
Exploración, Transformación y Análisis de Componentes Principales de Imágenes de Flores

Alejandro Uribe
Universidad de Buenos Aires
jauriberamirez@gmail.com

Abstract

El estudio exploró imágenes florales mediante transformaciones de datos y técnicas de agrupamiento sin usar modelos preentrenados. Este enfoque permitió desentrañar relaciones, identificar patrones y categorizarlas. Los hallazgos ofrecieron valiosas perspectivas sobre la diversidad y similitudes en estas imágenes, sentando así bases sólidas para futuras investigaciones en visión por computadora y clasificación floral. Destacó la posibilidad de extraer información relevante incluso sin entrenamiento de modelos.

1. Introducción

El procesamiento de imágenes ha evolucionado desde los métodos analógicos utilizados en los inicios de la fotografía hasta las técnicas digitales de hoy en día. Con el avance de la computación, se ha vuelto posible aplicar transformaciones digitales para mejorar la calidad y características de las imágenes, superando los antiguos métodos físicos y químicos. Sin embargo, el aumento en la resolución de las cámaras ha generado imágenes digitales de alta dimensionalidad, planteando un desafío en su procesamiento.

En el corazón de una imagen digital se encuentra una matriz de píxeles que representan colores a través de valores de intensidad de luz en un rango de 0 a 255. Para representar imágenes a color, se utiliza el modelo RGB, basado en los componentes primarios del color: Rojo, Verde y Azul. Este enfoque propuesto se centra en explorar el procesamiento de imágenes a través de transformaciones, con el objetivo de familiarizar al científico de datos con el proceso, ofreciendo una vía para explorar y buscar posibles agrupamientos naturales. Esta exploración no solo se limita a aspectos técnicos y comparativos de las imágenes, sino que también incluye la búsqueda de características distintivas a través de análisis estadístico descriptivo [1, 2].

2. Trabajo relacionado

En la última década, el desarrollo de técnicas de aprendizaje profundo ha permitido grandes avances en el campo de la visión por computadora, se han desarrollado numerosos modelos de aprendizaje profundo, como AlexNet, VGG, ResNet, Inception, etc. Estos modelos se entrenan con conjuntos de datos de gran tamaño, como ImageNet, que contiene más de 14 millones de imágenes etiquetadas [1]. No obstante, existen otras áreas en las que el procesamiento de imágenes se aplica de una forma más sencilla y sin recurrir a modelos de inteligencia artificial para procesar imágenes, como los filtros de Instagram o Snapchat.

Visión por computadora El procesamiento de imágenes es esencial para tareas como reconocimiento de objetos, detección de rostros, segmentación de imágenes, seguimiento de objetos en movimiento, etc [1].

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

Redes sociales Las redes sociales como Instagram, Facebook, Snapchat, etc permiten a sus usuarios compartir imágenes. Estas aplicaciones permiten procesar imágenes, como aplicar filtros, antes de compartirlas.

3. Preprocesamiento de imágenes

El proceso incluyó la descarga del conjunto de imágenes que se encontraban en la plataforma Kaggle, la carga de las imágenes, la estandarización de las imágenes, la exploración de las imágenes y la aplicación de técnicas de agrupamiento.

3.1. Datos

El conjunto de datos base estaba conformado por 210 imágenes a color de flores pertenecientes a 10 especies diferentes, acompañadas de etiquetas que indicaban la especie a la que pertenecía cada imagen [3]. Las imágenes se encontraban en formato **png** y, en su mayoría, tenían dimensiones 128×128 . La Figura 1 muestra ejemplos de imágenes del conjunto de datos.



Figura 1: Ejemplos por clase.

3.2. Estandarización

Aquellas imágenes que contaban con dimensiones distintas a 128×128 , fueron redimensionadas esperadas mediante un recorte desde el centro de la imagen. Existen otros medios para redimensionar imágenes, pero se optó por el recorte debido a que las imágenes no suelen contener información relevante en los bordes.

La Figura 2 muestra el total de imágenes por clase. Se observa que la clase con mayor cantidad de imágenes es la de **peony**, mientras que la clase con menor cantidad de imágenes es la de **Goldquelle**.

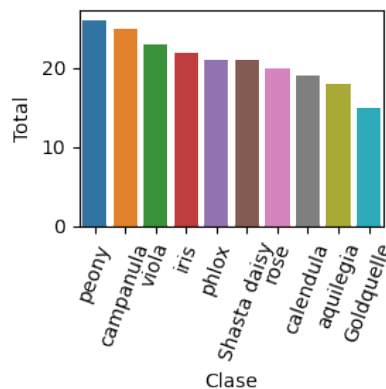


Figura 2: Total imágenes por clase.

3.3. Transformación de imágenes

A lo largo del trabajo, se abordaron distintas etapas de preprocesamiento, manipulación y análisis de datos, con el objetivo de proporcionar un enfoque completo y efectivo para la detección de patrones naturales en este conjunto de imágenes.

3.3.1. Procesadores de imágenes

Las imágenes se representan como matrices de dimensiones $N \times M$. Sin embargo, para poder manipularlas, era necesario convertirlas a un formato que pudiera ser procesado por los algoritmos. Para ello, se utilizaron las librerías PIL (*Python Imaging Library*), Numpy (*Numerical Python*) y OpenCV (*Open Source Computer Vision Library*). Dichas librerías permitían cargar las imágenes y convertirlas a un formato de **array** de **Numpy**, que podía ser manipulado con facilidad. La Figura 3 muestra un ejemplo de la conversión de una imagen a color a escala de grises.

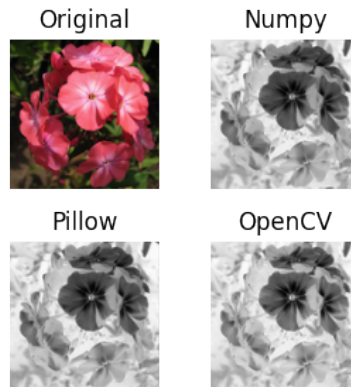


Figura 3: Procesadores.

De la Figura 3 se puede observar que independiente de la librería se puede obtener el mismo resultado. Sin embargo, la librería **OpenCV** permite obtener el resultado en una sola línea de código, mientras que las otras librerías más que eso.

3.3.2. Transformaciones

Se llevaron a cabo distintas transformaciones como parte del preprocesamiento de los datos. Estas transformaciones incluyeron cambios de color, cambios en la intensidad del color y aplicación de filtros. La Figura 7 muestra ejemplos de las transformaciones que se llevaron a cabo.



(a) Brillo.

(b) Superposición total/parcial.

(c) Filtro.

Figura 4: Transformaciones.

El recorte fue otra de las transformaciones que se llevaron. La Figura 5 muestra un ejemplo de esta transformación.



Figura 5: Recortes.

Para llevar a cabo esta transformación se llevó a cabo la multiplicación de la matriz de la imagen por una matriz de ceros y unos. Dicha matriz se construyó de acuerdo al siguiente pseudocódigo, el cual dibuja un círculo de radio R centrado en las coordenadas (c_x, c_y) :

```

if  $(i - c_x)^2 + (j - c_y)^2 < R^2$  :
    mask[i, j] = 1
else :
    mask[i, j] = 0

```

4. Análisis de imágenes

Se realizó un análisis descriptivo por cada canal (*Rojo*, *Verde* y *Azul*) y por cada clase. Para ello, se calculó la media, de cada canal.

Histogramas La Figura 6 muestra los histogramas de cada canal por clase. Se observa que los histogramas del canal *Azul* se encuentra concentrado cerca a la intensidad de 0 en todas las clases. Ahora bien, los histogramas del canal *Verde* se encuentran similares entre las clases, esto puede deberse a que la mayoría de imágenes se encuentran rodeadas de verde, sean hojas y/o tallos. Por otro lado, los histogramas del canal *Rojo* son distintos entre las clases, esto puede deberse a que las flores tienen distintos colores. En aquellas flores que tienen el color rojo predominante como **calendula**, **rose** o **peony** se aprecia que el histograma se encuentra concentrado cerca a la intensidad de 255.

y *Verde* son similares, mientras que el histograma del canal *Azul* es distinto.

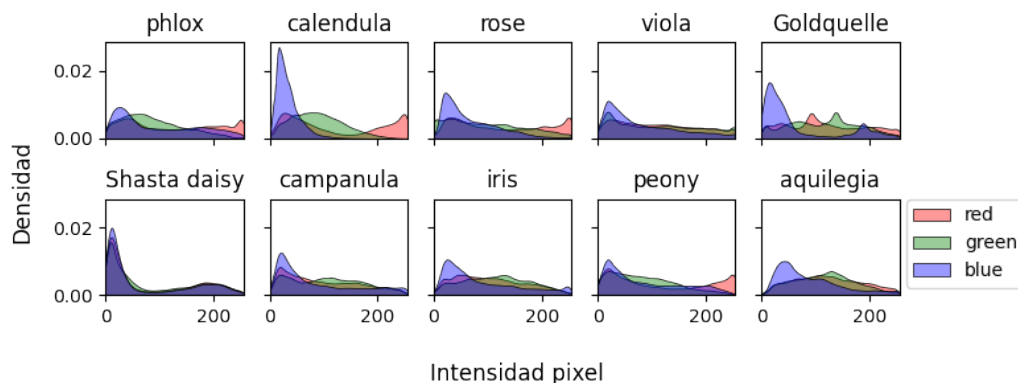
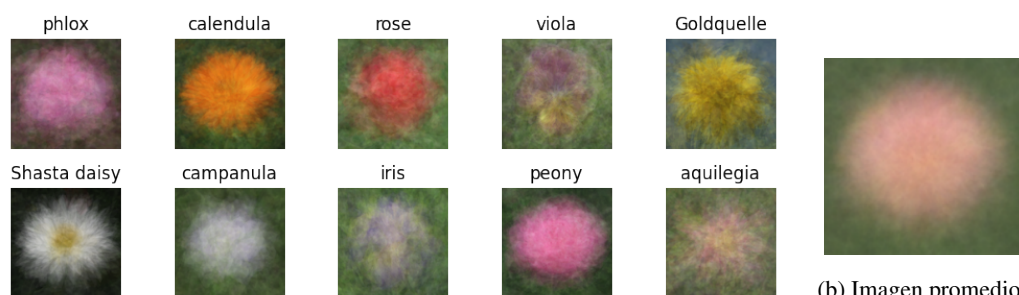


Figura 6: Distribución intensidad pixels por canal y clase.

Medias Para cada una de las clases, se calculó la media de cada canal. La Figura 7a muestra la imagen promedio por clase, mientras que la Figura 7b muestra la imagen promedio de todas las clases.



(a) Imagen promedio por clase.

Figura 7: Transformaciones.

Se observa que la imagen promedio de las clases **phlox**, **calendula**, **rose**, **Goldquelle**, **Shasta daisy**, **campanula** y **peony** son similares, todas corresponden a un círculo con un color predominante. En consecuencia, la imagen promedio de todas las clases es similar a la combinación de las imágenes promedio de las clases en cuestión, hecho que coincide con que la mayoría de flores tienen una forma circular y se encuentran en el centro de la imagen. O bien, la imagen promedio corresponde a la combinación de las flores **phlox** y **calendula**. Por otro lado, las imágenes promedio de las clases **iris**, **viola** y **aquilegia** son distintas, esto puede deberse a que estas clases tienen flores con formas distintas a las de las otras clases.

5. Análisis de componentes principales

Se llevó a cabo un análisis de componentes principales (**PCA**, por sus siglas en inglés) para reducir la dimensionalidad de los datos y facilitar la visualización de sus atributos. Previo a este análisis, se estandarizaron los datos, es decir, se dividió cada valor entre 255 y se restó 0.5. La Figura 8 muestra la proyección de los datos en el plano de los tres primeros componentes principales [4].

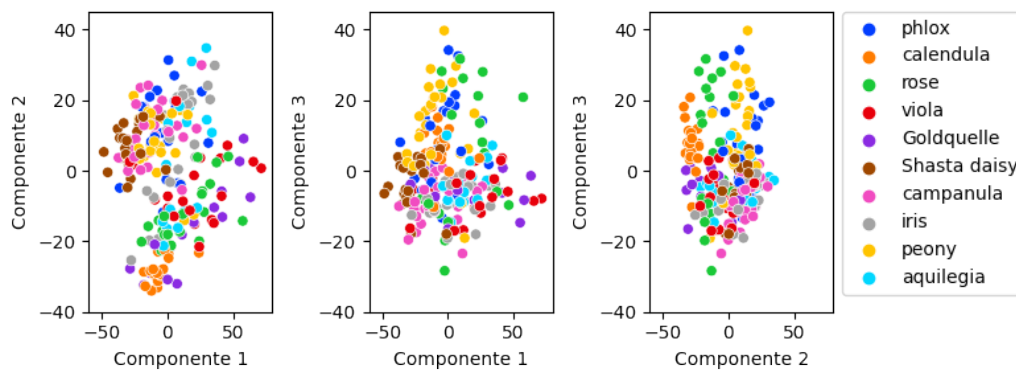


Figura 8: Análisis de componentes principales.

De la Figura 8 se concluyó que no hay un patrón específico que permita agrupar las imágenes por clase.

Por otro lado, se calculó la varianza explicada por cada componente principal. La Figura 9 muestra la varianza explicada por cada componente principal.

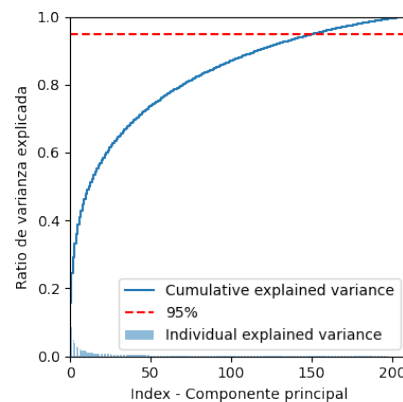


Figura 9: Varianza explicada.

Finalmente, el Análisis de Componentes Principales permitió reducir la dimensionalidad de los datos de 49152 a 210, es decir, se redujo la dimensionalidad en un 99.57 %. Además un análisis como el de la varianza explicada permite determinar cuántos componentes principales son necesarios para explicar la mayor parte de la varianza de los datos. Para un trabajo futuro se propone comprimir las imágenes a través del análisis de varianza explicada, por ejemplo, tomar aquellos atributos que

expliquen el 95 % de la varianza y comparar el rendimiento de un modelo de clasificación con los datos comprimidos y sin comprimir [5].

Referencias

- [1] A. Geron, *Hands-on machine learning with Scikit-Learn and TensorFlow : concepts, tools, and techniques to build intelligent systems*. Sebastopol, CA: O'Reilly Media, 2017.
- [2] R. Liu, Z. Wu, S. Yu, and S. Lin, "The emergence of objectness: Learning zero-shot segmentation from videos," in *Advances in Neural Information Processing Systems* (M. Ranzato, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, eds.), vol. 34, pp. 13137–13152, Curran Associates, Inc., 2021.
- [3] O. Belitskaya, "Flower color images | kaggle," 2017.
- [4] I. T. Jolliffe, "Principal component analysis," p. 271, 1986.
- [5] S. M, "Image compression using pca | kaggle."