



CHARACTER CLASSIFICATION

Implementación de un modelo de deep learning

Inteligencia artificial avanzada para la ciencia de datos II

Oswaldo Del Valle

A01275702

2/11/2024

Índice

Introducción	3
Dataset	3
Modelo	4
Implementación	4
Resultados	5

Introducción

Este proyecto se centra en el desarrollo de un modelo de deep learning capaz de clasificar imágenes de personajes específicos de la serie *The Quintessential Quintuplets*, donde cada una de las cinco protagonistas presenta rasgos físicos similares, para realizar esta clasificación se ha empleado un modelo basado en redes neuronales profundas utilizando la arquitectura EfficientNetB0, pre entrenada en *ImageNet* y ajustada para capturar las diferencias específicas entre los personajes.

Con un conjunto de datos de imágenes etiquetadas de cada una de las quintillizas, el modelo busca automatizar el proceso de identificación visual de cada hermana, optimizando así la precisión

Dataset

El dataset utilizado en este proyecto está compuesto por imágenes de cinco personajes principales de la serie *The Quintessential Quintuplets*, cada uno correspondiente a una de las hermanas, el conjunto de datos está organizado en cinco carpetas cada una etiquetada con el nombre de la respectiva hermana y contiene entre 600 y 700 imágenes aproximadas por carpeta, este diseño de carpetas facilita la carga de datos para la clasificación permitiendo que cada imagen esté directamente relacionada a una etiqueta de clase correspondiente al personaje

En cuanto al procesamiento se ajustó el tamaño de las imágenes a dimensiones más pequeñas para un procesamiento más eficiente por el modelo durante el entrenamiento, se usó las dimensiones de 300 x 150 para a su vez conservar la relación de aspecto de las imágenes. Luego se dividió el conjunto de datos en tres subconjuntos: entrenamiento (train), validación (val) y prueba (test), añadiendo el parámetro para conservar la proporción de imágenes de cada clase en cada conjunto lo cual ayuda a mantener una representación proporcional de los datos a lo largo del entrenamiento y evaluación

Modelo

Se seleccionó EfficientNetB0 una arquitectura de red neuronal convolucional conocida por su eficiencia y precisión en tareas de clasificación de imágenes, el modelo se utiliza con *transfer learning* ya que se carga con pesos pre entrenados en *ImageNet* un conjunto de datos extenso que permite captar características visuales generales, al emplear esta técnica el modelo aprovecha las representaciones aprendidas previamente facilitando la diferenciación entre los personajes a partir de rasgos específicos, ajustando EfficientNetB0 específicamente para el conjunto de imágenes se busca lograr una clasificación precisa de cada personaje sin requerir una cantidad masiva de datos o un alto costo computacional

Implementación

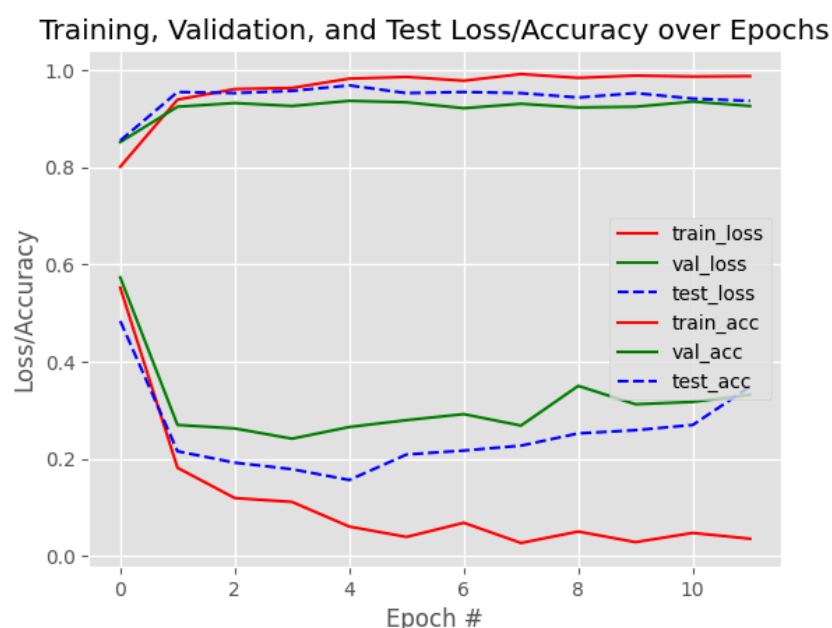
En el modelo se han añadido capas adicionales a EfficientNetB0 para optimizar su rendimiento en la clasificación específica del dataset. Primero se incorpora una capa de *GlobalAveragePooling2D* que reduce la dimensionalidad de las características extraídas por EfficientNetB0 al condensar la información relevante de cada imagen buscando mejorar así la eficiencia computacional y evitando redundancias, se agregan capas densas (*Dense layers*) con 128 neuronas y la función de activación ReLU que permite al modelo aprender relaciones no lineales complejas en las características lo cual es esencial para captar las sutiles diferencias entre los personajes, en cuanto a una medida para mitigar el sobreajuste se implementan capas de *Dropout* con una tasa del 20% eliminando conexiones aleatorias durante el entrenamiento y mejorando la capacidad de generalización del modelo

La ultima capa es una capa densa con 5 neuronas y activación *softmax* que convierte las salidas en probabilidades para cada clase, una neurona para cada clase, facilitando la interpretación de los resultados y la clasificación en una de las cinco categorías correspondientes a cada personaje. Además el modelo se entrena usando el optimizador *Adam* conocido por su capacidad de ajuste adaptativo de la tasa de aprendizaje lo cual acelera la convergencia sin necesidad de intervención manual. Finalmente la función de pérdida utilizada es la *categorical crossentropy*,

que es buena para problemas de clasificación multicategoría lo que permite evaluar la precisión del modelo en la asignación correcta de probabilidades a cada clase

Resultados

En las 12 épocas que el modelo se entrenó mostró una tendencia constante de mejora tanto en el conjunto de entrenamiento como en el conjunto de prueba alcanzando una alta precisión en ambos rápidamente, en la primera época la precisión en el conjunto de entrenamiento fue del 67% con una pérdida de 0.84, mientras que el conjunto de prueba mostró una precisión del 85% y una pérdida de 0.48, lo cual indica que el modelo logró captar características clave de cada clase desde el inicio, y a medida que el entrenamiento avanzó la precisión en el conjunto de prueba alcanzó un pico del 96.8% en la época 5 mientras que el conjunto de validación se mantuvo alrededor del 93% mostrando un ligero sobreajuste en las épocas posteriores, a partir de la época 6 las pérdidas comenzaron a aumentar levemente mientras que la precisión se mantuvo estable especialmente en el conjunto de prueba, por lo que las últimas épocas observadas la pérdida en el conjunto de prueba se incrementó hasta 0.34 con una precisión del 93% indicando que aunque el modelo sigue siendo muy preciso en la clasificación podría beneficiarse de técnicas adicionales de regularización para mejorar su capacidad de generalización y evitar una posible caída en la precisión en nuevas muestras, además considerar que el modelo pasó por un parado temprano que fue configurado para que si al paso de 7 épocas no existía una mejoría en el valor de precisión del conjunto de validación para el entrenamiento guardando los mejores parámetros obtenidos, esto para evitar un entrenamiento más largo que no estuviera generando mejoría y asegurar guardar los mejores pesos



Apoyando los resultados con la matriz de confusión se puede apreciar que en efecto el modelo cuenta con una buena precisión al clasificar las imágenes de los personajes con su clase correcta, lo que se puede apreciar observando que la diagonal es la que más peso tiene e identificando que los errores con otras clases son mínimos

