Image-specific Fine-tuning and Uncertainty Estimation for Medical Image Segmentation

Guotai Wang

Translational Imaging Group, Centre for Medical Image Computing
Wellcome / EPSRC Centre for Interventional and Surgical Sciences
University College London
Deep Learning with Convolutional Neural Networks (CNN)

- **Advantages and Disadvantages**
  - Automatic feature learning, end-to-end
  - Requires large amounts of annotated data

- **Fine-tuning of CNNs**
  - Adapt a pre-trained model to a given dataset
1, Image Specific Fine-tuning

Fine-tuning and Image-specific Model

• Fine-tuning of CNNs
  – Adapt a pre-trained model to a given dataset

• Image specific models
  – Self-adaption, e.g. GrabCut

Image-specific CNNs
• Better accuracy?
• Deal with unseen objects?
Fine-tuning and Image-specific Model

- Fetal MR image segmentation
  - Multiple-organs
  - Annotation for all organs?

Placenta | Fetal Brain | Fetal Lung | Maternal Kidney
---|---|---|---
Annotated for training | Unseen during training | Testing

1, Image Specific Fine-tuning
1, Image Specific Fine-tuning

Fine-tuning and Image-specific Model

• Method
  – 1, Use CNN to get an initial segmentation inside a bounding box
  – 2, Fine-tune the CNN with/without scribbles
  – 3, Deal with previously unseen objects

G. Wang et al. 2018, Interactive Medical Image Segmentation using Deep Learning with Image-specific Fine-tuning
Fine-tuning and Image-specific Model

Only the classifier part is fine-tuned for fast update

G. Wang et al. 2018, Interactive Medical Image Segmentation using Deep Learning with Image-specific Fine-tuning
1, Image Specific Fine-tuning

Mathematical Formulation

- **Joint optimization:**
  \[
  \min_{Y, \theta} \left\{ E(Y, \theta) = \sum_i \phi(\hat{y}_i|\hat{X}, \theta) + \lambda \sum_{i,j} \psi(\hat{y}_i, \hat{y}_j|\hat{X}) \right\}
  \]
  subject to: \(\hat{y}_i = s_i\) if \(i \in S\)

- **When \(\theta\) is fixed**
  (Graph Cut problem)
  \[
  \min_Y \left\{ \sum_i \phi'(\hat{y}_i|\hat{X}, \theta) + \lambda \sum_{i,j} \psi(\hat{y}_i, \hat{y}_j|\hat{X}) \right\}
  \]
  \[
  \phi'(\hat{y}_i|\hat{X}, \theta) = \begin{cases} 
  +\infty & \text{if } i \in S \text{ and } \hat{y}_i = s_i \\
  0 & \text{if } i \in S \text{ and } \hat{y}_i \neq s_i \\
  -\log P(\hat{y}_i|\hat{X}, \theta) & \text{otherwise}
  \end{cases}
  \]

- **When \(\hat{Y}\) is fixed**
  (Back propagation)
  \[
  \min_{\theta} \left\{ -\sum_i \left( \hat{y}_i \log p_i + (1 - \hat{y}_i) \log (1 - p_i) \right) \right\}
  \]

G. Wang et al. 2018, Interactive Medical Image Segmentation using Deep Learning with Image-specific Fine-tuning
Weighted Loss Function for Fine-tuning

- Pixels with different confidence
  - Network-based confidence
  - Interaction-based confidence

\[
\text{arg} \min_{\theta} \left\{ -\sum_i w(i) \left( \hat{y}_i \log p_i + (1 - \hat{y}_i) \log(1 - p_i) \right) \right\}
\]

G. Wang et al. 2018, Interactive Medical Image Segmentation using Deep Learning with Image-specific Fine-tuning
Results with Fetal MR Images

- **Unsupervised Fine-tuning**

<table>
<thead>
<tr>
<th></th>
<th>Train</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Placenta</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Fetal Brain</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Fetal Lungs</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Maternal Kidneys</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Patient number</td>
<td>10</td>
<td>6</td>
</tr>
</tbody>
</table>

G. Wang et al. 2018, Interactive Medical Image Segmentation using Deep Learning with Image-specific Fine-tuning
Results with Fetal MR Images

- **Supervised Fine-tuning**
  - Guided by scribbles
  - Only few interactions
  - Real-time update

G. Wang et al. 2018, Interactive Medical Image Segmentation using Deep Learning with Image-specific Fine-tuning
Results with Fetal MR Images

• Quantitative evaluation

### Table II

<table>
<thead>
<tr>
<th></th>
<th>P-Net</th>
<th>P-Net+CRF</th>
<th>BIFSeg(-w)</th>
<th>BIFSeg</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dice (%)</td>
<td>85.57±3.37</td>
<td>84.87±8.14</td>
<td>82.74±10.91</td>
<td>86.41±7.50*</td>
</tr>
<tr>
<td>FB</td>
<td>86.55±6.52</td>
<td>89.09±8.08</td>
<td>90.39±6.44</td>
<td></td>
</tr>
<tr>
<td>FL</td>
<td>85.59±6.42</td>
<td>82.17±8.87</td>
<td>85.35±5.88*</td>
<td></td>
</tr>
<tr>
<td>MK</td>
<td>85.29±5.08</td>
<td>84.61±6.21</td>
<td>86.33±4.28*</td>
<td></td>
</tr>
<tr>
<td>Tm (s)</td>
<td>0.02±0.01*</td>
<td>0.71±0.12</td>
<td>0.72±0.12</td>
<td></td>
</tr>
</tbody>
</table>

P: Placenta, FB: Fetal brain, FL: Fetal lungs, MK: Maternal kidneys.
Uncertainty Estimation for CNNs

• Why uncertainty?
  – Consider the distribution of possible predictions
  – Indicate how reliable a prediction is
  – High uncertainty may related to incorrect prediction

Non-deterministic prediction

Atelectasis?

Aware of potential incorrect prediction?
Uncertainty Estimation for CNNs

Uncertainty types

**Epistemic Uncertainty**
- Model uncertainty
- Can be explained away given enough data
- Approximated by test-time dropout [2]

**Aleatoric Uncertainty**
- From image noises
- Inherent in the observations

Images from [1]

Test-time Augmentation and Uncertainty

- Test-time Augmentation
  - Predict multiple transformed version of the input

Variations in training set

Not mathematically formulated

More possible observations for a testing image
Test-time Augmentation and Uncertainty

• A proposed mathematical formulation

Image Acquisition Model

\[ X = T_\beta(X_0) + e \quad X_0 = T_\beta^{-1}(X - e) \]

• Assumptions

1. The transformation is invertible
2. \( X \) and \( X_0 \) follow the same distribution

Prior distributions of \( \beta \) and \( e \)

CNN prediction

\[ Y = f(\theta, X) \]

Monte Carlo Simulation

\[ E(Y) \approx \frac{1}{N} \sum_{n=1}^{N} y_n = \frac{1}{N} \sum_{n=1}^{N} T_\beta \left( f(\theta, T_\beta^{-1}(X - e_n)) \right) \text{, where } \beta_n \sim P_\beta, e_n \sim P_e \]

Uncertainty estimation

\[ H(Y^i) \approx - \sum_{m=1}^{M} \hat{p}_m^i \ln(\hat{p}_m^i) \]

G. Wang et al. 2018, Test-time augmentation with uncertainty estimation for deep learning-based medical image segmentation
2. Uncertainty Estimation

Results on Fetal Brain Segmentation

- **Test-time augmentation**

<table>
<thead>
<tr>
<th>Flipping</th>
<th>rotation</th>
<th>scaling</th>
<th>noise</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p = 0.5$</td>
<td>$r \sim U(0, 2\pi)$</td>
<td>$s \sim U(0.8, 1.2)$</td>
<td>$e \sim \mathcal{N}(0.0, 0.05)$</td>
</tr>
</tbody>
</table>

Dataset (stacks)
- Train: 120
- Valid: 12
- Test: 48

Monte Carlo simul.
- $N = 20$

G. Wang et al. 2018, Test-time augmentation with uncertainty estimation for deep learning-based medical image segmentation
2. Uncertainty Estimation

Results on Fetal Brain Segmentation

- Quantitative evaluation

- Error rate as a function of uncertainty
Results on 3D Brain Tumor Segmentation

- BraTS dataset

<table>
<thead>
<tr>
<th>Transformation</th>
<th>Probability</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flipping</td>
<td>$p = 0.5$</td>
<td></td>
</tr>
<tr>
<td>Rotation</td>
<td>$r \sim U(0, 2\pi)$</td>
<td></td>
</tr>
<tr>
<td>Scaling</td>
<td>$s \sim U(0.8, 1.2)$</td>
<td></td>
</tr>
<tr>
<td>Noise</td>
<td>$e \sim \mathcal{N}(0.0, 0.05)$</td>
<td></td>
</tr>
</tbody>
</table>

Multi-modal images
- FLAIR
- T1
- T1C
- T2

Dataset
- Train: 215
- Valid: 20
- Test: 50

Monte Carlo simul.
- $N = 40$

G. Wang et al. 2018, Test-time augmentation with uncertainty estimation for deep learning-based medical image segmentation
Results on 3D Brain Tumor Segmentation

- Segmentation Accuracy

<table>
<thead>
<tr>
<th></th>
<th>3D U-Net</th>
<th>V-Net</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>WT</td>
<td>TC</td>
</tr>
<tr>
<td>Baseline</td>
<td>87.69±5.65</td>
<td>78.72±17.96</td>
</tr>
<tr>
<td>TTD</td>
<td>88.22±5.87</td>
<td>79.25±17.90</td>
</tr>
<tr>
<td>TTA</td>
<td>88.39±5.74</td>
<td>79.54±17.11</td>
</tr>
<tr>
<td>TTA + TTD</td>
<td>88.52±5.95</td>
<td>79.61±17.02</td>
</tr>
</tbody>
</table>

- Error rate as a function of uncertainty
Conclusion

• Image-specific fine-tuning
  – Adaptive to a given test image
  – Learn from specific image context
  – Improve segmentation accuracy (supervised or not)
  – Deal with unseen objects

• Uncertainty Estimation
  – Indicate how reliable the result is
  – Model-based and image-based
  – Test-time augmentation and uncertainty
Thanks

Q&A

http://niftynet.io/