Experiment is carried out according to the approach in the Nature paper. Trained on three different hosts within the same network domain

Dataset we are using open-source dataset: Duke-Breast-Cancer-MRI
We split the dataset into 4 parts. 40% data is on host1, 30% data is on host2, 10% data is on host3, remaining 20% data is separated as test dataset and stored on all hosts for prediction evaluation.

We are using the sub sequence which is and first post-contrast - pre-contrast sequence
post1 - pre

Swarm-learning params for experiments:

Model: Att-mil

    sync interval: 32
    training batch size: 32
    host1
        num of patients:368 276 92
        epoch:50
        nodeWeightage:100
    host2
        num of patients:276
        epoch:67
        nodeWeightage:75
    host2
        num of patients:92
        epoch:200
        nodeWeightage:25
    test
        num of patients:186

network structure and hparams:

    imagenet feature extraction + Att mil learning
    size of params: about 1mb
Optional feature extraction: imagenet, vit, swint, radimagenet
When performing the cross validation locally the feature extracted by imagenet performs the best


Model: 3D-CNN

sync interval: 512
    training batch size: 1
    host1
        num of patients:368

epoch:100

nodeWeightage:100

host2

num of patients:276

epoch:167

nodeWeightage:75

host2

num of patients:92

epoch:400

nodeWeightage:25

test

num of patients:186

network structure and hparams:

size of params: about 700mb
acc_kwargs:
  task: binary
aucroc_kwargs:
  task: binary
block: basic
block_inplanes:
- 64
- 128
- 256
- 512
feed_forward: true
in_ch: 1
layers:
- 3
- 4
- 6
- 3
loss: !!python/name:torch.nn.modules.loss.BCEWithLogitsLoss "
loss_kwargs: {}
lr_scheduler: null
lr_scheduler_kwargs: {}
optimizer: !!python/name:torch.optim.adamw.AdamW "
optimizer_kwargs:
  lr: 0.0001
out_ch: 1
spatial_dims: 3

Training results:

Malignant or Benign tumor prediction performance on breast MRI
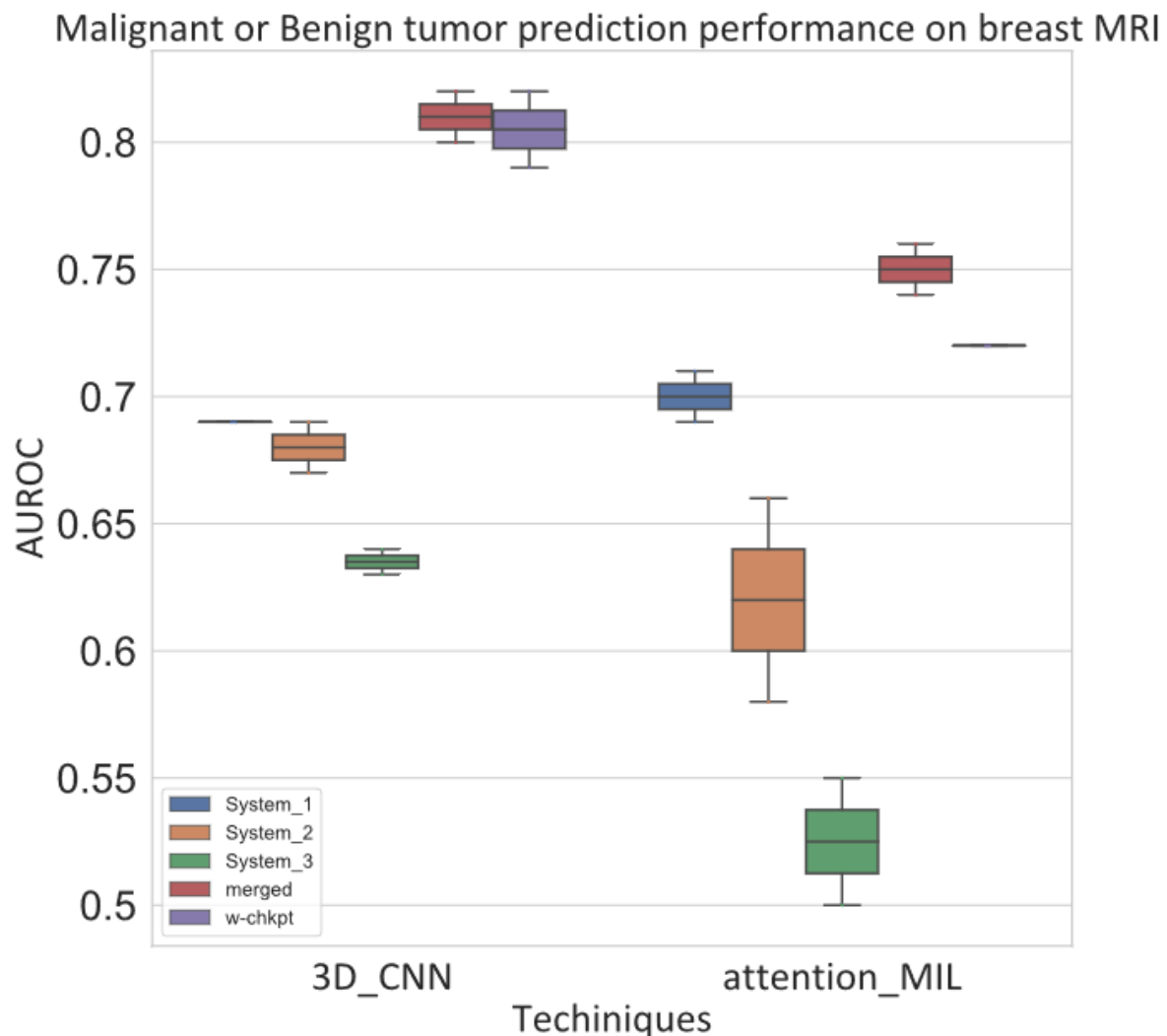
When observing the training/validation curve and results we could observe that Att mil is more data hungry and easy to overfit on small datasets. While it takes significantly longer for 3dcnn to converge and finish the training. The overall performance of 3dcnn is better than att mil. Reasons could be:
1, 700 times larger model params
2, The attention based mechanism could only give a score to the slices of mri data. While the tumor only overtakes a small portion of small slices. So the attention and the ground truth is like on two different dimensions(although we are not using annotation boxes with in sequences for this exp). The mil approach is optimal for histoimage but might not for mri data.

Future work:
1, We can use marugoto for swarm learning for histoimages. We are achieving a significant boost in the performance when training marugoto in swarm learning other than locally on small datasets.

2, Pretrain the feature extraction model with radi images.
3, It's really unstable for the communication between blockchain-based swarm network nodes when carrying out the experiments on att-mil. Should investigate the reasons.
4, Sync frequency plays a predominant role in improving the training results. Because the overfitting could easily happen when hosts with small datasets are looping in the local data training.
5, Swarm learning should be easy to adapt as long as the ml is under major frameworks like tensorflow, pytorch, pytorch lightening

Supplementary info:

When a node first reaches training end. Specifies the merge behavior of a SL node after it has achieved stopping criterion and it is waiting for all other peers to complete their training. During this period this SL node does not train the model with local data. This parameter decides the nature of the weights that this SL node contributes to the merge process. We are using:

`snapshot`: Node always contributes the weights that it had when it reached the stopping criterion, it does not accept merged weights.

`nodeWeightage`	A number between 1–100 to indicate the relative importance of this node compared with others during the parameter merge process.
	By default, all nodes are equal and have the same weight-age of one.