Cell Density Estimation in H&E Stained Whole-Slide-Images

CS904 Assignment 2, Annika Stechemesser

I. Introduction

In this assignment we are looking at the Whole-Slide-Image (WSI) shown in figure 1, which was stained with Hematoxilyn and Eosin (H&E). All the code used to produce these results was written in MATLAB. Details about the implementation can be found as comments in the accompanying code.

Figure 1: Whole Slide Image at level 2.

The image has a width of 79328 pixels and a height of 199368 pixels. It comes with 11 wavelet decomposition levels. We pick a patch on the image of size of 512x512 at level 0 and display it at all reduction levels to inspect the quality(Figure 2). We see that the resolution quality declines rapidly past level 3. Because of the size of the WSI we always have to make a trade-off between the resolution (level) at which we work and the computational effort/resources required. To get significant results in a reasonable computing time I carry out WSI level calculations (e.g. segmenting cells over the whole slide) at level 3.

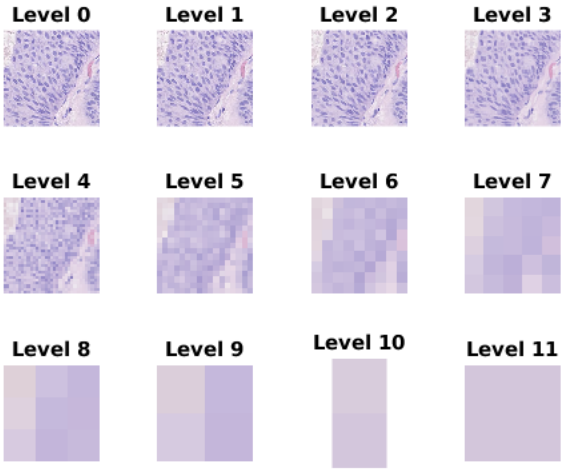


Figure 2: Resolution at different levels

II. Task 1: Tissue Segmentation

In this task we want to segment tissue regions in the image shown in figure 1. Two different methods are presented.

A. Tissue Segmentation using a local entropy filter

In this context, we consider entropy as a measure of activity, meaning a variation of pixel intensities. We use a 15x15 sliding window over the WSI at level 3 and compute the local entropy of the pixel intensities. We assume a bimodal distribution, meaning we assume the data to be separable in two clearly distinct groups (tissue or no tissue) where a high entropy corresponds to tissue and a low entropy corresponds to no tissue. To set the threshold we use Otsu thresholding. Otsu's method calculates the optimum threshold separating the two classes such that their intra-class variance is minimal. Using this threshold we compute the binary mask, where all pixels with a local entropy greater than the threshold get classified as tissue and everything else as background.

Finally, we want to prune out all regions that are smaller than a certain size (i.e. noise in the segmentation mask). To do that, we perform morphologically opening and closing (where the structuring element is in both cases a disk of radius 25) which helps remove small bright and dark spots respectively. The difference before and after the pruning is shown in figure 3. Every step of the segmentation process is visualized in figure 4.

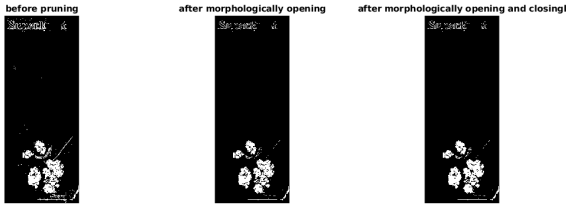


Figure 3: Effect of pruning (local entropy filter mask).

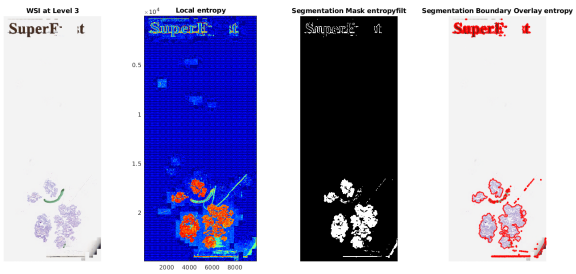


Figure 4: Cell segmentation process using a local entropy filter

The plots suggest that the method was successful in separating the tissue from the background. As it is a binary method,

it also classified some other variation rich structures on the slide as tissue, such as the edge of the writing in the top or parts of the green mark. This could be avoided by setting a higher pruning threshold, but in this case we would risk losing some actual tissue.

B. Tissue Segmentation using local standard deviation

We use a 15x15 (level 3) sliding window over the whole slide and calculate the local standard deviation in each of these windows. If the color variety, corresponding to the variety in pixel intensities, is high the local standard deviation will be large. If the variety of pixel intensities is low, the local standard deviation will be small. Most of the background is uniformly white and will therefore have a small local standard deviation. The actual tissue is due to the staining in more vibrant shades of pink and purple and will therefore have a large local standard deviation. We again use Otsu thresholding and then produce a binary mask according to this threshold, where white corresponds to tissue and black corresponds to no tissue. On our final mask, we prune out out all regions that are smaller than a certain size or do not contain any tissue in the same manner as before (structural element: disc with radius 25). The difference before and after the pruning is shown in figure 5, every step of the segmentation process using a local standard deviation filter is shown in figure 6.

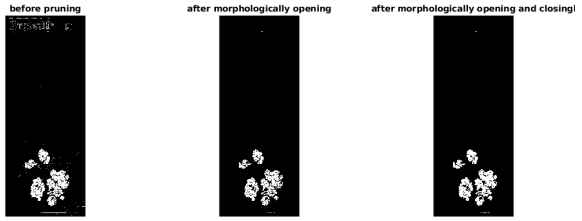


Figure 5: Effect of pruning (local standard deviation filter mask).

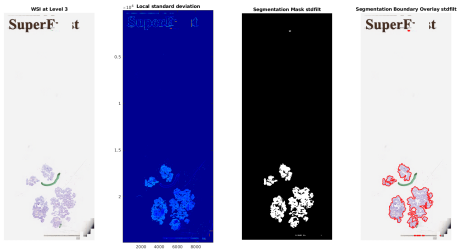


Figure 6: Cell segmentation process using a local standard deviation filter

Using a local standard deviation filter, we can successfully separate tissue from background. Overall this filter seems a lot less sensitive than the local entropy filter. The heat map [Figure 6, subplot 2] suggests that the local standard deviation is varying less between tissue and background than the local entropy [Figure 4, subplot 2]. The tissue regions still get segmented correctly, but the method is less sensitive for noise, as only small parts of the writing and the green mark on the slide get picked up and are not visible anymore after pruning. Overall, using a local standard deviation filter gave in this case the better result.

III. Task 2: Stain Separation and Nuclei Region Segmentation

A. The Ruifrok-Johnston Stain Deconvolution Method

Now we want to use the Ruifrok-Johnston stain deconvolution method to separate the Hematoxylin and Eosin stain channel images. Hematoxylin stains regions with a low pH value, for example DNA (which is an acid), as dark blue or purple. Eosin stains regions with a high pH value, for example cytoplasm or connective tissue, as pink.The RGB channels of a stained image as we have it here can't be directly used for stain measurement as the relationship between the RGB values and the stain concentration is nonlinear. Therefore we want to consider the problem in optical density space. The optical density d is defined as $d = -log_{10}(I/I_0)$ where I_o is the intensity of the light entering the specimen. We use the stain matrix proposed by Ruifrok and Johnston:

$$M = \begin{bmatrix} 0.6442 & 0.7166 & 0.2668 \\ 0.0928 & 0.9541 & 0.2831 \\ -0.0517 & -0.1576 & 0.5482 \end{bmatrix}$$

The rows in the matrix M correspond to RGB colors of Hematoxylin, Eosin and background in the optical density space. The observed optical density $d = [d_{\text{red}}, d_{\text{green}}, d_{\text{blue}}]$ is then a linear combination of the stain concentrations $c = [c_{\text{hematoxylin}}, c_{\text{eosin}}, c_{\text{background}}]$ in the sample: $d = c * M$. Using this information, we can now estimate c as $c = d * M^{-1}$. Figure 7 shows the results of the Ruifrok-Johnston stain deconvolution method on a patch of the image of size 512x512 at level 0.

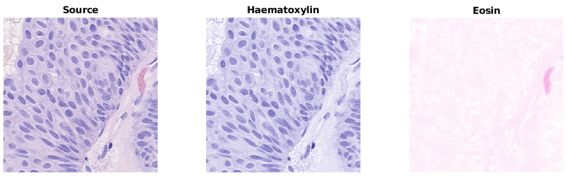


Figure 7: Ruifrok-Johnston stain deconvolution method - results

We can clearly see the separation of the stains. In the Hematoylin image the nuclei are colored dark, in the Eosin image the surrounding tissue is colored pink. While the RJ method gave satisfying results in this case, it is limited by the fact that the stain matrix needs to be estimated. Even though the computational complexity of the algorithm itself is quite low, the estimation of the stain matrix can be tedious and inefficient.

B. Nuclei Region Segmentation using Otsu Thresholding on the Hematoxylin Channel

Next, we want to use the results of the stain separation to identify regions in the image containing nuclei. To do that,

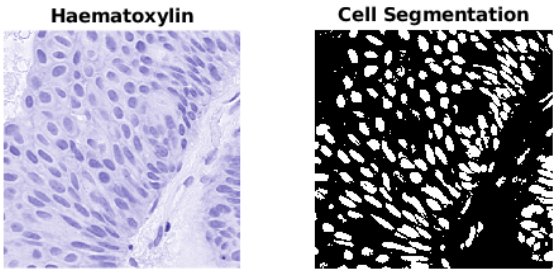
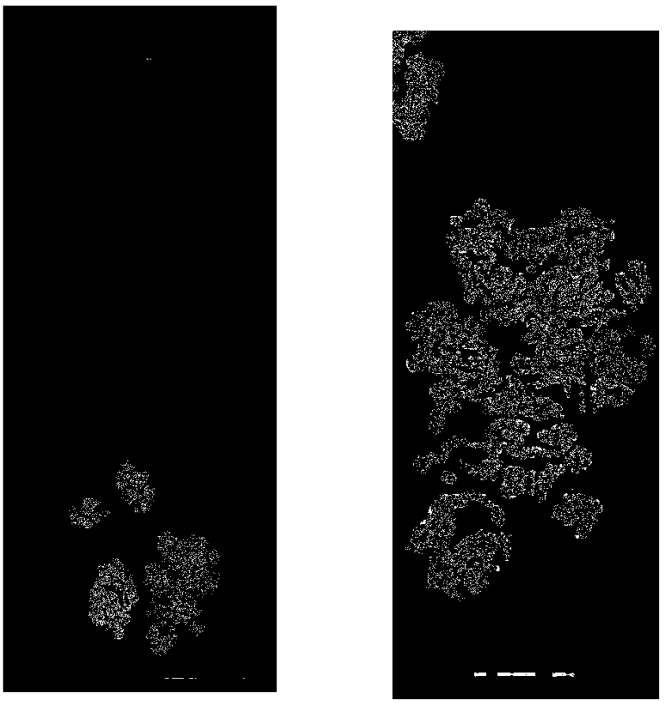


Figure 8: Cell segmentation using Otsu thresholding on the Hematoxylin channel

we perform Otsu thresholding on the hematoxylin channel. As mentioned before, Otsu's method calculates the optimum threshold separating the two classes so that their intra-class variance is minimal. We again assume the image is bimodal and attempt to separate into nuclei and background, where a high hematoxylin concentration indicates nuclei (as DNA is an acid and Hematoxylin stains acids). According to the threshold we generate a binary mask, where white regions correspond to the nuclei. The results are displayed in figure 8 for a 512x512 patch at level 0.



(a) Segmented nuclei on the WSI

(b) Segmented nuclei on the WSI (zoomed in)

Figure 9: Segmented nuclei using thresholding on the H-channel

We see that the method successfully separates the nuclei from the background. Now we want to segment nuclei in the whole slide. To do that we decompose the image on level 3 into block of size 64x64 (so they correspond to blocks of size 512x512 on level 0) and perform the stain deconvolution and the cell segmentation in the same manner as before on each individual block using the MATLAB function "blockproc". We then reconstruct the binary mask for the whole image from the masks for the individual blocks (details regarding the implementation can be found in the code). Some of the blocks don't contain any tissue so the algorithm can't generate a sensible output. Therefore, to get a meaningful result, we use the tissue segmentation mask we produced in Task 1 (local std. filter version) to make sure we only consider regions with tissue. The results are shown in figure 9.

C. Nuclei Region Segmentation using Thresholding on the Blue Ration Image

An alternative method we can use to achieve a rough segmentation of regions in the image containing nuclei is to use Otsu thresholding on the blue ratio image. Our whole slide image is an RGB image. If we extract the channels, we can compute the blue ratio image according to the following formula:

$$BR = 100\left(\frac{B}{1+R+G}\right)\left(\frac{256}{1+R+G+B}\right)$$

The blue ratio measures relative blueness (i.e. relative to red and green). The staining colors nuclei 'bluer' than the surrounding image, therefore they have higher relative blueness and appear lighter on the blue ratio image. We then do Otsu thresholding on the blue ratio image and generate the binary mask. Figure 10 shows the result for the same patch as before. Using the blue ratio image also successfully separated the nuclei from the background. We apply the method to the whole slide image on level three in the same manner as before. The results are shown in figure 11. Both method detect nuclei in the same regions of the picture.

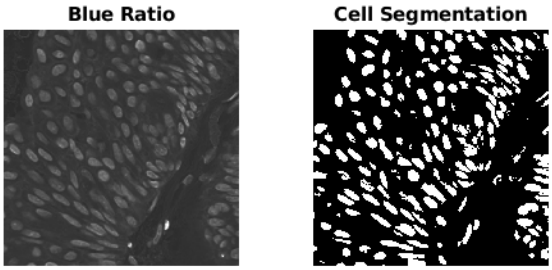
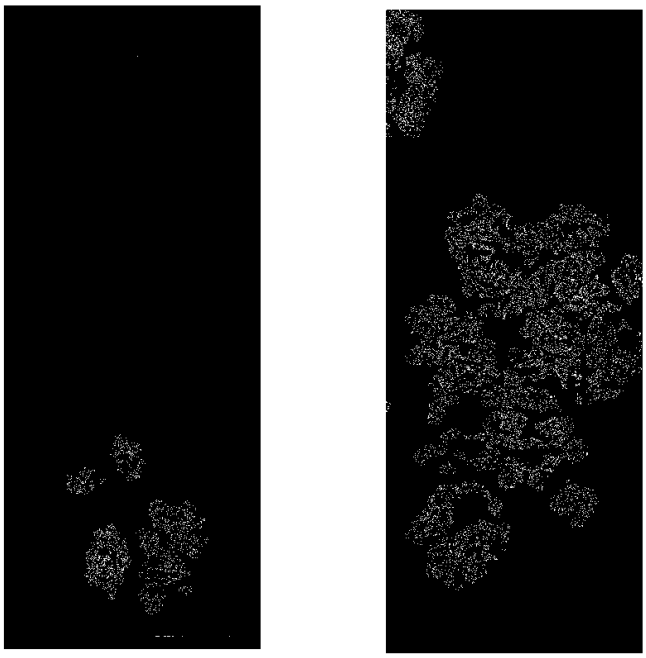


Figure 10: Cell segmentation using Otsu thresholding on the Blue Ratio Image

A third method to detect individual cells is the LIPSyM method. LIPSym is short for Local Isotropic Symmetry Measure. It is based on the idea that most cells form an isotropic symmetric structure, which basically means that whichever direction through the nucleus is chosen as a an axis, the cell will be symmetric on this axis. The LIPSym method computes the symmetry pixel wise and measures which ones are isotropic. Local maxima in the symmetry image correspond to the locations of cell centers, i.e. nuclei. This method wasn't implemented here and is only included for completeness.



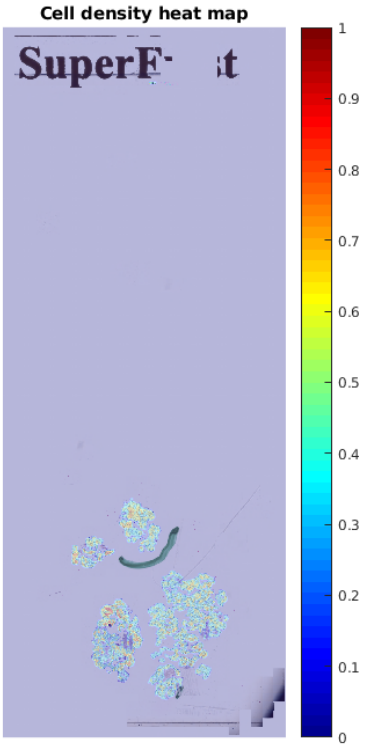
(a) Segmented nuclei on the WSI

(b) Segmented nuclei on the WSI (zoomed in)

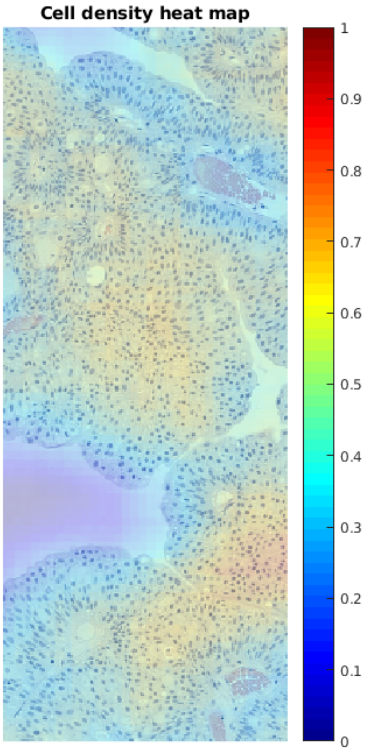
Figure 11: Segmented nuclei on the whole slide image (Thresholding on BR Image)

IV. Task 3: Local Cell Density Estimation

We now want to use the results from the previous task to measure the local cell density across the whole image. We define the local cell density as number of cells in a local region. To compute this we compute the number of connected components (white regions) in a local region. On the one hand we want to decompose the picture into very small boxes to get a precise result and a high resolution in the density image. On the other hand the boxes also need to be big enough to actually contain some cells, else the local density will always be zero or one. To solve this issue we let the boxes be of size 65x65 at level 3 ($\approx$ 512x512 at level 0) and let them be overlapping (stride 5). Processing the individual (overlapping) boxes, we compute the local cell density. Finally we normalize and smooth the results using a 2D Gaussian smoothing kernel with standard deviation of 0.5. To visualize our results we use alpha blending to overlay the original whole slide image with the densities, which we plot as a heat map (Figure 12). We consider the whole slide image as well as a zoomed in region. The density measure seems to successfully point out the regions with high cell density.



(a) Cellularity (whole slide image)



(b) Cellularity (zoomed in)

Figure 12: Cell density heat maps