Combining Conditional Random Fields with Deep Neural Networks for Semantic Segmentation

Rocherno F.M. de Jongh

University of Minnesota, Morris

UMM CSci Senior Seminar Conference, April 12 2018

Introduction

- It is easy for humans to see an image and immediately classify and understand what is on the image

Introduction

- It is easy for humans to see an image and immediately classify and understand what is on the image

Figure: http://www.pawbuzz.com/wp-content/uploads/sites/551/2014/11/corgi-puppies-21.jpg

Introduction

- It is easy for humans to see an image and immediately classify and understand what is on the image

Figure: http://www.pawbuzz.com/wp-content/uploads/sites/551/2014/11/corgi-puppies-21.jpg

- Can we make computers emulate this?

Introduction

Example

- In radiation treatment planning, radiologists need to compute best path for applying radiation

Introduction

Example

- In radiation treatment planning, radiologists need to compute best path for applying radiation
- Radiologists manually trace outlines on CT or MR images

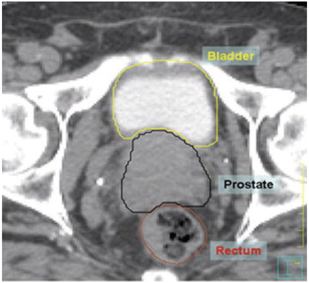


Figure: https://radiologykey.com/segmentation-of-pelvic-structures-from-ct-scans-for-planning-in-prostate-cancer-radiotherapy/

Outline

Outline

Introduction to Image Segmentation

Introduction to Image Segmentation

- In computer vision, image segmentation is the process of partitioning a digital image into multiple segments

Introduction to Image Segmentation

- In computer vision, image segmentation is the process of partitioning a digital image into multiple segments
- The goal of image segmentation is to cluster pixels into relevant image segments

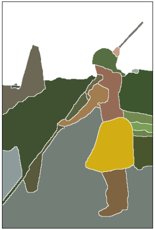


Figure: https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/resources.html

Introduction to Image Segmentation

- What if we want to partition the image in a more meaningful way?

Introduction to Image Segmentation

- What if we want to partition the image in a more meaningful way?
- What if we also want to *understand* the image?

Introduction to Semantic Segmentation

Introduction to Semantic Segmentation

- Semantics means the study of meanings

Introduction to Semantic Segmentation

- Semantics means the study of meanings
- That is what we are trying to do, study of meanings in images

Introduction to Segmentation

- **Semantic segmentation** is the understanding of an image at pixel level

Introduction to Segmentation

- **Semantic segmentation** is the understanding of an image at pixel level
- Semantic segmentation assigns a predefined object class to each pixel within the image.

Introduction to Segmentation

- **Semantic segmentation** is the understanding of an image at pixel level
- Semantic segmentation assigns a predefined object class to each pixel within the image.
  - Takes an image as an input

Introduction to Segmentation

- **Semantic segmentation** is the understanding of an image at pixel level
- Semantic segmentation assigns a predefined object class to each pixel within the image.
  - Takes an image as an input
  - Returns an image with all of the pixels in the image labeled

Examples of Semantic Segmentation

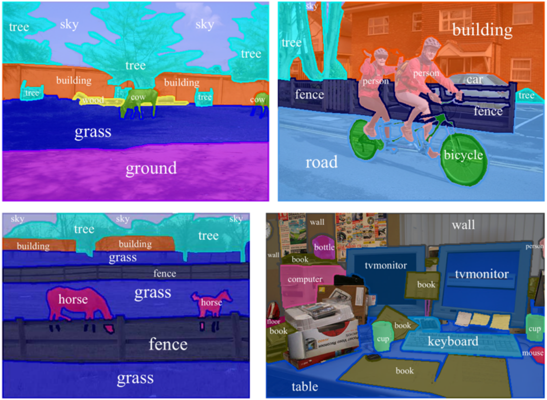


Figure: http://www.stat.ucla.edu/ xianjie.chen/

Applications of Semantic Segmentation

Semantic segmentation has numerous applications such as: Road Segmentation for Autonomous Driving

Figure: http://tex.stackexchange.com/

Introduction to Semantic Segmentation

- Semantic segmentation has numerous applications such as: synthetic "shallow depth-of-field effect shipped in the portrait mode" of the Pixel 2 and Pixel 2 XL smartphones.

Figure: https://research.googleblog.com/2017/10/portrait-mode-on-pixel-2-and-pixel-2-xl.html

Introduction to Semantic Segmentation

- Semantic segmentation has numerous applications such as: Mobile Real-Time Video Segmentation

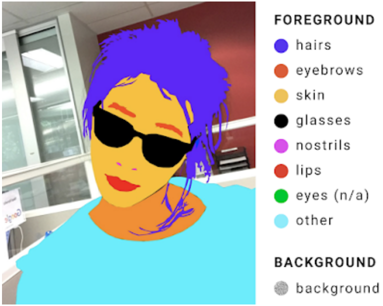


Figure: https://research.googleblog.com/2018/03/mobile-real-time-video-segmentation.html

Introduction to Semantic Segmentation

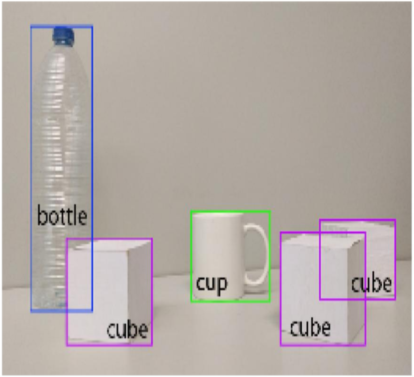
- Image classification

Introduction to Semantic Segmentation

- Image classification
- What's in the image?

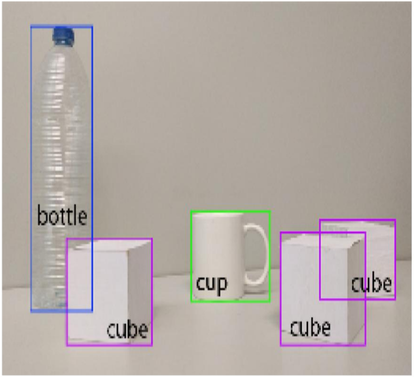
Source: http://arxivst.com/papers/2017/04/22/a-review-on-deep-learning-techniques-applied-to-semantic-segmentation/

Introduction to Semantic Segmentation



- Object detection

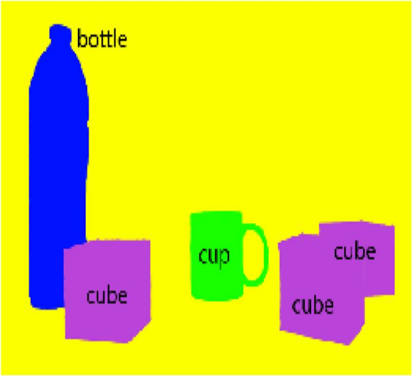
Introduction to Semantic Segmentation



- Object detection
- Box each object in image

Source:
http://arxivst.com/papers/2017/04/22/a-review-on-deep-learning-techniques-applied-to-semantic-segmentation/

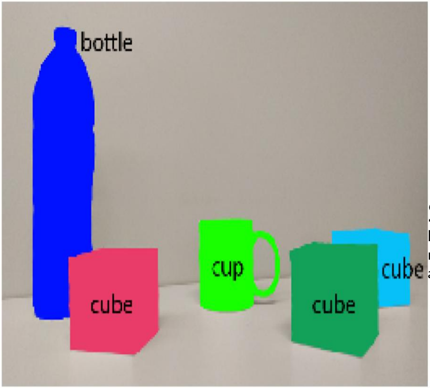
Introduction to Semantic Segmentation



- Semantic Segmentation

Source:
http://arxivst.com/papers/2017/04/22/a-review-on-deep-learning-techniques-applied-to-semantic-segmentation/

Introduction to Semantic Segmentation

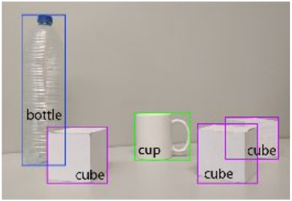


- Instance Segmentation

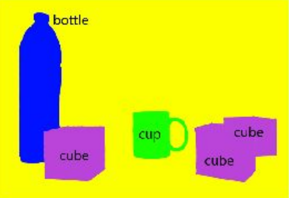
Source:
http://arxivst.com/papers/2017/04/22/a-review-on-deep-learning-techniques-applied-to-semantic-segmentation/

Introduction to Semantic Segmentation

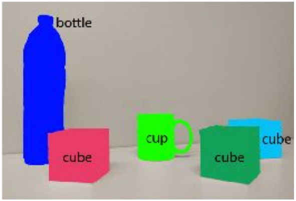
(a) Image classification



(b) Object detection



(c) Semantic segmentation



(d) Instance segmentation

Outline

Statistical Background

Statistical Background

- **Random variable**
  - A variable that its value depends on a random phenomenon

Statistical Background

- **Random variable**
  - A variable that its value depends on a random phenomenon
- **Probability distribution**
  - A distribution of the probabilities of the possible events that a random variable can take

Structured Probabilistic Models

Structured Probabilistic Models

- **Structured Probabilistic Models** is a way of describing a probability distribution, using a graph

Structured Probabilistic Models

- **Structured Probabilistic Models** is a way of describing a probability distribution, using a graph
- In a probabilistic graphical model, each node represents a random variable and the edges represent a probabilistic relationship between these random variables.

Structured Probabilistic Models

- **Directed graphical models** use graphs with directed edges

Structured Probabilistic Models

- **Directed graphical models** use graphs with directed edges
- **Undirected graphical models** use graph with undirected edges

Structured Probabilistic Models

- **Directed graphical models** use graphs with directed edges
- **Undirected graphical models** use graph with undirected edges
- We care about undirected graphical models, also called **Markov Random Fields**

Markov Random Fields

- Each pixel in an image will correspond to a node in a graph

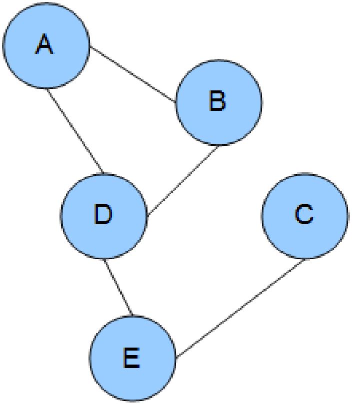
Markov Random Fields

- Each pixel in an image will correspond to a node in a graph
- Variables, such as colors in an image, are introduced to explain values of the nodes

Markov Random Fields

- Each pixel in an image will correspond to a node in a graph
- Variables, such as colors in an image, are introduced to explain values of the nodes
- Probability two nodes having the same value

Markov Random Fields



In this example: A depends on B and D. B depends on A and D. D depends on A, B, and E. E depends on D and C. C depends on E.

Source:
http://arxivst.com/papers/2017/04/22/a-review-on-deep-learning-techniques-applied-to-semantic-segmentation/

- **Conditional Random Fields** are an important special case of Markov Random Fields.

- **Conditional Random Fields** are an important special case of Markov Random Fields.
- Checks probability a node (pixel) is a certain value, given another know node value

Conditional Random Fields

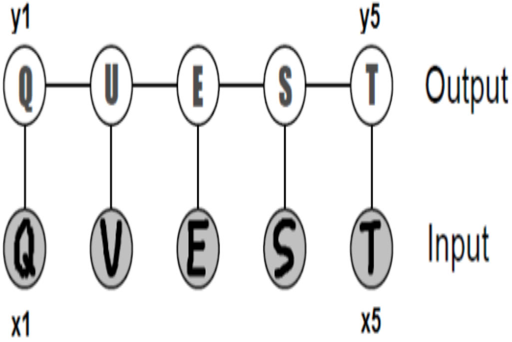


Figure: https://ermongroup.github.io/cs228-notes/representation/undirected/

Outline

Neural Networks

Neural Networks

- **Neural networks** are means for doing machine learning

Neural Networks

- **Neural networks** are means for doing machine learning
- Neural networks are generally comprised of layers of nodes

Neural Networks

- **Neural networks** are means for doing machine learning
- Neural networks are generally comprised of layers of nodes
- Each node contains an *activation function*

Neural Networks

- **Neural networks** are means for doing machine learning
- Neural networks are generally comprised of layers of nodes
- Each node contains an *activation function*
- Result in some form of output

Neural Networks

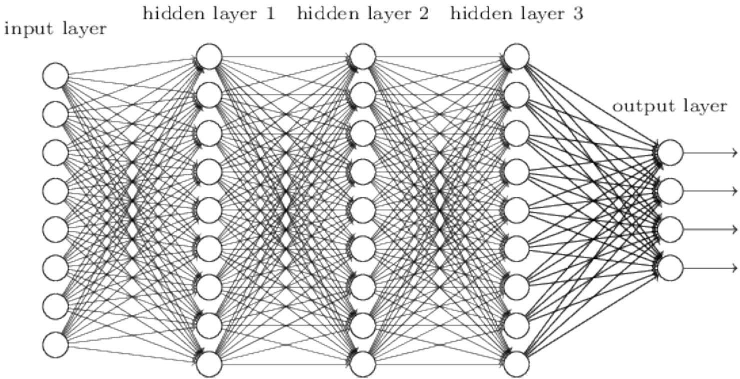


Figure: http://neuralnetworksanddeeplearning.com/chap6.html

Convolutional Neural Network

- Almost the same neural networks

Convolutional Neural Network

- Almost the same neural networks
- Convolutional Neural Networks (CNN) contains convolutional layer

Convolutional Layer

- Used to condense the input data into recognized patterns

Convolutional Layer

- Used to condense the input data into recognized patterns
- This layer uses a filter, which is an array recognizing certain feature

Convolutional Layer

- Used to condense the input data into recognized patterns
- This layer uses a filter, which is an array recognizing certain feature
- *Feature map* is an array containing all of the results of the convolutions

Convolutional Layer

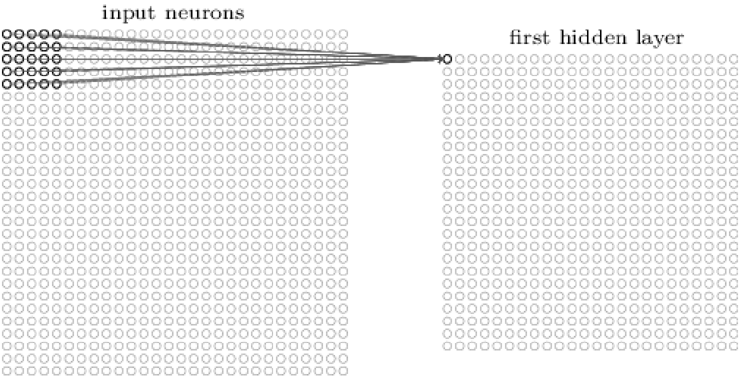


Figure: http://neuralnetworksanddeeplearning.com/chap6.html

CNN

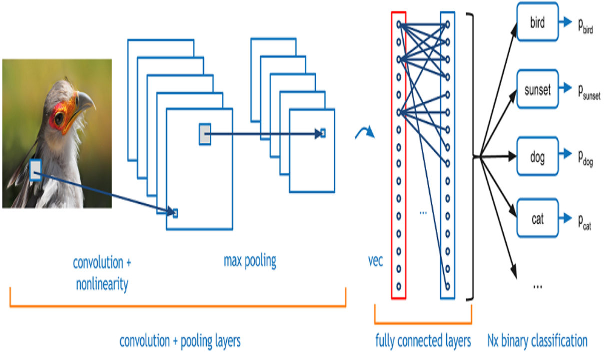


Figure: https://adeshpande3.github.io/A-Beginners-Guide-To-Understanding-Convolutional-Neural-Networks/

Fully Connected Layer

- Usually last layer

Fully Connected Layer

- Usually last layer
- Take an input from the feature map, and returns a vector of label probabilities

CNN

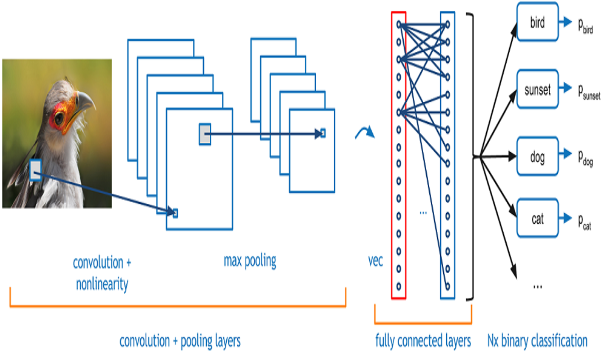


Figure:
https://adeshpande3.github.io/A-Beginner27s-Guide-To-Understanding-Convolutional-Neural-Networks/

CNN for pixel prediction

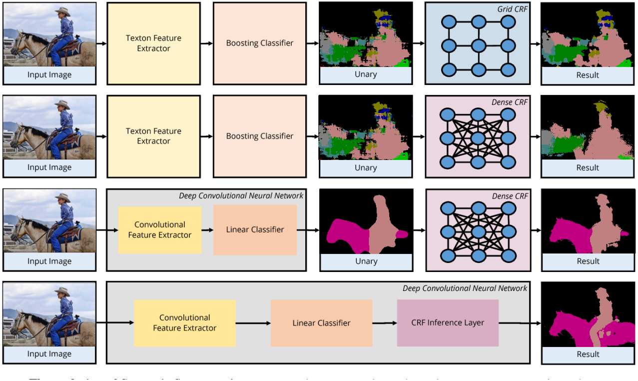


Figure: [3]

Combining DNN with CRF

- Let's look at a demo!
- http://www.robots.ox.ac.uk/ szheng/crfasrnndemo

Results

Method	IoU [%]	Base Network
Methods not using deep learning		
O2P [36]	47.8	–
Methods not using a CRF		
SDS [37]	51.6	AlexNet
FCN [6]	67.2	VGG
Zoom-out [38]	69.6	VGG
Methods using CRF for postprocessing		
DeepLab [5]	71.6	VGG
EdgeNet [39]	73.6	VGG
BoxSup [40]	75.2	VGG
Dilated Conv [27]	75.3	VGG
Centrale Boundaries [41]	75.7	VGG
DeepLab Attention [42]	76.3	VGG
LRR [30]	79.3	ResNet
DeepLab v2 [43]	79.7	ResNet
Methods with end-to-end CRFs		
CRF as RNNs [7]	74.7	VGG
Deep Gaussian CRF [8]	75.5	VGG
Deep parsing network (DPN) [44]	77.5	VGG
Context [32]	77.8	VGG
Higher-order CRF [33]	77.9	VGG
Deep Gaussian CRF [8]	80.2	ResNet

Figure: [3]

Acknowledgements

Special thanks to Professor Peter Dolan and Professor Elena Machkasova for their guidance and feedback.

References I

[1] Eclipse Deeplearning4j Development Team. Deeplearning4j: Open-source distributed deep learning for the JVM, Apache Software Foundation License 2.0. http://deeplearning4j.org.

[2] Anurag Arnab et al. "Higher Order Potentials in End-to-End Trainable Conditional Random Fields". In: *CoRR* abs/1511.08119 (2015). arXiv: 1511.08119. URL: http://arxiv.org/abs/1511.08119.

[3] A. Arnab et al. "Conditional Random Fields Meet Deep Neural Networks for Semantic Segmentation: Combining Probabilistic Graphical Models with Deep Learning for Structured Prediction". In: *IEEE Signal Processing Magazine* 35.1 (Jan. 2018), pp. 37–52. ISSN: 1053-5888. DOI: 10.1109/MSP.2017.2762355.

[4] Siddhartha Chandra and Iasonas Kokkinos. "Fast, exact and multi-scale inference for semantic image segmentation with deep gaussian crfs". In: *European Conference on Computer Vision*. Springer. 2016, pp. 402–418.

References II

[5] L. C. Chen et al. "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40.4 (Apr. 2018), pp. 834–848. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2017.2699184.

[6] Alberto Garcia-Garcia et al. "A review on deep learning techniques applied to semantic segmentation". In: *arXiv preprint arXiv:1704.06857* (2017).

[7] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

[8] Larry Hardesty. *Explained: Neural networks*. http://news.mit.edu/2017/explained-neural-networks-deep-learning-0414. Apr. 2017.

References III

[9] Roman Klinger and Katrin Tomanek. *Classical Probabilistic Models and Conditional Random Fields*. Department of Computer Science, Dortmund University of Technology, 2007.

[10] Volodymyr Kuleshov and Stefano Ermon. *CS 228: Probabilistic Graphical Models, Lecture Notes*. https://cs.stanford.edu/ ermon/cs228/index.html. 2017.

[11] John Lafferty, Andrew McCallum, and Fernando CN Pereira. "Conditional random fields: Probabilistic models for segmenting and labeling sequence data". In: (2001).

[12] E. Shelhamer, J. Long, and T. Darrell. "Fully Convolutional Networks for Semantic Segmentation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.4 (Apr. 2017), pp. 640–651. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2016.2572683.

References IV

[13] S. Zheng et al. "Conditional Random Fields as Recurrent Neural Networks". In: *2015 IEEE International Conference on Computer Vision (ICCV)*. Dec. 2015, pp. 1529–1537. DOI: 10.1109/ICCV.2015.179.

Discussion

Questions?