Investigation of Weighted Binary Cross Entropy on Segmentation Performance of Pleural Mesothelioma on CT Scans

Samuel J. Fulton[1], Christopher L. Valdes, MSc[1], Feng Li, MD, PhD[1], Owen Mitchell[2], Christopher M. Straus, MD[1], Hedy L. Kindler, MD[2] and Samuel G. Armato III, PhD[1]

[1]Department of Radiology and [2]Department of Medicine, The University of Chicago, Chicago, IL

Abstract

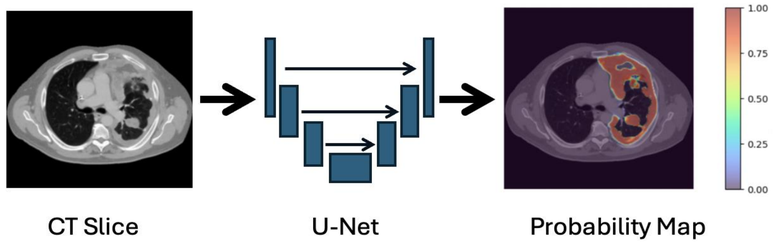
Pleural mesothelioma is an aggressive cancer of the pleural lining. Prior work has shown that no single probability threshold applied to CNN output maps simultaneously optimizes tumor-volume agreement and spatial-overlap metrics. In this study, we show that replacing binary cross-entropy (BCE) with weighted binary cross-entropy (WBCE) aligns the optimal thresholds for these two families of metrics. These findings indicate that WBCE can improve the agreement between complementary evaluation metrics.

Data Set

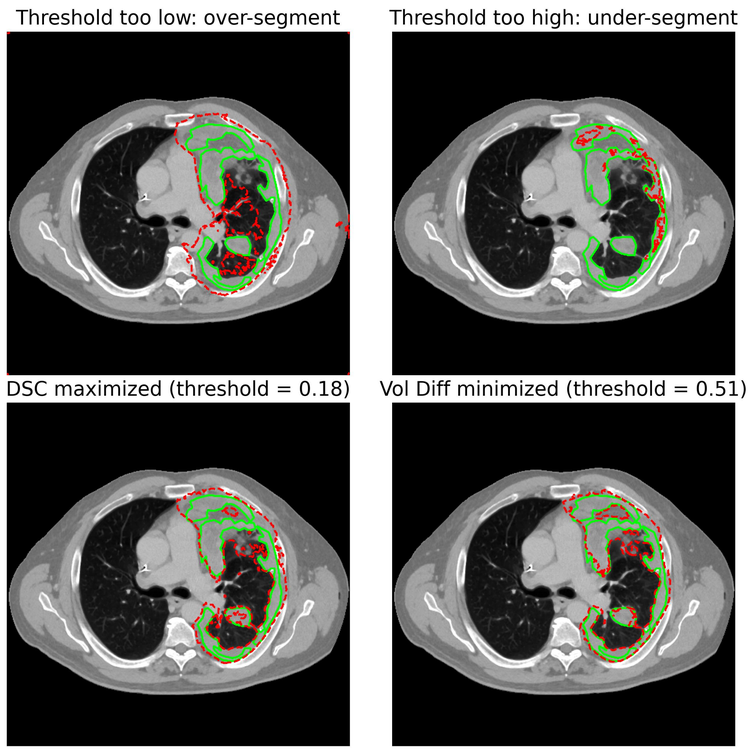
- CT scans from 126 patients with pleural mesothelioma
- Tumor regions manually contoured by a radiologist
- 96 patients for training and validation
- 30 patients held out as an independent test set

Methods

- U-Net–based convolutional neural network was used for automated tumor segmentation
- Encoder initialized with a pretrained VGG19 or ResNet50 backbone
- Network outputs a voxel-level probability map
- Probability threshold are varied to optimize segmentation metrics

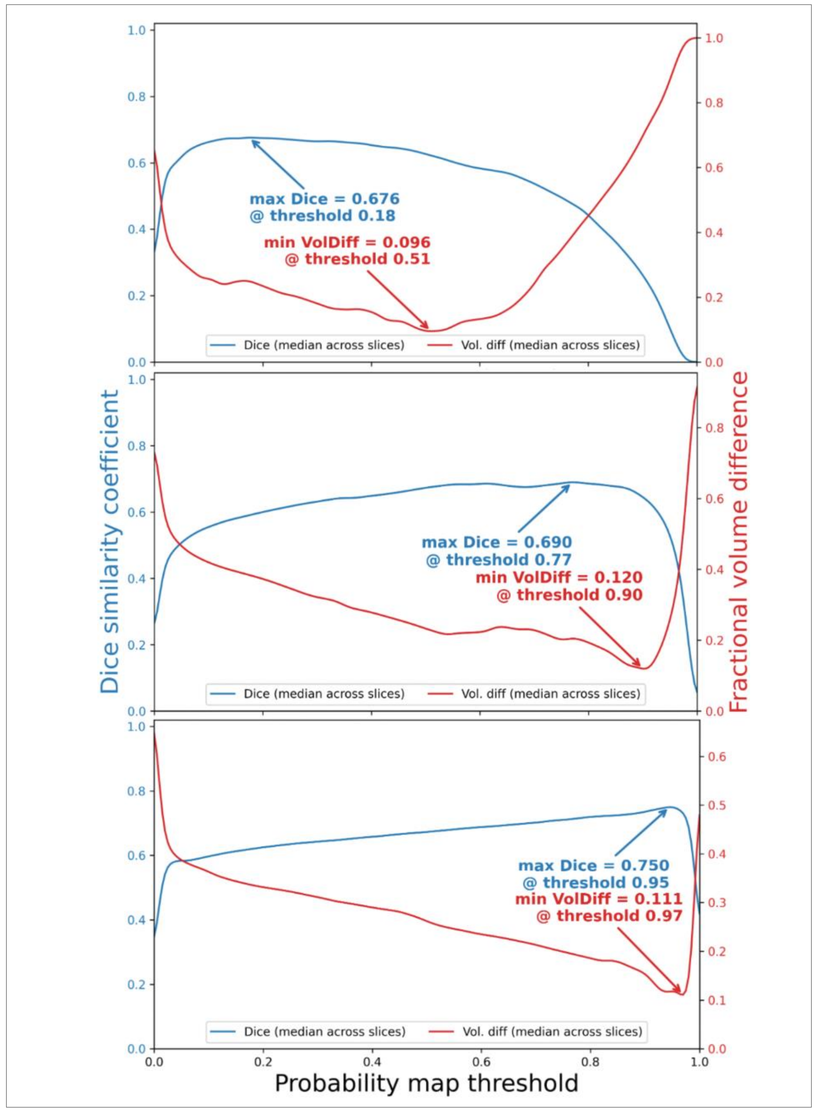


Methods



- Objective: identify a probability threshold that jointly maximizes spatial overlap (DSC) and minimizes fractional volume difference (FVD).
- **DSC =** $\frac{2|A \cap B|}{|A| + |B|}$, quantifies the spatial overlap between the reference standard A and the CNN segmentation B.
- **FVD =** $\frac{|Vol_{pred} - Vol_{ref}|}{Vol_{pred} + Vol_{ref}}$, measures the relative discrepancy between the predicted and reference tumor volumes.
- Models were trained under two loss functions:
  - $\mathbf{L_{BCE}} = -[\mathbf{y}\,\mathbf{log}(\mathbf{p}) + (\mathbf{1} - \mathbf{y})\,\mathbf{log}(\mathbf{1} - \mathbf{p})]$
  - $\mathbf{L_{WBCE}} - [\mathbf{w_1}\,\mathbf{y}\,\mathbf{log}(\mathbf{p}) + \mathbf{w_0}\,(\mathbf{1} - \mathbf{y})\,\mathbf{log}(\mathbf{1} - \mathbf{p})]$
  - where y is the reference label (0 or 1) and p is the predicted probability.
- Under BCE, the DSC-optimal and FVD-optimal thresholds do not coincide.
- Under WBCE, the two optimal thresholds converge to the same value.

Results



Conclusions

Training with weighted BCE aligns optimal thresholds, allowing a single threshold to achieve both accurate spatial overlap and volumetric agreement.

Acknowledgments

This study was funded by the Mesothelioma Applied Research Foundation. CT scans were provided and anonymized by the Human Imaging Research Office at the University of Chicago.

References

Shenouda et al., Journal of Imaging Informatics in Medicine, 2024.