

Desafío 10bis

**Segmentación de imágenes metalúrgicas
mediante redes semánticas**

Iyán Álvarez



Universidad
del País Vasco

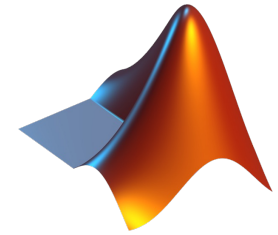
Euskal Herriko
Unibertsitatea

Trabajo realizado

Un script en Matlab.

El script “SemanticaMetalurgicas_main” carga los datos y los convierte al tamaño 700x1024, carga el conjunto de imágenes y segmentación, las muestrea y se entrena una red semántica. Finalmente, se reporta el accuracy obtenido.

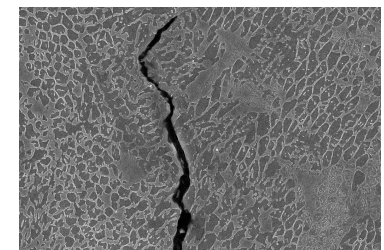
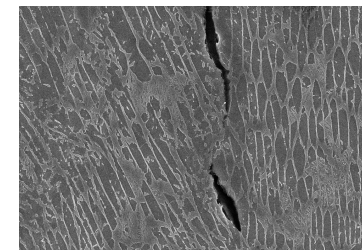
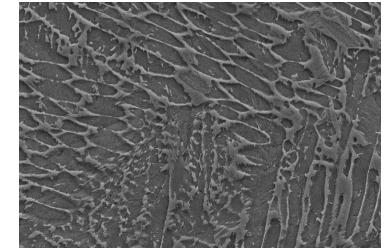
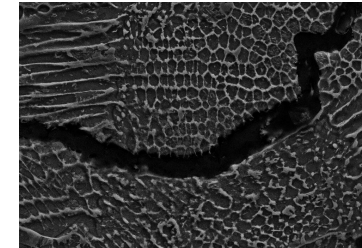
 SemanticaMetalurgicas_main.m



Datos

El dataset MetalDAM-DASCI es un conjunto de datos diseñado específicamente para la detección automática de grietas en estructuras metálicas mediante técnicas de visión por computadora y aprendizaje profundo. Las imágenes están etiquetadas con información sobre la ubicación y extensión de las grietas.

Se ha adaptado el tamaño de las imágenes.



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 de tamaño 64x64 en escala de grises
- Capas de Submuestreo (Downsampling)
- Capas de Sobremuestreo (Upsampling)
- Capa Final compuesta por
 - Capa de convolución 1x1 para asignar a cada píxel una de las 5 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 = [64 64 1];
imgLayer = imageInputLayer(inputSize);
...
downsamplingLayers = [
    conv
    relu
    maxPoolDownsample2x
    conv
    relu
    maxPoolDownsample2x
];
...
upsamplingLayers = [
    transposedConvUpsample2x
    relu
    transposedConvUpsample2x
    relu
];
...
finalLayers = [
    conv1x1
    softmaxLayer()
    pixelClassificationLayer()
];

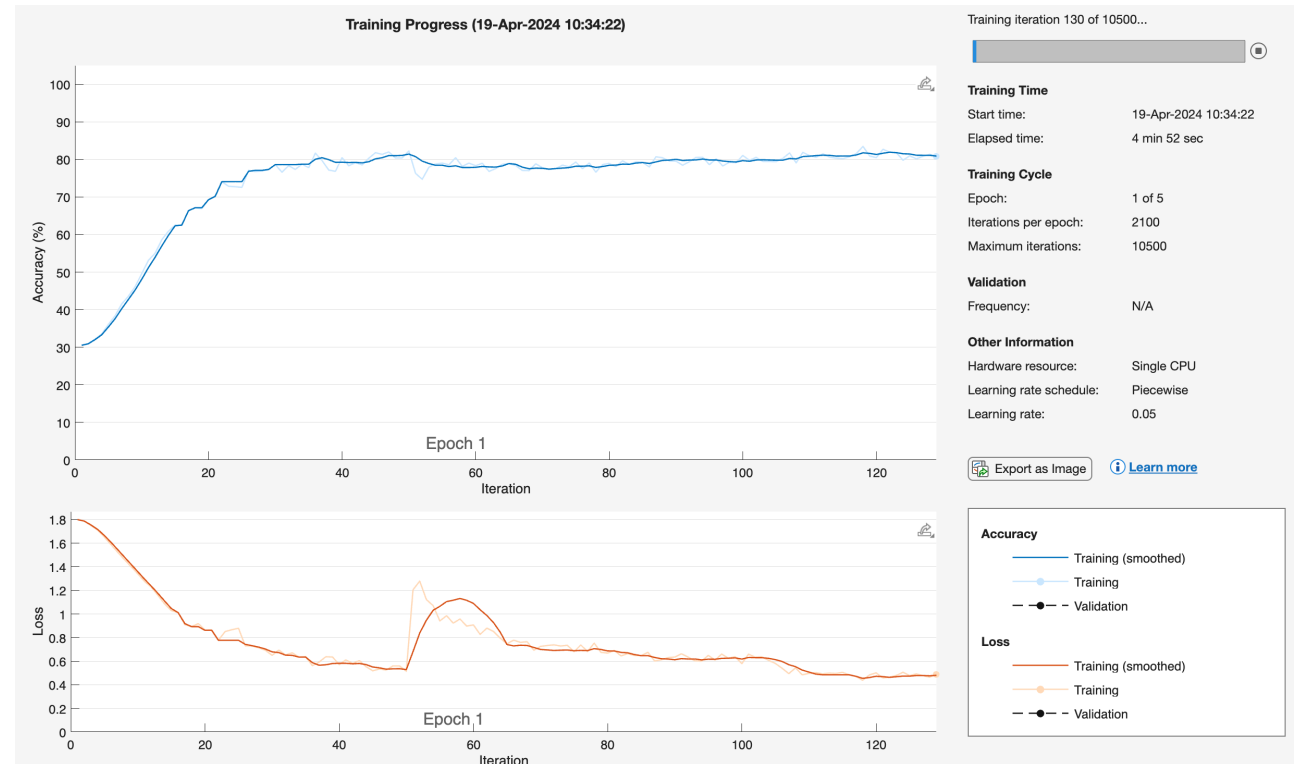
net = [
    imgLayer
    downsamplingLayers
    upsamplingLayers
    finalLayers
];
```

Entrenamiento de la red semántica

Para el proceso de entrenamiento se ha realizado la extracción aleatoria de parches de tamaño 64x64 de las imágenes de entrenamiento para aumentar la variabilidad.

No se ha finalizado el entrenamiento ya que necesitaba alrededor de 6 horas.

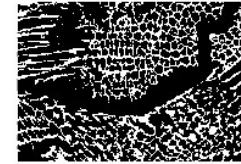
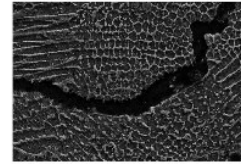
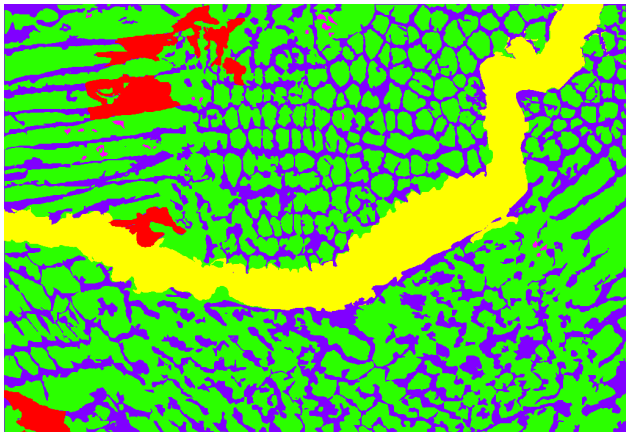
```
options = trainingOptions('sgdm',...  
    'InitialLearnRate', 0.05, ...  
    'Momentum', 0.9,...  
    'L2Regularization', 0.0001,...  
    'MaxEpochs', 5,...  
    'MiniBatchSize', 320,...  
    'LearnRateSchedule', 'piecewise',...  
    'Shuffle', 'every-epoch',...  
    'GradientThresholdMethod', 'l2norm',...  
    'GradientThreshold', 0.05, ...  
    'Plots', 'training-progress', ...  
    'Verbose', 0);
```



Evaluación de la red semántica

La accuracy promedio obtenida durante el entrenamiento es del 80%.

La red semántica es capaz de encontrar algunas clases con facilidad debido a la distinción clara en las características visuales entre esas clases, mientras que otras las confunde o no las detecta.



morado

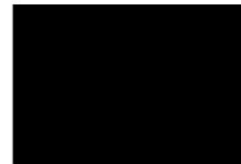
verde



rojo

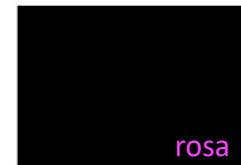
rosa

amarillo



morado

verde



rojo

rosa

amarillo