TIGER Grand Challenge

1st Aleksandar Dimitrievikj
Radboud University
aleksandar.dimitrievikj@ru.nl

2nd Simon Reichert
Radboud University
simon.reichert@ru.nl

3rd Richard Schwartzkopf
Radboud Univerity
richard.schwartzkopf@ru.nl

4th Neus Vegas Morales
Radboud University
neus.vegas@ru.nl

***Abstract*—The TIGER Challenge focuses on automating tumor-infiltrating lymphocyte (TIL) assessment in breast cancer. Our approach uses segmentation methods to identify key TIL components. Inspired by successful studies, particularly the nnU-Net approach, we integrate it into our strategy. We evaluate example codes and aim to enhance model efficiency with novel cell counting methods. Our goal is to refine boundary delineation and streamline development for improved breast cancer care.**

***Index Terms*—TIGER Challenge, breast cancer, tumor-infiltrating lymphocytes (TILs), automated assessment, nnU-Net, segmentation techniques, histopathology slides**

I. Introduction

In this project we are tackling the TIGER challenge. The TIGER (Tumor-Infiltrating Lymphocytes in Breast Cancer) Challenge aims to automate the assessment of TILs in histopathology slides of breast cancer. The Diagnostic Image Analysis Group (DIAG) of Radboud University Medical Center collaborates with the International Immuno-Oncology Biomarker Working Group to develop and validate AI algorithms to generate prognostically valuable TIL scores for Her2 positive and Triple Negative breast cancers. TIGER aims to improve clinical practices by enabling precise, automated TIL quantification, potentially improving treatment outcomes. While consisting of three subchallenges: 1) detecting lymphocytes and plasma cells, 2) segmenting invasive tumor and stroma, and 3) computing a TIL score per slide, our newly developed approach focused solely on the first subchallenge: detection of lymphocytes and plasma cells. These cell types are the principal components of tumor-infiltrating lymphocytes (TILs), which are essential for assessing the immune response in breast cancer histopathology slides. By utilizing state-of-the-art AI approaches to reliably detect these cells, we hope to contribute to the larger goal of improving automated TIL assessment. This detection approach utilizes image segmentation techniques to identify and quantify the different TIL components. Drawing inspiration from the successful nnU-Net segmentation method, we incorporate an adapted approach into our strategy. We evaluate and optimize the provided example code, with the goal of enhancing model detection. Our ultimate objective is to refine this boundary delineation process and streamline the development of automated lymphocyte and plasma cell detection tools, which could contribute to improved breast cancer diagnosis and treatment planning.

II. Related work

In our approach, we draw inspiration from recent studies in the field.

The U-Net approach, initially designed for image segmentation tasks, has demonstrated exceptional performance in automatically detecting lymphocytes in immunohistochemically stained tissue sections of breast, colon, and prostate cancer [1] . Impressively, it even outperformed human observers in this task [1]. This evidence highlights the effectiveness of integrating U-Net into our strategy for cell detection in breast cancer histopathology slides.

In recent research, a dataset of 171,166 annotated CD3 and CD8 cells was used to train deep learning algorithms for lymphocyte detection in histopathology images [2]. The algorithms were evaluated on tissue samples from breast, colon, and prostate cancer, demonstrating high performance in lymphocyte detection with U-Net achieving the highest F1-score. The results of the study also highlighted the importance of differentiating between nucleus and border labels. In this project, we incorporated these insights to enhance detection precision in breast cancer histopathology slides.

For optimally utilizing the U-Net architecture for cell detection, we are exploring nnU-Net, a self-adapting framework for biomedical image segmentation, to segment cells in histopathology slides [3]. Ultimately, we want to accomplish precise and efficient lymphocyte detection through initial cell segmentation using nnU-Net's automated configuration and training pipeline, thereby leading to improved automated assessment of tumor-infiltrating lymphocytes in breast cancer.

III. Experimental work

A. Dataset

This study's dataset contains training and test data comprising whole-slide images (WSIs) with detailed manual annotations. The training data is sourced from TCGA-BRCA, Radboud University Medical Center (RUMC), and Institut Jules Bordet (JB), while the test data includes private test sets for two leaderboards.

The dataset comprises selected and manually annotated whole-slide images of breast cancer, offering detailed insights into various tissue compartments and cell types present in the WSIs. These annotations cover invasive tumour regions, tumour-associated stroma, in-situ tumour lesions, healthy glands, necrotic tissue, inflamed stroma, and other tissue components.

Moreover, annotations of lymphocytes and plasma cells within the ROIs provide information for immune cell detection tasks. For our study, we focus on leveraging these manual annotations, particularly the regions of interest (ROIs), to train our segmentation model for lymphocytes, ensuring our model learns from relevant regions for accurate cell detection.

B. Segmentation approach

Inspired by the related literature, we are utilizing a segmentation approach for the present detection task. Namely, we use and adapt the approach mentioned in **Learning to detect lymphocytes in immunohistochemistry with deep learning** [2]. From the information of the bounding boxes we create segmentation masks such that for every bounding box in the training data we construct a mask for a cell with body and a membrane. This is also true for post-processing, where we use the segmentation mask produced by the U-Net architecture to create the final bounding boxes. In the next few sections we explain the approach in detail.

C. Mask Generation for Segmentation Model

To adapt the detection bounding boxes annotations for a segmentation problem, we implemented a process to create segmentation masks from the ROI images and corresponding bounding box information. The preprocessing pipeline involved the following steps.

First, we are using PNGs of the ROI images which have been extracted from the whole-slide images. Each ROI image had a corresponding entry in a COCO-formatted JSON file, which provided the bounding box coordinates for each cell within the image.

We generated segmentation masks for each ROI based on the information in the bounding box. We created the masks with three labels: the background was labelled as 0, the membrane of the cells was labelled as 1, and the body of the cells was labelled as 2. To standardize the cell annotations, each cell's mask was set to have a diameter of 4 mm, as referenced in previous studies [2]. This labelling allowed us to distinguish between different parts of the cells and the surrounding tissue.

Then, we used the generated segmentation masks to train our segmentation model, ensuring that the model could accurately identify and differentiate between the different components within the ROIs. This preprocessing step was needed for converting the bounding box detection annotations into a format suitable for training a segmentation algorithm, allowing us to perform lymphocyte segmentation within the annotated regions.

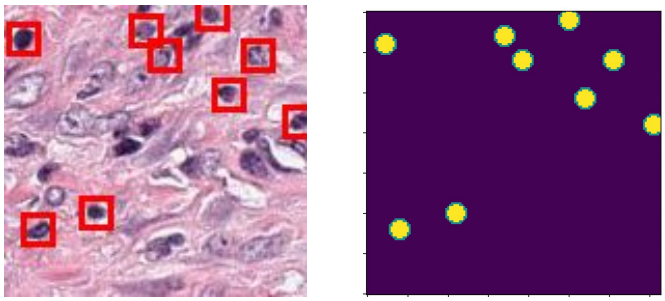

Fig. 1. Comparison of Original ROI Image (left) and Generated Segmentation Mask (right). The segmentation mask differentiates between the background (label 0), cell membrane (label 1), and cell body (label 2), facilitating precise lymphocyte segmentation.

D. Lymphocytes Segmentation Model

To segment lymphocytes within the annotated regions of interest (ROIs), we adopted nnU-Net, an advanced method for biomedical image segmentation tasks. nnU-Net extends the original U-Net architecture with additional features and optimizations specifically designed to address the challenges of working with medical imaging datasets [3].

The U-Net architecture, which was established in 2015, immediately became a standard in medical picture segmentation due to its simplicity and effectiveness. However, adapting U-Net to new situations involves a mutitude of different design choices, including architecture, preprocessing, training, and inference methods, considerably effecting performance and generalizability. The nnU-Net paradigm seeks to simplify network design by emphasizing critical features that impact performance and generalizability while eliminating superfluous complexities. A recent work [3] tested nnU-Net in the context of the Medical Segmentation Decathlon competition. The challenge tests segmentation performance across ten disciplines using heterogeneous datasets. nnU-Net outperformed across all datasets, earning high Dice scores and winning the challenge decisively. Notably, nnU-Net adapts to new datasets without operator intervention or fine-tuning, demonstrating its versatility and effectiveness.

IV. Training

As previously mentioned, the training process was conducted on the segmentation masks on the "Snellius" super-

computer. The dataset comprised 1878 samples. The default settings for nnU-Net are that it trains 5 different folds and then uses an ensemble of them to perform predictions. We used all the default setting for nnU-Net. During training, the best performing model is saved after each epoch. The weights that performed best on the validation set were saved and used during inference on the grand-challenge test set.

A. Segmentation Postprocessing and Bounding Box Generation

To integrate the segmentation results into the existing detection model, we implemented a postprocessing step that converts segmentation masks back into bounding boxes. This step is necessary for integrating the segmentation approach into our detection pipeline.

For this, we first compute the centre of mass for each connected component labelled as a lymphocyte in the predicted segmentation mask. By doing that, we calculate each lymphocyte's centroid, giving us the centre's x and y coordinates. Next, we extract the probability for each lymphocyte-labeled-pixel within this component from the nnU-Net output and calculate the average probability over all these pixels. Finally, we assign this average probability to the corresponding bounding box.

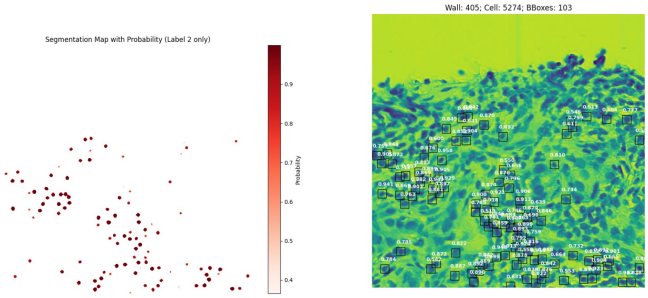

Fig. 2. Left: Output of nnU-Net, with each pixel colored according to its probability for label 2 (lymphocyte). Right: Original image with bounding boxes applied, showing the corresponding calculated probabilities.

V. Results

A. Segmentation results

For the challenge we only had access to the training data. Our original idea was to train the model with the whole training dataset and obtain test results from Grand Challenge. However, we had some problems with submitting the algorithm to Grand Challenge as discussed in the following section. After we realized that we could not obtain results, our next idea was to train the model with part of the training dataset and use the rest for testing locally. Unfortunately, due to time constraints and limited computing resources, we were not able to retrain and test the model.

Nonetheless, we achieved some promising results. During the training process, the segmentation achieved a validation mean Dice coefficient of 0.97, indicating a high level of accuracy in segmenting lymphocytes. This metric reflects the model's ability to closely match the predicted lymphocyte segmentations with the converted ground truth annotations.

It´s important to note that our validation metric is derived from the training dataset itself, rather than an independent test set. As a result, the Dice coefficient of 0.97 might not fully represent the model's performance on unseen data, as it could be overly optimistic. In an ideal scenario, we would have performed a proper data split and retrained the model to obtain a more reliable measure of its generalizability.

Despite these constraints, the initial results are promising and suggest that nnUNet has the potential to effectively segment lymphocytes in histopathology slides. As future work, we would recommend to address the data split issue and perform more comprehensive evaluations to ensure the robustness and accuracy of our model in diverse and unseen datasets. This will help in establishing a more representative performance metric and further validate the efficacy of nnUNet for this critical task in computational pathology.

B. Grand Challenge results

Despite successful local implementation, we encountered several initial obstacles in getting the algorithm to work on the Grand Challenge platform which partially hindered the development process. One of the main challenges was related to Docker; the difficulties we found in configuring the Docker environment and running it on the Grand Challange platform resulted in our algorithms failing to run as expected. However, despite these challenges, we successfully submitted our algorithm.

However, we still need to determine whether the algorithm will produce any sensible output because the challenge admins are still reviewing the evaluation before the results are made visible. Hence, our model's actual performance on the Grand Challenge validation remain to be confirmed.

VI. Conclusion and Discussion

Our initial integration of nnU-Net for lymphocyte segmentation in histopathology slides has shown considerable promise, with a validation mean Dice coefficient of 0.97 achieved on the Snellius supercomputer. Despite these encouraging results, the absence of a clear train-test split is a notable limitation that must be addressed to assess the model's performance accurately.

Future work might focus on implementing a proper data split to obtain a more reliable evaluation of the nnU-Net model. This will involve retraining and validating the model on an independent test set to ensure its generalizability and robustness. By doing so, we aim to establish a more representative performance metric that reflects the model's potential in clinical applications.

The promising initial results underscore the potential of nnU-Net for effective lymphocyte segmentation, paving the way for more accurate and automated assessment of tumour-infiltrating lymphocytes in various types of cancer and potentially other histopathological conditions.

This work lays the foundation for further advancements in computational pathology, contributing to improved diagnostic accuracy and personalized treatment strategies.

VII. Github Repository

nnUnet-Training+Inference.

Acknowledgment

We want to thank Francesco Ciompi for the opportunity to participate in this challenging task by implementing a completely new strategy to the baseline model with such promising results. We also thank him for the support provided through his participation in our weekly meetings, which has been valuable to the success of this project.

References

[1] Van Rijthoven, M., Aswolinskiy, W., Tessier, L., Balkenhol, M., Bogaerts, J., Van Der Laak, J., Salgado, R., & Ciompi, F. (2022). TIGER training dataset (ROI-level annotations of WSIROIS subset) (1.0) [dataset]. Zenodo. https://doi.org/10.5281/ZENODO.6014420

[2] Swiderska-Chadaj, Z., Pinckaers, H., van Rijthoven, M., Balkenhol, M., Melnikova, M., Geessink, O., Manson, Q., Sherman, M., Polonia, A., Parry, J., Abubakar, M., Litjens, G., van der Laak, J., & Ciompi, F. (2019). Learning to detect lymphocytes in immunohistochemistry with deep learning. In Medical Image Analysis (Vol. 58, p. 101547). Elsevier BV. https://doi.org/10.1016/j.media.2019.101547

[3] Isensee, F. et al. (2019). Abstract: nnU-Net: Self-adapting Framework for U-Net-Based Medical Image Segmentation. In: Handels, H., Deserno, T., Maier, A., Maier-Hein, K., Palm, C., Tolxdorff, T. (eds) Bildverarbeitung für die Medizin 2019. Informatik aktuell. Springer Vieweg, Wiesbaden. https://doi.org/10.1007/978-3-658-25326-4_7