



# [논문리뷰] DeepLabV1 [2014]

## Introduction

☀ 본 논문은 DCNN과 fully-Conncted CRF(conditional random field)를 결합한 **DeepLab**을 제안  
DCNN은 Image Classification과 Object Detection에서 좋은 성능을 보여줍니다. 하지만, Semantic Segmentation에서 **DCNN을 적용하는데 두 가지 기술적 한계**가 존재합니다.

### 1. Signal DownSampling

- 기존 DCNN의 각 계층에서 Max Pooling, Stride로 인하여, **Singal Resolution이 감소**합니다.  
**Atrous Alortihm(hole algorithm)을 적용하여 해결**하였습니다.

### 2. Spatial Insensitivity(Invariance)

- DCNN의 invariance 특성은 공간의 추상적인 정보를 필요로 하는 Image Classification에서는 장점이었지만, **정확한 Localization을 요구하는 Semantic Segmentation에서는 정확도를 제한** 하였습니다.

**DCNN-based pixel-level classifier와 CRF를 결합함으로써 해결**하였습니다.

DeepLab에는 3가지 장점이 존재합니다.

- Atrous Algorithm과 CRF를 적용함으로써, **Speed**가 improve되었습니다.
- Accuracy**가 PASCAL semantic segmentation challenge에서 Second를 달성하였습니다.
- DCNNs와 CPFs만 사용하므로, 상당히 모델이 **Simplicity**합니다.

## CNN for Dense Image Labeling

☀ Segmentation을 위하여, VGG16을 Dense feature extractor로 사용하였습니다.

### Dense Sliding Window feature extraction with the Hole Algorithm

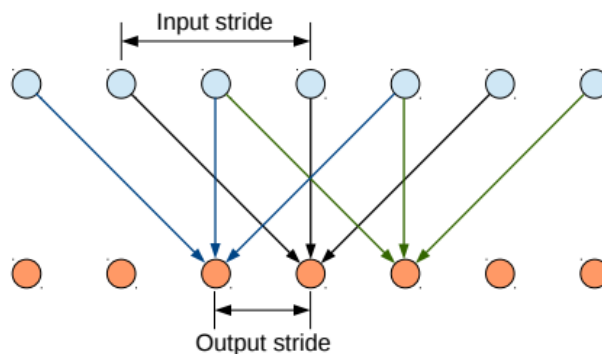


Figure 1: Illustration of the hole algorithm in 1-D, when  $kernel\_size = 3$ ,  $input\_stride = 2$ , and  $output\_stride = 1$ .

Semantic Segmentation에서는 **Dense Spatial evaluation가 중요**합니다. VGG-16의 fully connected layer를 Convolutional로 변경하여도, **sparsely(희소한 부분)을 생성**하기에는 무언가 필요합니다.

신경망에서 마지막 두 개의 max-pooling layer를 제거합니다. 그리고 Conv Filter를 수정합니다.

## Controlling the receptive field size and accelerate dense computation with convolutional Net

dense score computation을 위해서는 신경망의 receptive field size를 조절해야 합니다. 신경망의 fully connected layer를 fully Convolutional로 변경할 때, 7x7 receptive field에서 4096개의 filter를 얻습니다.

그리고, 이것은 **Computational Bottleneck**이 됩니다.

FC를 통과하기 전 마지막 feature map의 크기를 4x4로 사용하며, **마지막 feature map크기를 변경**합니다.

## Detailed boundary recovery

### DNN and localization challenge

DCNN의 score map은 Image에서 **객체의 존재와 대략적 위치를 안정적으로 예측**하지만, **정확한 윤곽을 예측하는 데는 적합하지 않**다는 단점이 있습니다.

Recognition Capactiy of DCNNs와 fine-grained localization accuracy of fully connected CRFs를 결합함으로써, **새로운 해결방안을 추구**하였습니다.

☀ DeepLabvV1은 기존 방법을 뛰어넘는 수준으로 성공했음을 보여줍니다.

(Localization Challenge 해결 및 Semantic Segmentation 결과를 생성, 뛰어난 객체 경계 복구 성능)

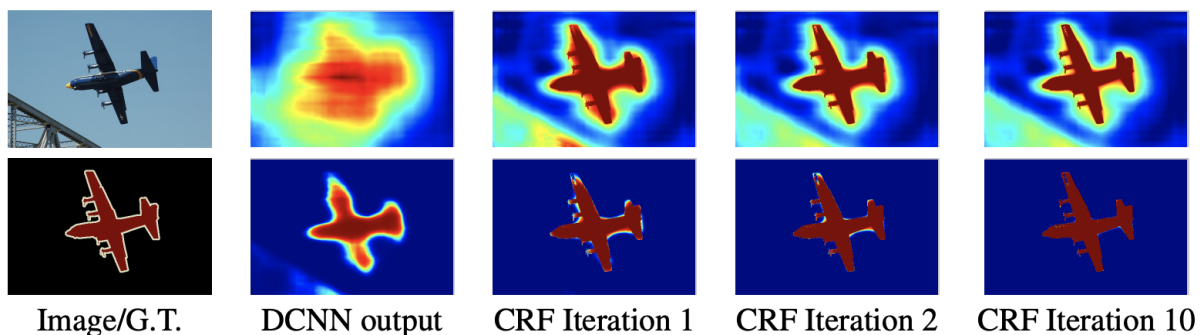
### FC Conditional Random Fields for Accurate Localization

전통적으로, short-range CRFs는 Segmentation map의 노이즈를 매끄럽게 하기 위해 사용합니다.

하지만, DCNN의 **Max-Pooling**과 **Upsampling**을 진행하면서 작아진 feature map의 결과를 **bilinear intepolation**으로 기존 크기로 확대하면, score map은 충분히 매끄러워집니다.

local-range CRFs와 함께 Contrast-sensitive potential을 활용하면, localization을 잠재적으로 개선할 수 있지만, 여전히 얇은 구조를 놓칠 수 있으며,

**고비용의 이산 최적화 문제(discrete optimization problem)**를 해결해야 합니다.



Testing에서 **bilinear intepolation**(이중 선형 보간법)을 사용하여, **DeepLab모델의 출력 resolution을 8배 증가 시켜, original image resolution의 class score map을 생성**합니다.

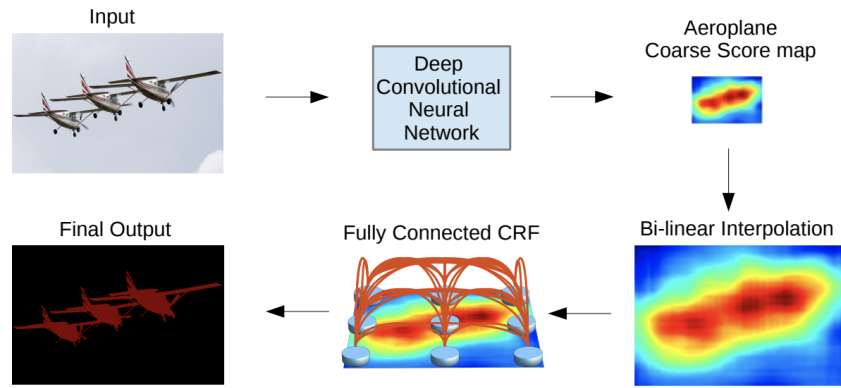


Figure 3: Model Illustration. The coarse score map from Deep Convolutional Neural Network (with fully convolutional layers) is upsampled by bi-linear interpolation. A fully connected CRF is applied to refine the segmentation result. Best viewed in color.

## Multi-Scale Prediction

DeepLab은 추가적으로 Boundary Localization의 정확도를 높이기 위해, multi-scale prediction을 시도해왔습니다. input image와 첫 번째 4개의 max-pooling layer의 output을 메인 네트워크의 마지막 layer의 feature map에 연결된 2-layer MLP에 연결했습니다.

Method	mean IOU (%)
DeepLab	59.80
DeepLab-CRF	63.74
DeepLab-MSc	61.30
DeepLab-MSc-CRF	65.21
DeepLab-7x7	64.38
DeepLab-CRF-7x7	67.64
DeepLab-LargeFOV	62.25
DeepLab-CRF-LargeFOV	67.64
DeepLab-MSc-LargeFOV	64.21
DeepLab-MSc-CRF-LargeFOV	68.70

(a)

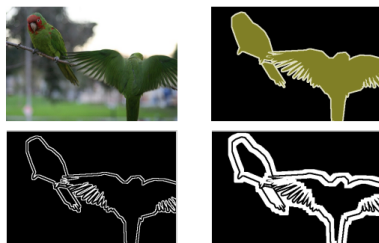
Method	mean IOU (%)
MSRA-CFM	61.8
FCN-8s	62.2
TTI-Zoomout-16	64.4
DeepLab-CRF	66.4
DeepLab-MSc-CRF	67.1
DeepLab-CRF-7x7	70.3
DeepLab-CRF-LargeFOV	70.3
DeepLab-MSc-CRF-LargeFOV	71.6

(b)

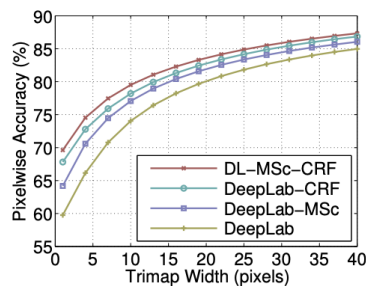
## Performance & Test and result

☀ DeepLab과 fully-connected CRFs를 통합한 모델은 성능이 크게 향상합니다.

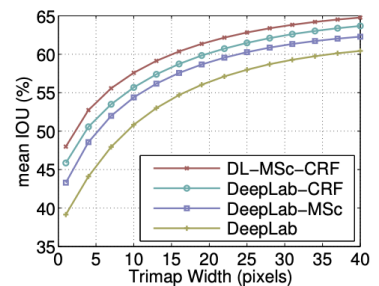
(복잡한 객체 경계를 정확히 포착합니다.)



(a)



(b)



(c)


DeepLab에 multi-scale feature와 large-FOV를 활용함으로써, 복잡한 객체 경계를 빠르고 정확히 예측해낼 수 있습니다.

DeepLab-CRF 및 DeepLab-MS-CRF 모델은 각각 평균 IOU 66.4%, 67.1%의 성능을 보입니다.  
이는 다른 최신 모델과 비교해 큰 차이를 보여줍니다.

multi-scale feature와 large-FOV를 모두 사용한 DeepLab-MS-CRF-LargeFOV는 Best 성능을 보인다 .

#### Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs


Deep Convolutional Neural Networks (DCNNs) have recently shown state of the art performance in high level vision tasks, such as image classification and object detection. This work brings together methods from DCNNs and probabilistic graphical models for addressing the task of pixel-level

 <https://arxiv.org/abs/1412.7062>



#### DeepLab\_v1/DeepLab\_v1 - SBD, VOC 2012.ipynb at main · Jasonlee1995/DeepLab\_v1

DeepLab v1 Implementation with Pytorch. Contribute to Jasonlee1995/DeepLab\_v1 development by creating an account on GitHub.

 [https://github.com/Jasonlee1995/DeepLab\\_v1/blob/main/Implementation/DeepLab\\_v1%20-%20SBD%2C%20VOC%202012.ipynb](https://github.com/Jasonlee1995/DeepLab_v1/blob/main/Implementation/DeepLab_v1%20-%20SBD%2C%20VOC%202012.ipynb)

#### Jasonlee1995/ DeepLab\_v1

DeepLab v1 Implementation with Pytorch

1 Contributor 0 Issues 3 Stars 2 Forks