

**DESARROLLO DE UNA RED NEURONAL DE HOPFIELD PARA EL RECONOCIMIENTO DE
CARACTERES ESCRITOS A MANO**

ALEXANDER DE JESUS MAESTRE TORRES

**UNIVERSIDAD POPULAR DEL CESAR
FACULTAD DE INGENIERIAS Y TECNOLOGIAS
VALLEDUPAR CESAR
NOVIEMBRE 2023**

TABLA DE CONTENIDO

- 1. INTRODUCCIÓN**
- 2. ESTADO DEL ARTE**
- 3. PLANTEAMIENTO DEL PROBLEMA**
- 4. OBJETIVOS**
 - 4.1. Objetivo general
 - 4.2. Objetivo específicos
- 5. JUSTIFICACIÓN**
- 6. METODOLOGÍA**
- 7. EXPLICACIÓN DETALLADA DE LA REALIZACIÓN DEL PROYECTO**
- 8. FASE DE ANÁLISIS**
- 9. FASE DE DISEÑO**
- 10. FASE DE PRUEBA**
- 11. FASE DE IMPLEMENTACIÓN Y SIMULACIÓN**
- 12. DEMOSTRACIÓN**
- 13. INTERFACES DEL SISTEMA**
- 14. REFERENCIAS BIBLIOGRÁFICAS**

INTRODUCCIÓN

El reconocimiento de caracteres escritos a mano es un procedimiento que consiste en identificar un símbolo, ya sea bien un número o una letra, a partir de la imagen de cierta escritura por medio del uso de herramientas de inteligencia artificial. Actualmente existen múltiples aplicaciones para el procedimiento de reconocimiento de caracteres. Muchas de estas aplicaciones incluyen: la automatización de la redirección de cartas en el correo postal, la identificación de las matrículas de los coches en los radares o la digitalización de los apuntes escritos a mano.

En la era digital, la capacidad de las computadoras para comprender y procesar la escritura a mano es una habilidad fundamental con aplicaciones en campos que van desde el reconocimiento de texto en documentos hasta la mejora de la interfaz hombre-máquina. El reconocimiento de caracteres escritos a mano ha sido un desafío continuo en el campo de la inteligencia artificial y el procesamiento de lenguaje natural. Para abordar este desafío, se han desarrollado diversas técnicas, y una de las más notables es la red neuronal de Hopfield.

Nuestro proyecto consiste en desarrollar una red neuronal de Hopfield para la identificación de caracteres numéricos a partir de un conjunto de píxeles que representan estos caracteres escritos a mano. El desarrollo de este proyecto lo pretendemos realizar por medio de una metodología general compuesta de 3 fases para la identificación de dichos caracteres. Para ello, realizamos la adquisición previa de la imagen, el pre-procesamiento de la imagen, y el reconocimiento o clasificación de la imagen.

ESTADO DEL ARTE

Internacional

Título: Reconocimiento de dígitos escritos a mano mediante métodos de tratamiento de imagen y modelos de clasificación

Año:2015

Autor: Luis Miralles Pechuán, Dafne Rosso Pelayo, Jorge Brieva

Lugar: México

El ROC (Reconocimiento Óptico de Caracteres) es una línea de investigación dentro del procesamiento de imágenes para la que se han desarrollado muchas técnicas y metodologías. Su objetivo principal consiste en identificar un carácter a partir de una imagen digitalizada que se representa como un conjunto de píxeles. En este trabajo se realizó para el ROC un proceso iterativo que consta de cinco fases. Para ello aplicaron varios métodos de tratamiento de imagen, dos métodos de selección de variables y exploraron diversos métodos supervisados de aprendizaje automatizado. Entre los modelos de clasificación destacamos los de Deep Learning por su novedad y su enorme potencial.

Internacional

Título: Análisis cualitativo del modelo generalizado de redes neuronales de Hopfield

Año:2023

Autor: DENISSE DEL PILAR ITA RAMÓN

Lugar: Perú

En la presente tesis el problema radica en el análisis del comportamiento asintótico de las redes neuronales de Hopfield a partir del análisis cualitativo. Asimismo, para las redes en estudio establecemos condiciones que garantizan la existencia de equilibrios aislados, que provean límites superiores para el número total de equilibrios asintóticamente estables. Con esto pretendemos establecer resultados que nos permitan localizar asintóticamente todos los equilibrios y así determinar las propiedades de estabilidad local. J.J. Hopfield (1984) más conocido por su invención de la red neuronal asociativa que ahora actualmente lo conocemos como red tipo Hopfield, describe que la red neuronal artificial está dada por una ecuación diferencial matricial que permita que la red sea dinámica. La red neuronal de Hopfield está dada por: $\dot{x} = -Ax + Ty + I$ y $y = S(x)$ donde \dot{x} representa el cambio de estado de la red neuronal, $-Ax$ representa la polarización del estado de la red neuronal (estado de reposo), T es la interconexión neuronal, I consiste en el umbral de activación de la red neuronal teniendo como condición inicial a $y = S(x)$ que representa la función logística o llamada función sigmoide.

El resultado principal consiste en mostrar la generalización del modelo de Hopfield, el cual presenta un nuevo parámetro $H(x)$ que viene a ser la función amplificación que representa el estado de despolarización de la red neuronal llamado también potencial de acción que está dado por: $\dot{x} = -H(x)(-Tx + S(x) - I)$. En la presente investigación presentaremos dos teoremas que satisfacen el sistema de ecuaciones, los teoremas I y II, es decir la existencia de soluciones y la convergencia de sus puntos de equilibrio.

Nacional

Título: Algoritmo para la Lectura por medio de Reconocimiento Óptico de Caracteres (OCR) de Etiquetas Nutricionales y la Generación de un Tipo de Sellos Frontales

Año: 2020

Autor: Eliana Martínez Cortes, Fredy Andrés Araujo Mosquera

Lugar: Neiva

En Colombia no se emplean sellos frontales de advertencia, y las tablas nutricionales de los productos alimenticios procesados son difíciles de interpretar sin conocimientos específicos en nutrición. El reconocimiento óptico de caracteres (OCR) es un proceso orientado a la lectura digital de una imagen textual de la que se obtienen los diferentes símbolos y caracteres que pertenecen a un determinado alfabeto (ABBY, 2019). En este trabajo se propone un algoritmo para generar sellos frontales que sean pertinentes para Colombia, a partir de la información de las etiquetas nutricionales obtenida por medio de Motor OCR Tesseract.

Todos los algoritmos desarrollados en el proyecto fueron implementados en Python. La metodología implementada inicia con el pre-procesamiento de las imágenes de las tablas nutricionales, continuando con la detección y el reconocimiento de la misma. Se obtienen las regiones de interés (ROI), se extrae la información para los sellos y por último se generan los sellos frontales GDA y Octagonal. El algoritmo presentó una precisión de 49% para la realización de los sellos. Los errores más frecuentes son confundir la g de los gramos con el nueve y no reconocer la palabra de interés.

Nacional

Título: Identificación Automática de Cilindros de Almacenamiento de Gas Utilizando Redes Neuronales Tipo Hopfield

Año: 2012

Autor: Luis Carlos Maldonado, César Augusto Peña, Oscar Gualdrón

Lugar: Pamplona

Las compañías que se dedican a la fabricación, comercialización y mantenimiento de cilindros para el gas licuado de petróleo en Colombia, estampan en placas de acero y soldadas al producto un código serial único para ser identificado dentro del parque de cilindros del país. Actualmente el proceso de identificación es manual y se revisan aproximadamente 7000 cilindros al día en una sola fábrica. El objetivo principal de este documento es presentar un sistema de visión artificial que utiliza redes neuronales artificiales para reconocer dicho código. Este sistema se compone físicamente de un dispositivo portátil de ambiente controlado en iluminación y escena para la adquisición de las imágenes. Otro componente del sistema es el ajuste de la imagen. El ajuste se realiza a partir de filtros de mediana, binarización, etiqueta, y segmentación; este procesamiento permite obtener mayor información significativa y discriminación de la imagen. Por último, el componente inteligente para la identificación se realiza con redes neuronales artificiales tipo Hopfield y un algoritmo que verifica la evolución del reconocimiento de la imagen. La efectividad del sistema se reporta con los resultados experimentales obtenidos con base al error con una cantidad significativa de muestras.

PLANTEAMIENTO DEL PROBLEMA

El reconocimiento de caracteres escritos a mano es una tarea crucial en numerosos campos de aplicación, desde la digitalización de documentos hasta la autenticación de usuarios en dispositivos móviles. A pesar de los avances tecnológicos en el procesamiento de imágenes y la inteligencia artificial, el reconocimiento preciso y eficiente de caracteres escritos a mano sigue siendo un desafío significativo debido a una serie de obstáculos clave.

Uno de estos desafíos es como la escritura a mano varía ampliamente de una persona a otra, y como una sola persona puede variar su escritura en diferentes momentos. Esto crea una amplia gama de estilos de escritura que dificultan la creación de modelos robustos de reconocimiento. Los caracteres escritos a mano también pueden variar en tamaño y resolución, lo que agrega complejidad al proceso de reconocimiento, ya que los modelos deben ser capaces de lidiar con estas variaciones.

También podemos destacar que las imágenes de caracteres escritos a mano pueden contener ruido y distorsiones debido a factores como la calidad del escaneo, la inclinación de la escritura y la calidad del papel. Estas imperfecciones pueden dificultar la identificación precisa de caracteres.

El reconocimiento de caracteres escritos a mano debe ser capaz de identificar una amplia variedad de caracteres, incluyendo letras, números, símbolos y caracteres especiales, lo que aumenta la complejidad del problema. En muchas aplicaciones, como sistemas de autenticación en tiempo real, es fundamental lograr un reconocimiento rápido y eficiente.

El principal propósito de nuestro proyecto es abordar algunos de estos desafíos a través del desarrollo de una red neuronal de Hopfield para el reconocimiento de caracteres escritos a mano. La red neuronal de Hopfield es una herramienta prometedora debido a su capacidad para aprender y generalizar patrones a partir de ejemplos. Sin embargo, todavía existe la necesidad de investigar cómo adaptar y optimizar esta red para abordar de manera efectiva los desafíos específicos relacionados con el reconocimiento de caracteres escritos a mano.

Para concluir, el problema radica en la necesidad de desarrollar un sistema de reconocimiento de caracteres escritos a mano que sea preciso, eficiente y capaz de lidiar con la variabilidad inherente en la escritura a mano. Este proyecto busca ofrecer una solución a esta problemática mediante el desarrollo y la optimización de una red neuronal de Hopfield para la identificación de caracteres escritos a mano.

OBJETIVOS

Objetivo general

Desarrollar una red neuronal de Hopfield para el reconocimiento de caracteres escritos a mano.

Objetivos específicos

- Recopilar un conjunto de datos adecuado de caracteres escritos a mano que incluya una variedad de estilos de escritura y caracteres.
- Investigar técnicas de pre-procesamiento de datos para mejorar la calidad de las imágenes de caracteres escritos a mano.
- Realizar el proceso de entrenamiento de la red neuronal de Hopfield utilizando el conjunto de datos preparado.
- Evaluar la precisión y el rendimiento de la red neuronal de Hopfield en la identificación de caracteres escritos a mano.
- Desarrollar una red neuronal de Hopfield que sea capaz de identificar patrones de caracteres escritos a mano.

JUSTIFICACIÓN

El reconocimiento de caracteres escritos a mano cumple un papel esencial en una variedad de aplicaciones tecnológicas y comerciales en la actualidad. La importancia del desarrollo de una red neuronal de Hopfield para esta tarea implica un amplio campo de posibilidades en la actualidad.

El reconocimiento de caracteres escritos a mano se aplica en una variedad de campos, desde el procesamiento de documentos y la digitalización de archivos históricos hasta sistemas de autenticación biométrica en dispositivos móviles. El desarrollo de un método preciso y eficiente es esencial para mejorar la eficiencia en estas aplicaciones.

También podemos destacar que la escritura a mano humana es altamente variable, y un solo carácter puede ser representado de muchas formas diferentes. Esto exige un enfoque de reconocimiento que sea capaz de adaptarse a la diversidad en la escritura de diferentes individuos.

La red neuronal de Hopfield se basa en principios biológicos, lo que la hace especialmente interesante para abordar el reconocimiento de caracteres escritos a mano como lo que hemos planteado, ya que se asemeja al funcionamiento de la memoria humana y la capacidad de recuperar patrones a partir de información parcial o ruidosa.

La capacidad de la red neuronal de Hopfield para aprender y generalizar patrones a partir de ejemplos es una ventaja significativa en el reconocimiento de caracteres escritos a mano, donde la variabilidad y la adaptación a diferentes estilos de escritura son fundamentales. La optimización de la red de Hopfield para el reconocimiento de caracteres escritos a mano en tiempo real es crucial para aplicaciones en las que se requiere una respuesta rápida, como la autenticación de usuarios en dispositivos móviles.

Un sistema de reconocimiento de caracteres escritos a mano altamente preciso y eficiente puede tener un impacto significativo en la eficiencia de numerosos procesos, desde la digitalización de documentos históricos hasta la accesibilidad de tecnologías de asistencia para personas con discapacidades.

En conclusión, podemos destacar que el desarrollo de una red neuronal de Hopfield para la identificación de caracteres escritos a mano posee un potencial para abordar un problema tecnológico crucial, su capacidad para aprender y adaptarse a la variabilidad en la escritura y su aplicabilidad en una variedad de contextos prácticos. Este proyecto representa una oportunidad de innovar en el campo del reconocimiento de caracteres escritos a mano y contribuir al avance de la inteligencia artificial y el procesamiento de lenguaje natural.

METODOLOGÍA

Para la realización de este proyecto que consiste en efectuar la identificación de caracteres a partir de un conjunto de píxeles que representan caracteres escritos a mano, aplicaremos una metodología de cinco pasos propuesta en un artículo que identificamos en el estudio del arte, [fuente 1](#). En esta metodología se aplica la reducción de dimensionalidad, la extracción de características de la imagen original, la selección de variables y la evaluación de varios modelos supervisados de aprendizaje automatizado. También hay que mencionar que como en esta metodología no se presenta una fase donde se identifique una fuente de datos de manera explícita hemos decidido incorporar esta fase en el proyecto.

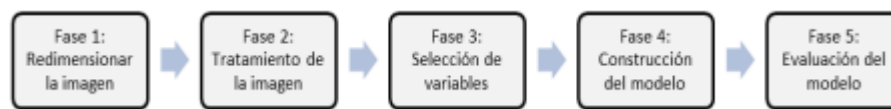


Fig. 1. Fases de la metodología. [Fuente 1](#).

Para la selección de variables se utilizan técnicas de PCA (Análisis de componentes principales) y RFE (Eliminación Recursiva de Características) que permiten crear modelos más eficientes para la identificación de caracteres. Con estas técnicas el tamaño del conjunto de datos decrece y el tiempo de clasificación de los caracteres se reduce. La metodología conformada por 5 fases (Figura 1) incorpora el tratamiento de imágenes, así como las fases correspondientes a la creación de modelos de aprendizaje automatizado.

Fase 1: Descripción de los datos

Normalmente, el proceso de reconocimiento óptico de caracteres inicia con la digitalización de las imágenes de escritura realizada a mano que contienen los dígitos, pero como hemos ya mencionado en este proyecto partimos de la necesidad de identificar o encontrar un conjunto de datos público que sea lo suficientemente robusto.

Este conjunto de datos será creado por nuestro propio criterio, utilizando la captura de imágenes de nuestra propia escritura que luego hemos separado en las muestras necesarias. Cada muestra tiene 1601 campos. Un campo que consideramos la salida del modelo representa el carácter y el resto de los campos son las entradas. Las entradas son los píxeles de la imagen que representa el dígito escrito a mano. Cada imagen tiene una resolución de 40x40, lo que hace un total de 1600 píxeles. Cada pixel se representa con un número entre el intervalo de 0 y 255 que indica el nivel de gris. De forma que el valor 0 representa el color blanco y el valor 255 el color negro.

Fase 2: Redimensionar la imagen

La resolución de una imagen se refiere a la cantidad de detalles que contiene. Cuantos más píxeles se utilicen para representar una imagen, mayor será su resolución. Cada imagen se compone de una cuadrícula de píxeles, y si la imagen es en color, cada píxel se describe mediante tres valores que representan los colores rojo, verde y azul (RGB). En nuestro caso, la imagen es en escala de grises, por lo que cada píxel tiene un valor entre 0 y 255.

Para que las imágenes se ajusten a los modelos, que tienen un número fijo de entradas, es necesario que todas las imágenes tengan el mismo tamaño. En algunos casos, esto implica reducir el tamaño de la imagen, lo que puede resultar en la pérdida de información, mientras que, en otros casos, se debe ampliar la imagen y rellenar los píxeles adicionales a través de un proceso de interpolación.

Una vez que hemos procesado la imagen, podemos extraer características de ella que son útiles para mejorar la precisión de las predicciones del modelo. Esto implica medir cosas como el promedio de los valores de los píxeles, contar cuántos píxeles tienen un valor por encima de cierto umbral o identificar regiones dentro de la imagen. También podemos aplicar algoritmos para detectar características como la textura o determinar cuántos agujeros hay en ciertos patrones. Estas características adicionales ayudan al modelo a comprender mejor la imagen y a realizar predicciones más precisas.

Fase 3: Tratamiento de la imagen

El tratamiento de imágenes es un conjunto de técnicas que se utiliza para mejorar la calidad de la imagen o para facilitar la búsqueda de información en ellas. Algunos ejemplos podrían ser eliminar el ruido, localizar bordes o suavizar la imagen.

Las técnicas más utilizadas y que podemos aplicar a nuestro problema y que aplicaremos en la fase dos son:

- **Binarización:** Consiste en convertir todos los píxeles de una imagen originalmente tengan varios valores a solamente dos tonos: blanco y negro. Para ello se establece un umbral por encima del cual convertiremos en negro y por debajo del cual será blanco. Esta operación es muy frecuente puesto que algunos algoritmos trabajan en esta escala. Las imágenes ocupan un tamaño en memoria muy pequeño puesto que cada píxel se puede representar con un píxel.
- **Fragmentación o segmentación:** Esta técnica consiste en seleccionar una parte de la imagen original. Para ello se desarrollan algoritmos que se basan en la detección de bordes que son pequeños cambios en la intensidad de los colores y otras propiedades. Existen muchos métodos genéricos de fragmentación, aunque para cada tipo de imagen se suele aplicar un algoritmo más específico. Por ejemplo, para detectar un cáncer de mama primeramente se aplica un filtro sobre las regiones sospechosas y posteriormente se aplica un algoritmo que evalúa ese conjunto de píxeles.

Fase 4: Selección de variables

Una vez completadas las fases anteriores utilizaremos PCA y RFE para seleccionar las características más importantes y descartar las menos relevantes. El PCA es una técnica estadística empleada en el descubrimiento de patrones en conjuntos de datos con una alta dimensionalidad. Consiste en la selección de componentes que otorgan una mayor aportación de información para el modelo.

En el caso de la clasificación de los dígitos manuscritos, las variables más importantes para generar los modelos de clasificación deben coincidir con los píxeles que reducen la incertidumbre de la clase. Esto es, las variables de los píxeles que tengan mayor información mutua. Al analizar este proceso se observa que algunas de las variables descartadas coinciden con los píxeles de los bordes de la imagen, esto es lógico debido a que los bordes generalmente no contienen información importante.

Fase 5: Construcción del modelo

En este proyecto, nos proponemos desarrollar una red neuronal de Hopfield para la identificación de caracteres escritos a mano. La red aprenderá a reconocer y distinguir caracteres individuales, lo que es esencial en aplicaciones como el procesamiento de formularios, digitalización de texto y sistemas de reconocimiento de escritura a mano en tabletas y dispositivos móviles.

La red neuronal de Hopfield, propuesta por John Hopfield en 1982, es un tipo de red neuronal recurrente que se destaca por su capacidad de almacenar y recuperar patrones. Su aplicación en el reconocimiento de caracteres escritos a mano ha demostrado ser prometedora debido a su capacidad para aprender y generalizar a partir de ejemplos de escritura manual. Esta red se basa en principios biológicos y utiliza conexiones ponderadas para representar y recuperar patrones de entrada.

Fase 6: Evaluación del modelo y muestra de resultados

La evaluación consiste en medir el grado de precisión de cada modelo generado. Esta fase informa si la dimensionalidad elegida es adecuada para la construcción del modelo. Para medir la calidad de los modelos utilizamos el accuracy de cada método.

La precisión (accuracy) es una métrica comúnmente utilizada en el contexto de problemas de clasificación en aprendizaje automático e inteligencia artificial. Representa la proporción de predicciones correctas que un modelo hace en relación con el total de predicciones realizadas.

En otras palabras, la precisión mide la fracción de ejemplos de un conjunto de datos que un modelo clasifica correctamente. Cuanto más alta sea la precisión, mejor es el rendimiento del modelo en términos de clasificación.

Explicación detallada de la realización del proyecto

El reconocimiento de caracteres numéricos escritos a mano es una tarea común en el campo de la visión por computadora y el procesamiento de lenguaje natural. Una red neuronal Hopfield es un tipo de red neuronal recurrente utilizada para problemas de asociación y recuperación de patrones. En este proyecto, desarrollaremos una red neuronal Hopfield para reconocer caracteres numéricos escritos a mano.

El proyecto consiste en implementar un sistema de reconocimiento de caracteres numéricos escritos a mano a partir de la implementación de una red neuronal de Hopfield. El funcionamiento del sistema se basa en que, a partir de una imagen de un carácter numérico escrito a mano, se pueda reconocer e identificar el dígito correspondiente a dicha imagen. Para ello, se deben obtener los segmentos de la imagen que se desee analizar, esto se realiza con la segmentación de la imagen capturada y procesada. Luego, se capturan los caracteres que componen cada segmento. Para poder reconocer los caracteres numéricos de las imágenes, primero se debe tener una gran cantidad de ejemplares de cada número escrito. Con esto se entrenaría correctamente la red neuronal porque estaría aprendiendo diferentes formas y estilos para escribir un mismo número.

Para el reconocimiento de los caracteres, se realiza el mismo procedimiento de segmentación de las imágenes y captura de los caracteres de cada segmento. Con los resultados obtenidos del entrenamiento, se procede a identificar el carácter. El producto final se planea implementar en Python.

Fases del Proyecto

- **Recopilación de Datos:** Adquirir un conjunto de datos que contenga imágenes de caracteres numéricos escritos a mano, como el conjunto de datos MNIST, que es ampliamente utilizado para este propósito. Pre procesar las imágenes, normalizándolas y escalándolas para que tengan un tamaño uniforme.
- **Creación de la Red Neuronal Hopfield:** Diseñar y programar una red neuronal Hopfield en un lenguaje de programación como Python. Puedes utilizar bibliotecas como NumPy para facilitar la implementación. La red neuronal Hopfield consta de una capa de neuronas, donde cada neurona representa un píxel en las imágenes de entrada.
- **Entrenamiento de la Red:** Utilizar el conjunto de datos de entrenamiento para entrenar la red neuronal Hopfield. Durante el entrenamiento, las imágenes de los caracteres numéricos se presentarán a la red, y la red aprenderá a asociar patrones con ellos.
- **Reconocimiento de Caracteres:** Utilizar el conjunto de datos de prueba para evaluar el rendimiento de la red. Presentar imágenes de caracteres numéricos escritos a mano a la red y observar cómo recupera los patrones asociados.

- **Evaluación y Métricas:** Calcular métricas de rendimiento, como la precisión, la tasa de error y la matriz de confusión, para evaluar la eficacia de la red en el reconocimiento de caracteres.
- **Ajuste y Optimización:** Si es necesario, ajustar parámetros de la red, como el número de neuronas o la tasa de aprendizaje, para mejorar el rendimiento. Experimentar con diferentes técnicas de pre procesamiento de datos para ver cómo afectan al rendimiento.
- **Interfaz de Usuario:** Desarrollar una interfaz de usuario que permita a los usuarios dibujar caracteres numéricos a mano y obtener la predicción de la red neuronal Hopfield.
- **Documentación y Presentación:** Documentar todo el proceso, incluyendo el código fuente, los resultados y las conclusiones. Preparar una presentación o informe que resuma el proyecto, los desafíos encontrados y las lecciones aprendidas.

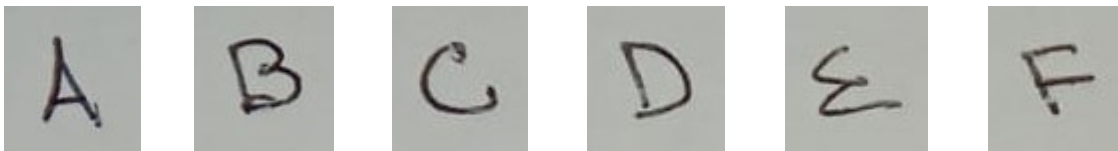
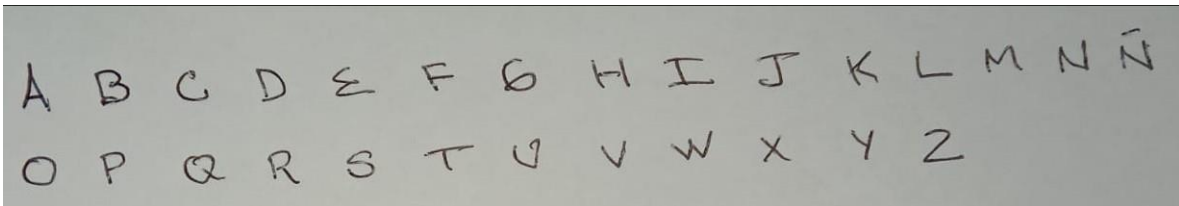
Para concluir, es importante resumir los resultados del proyecto y destacar los logros alcanzados. También es importante discutir posibles mejoras o extensiones futuras del proyecto.

Este proyecto proporcionará una comprensión básica de cómo desarrollar una red neuronal Hopfield para el reconocimiento de caracteres numéricos escritos a mano y cómo evaluar su rendimiento. Además, permitirá explorar áreas como el pre procesamiento de datos, la optimización de hiperparámetros y la interfaz de usuario, según las necesidades y objetivos específicos del proyecto.

Fase de análisis

La fase de análisis es fundamental para establecer una base sólida para el proyecto y garantizar que se cumplan los objetivos y requisitos. Una vez completada esta fase, se puede pasar a la implementación, el entrenamiento y la evaluación de la red neuronal. Destacando la importancia de la fase de análisis, se han identificado un conjunto de sub fases que nos ayudaran a llevar a cabo el desarrollo del proyecto.

- **Selección de Conjunto de Datos:** Identificar y seleccionar el conjunto de datos adecuado para el proyecto. En este caso, hemos creado nuestro propio conjunto de datos, utilizando la captura de imágenes de nuestra propia escritura que luego hemos separado en las muestras necesarias.



- **Exploración de Datos:** Realizar una exploración inicial de los datos para comprender su estructura y contenido. Esto incluye verificar el número de muestras, las dimensiones de las imágenes y la distribución de las etiquetas. Visualizar ejemplos de imágenes de caracteres para comprender la variabilidad de los datos y las posibles dificultades. En este caso se detectaron que las imágenes tienen dimensiones distintas, también se identificaron un total de 1117 muestras distintas, las cuales incluyen imágenes de letras en mayúscula, letras en minúscula y caracteres numéricos.

- **Pre procesamiento de Datos:** Identificar y llevar a cabo las tareas de pre procesamiento necesarias para preparar los datos para el entrenamiento de la red neuronal. Esto puede incluir la normalización de píxeles, la conversión de etiquetas a formato adecuado y la división del conjunto de datos en entrenamiento y prueba. En nuestro caso, primero redimensionamos cada una de las imágenes a 40 x 40 píxeles cada una. Luego transformamos las imágenes en una escala de grises, para luego binarizar las imágenes aplicando un umbral. Finalmente, vectorizamos la imagen y le asignamos una etiqueta. Para la división de los datos de prueba y entrenamiento, optamos por una división de 25% para los datos de prueba y un 75% para los datos de entrenamiento.



0000001111000010001111000110000110000110111111111	A
---------------------------------------------------	---

Fase de diseño

En esta fase, se debe definir la arquitectura de la red neuronal. Esto incluye la determinación del número de neuronas en la capa, la función de activación, y otros parámetros clave. Para nuestro caso hemos seleccionado una red neuronal Perceptrón Multicapa, que hemos diseñado a través de la librería `MLPClassifier` que nos provee Python, la cual hemos configurado de la siguiente manera:

```
mlp = MLPClassifier(random_state=1,  
                    learning_rate='invscaling',  
                    hidden_layer_sizes=(850, 850, 850),  
                    activation='logistic', max_iter=1000)
```

Número de capas ocultas	3
Número de neuronas de cada capa	850
Función de activación de las capas ocultas	Logistic
Error	0.0001
Número máximo de iteraciones	1000

- **MLPClassifier**: Este es el clasificador MLP en scikit-learn que se utiliza para construir y entrenar modelos de perceptrón multicapa.
- **random_state**: Este parámetro establece una semilla para la generación de números aleatorios. Esto es útil para reproducibilidad. Si usas la misma semilla, deberías obtener los mismos resultados cada vez que ejecutes el modelo.
- **learning_rate**: Este parámetro especifica la tasa de aprendizaje. En este caso, se utiliza 'invscaling', que indica una tasa de aprendizaje inversamente proporcional al tiempo.
- **hidden_layer_sizes**: Este parámetro define la arquitectura de la red, especificando el número de neuronas en cada capa oculta. En este caso, hay tres capas ocultas con 850 neuronas en cada una.
- **activation**: Este parámetro establece la función de activación para las capas ocultas. 'Logistic' se refiere a la función logística, también conocida como la función sigmoide.
- **max_iter**: Este parámetro establece el número máximo de iteraciones (épocas) durante el entrenamiento. Si el modelo no converge antes de alcanzar este número de iteraciones, el entrenamiento se detendrá.

Fase de prueba

La fase de prueba o evaluación consiste en medir el grado de precisión de nuestro modelo generado. Esta fase nos informa si la dimensionalidad elegida es adecuada para la construcción del modelo.

Selección de Métricas de Rendimiento: Identificar las métricas que se utilizarán para evaluar el rendimiento de la red. En el caso del reconocimiento de caracteres, las métricas comunes incluyen la precisión, la tasa de error y la matriz de confusión. Para nuestro caso hemos escogido la precisión, cuyo resultado fue:

Precisión	77%
-----------	-----

Esta cifra significa que la red ha logrado clasificar correctamente el 77% de las muestras en el conjunto de datos de prueba, un logro que destaca la capacidad del modelo para comprender y generalizar patrones en los datos.

El hecho de que la red haya alcanzado un nivel de acierto del 77% sugiere una eficacia en su capacidad para aprender y adaptarse a las complejidades de los datos de entrada. Sin embargo, para una evaluación más completa, es esencial contextualizar este resultado dentro de la naturaleza específica del problema que la red está abordando.

En términos de éxitos y desafíos, la precisión del 77% puede ser un indicador prometedor, pero es crucial considerar factores como el equilibrio de clases en el conjunto de datos y las posibles implicaciones prácticas de errores específicos, como falsos positivos o falsos negativos.

También destacamos la oportunidad de mejorar aún más el rendimiento del modelo. La experimentación con diferentes configuraciones de red, funciones de activación y tasas de aprendizaje podría ser el siguiente paso para aumentar la precisión. Además, estrategias como la ingeniería de características y la implementación de técnicas de regularización podrían contribuir significativamente a la robustez del modelo.

En conclusión, una precisión del 77% es un logro prometedor que refleja la capacidad de la red para aprender y generalizar patrones en datos complejos. Sin embargo, la mejora continua y la atención a los detalles específicos del problema son esenciales para llevar el rendimiento del modelo al siguiente nivel.

Fase de implementación y simulación

Las simulaciones realizadas para este proyecto incluyeron caracteres tanto de dígitos numéricos como de letras mayúsculas y minúsculas. En la Figura 1 se muestra la colección de imágenes utilizadas para el entrenamiento del sistema OCR. En cada prueba se entrenó al sistema OCR una sola vez por cada uno de los patrones de números o letras.

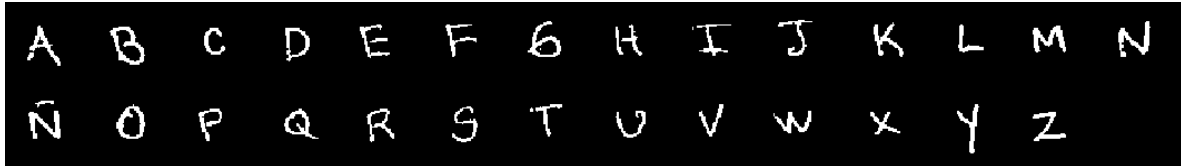


Fig. 1. Base de datos de imágenes utilizadas para el entrenamiento.

En las Figura 2 y 3 se muestran diversas pruebas realizadas en el sistema indicando, en cada caso. Para la primera prueba, fue dibujado a mano el carácter A. En este caso, el sistema realizó un reconocimiento exacto de la letra como se muestra en la Figura 2.

Letra original	Letra predicha
A	A

En la figura 3 se presenta el reconocimiento de una letra E dibujada de forma asimétrica para dificultar el reconocimiento. En el sistema OCR original no logra un reconocimiento con certeza, logra confundir la letra E con la letra K.

Letra original	Letra predicha
E	K

Conclusiones

Después de trabajar con Python y las librerías ya mencionadas, evaluamos este entorno como altamente adecuado para abordar problemas relacionados con la construcción y evaluación de modelos predictivos. Las ventajas clave de esta herramienta radican en su capacidad para crear modelos con diversas configuraciones, realizar una selección automática de las combinaciones más efectivas y presentar resultados en varias métricas.

Hemos observado que el redimensionamiento de la matriz para trabajar con imágenes más pequeñas ha conducido a ahorros significativos en tiempo y complejidad en la creación de modelos.

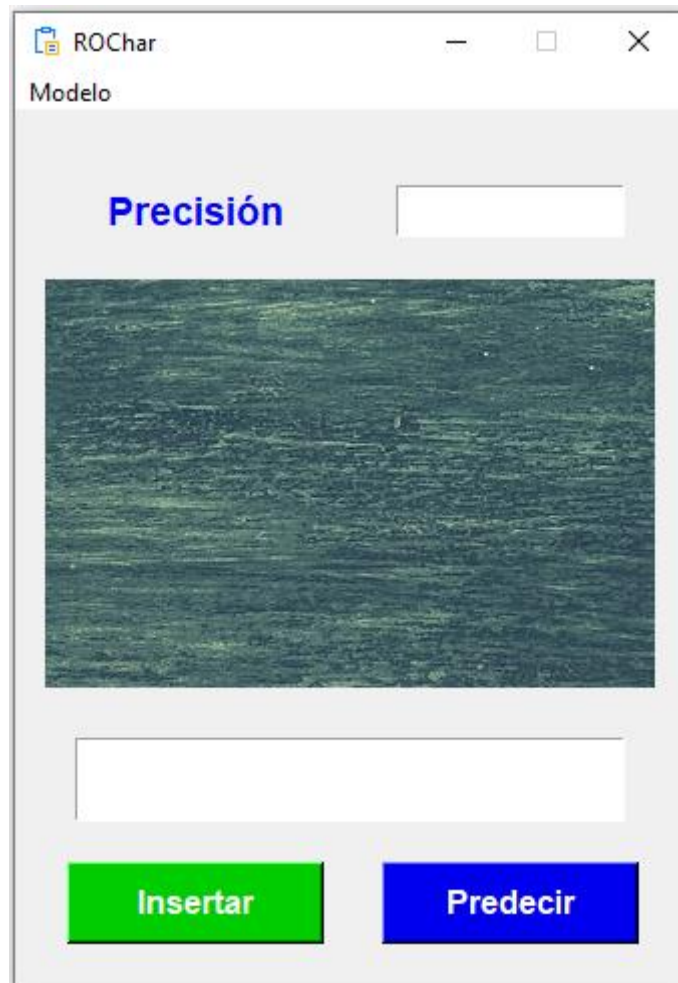
La modificación futura para el sistema OCR podría ser la unión de dos paradigmas que normalmente trabajan separados. Al agregar una red Hopfield para la reconstrucción de los datos de entrada, en este caso el carácter a reconocer, aumentaría el porcentaje de la eficacia del sistema en la etapa de reconocimiento.

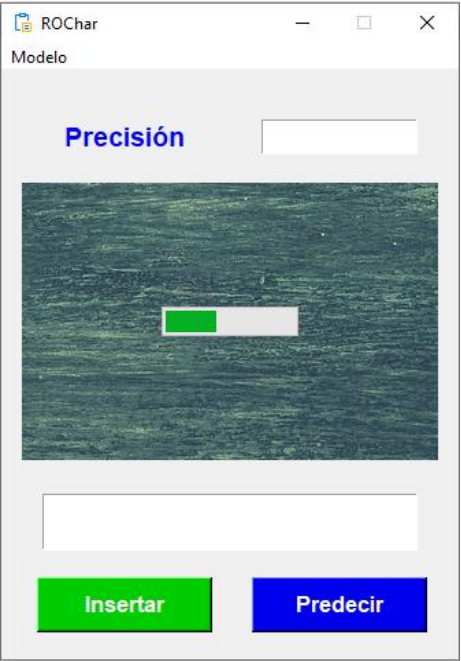
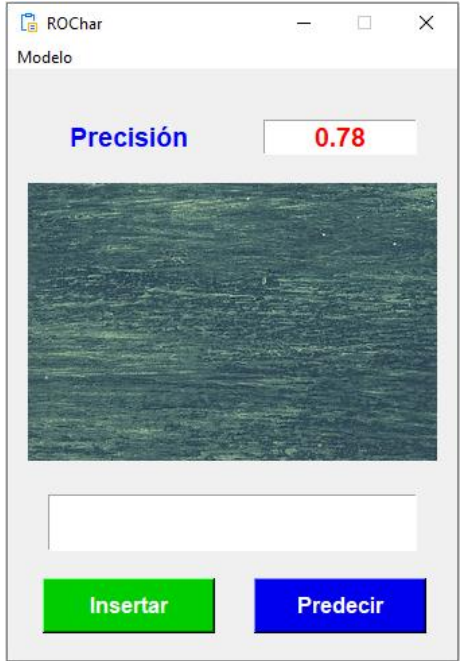
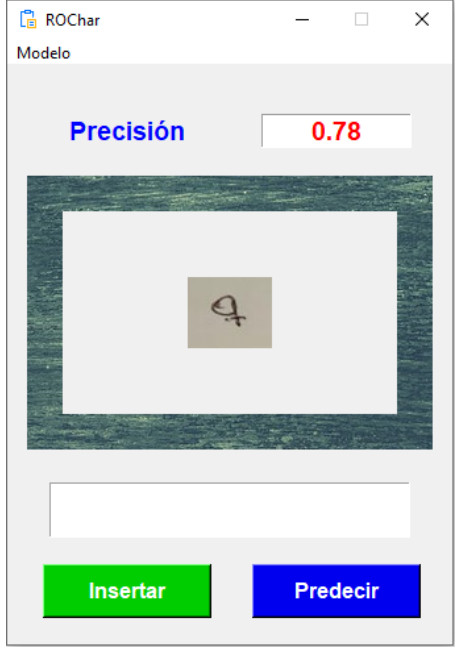
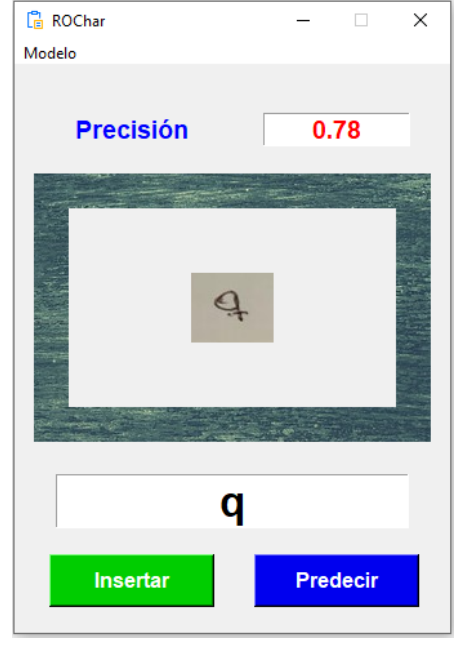
El sistema que hemos creado parte de una fase en la que se localizan los caracteres y se centran en una imagen. De forma que partimos de imágenes ya recortadas que facilitan en gran medida el reconocimiento. Pero en la vida real los caracteres no están aislados, sino que se encuentran dentro de textos. Una manera de segmentar este problema es descomponer los textos en pequeñas imágenes que reconozcan solamente un carácter.

Teniendo en cuenta los resultados obtenidos con los modelos de redes Perceptron Multicapa nos parece interesante seguir profundizando en este tipo de algoritmos. Pensamos que puede ser de interés general un estudio comparativo de estos modelos.

Interfaces del sistema

Nuestra aplicación está conformada por una interfaz simple que se compone de dos botones, dos campos de texto, una barra de menú y un panel para el posicionamiento de la imagen. En cuanto a los botones, uno de ellos es el botón Insertar, que es para insertar una imagen del carácter a predecir. El botón restante es el botón predecir, que es el responsable de ejecutar la predicción. Para las cajas de texto, la primera muestra la precisión del modelo y la segunda el resultado de la predicción. En cuanto a la barra de menú, está compuesta por dos botones uno para salir de la aplicación y otro para realizar el entrenamiento del modelo.



Entrenando el modelo	Modelo entrenado
	
Insertando un carácter	Realizando una predicción
	

REFERENCIAS BIBLIOGRÁFICAS

- [1] Miralles Pechuán, L., Rosso Pelayo, D., & Brieva, J. (2015). Reconocimiento de dígitos escritos a mano mediante métodos de tratamiento de imagen y modelos de clasificación. *Research in Computing Science*, 93(1), 83–94. <https://doi.org/10.13053/rcs-93-1-7>
- [2] Al Shalabi, H. M., Hasan, M. F., & Ali, A. M. (2005). A New Data Base Scheme Arabic Handwriting Recognition by Hopfield Neural Networks Algorithm. *Journal of Computer Science*, 1(2), 204–206. <https://doi.org/10.3844/jcssp.2005.204.206>
- [3] M. Butaev, M., Yu. Babich, M., Salnikovq, I. I., Martyshkin, A. I., Pashchenko, D. V., & Trokoz, D. A. (2020). Red neuronal para el reconocimiento de escritura a mano. *Nexo Revista Científica*, 33(02), 623–637. <https://doi.org/10.5377/nexo.v33i02.10798>
- [4] Martínez Cortés, Eliana Araujo Mosquera, Fredy Andres. (2020). Algoritmo para la Lectura por medio de Reconocimiento Óptico de Caracteres (OCR) de Etiquetas Nutricionales y la Generación de un Tipo de Sellos Frontales. <http://repositorio.uan.edu.co/handle/123456789/3154>