MA 598 Machine Learning Seminar
Summary 4

ID 5107

January 25, 2019

4 U-Net: Convolutional Networks for Biomedical Image Segmentation

4.1 Introductionss

CNNs have been performing very well

There are hurdles in some domains

Biomedical image processing needs localization segmentation

Hard to get lots of data

Hard to distinguish neighboriing cells

4.2 Past

Training in sliding window setup label for each pixel plus patch

Pros: localization, more training samples

Cons: slow, overlapping patches (which are inefficient)

Tradeoff: localization accurasy vs context

Large patches implies less localization, and smaller paches imply less context.

4.3 U-Net

A unet is a fully convolutional network. works well with very few images to train. It has better segmentation. The key ideas: starts with a set of contraction layers which caper the semantic/contextual information. Expansion layers which recover spatial information (use upsamlping to increase output resolution).

Overlap tile strategy: Arbitrarily large images, not limite by GPU size.

They also suggest ...

4.4 Training

Input images and corresponding segmentation maps are used for training.

Uses a stochastic gradient descent.

Energy function: pixel-wise softmax of feature map with cross entropy loss.

The pixel wise softmax equation isomorphic softmax

$$p_k(\mathbf{x}) = \exp(a_k(\mathbf{x}))/\left(\sum_{k'=1}^{K} \exp(a_{k'}(\mathbf{x}))\right)$$

$K = 2$

$\mathbf{x}$ pixel position in an image. $a_k(\mathbf{x})$ output in feature channel k at pixel $\mathbf{x}$

$p_k(\mathbf{x}) \approx 1$ for the channel with maximum $a_k(\mathbf{x})$ value. $p_k(\mathbf{x}) \approx 0$ for all other k.

Weighted loss

$$E = -\sum_{\mathbf{x}\in\Omega} w(\mathbf{x}) \log(p_{l(\mathbf{x})(\mathbf{x})})$$

$y = l(\mathbf{x})$ ground truth at pixel $\mathbf{x}$ (background vs foreground).

So

$$p_y(\mathbf{x}) = \begin{cases} 1 & \text{if corrected prediction,} \\ 0 & \text{otherwise.} \end{cases}$$

4.5 Data augmentation

Very few images.

Augmentation is essential for invariance and robustness.

Microscopic images have a lot of variations!

Shift and rotation invariance.

Deformations . Gray value variations.

4.6 Experiments

Neuronal structures in electron microscope.

30 training images.

Evaluation: warping, rand, pixel error

Better results than sliding-window CNN.

Cell segmentation in light microscopic images.

Cell segmentation in light micrscopic images.

35 training images. Average IOU of 92% (much better than the next best 83%.

Units outperform traditional methods. They are easy to extend. Need very few training images, and have a very reasonable training time.

4.7 What is their primary result?

4.8 Why is this important?

4.9 What are their key ideas?

4.10 What are the limitations, either in performance or applicability?

4.11 What might be an interesting next step based on this work?

4.12 What's the architecture?

4.13 How did they train and evaluate it?

4.14 Did they implement something?