

Machine Learning

Elían Fernando Mujica Armero
Octubre 2024

Corporación Universitaria Iberoamericana
Facultad de Ingeniería
Inteligencia Artificial

Tabla de Contenidos

Introducción	3
Metodología	3
Recolección de las imágenes	3
Aprendizaje no supervisado	4
Aprendizaje Supervisado	7
Resultados	10
Conclusiones	14
Lista de referencias	15

Introducción

En el presente documento se explica la metodología usada para el desarrollo de las actividades 3 y 4 relacionadas al Aprendizaje de Máquina. Para ello, se utiliza un conjunto de imágenes de 5 tipos de comida típica colombiana y se realiza una clusterización utilizando el algoritmo k-means, seguida de la aplicación de una red neuronal convolucional (CNN) para el aprendizaje supervisado.

Metodología

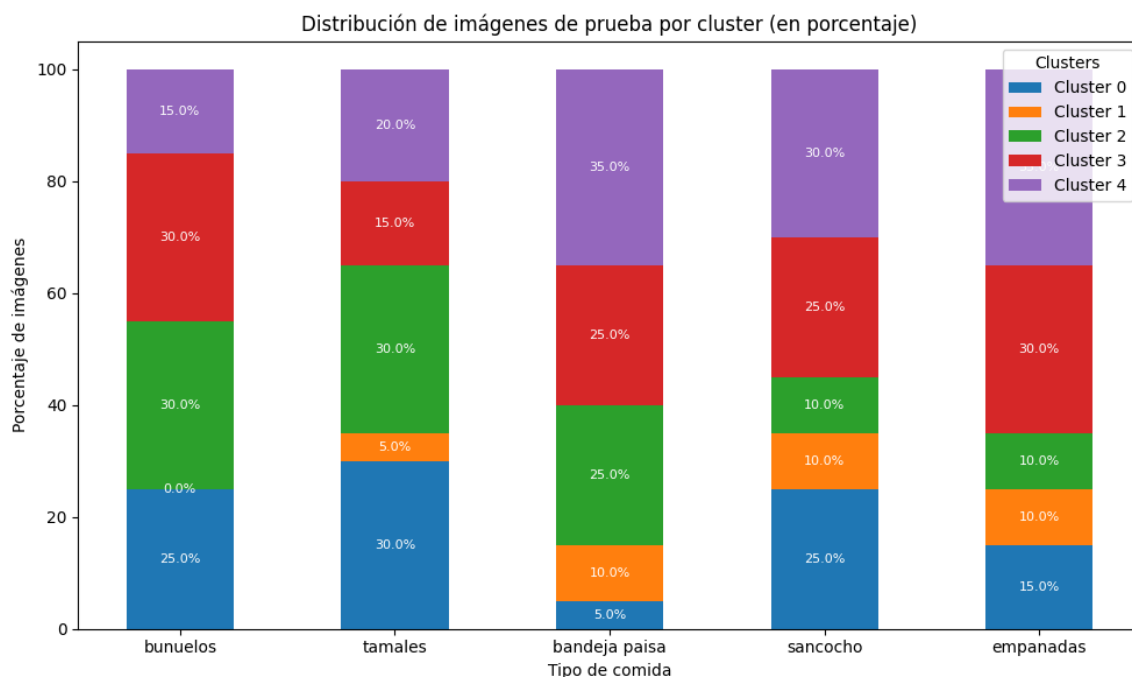
Recolección de las imágenes

Para la recolección de las imágenes, se desarrolló un script utilizando Selenium para obtener imágenes de Google. El script abre el navegador en la ruta de búsqueda de las imágenes y, mediante BeautifulSoup, extrae las imágenes identificándolas por los nombres de las clases, descargándolas automáticamente. Aunque estas imágenes son de baja calidad, resultan útiles para el entrenamiento del modelo.

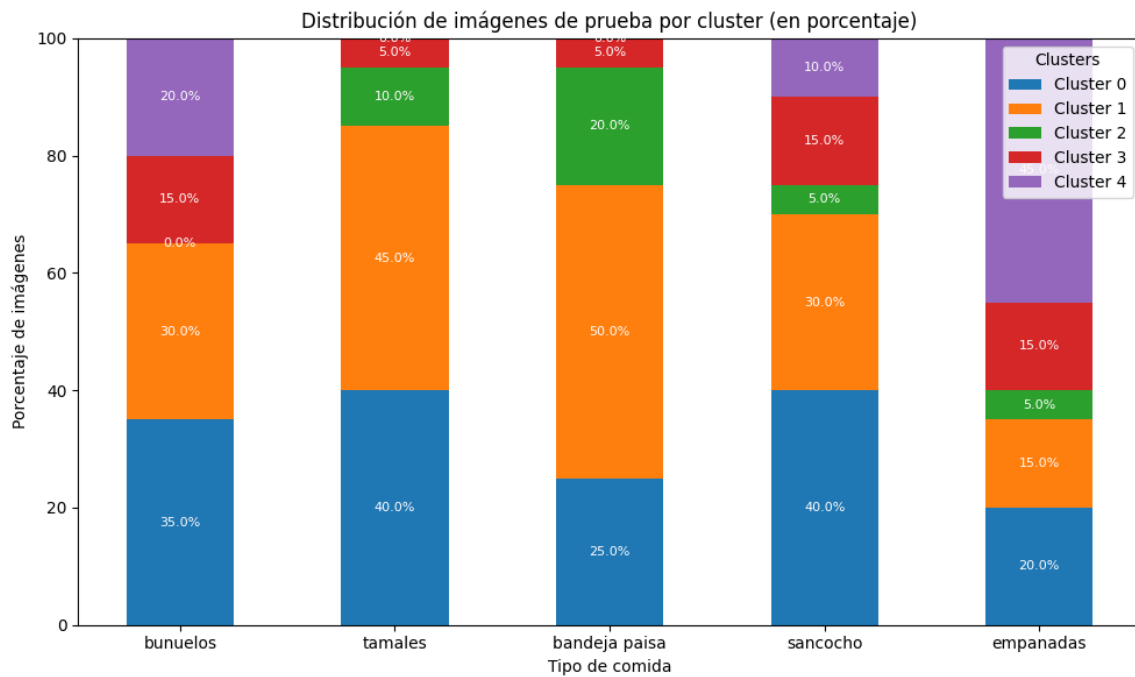
Las imágenes recolectadas corresponden a cinco categorías de comida típica colombiana: bandeja paisa, buñuelos, empanadas, sancocho y tamales. Inicialmente, se obtuvieron aproximadamente 6000 imágenes, pero se realizó una revisión manual para eliminar superficialmente aquellas que tuvieran contaminación en los datos. Además, se eliminaron algunas categorías de comida, quedando finalmente un total de 2142 imágenes, con aproximadamente 400 imágenes por cada tipo de comida. Adicionalmente, se apartaron 20 imágenes de cada categoría con el propósito de utilizarlas en la evaluación final de los modelos.

Aprendizaje no supervisado

Para la clasificación inicial, se consideró el uso del algoritmo k-means con el objetivo de separar las imágenes en múltiples clústeres. En primer lugar se realizó una implementación sencilla del algoritmo k-means, comenzando con 10 tipos de comida. Sin embargo, los resultados iniciales fueron bastante deficientes, ya que los clústeres no lograban distinguir de manera efectiva entre las distintas categorías de alimentos. Como respuesta, se redujeron las categorías a 5, excluyendo aquellas que eran muy similares entre sí o que presentaban una gran cantidad de elementos, pero los resultados seguían siendo de baja calidad:



Por esta razón, se aplicó una técnica basada en histogramas de color, lo cual permitió una ligera mejora en la calidad de la clasificación. No obstante, estos avances requerían aún de ajustes adicionales para lograr resultados óptimos:

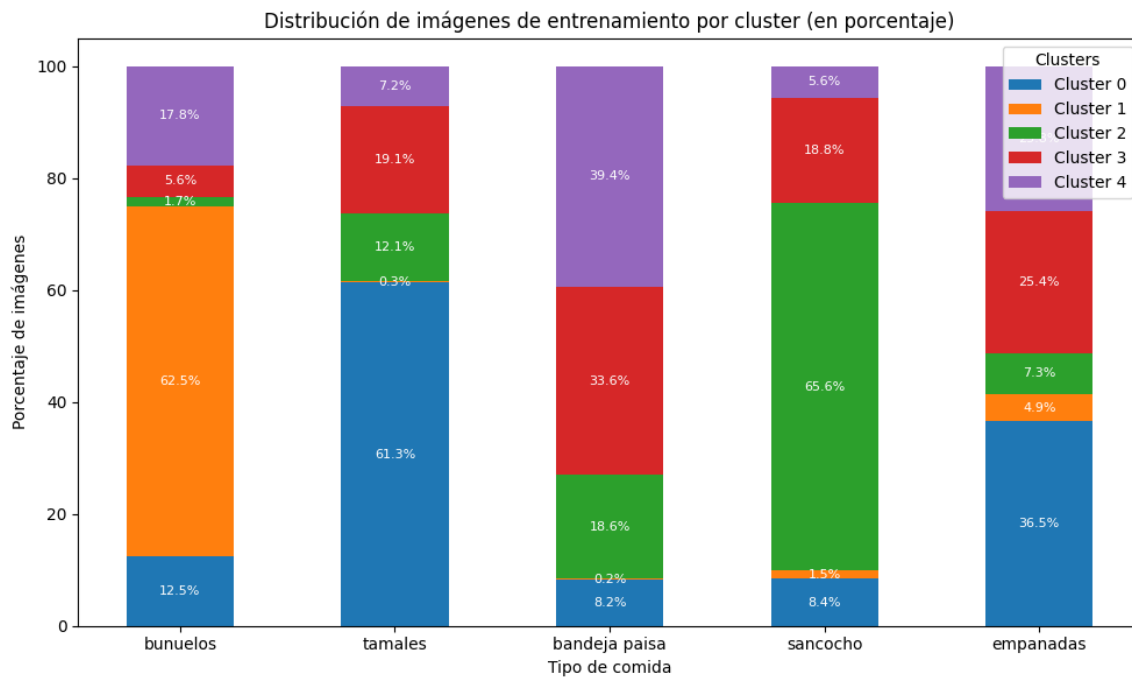


Se intentó aplicar una técnica de data augmentation, pero los resultados obtenidos fueron equivalentes a los anteriores, con la diferencia de un procesamiento significativamente más lento, lo que llevó a descartar esta técnica.

Dado que los histogramas de color mostraron cierto nivel de efectividad, se optó por una técnica similar pero mejorada, aplicando modelos supervisados para la extracción de características que permitieran obtener una mejor distribución mediante k-means.

Inicialmente se utilizó VGG16, seguido de ResNet, y finalmente se implementó

EfficientNetV2B0, con los siguientes resultados:



A partir del análisis de los resultados de la clusterización, se puede observar lo siguiente:

el clúster 1 (naranja) presenta una alta representación de buñuelos, alcanzando un 62.5% de las imágenes en dicha categoría. El clúster 0 (azul) se asocia principalmente con los tamales, con un 61.3% de las imágenes. Para el clúster 3 (verde), se observa una fuerte asociación con el sancocho, con un 65.6% de las imágenes correspondientes. Por otro lado, el clúster 4 (morado) agrupa en gran medida la bandeja paisa, con un 39.4% de las imágenes.

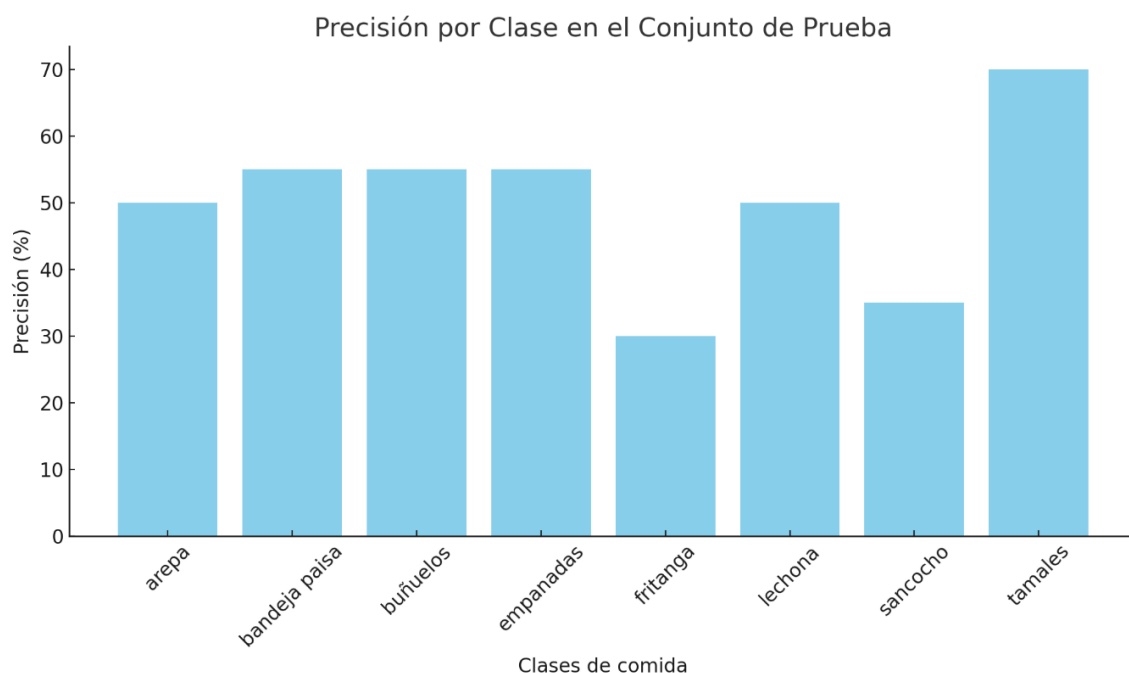
Sin embargo, el clúster 3 (rojo) no logra agrupar adecuadamente un tipo específico de comida, lo que indica que no generaliza correctamente. Además, la bandeja paisa y las

empanadas no logran agruparse en un clúster dominante, en contraste con los otros alimentos que tienen una clara asociación con un clúster específico, logrando una representación significativa de al menos un 60%.

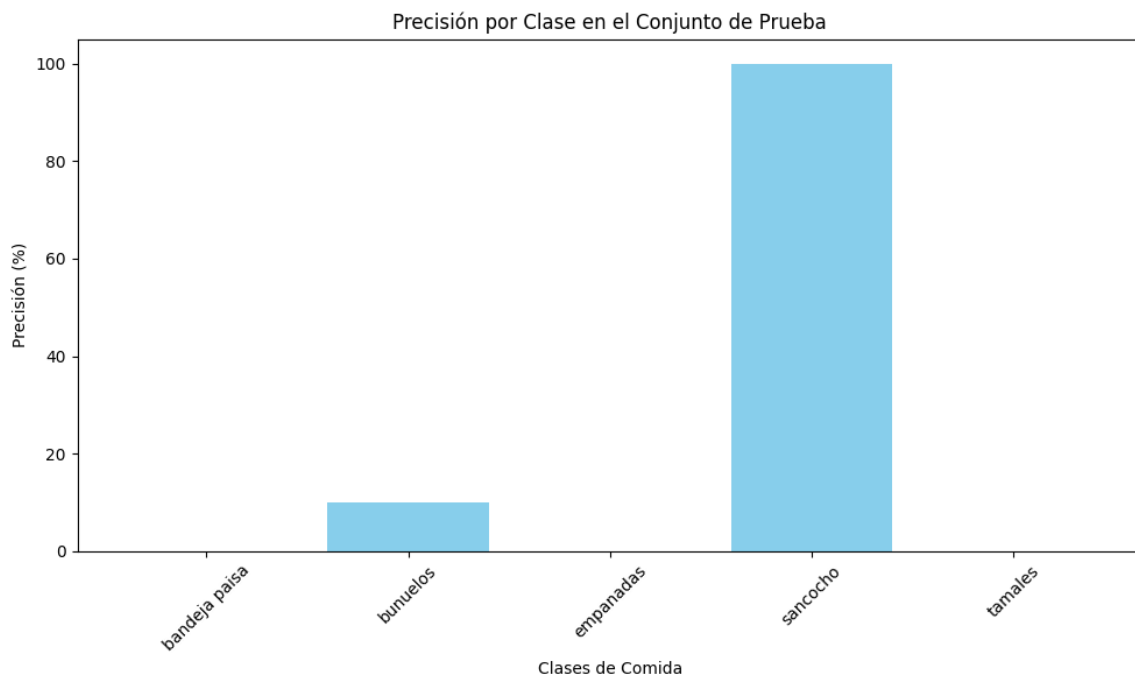
Es por esto por lo que, debido a los resultados inconsistentes obtenidos con las técnicas de clusterización no supervisadas, se decide, a continuación, utilizar un enfoque de aprendizaje supervisado para mejorar la clasificación de los tipos de comida.

Aprendizaje Supervisado

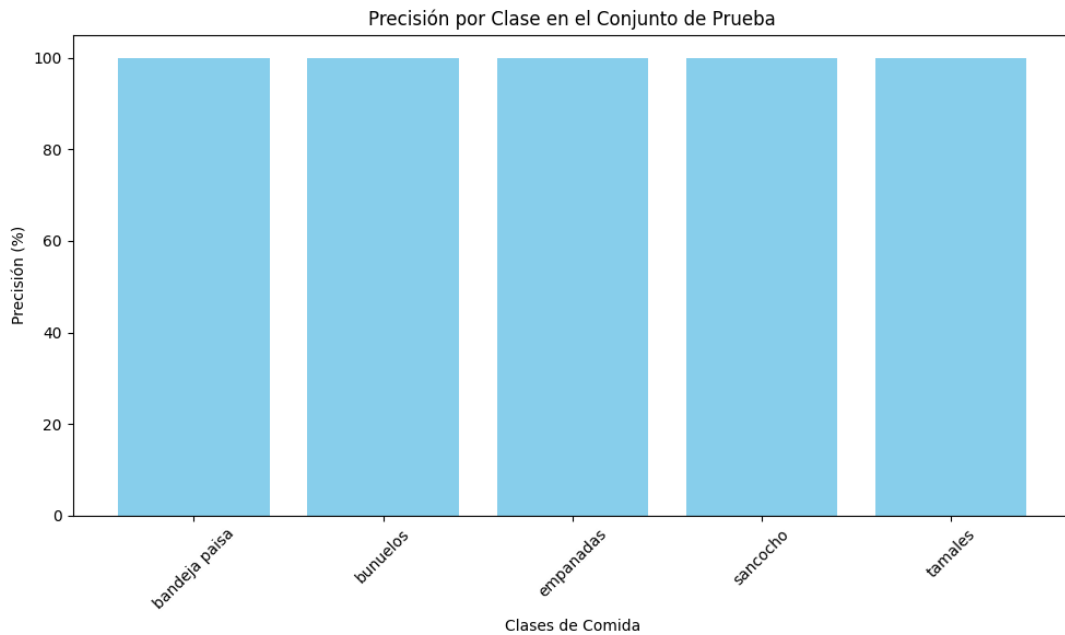
Para mejorar la clasificación de los tipos de comida, se decidió utilizar redes neuronales convolucionales (CNN). En un primer intento, se trabajó con las 10 categorías originales de comida, logrando una precisión global del 50%. Aunque se obtuvieron mejores resultados que con el algoritmo k Means, se evidenció la necesidad de continuar ajustando el modelo para mejorar su rendimiento.



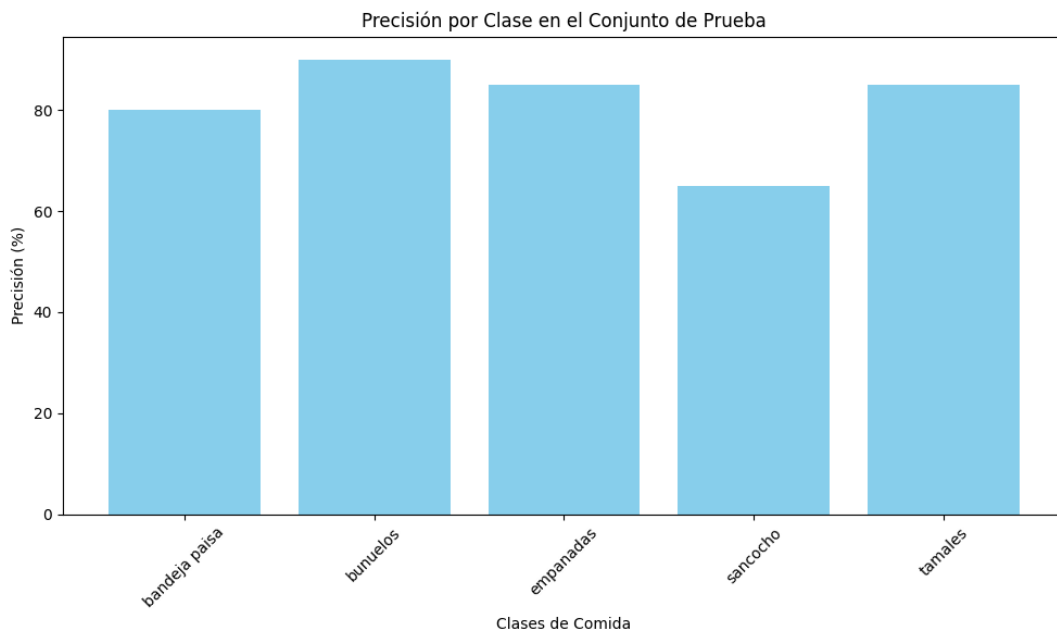
En un segundo intento, se redujeron las categorías a 5 tipos de comida, se entrenó el modelo con 20 epoch, se aumentó el tamaño del batch y se realizaron ajustes en la técnica de data augmentation. Sin embargo, la precisión global fue del 21%, mostrando un sobreajuste significativo hacia el sancocho, que alcanzó una precisión del 100%, mientras que las demás clases mostraron un rendimiento deficiente.



En un tercer intento, se buscaron arreglar los resultados obtenidos anteriormente. Se disminuyó el tamaño del batch, se añadió un mecanismo de early stopping, se modificaron las imágenes utilizando histogramas de color y se aumentó el número de epoch a 50. Aunque se estableció un mayor número de epoch, el modelo alcanzó una precisión global del 100% ya en la epoch, indicando nuevamente un problema de sobreajuste debido a un error en la configuración de los datos de prueba.

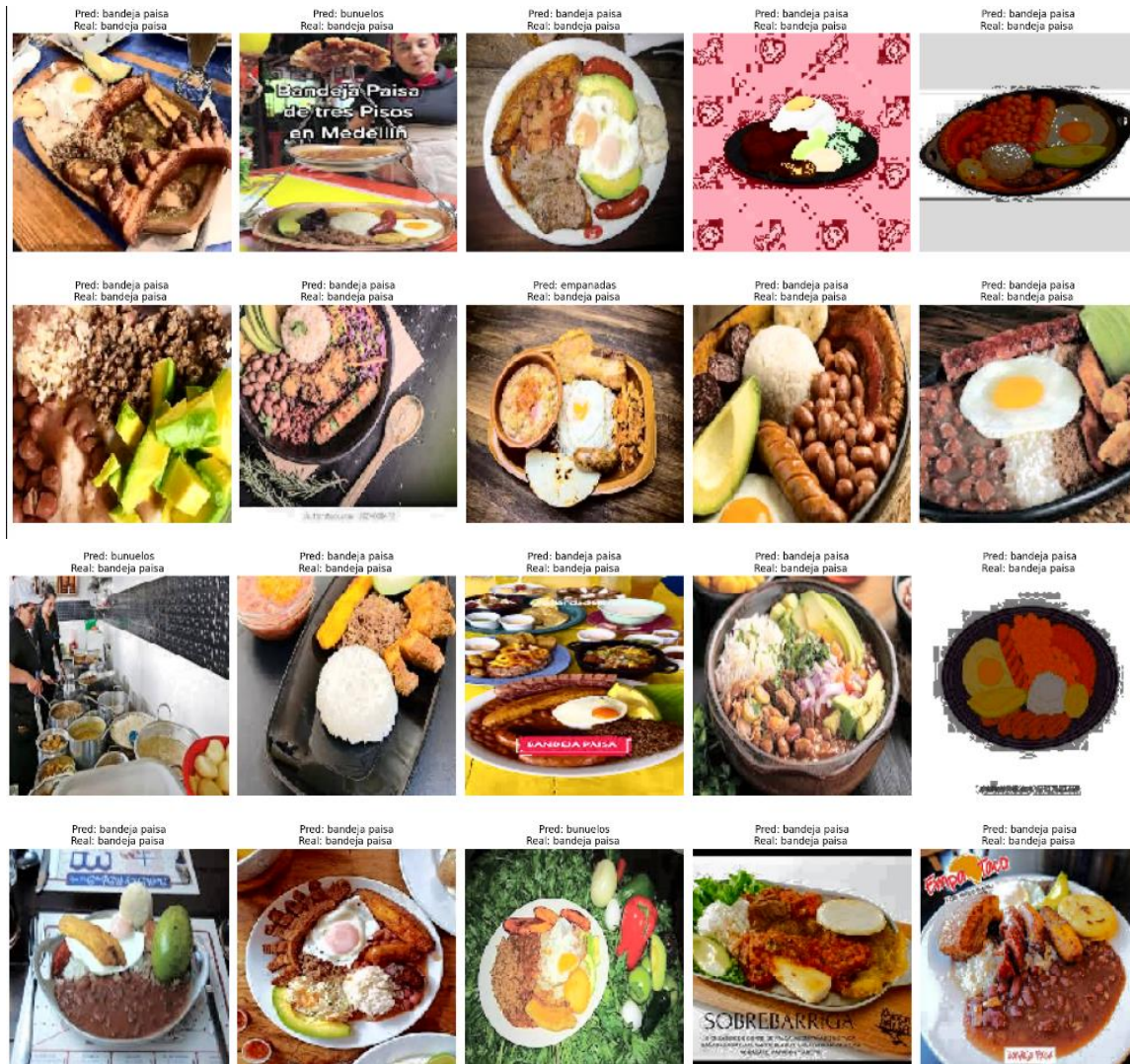


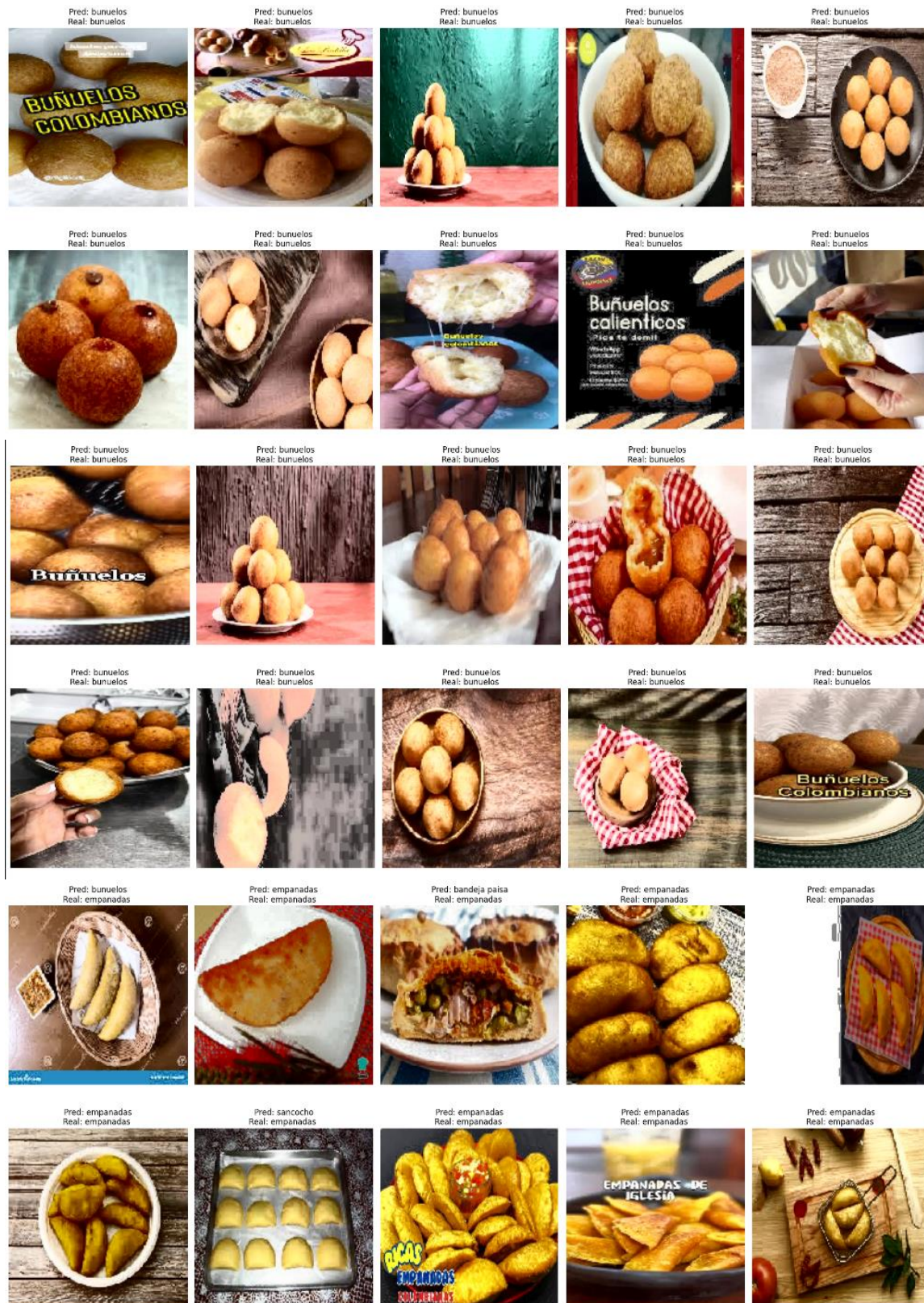
Finalmente, se corrigieron los problemas para evitar el overfitting. Se realizaron cambios en la técnica de data augmentation, se añadió batch normalization, dropout y el método ReduceLROnPlateau para optimizar el proceso de entrenamiento. Gracias a estos cambios, se logró alcanzar una precisión global del 82% evitando el overfitting:

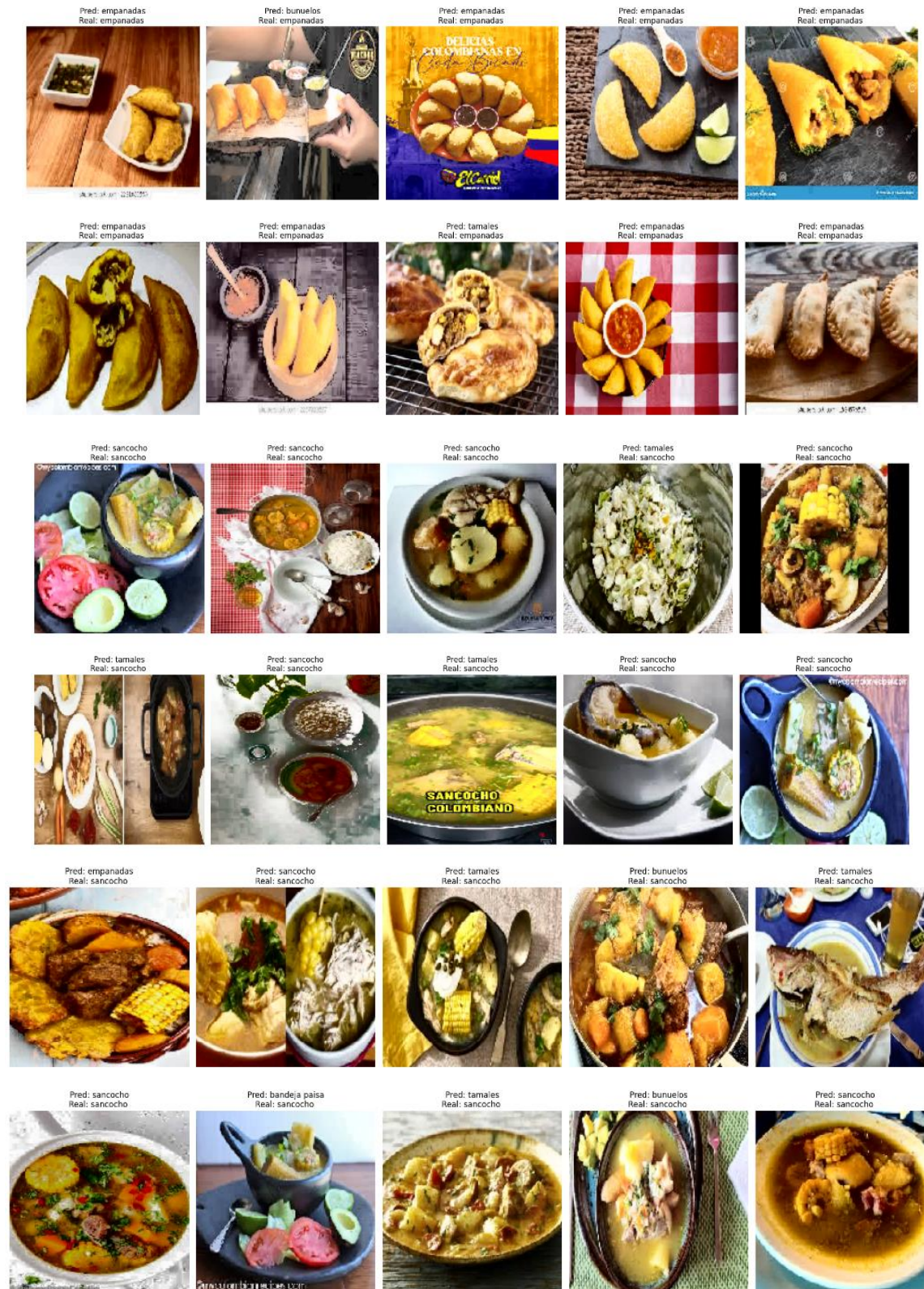


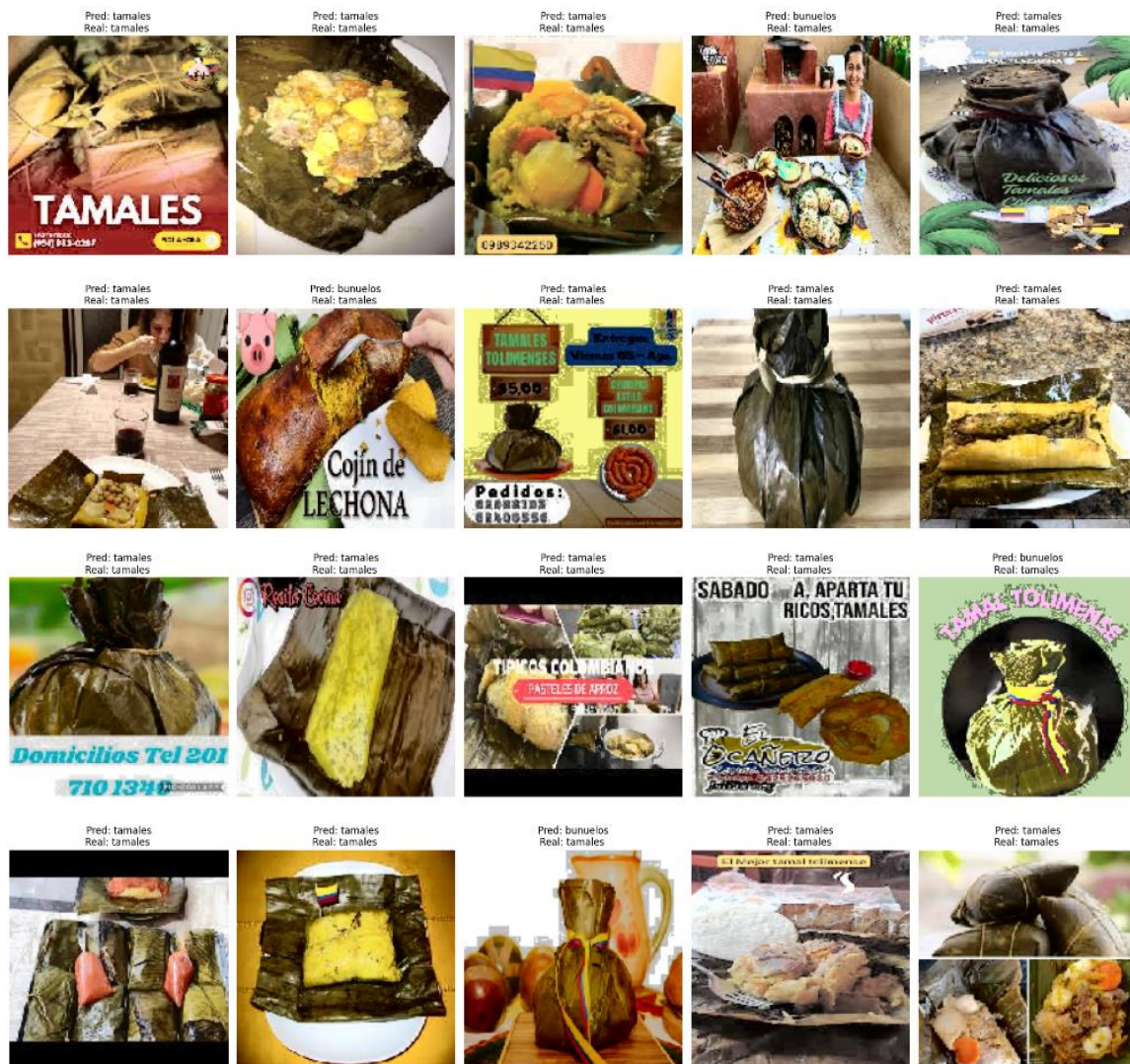
Resultados

Para evaluar el modelo, se generó un script basado en el modelo final y se pusieron a prueba las 100 imágenes que se apartaron al inicio del proyecto. Los resultados de estas pruebas se muestran a continuación:









Como se puede observar, las deficiencias del modelo se deben principalmente a la calidad de los datos de entrenamiento. Estas deficiencias incluyen la presencia de múltiples elementos en una misma imagen, lo cual dificulta la correcta identificación, así como colores, formas y texturas que resultan demasiado similares entre diferentes tipos de alimentos, generando confusión en la clasificación. Además, la superposición de texto en algunas imágenes con la combinación de otros elementos como ollas, imágenes de fondo,

personas, etc, introduce ruido que afecta la precisión del modelo. Estos factores, entre otros, contribuyen a que el modelo no pueda generalizar adecuadamente en algunos casos.

Conclusiones

Este proyecto mostró que la calidad de los datos de entrenamiento es crucial para el éxito de los modelos de aprendizaje automático. Las técnicas de aprendizaje no supervisado como k-means no fueron efectivas debido a la similitud entre las características visuales de las clases. El enfoque supervisado con CNN mejoró los resultados, aunque se enfrentaron problemas de sobreajuste que requirieron varios ajustes para mitigarlos. Finalmente, se alcanzó una precisión global del 82%, sin embargo, aún hay margen para mejorar el modelo. Implementar mejoras adicionales en el código, junto con un aumento en la capacidad de cómputo (dado que el entrenamiento inicial se realizó en una máquina local no apta para aprendizaje automático), podría proporcionar resultados superiores. Además, una mayor cantidad de datos de mejor calidad ayudaría a reducir el ruido presente en el conjunto de entrenamiento, permitiendo que el modelo generalice de forma más robusta. Estos aspectos reflejan el potencial de futuras iteraciones para alcanzar un rendimiento aún más alto y consistente.

Finalmente, esta actividad permitió explorar técnicas de aprendizaje automático, evidenciar la importancia de la calidad de los datos y comprender cómo los ajustes continuos en los modelos impactan significativamente en los resultados.

Lista de referencias

Benítez, R. (2014). Inteligencia artificial avanzada. Barcelona: Editorial UOC.

Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2023). Dive into deep learning. Cambridge University Press.

Hartigan, J. A., & Wong, M. A. (1979). A k-means clustering algorithm. *Applied statistics*, 28(1), 100-108.

Wu, J. (2017). Introduction to convolutional neural networks. National Key Lab for Novel Software Technology. Nanjing University. China, 5(23), 495.