

SegNet

요약

- Encoder - decoder 아키텍처

encoder = f.c layer를 제외한 VGG16 사용

decoder = 학습 파라미터가 필요없는 Un-maxPooling 이용하여 Upsampling한다.

- decoder에서 Upsampling된 feature map은 convolution layer를 통해 dense feature map으로 만든다.

- 성능이 우수하다.

• 메모리, 정확도 사이의 trade-off에서 FCN보다 좋은 퍼포먼스를 낸다.

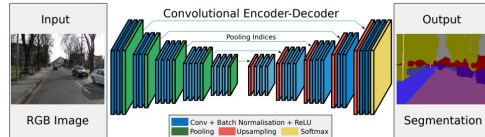
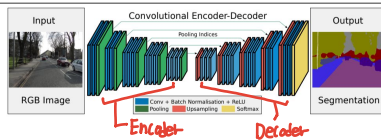


Fig. 2. An illustration of the SegNet architecture. There are no fully connected layers and hence it is only convolutional. A decoder upsamples its input using the transferred pool indices from its encoder to produce a sparse feature map(s). It then performs convolution with a trainable filter bank to densify the feature map. The final decoder output feature maps are fed to a soft-max classifier for pixel-wise classification.

구조

- Encoder - Decoder 구조



- Encoder

- Conv + Batch Norm + ReLU + Pooling을 사용하여 feature map을 준비한다.
- Pooling이 어떻게 이루어졌는지 기록을 저장한다.

- Decoder

- Upsampling 과정에서 FCN 파트에서 Un Pooling 과정을 사용하여 디메커이즈를 크게 한다.
- Soft max Layer에서는 클래스 개수만큼 채널이 존재해서 Soft max 과정을 진행한다.

특징

- 이미지 마다 모델을 바꿔야 하는 문제가 없다. 하나의 모델을 가지고 다양한 이미지를 예측할 수 있다.

- 논문에서는 메모리를 절약했다고 한다. (FC-Layer를 사용하지 않아서)