

Cross Validation of a Conditional Generative Adversarial Network for the Classification of Pleural Mesothelioma Histological Subtype on CT Scans

Christopher L. Valdes, MSc¹, Brookelyn Vermont¹, Feng Li, MD, PhD¹, Owen Mitchell², Christopher M. Straus, MD¹, Hedy L. Kindler, MD², Samuel G. Armato III, PhD¹

1. Department of Radiology, The University of Chicago 2. Department of Medicine, The University of Chicago

Abstract

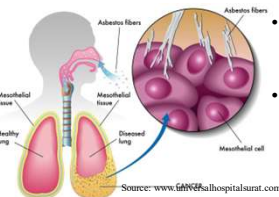
Purpose: This study aimed to cross validate a conditional generative adversarial network (cGAN) model for noninvasive histological subtype classification of pleural mesothelioma on computed tomography (CT) scans.

Methods: CT scans from 70 mesothelioma patients with epithelioid histology and 7 mesothelioma patients with sarcomatoid histology were retrospectively collected. A single axial series was selected from each scan to ensure consistency in reconstruction filter and slice thickness, and a radiologist indicated a single section that contained the largest cross section of disease. This section and its two inferior sections were designated as the "Primary Input," while the three superior sections were designated as the "Secondary Input." The 70 epithelioid scans were partitioned into 10 folds of 7 scans each. A cGAN, consisting of a generator and discriminator, was trained on 9 folds with augmentations consisting of 5° counterclockwise and clockwise rotations and mirror transformations. The generator was trained to generate the Secondary Input based on the Primary Input, yielding a "Synthetic Output." The trained generator was tested on the remaining fold and the sarcomatoid cases by generating Synthetic Outputs. L2 loss between the Synthetic Output and the Secondary Input of the test cases was used as the classification metric. Folds were iteratively swapped until each was tested by its respective generator. Model performance was quantified via receiver operating characteristic (ROC) analysis, and the average area under the curve (AUC) and 95% confidence interval (95% CI) were calculated across folds.

Results: The model achieved an average AUC value of 0.73 (95% CI: [0.68, 0.79]), with a maximum fold AUC value of 0.88.

Conclusion: The model showed potential for histological subtype classification of pleural mesothelioma on CT scans. Further studies will expand the data, apply more data augmentations, and evaluate classification performance between epithelioid and biphasic cases.

Background



- Pleural mesothelioma – rare and aggressive cancer that affects the pleural lining of the lungs.
- Histological subtype – classified as one of three categories:
 - Epithelioid
 - Sarcomatoid
 - Biphasic

- Biopsies only capture histological subtype at biopsy sites, while a non-invasive technique for subtype classification from CT scans would provide more global information.

Hypothesis/Objectives

- A conditional generative adversarial network (cGAN) can be trained, tested, and validated to reconstruct synthetic CT sections of pleural mesothelioma patients based on one specific subtype.
- Comparison of these synthetic sections with the corresponding real sections will allow for discrimination between sections that contain different histological subtypes.

Methods

- The CT scans of 70 pleural mesothelioma patients with epithelioid histology and 7 patients with sarcomatoid histology were retrospectively collected.
- A single CT series was selected to ensure consistency in slice thickness and reconstruction filter.
- Training data were augmented with mirror transformations and 5° clockwise and counterclockwise rotations.
- Epithelioid cases were sorted into 10 folds of 7 patients each for testing.

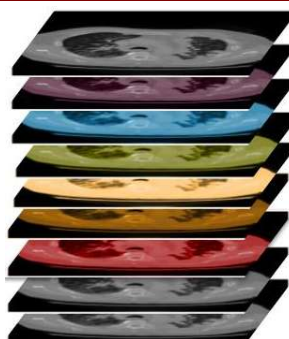


Figure 1. A single slice containing the largest tumor cross-section was selected by a radiologist, highlighted in yellow. Red, orange, and yellow slices were designated as the Primary Set, while the green, blue, and purple were designated as the Secondary Set.

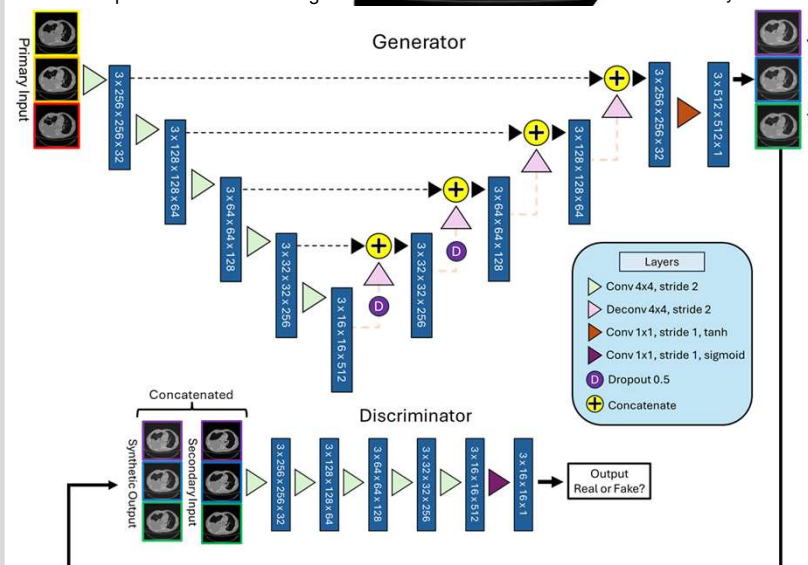
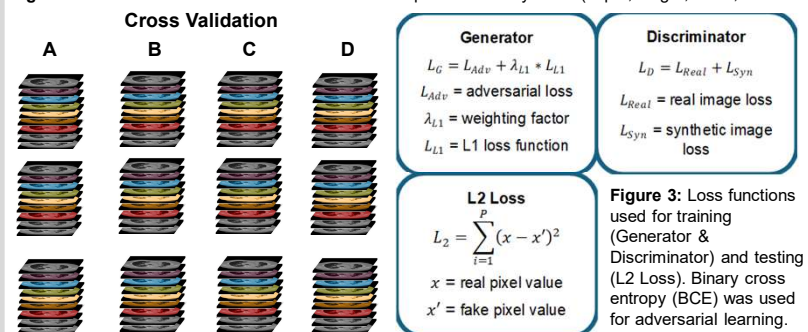


Figure 2. Network architecture of the cGAN. The shapes of each layer are (depth, height, width, channels).



Results

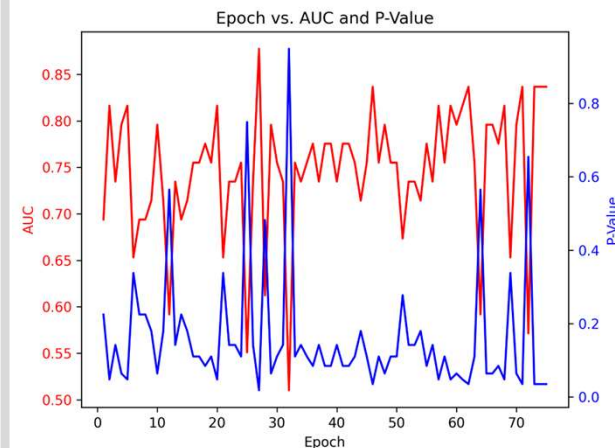


Figure 4: Sample plot of the AUC as a function of epoch for Fold 2. The maximum AUC for each curve was recorded for each corresponding fold. The p-value is based on L2 loss from the Wilcoxon Rank Sum Test.

Average AUC across Folds: 0.73
95% Confidence Interval: [0.68, 0.79]
Maximum AUC: 0.88

Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
0.76	0.88	0.63	0.69	0.67
Fold 6	Fold 7	Fold 8	Fold 9	Fold 10
0.71	0.69	0.78	0.86	0.67

Table 1: Fold-specific AUC values obtained from independently trained cGAN models during 10-fold cross validation. The AUC ranged from 0.63 to 0.88.

Future Directions

- Future work will expand the patient database and increase data augmentation to reduce overfitting.
- Preliminary results have shown promise for increasing AUC values.
- We will seek to incorporate a novel viscosity solution partial differential equation pipeline to smooth generator outputs and stabilize training.

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

- Han, Changhee, et al. "MADGAN: unsupervised Medical Anomaly Detection GAN using multiple adjacent brain MRI slice reconstruction." *BMC Bioinformatics*, 2021.