Clasificación de Imágenes con Deep Learning

Primera entrega

Vanessa Tocasuche Ochoa

Fundamentos de Deep Learning Universidad de Antioquia 2025

1. Contexto de aplicación.

A continuación, se presenta un ejemplo de aplicación de técnicas de **Deep Learning** orientado a la **clasificación de imágenes**, específicamente en la identificación de dos tipos de frutas: manzanas y tomates. Este caso práctico evidencia la relevancia de este tipo de soluciones en contextos industriales y agrícolas, donde la clasificación visual precisa de productos es una tarea esencial para optimizar procesos, reducir errores humanos y mejorar la eficiencia operativa.

Por ello, a partir de un data set obtenido en Kaggle, se desarrolla una solución orientada a este tipo de problemas, donde la optimización del modelo resulta fundamental para minimizar los falsos negativos en la clase de mayor costo —en este caso, las manzanas—, contribuyendo así a mejorar la precisión del sistema y la eficiencia operativa del proceso de clasificación.

2. Objetivo de Deep Learning

- Predecir si una imagen muestra una manzana o tomate mediante una plantilla de modelo para la extracción de características visuales, dada un conjunto de imágenes de tomates y manzanas clasificadas previamente.
- Procesar un conjunto de datos de imágenes mediante algunos temas de Fundamentos de Deep Learning como lo son Redes Neuronales Convolucionales para la extracción de características, dropout y pooling con la extracción de características más representativas para controlar el sobreajuste en datasets pequeños.

3. Dataset

Tipo de datos

Imágenes con extensión .jpeg

Tamaño (número de datos y tamaño en disco)

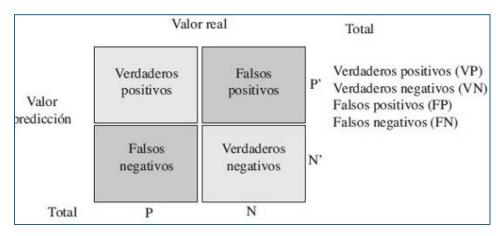
391 imágenes que ocupan 3,05 MB o 3,203,072 bytes

Distribución de las clases

Manzanas 218

Tomates 173

4. Métricas de desempeño



Fuente: Universidad Externado de Colombia

Métricas para clasificación

Métricas de mayor valor:

Accuracy (VP + VN) / (VP + VN + FP + FN)

• Presicion VP / (VP + FP)

RecallVP / (VP + FN)

Métricas complementarias

• F1-score 2 * (Precision * Recall) / (Precision + Recall)

• Matrix de confusión Tabla de imagen anterior

 AUC-ROC Capacidad del modelo para distinguir entre clases (Área bajo la curva ROC)

Métricas de Negocio:

- Tasa de error operativo [(FP + FN) / (VP + VN + FP + FN)] * 100
- Costo de error

(FP + FN) * Cu: Tomando Cu como costo promedio de tomates y manzanas.

Pero el costo de manzanas y tomates reales tiene mucha diferencia entre si, entonces se debe hacer una diferencia entre estos errores.

Definición

FPt: Número de tomates mal clasificados FNm: Número de manzanas al clasificadas. Ct y Cm

Costo total operativo Et = (FPt * Ct) + (FNm * Cm)
Costo medio Et / (VP + VN + FP + FN)

"Al optimizar el modelo para reducir falsos negativos en la clase de mayor costo (manzanas), el recall va a mejorar, además disminuye el gasto operativo generando un mejor retorno a la Inversión ROI"

5. Referencias y resultados previos

Samuel Cortinhas. (s.f.). Apples or tomatoes – Image Classification [Dataset]. Kaggle. https://www.kaggle.com/datasets/samuelcortinhas/apples-or-tomatoes-image-classification

Prthmgoyl. (s.f.). EfficientNet + ResNet + VGG: Tomato vs Apple [Notebook]. Kaggle. https://www.kaggle.com/code/prthmgoyl/efficientnet-resnet-vgg-tomato-vs-apple/notebook

Ahmed essam. (s.f.). *Apple-Tomatoes Classification* [Notebook]. Kaggle. https://www.kaggle.com/code/ahmedessammukhtar/apple-tomatoes-classification/notebook