# MIDRC XAI Challenge

## Submission #3 description – codename bestlr1_256

Siqi Sun, Team_Sunday

**Data:**

- Selection: Chest radiographs from the data.midrc.org website were exclusively used, and downloaded using the Gen3 agent as described in the COVIDx Challenge GitHub repo. The selection process involved:
  1) Ensuring a one-to-one correspondence between the CSV data and the images, and outputting the corresponding count.
  2) Based on this confirmation, checking the number of images that correspond to each score in the CSV. The output indicated cases where one image corresponded to one score, two images corresponded to one score, and three images corresponded to one score.
  3) A total of 1,625 chest radiographs were confirmed to have a one-to-one correspondence with their scores, forming the curated dataset of 1,625 images.
- Pre-processing:
  1) Lung segmentation was performed using a UNET + variational encoder model and weights from R. Selvan et al[1], the latter being free to use under the MIT licence.

**Model:**

This model is a multi-task neural network that combines EfficientNet with a Squeeze-and-Excitation (SE) attention mechanism, capable of performing both classification and segmentation tasks.

1) Base Model: It is built on the EfficientNet-B0 architecture. The first convolutional layer has been modified to accept 2-channel inputs (image + mask). The original classification head is removed to accommodate the new tasks.
2) SE Module: A SEBlock is added after the final convolutional layer of EfficientNet to introduce channel-wise attention. This block uses global average pooling and two fully connected layers to adjust the weights of each channel, enhancing feature representation.
3) Classification Head: Adaptive average pooling is applied to the feature map to reduce it to 1x1, followed by flattening and a fully connected layer for binary classification. A 50% Dropout is applied to prevent overfitting.
4) Segmentation Head: A 1x1 convolutional layer is used for pixel-level segmentation, generating an interpretability map that captures local features of the input.

The network outputs both a classification result and a pixel-wise segmentation map.

**Training:**

- Input: DICOM images and corresponding masks, resized to 512x512 pixels.
- Data Augmentation:
  1) RandomHorizontalFlip: Randomly flips images horizontally.

---

1. R. Selvan et al., *Lung Segmentation from Chest X-rays using Variational Data Imputation*, ICML Workshop on The Art of Learning with Missing Values, July 2020, https://arxiv.org/pdf/2005.10052.pdf

2) RandomRotation: Rotates images within a range of -30 to +30 degrees.
3) ColorJitter: Adjusts image brightness, contrast, saturation, and hue.
4) RandomResizedCrop: Random cropping and resizing with a scale range of 0.8 to 1.0.
5) ToTensor: Converts images to PyTorch tensors for model input.

- Training Configuration:
  1) Optimizer: Adam optimizer.
  2) Learning Rate: 0.001.
  3) Epochs: 15.
  4) Batch Size: 16.
- Loss Functions:
  1) Classification Loss: Binary Cross Entropy with Logits Loss (BCEWithLogitsLoss) for image-level classification.
  2) Segmentation Loss: BCEWithLogitsLoss for pixel-level segmentation, comparing model output to lung masks.
  3) Combined Loss: 70% Classification Loss + 30% Segmentation Loss.

     Total Loss: Total Loss = 0.7 * Classification Loss + 0.3 * Segmentation Loss.
- Output:
  1) Classification Results: A probability score is generated for each image.
  2) Segmentation Output: An explainability map is upsampled to the original image size using bilinear interpolation.

**<u>Inference</u>**

- Mask Generation: It loads DICOM images and generate lung segmentation masks.
- Model Inference: It loads the previously trained model and processes the generated masks and images for classification and segmentation. The classification results are saved to a CSV file, and the segmentation results are resized to the original image dimensions and saved as PNG images.