

Submission includes a report, run.py, score.py, 'TrainGAN' ipynb file for code inspection and results visualization. Generated models, losses and scores during training are saved under the 'logs' folder.

### 1. Problem definition (X as training data and Z as noise data)

- Train a GAN from lung cells images to fulfill these 3 criteria:
  - S1: D gives high score for original data
  - S2: G(z) images are close to the boundary of X
  - S3: Generated data needs to be similar to one of X

### 2. Method

- Train a GAN as usual with binary cross entropy losses on both discriminator and generator
- S1 and S3 are accounted from the discriminator and generator's default BCE loss function
- (Define S2 as boundary loss) For S2, change discriminator loss function from noise data Z as follow:  

$$D_{loss} = -[y_1 * \log(D(x)) + \gamma * (1 - y_2) * \log(1 - 4(D(G(x)) - 0.5)^2)]$$
, where  $y_1 = 1$  and  $y_2 = 0$
- Add a hyperparameter  $0 < \gamma < 1$  to the 2<sup>nd</sup> term of D's loss function for tuning purposes
- Train the GAN with high epoch, terminate when s1, s2 and s3 generate high values

### 3. Results

- The results of original GAN vs modified GAN is shown in figure 1 where s2 (red line) improves drastically while s1 and s3 remain the same in the first few epochs.

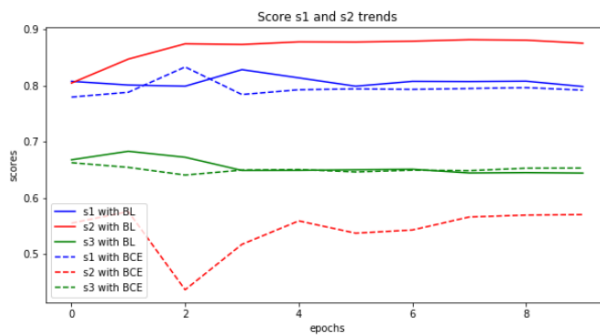


Figure 1 Results comparison between original GAN (dotted lines) vs modified GAN (solid lines)

BL: Boundary loss (S2 criteria)

BCE: Binary cross entropy loss

- Optimal value for  $\gamma$  is 0.8 after tuning
- After training with higher epochs, G and D losses reduced drastically and oscillates between 0 to 1
- After 60 epochs, s1 and s2 reached its optimal value of 0.9983 and 0.999 respectively (score history for s3 wasn't tracked due to longer computational time for s3)
- Best results obtained from given data:

```
{'s1': 0.9982962608337402, 's2': 0.9999310374259949, 's3': 0.6801146268844604, 'score': 0.6789090664113528}
```

