



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



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

Implementaciones básicas

Visión por Computador

Autor: Juan Antonio López Ramírez

Curso 2019-2020

Hemos partido del trabajo realizado en la asignatura de Redes Neuronales Artificiales y, concretamente, del modelo de Red Convolutiva para la tarea de CIFAR.

El *Data Augmentation* que se aplica es de un 20 % de altura y de anchura, junto con un rango de rotación de 20 grados y un zoom también del 20 %.

Cada bloque convolutivo cuenta con un *Batch Normalization*, un ruido gaussiano del 30 % y una función de activación ReLu.

El *learning rate annealing* está planificado para que el factor de aprendizaje varía en la *epoch* 75 (pasando a valer 0.01) y en la 125 (pasando a valer 0.001).

Finalmente, creamos un método al que llamamos *vgg_model()* y cuyo parámetro de entrada es el modelo de VGG que se desea utilizar. Por ejemplo, para usar el modelo A (VGG11), es necesario una línea de código similar a la siguiente:

```
1 model = VGG_model(11)
```

De esta forma, implementamos los modelos A, B, D y E de VGG (véase la figura 1). Debido a que el modelo C es muy similar al D, salvo que algunas convoluciones aplican un *kernel_size* de 1x1 en lugar de 3x3, se ha decidido omitir su implementación.

En nuestro caso, los resultados que se han obtenido han sido los siguientes:

Modelo de VGG	Resultados (%)
-	90.81
11	90.87
13	92.55
16	92.51
19	92.75

La primera fila representa el modelo que se usó en la práctica de CIFAR de la asignatura de Redes Neuronales Artificiales, sin emplear ningún tipo de VGG.

Como se puede apreciar, con cualquier modelo de VGG superamos la precisión del original, obteniendo el mejor resultado con VGG19, que es de **92.75 %**.

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224×224 RGB image)					
conv3-64	conv3-64 LRN	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 conv1-256	conv3-256 conv3-256 conv3-256	conv3-256 conv3-256 conv3-256 conv3-256
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 conv1-512	conv3-512 conv3-512 conv3-512	conv3-512 conv3-512 conv3-512 conv3-512
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Figura 1: Modelos de VGG. Se han implementado el A, B, D y E.