

Aprendizaje Automático II

Práctica 1: Clasificación de Imágenes mediante Redes Neuronales Convolucionales (CNN)

Autoras:

Laura Aguiar
Lucía Hernández

26 de diciembre de 2025

Introducción y Objetivo

Este informe detalla el desarrollo de un sistema de **Aprendizaje Profundo** diseñado para clasificar imágenes en seis categorías: edificios, bosques, glaciares, montañas, mar y calles. El objetivo central fue aplicar una metodología científica de experimentación para optimizar una CNN hasta superar el **85 % de precisión**.

Arquitectura del Sistema

El modelo se basa en una arquitectura de Red Neuronal Convolutacional, que procesa la información en tres etapas:

- **Capas Convolucionales:** Actúan como extractores de características (bordes, texturas y formas).
- **Dropout (0.5):** Técnica de regularización que .“apaga” neuronas aleatoriamente para obligar a la red a no depender de patrones específicos, evitando el sobreajuste.
- **Early Stopping:** Un monitor que detiene el entrenamiento si el modelo deja de mejorar en el conjunto de prueba, garantizando que guardamos siempre la mejor versión del modelo.

Metodología de Experimentación

Se han ejecutado cinco configuraciones, ajustando parámetros de forma aislada para comprender su impacto.

Cuadro 1: Tabla comparativa de configuraciones y resultados

Exp.	Estrategia	LR	Batch	Capas	Data Aug.	Accuracy
Config 1	Base	0.001	32	2	No	80 %
Config 2	Suavizado LR	0.0001	32	2	No	82 %
Config 3	Aumento Batch	0.0001	64	2	No	76 %
Config 4	Capacidad Visual	0.0001	64	3	No	83 %
Config 5	Robustez Final	0.0001	64	3	Sí	86 %

Análisis de la Configuración Ganadora (Config 5)

La Configuración 5 alcanzó el **86 % de precisión**, cumpliendo con los objetivos de la práctica.

4.1. El impacto del Data Augmentation

La diferencia clave radica en el **Aumento de Datos**. Al introducir variaciones aleatorias (rotación, zoom), el modelo deja de memorizar imágenes concretas y empieza a entender conceptos visuales abstractos. Esto explica por qué, con la misma arquitectura que la Config 4, logramos subir 3 puntos porcentuales.

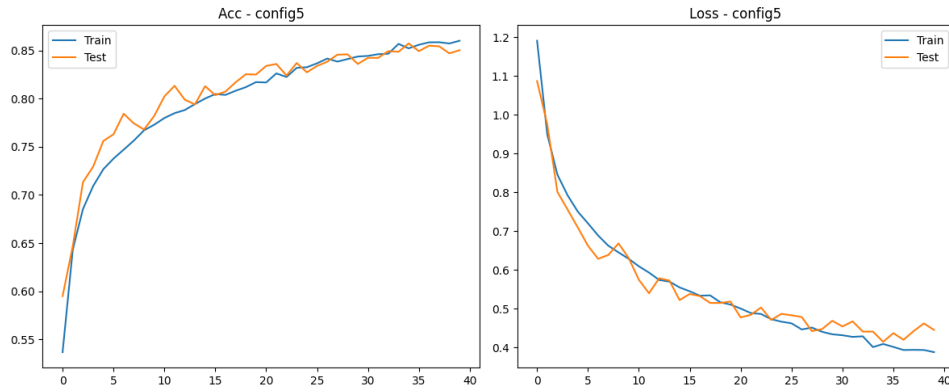


Figura 1: Curvas de Precisión y Pérdida. Se observa que el entrenamiento es estable y el error de prueba disminuye de forma constante.

4.2. Diagnóstico mediante Matriz de Confusión

Al analizar la matriz generada en `config5`, determinamos:

- **Fortalezas:** El modelo identifica con gran precisión los paisajes de tipo *forest* (98% de recall).
- **Confusiones Comunes:** Se detecta un solapamiento entre *buildings* y *street*, debido a la presencia de cemento y estructuras verticales en ambas clases.

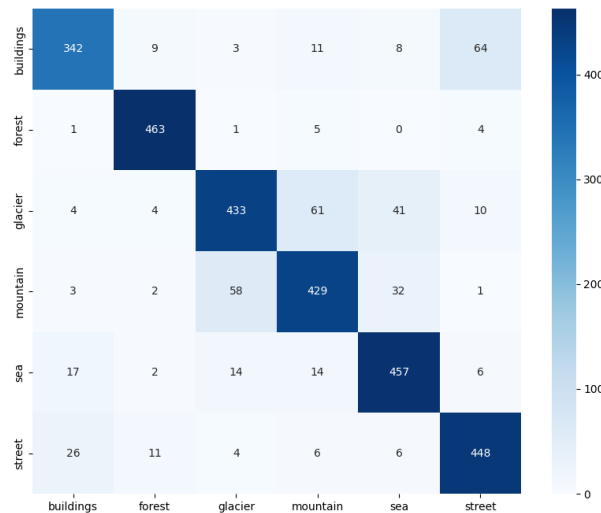


Figura 2: Matriz de confusión: el rendimiento en la diagonal principal valida la eficacia del modelo.

Conclusiones

La experimentación demuestra que un modelo de Deep Learning no solo depende de su tamaño, sino de la calidad del entrenamiento. El uso de un **Learning Rate bajo**, una **arquitectura de 3 capas** y, sobre todo, la **regularización mediante aumento de datos**, han sido los pilares para obtener un sistema robusto y preciso.