

Clasificación de Imágenes de Cartas usando Deep Learning

Gabriela Chimali Nava Ramírez | A01710530

09/11/2025

ABSTRACT La clasificación automática de imágenes es un desafío en la visión por computadora porque requiere arquitecturas capaces de aprender características complejas y muy variables. En este trabajo, se implementó una arquitectura de Deep Learning (aprendizaje profundo) con mejoras progresivas para clasificar un conjunto de 53 tipos de cartas de baraja. Los resultados demuestran cómo el Transfer Learning (v2) supera drásticamente a la CNN básica (v1), y cómo las técnicas de regularización (v4, v5) y optimización (v6) ayudan a estabilizar el entrenamiento, prevenir el sobreajuste y mejorar la capacidad para generalizar.

1. INTRODUCCIÓN

El uso de algoritmos de *Deep Learning*, específicamente las Redes Neuronales Convolucionales (CNNs), ha revolucionado el campo de la visión por computadora. Ya que permiten extraer patrones jerárquicos de las imágenes, desde bordes simples hasta formas complejas como números o letras, siendo fundamentales para tareas de clasificación. Una de las técnicas más efectivas y eficientes computacionalmente es el *Transfer Learning* (aprendizaje por transferencia), que permite aprovechar el conocimiento de modelos pre-entrenados en datasets masivos para aplicarlos a tareas más específicas, reduciendo significativamente el tiempo de entrenamiento y los datos necesarios.

En este reporte, se plantea la aplicación y mejora progresiva de un modelo de *Deep Learning* para la clasificación de un conjunto de datos real de 53 clases de cartas de juego. El objetivo es documentar la evolución del desempeño del modelo a través de seis versiones.

Iniciando con una CNN Básica (v1) que evolucionó a *Transfer Learning* con *MobileNetV2* (v2) para establecer una línea base más compleja.

Posteriormente, el modelo fue refinado buscando un mejor desempeño general. Primero ajustando la arquitectura con capas densas adicionales (v3), luego mayor variabilidad con *Data Augmentation* (v4), se modificó la regularización con *Dropout* (v5), y finalmente se optimizó el entrenamiento mediante *Learning Rate* adaptativo y *Early Stopping* (v6).

2. DESCRIPCIÓN DEL DATASET

Este proyecto se llevó a cabo utilizando el conjunto de datos “*Cards Image Dataset-Classification*”, el cuál contiene imágenes de cartas de baraja.

2a.- Estructura de Datos

El dataset está pre dividido en tres carpetas: *train* (entrenamiento), *valid* (validación) y *test* (prueba).

Todas las imágenes son de 224x224x3 píxeles en formato jpg y están recortadas de forma que sólo una carta sea visible, ocupando el 50% de los píxeles del espacio total. El modo es RGB.



Figura 1. Ejemplo de las cartas en train

Variable Objetivo (Clases)

La variable dependiente es categórica, con 53 clases únicas que representan cada tipo de carta. Ejemplo:

- ace of clubs
- king of diamonds
- seven of hearts
- joker

Volumen de Datos

- **Entrenamiento:** 7624 imágenes
- **Validación:** 265 imágenes (5 por clase)
- **Test:** 265 imágenes (5 por clase)

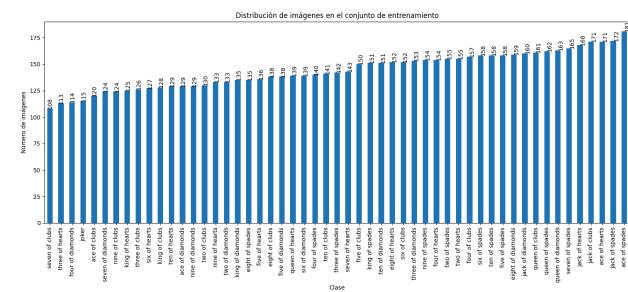


Figura 2. Distribución de clases en train

Como observamos, la representación de clase no es uniforme, por lo que posteriormente será importante contemplar técnicas para manejar este desbalance.

3. PREPROCESAMIENTO Y TRANSFORMACIÓN DE DATOS

3a.- Carga y exploración de datos TRAIN

Se utilizó *ImageDataGenerator* de *Keras* para adecuar las imágenes al formato requerido por la arquitectura y mejorar la capacidad de generalización del modelo.

4. REFERENCIAS

- [1] giose. (2025, October 29). *CardClassifier MobileNetV2 based.*
<https://www.kaggle.com/datasets/gpiosenka/cards-image-datasetclassification>
- [2] GeeksforGeeks. (2024, September 10). *Multiclass image classification using Transfer learning.*
GeeksforGeeks.
<https://www.geeksforgeeks.org/deep-learning/multiclass-image-classification-using-transfer-learning/>