DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs

Authors: L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A. L. Yuille

Abstract:

This paper introduces the DeepLab framework, a comprehensive approach to semantic image segmentation that leverages deep convolutional neural networks (CNNs), atrous (dilated) convolutions, and fully connected Conditional Random Fields (CRFs).

The goal of DeepLab is to achieve accurate and detailed pixel-wise labeling of images, effectively segmenting them into different semantic categories.

Introduction:

Semantic image segmentation involves assigning a class label to each pixel in an image. The authors highlight the challenges in achieving accurate segmentation due to the need for both local and global context information.

The DeepLab framework aims to address these challenges by combining deep learning techniques and graphical models.

Deep Convolutional Nets:

DeepLab uses a modified VGG16 or ResNet architecture as the backbone for feature extraction. The authors describe the architecture's layers and how it processes input images to generate feature maps that capture different levels of information.

Atrous Convolution:

Atrous convolution, also known as dilated convolution, allows for the effective adjustment of the receptive field size without increasing the number of parameters.

DeepLab utilizes atrous convolutions to capture multi-scale contextual information from the input images.

Fully Connected Conditional Random Fields (CRFs):

The final output of the convolutional network can still contain inconsistencies at the pixel level. To refine the segmentation results, DeepLab employs fully connected CRFs, which model the relationships between neighboring pixels and their labels.

This post-processing step helps improve the spatial coherence of the segmentation.

Experiments and Results:

The authors conducted experiments on various benchmark datasets to evaluate the performance of the DeepLab framework.

They compared the results with other state-of-the-art methods and demonstrated the effectiveness of their approach in achieving accurate and fine-grained semantic segmentation.

Conclusion:

The paper concludes by summarizing the contributions of the DeepLab framework in combining deep convolutional nets, atrous convolutions, and fully connected CRFs for semantic image segmentation.

The authors highlight the potential applications of their approach in various domains, including medical imaging, remote sensing, and more.