

Informe de seguimiento FASE 2

Identificación de huellas de calzado a partir de imágenes con redes neuronales convolucionales

Laura Rivera Sánchez

MU Ingeniería Computacional y Matemática

Área de Inteligencia Artificial

Nombre Tutor/a de TFM

Elena Álvarez de la Campa Crespo

Profesor/a responsable de la asignatura

Carles Ventura Royo



Universitat
Oberta
de Catalunya



Fecha Entrega 05/05/2023

Índice

Informe de seguimiento FASE 2.....	i
1. Descripción del avance del proyecto.....	1
2. Relación de las actividades realizadas	1
3. Desviaciones en la temporización.....	3
4. Resultados parciales	3
5. Trabajo restante	5
6. Comentarios y dudas	6

1. Descripción del avance del proyecto

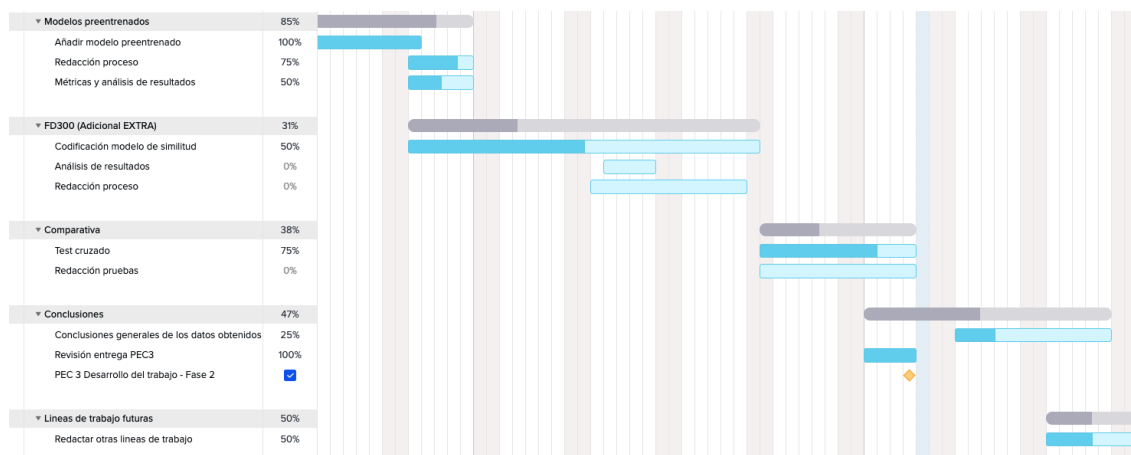
Para esta segunda entrega el objetivo principal era realizar los siguientes puntos:

- Mejorar los resultados del modelo.
- Realizar experimentos.
- Crear una primera versión del modelo de similitud.
- Crear comprobación cruzada.

Se han cumplido todos los objetivos, pero algunos experimentos aún están en proceso.

2. Relación de las actividades realizadas

Según la planificación definitiva, para esta segunda fase se realizarían las siguientes tareas:



Aunque se planificó crear una red neuronal siamesa para la búsqueda de similitud con las huellas de la segunda base de datos FID300, se ha priorizado que funcione correctamente la red neuronal para la predicción de la marca.

Por otro lado, se han solucionado los problemas con el modelo y se consiguen resultados satisfactorios, esto a permitido avanzar en los experimentos previstos.

Tareas realizadas hasta el momento:

- **Investigar** porque el resultado de la red neuronal de clasificación de marca obtiene una eficacia tan baja.
- **Parametrización** de la red neuronal de clasificación de la marca para realizar diferentes experimentos.
- **Investigar** el tipo de resultado que tendría más sentido en este caso, que devuelva la marca o un listado de % de afinidad a las marcas.
- Creación de funciones que devuelven el **top1** y **top3** de las marcas según el array que devuelve la capa *softmax* de la red neuronal.
- Creación de funciones para facilitar la ejecución del modelo con diferentes parámetros.
- Concreción de las **hipótesis** o preguntas a responder con los experimentos:
 - o ¿Qué preprocesado de imágenes funciona mejor?
 - o ¿Utilizar más epochs, resulta siempre en mejor resultado?
 - o ¿Un mayor tamaño de las imágenes utilizadas, resulta en un mejor resultado?
 - o ¿Cuál es el mínimo de muestras por marca aceptable para el modelo?
- Programación del **preprocesado** de las imágenes de huellas de calzado, este incluye:
 - o Aumentación.
 - o Transformación a escala de grises.
 - o Binarización (convertir a blanco y negro).
 - o Recortar los espacios en blanco de alrededor de la imagen (columnas y filas totalmente en blanco).
- Obtención de **resultados satisfactorios** para la detección de marca después de mejorar el preprocesado de las imágenes.
- Redacción de los experimentos en el documento de memoria.
- **Experimentos** con diferentes subconjuntos de datos, para comprobar si el resultado empeora con menos muestras por marca.
- Experimentos con diferentes **tamaños** de imagen.
- División de las tareas del proyecto en diferentes cuadernos para mejor lectura.

A continuación, las tareas **no planificadas** inicialmente pero que se han realizado:

- Al aprender que existen diferencias entre VGG16 y ResNet, se ha incorporado pruebas con ambas redes preentrenadas para ver si existen diferencias en el resultado.
- Experimentos desbloqueando las primeras capas de las redes anteriores.
- Reestructuración del código para añadir funciones y se comprenda mejor.

3. Desviaciones en la temporización

1. Se planificó tener la red neuronal de similitud para el segundo objetivo (adicional) del proyecto para esta segunda fase, pero se ha priorizado mejorar la red neuronal del primer objetivo de detección de marcas ya que se trata del eje central del proyecto.
2. También se planificó tener las redes preentrenadas, pero dado un error que aparecía al hacer la predicción con el subconjunto de test se retrasó ligeramente y está pendiente terminar los experimentos (previstos finalizar la semana del 08/05/2023)

4. Resultados parciales

Tanto el documento con la entrega parcial de la memoria en la que algunos apartados no están en su versión definitiva, como el código realizado están disponibles en el siguiente repositorio GitHub:

https://github.com/laurivsan/TFM_AI_2023

En concreto los ficheros:

[TFM IA MUECIM LauraRivera PEC3.pdf](#)

[TFM LauraRivera objetivo1 marca](#)

[TFM LauraRivera objetivo1 cross validation](#)

Por último, un resumen de los resultados obtenidos en los experimentados realizados sobre la red neuronal de **predicción de marca** con diferentes parámetros para el preprocesado de las imágenes:

Todas las pruebas en redes neuronales se han realizado con epoch=10

Shape	pre	Train				Test	
		Acc	Val_acc	Loss	Val_loss	Top 1	Top 3
(280,832,3)	-	0,30	0,33	1,85	1,85	0,33	0,62
(280,832,3)	A	0,49	0,47	1,30	1,45	0,45	0,59
(280,832,1)	G	0,30	0,33	1,83	1,84	0,33	0,62
(280,832,1)	GB	0,88	0,66	0,38	0,86	0,64	0,88
(280,832,1)	AG	0,58	0,55	1,11	1,14	0,51	0,51
(280,832,1)	GBC	0,97	0,76	0,13	0,84	0,74	0,97
(280,832,1)	AGBC	0,34	0,39	1,68	1,6	0,38	0,61

Los diferentes procesos aplicador en las imágenes se especifican con las siguientes siglas para simplificar:

A	Aumentación
G	Convertir a escala de grises
B	Convertir a blanco y negro (Binary)
C	Recortar filas y columnas en blanco

En estos experimentos se puede comprobar como el **preprocesado de las imágenes es muy importante** para obtener mejores resultados, ya que la eficacia aumenta considerablemente al convertir las imágenes a blanco y negro, por lo que el modelo aprende del patrón y no del color o tonalidad de gris.

El resto de las pruebas se encuentran en el documento de la memoria. Falta por finalizar algunos experimentos para completar todas las tablas de resultados, añadir las conclusiones y redactar algunos apartados.

5. Trabajo restante

- **Más experimentos** con la red neuronal de detección de marca de calzado, con diferentes tamaños de las imágenes y diferentes criterios en el momento de filtrar las muestras según el número de muestras por marca.
- Terminar y experimentar con el modelo de detección de imágenes similares con el conjunto de datos FD-300 correspondiente al Objetivo 2: Similitud FD-300.
- **Finalizar las pruebas cruzadas** en otro Notebook y documentar.
- **Completar la redacción** de la memoria con descripciones y explicaciones.
- Crear **presentación y video** para la entrega final.
- Redacción del resumen del trabajo de la primera página.
- Revisar formato de las referencias.

6. Comentarios y dudas

- En la memoria se menciona train, test y validación. Al tratarse de la documentación en español, ¿debería traducirlo a “entrenamiento”, “pruebas” y “validación”?
- Se ha inicializado el desarrollo de pruebas cruzadas. Se entrena el modelo con todos los datos de la base de datos que contiene la etiqueta de marca y posteriormente realiza predicciones con 10 imágenes de la otra base de datos FD300. ¿Cómo se mide la eficacia en este caso? La segunda base de datos no tiene la información de marca. ¿Realizando una comprobación visual? Por ejemplo, imprimiendo algunos ejemplos y su predicción.
- En las pruebas cruzadas se ha visto que siempre devuelve la marca “Nike”, la que más aparece en el conjunto de datos de entrenamiento. Las imágenes son menos nítidas en la segunda base de datos y eso podría afectar a la clasificación. Seguramente haga falta otro tipo de preprocesado en esta segunda base de datos en trabajo futuro para que las pruebas cruzadas funcionen.
- ¿Debería especificar en la memoria las librerías que utiliza el código?