



**Tecnológico
de Monterrey**

***Desarrollo de aplicaciones avanzadas de
ciencias computacionales***

**Modelado de una red neuronal
convolucional para una identificación
binaria: Espresso vs Longblack**

Integrantes

~

Alison Magie Yáñez Dávila

| A01423011

Fecha de entrega:
1 de Junio del 2023

Índice

Contexto	3
Modificaciones de imágenes	4
Modelo base	5
Experimento 1	5
Experimento 2	5
Conclusiones	5
Referencias	6

Contexto

En este proyecto, se desarrollará un modelo de red neuronal convolucional para realizar una clasificación binaria con el objetivo de distinguir entre dos tipos de café: espresso (café extraído en estado puro del grano) y longblack (café expreso mezclado con agua caliente). Para este propósito, se utilizará un dataset creado por "Philippe - Kangaroo.AI", que consta de 400 imágenes recortadas en una proporción de 1:1 y con un tamaño de 256x256 píxeles. Además, se usarán como referencia dos repositorios adicionales donde se ha llevado a cabo esta misma clasificación binaria utilizando el mencionado dataset, lo que proporcionará recursos valiosos para el desarrollo del proyecto.

Modificaciones de imágenes

Con la intención de tener una mayor variedad de imágenes para enriquecer el entrenamiento se utilizó una clase en la biblioteca de Keras de python llamada *ImageDataGenerator* el cual se utiliza para generar imágenes en tiempo real.

Los parámetros que se establecieron para generar las imágenes en tiempo real son:

- Rescala de los valores de los píxeles en 1./255
- Rotación aleatoria de imágenes en 40°
- Rangos de desplazamiento horizontal y verticalmente en 0.2
- Cortes aleatorios en las imágenes en 0.3
- Zoom aleatorio sobre las imágenes en 0.3

Esta librería se utilizó en cada uno de los entrenamientos realizados, por lo que los parámetros establecidos no cambian.

Modelo base

En el modelo base se utilizó una red neuronal convolucional con la arquitectura VGG16, en donde se especificaron los parámetros de pesos pre entrenados en conjunto de datos ImageNet, la especificación de no incluir capas totalmente conectadas al modelo y la forma de entrada de imágenes a la red. A partir de esto se creó un modelo secuencial, en donde se agregó esta arquitectura como una capa.

Posteriormente, en la compilación del modelo se especificaron la función de pérdida, el optimizador y las métricas a utilizar.

Finalmente, utilizando los generadores de datos 'train_generator' y 'val_generator' se realizó el entrenamiento a través del método fit, en el cual se establecieron los valores de 7 pasos por época, 5 pasos en cada época de validación, 30 épocas y un valor de 15 en el batch.

De acuerdo al modelo desarrollado, durante el entrenamiento los valores de precisión y pérdidas mostraron el aprendizaje del modelo con un resultado del 95%, sin embargo, al realizar la comparación con los datos de validación este mostró una deficiencia.

Aprendizaje del modelo base	
Pérdida	0.2465
Precisión	0.9579
Pérdida con datos de validación	18.7595
Precisión con datos de validación	0.5143

Tabla 1: Resultados del entrenamiento del modelo base

Tal como se puede observar en los resultados del entrenamiento, se puede deducir que el modelo está experimentando un comportamiento *underfitting*.

Al visualizar el desempeño mediante una matriz de confusión se obtienen los siguientes resultados:

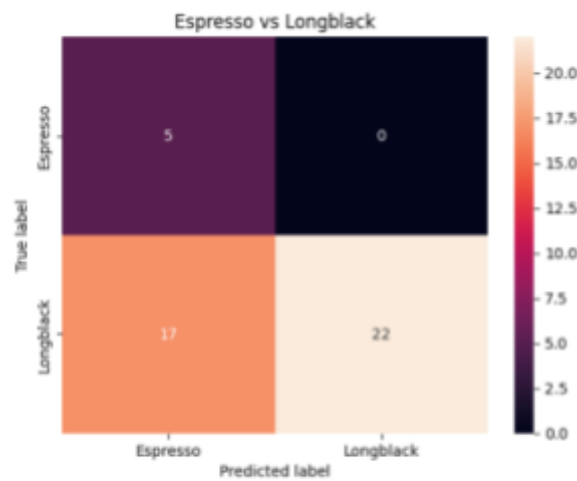


Figura 1: Matriz de confusión del modelo base

Tal como se puede observar, el modelo demuestra solamente tener la capacidad de clasificar correctamente las imágenes en un porcentaje del 61%, teniendo mayor exactitud en la clasificación de cafés longblack.

Con base a lo anterior, se puede concluir que el modelo es demasiado simple para abarcar la problemática de identificación binaria entre un espresso y un longblack, por lo que se llevará a cabo la estrategia de la implementación de hiper parámetros en función con el escrito "Espresso or Longblack?", el cual describe un modelo con 99% de precisión.

Experimento 1

Dentro del experimento 1, tal como se mencionó anteriormente, se llevó a cabo la implementación de hiper parámetros con la intención de aumentar la complejidad del modelo.

Las modificaciones desarrolladas durante este experimento fueron implementadas a partir de la comparativa con el escrito "Espresso or Longblack?" el cual especificaba un valor de $1e-5$ en el learning rate, validation split en 0.2. y 20 en número de épocas.

Además de lo anterior, se realizaron modificaciones en la matriz de confusión para mostrar correctamente la información obtenida durante el entrenamiento debido a que en el código descrito en el modelo el formato implementado era incorrecto.

Una vez ya realizado los cambios los resultados obtenidos fueron los siguientes:

Aprendizaje del experimento 1	
Pérdida	0.4476
Precisión	0.8190
Pérdida con datos de validación	1.7836
Precisión con datos de validación	0.7857

Tabla 2: Resultados del entrenamiento del experimento 1

Durante el experimento 1 en comparación al modelo base, este obtuvo una precisión de 81% durante el entrenamiento, además de que hubo una mejora durante el entrenamiento con los datos de validación.

Sin embargo, durante la evaluación del modelo utilizando un conjunto de datos de testeo, este arrojó un resultado del 90% de precisión.

Al visualizar el desempeño del modelo en la matriz de confusión se puede visualizar la mejora del modelo anterior en la clasificación binaria:

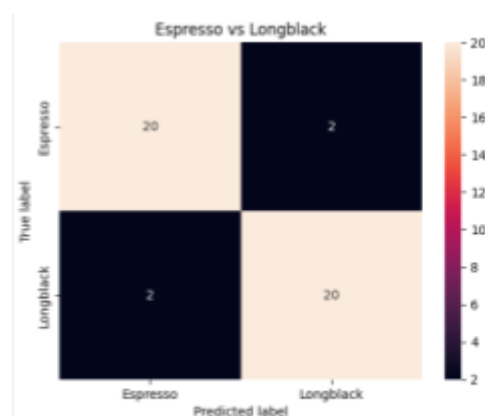


Figura 2: Matriz de confusión del experimento 1

En la matriz de confusión, se puede visualizar que en este nuevo modelo se demuestra una gran mejoría en la identificación binaria, por lo que se puede concluir que a partir de la modificación de hiper parámetros a los valores que se muestran en el escrito "Espresso or Longblack?", este modelo clasifica correctamente los tipos de cafés.

Experimento 2

Con la intención de encontrar una mejora en el modelo anterior, se realizó la modificación de los hiper parámetros descritos en el experimento 1, aumentando el número de épocas a 25, validation steps en 6 y validation split en 0.4. Estos valores se definieron así para garantizar un mayor procesamiento de los datos al aumentar la cantidad de épocas, número de pasos a considerar en cada época de validación y mayor cantidad de datos de entrenamiento que se utilizarán como validación.

Una vez ya realizado los cambios los resultados obtenidos fueron los siguientes:

Aprendizaje del experimento 1	
Pérdida	0.4036
Precisión	0.9338
Pérdida con datos de validación	2.3609
Precisión con datos de validación	0.6714

Tabla 3: Resultados del entrenamiento del experimento 2

En el experimento 2 se obtuvo una precisión de 93% durante el entrenamiento, que en contraposición del experimento 1 este demostró tener una mayor precisión de clasificación binaria durante el entrenamiento.

Sin embargo, durante la evaluación del modelo utilizando un conjunto de datos de testeo, este arrojó un resultado del 75% de precisión.

A través de la matriz de confusión se puede visualizar de forma más real el desempeño obtenido en este modelo:

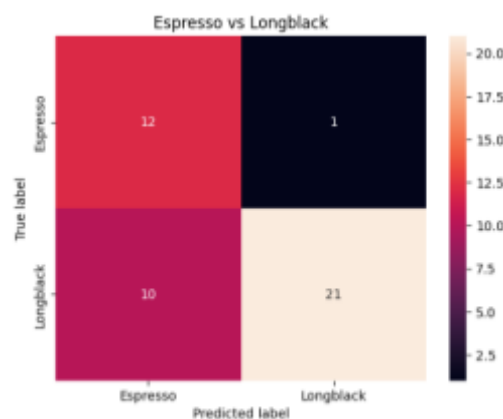


Figura 3: Matriz de confusión del experimento 2

En esta matriz, se puede percibir que la precisión de clasificación de los cafés disminuyó considerablemente. No obstante, al igual que el modelo base, el modelo realizado en el experimento 2 logró aprender a identificar los patrones necesarios para catalogar correctamente cafés de tipo Longblack.

Conclusiones

Basándonos en los resultados de cada experimento, se puede determinar que el modelo funcional que obtuvo una mejor clasificación fue el experimento 1 al ocupar hiper parámetros similares descritos en el repositorio Espresso or Long Black?. Por otra parte, en las salidas del entrenamiento del modelo base y del experimento 2 se puede deducir que obtuvieron un mejor desempeño al identificar cafés de tipo Long Black.

Referencias

Referencias: PHILIPPE. (2021). Coffee images. Kaggle: Your Machine Learning and Data Science Community. <https://www.kaggle.com/datasets/filhypedeeplearning/coffee-images>
stpeteishii. (2021, 16 de noviembre). Coffee Images Classify Torch Conv2d. Kaggle: Your Machine Learning and Data Science Community. <https://www.kaggle.com/code/stpeteishii/coffee-images-classify-torch-conv2d>
filhypedeeplearning. (2021, 24 de noviembre). Espresso or Longblack? Kaggle: Your Machine Learning and Data Science Community. <https://www.kaggle.com/code/filhypedeeplearning/espresso-or-longblack/input>