

# Integración de memristores en redes neuronales para la clasificación del dataset MNIST utilizando TensorFlow

1<sup>st</sup> Ricardo Aldair Tirado Torres  
*Centro de Investigación en Computación*  
*Instituto Politécnico Nacional*  
Ciudad de México, México  
rtiradot2023@cic.ipn.mx

2<sup>nd</sup> Ricardo Barrón Fernández  
*Centro de Investigación en Computación*  
*Instituto Politécnico Nacional*  
Ciudad de México, México  
rbarron@cic.ipn.mx

**Resumen**—El cómputo neuromórfico, con sus redes neuronales pulsantes, el aprendizaje STDP y los memristores, representa un paradigma revolucionario en la informática. La integración de estos conceptos en la tecnología CMOS no solo promete un salto significativo en la eficiencia y capacidad de las máquinas, sino que también nos acerca un paso más a replicar la asombrosa complejidad y eficiencia del cerebro humano. Esta tecnología abre un abanico de posibilidades en el campo de la inteligencia artificial y más allá, con el potencial de transformar la forma en que interactuamos con las máquinas y el mundo que nos rodea. En este trabajo se utiliza al nodo tecnológico de SKY130, un proceso de fabricación que integra los elementos necesarios para diseñar una red neuronal pulsante que integra el concepto del aprendizaje STDP no supervisado.

**Index Terms**—Redes neuronales pulsantes, sinapsis memristiva, plasticidad dependiente del tiempo de pulso (STDP), SKY130.

## I. INTRODUCCIÓN

Las redes neuronales artificiales (ANNs, por sus siglas en inglés) son modelos computacionales inspirados en la estructura y el funcionamiento del cerebro humano, diseñados para realizar tareas de reconocimiento de patrones, clasificación y predicción. Estas redes consisten en capas de neuronas artificiales conectadas entre sí, donde cada conexión tiene un peso asociado que se ajusta durante el proceso de entrenamiento. El entrenamiento de una ANN implica la optimización de estos pesos mediante algoritmos de aprendizaje, siendo el descenso de gradiente uno de los métodos más comunes. Una vez entrenada, la red puede realizar inferencias sobre datos nuevos, aplicando los pesos optimizados para generar predicciones o clasificaciones.

En el ámbito de la inteligencia artificial, las ANNs se utilizan extensamente para la clasificación de patrones, un proceso crucial en aplicaciones como el reconocimiento de imágenes, el procesamiento del lenguaje natural y la detección de fraudes. El aprendizaje profundo, una subrama del apren-

dizaje automático que se basa en ANNs con múltiples capas (redes neuronales profundas), ha revolucionado estos campos, logrando avances significativos en precisión y eficiencia.

Los memristores han emergido como una tecnología prometedora para implementar pesos sinápticos en redes neuronales artificiales. Un memristor es un componente electrónico cuya resistencia puede ser ajustada y memorizada, comportándose de manera análoga a una sinapsis biológica. En el contexto del aprendizaje máquina y el aprendizaje profundo, los memristores ofrecen una solución eficiente para el almacenamiento y ajuste de pesos sinápticos, permitiendo la creación de hardware neuromórfico que puede ejecutar tareas de inferencia y entrenamiento de manera más rápida y con menor consumo de energía en comparación con los sistemas tradicionales basados en transistores.

El memristor de HP, desarrollado por Hewlett-Packard, es uno de los primeros y más conocidos dispositivos de este tipo, capaz de recordar su estado resistivo incluso después de apagar la energía. Este atributo lo hace ideal para su uso en ANNs, donde puede representar y ajustar eficientemente los pesos sinápticos, facilitando así la implementación de redes neuronales en dispositivos de hardware especializados.

Para evaluar y entrenar redes neuronales, se utilizan conjuntos de datos estandarizados que permiten medir el rendimiento y la precisión del modelo. Uno de los conjuntos de datos más utilizados en la investigación de ANNs es el MNIST (Modified National Institute of Standards and Technology), que contiene 60,000 imágenes de entrenamiento y 10,000 imágenes de prueba de dígitos escritos a mano, cada una con una resolución de 28x28 píxeles. Este conjunto de datos ha sido fundamental para el desarrollo y benchmarking de nuevas arquitecturas de redes neuronales y técnicas de aprendizaje automático.

Este reporte explorará en detalle los conceptos mencionados, comenzando con una revisión de las redes neuronales artificiales, seguido de una discusión sobre el papel de los

memristores como pesos sinápticos, destacando el memristor de HP, y finalizando con un análisis del conjunto de datos MNIST y su relevancia en el campo de la inteligencia artificial, el aprendizaje máquina y el aprendizaje profundo.

A lo largo de este trabajo, discutimos la teoría subyacente de los memristores, detallamos la arquitectura de nuestra red neuronal y evaluamos el desempeño de nuestro modelo en la tarea de clasificación de imágenes. Los resultados obtenidos muestran que la integración de memristores en redes neuronales es una estrategia prometedora que puede contribuir significativamente al desarrollo de sistemas de IA más avanzados y eficientes.

## II. PRELIMINARES

### II-A. Redes neuronales artificiales

Las redes neuronales artificiales son modelos computacionales inspirados en la estructura y el funcionamiento del cerebro humano. Estas redes están compuestas por unidades interconectadas llamadas neuronas artificiales, distribuidas típicamente en tres tipos de capas:

- *Capa de entrada*: Recibe las señales iniciales del entorno externo. En el caso de la clasificación de imágenes, cada neurona de la capa de entrada representa un píxel de la imagen.
- *Capas ocultas*: Procesan las señales de entrada mediante una serie de transformaciones no lineales. Estas capas están compuestas por neuronas que aplican funciones de activación a sus entradas ponderadas.
- *Capa de salida*: Genera la salida final del modelo. Para la clasificación, cada neurona en la capa de salida puede representar una clase diferente.

Cada neurona en una red neuronal artificial realiza una operación matemática sobre sus entradas ponderadas y una función de activación. La salida de una neurona  $j$  se puede representar matemáticamente como:

$$y_j = \sigma \left( \sum_i w_{ij} x_i + b_i \right) \quad (1)$$

Donde:

- $y_j$  es la salida de la neurona  $j$ .
- $\sigma$  es la función de activación (por ejemplo, ReLU, sigmoide, tanh).
- $w_{ij}$  es el peso sináptico entre la neurona  $i$  de la capa anterior y la neurona  $j$ .
- $x_i$  es la entrada desde la neurona  $i$  de la capa anterior.

- $b_j$  es el sesgo de la neurona  $j$

El proceso de entrenamiento de una red neuronal implica ajustar los pesos sinápticos  $w_{ij}$  y los sesgos  $b_j$  para minimizar una función de pérdida, que mide la diferencia entre las predicciones de la red y los valores reales. Este ajuste se realiza mediante un algoritmo de optimización, como el gradiente descendente, junto con un proceso de retropropagación del error.

### II-B. Memristores

Los memristores (resistores de memoria) son componentes eléctricos pasivos que han sido teorizados durante décadas y que finalmente se desarrollaron experimentalmente en los últimos años. Estos dispositivos se caracterizan por su capacidad para recordar la cantidad de carga que ha pasado a través de ellos, lo que les permite retener un estado resistivo específico incluso después de que se ha eliminado la fuente de energía. Esta propiedad de memoria los hace especialmente adecuados para aplicaciones en almacenamiento de datos no volátil y procesamiento neuromórfico.

La idea de los memristores fue propuesta por primera vez por Leon Chua en 1971. Chua predijo la existencia de un cuarto elemento pasivo fundamental en los circuitos eléctricos, junto con el resistor, el capacitor y el inductor. Este cuarto elemento, el memristor, tiene una relación directa entre el flujo magnético y la carga eléctrica, lo que le confiere sus propiedades únicas. El memristor se define matemáticamente por la ecuación 2 y 3.

$$V = R(\omega) \cdot I \quad (2)$$

$$\frac{d\omega}{dt} = f(I) \quad (3)$$

Donde:

- $V$  es la tensión a través del memristor.
- $I$  es la corriente que pasa a través del memristor.
- $R(\omega)$  es la resistencia dependiente del estado  $\omega$ , que es una función de la carga histórica.

La capacidad de retener un estado resistivo específico sin necesidad de energía continua. Analogía con las sinapsis: Su comportamiento es similar al de las sinapsis neuronales, donde la resistencia del memristor puede ajustarse de manera análoga a la fuerza sináptica en las redes biológicas. Densidad y eficiencia energética: Ofrecen alta densidad de almacenamiento y baja energía de operación en comparación con los dispositivos convencionales.

La integración de memristores en redes neuronales artificiales ofrece una solución prometedora para mejorar la eficiencia energética y la densidad de almacenamiento de estos sistemas. Los memristores pueden emular las sinapsis biológicas al proporcionar una forma no volátil y analógica de almacenamiento y procesamiento de información. En la siguiente sección, describimos nuestro enfoque para integrar un modelo de memristor de HP en una red neuronal artificial para la clasificación del dataset MNIST, utilizando TensorFlow.

## REFERENCIAS

### III. SISTEMA PROPUESTO

#### III-A. Neurona LIF

Cuadro I: Parámetros  $\frac{W}{L}$  del sistema

Neurona		Sinapsis	
Parámetro	Valor	Parámetro	Valor
$W_1$	$0,2\mu m$	$W_1$	$5\mu m$
$W_2$	$0,2\mu m$	$W_2$	$5\mu m$
$W_3$	$1\mu m$	$W_3$	$1,5\mu m$
$W_4$	$0,2\mu m$	$W_4$	$15\mu m$
$W_5$	$0,2\mu m$	$W_5$	$1,5\mu m$
$W_6$	$0,2\mu m$	$W_6$	$15\mu m$
$W_7$	$0,2\mu m$		
$W_8$	$0,2\mu m$		
$W_9$	$1\mu m$		

\*  $L_{min} = 0,1\mu m$  en ambos circuitos.

### CONCLUSIÓN

Este trabajo presenta una implementación de neuronas LIF interconectadas entre sí, a través de un modelo de sinapsis 6T1R, cuyo componente principal era el memristor, el cual proporcionó diferentes pesos sinápticos de acuerdo con la regla de aprendizaje STDP. La construcción del sistema se hizo por medio de la herramienta de diseño esquemático Xschem y la simulación con NGSPICE. Cabe señalar que se utilizó la tecnología de SKY130, desarrollado por SkyWater Technology en colaboración con Google, que es un proceso de fabricación de semiconductores que ha sido puesto a disposición de la comunidad de diseño de circuitos de manera abierta y gratuita. Esto facilita la innovación y reduce significativamente los costos asociados con el diseño y la producción de chips, haciendo más accesible la experimentación y el desarrollo de nuevas tecnologías. Dentro de los elementos proporcionados por este nodo tecnológico, se encuentra el memristor, una tecnología clave que facilita la implementación de sinapsis artificiales en sistemas neuromórficos, soportando redes neuronales pulsantes y el aprendizaje STDP. Su capacidad para cambiar de estado de manera no volátil y su integración en la tecnología CMOS permiten la creación de sistemas de cómputo avanzados que no solo son más eficientes energéticamente, sino también más apto de emular las capacidades adaptativas y de aprendizaje del cerebro humano. La convergencia de estas tecnologías promete revolucionar el campo de la inteligencia artificial, llevando el rendimiento y la eficiencia del cómputo a nuevos niveles.