Kendall Weihe

CS 463G Midterm – draft #1

Notes:

* General outline
  + Abstract
  + Intro
  + Related work
  + The paper’s implementation of a RNN-CRF
  + My implementation
  + Results
  + Conclusion
  + References
* **Intro:**
  + Success of CNN
    - Original/big papers
    - Drawbacks of CNN
      * *Coarse outputs*
      * *smoothness*
  + Success of RNN
    - Original/big papers
    - Drawbacks of RNN – LSTM
      * Vanishing/exploding gradient
  + Pixel-wise classification (segmentation)
    - Successful papers
    - Difficult obstacles
  + Potential applications
* **Related work:**
  + **…**
* **The paper’s implementation:**
  + What a CRF is
    - Formulas (use previous papers formulas)
  + What a mean-field iteration is
  + Each layer
    - Initialization
    - Message passing
    - Weighting filter outputs
    - Compatibility transform
    - Adding unary potentials
    - Normalization
  + CRF as RNN
* **My implementation**
  + Tensorflow
  + Learning rate, dropout
  + CNN
    - X conv layers
    - Y deconv layers
  + CRF RNN
    - For loop
    - Conv layers
    - Code snippet
  + Cost function
  + Loss computation
  + *Overlapping predictions – smoothing function*
* **Results:**
  + Dataset
    - Input data
      * Dimensions
      * Cropped volume
    - Ground truth data
      * How was it collected?
      * From VC app
      * Fixed width – gradient
  + Results graph
    - Loss value / time
  + Sample prediction?
* **Conclusion:**
  + Recap intro
  + Recap my implementation
  + Compare results to strictly CNN
  + Potential improvements
* **References:**
  + **…**

# Abstract:

*Deep Convolutional Neural Networks (CNN) have shown state of the art performance in the previous five years. Recurrent Neural Networks (RNN) have recently resurfaced as a popular deep learning architecture. While there has been various works on image segmentation over the last five years, the field of applying machine learning to volumetric image segmentation remains relatively undiscovered. In this paper, I apply part of the network outlined in [PAPER] to achieve precise pixel-wise classification on microcomputed tomography volumetric scans of ancient scrolls. The network combines the strengths of Convolutional Neural Networks, the* ***reinforcement*** *of Recurrent Neural Networks, and Condition Random field based probabilistic model. The dataset is part of [REFERENCE] the EDUCE project, where we virtually unwrap scrolls from antiquity in order to conserve the physical structure. The goal of applying machine learning to the EDUCE project is to automate the current bottleneck – the segmentation component. In terms of this paper, image segmentation is the process of isolating voxels located on scroll surfaces.*

# Introduction:

**[WHAT ARE CONVOLUTIONAL NEURAL NETWORKS?]** Convolutional Neural Networks (CNNs) have become the popular choice for vision tasks such as classification, object detection, and image segmentation. The success of CNNs is attributed to the networks ability to learn hierarchical representations of input images without handcrafted techniques (reference?). One technical hurdle of a CNN is signal down-sampling – a result of reducing image dimensions through max-pooling or non-unit strides. [REFERENCE] proposed a solution to this problem by introducing Convolutional Transpose layers – or the process of upsampling, commonly incorrectly labeled deconvolution. Convolutional Transpose layers have shown to be very useful for image segmentation [REFERENCE] – allowing the network to capture local information in downsampling and global information in upsampling [REFERENCE – Vnet].

**[WHAT ARE CRFS?]** A Conditional-Random Field (CRFs) is a statistical model typically used in machine learning for structured predictions. In the context of pixel-wise classification, a CRF contains a series of weights applied to each pixel to obtain a global observation. [10 from paper] showed how the output of a CNN can serve an input to a CRF as a post-processing step. [PAPER] combined the to two disjoint components in an end-to-end trainable system, claiming that the combination would outperform a disjoint system. Their results achieved state-of-the-art accuracy on the Pascal VOC benchmark. [REFERENCE] showed how Convolutional Layers can be used to implement a Conditional-Random Field for image segmentation. Furthermore, [REFERENCE] applied a Recurrent Neural Network (RNN) to the network to form a what they call a RNN-CRF end-to-end network.

The dataset used in training and testing for this paper comes from Dr. Breant Seales' (University of Kentuky) *Enhanced Digital Unwrapping for Convservation and Explorations* (EDUCE) project. The goal of the EDUCE project is to apply modern microcomputed tomography scanning techniques to discover unreadable works of ancient literature – typically ancient scrolls. One component of the unwrapping pipeline is the segmentation process. In order to digitally unwrap a volumetric scan of a scroll, voxels located on the surfaces of scrolls must be isolated. Where subsequently a three dimensional mesh can be generated and unwrapped. The goal of this paper is to apply parts of the network developed in (RNN-CRF paper) to a network capable of segmenting voxels located on scroll surfaces.

# Related Work:

[REFERENCE Seales' paper] applied [GREEDY APPROACHES & EIGENVECTORS] to the scroll segmentation process.

To date, no machine learning techniques have been applied to automate the segmentation process of uncovering volumetric scans of ancient literature.

Vnet, upsampling papers, fully connected Convs

Fully connected CRF's for semantic image segmentation

Conditional Random Fields as a Recurrent Neural network

To date, there are no cases where in which machine learning techniques have been applied to automating the segmentation component of uncovering volumetric scans of ancient literate. However, [Seales & people] showed how

# 3. CRFs as RNN:

## 3.1 Conditional Random Fields

…

## 3.2 Mean-field Iteration

* “based on the observation that filter-based approximate mean-field inference approach for dense CRFs relies on applying Gaussian spatial and bilateral filters on the mean-field approximates in each iteration”
* implemented as a stack of convolutional layers
  + trainable CRF parameters are the same as CNN filters
* a mean-field iteration can be implemented via a RNN
* in doing so, the network can back-propagate the error differentials

## 3.21 Message Passing

* reflects how strongly a pixel is related to another pixel
  + pixel location
  + pixel value
* they implement a fully connected (fully connected MLP) CRF
* permutohdral lattice implementation
* (talk about this in the *my model* section) – due to time constraints I did not do fully connected, I did locally connected – there are drawbacks because of padding

…

## 3.22 Weighting Filter Outputs

* taking the weighted sum of the M filter outputs from the previous step
* 1x1 filter – M input channels and 1 output channel
* the intuition is that…

…

## 3.23 Compatibility Transform

* assigns penalty if different labels are assigned to pixels with similar properties

…

## 3.24 Adding Unary Potentials

…

## 3.25 Normalization

…

## 3.26 Implementing mean-field iteration as a RNN

…