Semantic Segmentation

TINGWU WANG
MACHINE LEARNING GROUP,
UNIVERSITY OF TORONTO
Contents

1. What is semantic segmentation?
   1. What is segmentation in the first place?
   2. What is semantic segmentation?
   3. Why semantic segmentation

2. Deep Learning in Segmentation
   1. Semantic Segmentation before Deep Learning
   2. Conditional Random Fields
   3. A Brief Review on Detection
   4. Fully Convolutional Network

3. Discussions and Demos
   1. Demos of CNN + CRF
   2. Segmentation from Natural Language Expression
   3. Make CRF Great Again?
What is semantic segmentation

1. What is segmentation in the first place?
   1. Input: images
   2. Output: regions, structures
      1. line segments, curve segments, circles, etc.
What is semantic segmentation

1. What is segmentation in the first place?
   1. Input: images
   2. Output: regions, structures
   3. Most of the time, we need to "process the image"
      1. filters
      2. gradient information
      3. color information
      4. etc.

That's not quite so human.

What if we want to understand the image?

Arbelaez, Pablo, et al. [1]
What is semantic segmentation

1. What is semantic segmentation?
   1. Idea: recognizing, understanding what's in the image in pixel level.

   "Two men riding on a bike in front of a building on the road. And there is a car."
What is semantic segmentation

1. What is semantic segmentation?
   1. Idea: recognizing, understanding what's in the image in pixel level.
   2. A lot more difficult
   (Most of the traditional methods cannot tell different objects.)
What is semantic segmentation

1. What is semantic segmentation?
   1. Idea: recognizing, understanding what's in the image in pixel level.
   2. A lot more difficult
      (Most of the traditional methods cannot tell different objects.)

No worries, even the best ML researchers find it very challenging.

3. Output: regions with different (and limited number of) classes
   1. COCO detection challenge: 80 classes.
   2. PASCAL VOC challenge: 21 classes
What is semantic segmentation

1. Why semantic segmentation?
   1. robot vision and understanding
   2. autonomous driving (remember your assignment?)
What is semantic segmentation

1. Why semantic segmentation?
   3. medical purposes (ISBI Challenge)
Contents

1. What is semantic segmentation?
   1. What is segmentation in the first place?
   2. What is semantic segmentation?
   3. Why semantic segmentation

2. Deep Learning in Segmentation
   1. Semantic Segmentation before Deep Learning
   2. Conditional Random Fields
   3. A Brief Review on Detection
   4. Fully Convolutional Network

3. Discussions and Demos
   1. Demos of CNN + RCF
   2. Segmentation from Natural Language Expression
   3. Make CRF Great Again?
Deep Learning in semantic Segmentation

1. Semantic segmentation before deep learning
   1. relying on conditional random field.
   2. operating on pixels or superpixels
   3. incorporate local evidence in unary potentials
   4. interactions between label assignments

J Shotton, et al. [3]
Deep Learning in semantic Segmentation

1. What is conditional random field?
   1. probabilistic framework for labeling and segmenting structured data
   2. no need to understand the math, just know the idea

what it tries to model is the relationship between pixels, e.g.:
1. nearby pixels more likely to have same label
2. pixels with similar color more likely to have same label
3. the pixels above the pixels "chair" more likely to be "person" instead of "plane"
4. refine results by iterations
Deep Learning in Semantic Segmentation

1. A Brief Review on Classification
   0. Again, it is totally fine if you don't understand the deep neural network.
      
      imagine it as a black magic box if you want :)
      1. input: the whole image
      2. output: the probability of each class (person, dog, cat, ...)
      3. not appliable on semantic segmentation

Deep Learning in Semantic Segmentation

1. How to move from classification to semantic segmentation?
   1. Remember traditionally we use superpixels (Polygon)?
Deep Learning in Semantic Segmentation

1. Transition to segmentation; early ideas
   1. superpixel proposals
   2. do classification on each superpixel.

M Mostajabi, et al. [6]
Deep Learning in Semantic Segmentation

1. Fully Convolutional Networks for Semantic Segmentation
   1. forget about pixels/superpixel input

Long, J., et al. [8]
Deep Learning in Semantic Segmentation

1. Fully Convolutional Networks for Semantic Segmentation

Long, J., et al. [8]
Deep Learning in Semantic Segmentation

1. Fully Convolutional Networks + CRF
   1. the output from DCNN is blurry and inaccurate
   2. rediscovery of CRF

---

Image/G.T.  DCNN output  CRF Iteration 1  CRF Iteration 2  CRF Iteration 10

LC Chen, et al. [9]
Deep Learning in Semantic Segmentation

1. Conditional Random Fields as Recurrent Neural Networks
   1. end-to-end training
      \[ \text{optimize}(A) + \text{optimize}(B \text{ given } A) < \text{optimize}(A, B \text{ together}) \]
1. What is semantic segmentation?
   1. What is segmentation in the first place?
   2. What is semantic segmentation?
   3. Why semantic segmentation

2. Deep Learning in Segmentation
   1. Semantic Segmentation before Deep Learning
   2. Conditional Random Fields
   3. A Brief Review on Detection
   4. Fully Convolutional Network

3. Discussions and Demos
   1. Demos of CNN + RCF
   2. Segmentation from Natural Language Expression
   3. Make CRF Great Again?
Discussions and Demos

1. Online Demos about CRF as RNN semantic segmentation
Discussions and Demos

1. Segmentation from Natural Language Expression
   1. what does it mean?
   
   e.g., the phrase "two men sitting on the right bench" requires segmenting only the two people on the right bench and no one standing or sitting on another.

(a) input image  (b) object class segmentation of class *people*  (c) object instance segmentation of class *people*  (d) segmentation from expression “people in blue coat”
Discussions and Demos

1. Segmentation from Natural Language Expression
Discussions and Demos

1. Make Probabilistic Graphical Model Great Again?
   1. what happened to DPM [12]
      1. mixtures of multiscale deformable part models
      2. later people found DPM could be placed by a CNN layer [13]
      3. no one uses dpm now.
   2. what happened to object proposals in detection
      1. Human designed proposals (selective search, edge box, ...) [14]
      2. later people found proposal generating could be replaced by a CNN layer [15, 16]
      3. no one (well, maybe still many people) uses human designed proposals now.

3. what is happening to CRF in semantic segmentation
   1. pairwise relationship between pixels
   2. later people find CRF could be replaced by a CNN layer
   3. no one uses CRF? well, we don't know future
Discussions and Demos

1. The powerfulness of deep learning

Agent Smith: If you can't beat us...

Agent Smith Clone: Join us!
References


References


Q&A

For those who are interested in CRF and want to know the math, I recommend this tutorial:


(might take a long time to understand. good luck ;P)