

GRADUATE CERTIFICATE INTELLIGENT SENSING SYSTEMS PRACTICE MODULE REPORT - EARLY DETECTION AND DIAGNOSTIC OF PROSTATE CANCER (HIGH QUALITY SEGMENTATION)

Zhang Yu (A0213498X), Natarajan Anandan(A0213514U)

Group 9, Institute of Systems Science, National University of Singapore, Singapore 119615

ABSTRACT

Cancer detection has always been a major issue for the pathologists and medical practitioners for diagnosis and treatment planning. The manual identification of cancer from microscopic biopsy images is subjective in nature and may vary from expert to expert depending on their expertise and other factors which include lack of specific and accurate quantitative measures to classify the biopsy images as normal or cancerous one. The automated identification of cancerous cells from microscopic biopsy images helps in alleviating. As part of this practice module project, we aim to provide a system to automate the identification of the cancerous cells from microscopic biopsy images which can be used by doctors and pathologist for early detection and treatment. We have proposed a method of segmentation and classification technique based on the features and utilized Unet model for segmentation.

Video: <https://youtu.be/F6nC36dUT6I>

Index Terms— early detection, identification, segmentation technique

1. INTRODUCTION

Cancer detection has always been a major issue for the pathologists and medical practitioners for diagnosis and treatment planning. The manual identification of cancer from microscopic biopsy images is subjective in nature and may vary from expert to expert depending on their expertise and other factors which include lack of specific and accurate quantitative measures to classify the biopsy images as normal or cancerous one. For the detection and diagnosis of cancer from microscopic biopsy images, the histo, pathologists normally look at the specific features in the cells and tissue structures. The various common features used for the detection and diagnosis of cancer from the microscopic biopsy images include shape and size of cells, shape and size of cell nucleus, and distribution of the cells.

Prostate cancer is also given a grade called a Gleason score. This score is based on how much the cancer looks like healthy tissue when viewed under a microscope. Less aggressive tumours generally look more like healthy tissue. Tu-

mours that are more aggressive are likely to grow and spread to other parts of the body. They look less like healthy tissue. The Gleason scoring system is the most common prostate cancer grading system used. The pathologist looks at how the cancer cells are arranged in the prostate and assigns a score on a scale of 3 to 5 from 2 different locations. Cancer cells that look like healthy cells receive a low score. Cancer cells that look less like healthy cells or look more aggressive receive a higher score. To assign the numbers, the pathologist determines the main pattern of cell growth, which is the area where the cancer is most obvious, and then looks for another area of growth. The doctor then gives each area a score from 3 to 5. The scores are added together to come up with an overall score between 6 and 10. Gleason scores of 5 or lower are not used. **The lowest Gleason score is 6, which is a low-grade cancer. A Gleason score of 7 is a medium-grade cancer, and a score of 8, 9, or 10 is a high-grade cancer.** A lower-grade cancer grows more slowly and is less likely to spread than a high-grade cancer. Here is the current manual process.

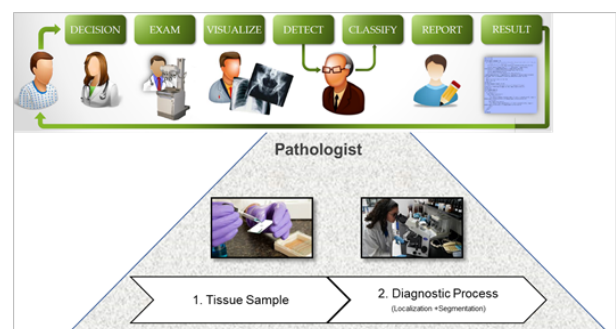


Fig. 1. Current Manual Process.

Challenges with the current process are as follows :

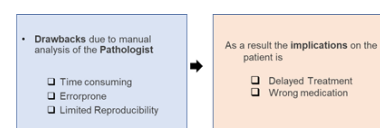


Fig. 2. Challenges in the Current Manual Process.

Our team had to understand the medical area of cancer and how it is graded. We also went through several papers to know the current areas of research both in the field of automated cancer identification and chemical processing of the tissue samples for better determination of the cancer cells. We have developed and trained the model,

1. The **input is the image** of the cancerous cell for pathologist to upload.
2. For training purposes, we have created several samples from the training data available by different augmentation methods.
3. Built the segmentation model and currently performing training to improve the dice coefficient results.
4. The **output is the image of the segmentation**, shown in the user interface as an image adjacent to the input.

2. LITERATURE REVIEW

Having defined the business problem, our group's aim is to a novel approach to provide High Quality Segmentation providing extremely helpful in clinical practice for Early Detection and Diagnostic of Prostate cancer.

To develop a model to classify and segmentation, Our team had to understand the medical are of cancer and how it is graded. We also went through several papers to know the current areas of research both in the field of automated cancer identification and chemical processing of the tissue samples for better determination of the cancer cells.

The first step is usually a visit to the general practitioner, for example because the patient suffers from pain during urination or from frequent urges to urinate. The GP refers the man to a urologist who starts the hospital-based diagnosis trajectory, including measuring prostate-specific antigen (PSA) levels, increasingly including an MRI exam, and pathology testing. After prostate cancer has been diagnosed, a treatment plan is made. This can be a more direct approach, usually radiation therapy or surgery), or a watchful approach, the active surveillance pathway.[1]

Deep learning methods[2] have shown promising results in a variety of computer vision tasks such as segmentation, classification, and object-detection. These methods consist of convolution layers that are able to extract different features from low-level local features to high-level global features from input images. A fully connected layer at the end of the convolutions neural layers converts convoluted features into the probabilities of certain labels. Different types of layers, such as batch normalization layer, which normalizes the input of a layer with a zero mean and a unit variant, and dropout layer, which is one of regularization techniques that ignores randomly selected nodes, have been shown to improve the performance of deep learning-based methods. Nevertheless,

to achieve convincing performance, an optimal combinations and structures of the layers as well as precise fine-tuning of hyper-parameters are required. This remains as one of the main challenges of deep learning-based methods when applied to different fields such as medical imaging.

It is not hard to see that AI can provide similar support in the pathology area, as we previously described in the imaging section. AI has been amply deployed to detect cancer in pathology images. Research has shown that algorithms are up to this task by providing impressive accuracy scores. [1]Tolkach et al., for example, claimed an accuracy of 98

To make this Gleason score-setting algorithm less of a black box, a step in between can be letting the algorithm determine which areas in the sample show a certain Gleason pattern. In other words, the algorithm localizes and segments patches with the same Gleason score, creating a visual way for pathologists to assess the AI results.[3]

3. DATASET

Given the confidentiality limitations in accessing the patient's data, we have used the publicly available data set available from various images from internet (public dataset) and primarily using Harvard Dataverse dataset. [4]

HE stained images [5] from five prostate cancer Tissue Microarrays (TMAs) and corresponding Gleason annotation masks. In the masks, pixel indices correspond to classes as follows: 0=Benign (green), 1=Gleason₃(blue), 2 = Gleason₄(yellow), 3 = Gleason₅(red), 4 = unlabelled(white)

The images are pre-processed, ie., during this step the data gets transformed, or Encoded, to bring it to such a state that now the machine can easily parse it.

Adaptive Histogram Equalization : We have used this to improve the local contrast and enhancing the definitions of edges in each region of an image

Image Augmentation: Creation of training images through different ways of processing or combination of multiple processing, such as random rotation, shifts, shear and flips, etc. Mentioned here are the different methodologies used for Augmentation.

The input files are of size 3100x3100 before processing. Post pre-processing with above methods, the size are constantly kept as 256 x 256 resolution.

Index Terms— HSV, Flipping, Colour Jitter, HED, Random Scaling

A sample image from data set below.

The dataset consists of 888 images, with following details. TMA 76 is Validation Set TMA 80 is Test Set TMA 111, 199 204 are used as Training Set.



Fig. 3. Sample Dataset Image

Table 17. prostate cancel data.

Name	Benign	6(3+3)	7(3+4,4+3)	8(4+4,3+5,5+3)	9(4+5,5+4)	10(5+5)
TMA76	42	35	25	15	2	14
TMA80	12	88	38	91	3	13
TMA111	0	95	39	69	16	8
TMA199	61	69	17	26	2	1
TMA204	0	1	2	25	8	69

Fig. 4. Dataset

4. PROPOSED SYSTEM

Our proposed methodology for the detection of the cancerous cells involves various stages as shown in the below model. For the detection and diagnosis of cancer from microscopic biopsy images, the histopathologists normally look at the specific features in the cells and tissue structures. The various

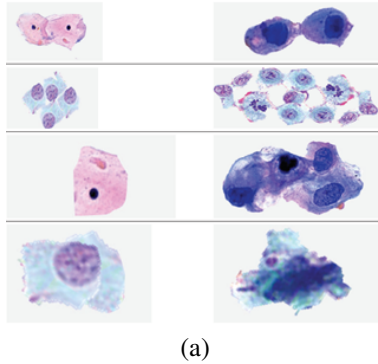


Fig. 5. View of Normal and Cancerous cells

common features used for the detection and diagnosis of cancer from the microscopic biopsy images include shape and size of cells, shape and size of cell nuclei, and distribution of the cells. Our aim is to design an application which can automate the process of Cancer detection. After delving through the various methodologies and performing the comparative analysis we have finalized on the methods explained below.

4.1. Preprocessing

During this stage, the microscopic images are huge in size and the samples required for the training are usually limited. Hence, we performed image enhancement and augmentation to improve the image quality required for the segmentation

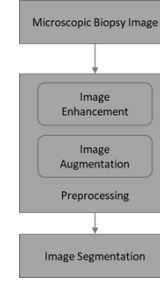


Fig. 6. Model for detection of Cancerous cells

and for the training of the model.

Adaptive Histogram Equalization :

This methodology was used to enhance the picture quality. This approach is suitable for improving the local contrast and enhancing the definitions of edges in each region of an image. This method was chosen to enhance the im-

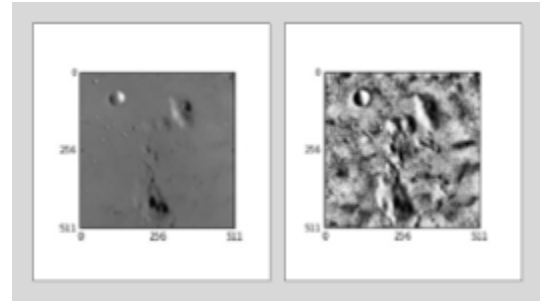


Fig. 7. Comparison of the Original image and the Adaptive Histogram Equalized image

age quality since is adaptive method computes several histograms, each corresponding to a distinct section of the image, and uses them to redistribute the lightness values of the image. It is therefore suitable for improving the local contrast and enhancing the definitions of edges in each region of an image.

Image Augmentation :

This methodology is required since we required a good amount of training data to improve the performance of the segmentation model. Hence, we had to rely upon creating training images through different ways of processing explained below. Stained colour augmentation is proven to reduce the generalization error by CNN model by simulating variations of the training data. This kind of augmentation leaves morphological features intact and focuses on simulating stain colour variations instead. Brightness contrast (BC): Morphological transformations was extended by random brightness and contrast image perturbations. Hue-Saturation-Value (HSV): This method works by randomly shifting the hue and saturation channels in the HSV colour space. This transformation produced substantially different colour distributions when applied to the training images.

Haematoxylin-Eosin-DAB (HED): This method followed three steps as explained below:

1. Disentangled the haematoxylin and eosin colour channels by means of colour deconvolution using a fixed matrix.
2. Perturb the haematoxylin and eosin stains independently.
3. Transform the resulting stains into regular RGB colour space.

During training, we selected the value of the augmentation hyper-parameters randomly within certain ranges to achieve stain variation. We tuned all ranges manually via visual examination. We used a scaling factor between [0.8, 1.2], elastic deformation parameters [80, 120] and [9.0, 11.0], additive Gaussian noise with [0, 0.1], Gaussian blurring with [0, 0.1], brightness intensity ratio between [0.65, 1.35], and contrast intensity ratio between [0.5, 1.5]. Using HED proved to provide better performance when compared with other data augmentation techniques such as HSV, Brightness and Contrast methods.

4.2. Image Segmentation

Labelling of each pixel of an image with a corresponding class of what is being represented. So the output of segmentation is the determination of each in to a particular class. U-net is a popular CNN model that was developed for medical image segmentation, it is able to work with fewer training images to yield more precise segmentation. U-net architecture contains two stages.

1. Contraction path /encoder: which is used to capture the context in the image. The encoder is just a traditional stack of convolutional and max pooling layers.
2. Symmetric expanding path /decoder: which is used to enable precise localization using transposed convolutions.

Hence it is an end-to-end fully convolutional network, i.e. it contains Convolutional layers and does not contain any Dense layer due to which it can accept image of any size.

We have used dice score methodology to quantify the performance of image segmentation. It is a method to evaluate the similarity between 2 sets of data, which in our case we evaluate the ground truth region with the output of the segmentation model.

A pixel can be classified as 1 or 0. If there is a mask in a pixel we state 1, if there is not a mask we state 0. Dice coefficient is a measure of overlap between two masks. 1 indicates a perfect overlap while 0 indicates no overlap. For eg, lets say the ground truth is mask A, the model generated a mask B, then

2

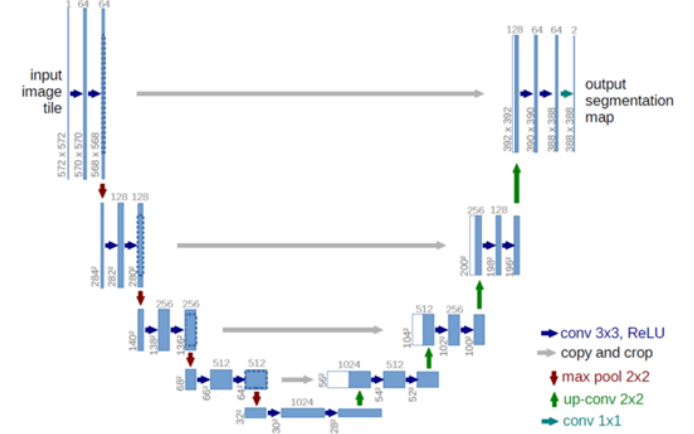


Fig. 8. U-net Architecture

$$DSC = \frac{2|A \cap B|}{|A| + |B|}$$

DSC: Dice similarity coefficient

Fig. 9. DSC Dice Similarity Coefficient

- • Number of positives is the total number of pixels that have intensity 1 in image A
- • Number of true positives is the total number of pixels which have the value 1 in both A and B. So it the intersection of the regions of ones in A and B. It is the same as using the AND operator on A and B.
- • Number of false positives is the number of pixels which appear as 1 in B but zero in A.

So the dice score is calculated as mentioned here,

$$\text{Dice score} = \frac{2 \cdot \text{number of true positives}}{2 \cdot \text{number of true positives} + \text{number of false positives} + \text{number of false negatives}}$$

Fig. 10. Calculation

The Dice score is not only a measure of how many positives you find, but it also penalizes for the false positives that the method finds, like precision. The only difference is the denominator, where you have the total number of positives instead of only the positives that the method finds. So, the Dice score is also penalizing for the positives that your algorithm/method could not find.

4.3. User Interface

We developed a simple User Interface for Pathologists to upload the image of the cell sample, and run the model and get the 'segmentation' result.

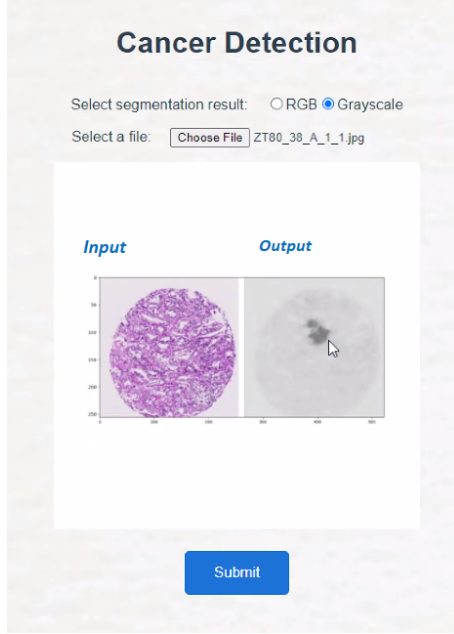


Fig. 11. User Interface

5. EXPERIMENTAL RESULTS

We evaluated various approaches and documented the quantitative and qualitative results of each approaches.

5.1. DeeplabV3 Model Evaluation

We have trained **DeeplabV3 with cityscapes dataset on resnet v2**. [6] Following are the (hyper)parameters used and the results are below. we were not able to proceed further due to low computation power. Find the result in Figure.

5.2. Approach 1: Unet segmentation with MobileNetV2

Inference speed is big concern due to low computation power, so MobileNetV2[7] has been chosen. Dice coefficient works well on imbalanced data, so used as loss function in Equation 1

$$DiceLoss = 1 - \frac{2 \sum_{pixels} (y_{true} * y_{pred})}{\sum_{pixels} (y_{true}) + \sum_{pixels} (y_{pred})} \quad (1)$$

Find the (hyper)parameters in table

Train Epochs	26
Batch Size	10
Decay	2e-4
max iteration	30000
initial learning rate	7e-3

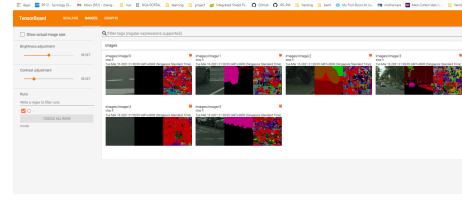



Fig. 12. DeeplabV3 on cityscapes dataset


Table 1. Approach 1 : RESULTS / OBSERVATION

train epochs	36
batch size	8
Learning Rate	1e-4
	
Dice coefficient	0.7871

5.3. Approach 2: Unet segmentation with Patch Pre-processing

We have small dataset with **large image size 3100*3100** and it is measure region by region as well, so created patch with **size 200*200**, the training size is increased to **73152 from 508**, so increased the batch size and reduced the epoch to speed up. we should train more, although the dice coefficient value increased but when look at some sample out, it is sparse. We use decay to adjust learning rate after every batch update. Find the (hyper)parameters in table

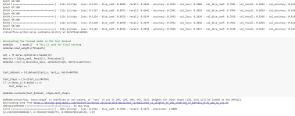
Table 2. Approach 2 : RESULTS / OBSERVATION

train epochs	16
batch size	32
Learning Rate	1e-4
momentum	0.9
decay	1e-4/25
	
Unet Coefficient (Approach1)	0.7871
Unet patch (Approach2)	0.8281

5.4. Approach 3: Unet segmentation by Increasing Epoch

Following is the setup by increasing Epoch

Table 3. The (hyper)parameters

train epochs	100
batch size	8
Learning Rate	1e-4
	
Unet Coefficient (Approach1)	0.7871
Unet patch (Approach2)	0.8281
Unet increased epoch (Approach3)	0.7858


5.5. Approach 4: Unet segmentation with HSV augmentation

1 HSV color space [8] augmented image is generated from each original image, trained on top of pre-trained unet model in Approach 1 and HSV augmentation function is provided below, subscript i represents hue or saturation channel, α is drawn from a uniform distribution $[-1, 1]$ [9]

$$S'_i = \alpha_i S_i \quad (2)$$

The reasons as for why our model was unable to fully segment the epidermal area in these images, varies from image to image. One image has a piece of folded tissue in the epidermis area, which our system didn't recognize. Another image comprise of an epidermis broken up into several parts, where some of these were deemed as too small and removed in the post-processing stage.

Table 4. The (hyper)parameters

train epochs	36
batch size	8
Learning Rate	1e-4
	
Unet Coefficient (Approach1)	0.7871
Unet patch (Approach2)	0.8281
Unet increased epoch (Approach3)	0.7858
Unet with HSV augmentation (Approach4)	0.8247

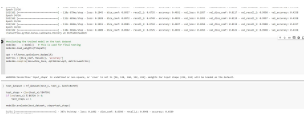
5.6. Approach 5: Unet segmentation with HED augmentation

3 HED augmented images are generated from each original image, the training size is increased to 1289 from 508 and

HED augmentation function is provided below in Equation, i represents each stain channel, α is drawn from a uniform distribution $U(1 - \alpha; 1 + \alpha)$, β is drawn from a uniform distribution $U(\alpha; \alpha)$, and typically $\alpha = 0.05$ [10]

$$S'_i = \alpha_i S_i + \beta_i \quad (3)$$

Table 5. The (hyper)parameters

train epochs	36
batch size	8
Learning Rate	1e-4
	
Unet Coefficient (Approach1)	0.7871
Unet patch (Approach2)	0.8281
Unet increased epoch (Approach3)	0.7858
Unet HSV (Approach4)	0.8247
Unet HED (Approach5)	0.8398

The result clearly shows that HSV and HED did not make any difference in the efficiency.

This would potentially be a limitation with the data set being used.

5.7. Approach 6: Unet multi-class segmentation

Try Multidimensional Dice coefficient since we have RGB masks Multidimensional Dice coefficient [11] is used as loss function and is provided in Figure 13 The evaluation results

```
# dice coefficient loss
smooth = 1e-15
def dice_coef(y_true, y_pred, smooth=1):
    intersection = K.sum(y_true * y_pred, axis=[1,2,3])
    union = K.sum(y_true, axis=[1,2,3]) + K.sum(y_pred, axis=[1,2,3])
    return K.mean((2. * intersection + smooth) / (union + smooth), axis=0)

def dice_loss(y_true, y_pred):
    return -dice_coef(y_true, y_pred)
```

Fig. 13. Unet multi-class segmentation

shows greater efficiency.



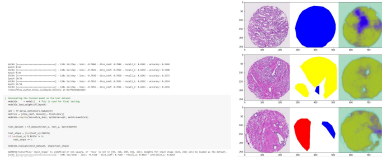
Table 6. The (hyper)parameters

train epochs	36
batch size	8
Learning Rate	1e-4
Unet Coefficient (Approach1)	0.7871
Unet patch (Approach2)	0.8281
Unet increased epoch (Approach3)	0.7858
Unet HSV (Approach4)	0.8247
Unet HED (Approach5)	0.8398
Unet multi-class segmentation (Approach 6)	0.5514

5.8. Approach 7: Unet multi-class HSV

try multi-class segmentation with HSV augmented data on top of pre-trained multi-class segmentation in Approach 6. The evaluation results shows greater efficiency.

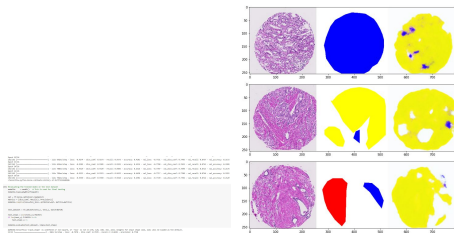
Following is the result (segmented) image.

**Table 7.** The (hyper)parameters

train epochs	36
batch size	8
Learning Rate	1e-4
Unet Coefficient (Approach1)	0.7871
Unet patch (Approach2)	0.8281
Unet increased epoch (Approach3)	0.7858
Unet HSV (Approach4)	0.8247
Unet HED (Approach5)	0.8398
Unet multi-class segmentation (Approach 6)	0.5514
Unet multi-class HSV (Approach 7)	0.7146

5.9. Approach 8: Unet multi-class HED

Try multi-class segmentation with HED augmented data

**Table 8.** The (hyper)parameters

train epochs	36
batch size	8
Learning Rate	1e-4
Unet Coefficient (Approach1)	0.7871
Unet patch (Approach2)	0.8281
Unet increased epoch (Approach3)	0.7858
Unet HSV (Approach4)	0.8247
Unet HED (Approach5)	0.8398
Unet multi-class segmentation (Approach 6)	0.5514
Unet multi-class HSV (Approach 7)	0.7146
Unet multi-class HED (Approach 8)	0.7978

6. CONCLUSIONS AND FUTURE WORK

6.1. Conclusion

Our team had a great time working on this project, and we definitely picked up some useful skills along the way.

The problem statement is to show the cancerous / non-cancerous cell segmentation as soon as possible, to the pathologist as a recommendation. With multiple segmentation approaches we tried in this practice module, we conclude the following.

1. Approach 2 (Unet Patch), Approach 4 (HSV) and Approach 5 (HED) are more than 80% accuracy. Color space augmentation (HSV or HED) can help increase the accuracy.
2. Classical Dice Coefficient Loss function is efficient (approach 1, 2,3,4 5) and have more accuracy than Multi-dimensional Dice Coefficient function (approach 6, 7 and 8).
3. Patch Processing with half of the epoch (epoch=16) can reach the higher accuracy than Processing without Patching (epoch =36), or patch processing with Higher Epoch (epoch=36).
4. More epoch does not mean higher accuracy. For eg., Approach 3 (epoch=100) is not having good accuracy than Approach 1 (epoch=36)

6.2. Future Work

1. More analysis can be done with combining multiple approaches to increase the accuracy of the model and we could use Otsu's threshold in Figure 14 to mask off some artifacts/imperfection in HE stained images. [12] in Figure
2. Classification: Classification of the segmented image, can help to arrive an accurate recommendation to pathologist.

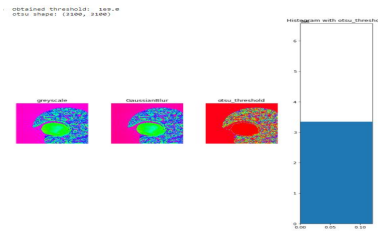


Fig. 14. otsu threshold

3. Use more geographical specific Private data-set to create and train the models to increase the usage of the model for specific institution (like major hospital).

7. CONTRIBUTIONS

7.1. Zhang Yu - A0213498X

Development, Integration, Training & Documentation : Research & understanding of the domain, Data-set Identification, Selection, and understanding. Technical Design. UI and model development , training & testing. Documentation & presentation & .

7.2. Natarajan Anandan - A0213514U

Project Managmeent, Busienss Analysis & Domain Analysis, Documentation & Presentations: Project Management, Research & understanding of the domain, Study the previous research and models already in the Technology area for this domain, UI Design & Testing. Documentation & presentation. Initial proposal, mid project presentation and final project documentation.

8. REFERENCES

- [1] Toma M. Tolkach, Y. and G Kristiansen, "V high-accuracy prostate cancer pathology using deep learning," *Nat. - Mach. Intel.*, p. 411–418, 2020.
- [2] Masoom A. Haider Farzad Khalvati Sunghwan Yoo, Isha Gujrathi, "Prostate cancer detection using deep convolutional neural networks," <https://www.nature.com/articles/s41598-019-55972-4>, 20 December 2019.
- [3] L. et al. Pantanowitz, "An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: a blinded clinical validation and deployment study," *Lancet Digit. Hea*, 2020.
- [4] Eirini Arvaniti, Kim Fricker, Michael Moret, Niels Rupp, Thomas Hermanns, Christian Fankhauser, Norbert Wey, Peter Wild, Jan Hendrik Rüschoff, and Manfred Claassen, "Replication data for: Automated gleason grading of prostate cancer tissue microarrays via deep learning.," 2018, Harvard Dataverse.
- [5] Bulten W., Bándi P., and Hoven J. et al, "Deep learning methods in transportation domain: a review," *Epithelium segmentation using deep learning in HE-stained prostate specimens with immunohistochemistry as reference standard.*, vol. 864, no. 9, 2019.
- [6] Liang Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, and Hartwig Adam, "Encoder-decoder with atrous separable convolution for semantic image segmentation," *Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation.*, 2018.
- [7] Nikhil Tomar, "Unet segmentation with pretrained mobilenetv2 as encoder," in *UNET Segmentation with Pre-trained MobileNetV2 as Encoder*, Jun. 2020.
- [8] David Telleza, Geert Litjens, Peter B., Wouter Bultena, John-Melle Bokhorsta, Francesco Ciompi, and Jeroen van der Laak, "Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology," *Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology.*, 2020.
- [9] Alvaro Fuentes, Dong Im, Sook Yoon, and Dong Park, "Spectral analysis of cnn for tomato disease identification," 05 2017, pp. 40–51.
- [10] David Tellez, Maschenka Balkenhol, Irene Otte-Höller, Rob Loo, Rob Vogels, Peter Bult, Carla Wauters, Willem Vreuls, Suzanne Mol, Nico Karssemeijer, Geert Litjens, Jeroen van der Laak, and Francesco Ciompi, "Whole-slide mitosis detection in he breast histology using phh3 as a reference to train distilled stain-invariant convolutional networks," *IEEE Transactions on Medical Imaging*, vol. PP, pp. 1–1, 03 2018.
- [11] Sohini Roychowdhury, "A machine learning engineer's tutorial to transfer learning for multi-class image segmentation using u-net.," 2021.
- [12] Taqi SA, Sami SA, Sami LB, and Zaki SA, "A review of artifacts in histopathology," *A review of artifacts in histopathology.*, 2018.