



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



Máster en Inteligencia Artificial, Reconocimiento de Formas e Imagen Digital  
Universitat Politècnica de València

## Reconocimiento de Género

Visión por Computador

*Autor:* Juan Antonio López Ramírez

Curso 2019-2020



## Introducción

---

En esta memoria se va a explicar el procedimiento seguido para la correcta realización del ejercicio de reconocimiento de género a partir de un dataset de imágenes mediante técnicas de Redes Neuronales Convolucionales.

Para ello, se han implementado diferentes Residual Networks para alcanzar los objetivos propuestos, que son:

- Superar el 95 % de precisión en el conjunto de test.
- Superar el 90 % de precisión en el conjunto de test utilizando menos de 100K parámetros.

El dataset empleado es *LFW*, que contiene 13233 imágenes de caras (cada una de 100x100 píxeles en formato RGB). Usaremos 10585 imágenes para entrenamiento y 2648 para test.



**Figura 1:** *Labeled Faces in the Wild Home*, dataset empleado para el reconocimiento de género.

## Implementación del modelo para 95 % de precisión

---

Para empezar, hemos partido del modelo de *ResNet* que hemos visto en las clases de teoría de la asignatura.

En nuestro caso, el modelo está formado por cuatro bloques convolucionales y un *deep fully connected*.

Cada bloque tiene la mitad de convoluciones que el siguiente, componiendo la red unidades convolucionales de 1, 2, 4 y 8.

Cada uno de estos cuenta con una capa de Max Pooling de 2x2, un operador de 32 filtros, una capa de ruido Gaussiano del 30 % y otra de *Batch Normalization*.

Los bloques de 2, 4 y 8 unidades tiene, en adición a lo anterior, otro operador convolucional, también de 32 filtros; una función de activación ReLu y un operador de adición entre el filtro anterior y el final de la unidad básica.

Finalmente, la *fully connected* cuenta con 512 neuronas, a la que le aplicamos, igual que las convolucionales, un *Batch Normalization* y una capa de ruido Gaussiano del 30 %. Por último, le añadimos la función de activación *softmax*.

El *Data Augmentation* aplicado ha consistido en modificar las muestras en un 20 % de altura y anchura, además de añadirle un rango de rotación de 20 grados y un zoom del 20 %.

## Implementación del modelo con menos de 100K parámetros

Para este modelo, hemos partido del anterior y se ha reducido a la mitad el número de bloques convolucionales (hemos pasado de tener 4 *conv blocks* a solo 2), junto con los filtros aplicados en los operadores de esas convoluciones, que han pasado de 32 a 8.

De esta manera conseguimos un modelo muy similar al anterior pero que es más liviano debido a sus 89090 parámetros (véase la figura 2), pero que sigue siendo capaz de realizar la tarea de reconocimiento de género.

```
=====
Total params: 89,090
Trainable params: 88,930
Non-trainable params: 160
```

**Figura 2:** Se ha hecho un *summary* al modelo y se puede apreciar que el número de parámetros es inferior a 100K.

## Resultados

Cuando ya tenemos nuestros modelos de *ResNet*, se ha ido modificando el tamaño del *batch* y las *epochs* del training. Los resultados obtenidos han sido los siguientes:

Tamaño del batch	Epochs	Resultados gender95 (%)	Resultados gender90-100K (%)
100	100	95.85	92.82
100	200	<b>95.96</b>	<b>94.15</b>
100	300	95.81	94.11
150	100	94.22	92.90
150	200	95.54	91.92
150	300	94.41	93.09

Por un lado, en los experimentos de *gender95*, salvo en el cuarto y sexto, superamos la precisión que se pedía en el enunciado del problema, siendo el mejor resultado un **95.96 %**, obtenido con un tamaño de *batch* de 100 y 200 *epochs*.

Por otro lado, en *gender90-100K*, con la misma configuración que el anterior modelo, obtenemos como mejor resultado un **94.15 %**, superando la barrera del 90 %.