

Modelo de clasificación para obras de arte

Proyecto final Big Data

Daniel Pombo^a

Juan Camilo Jaramillo^a

Santiago Ramírez Castañeda ^a

^a*Universidad de La Sabana*

1. Marco Teórico -SRC

Uso, procesamiento y clasificación de imágenes

Una imagen en un plano bidimensional se puede definir como una función de dos variables, un ejemplo de esto podría ser $a(x, y)$ siendo a el brillo de una imagen en la coordenada (x, y) , el caso de una imagen digital no es fundamentalmente diferente ya que esta se divide en N filas y M columnas, donde sus intersección es definida como un pixel (Young, Gerbrands, & Vliet, 1995).

Una de las variables que considera la función de una imagen es el color o longitud de onda (λ), esta es percibida por los seres humanos gracias los tres pigmentos de colores que se encuentran en la retina (Young et al., 1995). En imágenes digitales, su representación se da de una manera similar en términos de las combinaciones de colores primarios -generalmente siendo rojo, verde y azul- donde cada canal representa la intensidad del respectivo color entre 0 y 255. En otras palabras, cada imagen en RGB está realmente compuestas por 3 imágenes.

Por otro lado, la clasificación de imágenes es un proceso que busca que la maquina pueda “leer” una nueva información que recibe en forma de imagen y mediante un proceso de entrenamiento previo pueda clasificarla en una determinada clase, para el caso de aprendizaje supervisado, las clases de un set de entrenamiento son dadas inicialmente a la máquina Raut (2023).

Este proceso de clasificación de imágenes ha sido abordado desde las redes neuronales convolucionales en los años 90 (LeCun et al., 1989), encontrando altos indicadores del desempeño de estas (Zeiler & Fergus, 2014). El funcionamiento fundamental de estas parte por un mapeo de imagen 2D x_i por una serie de capas a un vector de probabilidad \hat{y}_i para todas las posibles clases, siendo la capa final un clasificador de tipo “softmax” (Zeiler & Fergus, 2014).

Antecedentes -SRC

Un ejemplo de estas aplicaciones puede verse en el mundo del comic, donde las redes neuronales han sido identificadas en un estudio sistematico como la segunda estrategia de vision computacional más comun para analizar e interpretar los datos (Sharma & Kukreja, image segmentation, classification and recognition methods for comics)

2. Contexto y Objetivo del Proyecto -SRC

Objetivo de clasificar imágenes relacionadas a elementos artísticos

El objetivo principal de este proyecto es desarrollar un modelo capaz de clasificar imágenes con obras artísticas. Este sistema utilizará técnicas avanzadas de aprendizaje automático, específicamente redes neuronales

convolucionales, para identificar y categorizar diversos tipos de obras de arte, incluyendo pinturas, esculturas, iconografías, grabados y dibujos. La clasificación precisa y eficiente de estas imágenes no solo facilitará la organización a grandes bases de datos de arte, sino que también permitirá nuevas formas de análisis y estudio de las obras artísticas desde una perspectiva de la vision computacional.

Justificación y uso para la clasificación automatizada de piezas artísticas, casos de uso

La automatización de este proceso no solo aumentará la eficiencia del mismo, sino que también permitirá manejar volúmenes mucho mayores de información, algo especialmente útil en la era digital. Además, la clasificación automatizada puede abrir nuevas vías para la investigación académica y el estudio del arte. Por ejemplo, puede facilitar el análisis comparativo entre diferentes periodos o estilos, así como la identificación de patrones y tendencias que podrían no ser evidentes mediante el análisis manual.

3. Colección de Datos -SRC

Descripción de dónde salieron los datos y qué contienen

La información proviene de: Kaggle. Como indica su documentación, la base original cuenta con 12.800 imágenes provenientes de google images, yandex images y el “Virtual Russian Museum [VRM]”

Formato	Cantidad
.jpg	7,583
.jpeg	5,145
.png	56

Esta viene distribuida en dos carpetas con información separada “dataset” y “musemart” cada una con un respectivo “training_set” y “validation_set”, no se reportó una razón para esta separación inicial en “dataset” y “musemart” y por eso esta fue integrada inicialmente en dos únicas carpetas “training_set” y “validation_set”.

4. Análisis Exploratorio -SRC

Highlights y procedimiento corto de EDA con hallazgos

Como parte de la extracción inicial de las imágenes en el training_set, se identificaron 108 que no pudieron ser procesadas por errores en formato, y al revisar la conversión de estas en arrays se encontraron 1567 imágenes duplicadas. Finalmente, tras la limpieza de datos, se procesaron 6621 imágenes correctamente como parte del training set y con la siguiente distribución de clases:

Para el validation_set, 16 imágenes no pudieron ser procesadas por errores en formato, y se encontraron 152 duplicadas, es decir se contó con 955 imágenes para el proceso de validación.

Las imágenes cargadas fueron transformadas a arrays con un shape (150, 150, 3) para que pudieran ser operadas por la RNC. Adicionalmente, se analizaron los histogramas correspondientes para la intensidad promedio de los canales RGB de las clases de imágenes. Estos pueden observarse a continuación:

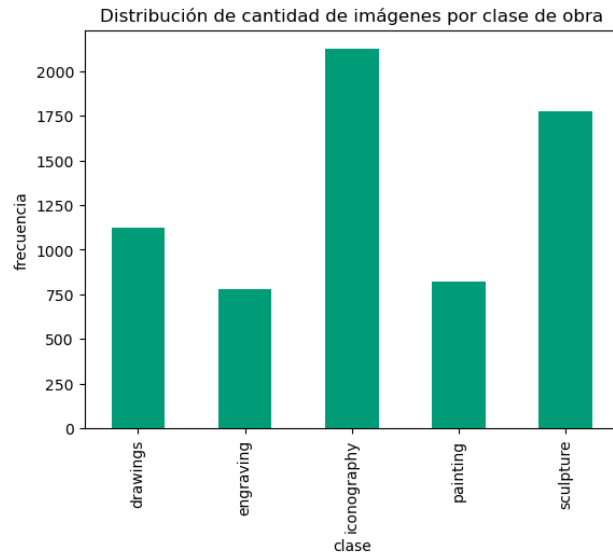
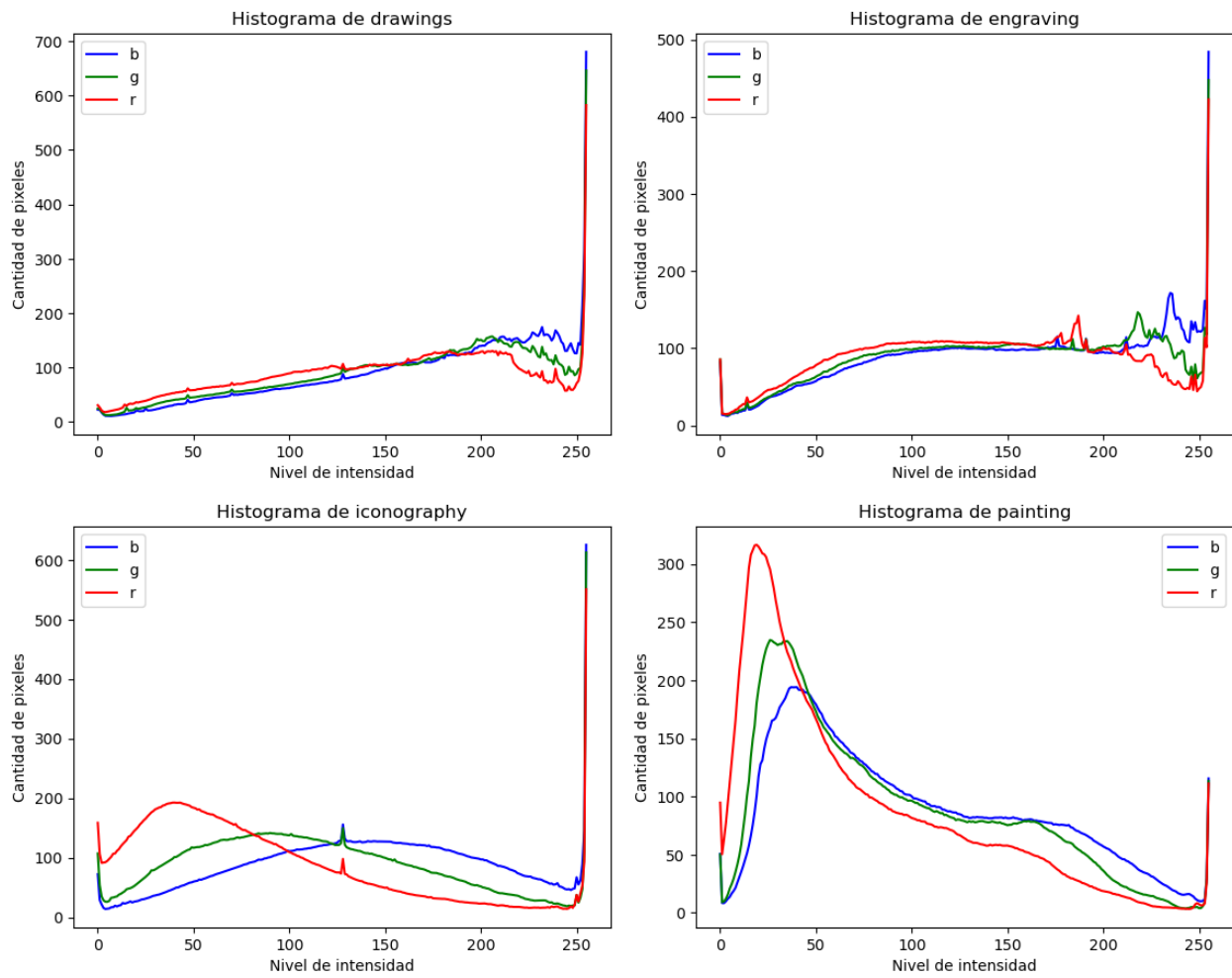
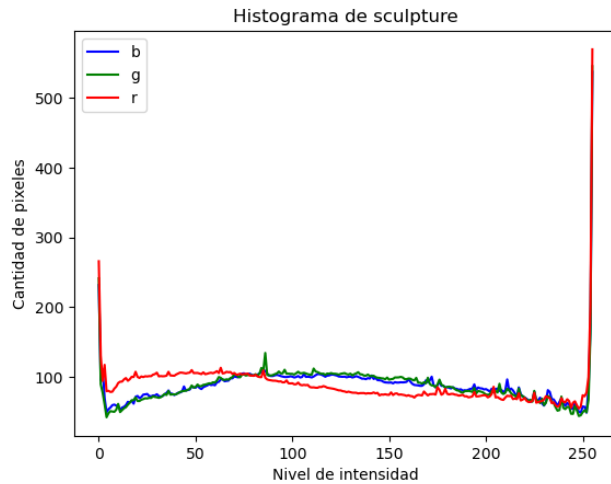


Figure 1: Distribución de clases





Algunos análisis que se pueden realizar de estos es la similitud entre los histogramas de dibujos y grabados que puede entenderse dado que estos se caracterizan por ser imágenes en escala de grises. Por otro lado, la mayor diferencia en la distribución de pixeles entre canales de color puede verse en las iconografías, esto se explica dada que en estas imágenes se pueden observar el uso de colores con un alto contraste entre sí en la creación y diseño de las mismas.

También se puede notar los picos en las zonas claras de intensidad de la imagen, esto se identificó que es debido a las imágenes que originalmente venían con “bordes” en blanco a su alrededor, se asume que estas originalmente correspondían a imágenes tipo .png que fueron transformadas a .jpg asignándoles este color blanco a los pixeles que no se registraban.

5. Modelo de Clasificación -JC

Explicación de Modelo Neuronal

Visualización 3D

Embedding de imágenes

6. Carga de Datos -DP

Explicación de RDS Amazon y su uso

Carga de embeddings de imágenes

Conexión con sistema

7. Infraestructura y Despliegue -DP & JC

Diagrama y explicación de procesamiento

Explicación breve de cómo procede FastAPI (Backend)

Explicación breve de Dash

8. Solución / Herramientas - DP

Demo

Pruebas Unitarias

Explicación de uso

9. Discusión y Conclusiones - UN POCO TODOS

Los resultados encontrados en el análisis exploratorio nos permiten visualizar algunas de las diferencias que existen entre las distribuciones de intensidad por canal de colores para cada clase de imagen. Por ejemplo, se pudo visualizar las semejanzas en esta distribución entre las clases de dibujo y

10. Limitaciones - UN POCO TODOS

11. Referencias

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural Computation*, 1, 541–551. <https://doi.org/10.1162/neco.1989.1.4.541>

- Ponnusamy, R., Sathyamoorthy, S., & Manikandan, K. (2017). A review of image classification approaches and techniques. *International Journal of Recent Trends in Engineering and Research*, 3, 1–5. <https://doi.org/10.23883/IJRTER.2017.3033.XTS7Z>
- Raut, R. (2023). An analytical approach on various deep learning models for image classification. *International Journal of Intelligent Systems and Applications in Engineering*, 11, 593–605.
- Sharma, R., & Kukreja, V. (image segmentation, classification and recognition methods for comics). Image segmentation, classification and recognition methods for comics: A decade systematic literature review. *Engineering Applications of Artificial Intelligence*, 131, 107715. <https://doi.org/10.1016/j.engappai.2023.107715>
- Young, I. T., Gerbrands, J. J., & Vliet, L. J. V. (1995). Fundamentals of image processing.
- Zeiler, M. D., & Fergus, R. (2014). Visualizing and understanding convolutional networks. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 8689 LNCS, pp. 818–833). Springer Verlag. https://doi.org/10.1007/978-3-319-10590-1_53