

Autores: Paula Barrillas, 22764 Diego Duarte, 22075	Docente: Luis Roberto Furlan Laboratorio 4
Sección: 10	Fecha: 11/08/2025

Laboratorio 4.

Clasificación de Imágenes con CNN (CIFAR-10)

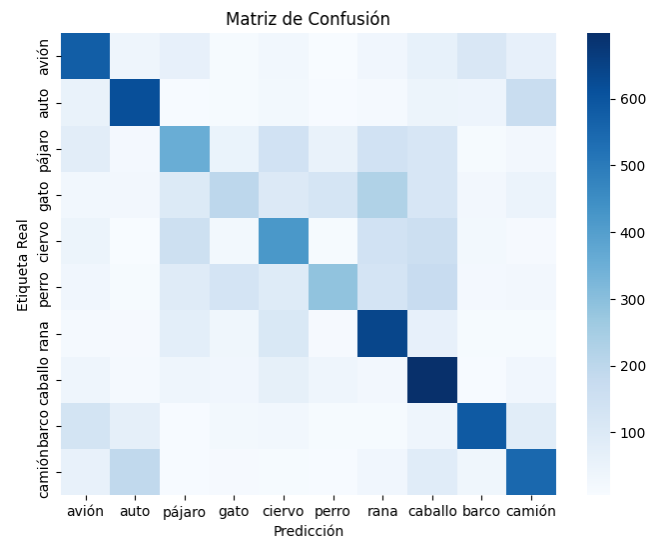
Repositorio Github: <https://github.com/paulabaal12/LAB4-DS>

ANN

El modelo tiene una arquitectura relativamente sencilla: una capa de aplanamiento para convertir la imagen en un vector, seguida de dos capas densas con 256 y 128 neuronas respectivamente (ambas con activación ReLU), y una capa de salida con 10 neuronas y activación softmax para la clasificación multiclase.

Exactitud Entrenamiento	Exactitud Validación	Tiempo Entrenamiento	Exactitud Prueba
0.57	0.49	302.44 s	0.49

Los resultados obtenidos muestran una exactitud en entrenamiento de 0.57 (57%) y una exactitud en validación de 0.49 (49%), mientras que en la prueba final se alcanzó también 0.49 (49%). Esto sugiere que el modelo logra aprender algunos patrones de los datos, pero aún presenta limitaciones en su capacidad de generalización, quedando por debajo de lo esperado para este tipo de tarea. A continuación, se presentan la matriz de confusión y ejemplos en situaciones donde le modelo se equivocó:

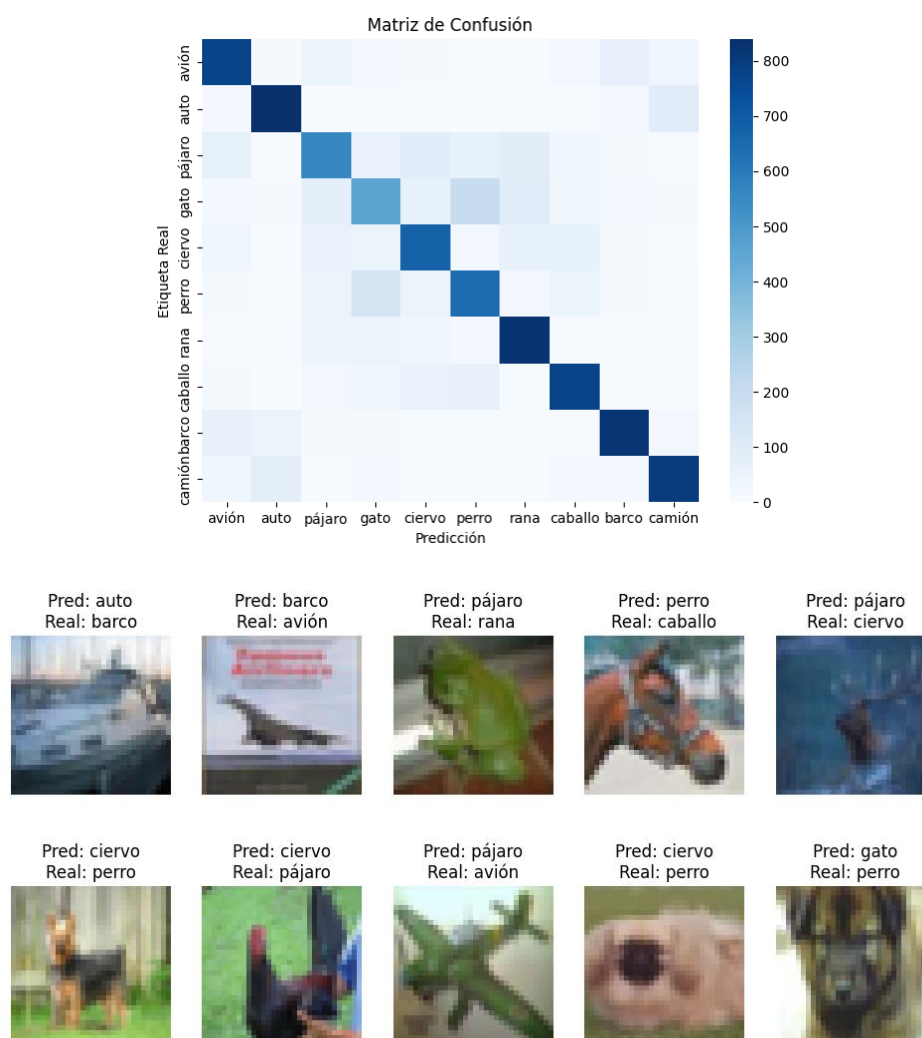


CNN

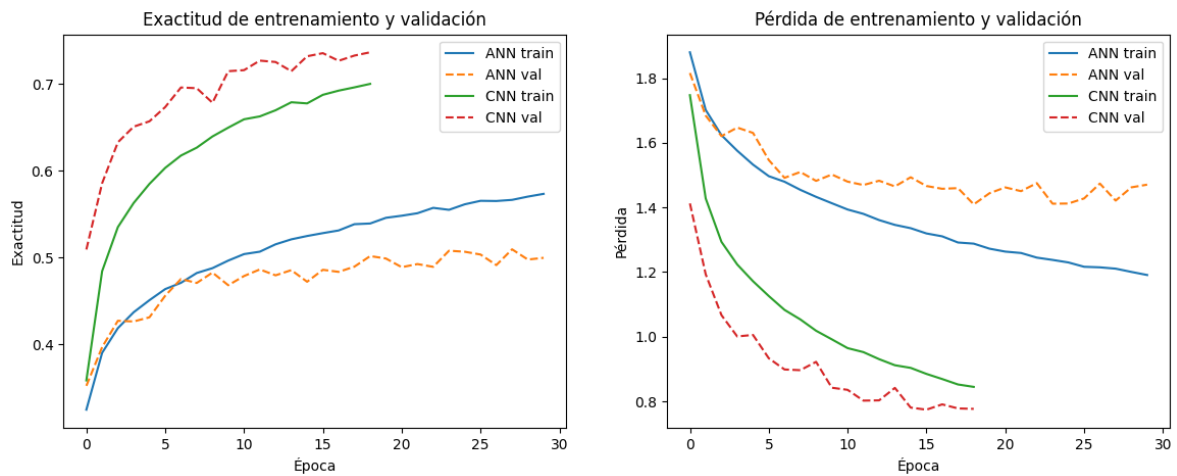
La arquitectura está compuesta por dos bloques principales de convolución + max pooling + dropout, que permiten extraer características locales de las imágenes y al mismo tiempo reducir el riesgo de sobreajuste. Luego, los mapas de características se aplanan y se pasan a una capa densa de 128 neuronas con activación ReLU y una capa final softmax con 10 salidas para la clasificación multiclase. Además, se incluyó la técnica de early stopping.

Exactitud Entrenamiento	Exactitud Validación	Tiempo Entrenamiento	Exactitud Prueba
0.69	0.73	765.27 s	0.71

Los resultados obtenidos fueron una exactitud en entrenamiento de 0.69 (69%), una exactitud en validación de 0.73 (73%) y una exactitud en prueba de 0.71 (71%), lo que indica que el modelo logró un buen nivel de generalización, siendo capaz de reconocer patrones relevantes en los datos de prueba.



Corparacion



Las gráficas muestran que la red neuronal artificial (ANN) alcanza una exactitud máxima cercana al 57% en entrenamiento y alrededor del 49% en validación, con una marcada separación entre ambas curvas, lo que evidencia dificultades para generalizar y un aprendizaje limitado en la tarea de clasificación de imágenes. En contraste, la red convolucional (CNN) obtiene resultados superiores, con aproximadamente 70% de exactitud en entrenamiento y 73% en validación, manteniendo curvas mucho más cercanas entre sí, lo que refleja una mejor capacidad de generalización. En cuanto a la pérdida, la ANN presenta una disminución gradual pero con valores de validación consistentemente más altos que los de entrenamiento, mientras que la CNN logra reducirla de manera rápida y estable, manteniéndose en niveles bajos y similares en ambos conjuntos. En conjunto, los resultados confirman que la CNN supera a la ANN tanto en exactitud como en estabilidad del entrenamiento, mostrando que las redes convolucionales son más adecuadas para problemas de clasificación de imágenes al aprovechar mejor la estructura espacial de los datos.

Reflexión

En este laboratorio permitió comparar dos aproximaciones fundamentales para la clasificación de imágenes: la red neuronal artificial (ANN) y la red neuronal convolucional (CNN), utilizando el conjunto de datos CIFAR-10. En donde los resultados obtenidos evidencian una diferencia significativa en el desempeño de ambos modelos, siendo la CNN notablemente superior en términos de exactitud, generalización y estabilidad durante el entrenamiento.

La principal razón por la que la CNN supera al modelo ANN en este problema radica en la naturaleza de los datos y la arquitectura de cada red. Debido a que las imágenes contienen información espacial y patrones locales que son difíciles de capturar por una ANN tradicional, ya que esta trata cada píxel como una entrada independiente y pierde la relación espacial entre ellos. En cambio, la CNN está diseñada específicamente para aprovechar la estructura bidimensional de las imágenes. En donde sus capas de convolución aplican filtros que recorren la imagen y extraen automáticamente características locales como bordes, texturas y formas, construyendo representaciones jerárquicas cada vez más complejas a medida que se avanza en la red. Esto permite que la CNN aprenda de manera más eficiente y efectiva los patrones relevantes para la clasificación, mientras que la ANN se ve limitada por su incapacidad para captar estas relaciones espaciales. Por otro lado, las capas de convolución y pooling juegan un papel fundamental en el éxito de las CNN. En donde las capas de convolución actúan como detectores de patrones locales permitiendo que el modelo identifique elementos clave en diferentes regiones de la imagen sin necesidad de intervención manual. Por su parte, las capas de pooling, como MaxPooling, reducen la dimensionalidad de los mapas de características y ayudan a que el modelo sea más robusto frente a pequeñas variaciones o desplazamientos en la imagen. Además, el pooling contribuye a disminuir la cantidad de parámetros y el costo computacional, facilitando el entrenamiento de redes más profundas y complejas.

Para mejorar aún más el desempeño de la CNN, existen varias estrategias que podrían implementarse. Una de las más efectivas es el data augmentation, que consiste en aumentar artificialmente el conjunto de datos de entrenamiento mediante transformaciones como rotaciones, traslaciones, volteos y cambios de escala. Esto ayuda a que el modelo generalice mejor y reduzca el sobreajuste. Otra mejora importante sería emplear arquitecturas más profundas y avanzadas, como VGG, ResNet o DenseNet, que han demostrado un rendimiento sobresaliente en tareas de visión por computadora. Asimismo, la incorporación de técnicas como batch normalization puede estabilizar y acelerar el proceso de entrenamiento, además de mejorar la capacidad de generalización del modelo.

En conclusión, la experiencia de este laboratorio confirma que las redes convolucionales son considerablemente más adecuadas para la clasificación de imágenes que las redes neuronales artificiales tradicionales. La capacidad de las CNN para extraer y combinar características

espaciales relevantes se traduce en un mejor desempeño, mayor exactitud y capacidad de generalización en el conjunto de prueba.