



Generación de paleta de colores a partir de imágenes con técnicas de machine learning no supervisado

Una paleta de colores es un conjunto de tonos que se utilizan juntos para crear un efecto armonioso, coherente y atractivo en las representaciones visuales (diseño gráfico, cine, fotografía, pintura, etc.), así como para transmitir mensajes y generar emociones determinados, de acuerdo con la intencionalidad del autor de la obra.

En un contexto de amplia demanda de herramientas que puedan asistir a diseñadores gráficos, directores de arte, pintores y creadores de contenido, entre otros, para la selección rápida y adecuada de los colores apropiados para sus obras, una aplicación para configurar paletas de colores a partir de imágenes sería de mucha utilidad.

El problema central radica en desarrollar un método automatizado que pueda generar una paleta coherente y estéticamente atractiva partiendo de los colores identificados en una imagen. Esto implica no solo reconocer tonos dominantes, sino también comprender la relación existente al interior de la gama cromática empleada para crear combinaciones visuales agradables y funcionales. Abordar este problema se traduciría en varias ventajas:

- Eficiencia en el proceso creativo, al liberar tiempo que los autores podrían emplear en otros aspectos estratégicos de sus proyectos.
- Consistencia interna en los diseños, manteniendo una identidad visual constante en diferentes aplicaciones y proyectos.
- Facilidad de uso y recurso de aprendizaje para personas con menos experiencia en diseño, fomentando la creatividad en diversos campos.
- Optimización de la calidad y del impacto de la comunicación visual.

Una vía para lograr una herramienta como esta es utilizar técnicas de machine learning no supervisado sobre imágenes para visualizar la distribución de los colores presentes en estas y generar, de manera automática, modelos de paletas de colores. En esta primera aproximación el método va a permitir identificar los colores en una imagen para construir un muestrario por

similitud de píxeles, lo cual puede ser muy útil para estudios de marketing, psicología, medicina, arte, ambiente, entre otros. Por ejemplo, se podrían extraer los colores de las diferentes superficies de la Tierra en una imagen satelital para estudiar la distribución de la vegetación o de la contaminación

A. Objetivo.

- Desarrollar un método, basado en técnicas de agrupación, que permita extraer los tonos de una imagen y generar un muestrario de los colores presentes en esta.

B. Conjunto de datos.

Los datos están asociados con imágenes de obras de arte. Pueden ser descargados a partir de este [enlace](#).

C. Actividades para realizar.

1. Recopilación de las imágenes a partir del repositorio. La idea es seleccionar un conjunto diverso de muestras en diferentes estilos artísticos.
2. Preparación de las imágenes para el entrenamiento y prueba del modelo. Para este paso construir un pipeline que integre las transformaciones que se consideren adecuadas.
3. Desarrollo del modelo de agrupación para identificar los colores presentes en las imágenes.
4. Creación de un modelo que transforme los grupos de colores identificados en un muestrario representativo. Adicionalmente, se debería mostrar la distribución de los colores de la imagen en un espacio de dos dimensiones.

D. Consideraciones.

Un diagrama general del método que se quiere desarrollar se muestra en la siguiente la Fig. 1:

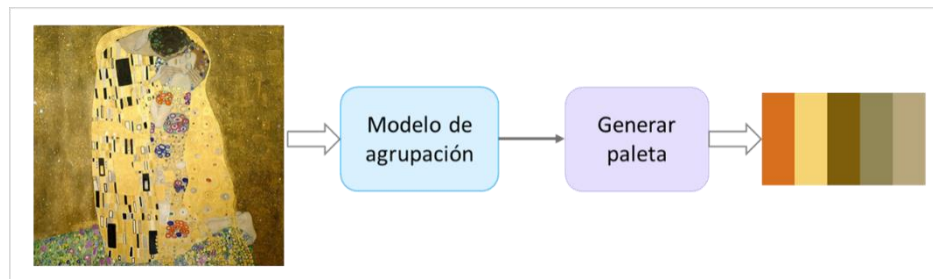


Fig. 1. Esquema del método de generación de paletas.

El algoritmo de agrupación a utilizar queda a consideración de cada grupo, pero es importante justificar la elección.

El número de colores de una paleta generada a partir de una imagen debería estar entre 5 y 7.

Para la visualización de la distribución de colores seleccionar un método como t-SNE.

E. Entregable.

Notebook (*.ipynb y *.html) del método desarrollado. El Notebook debe estar documentado con las justificaciones de las decisiones tomadas en cada paso. Además, deben ser visibles las ejecuciones de cada celda. Para evidenciar el desempeño del método construido el notebook debe mostrar la paleta generada para al menos cuatro (4) imágenes de diferentes estilos, con la visualización de la distribución de colores en el espacio de dos dimensiones

Esta entrega debe realizarse al final de la semana 5, en donde encontrarás un espacio para adjuntar los dos archivos.

F. Criterios de evaluación.

Actividad	Porcentaje	Objetivos de aprendizaje
Recopilación de un conjunto diverso de imágenes en diferentes estilos artísticos, a partir del repositorio.	10%	Caracterizar las fases del proceso de desarrollo de un proyecto de machine learning. Reconocer la importancia de la preparación de los datos como etapa previa a la aplicación de los algoritmos de aprendizaje.
Preparación de los datos, justificando las decisiones tomadas.	10%	
Construcción del pipeline de preparación de datos.	10%	
Construcción del modelo de agrupación con el algoritmo seleccionado con búsqueda de hiperparámetros, validándolo con medidas de evaluación adecuadas. Se justifica la selección del algoritmo y las métricas empleadas.	30%	Caracterizar las técnicas para resolver problemas de agrupación. Aplicar algoritmos basados en densidad, jerárquicos e "híbridos" para resolver un problema de agrupación. Identificar las métricas que pueden ser utilizadas para evaluar modelos de agrupación con base en las características del algoritmo empleado.

Construcción de la función que transforma los grupos de colores identificados en un muestrario representativo	15%	Caracterizar las fases del proceso de desarrollo de un proyecto de machine learning.
Evidencia del desempeño del método construido mostrando la paleta generada (entre 5 y 7 colores) para al menos cuatro (4) imágenes de diferentes estilos, con la visualización de la distribución de colores en el espacio de dos dimensiones	25%	Aplicar los algoritmos PCA y t-SNE para la reducción de la dimensionalidad y visualización de conjuntos de datos