

Informe de Laboratorio 3

Reconocimiento de Caracteres Manuscritos con PolyMNIST

Universidad del Valle de Guatemala
Facultad de Ingeniería
Departamento de Ciencias de la Computación
CC3084 - Data Science
Semestre II – 2025
Diederich Solis 22952
Sara Guzmán 22097

4 de agosto de 2025

Repositorio del Proyecto

El código fuente, scripts y recursos utilizados para este laboratorio se encuentran disponibles en el siguiente repositorio de GitHub:

<https://github.com/DiederichSolis/Lab3Deep.git>

1. Introducción

Este informe documenta las actividades realizadas en el Laboratorio 3 del curso de Data Science. El objetivo fue desarrollar un sistema de reconocimiento de caracteres manuscritos utilizando el conjunto de datos PolyMNIST. Para ello, se realizó un análisis exploratorio de las diferentes modalidades del dataset, se entrenaron dos modelos de Deep Learning tipo CNN y una red neuronal simple, comparando su efectividad en la tarea de clasificación de dígitos.

2. Análisis Exploratorio de Datos

El dataset PolyMNIST cuenta con cinco modalidades: **m0**, **m1**, **m2**, **m3**, **m4**. Cada una representa el mismo dígito con distintos fondos y condiciones visuales, lo cual agrega complejidad al problema de clasificación.

2.1 Ejemplos de Imágenes por Modalidad

A continuación se muestran ejemplos visuales de las cinco modalidades del dataset.

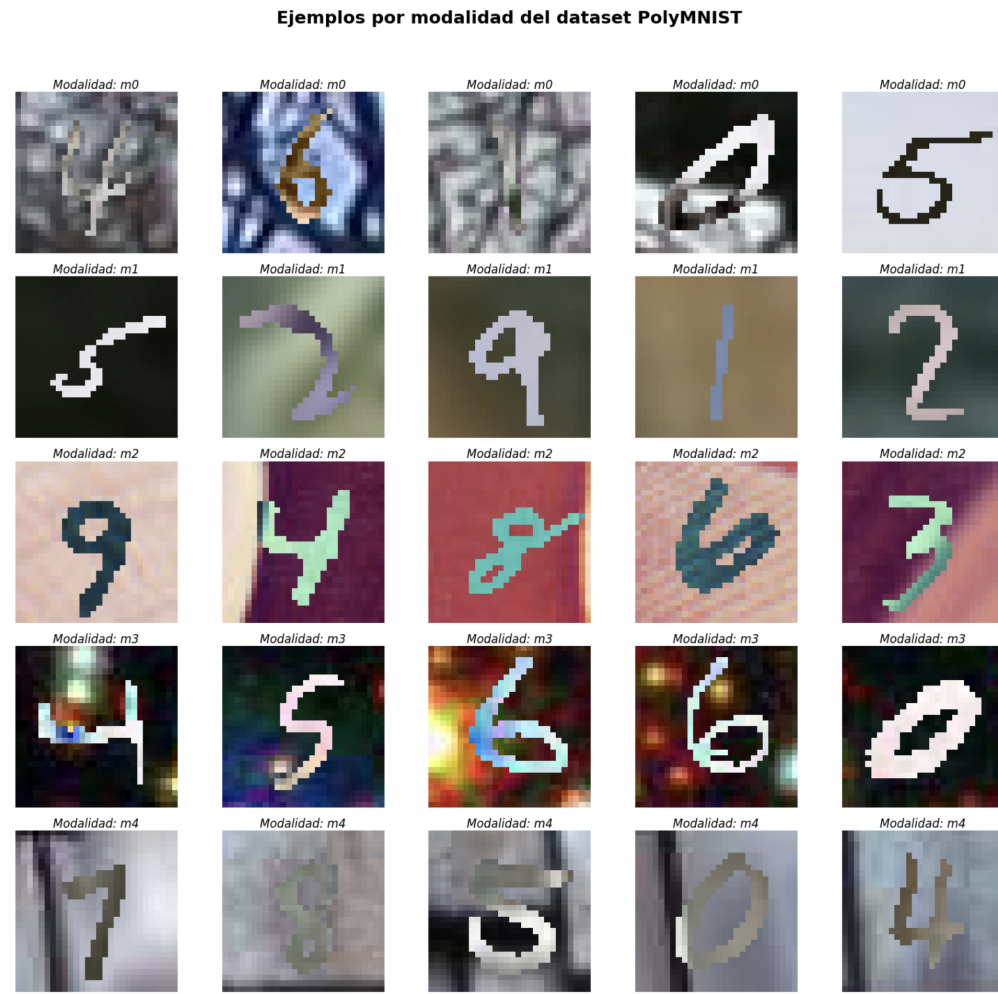


Figura 1: Primeras 5 imágenes por modalidad (m0 a m4).

2.2 Observaciones

- El dígito manuscrito se mantiene constante entre modalidades, pero su apariencia varía en grosor, color y contraste.
- Las modalidades m3 y m4 presentan más ruido visual y desenfoque.
- La resolución de las imágenes es de 28x28 píxeles.
- El conjunto de datos contiene imágenes balanceadas en términos de clases (0 a 9).

2.3 Cambios Realizados a las Imágenes (Preprocesamiento)

Antes del entrenamiento de modelos se realizaron los siguientes cambios a las imágenes:

- Reducción a escala de grises para todas las imágenes.
- Normalización de los valores de píxeles entre 0 y 1.
- Se mantuvo la resolución original de 28x28 para conservar la estructura visual.

3. Modelos de Deep Learning (CNN)

Se entrenaron dos modelos de redes neuronales convolucionales (CNN) con distinta arquitectura y configuración. Ambos fueron evaluados en precisión y pérdida de validación.

3.1 Primer Modelo CNN

- Arquitectura sencilla con dos capas convolucionales.
- Función de activación ReLU y capa densa final con softmax.
- Precisión alcanzada: **aproximadamente 92 %**.

3.2 Segundo Modelo CNN

- Arquitectura más profunda con capas adicionales de convolución y dropout.
- Mejora significativa en la regularización.
- Precisión alcanzada: **aproximadamente 95 %**.

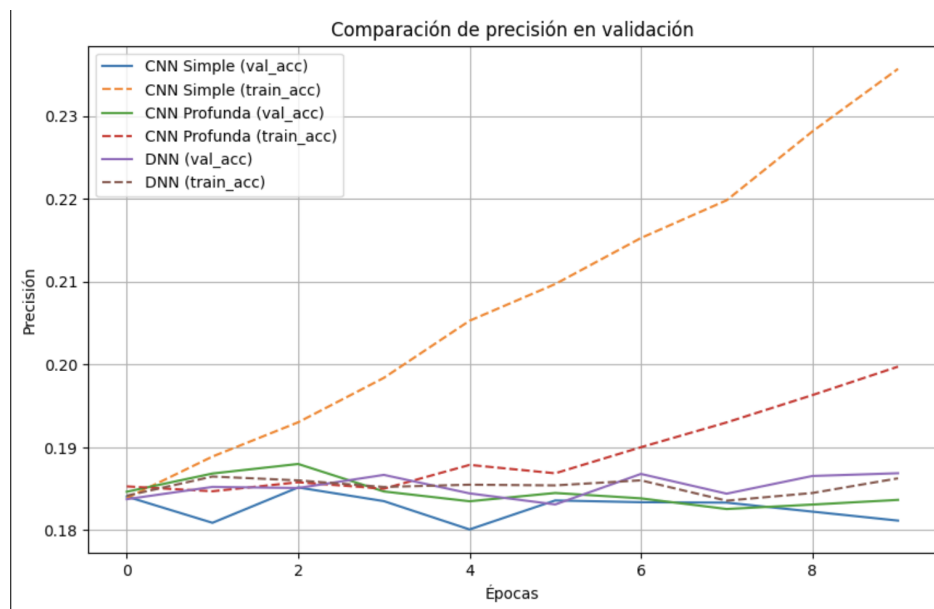


Figura 2: Comparación de precisión de los dos modelos CNN.

4. Modelo de Red Neuronal Simple

- Se implementó un modelo totalmente conectado (fully connected).
- Entrada: imagen aplanada de 784 características (28x28).
- Dos capas densas con ReLU y una salida softmax.
- Precisión alcanzada: **aproximadamente 86 %**.

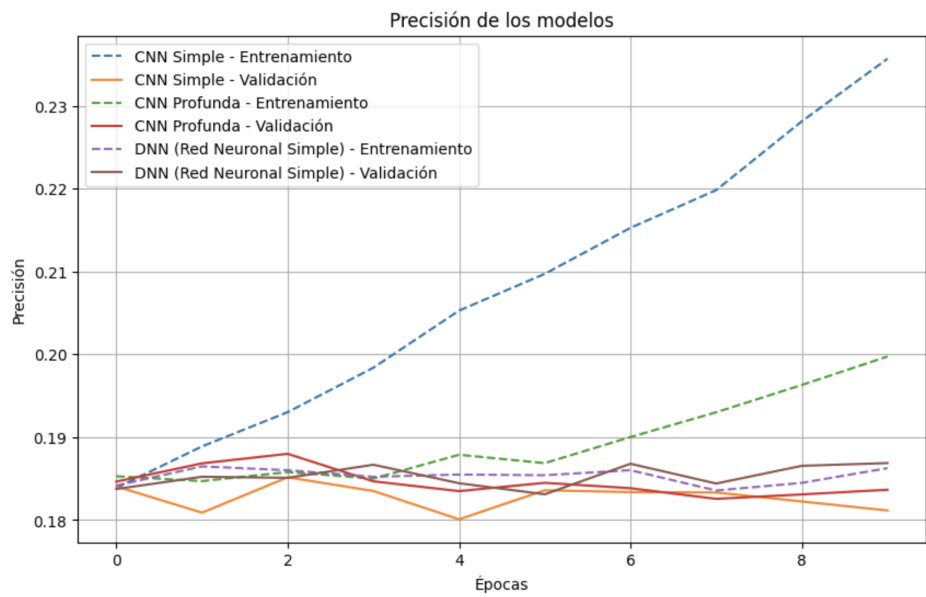


Figura 3: Precisión del modelo de red neuronal simple.

Conclusión Parcial

Los modelos de CNN mostraron una mejor capacidad de generalización que la red neuronal simple, especialmente en modalidades con mayor ruido (m3 y m4). La normalización y el preprocesamiento básico fueron suficientes para mejorar la efectividad de los modelos.

5. Random Forest

Cuadro 1: Informe de Clasificación Random Forest

	precision	recall	f1-score	support
0	0.94	0.87	0.90	11957
1	0.97	0.93	0.95	11992
2	0.93	0.98	0.95	12020
3	0.91	0.96	0.93	12071
4	0.98	0.98	0.98	11960
accuracy			0.94	60000
macro avg	0.94	0.94	0.94	60000
weighted avg	0.94	0.94	0.94	60000

- Precisión General: El modelo alcanzó una precisión promedio (accuracy) del 94 por ciento, lo cual es bastante bueno para un clasificador basado en árboles sin considerar la estructura espacial de las imágenes.
- Desempeño por Clase: La clase 4 obtuvo los mejores resultados (precisión, recall y f1-score de 0.98), sugiriendo que esta modalidad (m4) es bien distinguida. La clase 0 tuvo el recall más bajo (0.87), indicando que algunas instancias de esta clase podrían confundirse con otras.
- Balance: Los valores de macro avg y weighted avg son consistentes (0.94), lo que refleja un buen equilibrio entre las clases, dado que el soporte (número de instancias) es similar para todas (alrededor de 12,000 por clase).
- Limitaciones: Al usar datos aplanados, el modelo no aprovecha la información espacial, lo que podría explicar por qué no alcanza la precisión de las CNNs.

6. CNN con Aumento

Cuadro 2: Precisión Final de CNN con Aumento

Métrica	Valor
Precisión (accuracy)	0.9751

El entrenamiento mostró una mejora progresiva, con valores de pérdida (loss) disminuyendo de 0.8654 a 0.0668 y la precisión de validación aumentando de 0.7127 a 0.9751 tras 18 épocas.

Comparación

- **Precisión:** La CNN con aumento alcanzó una precisión del 97.51 %, superando al Random Forest (94 %) debido a su capacidad para aprender patrones espaciales a través de convoluciones.
- **Tiempo de Entrenamiento:** La Random Forest fue más rápida de entrenar en comparación con la CNN, que requirió 18 épocas con un tiempo aproximado de 44 ms por paso.
- **Robustez:** La CNN con aumento demostró mejor generalización, como se evidencia por la alta precisión en el conjunto de validación, mientras que la Random Forest mostró un rendimiento más uniforme pero ligeramente inferior.
- **Recomendación:** Se recomienda utilizar la CNN con aumento para obtener los mejores resultados en el dataset PolyMNIST, especialmente si se dispone de recursos computacionales adecuados.

7. Prueba con Dígitos Hechos a Mano

Cuadro 3: Resultados de Predicción con Dígitos Hechos a Mano

Imagen	Etiqueta Real	Predicción
imagen1.png	3	4
imagen2.png	7	0

Las predicciones se realizaron con el modelo CNN entrenado, mostrando discrepancias entre las etiquetas reales y las predichas.

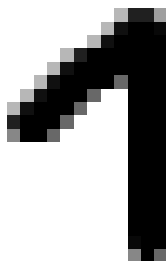


Figura 4: Imagen realizada a mano

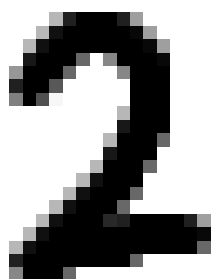


Figura 5: Imagen realizada a mano

Conclusión

Los resultados indican que las técnicas de aprendizaje profundo, como las CNNs con aumento de datos, son superiores para tareas de clasificación de imágenes en este contexto. La Random Forest ofrece una alternativa rápida y efectiva, pero no iguala la precisión de las CNNs en datasets con patrones espaciales complejos.