

Aprendizaje Automático II

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

Autoras:

Laura Aguiar
Lucía Hernández

Repositorio del Proyecto:

https://github.com/Lauuaguiar/AA2_Pr-ctica_1

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 fundamentales:

- **Capas Convolucionales:** Actúan como extractores de características visuales (bordes, texturas y formas complejas).
- **Dropout (0.5):** Técnica de regularización que “apaga” neuronas aleatoriamente durante el entrenamiento. Esto evita que la red dependa en exceso de píxeles específicos, reduciendo el sobreajuste.
- **Early Stopping:** Mecanismo que monitoriza el error en el conjunto de prueba y detiene el proceso si no hay mejora en 5 épocas, asegurando la mejor versión del modelo.

Metodología de Experimentación

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

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 como rotaciones y zooms, el modelo se ve obligado a aprender conceptos visuales abstractos en lugar de memorizar imágenes. Esto explica por qué, con la misma arquitectura que la Config 4, se logró subir un 3 % la precisión.

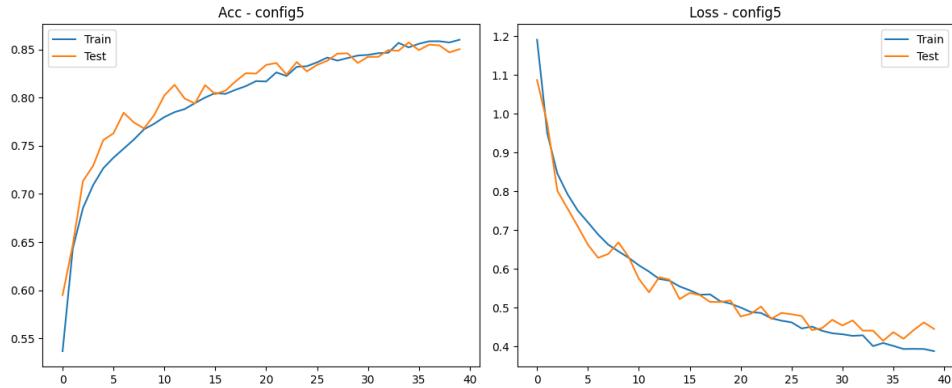


Figura 1: Curvas de Precisión y Pérdida (Config 5). La estabilidad de las curvas valida el uso de un Learning Rate bajo (0.0001).

4.2. Diagnóstico mediante Matriz de Confusión

Al analizar la matriz generada, determinamos:

- **Fortalezas:** Excelente identificación de paisajes de tipo *forest* (98 % de recall).
- **Confusiones Comunes:** Ligero solapamiento entre *buildings* y *street*, debido a la similitud de materiales geométricos (asfalto y concreto).

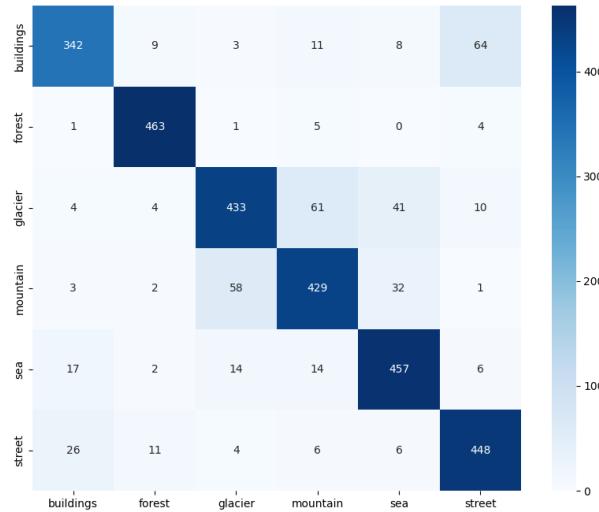


Figura 2: Matriz de confusión final del mejor modelo.

Conclusiones

La experimentación demuestra que la calidad del entrenamiento es tan relevante como la arquitectura. El uso de un **Learning Rate bajo**, una **arquitectura de 3 capas** y la **regularización mediante aumento de datos**, han sido los pilares fundamentales para obtener un sistema robusto y capaz de generalizar con éxito.