Psi-Net: Shape and boundary aware joint multi-task deep network for medical image segmentation

Balamurali Murugesan, Kaushik Sarveswaran, Sharath M Shankaranarayana, Keerthi Ram, Jayaraj Joseph, Mohanasankar Sivaprakasam

Healthcare Technology Innovation Centre, IIT Madras
June 28, 2019
Outline

Slide 3: Deep learning : Medical image segmentation challenges
Slide 5: Related Work
Slide 8: Comparison of related work
Slide 9: Multi task learning (MTL)
Slide 10: Proposed Network - Psi-Net
Slide 12: Dataset and Preprocessing
Slide 14: Quantitative results
Slide 16: Qualitative results
Deep learning: Medical image segmentation challenges

- The segmentation network should be shape-aware.

- The segmentation output produced by the network should ideally have no outliers.

- The boundary of segmentation outputs produced by networks should be smooth.

- The network should be able to handle class-imbalance.

- The network should work for multi-instance object segmentation.
Deep learning: Medical image segmentation challenges

- Outliers
- Non-smooth boundaries.
- Multi-instance object segmentation
- Shape information
Related Work (U-Net) [1]

- U-Net is the most commonly used deep learning network.
- **U-Net**
  - Encoder decoder type of network
  - Input - Image, Output - Probability segmentation map
  - Loss function - Cross entropy

- Shape information: Network not aware of shape
- Outliers/non smooth boundaries: Pixel-wise classification
- Class imbalance - Cross entropy is used
Related work (DCAN: Deep Contour-Aware Networks for Accurate Gland Segmentation) [2]

- Architecture: Single encoder and two parallel decoders
- Decoders are used for mask and contour prediction
- Loss function - Cross entropy for mask and contour
- Shape information - Contour is used to provide shape information
- Class imbalance - Cross entropy has difficulty in handling class imbalance
- Outliers/uneven mask boundaries - Pixel wise classification will result in outliers.
Related work (Deep multi-task and task-specific feature learning network for robust shape preserved organ segmentation (DMTS)) [3]

- Architecture: Single encoder and two parallel decoders
- Decoders are used for mask prediction and distance estimation
- Loss function - Cross entropy for mask and Mean square error for distance map
- Shape information - Distance map is used to provide shape information
- Class imbalance - Learning regression map helps to alleviate the problem
- Outliers/non smooth boundaries - Distance map acts as regularizer to provide smooth boundaries with reduced outliers
Comparison of related work

Multiple object instances - DMTS treats small object as outliers and removes the correct prediction

<table>
<thead>
<tr>
<th></th>
<th>Shape information</th>
<th>Class imbalance</th>
<th>Smooth boundary</th>
<th>Multiple object instances</th>
</tr>
</thead>
<tbody>
<tr>
<td>U-Net</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>✓</td>
</tr>
<tr>
<td>DCAN</td>
<td>✓</td>
<td>X</td>
<td>X</td>
<td>✓</td>
</tr>
<tr>
<td>DMTS</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>X</td>
</tr>
<tr>
<td>Ours</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
Multi task learning (MTL) [4]

- Learning multiple related tasks together will help the model generalize better for the intended task.

- The generalization performance for tasks 1, 2, 3 and 4 is better for model trained together compared to training it for each task separately.
- MTL is a form of inductive transfer.
- L1 regularization is common form of inductive bias
- In case of MTL, inductive bias is provided by auxiliary tasks.
Proposed Network - Psi-Net

- Single encoder, three decoders
- Three decoders learn three tasks: Mask, Contour, and Distance in parallel
- Intended task: Mask, Auxiliary task: Contour, Distance
Psi-Net Loss function

\[ \mathcal{L}_{total} = \lambda_1 \mathcal{L}_{mask} + \lambda_2 \mathcal{L}_{contour} + \lambda_3 \mathcal{L}_{distance} \]

\[ \mathcal{L}_{mask} = \sum_{x \in \Omega} \log p_{mask}(x; l_{mask}(x)) \]

\[ \mathcal{L}_{contour} = \sum_{x \in \Omega} \log p_{contour}(x; l_{contour}(x)) \]

\[ \mathcal{L}_{distance} = \sum_{x \in \Omega} (\hat{D}(x) - D(x))^2 \]
Dataset and Preprocessing

Optic cup and disc segmentation [5]
- ORIGA dataset for the task of optic disc and cup segmentation
- 650 color fundus image with train and test
- Image dimension: 256 x 256

Polyp segmentation [6]
- MICCAI 2018 Gastrointestinal Image Analysis (GIANA)
- 912 images with train and test
- Image dimension: 256 x 256

Preprocessing
- Contour map - Estimating boundary of connected components
- Distance map - Euclidean distance transform to the mask
Evaluation metrics

Segmentation evaluation

\[ J \text{accard}(A, B) = \frac{|A \cap B|}{|A \cup B|} \]

\[ \text{Dice}(A, B) = \frac{2|A \cap B|}{|A| + |B|} \]

Shape similarity

\[ H(A, B) = \max \left\{ \sup_{x \in A} \inf_{y \in B} \|x - y\|, \sup_{y \in B} \inf_{x \in A} \|x - y\| \right\} \]
Evaluation: Segmentation and shape similarity

<table>
<thead>
<tr>
<th>Architecture</th>
<th>Cup</th>
<th>Disc</th>
<th>Polyp</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dice</td>
<td>Jaccard</td>
<td>Hausdorff</td>
</tr>
<tr>
<td>U-Net</td>
<td>0.8655</td>
<td>0.7712</td>
<td>14.832</td>
</tr>
<tr>
<td>DCAN</td>
<td>0.8715</td>
<td>0.7803</td>
<td>14.775</td>
</tr>
<tr>
<td>DMTS</td>
<td>0.8723</td>
<td>0.7807</td>
<td>14.814</td>
</tr>
<tr>
<td>Psi-Net (Ours)</td>
<td><strong>0.8745</strong></td>
<td><strong>0.7848</strong></td>
<td><strong>14.541</strong></td>
</tr>
</tbody>
</table>

Note: The performance improvement is significant in Polyp followed by Cup and Disc. This corresponds to the difficulty of dataset.
Evaluation: Segmentation around boundaries [7]
Qualitative results

Image, GT, U-Net, DCAN, DMTS, Psi-Net(ours)
References


