91% de aciertos, asimismo uno de los mayores problemas que se pudieron notar fue el reconocimiento de figuras en el fondo de la imagen.

Este problema se presenta en casi todas las pruebas en tiempo real, dicho detalle puede mejorarse con un mejor entrenamiento de la red, es decir, agregar más ruido de fondo para la clase “basura” para así poder tener un mejor umbral de diferencia entre lo que es una figura geométrica y el ruido de fondo.

REFERENCIAS

- [1] Sánchez E, Alanis A. 2006. Redes Neuronales: Conceptos fundamentales y aplicaciones a control automático. Madrid. Prentice-Hall. 210..
- [2] Hubel H, Wiesel, Torsten (2005). Brain and visual perception: the story of a 25-year collaboration. US. Oxford University Press .
- [3] Matich J. (2001). Redes Neuronales: Conceptos Básicos y Aplicaciones. (PDF). Buenos Aires, Argentina. p. 6. Consultado el 26 de octubre de 2018.
- [4] Richard O. Duda, Peter E. Hart, David G. Stork. (2001). Clasificación de patrones. New York . Wiley .
- [5] University of Southampton. (2019). OpenIMAJ (Version 1.3.9) [Online]. Disponible: <http://www.openimaj.org>
- [6] Academic Press, 1990, Chap. 6.
- [7] Hunt K.(2010) The Art of Image Processing with Java, Massachusetts. CRC Press.
- [8] Gómez W. 2015. Reconocimiento de objetos en fotografías, Tamaulipas, Cinvestav.
- [9] Ballard D. (1982) Computer Vision, Prentice-Hall, Chap. 2.
- [10] Davies E. Machine Vision: Theory, Algorithms and Practicalities,