

Karlsruhe Institute of Technology and Toyota
Technological Institute at Chicago



- Abraham Rodriguez
- Cecilia Sanchez

DESCRIPCIÓN DEL PROBLEMA

Los vehículos autónomos necesitan un sistema de percepción de objetos y de calles en forma confiable para su funcionamiento.

Nuestro objetivo en base al dataset KITTI es entrenar diferentes modelos de segmentación para detectar las calles en las imágenes.

El dataset se encuentra en:

<https://www.kaggle.com/datasets/sakshaymahna/kittiroadsegmentation>



SOLUCION PROPUESTA

- Entrenar un modelo que permita identificar las calles en las imágenes.
- La métrica que se utilizará para evaluar los modelos sera Dice Loss.
- La solución debera devolver una máscara que corresponda con el área de la calle en la imagen de entrada.



CONJUNTO DE DATOS

- El dataset está compuesto por 289 imágenes RGB con una resolución de 1242x375 que llevamos a un tamaño de 128 x 128.
- Dividimos el dataset en entrenamiento 80%, validación 10% y test 10%.

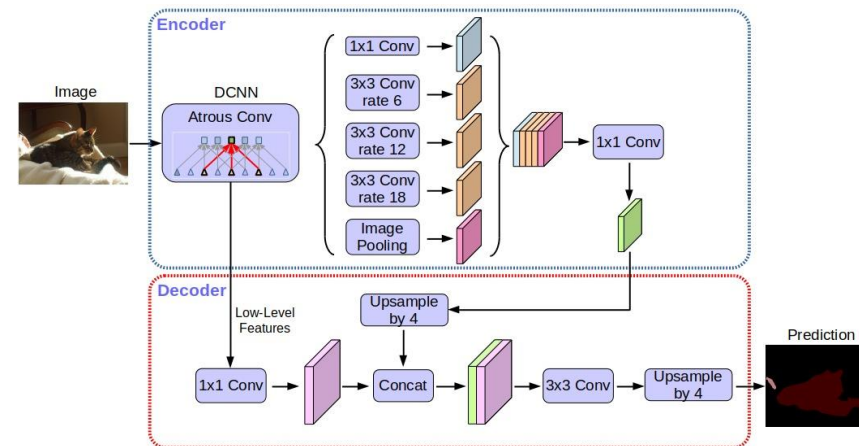
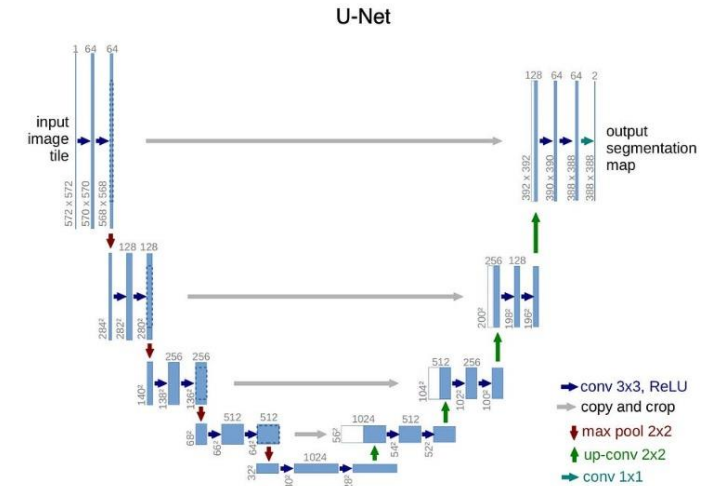
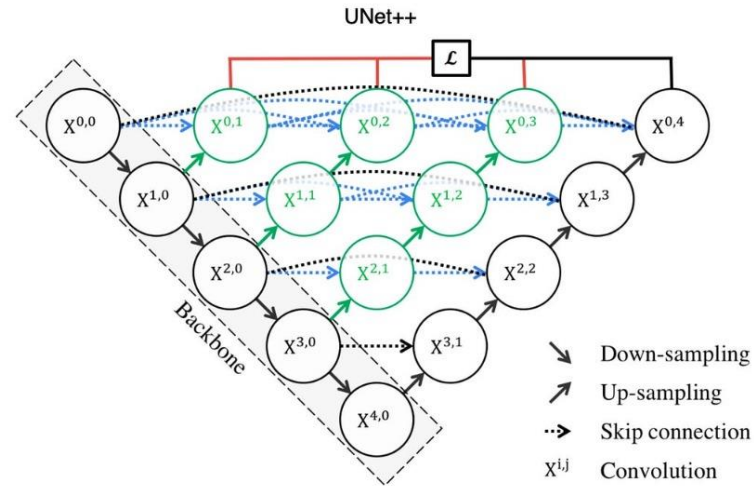


ARQUITECTURAS

Arquitecturas elegidas

- U-NET
- U-NET ++
- Deeplab

Tanto UNET ++ como Deeplab son modelos preentrenados, Deeplab utiliza Resnet y UNET++ xception.

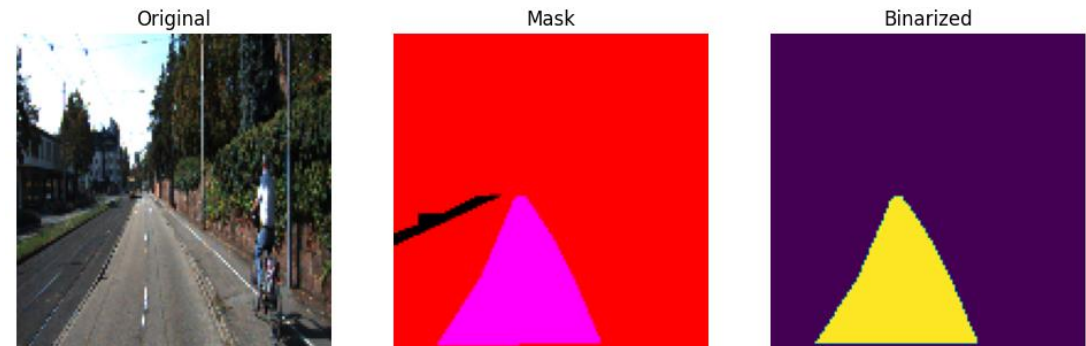


ANALISIS EXPLORATORIO

Data Augmentation

- Horizontal Flip (50%)
- Rotate(-10%, 10%)

Binarización de las mascararas

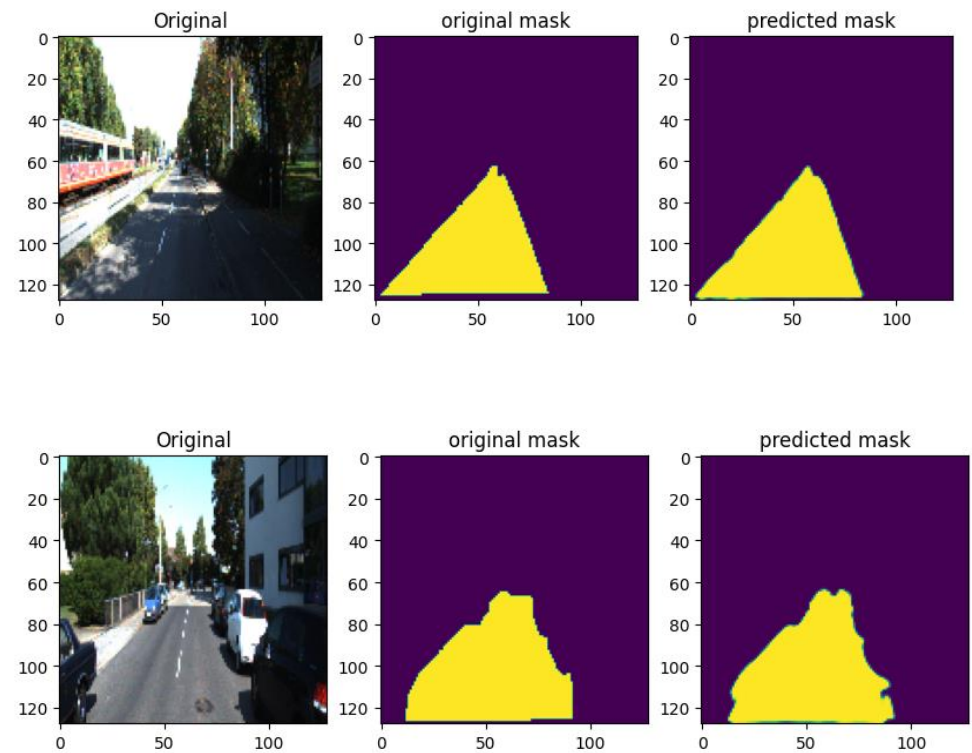
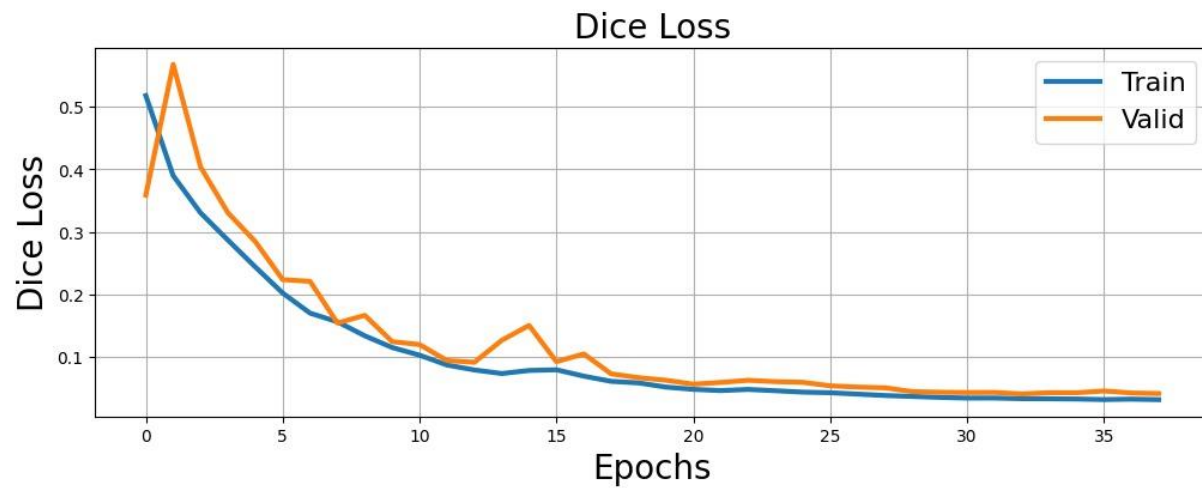


HIPERPARAMETROS

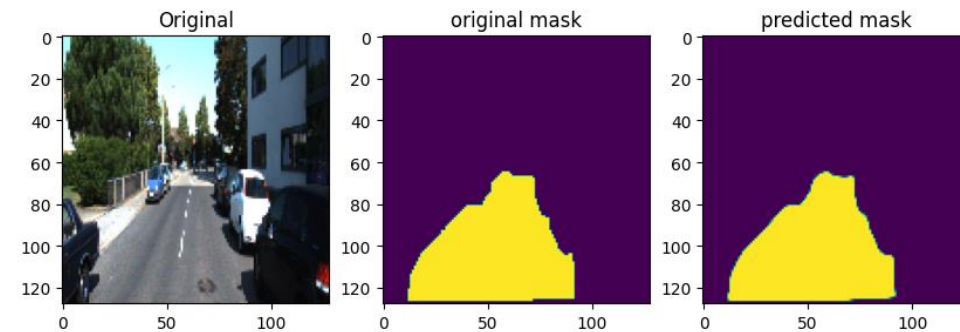
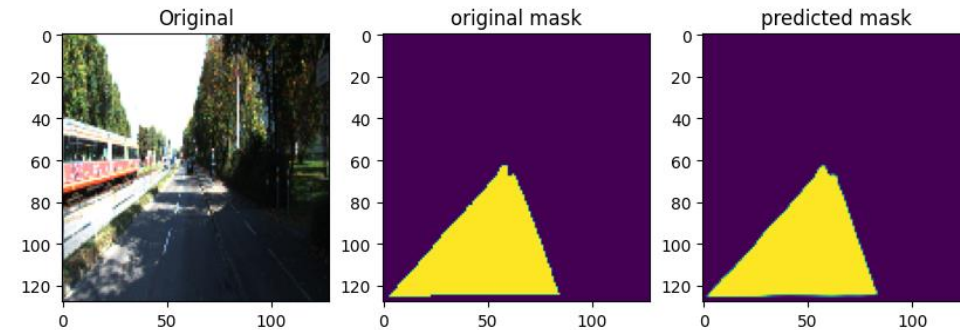
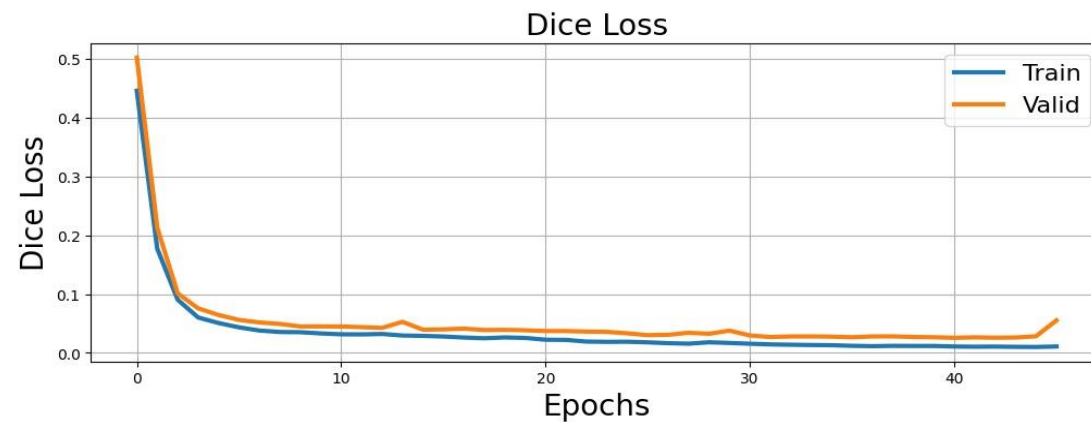
- U-NET, U-NET ++, Deeplab
 - Learning Rate = $1e-3$ y $3e-3$
 - Optimizer = Adam
 - Early Stop patience=5



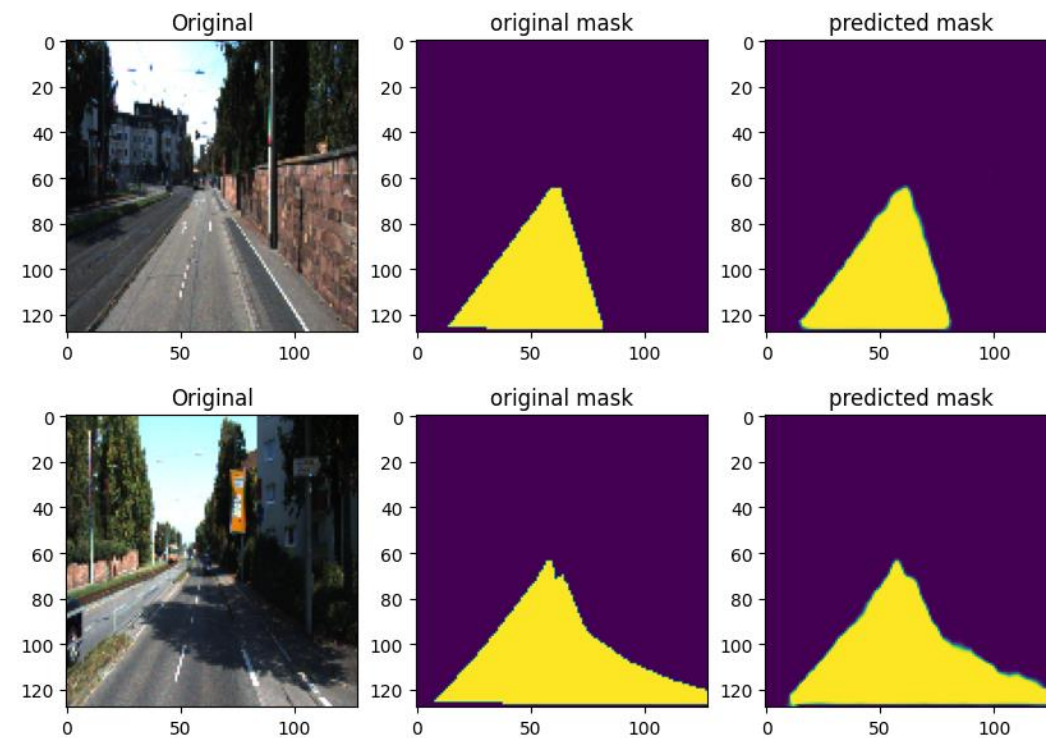
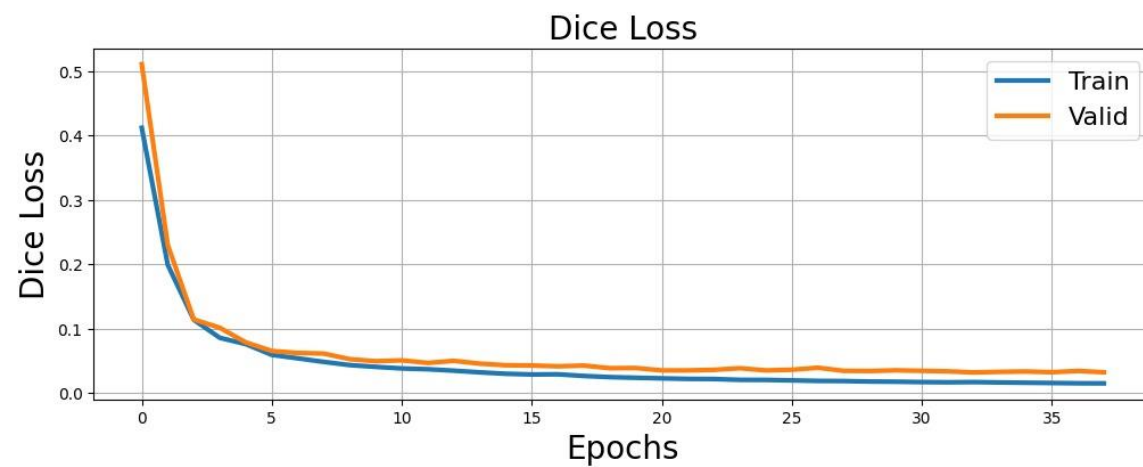
ENTRENAMIENTO U-NET



ENTRENAMIENTO U-NET++



DEEPLAB V3



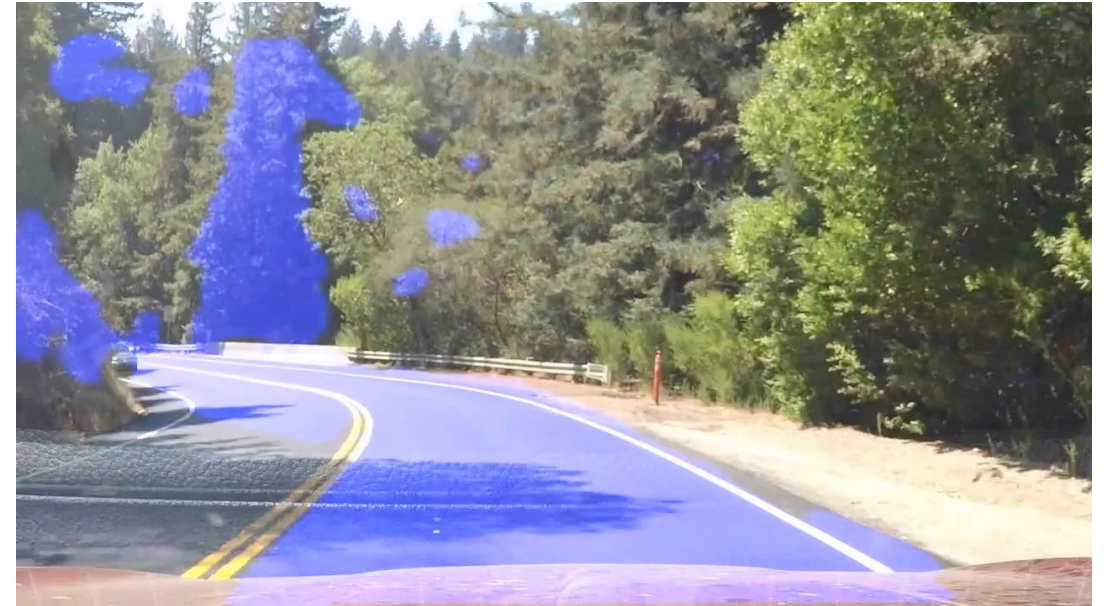
METRICAS

Arquitectura	Learning Rate	Dice Loss	IOU Score	Total parametros entrenables	Epocas de entrenamiento
U-NET	1e-3	0.030842	0.945116	31.037.633	38
	3e-3	0.046337	0.912659		28
U-NET ++	1e-3	0.007389	0.987283	34.006.905	49
	3e-3	0.010960	0.979075		46
DEEPLAB	1e-3	0.018328	0.968892	45.669.713	39
	3e-3	0.025653	0.951807		31



ALGUNAS PRUEBAS CON VIDEO

- Videos del dataset KITTY



- Video propio



CONCLUSIONES

Despues de entrenar los distintos modelos podemos observar que:

- Unet y Unet ++ suelen devolver mascaras con bordes mas definidos que Deeplab
- Unet ++ necesitó mas épocas de entrenamiento
- En Unet ++ se consigue un menor loss.

