

Desafío 10

**Segmentación de imágenes de carreteras
mediante redes semánticas**

Iyán Álvarez



Universidad
del País Vasco



Euskal Herriko
Unibertsitatea

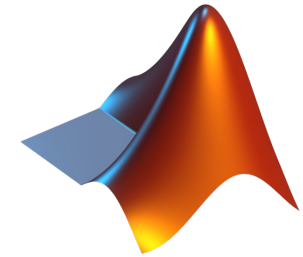
Trabajo realizado

Dos script en Matlab.

El script “SemanticaCarreteras_main” carga el conjunto de imágenes y segmentación, las muestrea, se entrena una red semántica y se evalúan y exponen los resultados obtenidos. Finalmente, se reporta el accuracy obtenido.

El script “UNetCarreteras_main” carga el conjunto de imágenes y segmentación, las muestrea, se entrena una red UNet y se evalúan y exponen los resultados obtenidos. Finalmente, se reporta el accuracy obtenido.

 SemanticaCarreteras_main.m
 UNetCarreteras_main.m



Datos

El conjunto de datos proporcionado esta compuesto por 18 imágenes satélite para la detección automática de carreteras mediante técnicas de visión por computador y aprendizaje automático.



Arquitectura de la red semántica

Una red neuronal semántica es un tipo de arquitectura de red neuronal convolucional (CNN) diseñada específicamente para la tarea de segmentación semántica en imágenes. La segmentación semántica implica asignar una etiqueta a cada píxel de una imagen, lo que permite identificar y clasificar cada región de la imagen en función de su contenido semántico.

- Capa de entrada que acepta imágenes RGB de tamaño 512x512
- Capas de Submuestreo (Downsampling)
- Capas de Sobremuestreo (Upsampling)
- Capa Final compuesta por
 - Capa de convolución 1x1 para asignar a cada píxel una de las 2 clases posibles.
 - Capa de activación softmax para convertir las salidas en una distribución de probabilidad.
 - Capa de segmentación de píxeles para asignar la clase con la probabilidad más alta a cada píxel.

```
inputSize = [512 512 3];
imgLayer = imageInputLayer(inputSize);
...
downsamplingLayers = [
    conv
    relu
    maxPoolDownsample2x
    conv
    relu
    maxPoolDownsample2x
];
...
upsamplingLayers = [
    transposedConvUpsample2x
    relu
    transposedConvUpsample2x
    relu
];
...
finalLayers = [
    conv1x1
    softmaxLayer()
    pixelClassificationLayer()
];

net = [
    imgLayer
    downsamplingLayers
    upsamplingLayers
    finalLayers
];
```

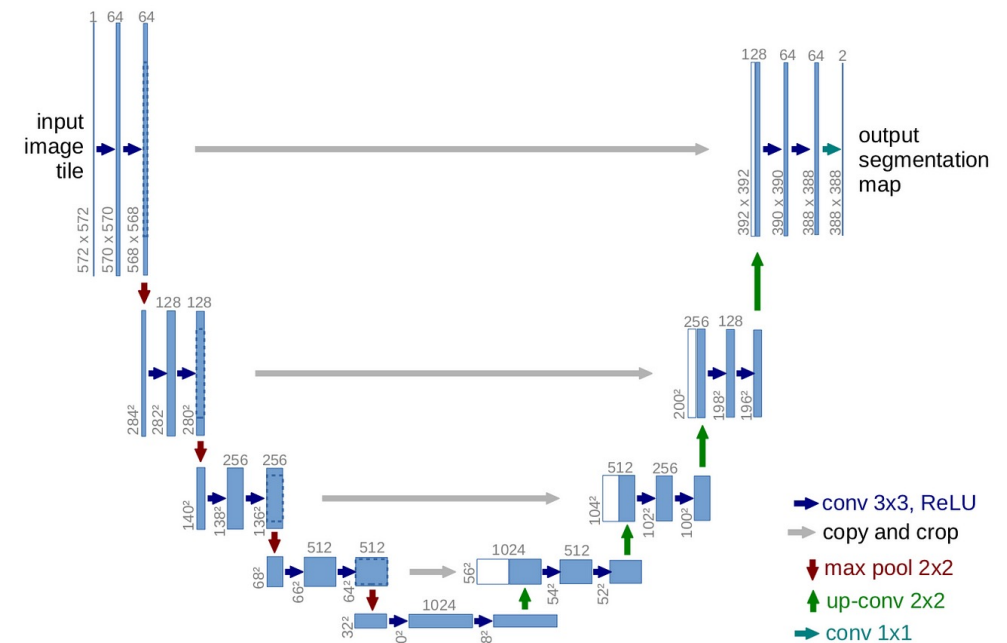
Arquitectura de la red UNet

UNet es una arquitectura de red neuronal convolucional (CNN) ampliamente utilizada para tareas de segmentación de imágenes. Esta diseñada originalmente para la segmentación de imágenes biomédicas, pero es aplicable a diversas aplicaciones de visión por computador.

Su nombre proviene de su forma en U, que se asemeja a la estructura de su arquitectura.

- Codificador (Encoder)
- Decodificador (Decoder)
- Conexiones Residuales

```
input_size = [512 512 3];  
num_classes = 2;  
encoder_depth = 3;  
lgraph = unetLayers(input_size,num_classes,  
    'EncoderDepth',encoder_depth);
```



Entrenamiento de la red semántica y UNet

Para el proceso de entrenamiento se ha realizado la extracción aleatoria de 10 parches de cada imagen tamaño 512x512 para aumentar la variabilidad.

La opción propuesta de utilizar 1000 parches de 64x64 de cada imagen no me parecía adecuada para poder capturar las carreteras correctamente.

```
options = trainingOptions('sgdm',...  
    'InitialLearnRate', 0.05, ...  
    'Momentum', 0.9,...  
    'L2Regularization', 0.0001,...  
    'MaxEpochs', 10,...  
    'MiniBatchSize', 32,...  
    'LearnRateSchedule', 'piecewise',...  
    'Shuffle', 'every-epoch',...  
    'GradientThresholdMethod', 'l2norm',...  
    'GradientThreshold', 0.05, ...  
    'Plots', 'training-progress', ...  
    'Verbose', 0);
```

No se ha realizado el entrenamiento ya que necesitaba alrededor de 10 horas cada red y el hardware disponible no era suficiente.



A pesar de esto, se han analizado las arquitecturas de la red y se ha comprobado el funcionamiento de los scripts.