

Quantitative Spatial Analysis on Whole Slide Images Using U-Net

Sanghoon Lee^{1, *}, Yanjun Zhao², Mohamed Masoud³, Saeid Belkasim⁴

¹Department of Computer Sciences and Electrical Engineering, Marshall University, Huntington, United States of America

²Department of Computer Science, Troy University, Troy, United States of America

³Department of Neurology, Emory University, Atlanta, United States of America

⁴Department of Computer Science, Georgia State University, Atlanta, United States of America

Email address:

leesan@marshall.edu (S. Lee)

*Corresponding author

To cite this article:

Sanghoon Lee, Yanjun Zhao, Mohamed Masoud, Saeid Belkasim. Quantitative Spatial Analysis on Whole Slide Images Using U-Net. *Computational Biology and Bioinformatics*. Vol. 8, No. 2, 2020, pp. 90-96. doi: 10.11648/j.cbb.20200802.18

Received: October 22, 2020; **Accepted:** November 18, 2020; **Published:** December 4, 2020

Abstract: Advances in whole slide imaging technology have promoted a high use of digital slide images and generated a large volume of image data that is reliable and useful in determining treatment outcome. Recent technologies closely related to machine learning and deep learning algorithms have contributed to the success of digital histopathology by analyzing the digitized slide images providing quantitative information that are useful for faster turnaround times and effective treatment for the patient. The digital histopathological image analysis has received much attention due to its capability of mitigating the problem of the hand-crafted features. Features directly learned from raw data are trainable within the deep learning procedure and can be used for the histopathology image classification task. However, understanding the spatial context of cancer cells is still a challenging issue because of the heterogeneity of the tumor microenvironment which varies greatly, preventing successful diagnosis and leads to inappropriate therapeutic approaches for cancer patients. In this paper, we present a spatial analysis method for tumor microenvironment analysis using the U-Net architecture, a semantic segmentation deep-learning model, for a better understanding of the spatial relations between tissue types. We demonstrate the effectiveness of the U-Net architecture using a dataset created by an international crowdsourcing study. Moreover, we show that the quantitative estimates can be derived from the univariate spatial analysis.

Keywords: Cancer, Whole Slide Images, Spatial Analysis, U-Net, Machine Learning

1. Introduction

Whole slide imaging is an imaging technology that produces high-resolution digital images from microscopic slides [1]. Recent advances in whole slide imaging technology have enabled microscopes to digitize an entire glass slide within a few minutes, generating a large volume of image data that is highly reliable and useful in determining treatment outcome [2]. As a result, whole-slide image (WSI) analysis has become increasingly common in clinical and basic science investigations.

Machine learning (ML) techniques have been successfully employed in the detection, diagnosis, treatment, and prevention of human cancers, and promoted the transition of

modern pathology [3]. ML-based approaches for analyzing tissue samples obtained for the microscopic examination can provide satisfactory outcomes and validated clinical prediction, while specialized software tools with onboard ML can assure faster turnaround times in a quantitative analysis of WSIs [4-7]. The ML-based approaches help to provide definitive diagnoses and effective treatments in digital histopathology.

Recently, deep learning features using deep convolutional neural networks (DCNNs) have received much attention in the digital histopathological image analysis due to their capability of mitigating the problems of the hand-crafted features [8, 9]. While hand-crafted features rely on explicit algorithms requiring large efforts for novel datasets, features directly

learned from raw data are trainable within DCNNs and can be automatically transformed for use in the histopathological image analysis. Features trained by DCNNs provide better

accuracy through convolution and pooling layers, followed by a fully connected layer for the histopathology image classification task [10-13].

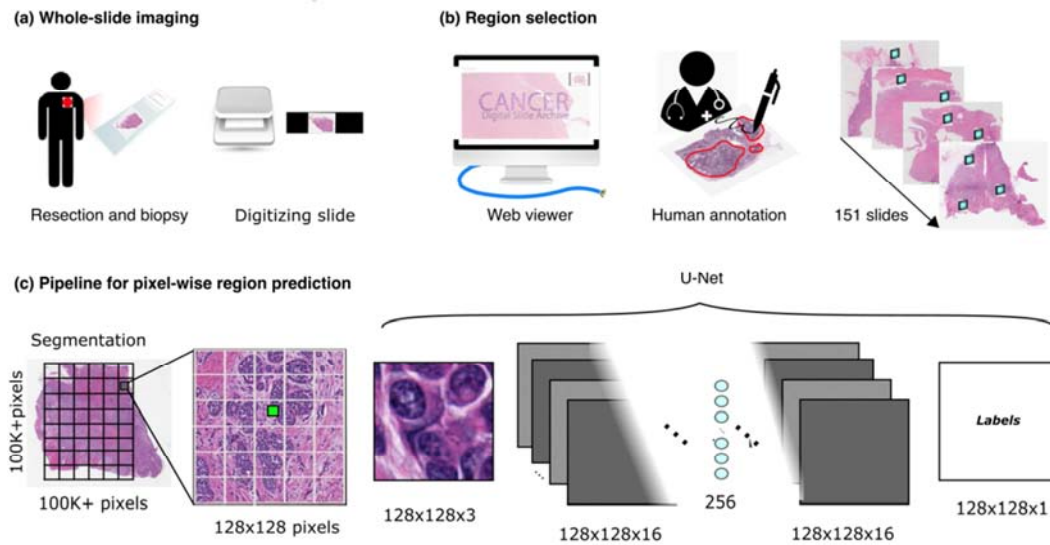


Figure 1. Overview of the U-Net based pixel-wise region prediction. (a) Whole slide imaging process to produce multi-resolution pyramid style images. (b) Region of interest selection. An international crowdsourcing study was adopted for the region selection. (c) Pipeline for pixel-wise region prediction. U-Net was used for training and the pixel-wise region prediction.

Involving domain experts in validating DCNN models is associated with human error or bias [14, 15]. Furthermore, labeling sufficient training samples to ensure accurate classification for hundreds-of-millions of histologic structures is a very time-consuming or laborious task, causing a significant delay in the diagnostic process [16, 17]. Key challenges in the WSI analysis using ML are 1) developing an automatic DCNN method to minimize human bias 2) enhancing the DCNN capabilities through spatial analysis.

Meanwhile, cancer researchers have investigated malignant tumors in microenvironmental conditions for effective cancer detection [18, 19]. With the advances in the whole slide imaging technologies, cancer researchers can now create numerous digitized high-resolution quickly, improving the cancer detection and diagnosis for precise treatment and functional monitoring of cancer. However, understanding the spatial context of cancer cells is still a challenging issue because of the heterogeneity of the tumor microenvironment that varies greatly, preventing successful diagnosis and leading to wrong therapeutic approaches for cancer patients.

In this paper, we present a spatial analysis method for tumor microenvironment analysis using the U-Net architecture, a semantic segmentation deep-learning model, for a better understanding of the spatial relations between tissue types. The main contributions of this paper are; 1) we present a U-Net based cancer detection method and 2) we propose a univariate spatial analysis method that can be used for an effective prognostic biomarker. The rest of the paper will be organized as below. Section 2 describes the detection procedure of tumor and tumor-infiltrating lymphocytes (TILs) using the U-Net. Section 3 presents our spatial analysis method of the detected regions in whole slide images. Section 4 shows the performance results of the tumor and TIL

detection and demonstrates the advantages of the quantitative estimates derived from the univariate spatial analysis.

2. U-Net Based Cancer Classification

2.1. Whole Slide Imaging

Tissue specimens obtained in the surgical pathology laboratory are usually mounted on a thin flat piece of glass. The glass slide is then digitized by the microscope slide scanners producing reliable and high-resolution images of tissue samples in a few minutes and potentially aiding biological scientists in interpreting histopathological whole slide images. Whole slide image analysis has been successfully exploited in the histological image interpretation throughout this procedure. The digitized whole slide images from the microscope slide scanners can be employed to obtain meaningful features prognostically related to tissue specificity can be stimulated by histological, solving the problem of observer variability and reducing the workload of pathologists [20, 21]. In this paper, we obtained hematoxylin and eosin-stained formalin-fixed paraffin-embedded (FFPE) sections of histopathology whole slide images of breast invasive carcinoma from The Cancer Genome Atlas at the Genomic Data (TCGA) [22].

2.2. Region Selection

Traditional approaches to select interesting regions in whole slide images have been done by pathologists involved in the project. However, they often cause the problem of observer variability and are not generalized well because it requires many experiences related to the tissue section and is difficult to share their knowledge with others. In order to avoid the

problems, we selected an annotated dataset created by an international crowdsourcing study [23]. In the study, the regions were selected by the combined effort of a medical doctor and a coordinator, and then approved by senior resident. In this paper, we selected the regions for the evaluation performance.

2.3. Cancer Classification Using U-Net

U-Net is a U-shaped network model that performs a semantic segmentation of biomedical images by using a Fully Convolutional Network (FCN) [24]. U-Net was firstly introduced to perform a cell segmentation in microscopic images by Ronneberger et al. [25]. The basic idea of U-Net is to provide a supplement network in each successive layer so that these layers contain the local information of high-resolution features by modifying the FCN. The U-Net architecture is similar to a traditional convolutional neural network but it captures context information in high-resolution layers that promote accurate localization. Table 1 shows the modified U-Net architecture used in this paper. Our method adopted U-Net architecture to detect cancerous regions in whole-slide images. The overall process of the U-Net based cancer prediction is shown in Figure 1.

Table 1. Modified U-Net architecture layers [17].

Layer (type)	Output Shape	Parameter
Input Layer	(128,128,3)	0
Lambda	(128,128,3)	0
Conv2D	(128,128,16)	448
Dropout	(128,128,16)	0
Conv2D	(128,128,16)	2320
MaxPooling2D	(64,64,16)	0
Conv2D	(64,64,32)	4640
Dropout	(64,64,32)	0
Conv2D	(64,64,32)	9248
MaxPooling2D	(32,32,32)	0
Conv2D	(32,32,64)	18496
Dropout	(32,32,64)	0
Conv2D	(32,32,64)	36982
MaxPooling2D	(16,16,64)	0
Conv2D	(16,16,128)	73856
Dropout	(16,16,128)	0
Conv2D	(16,16,128)	147584
MaxPooling2D	(8,8,128)	0
Conv2D	(8,8,256)	295168
Dropout	(8,8,256)	0
Conv2D	(8,8,256)	590080
Conv2DTrans	(16,16,128)	131200
Concatenate	(16,16,256)	0
Conv2D	(16,16,128)	295040
Dropout	(16,16,128)	0
Conv2D	(16,16,128)	147584
Conv2DTrans	(32,32,64)	32832
Concatenate	(32,32,128)	0
Conv2D	(32,32,64)	73792
Dropout	(32,32,64)	0

Layer (type)	Output Shape	Parameter
Conv2D	(32,32,64)	36928
Conv2DTrans	(64,64,32)	8224
Concatenate	(64,64,64)	0
Conv2D	(64,64,32)	18464
Dropout	(64,64,32)	0
Conv2D	(64,64,32)	9348
Conv2DTrans	(128,128,16)	2064
Concatenate	(128,128,32)	0
Conv2D	(128,128,16)	4624
Dropout	(128,128,16)	0
Conv2D	(128,128,16)	2320
Conv2D	(128,128,1)	17

3. Spatial Analysis Using U-Net

The completion of the semantic segmentation deep learning approach can leverage the spatial analysis for the quantified whole slide images. Common approaches on the spatial analysis methods merely generate a density map for each image with all features already predicted. However, it is not common to analyze the spatial characteristics in the entire regions of the whole slide images. The challenging issues are 1) how to generalize the pixel-wise predicted regions beyond uniform maps of whole slide images and 2) how to create a density map representation using the predicted regions. We use the whole slide regions predicted by the U-Net and consider the Getis-Ord-Gi* statistic to create a density map representation of spatial data [26, 27]. The Getis-Ord-Gi* statistic is a well-known hot spot analysis method that provides a statistically significant hotspot region with the relatedness of its neighbors. We compute the density of each square region by the Getis-Ord-Gi* statistic. The mathematical equation of the statistic Gi for the whole slide image is given by

$$G_i = \frac{\sum_{j=1}^n w_{i,j} x_j - W_i \bar{X}(i)}{S(i) \sqrt{((n-1)W_i^2 - (W_i)^2)/(n-2)}} \quad (1)$$

where x_j represents the attribute value for a square region j . $W_{i,j}$ represents the spatial weight between two regions (i, j) . n is the total number of square tessellation regions in a whole slide image. $W_i^2 = \sum_{j=1}^n w_{i,j}^2$, $W_i = \sum_{j=1}^n w_{i,j}$, $\bar{X}(i) = \sum_{j=1}^n x_j / (n-1)$, and $S(i) = \sqrt{(\sum_{j=1}^n x_j^2 / (n-1)) - (\bar{X}(i))^2}$ where $i \neq j$ and $w_{i,j} \neq 0$. Thus, the standardized Gi* for each whole slide image in the spatial analysis is given by

$$G_i^* = (\sum_{j=1}^n w_{i,j} x_j - W_i^* \bar{X}) / S \sqrt{(nW_i^{2*} - (W_i^*)^2)/(n-1)} \quad (2)$$

where $W_i^{2*} = \sum_{j=1}^n w_{i,j}^2$, $W_i^* = \sum_{j=1}^n w_{i,j}$, $\bar{X}(i) = \sum_{j=1}^n x_j / n$, and $S = \sqrt{(\sum_{j=1}^n x_j^2 / n) - (\bar{X})^2}$. Therefore, by equation (2), the Getis-Ord-Gi* statistic can be involved in univariate spatial analysis of whole slide images.

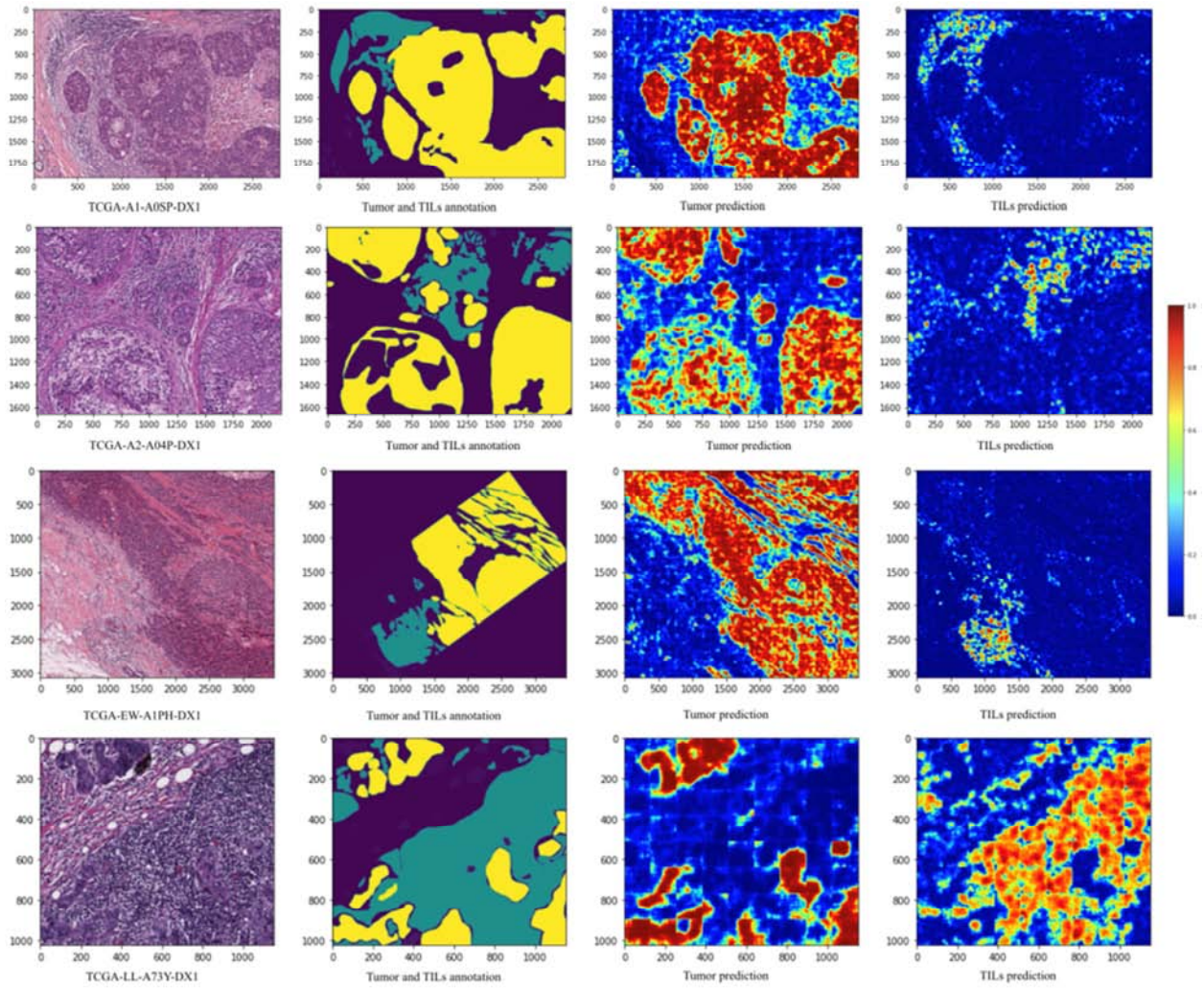
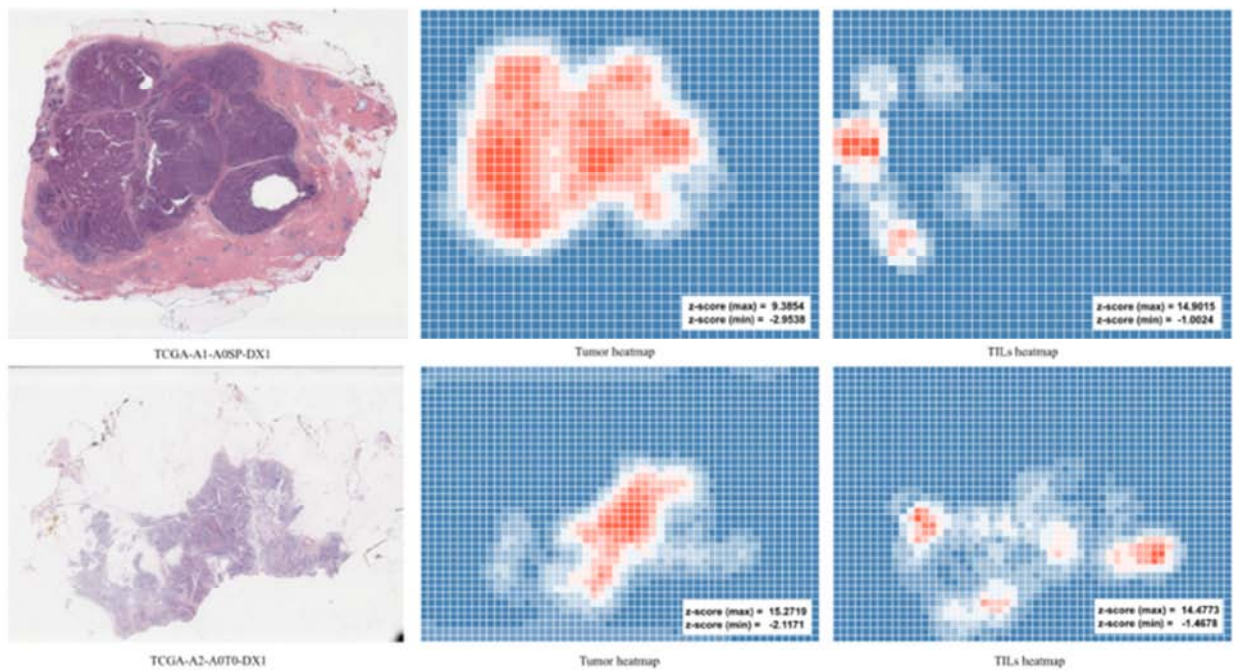


Figure 2. Prediction results on tumor and TILs. Left to right: slide name, annotated regions for tumor and TILs, tumor prediction, and TILs prediction.



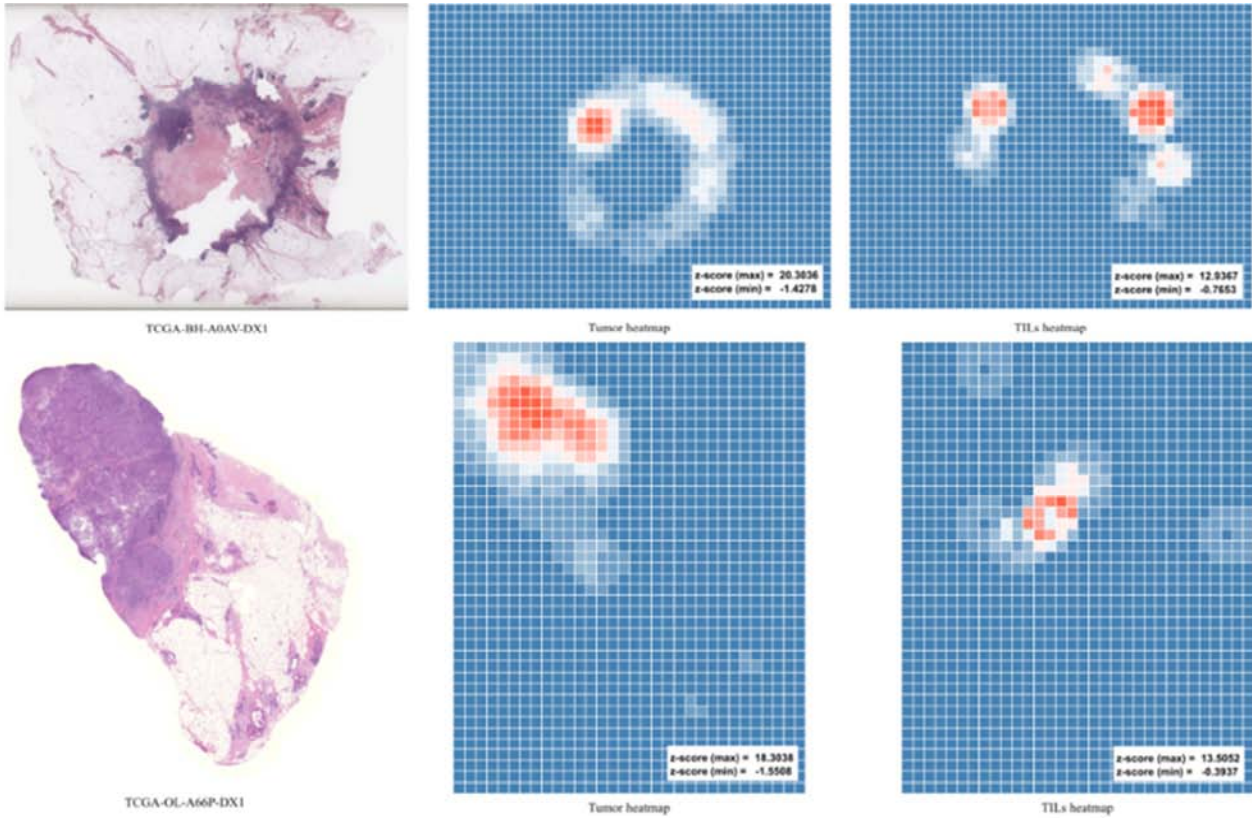


Figure 3. Heatmaps generated by Getis-Ord-Gi* Left to right: whole slide image, tumor heatmap, and TILs heatmap.

4. Performance Results

In this section, we demonstrate the effectiveness of our work. First of all, we perform the AUCROC measurement on

both Tumor and TILs predictions. Next, we perform spatial analysis using the prediction results. We used 151 hematoxylin and eosin-stained whole-slide images from [23] which are images of formalin-fixed paraffin-embedded tissues obtained from the Cancer Genome Atlas.

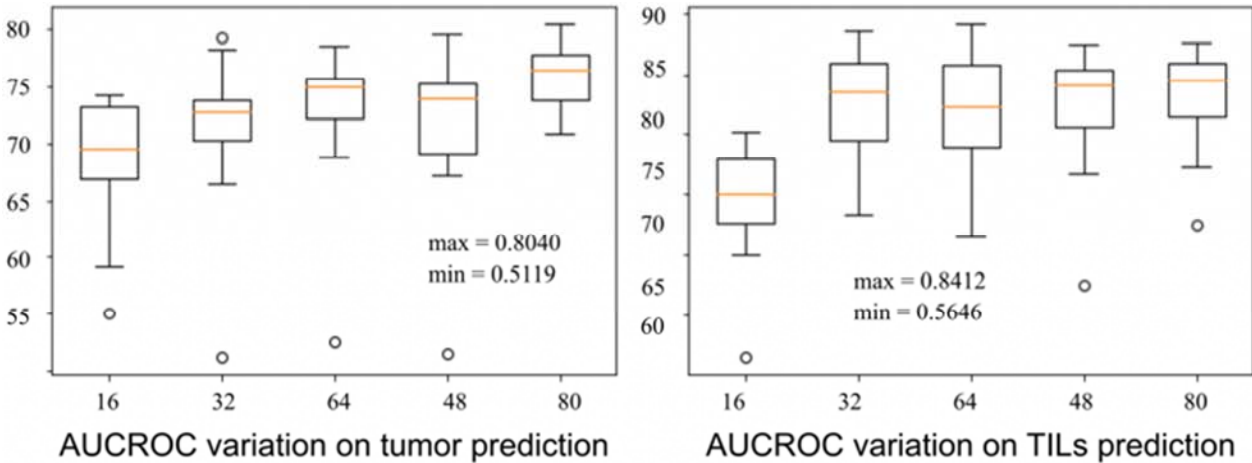


Figure 4. AUCROC results. Left: AUCROC variation on tumor prediction. Right: AUCROC variation on TILs prediction.

4.1. Results on Cancer Classification Using U-Net

The Area Under the Curve (AUC) and the Receiver Operating Characteristics (ROC) curve are well-known performance measurements for classification problems. In general, the ROC curve plots the True Positive Rate (TPR) with the False Positive Rate (FPR) in two-dimensional space such that the area under the curve plotted represents how the

classification model is suitable for distinguishing between classes. We selected 74 annotated hematoxylin and eosin-stained images for the training purpose and 42 annotated images for the test purpose among the 151 annotated images. Figure 2 shows the prediction results on tumor and TILs. In the second column, the yellow color represents the tumor annotation while the green color represents the TILs

annotation. Predictions are displayed as heatmaps in the third and the fourth columns. The heatmaps represent the prediction probabilities for tumor and TILs respectively. The AUCROC results are shown in Figure 4. Since the prediction results vary in the training set, we have investigated the prediction variation by selecting random samples in the training set and increasing the sample numbers. In Figure 4, the x-axis represents the number of square regions (128x128x3) selected for training and the y-axis represents the AUCROC results between 0 and 100. We found that the AUCROC results of the tumor prediction with a small number of training samples indicate the high AUCROC variance. However, in the case of the AUCROC results of the TILs prediction, the small number of training samples doesn't always indicate the high AUCROC variance.

4.2. Results on Spatial Analysis

In Section 4.1, we conducted a performance evaluation on tumor and TILs prediction rather than on the whole slide images but on the selected regions. In this section, we will show how our spatial analysis can be used on the whole-slide image predictions. In order to perform the whole slide image prediction, we created two models trained from the 151 images in the dataset [23]. Each model was created for the tumor and TILs predictions on the whole slide level respectively. After the prediction, the heatmap images were generated by the Getis-Ord-Gi* statistic. We used z-scores of the Getis-Ord-Gi* statistic to measure Spearman rank correlation between immune hot regions and immune cold regions. We used the proposed univariate spatial analysis method to distinguish two regions: immune hot and immune cold. The immune hot regions represent the large TILs prediction, while the immune cold regions represent the small TILs prediction. Figure 3 shows the spatial analysis results on the whole slide images. The second and third columns represent the immune hot/cold region corresponding to the tumor and TILs. The z-scores computed by Getis-Ord-Gi* statistic are displayed in the bottom right of each figure.

5. Conclusion

In this paper, we presented a pixel-wise region detection method on whole slide images using a semantic segmentation deep-learning model called U-Net. The U-Net architecture showed good results on the tumor and TILs classification measured in AUCROC. In addition to the prediction of the tumor and TILs, we proposed a spatial analysis method using Getis-Ord-Gi* for a better understanding of the spatial relations between tissue types. We demonstrated the effectiveness of the U-Net architecture using a dataset created by an international crowdsourcing study. Moreover, we showed that the quantitative estimates can be derived from the univariate spatial analysis.

Although we have presented the spatial analysis method on the whole slide images using U-Net, our method is limited to investigating TILs and didn't explore other types of cells in the tumor microenvironment. We plan to perform the spatial

analysis targeting tumor-stroma interactions that would improve treatment response.

References

- [1] Pantanowitz, L., Sinard, J. H., Henricks, W. H., Fatheree, L. A., Carter, A. B., Contis, L., Beckwith, B. A., Evans, A. J., Lal, A. and Parwani, A. V., 2013. Validating whole slide imaging for diagnostic purposes in pathology: guideline from the College of American Pathologists Pathology and Laboratory Quality Center. *Archives of Pathology and Laboratory Medicine*, 137 (12), pp. 1710-1722.
- [2] Abels, E. and Pantanowitz, L., 2017. Current state of the regulatory trajectory for whole slide imaging devices in the USA. *Journal of pathology informatics*, 8.
- [3] Komura, D. and Ishikawa, S., 2018. Machine learning methods for histopathological image analysis. *Computational and structural biotechnology journal*, 16, pp. 34-42.
- [4] Lee, S., Amgad, M., Chittajallu, D. R., McCormick, M., Pollack, B. P., Elfandy, H., Hussein, H., Gutman, D. A. and Cooper, L. A., 2020. HistomicsML2. 0: Fast interactive machine learning for whole slide imaging data. *arXiv preprint arXiv: 2001.11547*.
- [5] Nalisnik, M., Amgad, M., Lee, S., Halani, S. H., Vega, J. E. V., Brat, D. J., Gutman, D. A. and Cooper, L. A., 2017. Interactive phenotyping of large-scale histology imaging data with HistomicsML. *Scientific reports*, 7 (1), pp. 1-12.
- [6] Lee, S., Amgad, M., Masoud, M., Subramanian, R., Gutman, D. and Cooper, L., 2019, November. An Ensemble-based Active Learning for Breast Cancer Classification. In *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 2549-2553). IEEE.
- [7] Cruz-Roa, A., Gilmore, H., Basavanthally, A., Feldman, M., Ganesan, S., Shih, N. N., Tomaszewski, J., González, F. A. and Madabhushi, A., 2017. Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent. *Scientific reports*, 7, p. 46450.
- [8] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015 May 28; 521 (7553): 436-44. doi: 10.1038/nature14539. PMID: 26017442.
- [9] Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, Hulsbergen-van de Kaa C, Bult P, van Ginneken B, van der Laak J. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci Rep*. 2016 May 23; 6: 26286. doi: 10.1038/srep26286. PMID: 27212078; PMCID: PMC4876324.
- [10] Hou L, Samaras D, Kurc TM, Gao Y, Davis JE, Saltz JH. Patch-based Convolutional Neural Network for Whole Slide Tissue Image Classification. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2016 Jun-Jul; 2016: 2424-2433. doi: 10.1109/CVPR.2016.266. PMID: 27795661; PMCID: PMC5085270.
- [11] Sun W, Zheng B, Qian W. Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis. *Comput Biol Med*. 2017 Oct 1; 89: 530-539. doi: 10.1016/j.compbiomed.2017.04.006. Epub 2017 Apr 13. PMID: 28473055.

- [12] Murthy V, Hou L, Samaras D, Kurc TM, Saltz JH. Center-Focusing Multi-task CNN with Injected Features for Classification of Glioma Nuclear Images. *IEEE Winter Conf Appl Comput Vis*. 2017 Mar; 2017: 834-841. doi: 10.1109/WACV.2017.98. Epub 2017 May 15. PMID: 29881826; PMCID: PMC5988234.
- [13] Xu Y, Jia Z, Wang LB, Ai Y, Zhang F, Lai M, Chang EI. Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC Bioinformatics*. 2017 May 26; 18 (1): 281. doi: 10.1186/s12859-017-1685-x. PMID: 28549410; PMCID: PMC5446756.
- [14] Campanella, G., Hanna, M. G., Geneslaw, L., Miraflor, A., Silva, V. W. K., Busam, K. J., Brogi, E., Reuter, V. E., Klimstra, D. S. and Fuchs, T. J., 2019. Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nature medicine*, 25 (8), pp. 1301-1309.
- [15] Kothari, S., Phan, J. H., Stokes, T. H. and Wang, M. D., 2013. Pathology imaging informatics for quantitative analysis of whole-slide images. *Journal of the American Medical Informatics Association*, 20 (6), pp. 1099-1108.
- [16] Topol, E. J., 2019. High-performance medicine: the convergence of human and artificial intelligence. *Nature medicine*, 25 (1), pp. 44-56.
- [17] Vandenberghe, M. E., Scott, M. L., Scorer, P. W., Söderberg, M., Balcerzak, D. and Barker, C., 2017. Relevance of deep learning to facilitate the diagnosis of HER2 status in breast cancer. *Scientific reports*, 7, p. 45938.
- [18] Joyce JA. Therapeutic targeting of the tumor microenvironment. *Cancer cell*. 7 (6): 513-20, 2005.
- [19] Lorusso G, Ruegg C. The tumor microenvironment and its contribution to tumor evolution toward metastasis. *Histochemistry and cell biology*. 130 (6): 1091-103, 2008.
- [20] Yu KH, Zhang C, Berry GJ, Altman RB, Ré C, Rubin DL, Snyder M. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nat Commun*. 2016 Aug 16; 7: 12474. doi: 10.1038/ncomms12474. PMID: 27527408; PMCID: PMC4990706.
- [21] Ehteshami Bejnordi B, Balkenhol M, Litjens G, Holland R, Bult P, Karssemeijer N, van der Laak JA. Automated Detection of DCIS in Whole-Slide H&E Stained Breast Histopathology Images. *IEEE Trans Med Imaging*. 2016 Sep; 35 (9): 2141-2150. doi: 10.1109/TMI.2016.2550620. Epub 2016 Apr 5. PMID: 27076354.
- [22] Tomczak, K., Czerwińska, P., and Wiznerowicz, M.: 'The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge', *Contemp Oncol (Pozn)*, 2015, 19, (1A), pp. A68-A77.
- [23] Amgad, M., Elfandy, H., Hussein, H., Atteya, L. A., Elsebaie, M. A., Abo Elnasr, L. S., Sakr, R. A., Salem, H. S., Ismail, A. F., Saad, A. M. and Ahmed, J., 2019. Structured crowdsourcing enables convolutional segmentation of histology images. *Bioinformatics*, 35 (18), pp. 3461-3467.
- [24] Long, J., Shelhamer, E. and Darrell, T., 2015. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431-3440).
- [25] Ronneberger, O., Fischer, P. and Brox, T., 2015, October. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.
- [26] Jana M, Sar N. Modeling of hotspot detection using cluster outlier analysis and Getis-Ord G_i^* statistic of educational development in upper-primary level, India. *Modeling Earth Systems and Environment*. 2 (2): 60, 2016.
- [27] Junior GB, Paiva AC, Silva AC, de Oliveira AC. Classification of breast tissues using Getis-Ord statistics and support vector machines. *Intelligent Decision Technologies*. 3 (4): 197-205, 2009.