Paper Reading

Jinming Su

Last update: May 30, 2019

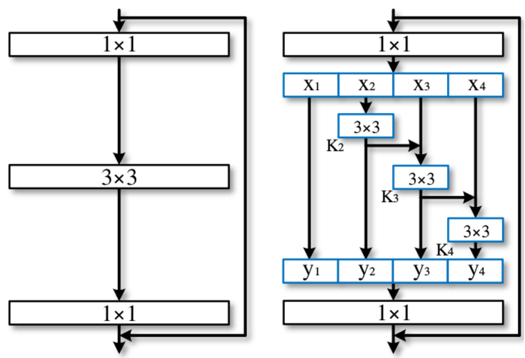
Contents

Todo list

Chapter 1

Classification

1.1 Res2Net: A New Multi-scale Backbone Architecture, arxiv, 2019.



(a) Bottleneck block (b) Res2Net module

Figure 1.1: Residual Learning.

This paper is mainly a promotion of residual network, which seems to be similar to cardinality, only adding more connections.

Chapter 2

Object Detection

2.1 Proposal

2.1.1 Multiscale Combinatorial Grouping for Image Segmentation and Object Proposal Generation. TPAMI, 2016

I don't read this paper, but the code is tested. Code can be found in `https://github.com/jponttuset/mcg`. When testing, run pre-trained/install.m, pre-trained/demos/demo_im2mcg.m and get the following results.

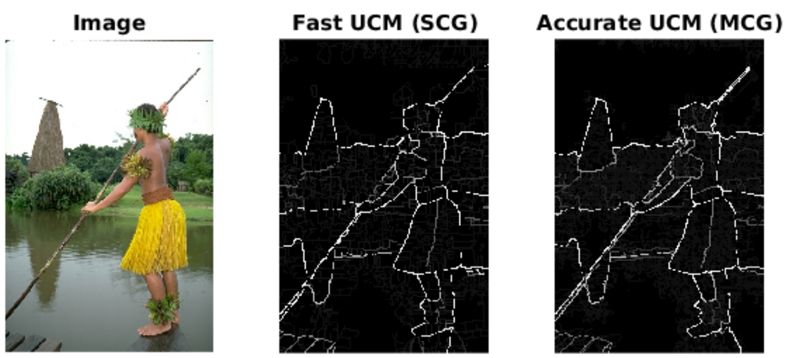


Figure 2.1: The UCM of MCG.

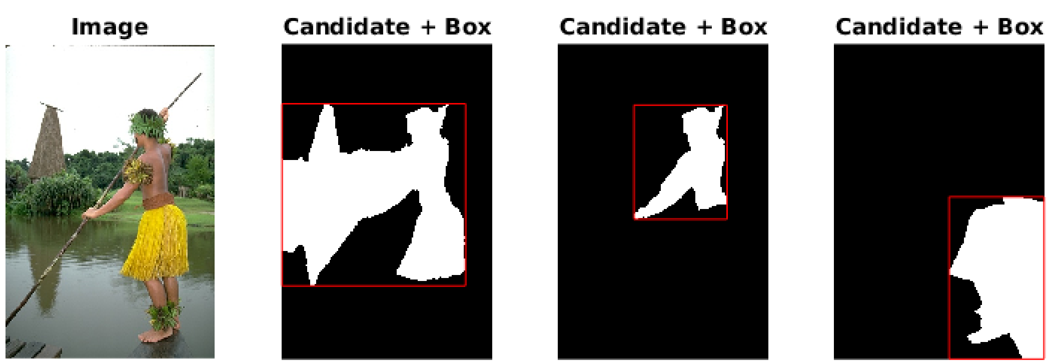


Figure 2.2: The results of MCG.

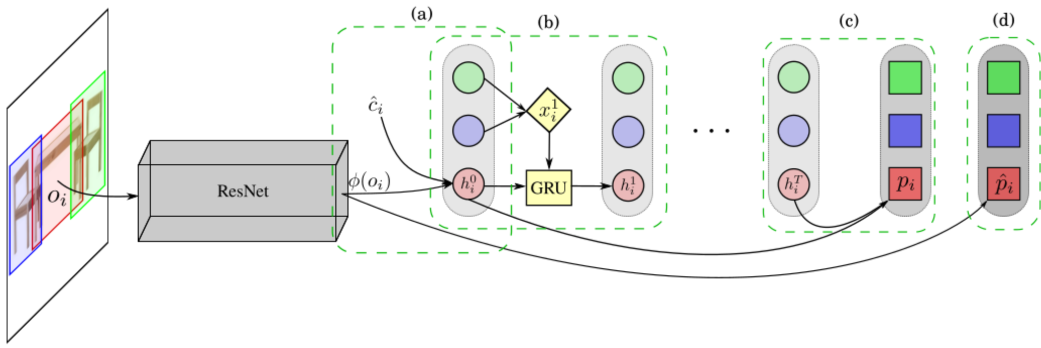


Figure 2.3: GGNN.

2.2 What Object Should I Use? - Task Driven Object Detection. CVPR, 2019.

This paper has two contributions: (1) construct a COCO-Tasks dataset which comprises about 40,000 images where the most suitable objects for 14 tasks have been annotated; (2) proposes a method buliding on Gated Graph Neural Network to detect the most suitable objects for a given task.

2.3 Feature Pyramid Networks for Object Detection. CVPR, 2017.

This paper propose a new backbone network.

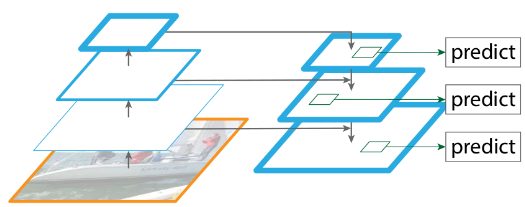


Figure 2.4: Structure of FPN.

2.4 Focal Loss for Dense Object Detection. ICCV, 2017

This paper propose a improved cross entropy loss.

$$\mathcal{FL}(p,y) = \begin{cases} -\alpha(1-p)^{\gamma}\log(p), & y=1 \\ -(1-\alpha)p^{\gamma}\log(1-p), & y=0 \end{cases} \tag{2.1}$$

where α is the weighting factor for class imbalance, and p_t is the confidence

Chapter 3

Salient Object Detection

3.1 DNA: Deeply-supervised Nonlinear Aggregation for Salient Object Detection. arxiv, 2019.

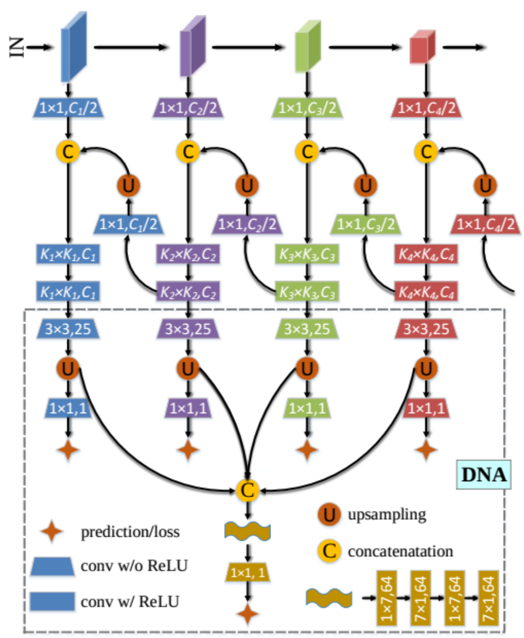


Figure 3.1: Residual Learning.

This paper has two contributions: (1) theoretically and experimentally analyzes the natural limitaion of traditional side-output aggregration which can only make limited use of multi-scale side-ouput informantion; (2) proposes Deeply-supervised nonlinear aggregration (DNA) for side-output features. (3) as experience, in DNA, convolution layers with kernels of $n \times 1$ and $1 \times n$ are used, which is proved to be effective. Moreover, authers claim that large kernel size in DNA can improve performance.

3.2 Instance-Level Salient Object Segmentation. CVPR, 2017.

This paper has contributions as follows: (1) introduce a new task of instance-level salient object detection, and construct a corresponding dataset with 1000 images that are collected from existing SOD datasets; (2) propose a method for this task as shown in figure.

Moreover, the code can be found in `http://www.sysu-hcp.net/instance-level-salient-object-segmentation/` while the code is only the part of SOD without contour and instance.

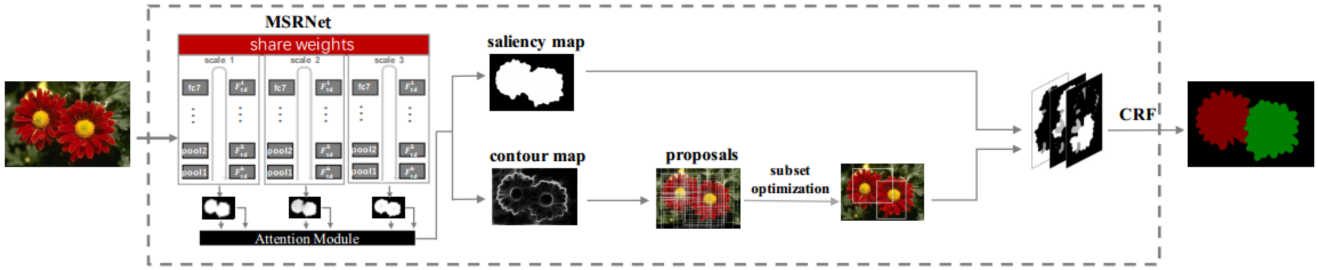


Figure 3.2: Instance-level salient object detection.

3.3 Semantic Instance Meets Salient Object: Study on Video Semantic Salient Instance Segmentation. WACV, 2019.

3.4 A Simple Pooling-Based Design for Real-Time Salient Object Detection. CVPR, 2019

3.5 S4Net: Single Stage Salient-Instance Segmentation. CVPR, 2019.

3.6 Contrast Prior and Fluid Pyramid Integration for RGBD Salient Object Detection. CVPR, 2019

Chapter 4

Semantic Segmentation

4.1 WebSeg: Learning Semantic Segmentation from Web Searches. arxiv, 2018.

This paper propose a method to learning semantic segmentaion web searching images. The method is mainly an online

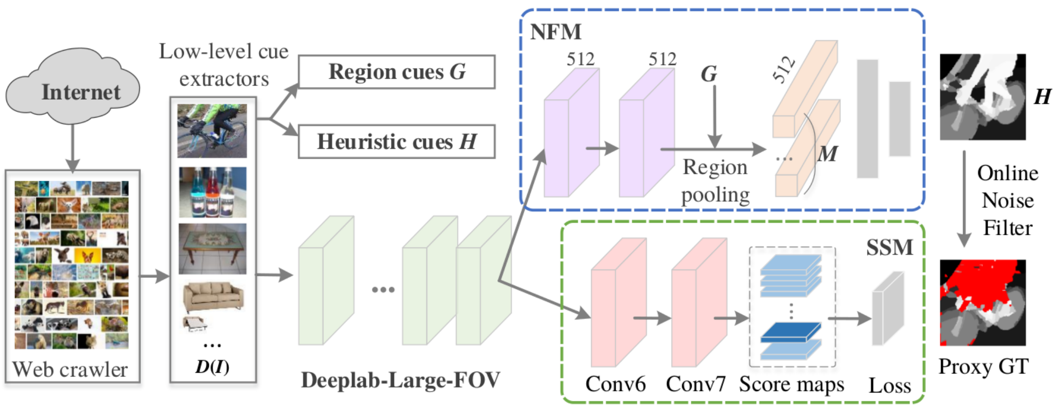


Figure 4.1: Framework.

noise filtering module, which is able to filter potentially noisy region labels to obtain the correct region labeled by the qury keywords.

4.2 ADVENT: Adversarial Entropy Minimization for Domain Adaptation in Semantic Segmentation. CVPR, 2019.

This paper focuses on the problem of domain adapation in semantic segmentation. In detail, the contributions are as follows: (1) propose to minimize the pixel-level entropy of target domain to penalize low-confident predictions on target domain; (2) propose a entropy-based adversarial traing approach to privide the structure adaptation; (3) extra two trick: a) training on specific entropy ranges and b) add class-ratio priors.

4.3 Efficient Ladder-style DenseNets for Semantic Segmentation of Large Images. arxiv, 2019.

This paper propose a ladder-style densenet for semantic segmentation, while I think this is not new idea, which has been widely used in present work such as FPN. Moreover, this paper present to reduce the training memory footprint by aggressive re-computation of imtermediate activations during convolutional backprop, which maybe is worth learning.

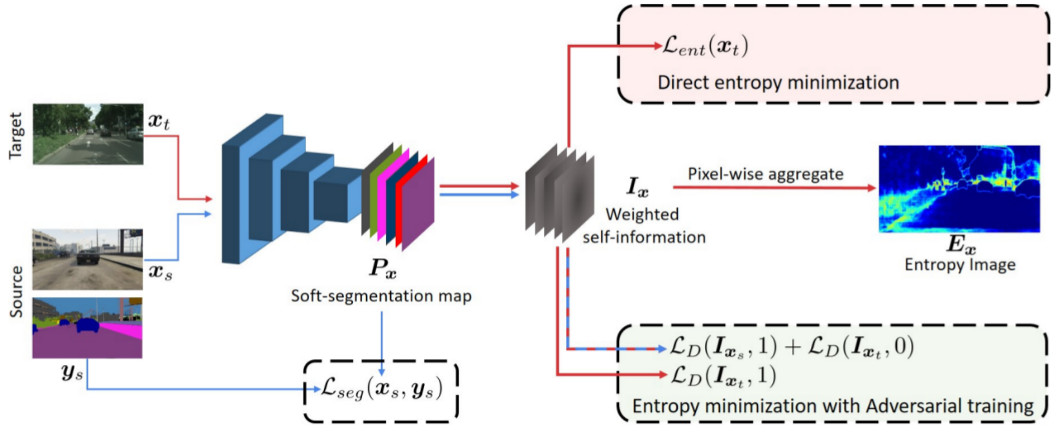


Figure 4.2: Advsarial entropy.

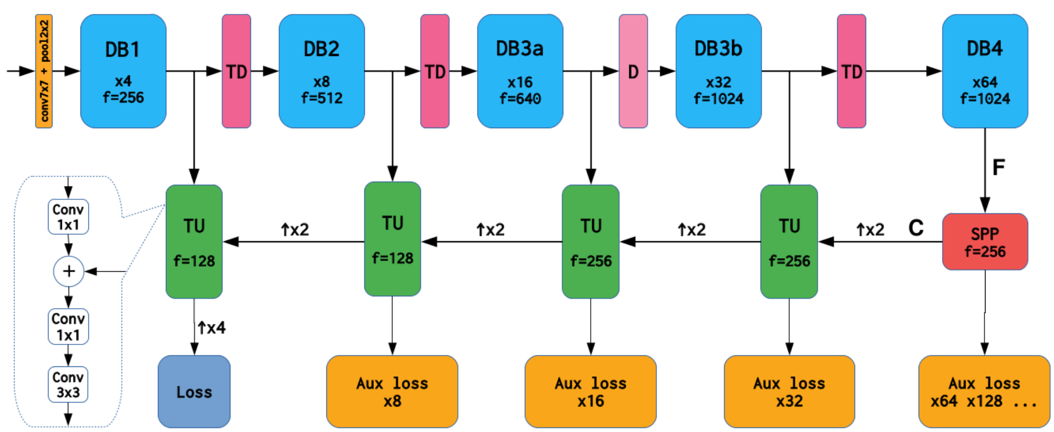


Figure 4.3: Ladder-style densenet for semantic segmentation.

Chapter 5

Pose Estimation

5.1 DeepPose: Human Pose Estimation via Deep Neural Networks. CVPR, 2014.

Figure 5.1: Deep pose.

This method proposes a method to regress key points. It first predict these approximate locations, and then the box is croped, upsampled and sent to the regressor for the finer prediction.

5.2 Human pose estimation via Convolutional Part Heatmap Regression. ECCV, 2016.

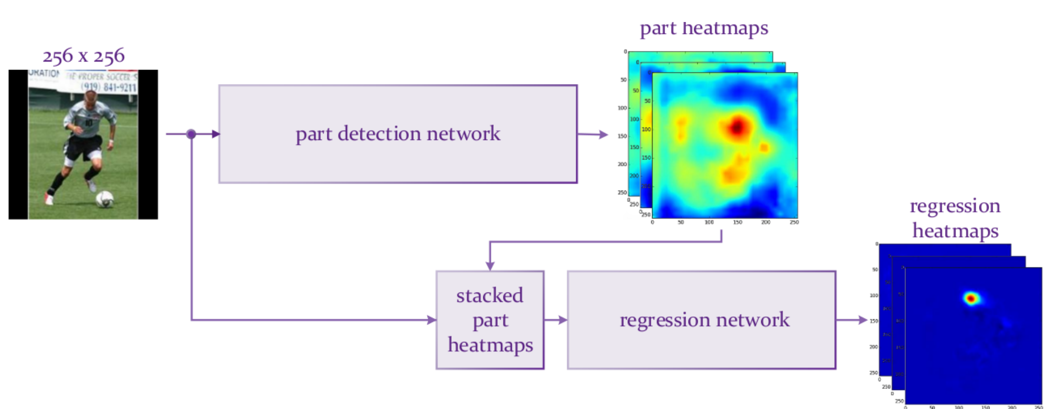


Figure 5.2: Convolutional part heatmap regression.

This paper proposed the modern mainstream architecture by part heatmaps to solve the human pose estimation.

5.3 Deep High-Resolution Representation Learning for Human Pose Estimation. CVPR, 2019.

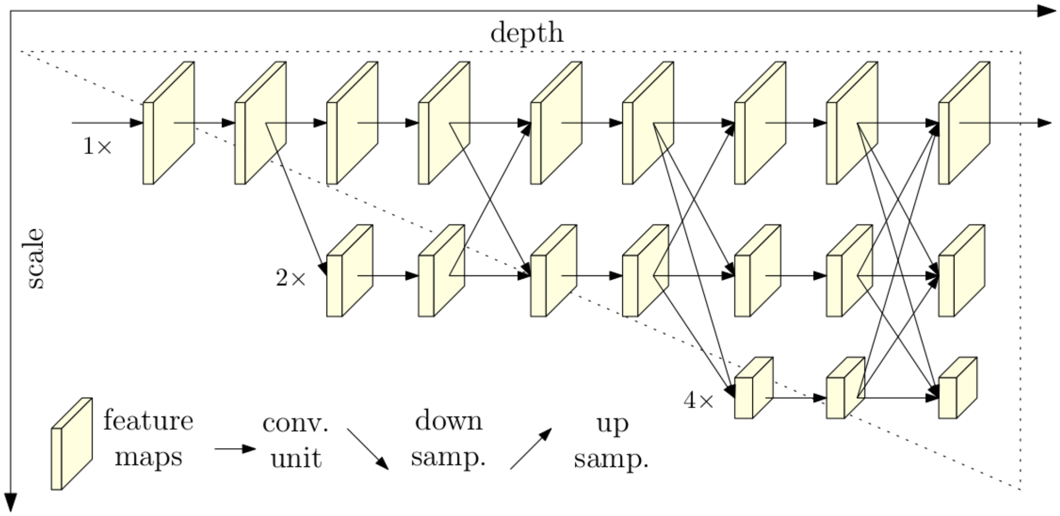


Figure 5.3: HRNet.

This paper propose a new backbone network for maintaining high-resolution.

Chapter 6

Others

6.1 Single Image Haze Removal Using Dark Channel Prior. CVPR, 2009. Best paper.

This paper introduces dark channel prior that is an observation -most local patches in haze-free outdoor images contain some pixels which have very low intensities in at least one color channel. This is the one of most famous papers in the domain of dehazing.

Some formulas in this paper are easy to understand. One can refer to `https://www.cnblogs.com/Imageshop/p/3281703.html` for more understanding.

The unofficial python code can be found in `https://github.com/su526664687/dark-channel-prior-dehazing`.

6.2 Single Image Dehazing Using Ranking Convolutional Neural Network. TMM, 2018.

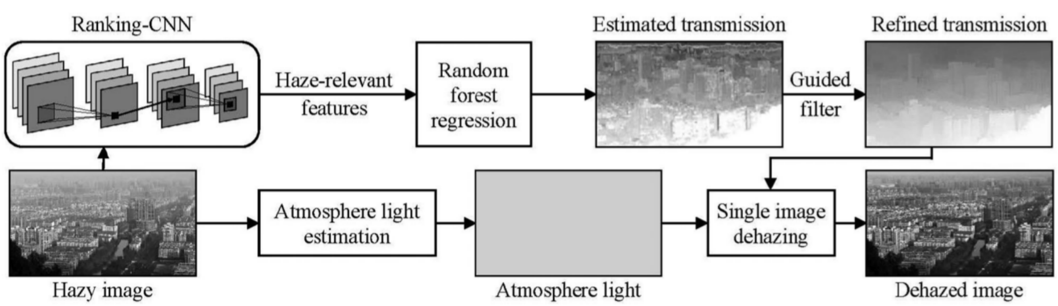


Figure 6.1: GGNN.

This paper presents a ranking cnn to deal with dehazing. The ranking-cnn mainly means a ranking layer. In this layer, the values in a feature map are ranked, and the same operation is conducted for each feature map. Moreover, this paper introduce a method to synthesize haze images.

[1]

Bibliography

[1] Jinming Su. For convenient compile. In Nothing, 2019.