The Spine Project

Prateek Garg
Department of Computer Science and Engineering
Indian Institute of Technology Delhi
prateekgarg.iitd@gmail.com

Lakshya Singhal
Department of Mechanical Engineering
Indian Institute of Technology Delhi
lakshyasinghaliitd@gmail.com

Abstract

Patients suffering from Spinal Cord injuries who have suffered injuries due to some kind of trauma often need urgent attention from a medical expert. They must be congregated early in the spinal units to receive better facilities and adequate attention, ultimately improving the outcome of the treatment and rehabilitation. Our literature review indicates that early surgery and comprehensive rehabilitation markedly reduces the overall morbidity of such patients by enabling them to lead an independent life. Therefore, to advance our studies in enabling accurate and quick identification of a person's state of damage of the vertebral column, we are presenting a classifier for this dataset. Also, we present segment the raw images of the vertebrae for better visualization.

1 Description of Dataset

The dataset consists of radiological images, X rays, from over 1000 thousand patients who visited the Indian Spinal Injuries Centre seeking for their treatment. The X-ray images of the patients focus or highlight the thoracic and lumbar regions of the vertebral column. For each patient in the dataset, there are two X-Ray images highlighting two different views namely AP and Lateral, that help in identifying the underlying abnormality vividly. This gives us a total of 2000 X-ray images, equally divided among the two views.

1.1 Training Data

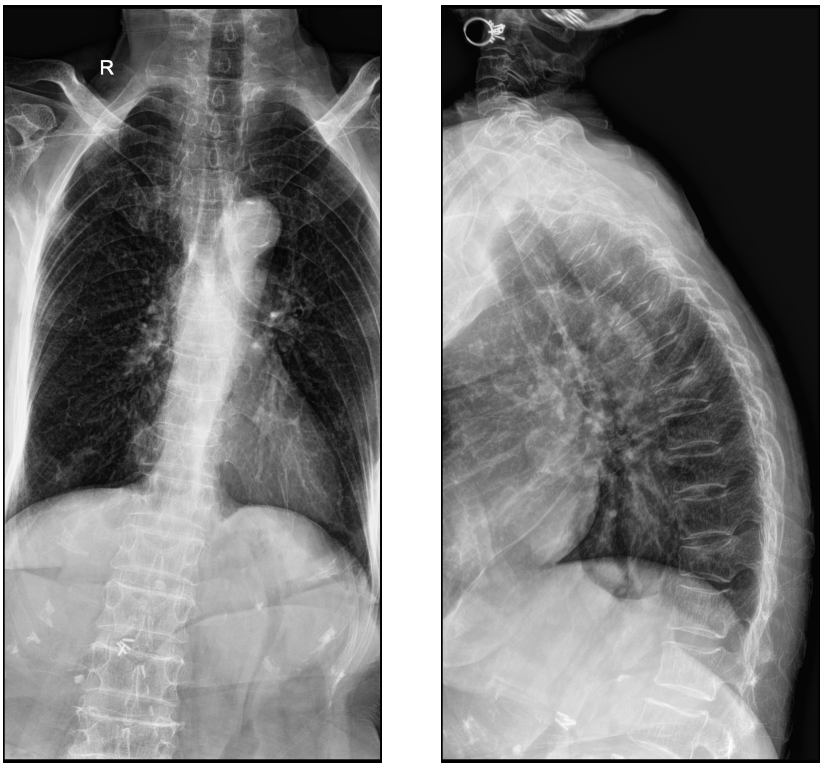
The training data had 2 images, the two views namely AP view and Lateral View, (shown in figure 1). Also, the segmentation masks for these images were also included. Some comments about the dataset:

- All the images were of different sizes.
- There were some images and masks missing.
- Two of the masks had very fine line, because of which the segmentation model had difficulty learning it. (In some cases the line was even 1 pixel wide.)

1.2 Preprocessing

- As the image sizes were different, we downsampled all of them to 224x224.
- We can notice that for all the LAT and AP images the respective RGB values were all same, that is, we need not take all three filters. Hence we separated out a single channel from both images.

Submitted to 33rd Conference on Neural Information Processing Systems (NeurIPS 2019). Do not distribute.



(a) AP View (b) Lateral View

Figure 1: Input Data

- We appended these two channels obtained. Further, as we wished to try some Transfer Learning on these for the classification task and most of the pretrained models were for RGB images, that is 3 channels, so we appended a zero channel layer. This made the final input to the neural network as a 224x224x3 image.

1.3 t-SNE and PCA visualization

We did the Principal Component Analysis on the concatenated single filter from both images (dim=224x224x2), and reduced their dimension to 3. The 3D and 2D views for the same are presented in Figure(2) and Figure(3).
We tried t-SNE visualization to also on the the same data as mentioned above. The output dimension in t-SNE was kept as 2. The visualization graph for the same is in Figure(4)

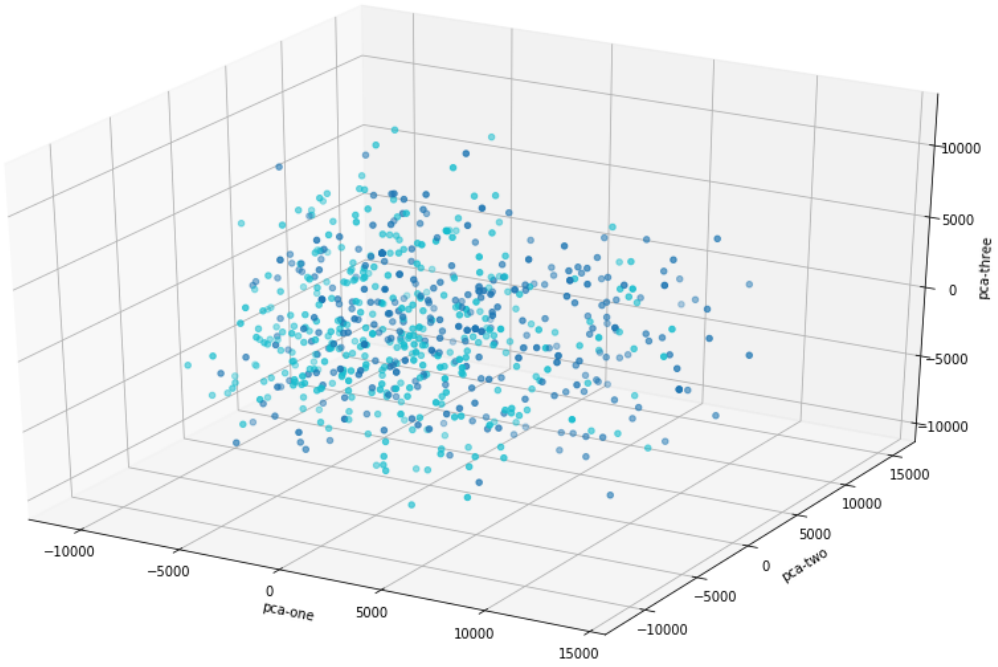


Figure 2: 3D PCA Visualization

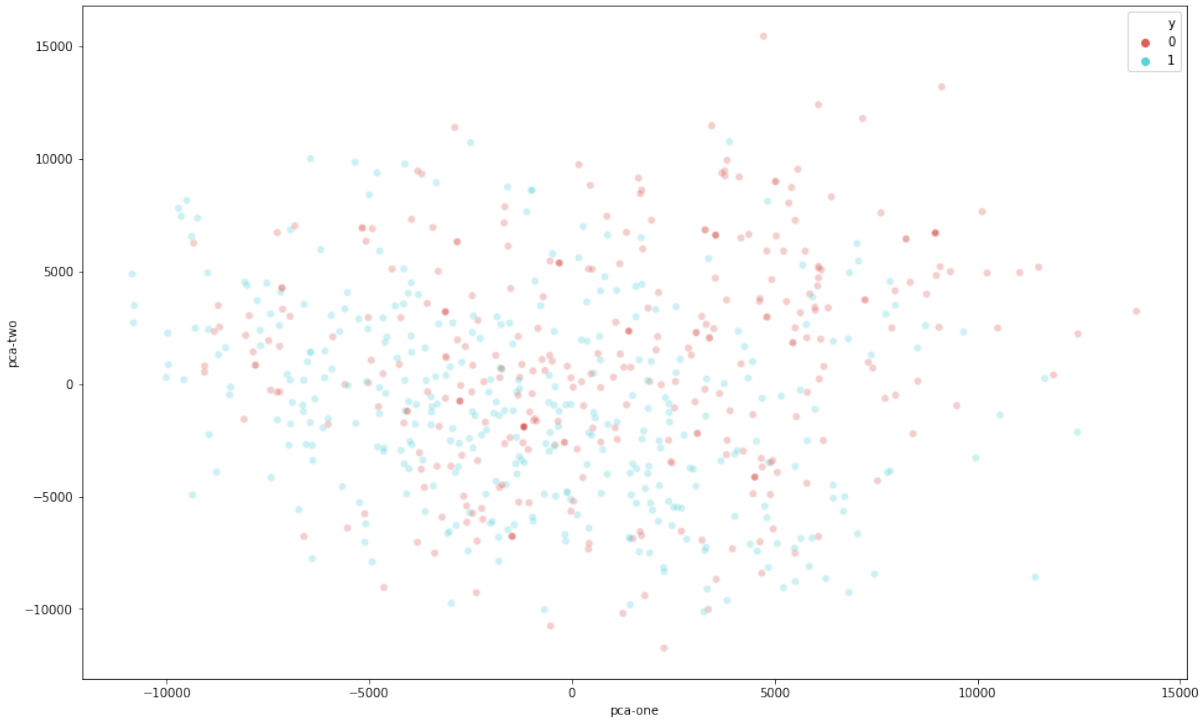


Figure 3: 2D PCA Visualization

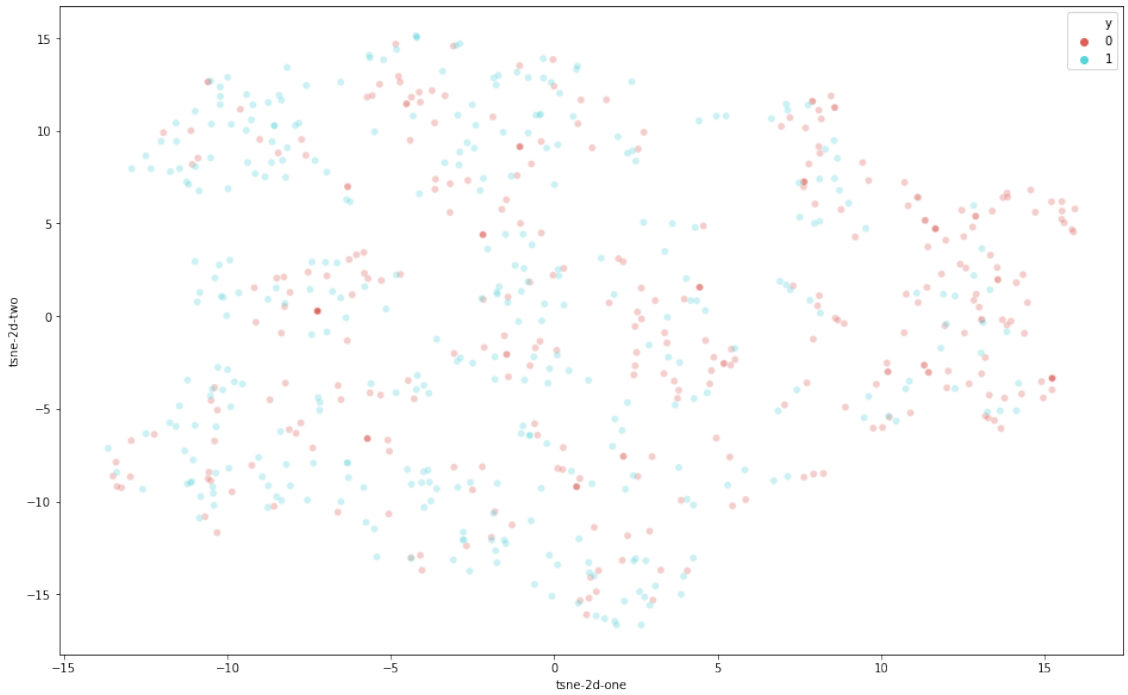


Figure 4: t-SNE Visualization

2 Classification Task

For classification task, we had to identify patients with damaged or normal spinal cord according to the characteristics of their AP and LAT images. In this case we take the original RGB image for both LAT and AP views and down sample them to 224x224x1. Then augment AP and LAT images together and input it to the neural network. The output of the network will be a label (Damaged or Normal).

2.1 Architecture

We have built our own architecture. Our implementation has 3 Convolutional layers, 3 Max Pooling Layers and fully connected layers. The architecture as mentioned in the paper is shown in Figure(5)

2.2 Analysis

2.2.1 Hyperparameters:

We used Batch Normalization, Dropout, Early Stopping in our model. These contributed to the hyper parameters. Further, ReduceLROnPlateau, num_epochs, batch_size are other hyper parameters. Primarily, we chose the hyper parameters, on hit-and-trial basis. We wanted to run a grid-search but the time constraint did not allow it.

2.2.2 Learning Curves:

The learning curves is shown below. The metric used while making these curves was "accuracy" and the loss function was "binary_crossentropy".

2.2.3 Data Augmentation:

We used data augmentation to expand the training dataset in order to improve the performance and ability of the model to generalize. We configured image data by Feature Standardization, Random Rotations, Random Shifts and Horizontal Flips.

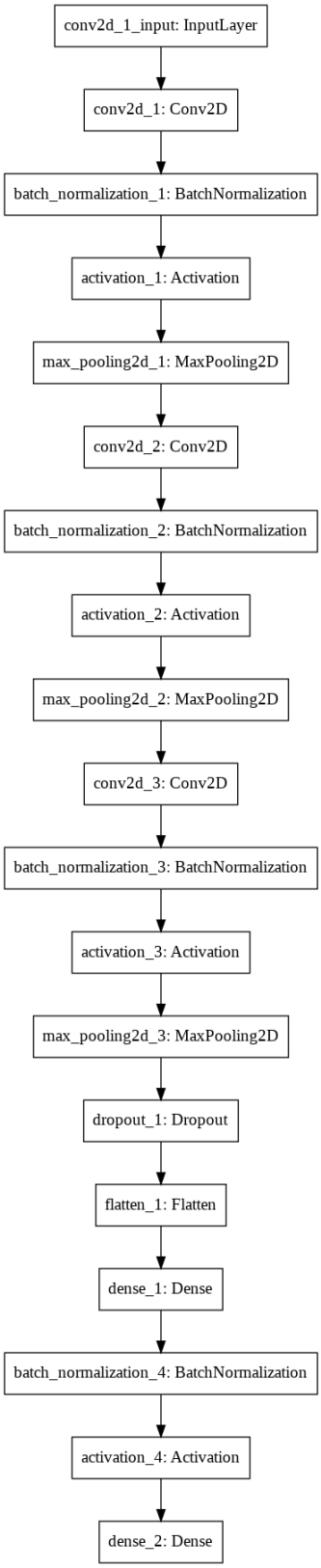


Figure 5: CNN Architecture

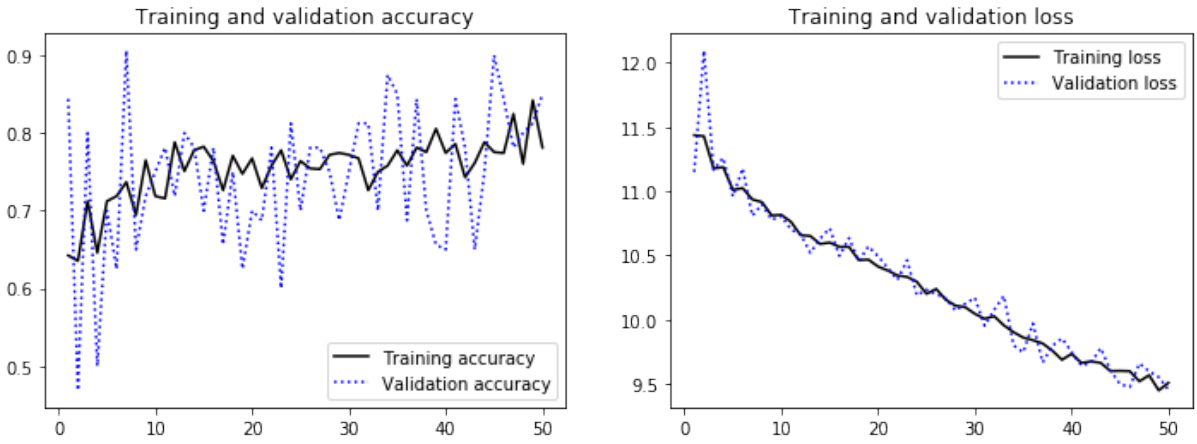


Figure 6: Learning Architecture

2.2.4 Validation Loss:

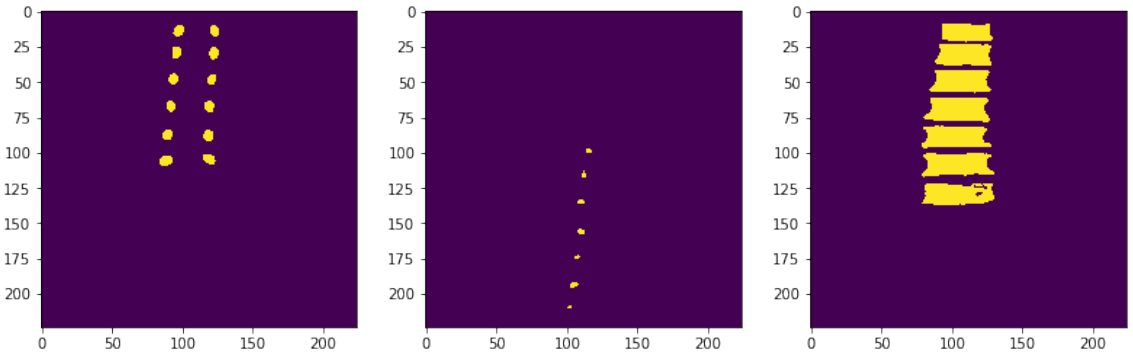
We adjusted the validation split in data to 0.1 ratio. Accuracy was used as metric and the validation acc is 0.85.

2.2.5 Observations:

We tried various deep learning models on our dataset. Some noticeable models are Inception-ResNet-v2 and Inception-v3. We trained Inception-ResNet-v2 on our dataset with pre-trained weights of imagenet, we observed that our model got overfitted though the training accuracy was 1 but validation accuracy was highly varying showing the overfitting nature of the model. After applying data augmentation on our training data, we observed a significant improvement towards generalization in our model but still we were not able to get good enough validation accuracy.

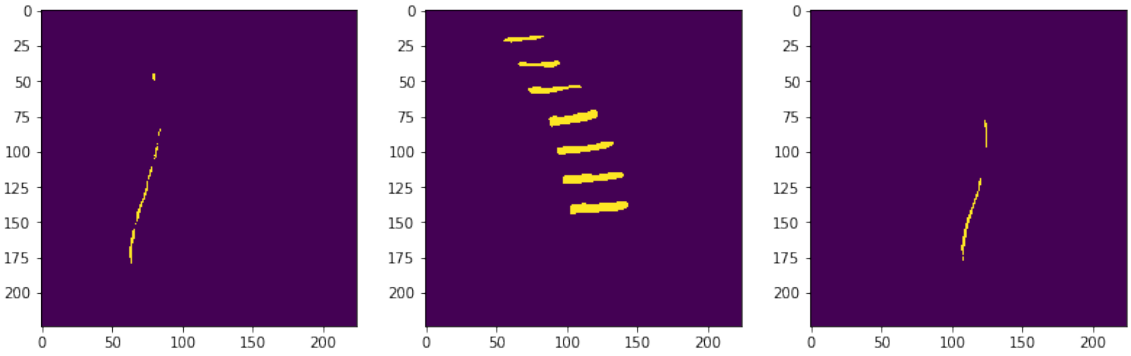
3 Segmentation Task

For the segmentation task, we had to create 3 masks for the AP image and 5 masks for the LAT image. In this case we take the original RGB image for the two cases separately, down sample it to 224x224x3 and then input it to the neural network. The output of the network will be 224x224 masks (single-channel). Hence we trained eight different models. Some test masks are shown below:

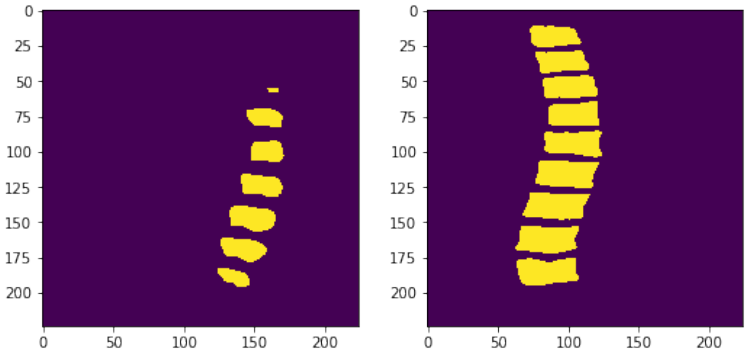


(a) AP Pedicle (b) AP Spinous Process (c) AP Vertrbra

Figure 7: AP Masks



(a) LAT Anterior Vertebral Line (b) LAT Disk Height (c) LAT Posterior Vertebral Line



(d) LAT Spinous Process; Line (e) LAT Vertebra; Line

Figure 8: LAT Masks

3.1 Architecture

We use the UNET Architecture. The architecture is one of the state of the art architectures for bio-medical image segmentation. Our U-Net implementation has 4 downblocks, 1 bottleneck layer and 4 upblocks. The architecture as mentioned in the paper is shown in Figure(8)

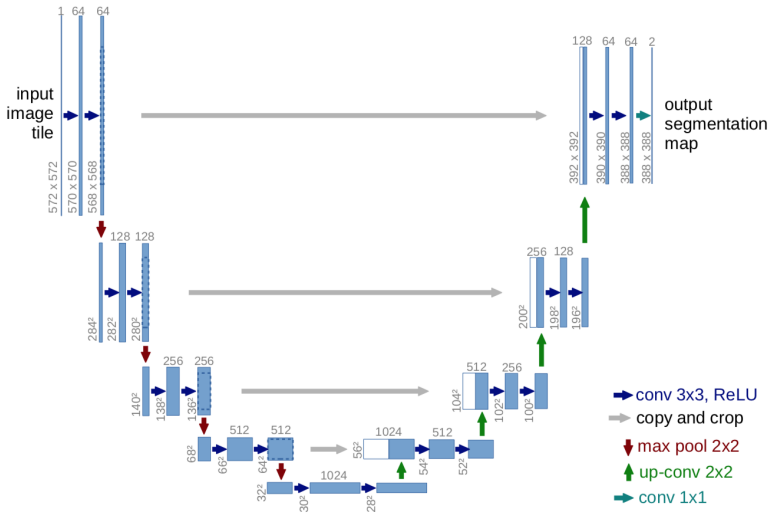


Figure 9: UNET Architecture

Table 1: Validation Data

Name	Validation Loss	Validation DICE Coeff.
AP Pedicle	0.0204	0.6350
AP Spinous Process	0.0072	0.2673
AP Vertrbra	0.0636	0.8294
LAT Anteror Vertebral Line	0.0049	0.0370
LAT Disk Height	0.0337	0.6670
LAT Posterior Vertebral Line	0.0050	0.0441
LAT Spinous Process Line	0.0505	0.6900
LAT Vertebral Line	0.0489	0.8472

Table 2: Accuracy on Test Data

Name	Avg. Dice Coeff.
AP Pedicle	0.612
AP Spinous Process	0.210
AP Vertrbra	0.7981
LAT Anteror Vertebral Line	0.029
LAT Disk Height	0.6937
LAT Posterior Vertebral Line	0.0050
LAT Spinous Process Line	0.6505
LAT Vertebral Line	0.8089

3.2 Analysis

3.2.1 Hyperparameters:

We used dropout, Early Stopping in our model. These contributed to the hyper parameters. Further, ReduceLROnPlateau, num_epochs, batch_size are other hyper parameters. Primarily, we chose the hyper parameters, on hit-and-trial basis. We wanted to run a grid-search but the time constraint did not allow it.

3.2.2 Learning Curves:

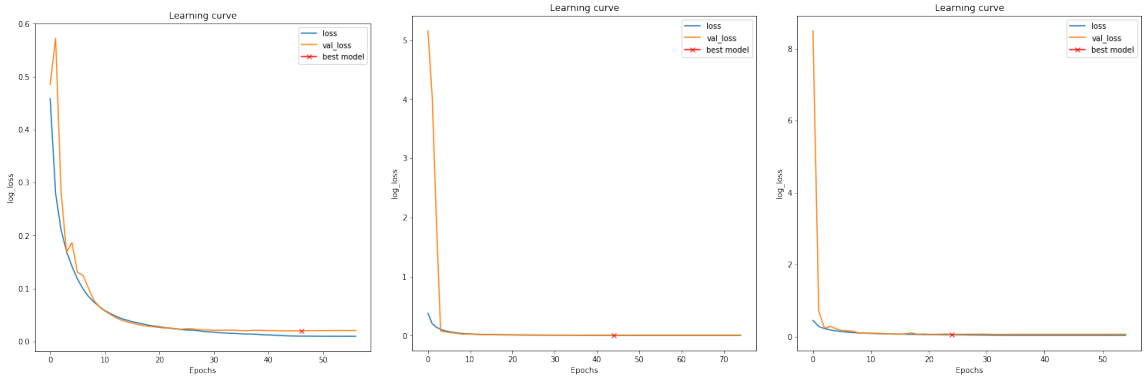
The learning curves for all masks are shown below. The metric used while making these curves was "DICE" and the loss function was "binary_crossentropy".

3.2.3 Validation Loss:

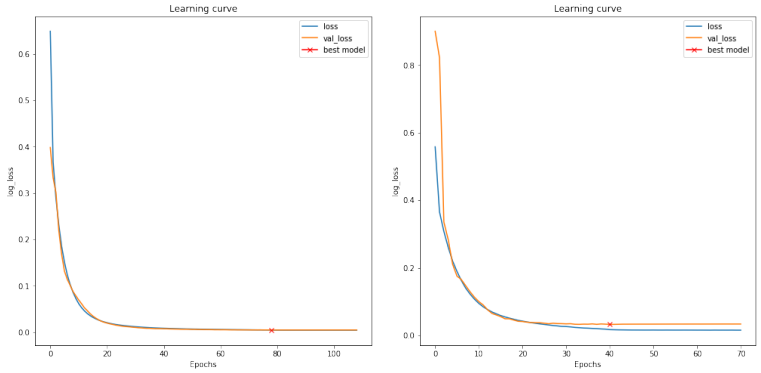
We adjusted the validation split in data to 0.1 ratio. The DICE coef. for the same and the validation loss is provided in Table 1.

3.2.4 Threshold:

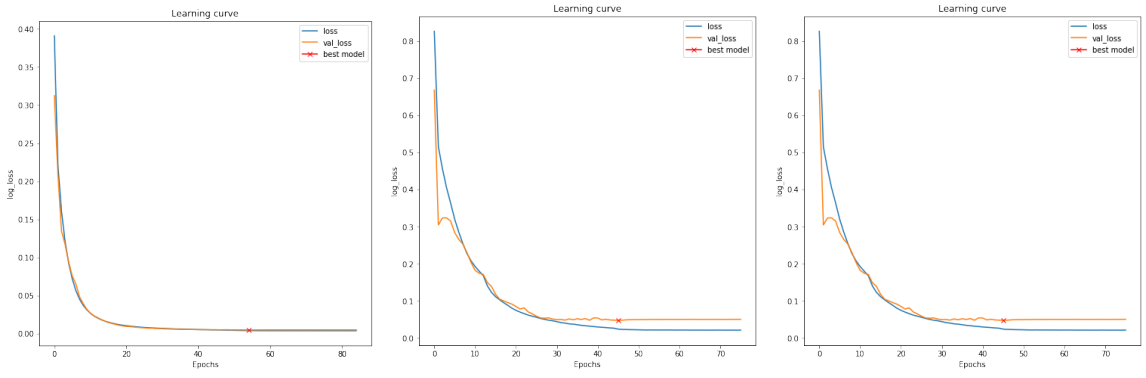
In some images, the segmentation masks were very sparse. In some cases, the masks were a line of width 1 pixel. To deal with such situation, we reduced the threshold (non-linearity) after output layer from 0.5 to 0.13. This helped to get us some image rather than getiing complete black image.



(a) AP Pedicle (b) AP Spinous Process (c) AP Vertebra



(d) LAT Anteror Vertebral Line (e) LAT Disk Height



(f) LAT Posterior Vertebral Line (g) LAT Spinous Process Line (h) LAT Vertebral Line

Figure 10: Learning Curves for all masks

Future Work

As future work, we can apply Principle Component Analysis on input data to get set of linearly uncorrelated features, then train the resulting features on models such as Support Vector Machine which work well on less amount of data.

Acknowledgments

I would like to thank Dr. Prathosh AP for guiding us throughout our assignment and providing this dataset. I would like to thank Mr. Arpit Jain for some fruitful discussion, Mr. Gantavya Bhatt for being an awesome teaching assistant. I would like to thank RIKEN Center for Artificial Intelligence, my former employer for providing the computation resources (RAIDEN).

References

[1] Ronneberger O., Fischer P., Brox T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention – *MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham*