

Assignment 3: Report

(u2151556)

Task-1

Approach -

1. Data preparation - Images of size **4500x4500**, are considered in the ORCA validation dataset for training the model, as all the test dataset images are of 4500x4500 and around 91% of the total images in the validation dataset holds the same. Further, the ORCA validation set is split in ratio "**0.9 : 0.1**" for training and validation of the model respectively. Non-overlapping patches of size **1000x1000** are extracted from the training, validation and test dataset images with appropriate padding and stored in the database. For better time efficiency, the **HDF5** file format is employed using **pytable** extension to compress and store the patches.

Each dataset(training,validation and test) consists of both images and their corresponding masks. The masks in the original datasets didn't have three distinct intensities for tumour, non-tumour and background pixels. Therefore, the intensity histogram is explored for all the images, and found that the majority of the pixel densities are concentrated at **0,129** and **255** intensity values(fig-1-a). So, to get the masks in the required format for training the model i.e., for three class segmentation problems, the pixels should be encoded as **0,1, & 2** values. We came up with a smart approach, by dividing the mask array with **129** and taking the round-off at each pixel, resulting in an encoded mask as desired for the model training with following mapping {0 → 'background' ; 1 → 'non-tumour' ; 2→ 'tumour' }. In (fig-1-b), the yellow, green and dark blue represents the tumour, non-tumor and background regions in a sample mask.

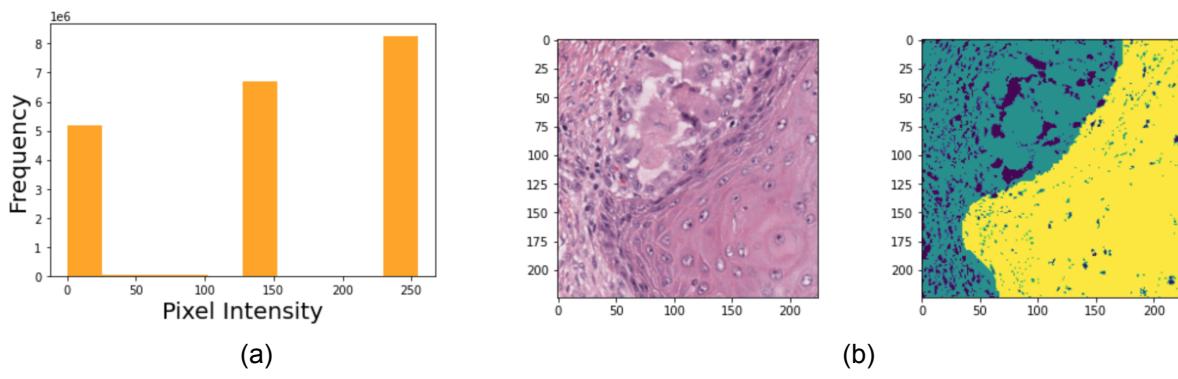
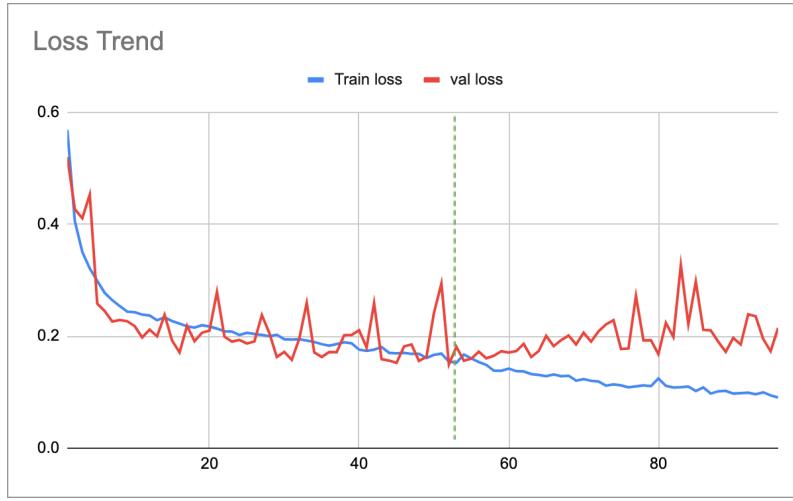


Figure-1

2. Model and results - A Unet architecture with “**depth:5**” and “**wf:2**” is first trained with Adam optimizer for 100 epochs because minimum loss for the validation dataset is observed around 50-60 epochs. Following chart shows the trend of training and validation loss to locate the optimal point (best model) after which the validation loss starts increasing and the train loss keeps decreasing.



The trained model resulted in lower **Dice score** (metric used for model evaluation), as shown in (figure-2-a) with mean $\rightarrow 0.55$ & standard deviation $\rightarrow 0.18$. After digging deeper in the data, we realised that the model was poorly segmenting the corner patches of test images, as shown in (figure-3-a), reason being, (i) the model had less complexity, (ii) the model couldn't learn such data points properly as it was not exposed to more of such data points, given an image only has four corner patches. So, to solve this problem, we increased the **depth** to **6** and the **wf** to **3** to add more complexity, and upsampled the corner patches by replicating in the training dataset. Bingo ! it resulted in better performance (dice score mean $\rightarrow 0.69$ & standard deviation $\rightarrow 0.16$, figure-2-b).

Other than this, the major challenge was the runtime in training the model, each epoch was taking around 4-5 minutes to execute. It got fixed by saving the patches post resizing them(*Data preparation* section), earlier which was happening while calling the Dataloader. The runtime for one epoch cut down to 30-50 seconds.

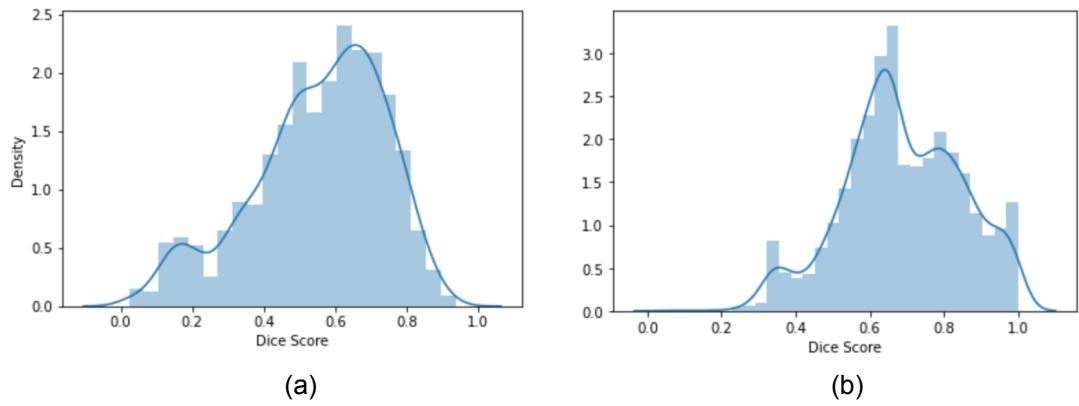


Figure-2

Further, after looking at the test patches and their prediction masks from top and bottom percentiles of dice scores, we observed that the model is able to segment those patches effectively, which has clear and distinctive visible boundaries between tumour and non-tumour regions(figure-3-c). And patches with low contrast and less clear boundaries were segmented poorly.(figure-3-b).

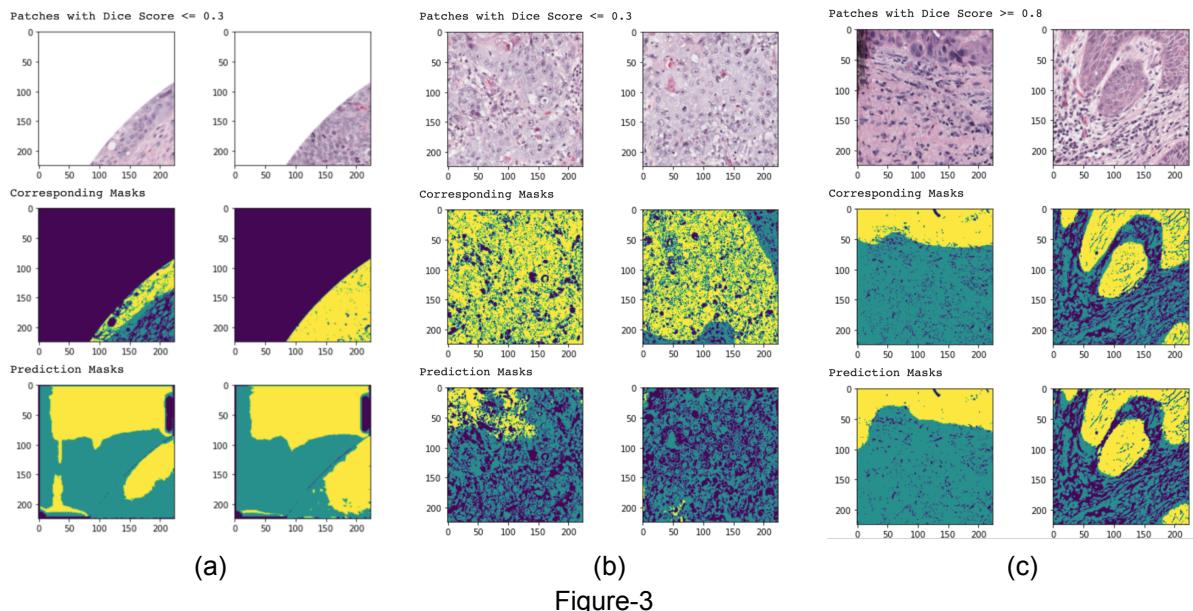
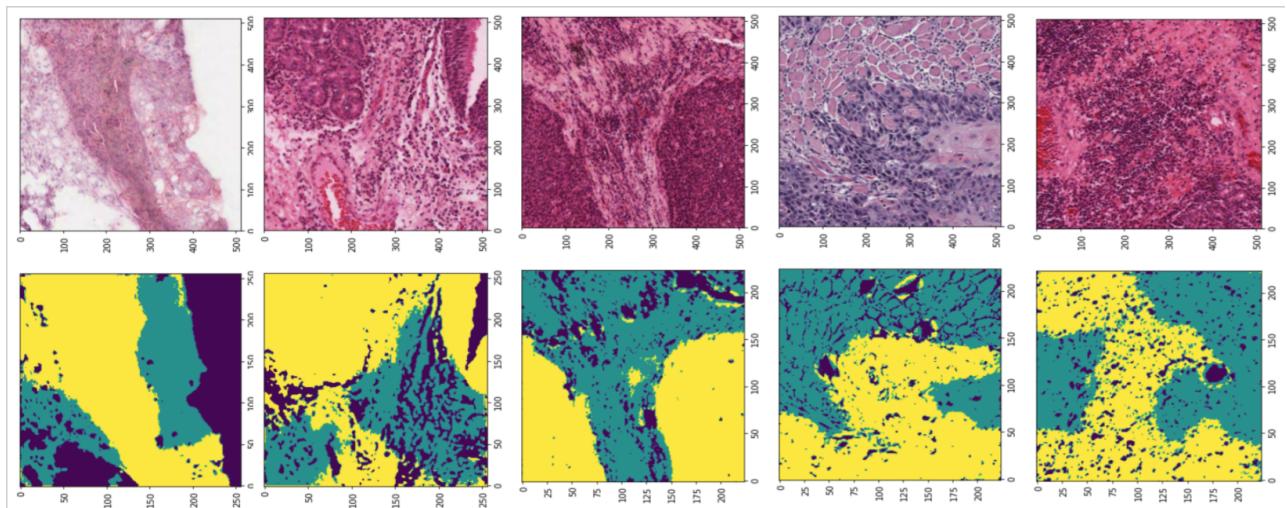


Figure-3

Task-2

Approach - To extract the HPFs from the provided WSIs, some random pixel positions are chosen from the binarised thumbnail (with least magnification level as they are fast to read) of a WSI. These random pixel positions(in the maximum level) are extrapolated in the level-0(maximum magnification level) image, as the **read_region** method in **OpenSlide** library only considers level-0 coordinates in the location argument. Then **512x512** patches are extracted from the desired level of WSIs of corresponding magnification of **0.5 mpp** (20X) using the extrapolated pixel positions. Further, a condition is applied while extraction, i.e the ratio of the non-background pixels in the patch should be greater than 0.5, to increase the likelihood of the patches resulting in high cellularity regions.

Using the trained segmentation model in Task-1, prediction masks are generated for the extracted HPFs. Below are a few sample patches and their predicted masks.



The model seems to perform well on the HPFs as it is accurately able to locate and segment the tumour, non-tumour and the background regions.

Task-3

Approach - Connected components are labelled using `sklearn.measure.label` on the binarised tumour(`mask == 2`) and non-tumour (`mask == 1`) images separately post removing small objects from the binarized images with closing, erosion and dilation methods from `skimage.morphology` library. Then all the three features RTNA, RCTN and AVCT are calculated for the patches extracted and segmented in Task-2 at both patch level and WSI level. At patch level, features are calculated as per asked in the question, except AVST which is simply considered as count of the number of tumour regions in a patch. At WSI level, RTNA, RCTN are calculated by aggregating(sum) the tumours and non-tumour metrics separately, then taking ratio of them, and AVST is the average number of tumour regions across patches of a WSI.

A correlation chart is plotted between calculated features to find any correlation amongst features at both patch and WSI level. RCTN, and AVCT feature appears to show positive correlation with value – and – at patch and WSI level respectively.

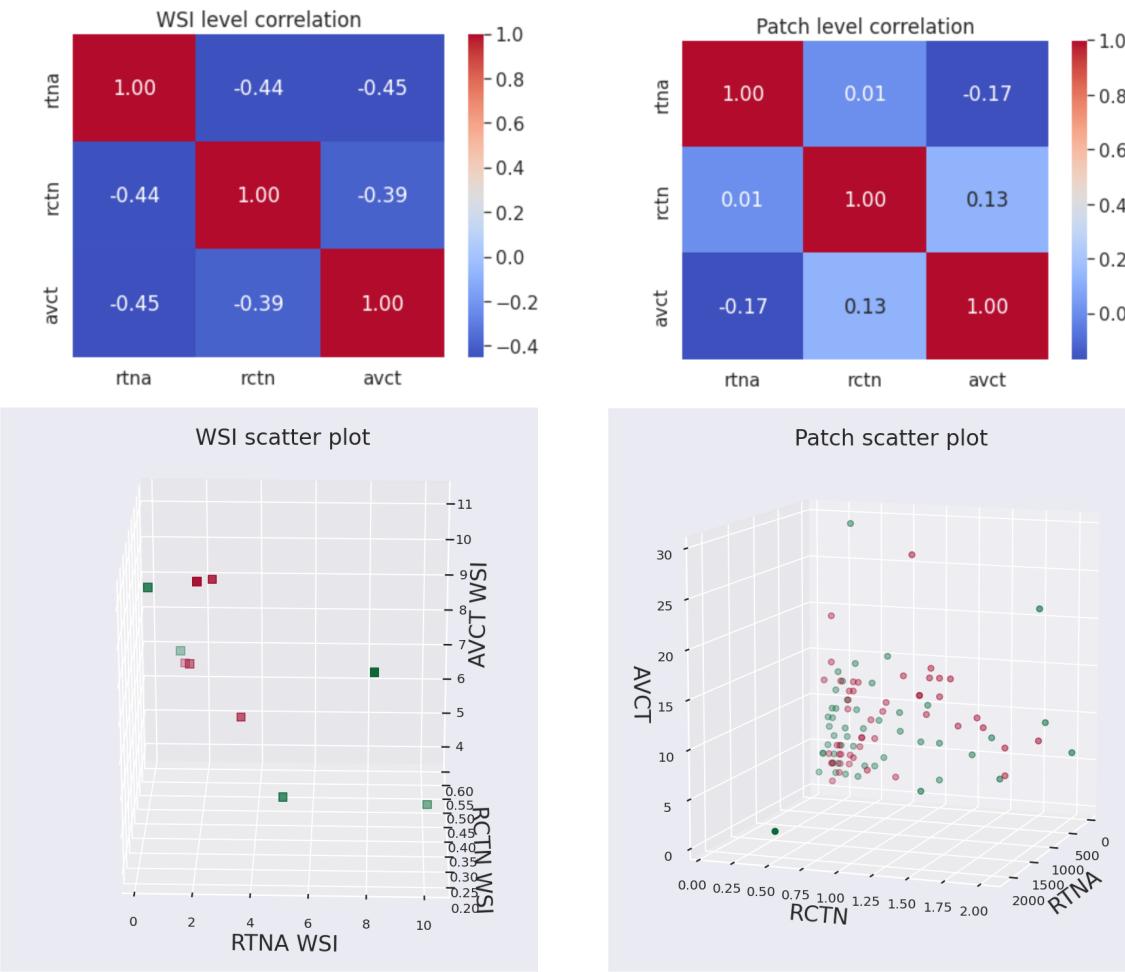


Figure-4(red dots represents HPV- and green dots represents HPV+)

At WSI level, RTNA is found to be correlated with RCTN and AVST. So, we can remove RTNA from the feature set but it might lead to information loss because correlation is not very strong, ~45% with both. But if we look at the correlation matrix patch level data, there is

no significant correlation among all the variables. So, it would be better to consider all the three features for classification.

From the scatter plot of patch, it is visible that the significant variation is along AVCT and RCTN, and very mild variation is visible along the RTNA axis. Whereas in WSI scatter plot, major variation is along AVCT and RTNA axes. There is a clear distinction visible in the WSI scatter plot but not in patch scatter plot, which could happen due to less data points in WSI plots and also the WSI plot is plotted using aggregated values from patch level which could be the reason for the ambiguity between both the scatter plots.