

UNIVERSIDAD DEL VALLE DE GUATEMALA

Data Science - Sección 10

Lynette García



Excelencia que trasciende

DEL VALLE
GRUPO EDUCATIVO

Laboratorio #3

Maria Ramirez #21342

Gustavo González #21438

25 de agosto de 2024

Repositorio Github:

<https://github.com/mariaRam2003/Laboratorio4-DataScience>

Introducción

En este laboratorio, se abordó el problema del reconocimiento de dígitos manuscritos utilizando tres enfoques distintos: una Red Neuronal Convolutiva (CNN), Transfer Learning con VGG16, y un clasificador de Máquinas de Vectores de Soporte (SVM). Se llevó a cabo un análisis exploratorio de los datos, seguido de la construcción, entrenamiento y evaluación de los modelos. Finalmente, se realizó una comparación entre ellos para determinar cuál es el más efectivo para esta tarea.

Análisis Exploratorio de Datos

- **Descripción del Dataset:**

El dataset utilizado contiene imágenes de dígitos manuscritos organizadas en modalidades, con cinco modalidades distintas (m0 a m4), cada una con 60,000 imágenes.

- **Tamaño y Resolución:**

Todas las modalidades presentan una resolución única de 28x28 píxeles, lo que asegura consistencia en el tamaño de las imágenes para el entrenamiento y prueba de modelos.

- **Balanceo de Datos:**

Se realizó un balance previo ya que los datos presentaban cierta discrepancia. Los datos se balancearon, con 50,000 imágenes por modalidad.

Descripción de los Modelos

1. Red Neuronal Convolutiva (CNN)

- **Arquitectura:** Se implementó una CNN desde cero, compuesta por varias capas convolucionales seguidas de capas de pooling y capas densas al final. Se utilizó Dropout para regularización.
- **Entrenamiento:** Se entrenó el modelo utilizando un conjunto de datos reducido debido a limitaciones de tiempo.
- **Resultados:** La CNN mostró una precisión de 92.53% en el conjunto de validación tras aplicar técnicas de aumento de datos (image augmentation).

2. Transfer Learning con VGG16

- **Arquitectura:** Se utilizó el modelo VGG16 preentrenado, con sus capas convolucionales congeladas. Se añadió un clasificador personalizado en la parte superior.
- **Entrenamiento:** Al igual que con la CNN, se empleó un conjunto de datos reducido. Se utilizó la técnica de aumento de datos para mejorar la generalización del modelo.
- **Resultados:** VGG16 obtuvo una precisión de 29.62% en el conjunto de validación, mostrando un rendimiento inferior comparado con la CNN.

3. Máquina de Vectores de Soporte (SVM)

- Descripción: Se usó un SVM con un kernel gaussiano (RBF) para clasificar las imágenes de dígitos. Se escalaron las imágenes antes de entrenar el modelo.
- Entrenamiento: Debido a limitaciones de memoria, se utilizó una muestra más pequeña de datos para el entrenamiento.
- Resultados: El SVM alcanzó una precisión de 76% en el conjunto de validación.

Comparación de Modelos

- CNN vs. VGG16: La CNN superó significativamente al modelo VGG16, lo que sugiere que para este conjunto de datos, una arquitectura personalizada entrenada desde cero es más efectiva que el uso de un modelo preentrenado como VGG16. La menor precisión de VGG16 podría atribuirse a la congelación de capas y la falta de datos suficientes para ajustar las capas superiores adecuadamente.
- CNN vs. SVM: La CNN también superó al SVM. Aunque el SVM mostró un rendimiento decente, no pudo igualar la precisión de la CNN. Esto resalta la fortaleza de las CNN en la extracción de características y clasificación en tareas de visión por computadora.

Limitaciones del Estudio: Un factor clave en los resultados es el número limitado de imágenes utilizadas para el entrenamiento debido a restricciones de tiempo. Este factor afectó particularmente al rendimiento de VGG16 y SVM. Para mejorar los resultados, sería ideal entrenar los modelos con un conjunto de datos más grande.

Conclusiones:

- La CNN fue el modelo más efectivo, alcanzando una precisión de 92.53%.
- Aunque el SVM mostró un rendimiento aceptable, quedó por detrás de la CNN.
- El uso de Transfer Learning con VGG16 no fue beneficioso debido a la falta de datos suficientes.
- Se recomienda el uso de CNNs para tareas de reconocimiento de manuscritos, asegurando un volumen adecuado de datos para un entrenamiento efectivo.

Referencias

Pambudi, A. (2024). MNIST Multiple Dataset Comprehensive Analysis [Data set]. Kaggle. <https://www.kaggle.com/datasets/agungpambudi/mnist-multiple-dataset-comprehensive-analysis/data>