# MultiMix: Sparingly-Supervised, Extreme Multitask Learning From Medical Images

Ayaan Haque[1], Abdullah-Al-Zubaer Imran[2,3], Adam Wang[2], Demetri Terzopoulos[3,4]

[1]Saratoga High School, [2]Stanford University, [3]University of California, Los Angeles, [4]VoxelCloud, Inc.

IEEE - ISBI 2021

## Introduction

- **Problem:** Especially in the medical imaging domain, deep learning from limited labeled data lacks generalizability and explainability.
- **Goal:** Improve model generalizability in semi-supervised multi-task learning, while preserving explainability.
- **MultiMix:** A novel, semi-supervised, multi-task learning model for the joint classification and segmentation of medical images, leveraging consistency augmentation and a saliency bridge module.
- **Code:** https://github.com/ayaanzhaque/MultiMix

## Method

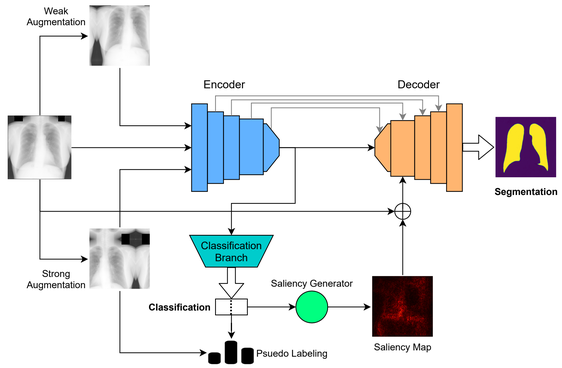


Figure 1: Schematic of the proposed *MultiMix* model.

- Classification: Using predictions on unlabeled weakly augmented images [2], pseudo-labels are generated with confidence, and loss is computed with strongly augmented versions of the images.
- Segmentation: Generated saliency maps from the encoder are concatenated via the saliency bridge module (Fig. 2) for improved segmentation.

- Classification loss:
  $L^c = L_l(\hat{c}_l, c_l) + \lambda L_u(\hat{c}_s, \arg\max(\hat{c}_w) \geq t)$
- Segmentation loss:
  $L^s = \alpha L_l(\hat{y}_l, y_l) + \beta L_u(\hat{y}_l, \hat{y}_u)$

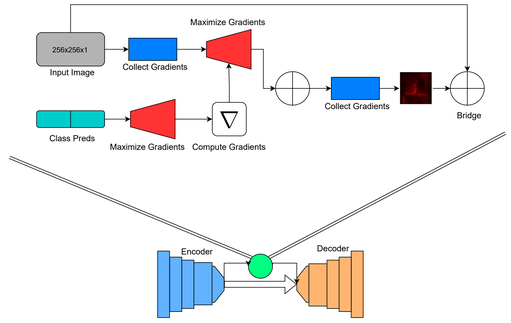


Figure 2: Proposed saliency bridge: Classification predictions are taken and saliency maps are produced through gradient maximization. They are then concatenated with the input images.

## Dataset

Table 1: Details of the datasets used for training and testing.

| Mode | Dataset | Total | Normal | Abnormal | Train | Val | Test |
|---|---|---|---|---|---|---|---|
| in-domain | JSRT | 247 | – | – | 111 | 13 | 123 |
| | CheX | 5,856 | 1583 | 4273 | 5216 | 16 | 624 |
| cross-domain | MCU | 138 | – | – | 93 | 10 | 35 |
| | NIHX | 4185 | 2754 | 1431 | – | – | 4185 |

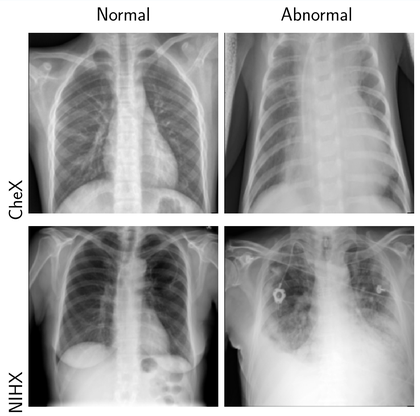


Figure 3: Example images from the CheX and NIHX datasets.

## Visualizations

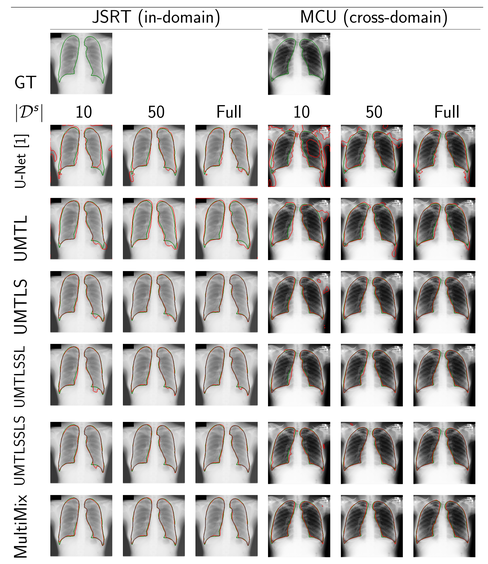


Figure 4: Boundary visualization of the predicted segmentations in a chest X-ray affirms the superiority of MultiMix. Color code: green (reference), red (predicted).

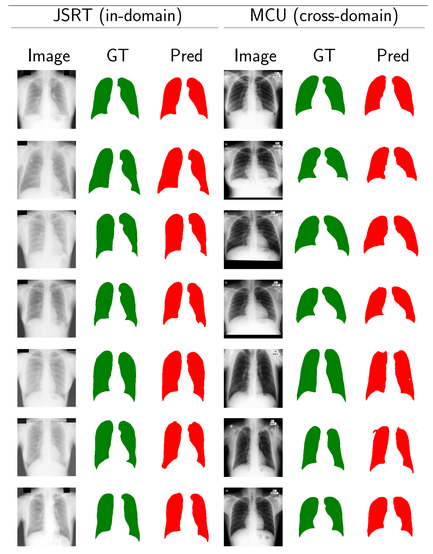


Figure 5: Visualizations of the segmented lung masks output by MultiMix-50-1000 on the in-domain JSRT dataset and cross-domain MCU dataset. The results show good agreement between the groundtruth and predicted masks.

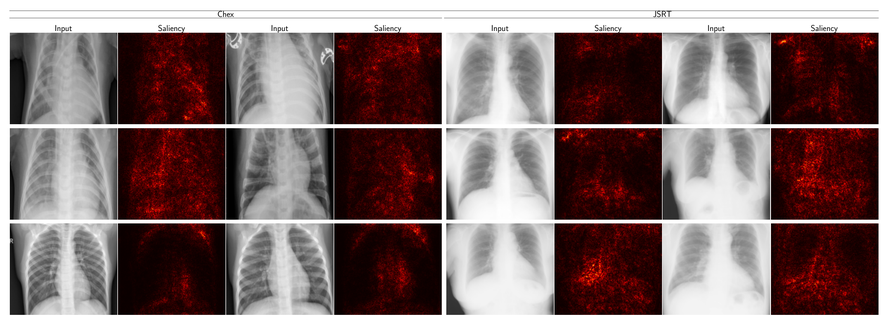


Figure 6: Saliency maps (from MultiMix-50-1000) based on the classification gradients of images on both the classification (left) and segmentation (right) datasets show highlighted lungs and visualizations. These maps hold information from the encoder that is important for segmentation.

## Graphs and Metrics

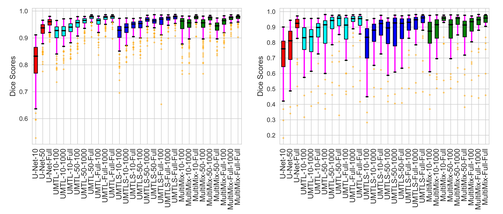


Figure 7: Distributions of the Dice scores demonstrate the segmentation superiority of the MultiMix model over the baseline models.

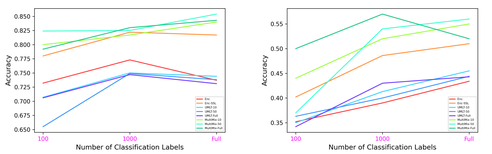


Figure 8: Classification accuracies of different supervised and semi-supervised baselines at different training datasizes shows the improved consistency and accuracy of MultiMix.

## References

[1] O. Ronneberger, P. Fischer, and T. Brox.
U-net: Convolutional networks for biomedical image segmentation.
In *MICCAI*, pages 234–241, 2015.

[2] K. Sohn, D. Berthelot, et al.
Fixmatch: Simplifying semi-supervised learning with consistency and confidence.
In *Advances in Neural Information Processing Systems*, volume 33, pages 596–608, 2020.