Contexte du projet

Segmentation d'images

Input

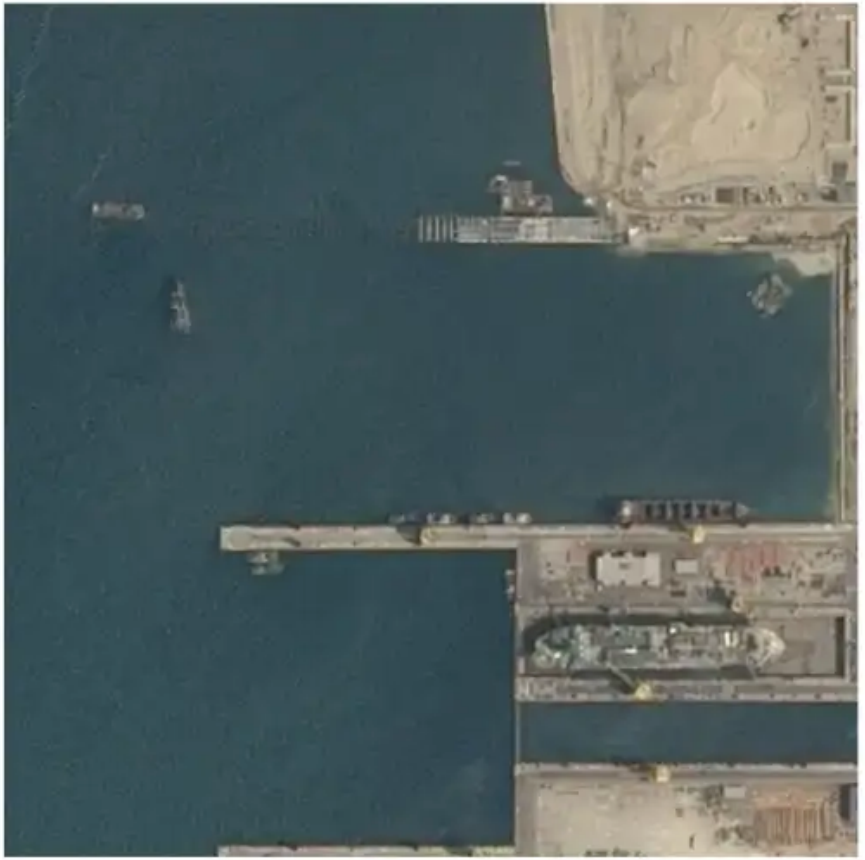
segmented

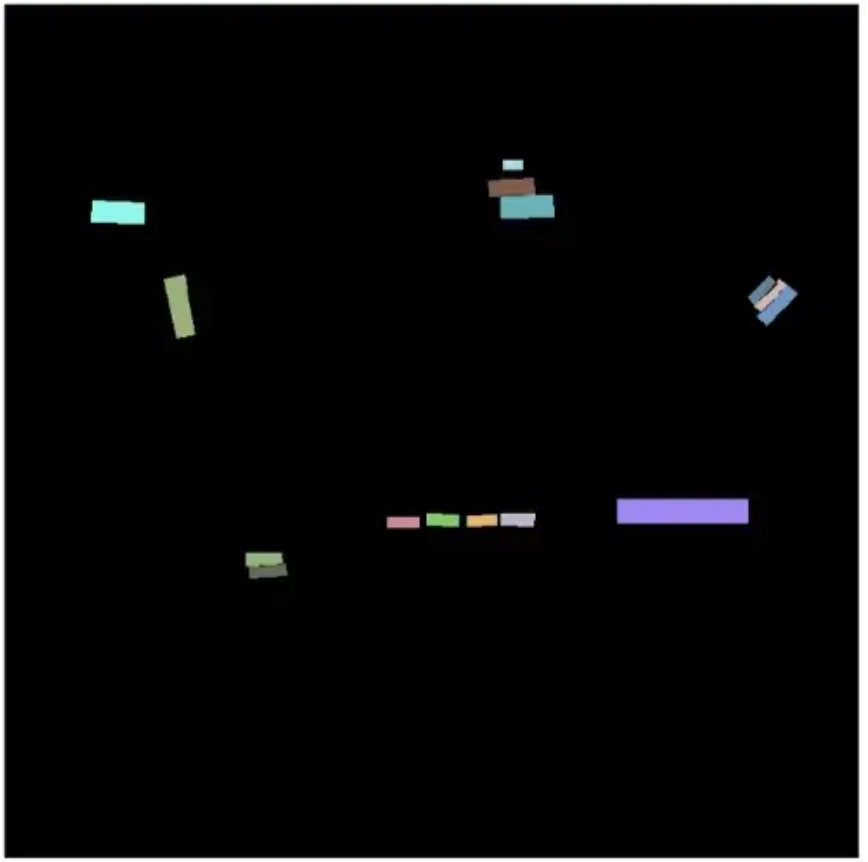
1: Person
2: Purse
3: Plants/Grass
4: Sidewalk
5: Building/Structures

3 3 3 3 3 3 3 3 3 3 3 3 5 5 5 5 5 5
3 3 3 3 3 3 3 3 3 3 3 3 5 5 5 5 5 5
3 3 3 3 3 3 1 1 3 3 3 3 5 5 5 5 5 5
3 3 3 3 3 1 1 1 1 3 3 3 5 5 5 5 5 5
3 3 3 3 3 3 1 1 3 3 3 5 5 5 5 5 5 5
5 5 3 3 3 3 1 1 3 3 5 5 5 5 5 5 5 5
4 4 3 4 1 1 1 1 1 1 4 4 4 5 5 5 5 5
4 4 3 4 1 1 1 1 1 1 4 4 4 4 4 5 5 5
4 4 4 1 1 1 1 1 1 1 1 4 4 4 4 4 4 4
3 3 3 1 1 1 1 1 1 1 1 4 4 4 4 4 4 4
3 3 3 1 2 2 1 1 1 1 1 4 4 4 4 4 4 4
3 3 3 1 2 2 1 1 1 1 1 4 4 4 4 4 4 4

Semantic Labels

Mesures de performances

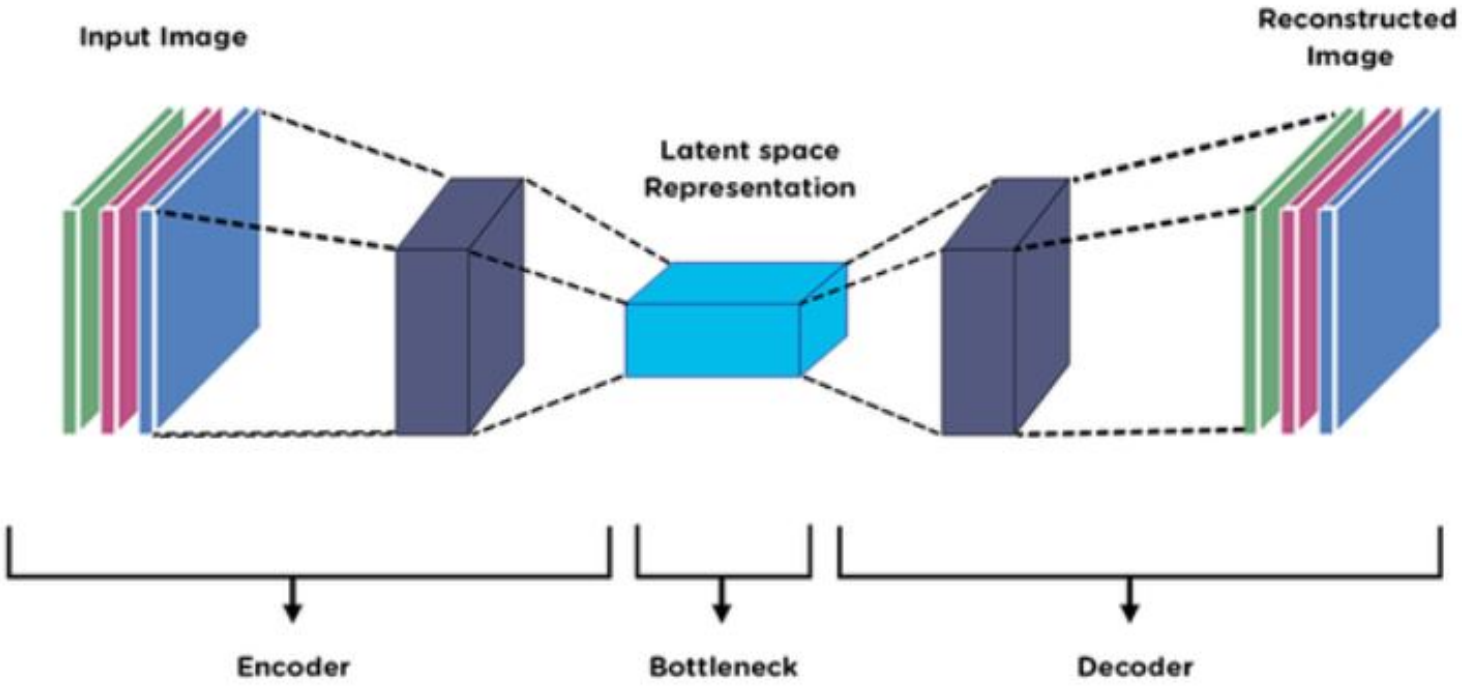




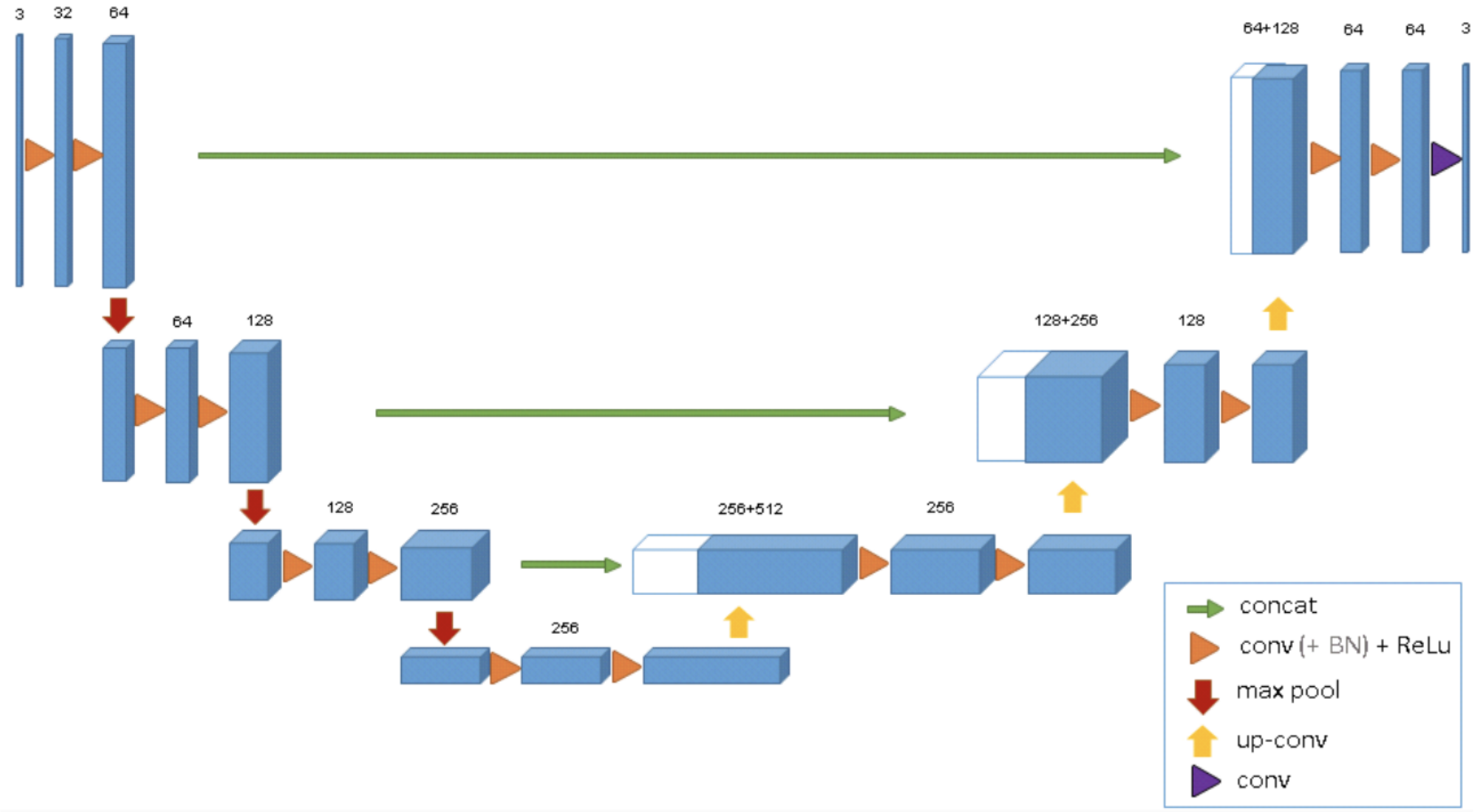
Mesures de performances

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$

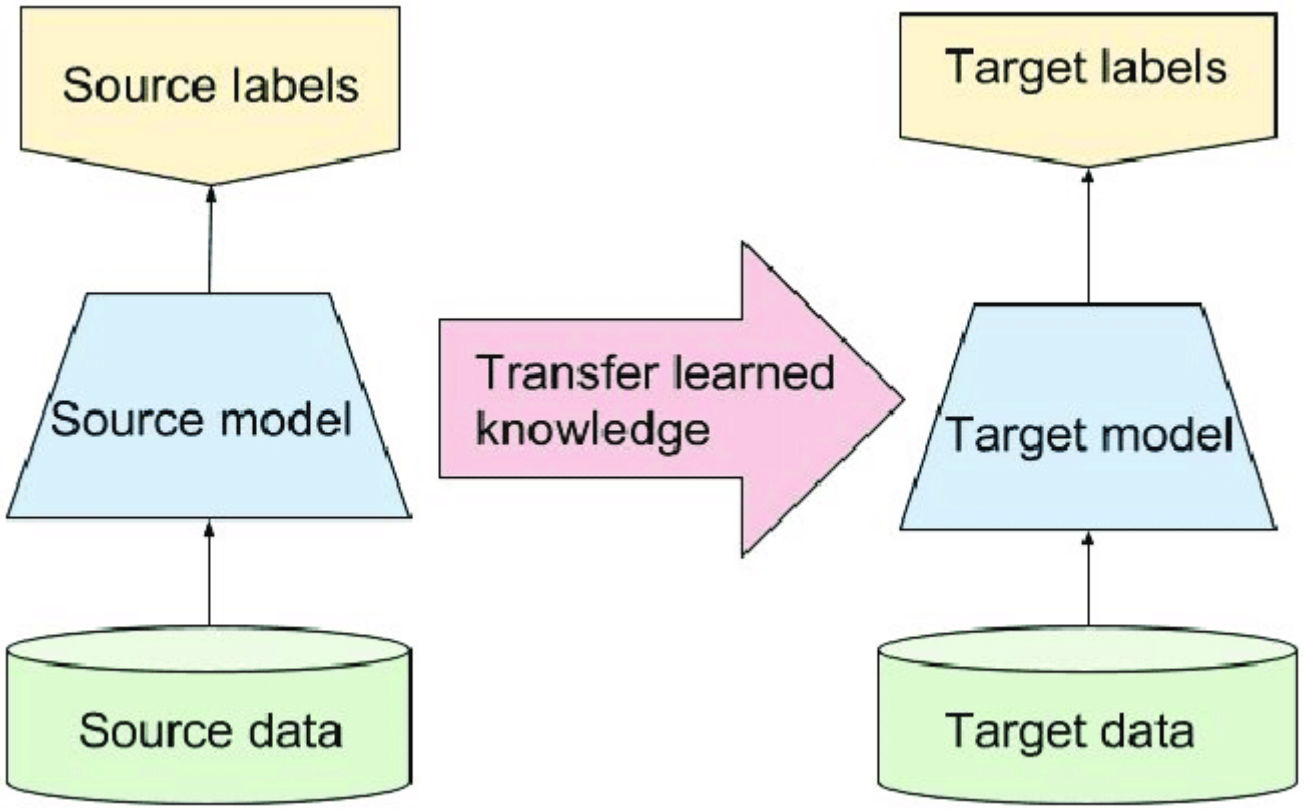
Modèles


Input Image
Reconstructed Image
Latent space Representation
Encoder
Bottleneck
Decoder

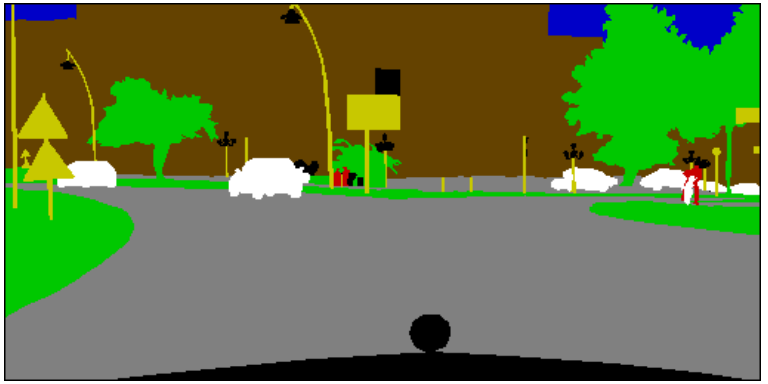
U-Net


3
32
64
64+128
64
64
3
64
128
128+256
128
128
256
256+512
256
256
concat
conv (+ BN) + ReLu
max pool
up-conv
conv

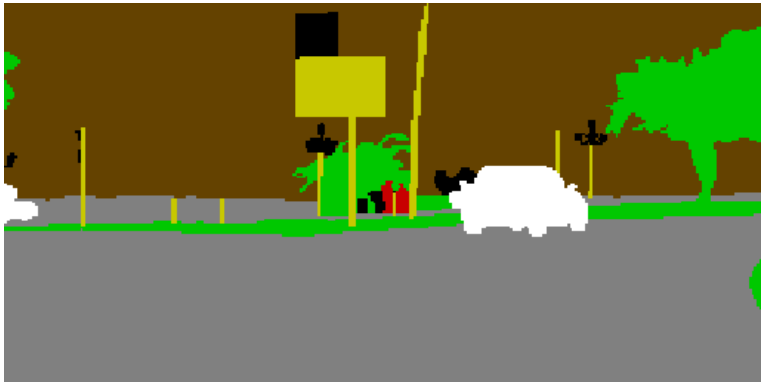
Transfer learning



Augmentations



augmentation



Résultats

model_name	img_size	n_augments	img_augments	mean_IoU	class_IoU	epoch_train_time
U-Net pretrained vgg16	(256, 512)	1	{'h_flip': 1, 'crop': 0.2}	0.625	[0.691 0.904 0.718 0.298 0.767 0.667 0.3 0.658]	643.0
U-Net pretrained vgg16	(256, 512)	0	{}	0.621	[0.693 0.904 0.713 0.293 0.782 0.679 0.279 0.621]	547.0
U-Net pretrained vgg16	(256, 512)	1	{'h_flip': 1, 'rotate': 0.2}	0.615	[0.687 0.902 0.711 0.25 0.773 0.66 0.275 0.657]	665.0
U-Net pretrained resnet50	(256, 512)	1	{'h_flip': 1, 'crop': 0.2, 'rotate': 0.2}	0.475	[0.594 0.748 0.542 0.226 0.733 0.276 0.124 0.555]	656.0
U-Net pretrained resnet50	(256, 512)	0	{}	0.455	[0.524 0.779 0.553 0.16 0.658 0.271 0.205 0.49]	546.0
U-Net	(512, 1024)	0	{}	0.592	[0.761 0.928 0.676 0.212 0.72 0.658 0.249 0.532]	1140.0
U-Net	(256, 512)	0	{}	0.473	[0.677 0.867 0.6 0.028 0.663 0.314 0.117 0.515]	500.0
U-Net	(128, 128)	0	{}	0.390	[0.518 0.809 0.52 0. 0.449 0.446 0.021 0.36]	370.0

Améliorations possibles

- résolution
- fonction de perte
- modèle
- canal d'image supplémentaire
- classe "void"

Déploiement cloud