

# Segmentación semántica de imágenes satelitales

Visión por computadora II

Especialización en inteligencia artificial

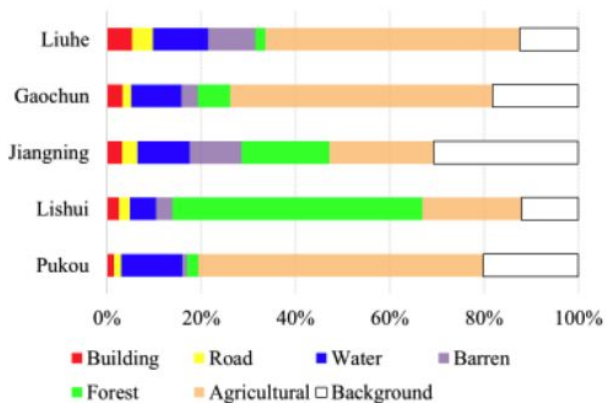
Edgar David Guarín Castro - Tatiana Arenas Suárez

# Desafío a abordar

Implementar algunos algoritmos de *segmentación semántica* en imágenes satelitales de áreas rurales para categorizar y etiquetar siete clases: fondo, carretera, edificio, bosque, agua, agricultura, y terreno estéril.

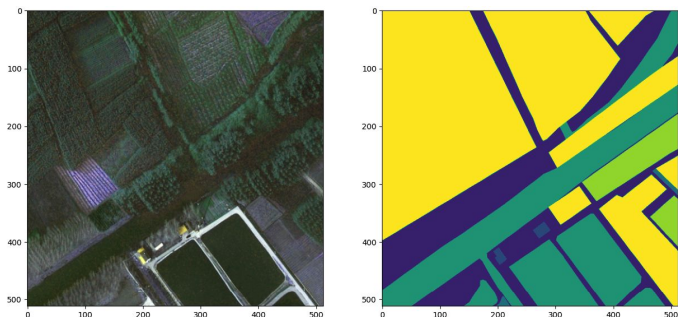


# Datos para la segmentación semántica adaptativa del dominio de la cobertura del suelo



LoveDA es un conjunto de datos para la segmentación semántica en teledetección de la cubierta terrestre capturadas en china.

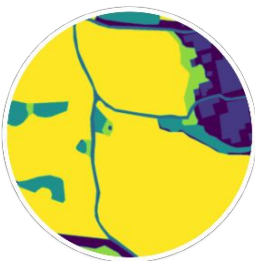
- **Tamaño de imagen:** 1024 X 1024 pixeles en png en 3 canales.
- **Resolución de imagen:** 0.3m por pixel.
- **Tamaño del dataset:** Zona rurales 2358 imágenes, zonas urbanas 1833.



# Pasos generales del modelo



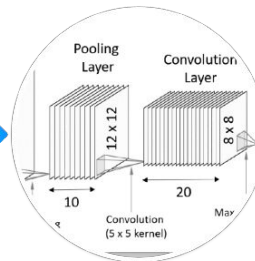
Cambio de tamaño  
512x512 px



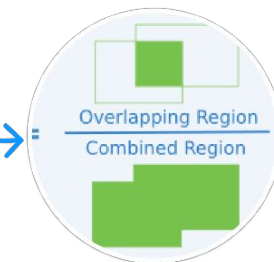
Asignación de  
etiquetas por  
clases



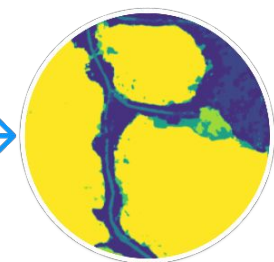
Transformación de  
características



Train/test split  
Construcción del  
modelo



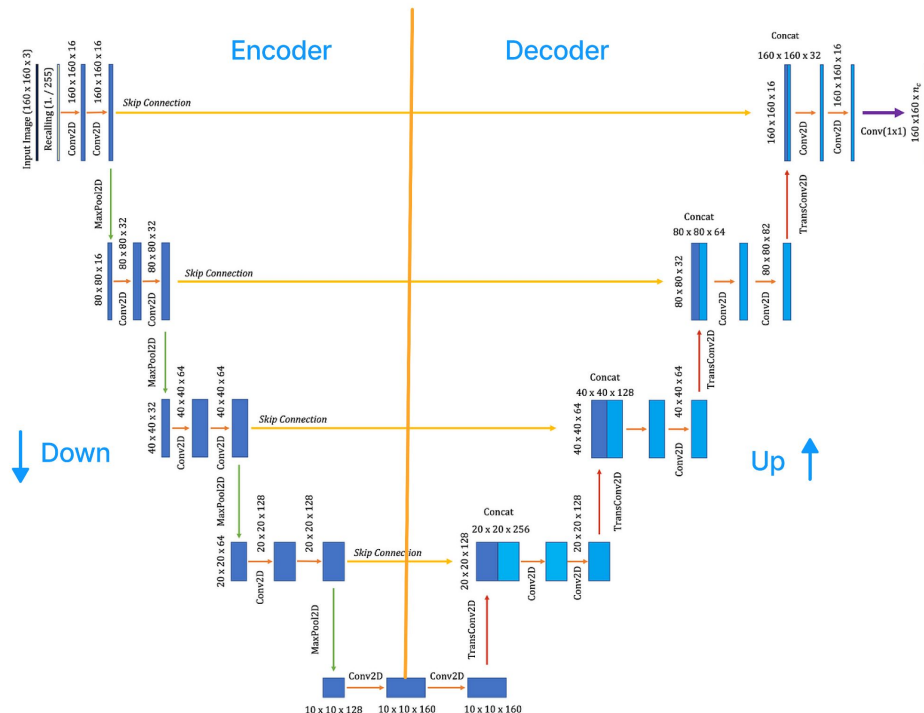
Ajuste y evaluación  
del modelo  
ACC, IoU.



Evaluación de la  
predicción

# U-NET : Arquitectura

Las **convoluciones** se combinan con operaciones de **downsampling** (**max-pooling**) para reducir las dimensiones espaciales mientras aumentan el número de canales.

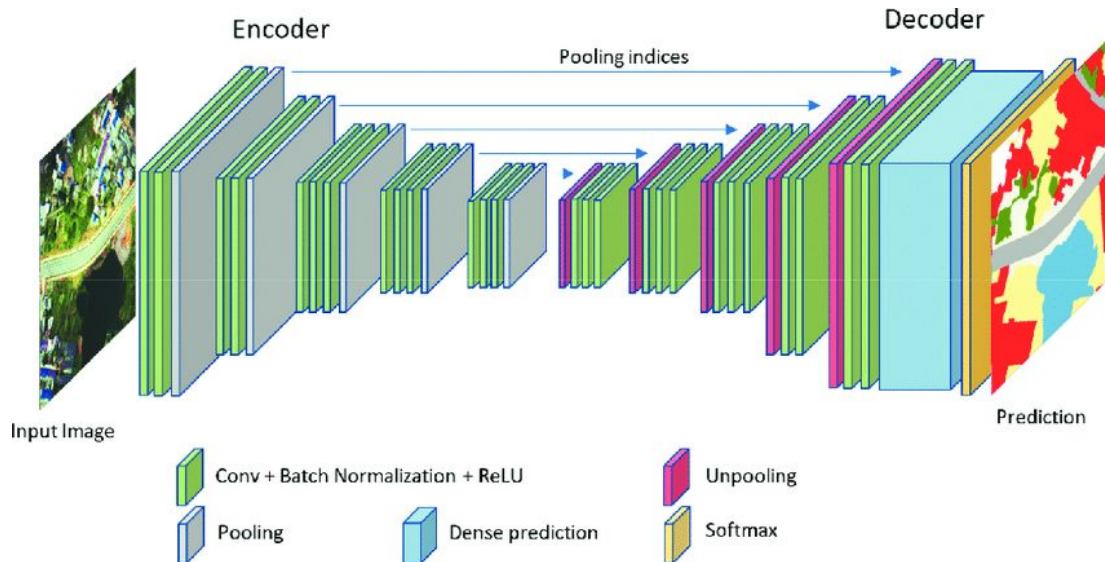


Se usan operaciones de **upsampling** junto con **convoluciones** para reconstruir la imagen en su resolución original (**TransCon**), pero ahora con la información semántica aprendida.

Las **skip connections** permiten transferir características espaciales finas desde el encoder al decoder.

# SegNet : Arquitectura

La información espacial se recupera a partir de los **índices del max-pooling**



El decoder realiza operaciones de **upsampling** utilizando los índices de las operaciones de max-pooling en el encoder para restaurar las características a la resolución original. Este proceso se llama **unpooling**.

# Entrenamiento de los modelos



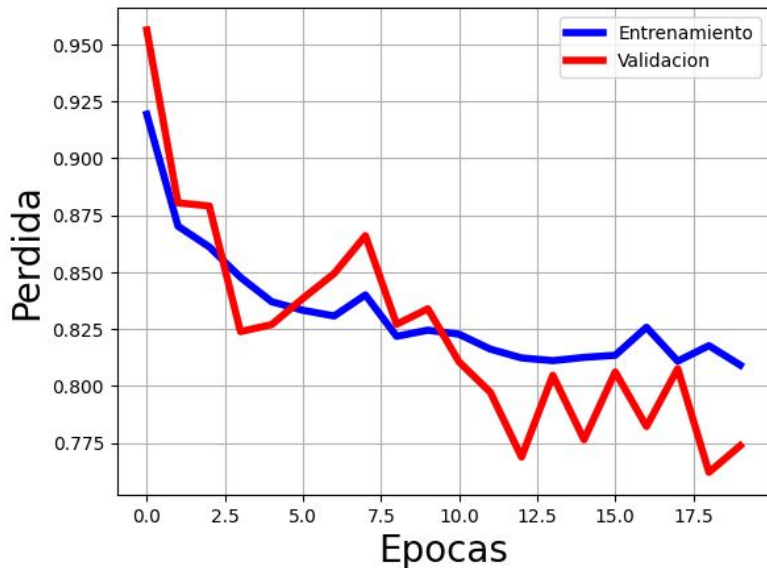
- ✓ Fueron usadas **201** imágenes y sus máscaras (train), **20** imágenes de prueba (test).
- ✓ **4** imágenes por lote de entrenamiento y prueba (batch size).
- ✓ Se entrenó con un total de **20** épocas
- ✓ Código base escrito en Python con **Pytorch**.
- ✓ Entrenamientos realizados en **Colab** con **GPU**.
- ✓ Variación de parámetros `step_size`, `gamma`, `bilinear`.

```
epoch 10 - loss : 0.88068 - acc : 0.44 - val loss : 0.86982 ·  
epoch 11 - loss : 0.87425 - acc : 0.46 - val loss : 0.86375 ·  
epoch 12 - loss : 0.87555 - acc : 0.44 - val loss : 0.86855 ·  
epoch 13 - loss : 0.87269 - acc : 0.45 - val loss : 0.88825 ·  
epoch 14 - loss : 0.87385 - acc : 0.44 - val loss : 0.85737 ·  
epoch 15 - loss : 0.86976 - acc : 0.45 - val loss : 0.85994 ·  
epoch 16 - loss : 0.87114 - acc : 0.45 - val loss : 0.85213 ·  
epoch 17 - loss : 0.87386 - acc : 0.45 - val loss : 0.86503 ·  
epoch 18 - loss : 0.87388 - acc : 0.44 - val loss : 0.85751 ·  
epoch 19 - loss : 0.85857 - acc : 0.47 - val loss : 0.84187 ·
```

# Métricas y funciones de pérdida

## U-Net

Perdida mIoU

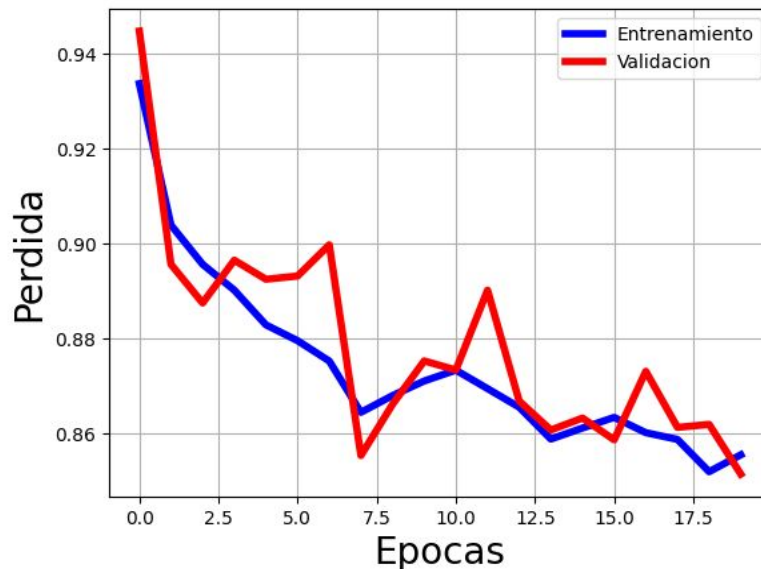


Train Acc = 0.53 | Val Acc = 0.61

**Scheduler:** StepLR → step\_size=2, gamma=0.6

## SegNet

Perdida mIoU

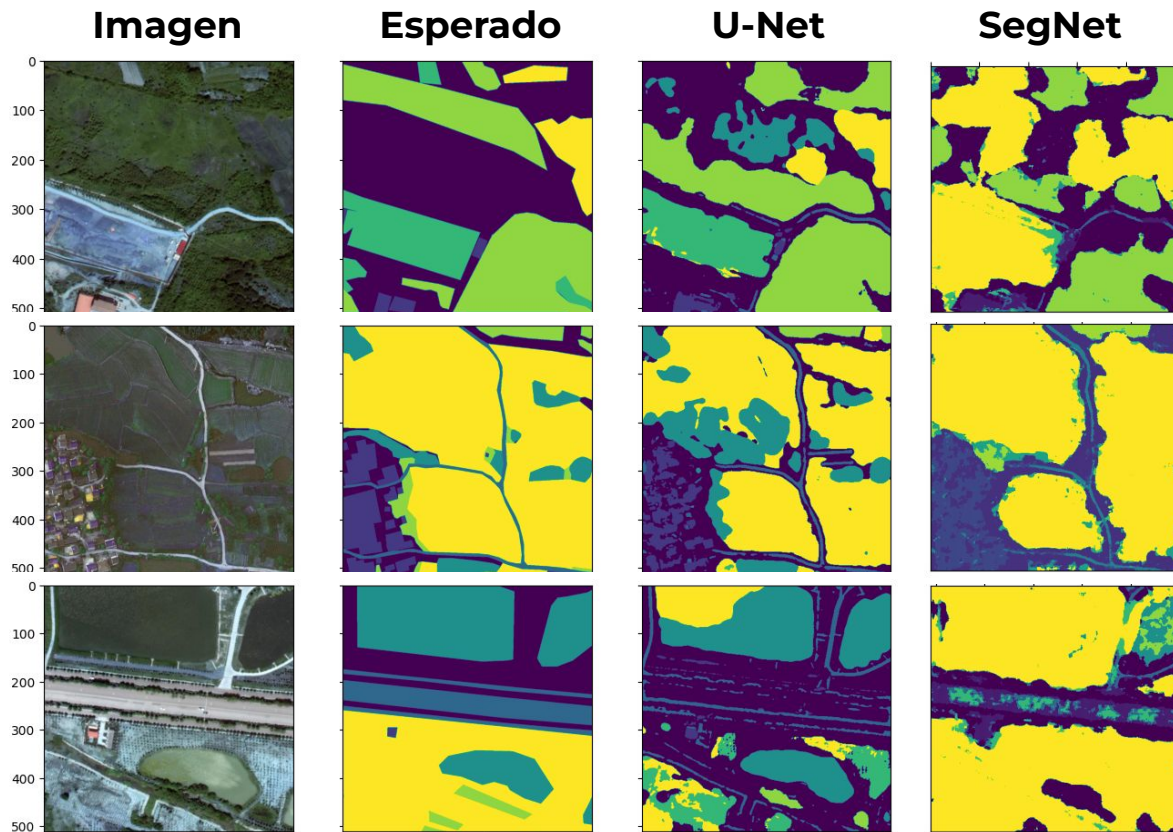


Train Acc = 0.49 | Val Acc = 0.50

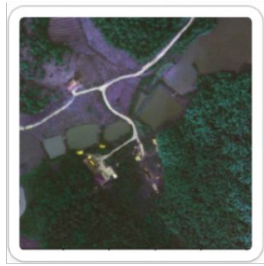
**Scheduler:** StepLR → step\_size=2, gamma=0.8



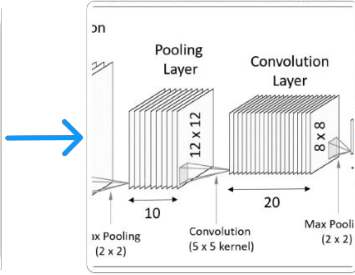
# Resultados



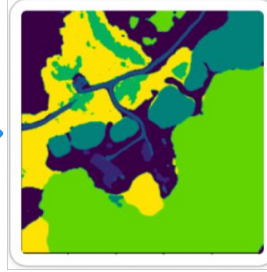
# Predicción y despliegue



Mapbox raster  
Api, 512x512  
1m/px



Se segmenta la  
imagen  
con el modelo  
seleccionado



Se genera  
la máscara  
de segmentación

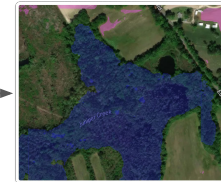


Los contornos  
se transforman  
archivos Geojson



Los resultados  
son desplegados  
en Mapbox

Mapbox predicciones de  
Bosque y terreno estéril



# Referencias

- Junjue Wang and Zhuo Zheng and Ailong Ma and Xiaoyan Lu and Yanfei Zhong, LoveDA: A Remote Sensing Land-Cover Dataset for Domain Adaptive Semantic Segmentation. 2021. <https://doi.org/10.5281/zenodo.5706578>