Transparent Deep Learning for Error Detection in Medical Imaging Segmentation

Sean Mullan, Milan Sonka

Nov 5, 2020
Work by Others: Deep Learning for Lung Segmentation

- Pulmonary U-Net (2018)\(^1\)
  - Single Slice
  - 30M+ Parameters
  - 0.95 Dice Coefficient on LIDC Images\(^2\)

- U-Net R231 (2020)\(^3\)
  - Single Slice
  - 31M+ Parameters
  - 0.97 ± 0.05 Dice Coefficient
Dense Pathologies  
(Tumors, Large Nodules)

→ High test scores, low nodule sensitivity
Approach: Semi-Transparent Architecture

→ Spatial Attention Layers
→ 434,640 parameters (1.4% of U-Net)
Approach: Attention Blocks

Dual Attention Decoder Block

Spatial Attention Path

Channel-wise Attention Path

Sun et al. 2020
Data and Evaluation

→ 862 Pulmonary CTs with nodules >= 3mm
→ Trained to 0.99 ± 0.03 Dice

→ Generate predictions and select edge cases
  • 5 with the lowest lung segmentation Dice Coefficient
  • 5 with the most false negative nodule voxels

→ Correlate locations of attention/attribution variance with segmentation error
Results

Source Image

Predicted Label

Attention Maps

Attribution Maps

D1

D2

D3

D4
Attention Results

Source Image

Predicted Label

Extracted Spatial Attention Maps

D1

D2

Slight Attention

D3

Distortion

D4

Ring of Increased Attention
Attribution Results

Source Image

Predicted Label

Generated Attribution Maps

D1

High Attribution

D2

Outlier Pixel

D3

Transition Area

D4

Refined Transition Area
Future Work
Dealing with Extreme Cases

Attention Maps

Attribution Maps
Future Work
Dealing with Extreme Cases

→ Extreme outliers result in confident errors
  • No Attention/Attribution anomalies

→ Potential Solutions
  • Active Contour Loss
  • Boundary Shape Network
Conclusions

- Transparent architecture improves interpretability of Deep Learning networks

- Network attention and attribution variance can be used to detect regions of error

- Extreme outliers can cause models to be confidently incorrect
Thank You!
Citations


