

## Examen de Problemas, Computer Vision. 1 de Junio de 2021

Puntuación: 3 puntos

Duración: 40 minutos

Nombre:

Apellidos:

### Problema 1 (1 punto)

Si empleamos la red VGG-19 como la que vemos en la tabla (E) y asumimos que todas las convoluciones son con padding.

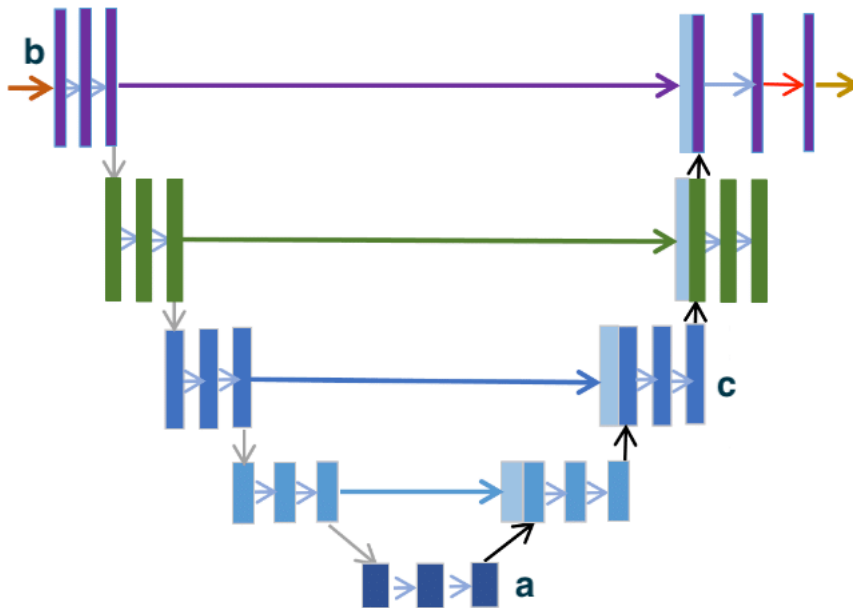
ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input ( $224 \times 224$ RGB image)					
conv3-64	conv3-64 <b>LRN</b>	conv3-64 <b>conv3-64</b>	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool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

Indica, sin tener en cuenta la parte Fully Connected:

- Cuál sería el tamaño mínimo de imagen que admitiría
- Calcula el número de parámetros

## Problema 2 (1 punto)































Queremos diseñar una U-Net donde el encoder sea una **vgg-19 pre-entrenada** con ImageNet. O sea después de la última **conv3-512** empezaría el decoder. Esta U-Net la vamos a emplear para segmentar la columna vertebral en radiografías. Las radiografías tienen el siguiente tamaño 256x256 y una profundidad de 1 dado que son en escala de grises. Recordad que en la subida (decoder) la profundidad de los mapas se divide por 2 en cada salto.



- Cuál sería el tamaño del mapa obtenido justo antes del decoder (a)
- Qué tendríamos que hacer antes de introducir nuestra radiografía (b)
- Cuál sería el tamaño del mapa en este punto (c)
- ¿ Qué error veis en la figura ?

### Problema 3 (0.5 puntos)

Queremos diseñar una red convolucional para poder hacer reconocimiento de caracteres del alfabeto Cuneiforme:

Alfabeto cuneiforme de Ugarit					
 'a	 b	 g	 h	 d	 h
 w	 z	 h	 t	 y	 k
 š	 l	 m	 d	 n	 z
 s	 c	 p	 š	 q	 r
 t	 g	 t	 i	 u	 s <sub>2</sub>

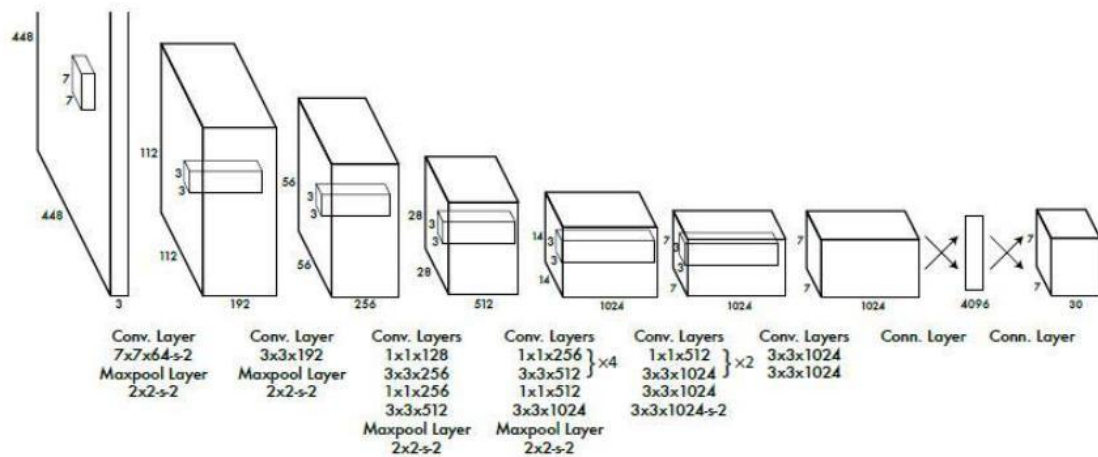
Después de diseñar la red neuronal implementamos el siguiente Data Generator para hacer data augmentation:

```
datagen=ImageDataGenerator(  
    rotation_range=10,  
    width_shift_range=0.2,  
    height_shift_range=0.2,  
    zoom_range=0.5,  
    horizontal_flip=True,  
    vertical_flip=False  
)
```

Indica qué problemas podríamos tener al emplear dicho data generator.

## Problema 4 (0.5 puntos)

Indica qué modificaciones tendríamos que realizar a esta topología Yolov1 si quisiéramos detectar 5 tipos de objetos diferentes en un grid de 10x10 con la posibilidad de tener 3 bounding box por celda.



Señala sobre la figura los cambios y descríbelos.