CAP 6412 Advanced Computer Vision


Boqing Gong
March 29, 2016
Today

• Recurrent Neural Networks (RNNs): Backpropagation
  • Whiteboard

• Biomedical imaging, by Dustin
Locality Sensitive Deep Learning For Detection And Classification Of Nuclei In Routine Colon Cancer Histology Images

Sirinukunwattana, K., Raza, S. E. A., Tsang, Y. W., Snead, D., Cree, I. A., & Rajpoot, N. M.

Presented by Dustin Morley
Outline

- Overview of Biomedical Computer Vision
  - Types of Problems
  - Research personnel and approaches
  - Unique Challenges
- Paper Presentation (Sirinukunwattana et al.)
  - Problem Statement
  - Methods
  - Results
Biomedical C.V. - Introduction

- Huge variety of interesting medical imaging modalities
  - Tomography (X-ray, MRI, PET, etc.), ultrasound, microscopy
  - Even conventional cameras (special case: ophthalmology)
- Work closely with medical doctors
- In U.S. – accountable to FDA
  - Device approvals – 510k or PMA → official claims
  - Software can get a little gray
    - FDA trying to find ways to avoid unnecessarily stifling key technological progress
  - Routinely audit companies with FDA-approved products
Computer Vision Problems in Medicine

- Recognition/Diagnosis
- Segmentation
  - 2D or 3D, depending on modality
- Registration
  - Same device, different exam
  - Different device but same modality
  - Most challenging – different modality
- 3D Reconstruction from 2D images/slices
  - With or without segmentation
- Calibration of imaging devices

Image Credit: LENSAR, LLC
Medical Research – M.D.’s and Scientists

- Medical doctors
  - Majority of papers – present measurement and/or procedure outcome data, and look for correlations (essentially clinical studies)
  - Important minority – present a newly discovered challenge and some basic manual technique for navigating the challenge

- Scientists
  - Like M.D.’s, some work presenting measurement data and looking for correlations
    - Statistical work a bit more complex (explore nonlinear relationships, etc.)
  - Present new techniques (i.e. imaging techniques)
  - Build comprehensive models
Medical Research - Engineers

- Specific implementation design of medical equipment
- Explore ways to automate parts of medical procedures
  - Robotics, image segmentation/registration, etc.
- Focus on solving well-defined problems that can be objectively evaluated
  - Ex. – segmentation and registration given problem-defining ground truth, improving SNR of imaging system, etc.
- Either working for/with a medical device company or doing research alongside a strong university medical program (e.g. Harvard)
Biomedical Engineering Approaches

- Doctors voice challenge → engineers build new technology to address this challenge
  - Ex. – image registration
- Engineers look at what doctors are doing → design technology to improve efficiency, repeatability, accuracy, etc.
  - Ex. – diagnostic algorithms, surgical robotics
- Often need to convert subjective or “big picture” ideas from doctors into objective goals, or “read between the lines” to figure out what they actually want/need
Example – Converting Doctors’ “Big Picture” Ideas into an Engineered System

Image Credit: LOCS Study Group

Image Credit: LENSAR, LLC
Medical Computer Vision – Unique Challenges

- No public datasets available for most tasks
  - Remember – variety of modalities and systems
  - Limits ability to directly compare results with other published work
- Obtaining data, especially large samples, is very difficult
  - Regulatory hurdles (FDA, HIPAA)
  - Data transportation (doctors and medical staff aren’t always technically skilled)
  - Scheduling (doctors are busy people!)
- For tasks directly impacting medical outcomes, extremely reliable performance required
  - Chances of algorithm making things worse for the patient than if your algorithm wasn’t used at all need to be virtually nonexistent
Nuclei Detection/Classification in Histology Images – Problem Statement

- Tissue sample from tumour → microscope slide (with staining) → histology image
- Several nuclei in each image, high degree of heterogeneity
- Goal – detect and classify every nucleus in the image
- Motivation – a new quantitative analysis tool available for pathologists in grading severity and planning treatment
Dataset

- **100** total stained histology images colorectal adenocarcinomas, cropped from 10 whole-slide images from **9** patients
- Each image 500x500 pixels at 0.55 µm/pixel → 275x275 µm
- Ground truth – manual annotation by an experienced pathologist
  - Total of 29,756 nuclei marked for detection
  - 22,444 also given a class label (remaining 7,312 nuclei unlabeled)
  - 7,722 epithelial, 5,712 fibroblast, 6,971 inflammatory, 2,039 miscellaneous
Spatially Constrained CNN

- Idea: CNN output as function of (x,y) should ideally be as continuous as possible, with a peak at location of each nucleus
- Assume maximum of M nuclei present
- 2 special layers after the fully connected layer realize this constraint
- Parameter Estimation Layer – 3M sets of weights to be learned
  - $u_m, v_m, h_m$ - 2D location and probability map height of $m^{th}$ nucleus center
- Spatially Constrained Layer – output Lorentzian probability distribution over distance to closest $u_m, v_m$ jointly multiplied by $h_m$
Spatially Constrained CNN

\begin{align*}
u_m &= (H' - 1) \cdot \text{sigm}(W_{L-1,u_m} \cdot x_{L-2} + b_{u_m}) + 1, \\
v_m &= (W' - 1) \cdot \text{sigm}(W_{L-1,v_m} \cdot x_{L-2} + b_{v_m}) + 1, \\
h_m &= \text{sigm}(W_{L-1,h_m} \cdot x_{L-2} + b_{h_m}),
\end{align*}

What I think they mean:

\[
\hat{y}_j = \begin{cases} 
\left( \frac{1}{1 + \left(\left\|z_j - \overline{z}_m^0\right\|_2^2 \right) / 2} \right) h_m, & \left\|z_j - \overline{z}_m^0\right\|_2^2 < d \text{ AND } \forall m \neq m' \left\|z_j - \overline{z}_m^0\right\|_2^2 \leq \left\|z_j - \overline{z}_{m'}^0\right\|_2^2 \\
0, & \text{otherwise}
\end{cases}
\]
Nucleus Classification

- Separate CNN trained for classification given detection
- Best method tried – Neighboring Ensemble Predictor
  - Try a few adjacent patches
  - Weighted collection of outputs forms final classification
## Network Architecture

<table>
<thead>
<tr>
<th>Layer</th>
<th>SC-CNN for detection</th>
<th>softmax CNN for classification</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Type</td>
<td>Filter Dimensions</td>
</tr>
<tr>
<td>0</td>
<td>I</td>
<td>27 $\times$ 27 $\times$ 1</td>
</tr>
<tr>
<td>1</td>
<td>C</td>
<td>4 $\times$ 4 $\times$ 1 $\times$ 36</td>
</tr>
<tr>
<td>2</td>
<td>M</td>
<td>2 $\times$ 2</td>
</tr>
<tr>
<td>3</td>
<td>C</td>
<td>3 $\times$ 3 $\times$ 36 $\times$ 48</td>
</tr>
<tr>
<td>4</td>
<td>M</td>
<td>2 $\times$ 2</td>
</tr>
<tr>
<td>5</td>
<td>F</td>
<td>5 $\times$ 5 $\times$ 48 $\times$ 512</td>
</tr>
<tr>
<td>6</td>
<td>F</td>
<td>1 $\times$ 1 $\times$ 512 $\times$ 512</td>
</tr>
<tr>
<td>7</td>
<td>S1</td>
<td>1 $\times$ 1 $\times$ 512 $\times$ 3</td>
</tr>
<tr>
<td>8</td>
<td>S2</td>
<td>11 $\times$ 11</td>
</tr>
</tbody>
</table>
Training Data Augmentation

- Arbitrary rotations in increments of 90 degrees and vertical/horizontal flipping
- Arbitrary perturbation of color distribution
  - Used HSV space; random adjustments up to 5% for H, up to 10% for S & V
- Extracted multiple differently located patches of same nucleus
  - Achieves some translation invariance
Training Details

- All weights initialized from zero-mean Gaussian with std. dev. = 0.01
- All biases initialized to zero
- SGD with momentum 0.9, weight decay 0.0005 for 120 epochs
- Learning rate $10^{-2}$ 60 epochs, $10^{-3}$ next 40 epochs, $10^{-4}$ final 20 epochs
- 20% of training data used for validation
Comparative Detection Results

- Compared against 5 other published methods
  - CRImage publicly available, original authors provided implementations for SR-CNN and LIPSyM, CP-CNN and SSAE had to be implemented by authors
  - Networks SC-CNN, SR-CNN and CP-CNN were built as identically as possible

- 2-fold cross validation (50 images/fold)

<table>
<thead>
<tr>
<th>Method</th>
<th>Precision</th>
<th>Recall</th>
<th>F1 score</th>
<th>Median Distance (Q1, Q3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC-CNN ($M = 1$)</td>
<td>0.758</td>
<td><strong>0.827</strong></td>
<td>0.791</td>
<td><strong>2.236</strong> (1.414, 5.099)</td>
</tr>
<tr>
<td>SC-CNN ($M = 2$)</td>
<td>0.781</td>
<td>0.823</td>
<td><strong>0.802</strong></td>
<td><strong>2.236</strong> (1.414, 5)</td>
</tr>
<tr>
<td>CP-CNN</td>
<td>0.697</td>
<td>0.687</td>
<td>0.692</td>
<td>3.606 (2.236, 7.616)</td>
</tr>
<tr>
<td>SR-CNN</td>
<td>0.783</td>
<td>0.804</td>
<td>0.793</td>
<td><strong>2.236</strong> (1.414, 5)</td>
</tr>
<tr>
<td>SSAE [7]</td>
<td>0.617</td>
<td>0.644</td>
<td>0.630</td>
<td>4.123 (2.236, 10)</td>
</tr>
<tr>
<td>LIPSyM [15]</td>
<td>0.725</td>
<td>0.517</td>
<td>0.604</td>
<td><strong>2.236</strong> (1.414, 7.211)</td>
</tr>
<tr>
<td>CRImage [9]</td>
<td>0.657</td>
<td>0.461</td>
<td>0.542</td>
<td>3.071 (1.377, 9.022)</td>
</tr>
</tbody>
</table>
Comparative Classification Results

- Superpixel method had to be implemented by authors
Future Directions

- Computer vision – apply same/similar methods on other medical imaging data
  - CNN for pixelwise classification seems to be a good generic framework that can even be used on small data samples encountered in medical imaging
- Medical – develop the “sophisticated tissue morphometry” methods that can make use of this work
  - Not actually so easy...
  - Ultimately, only useful if output can be mapped to treatment decisions
    - Seems feasible in this case, but don’t expect clinical studies to be finished any time soon