Mathieu Goulet

Centre régional intégré de cancérologie de Lévis (CISSS-CA), Team JMAI

Explainability pixel mapping network design

- See Figure 1. Not unlike Grad-CAM, the architecture aims to take the Classification network output/activation at different point between its convolution blocks, and to combine them back in a 2D map representative of the region that resulted in the actual classification result. Weights of the encoding part of the network are the same as the classifier. The decoder of the network is inspired by the UNET architecture.

Multi-stage design used for inference

- See Figure 2. Thresholds represent around 99% true negative rate for the classifier and around 50% opacity extent for the regressors (as estimated on the local validation dataset).
- As the loss function on the explainability maps score is VERY penalizing toward false negative pixels, multiple overestimations were performed on the resulting opacity maps to ensure minimal loss function penalties from missed (false negative) pixels.
- No opacity pixel permitted outside of the lungs segmentation.
- To avoid log loss maximas (log(0)), no pixels value was set under 10 bits (1 bits outside the lungs), nor over 245 bits.

Datasets used for training

- RSNA pneumonia detection challenge, stage 2 (https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data).

Image pre-processing steps

- Correct images inverted photometric interpretation.
- Resize Images to 512x512 (whole image) or 512x320 (split lung images).
- Rescale pixels values between -1 and 1.

Adjustments from previous submission (Test003)

- Retraining on the whole lung mapping weights using more data augmentation
- Overestimation size of opacity maps is now function of the mean regression score (two levels : high and low)

Training and classification/regression models details

Model	**Main binary classifier**	**Left and right lung extent regressor**	**Opacity mapping**	**Lung segmentation**
Base weights	First-place models of the mRALE mastermind challenge. Head replaced with either a binary classifier or single output regression.		The Main binary classifier, the Right lung regressor and the Left lung regressor were used for the whole image mapping, the right lung mapping and the left lung mapping, respectively	None
Ground truth	RSNA dataset labels	Surface of lung inside the RSNA dataset bounding boxes divided by the lung surface	Pixels inside the RSNA bounding boxes which were also inside the lungs	Manually drawn from a sample of 1000 opacity CXRs image from the RSNA dataset
"Frozen" weights during training	None		Classifier or regressor base weights	None
Data augmentation	±15deg CXRs rotation			
Backbones	EfficientNet-B4 and ConvNeXt-tiny			EfficientNet-B4
Optimizer	AdamW			
Loss function	Binary Cross-Entropy	Average of MSE and MAE	DICE	
Initial learning rate	1e-5	5e-5	1e-4	1e-4
Cosine annealing	Yes	Yes	Yes	No
Weight decay per epoch	50%	75%	75%	None
Epochs	12 to 20	20 to 25	20 to 30	100
Batch size	32	50	32 or 50	32

Figure 1: classifier (or regressor) and explainability pixel mapping network. The EfficientNetB4 backbone is represented, but similar architecture was devised for the ConvNeXt-tiny backbone.

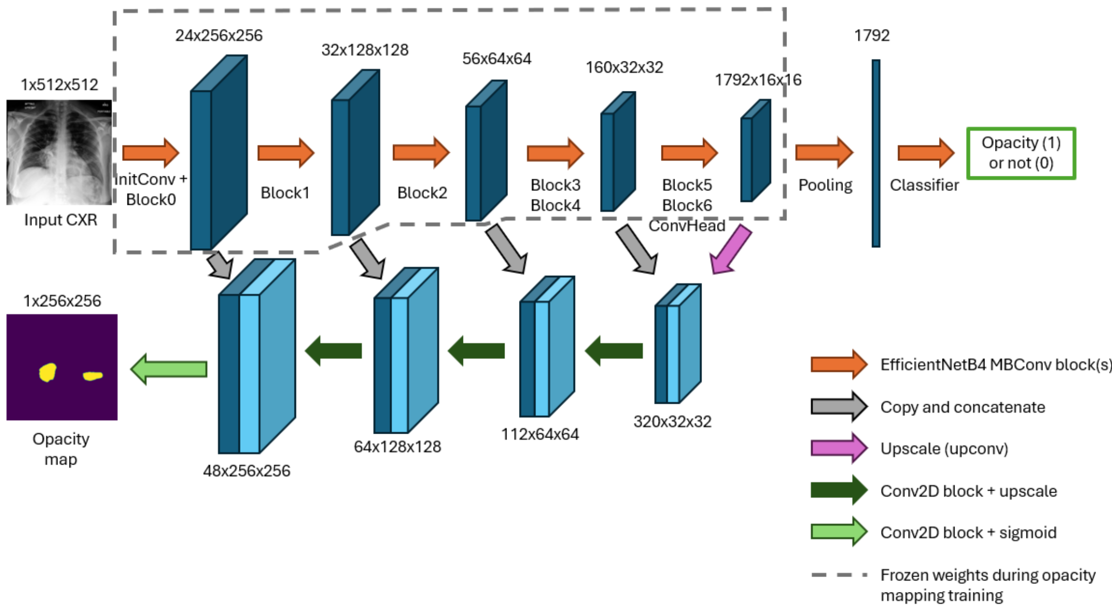

Figure 2: multi-stage decision process used for inference. Green boxes refer to neural networks.