International Journal of Environmental Sciences
ISSN: 2229-7359
Vol. 11 No. 17s, 2025
https://www.theaspd.com/ijes.php

Detection Of Tumor-Infiltrating Lymphocytes In Digital Oral Pathology Based On Computer Vision

Jing Yuanwei[1], Anand Ramanathan[2], Anis Salwa Mohd Khairuddin[3*], Aina Najwa Mohd Khairuddin[4], Uswah Khairuddin[5]

[1] Department of Electrical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Malaysia
[2] Department of Oral & Maxillofacial Clinical Sciences, Faculty of Dentistry, Universiti Malaya, Kuala Lumpur, Malaysia; Malaysia. Department of Oral Biology and Oral Pathology, Saveetha Dental College and Hospitals, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India,
[3] Department of Electrical Engineering, Faculty of Engineering, Universiti Malaya, Kuala Lumpur, Malaysia
[4] Department of Community Oral Health & Clinical Prevention, Faculty of Dentistry, Universiti Malaya, Kuala Lumpur, Malaysia
[5] Centre for Artificial Intelligence and Robotics, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
1st 23069190@siswa.um.edu.my; 2nd drranand@um.edu.my; 3rd anissalwa@um.edu.my;
4th ainanajwa@um.edu.my; 5th uswah.kl@utm.my
*Corresponding Author: anissalwa@um.edu.my

Abstract—*Oral cancer, with its rising incidence and mortality rates globally, represents a significant public health concern, making early diagnosis and effective treatment vital. Generally, Tumor-Infiltrating Lymphocytes (TILs) are critical immune cells that infiltrate tumors and are widely recognized as important biomarkers of tumor behavior and patient prognosis in oral cancer. However, despite their potential as a prognostic tool, the manual assessment of TILs remains a complex, labor-intensive process that is highly dependent on subjective judgment of pathologists. Considering these challenges, there has been a growing interest in utilizing computer vision algorithms to automate the quantification of TILs. Recent studies suggest that computer vision approach can offer a more objective and reproducible means of evaluating TILs, with the potential to provide real-time insights that could guide clinical decision-making. This study proposes a computer vision framework for the detection and analysis TILs in oral pathology, incorporating convolutional neural networks (CNNs). Firstly, the whole slide images (WSIs) were segmented into smaller patches using DenseNet169 to classify the tumor and non-tumor region. Then, U-Net and K-Net methods were applied for the TILs detection. The findings underscore the potential of automated TILs analysis to reduce diagnostic workload and improve prognostic assessments in oral cancer.*

Keywords— *CNN, Deep Learning, Objective Detection, Image***.**

I. INTRODUCTION

In recent years, significant advancements in digital pathology, coupled with the rapid evolution of artificial intelligence (AI) and deep learning (DL) algorithms, have revolutionized the landscape of cancer research, particularly in the analysis and quantification of tumor-infiltrating lymphocytes (TILs). TILs are critical immune cells that infiltrate tumors and are widely recognized as important biomarkers of tumor behavior and patient prognosis. Their presence, density, and spatial distribution within the tumor microenvironment have been correlated with various clinical outcomes, including disease progression, response to treatment, and overall survival across multiple cancer types. However, despite their potential as a prognostic tool, the manual assessment of TILs remains a complex, labor-intensive process that is highly dependent on the subjective judgment of pathologists. This variability in interpretation, referred to as inter-observer variability, is a significant barrier to the widespread adoption of TILs as a standard clinical biomarker[1].

Currently, there has been a growing interest in utilizing computational pathology, particularly image analysis and AI-driven algorithms, to automate and standardize the quantification of TILs. Recent studies suggest that computational methods, including deep learning, can offer a more objective and reproducible means of evaluating TILs, with the potential to provide real-time insights that could guide clinical decision-making. These AI-driven approaches can be trained to recognize patterns in histological images, enabling the precise identification and quantification of TILs within the tumor microenvironment[2].

Incorporating deep learning into the analysis of TILs could significantly enhance the objectivity and reproducibility of TIL quantification, which in turn could provide critical insights into the tumor microenvironment and its role in disease progression[3]. This study aims to investigate the application of computer vision approach to detect the spatial distribution of TILs in oral cancers, with the ultimate goal of predicting patients' prognoses. By developing an automated detection of TILs, this research seeks to improve the accuracy, consistency, and efficiency of TIL quantification, thereby reducing pathologists' workload and enhancing the diagnostic process.

II. PREVIOUS WORKS

A. Application of Deep Learning for Medical Image Analysis

Medical image classification is one of the primary applications of deep learning in the medical imaging domain. By training deep learning models, researchers can automatically classify medical images (such as X-rays, CT scans, and MRIs) as normal or abnormal[4]. Deep learning has also achieved significant results in early screening for various medical applications[4, 5]. Deep learning models, especially U-Net and its variants, have become the standard tools for medical image segmentation. In tumor segmentation tasks, U-Net and its improved versions, such as ResNet and Attention U-Net, have been widely used. These models effectively extract tumor region boundaries by fusing multi-level features, which has been applied to the automatic segmentation of brain, lung, and breast tumors[6]. Studies show that deep learning models not only improve accuracy in tumor segmentation tasks but also reduce the time cost of manual segmentation. Another important application of deep learning is the automatic detection and diagnosis of diseases. Studies have shown that deep learning models can accurately detect early lesions by analyzing fundus images, aiding doctors in developing timely treatment plans. Deep learning also demonstrates strong capabilities in the automatic detection of skin lesions[7]. By training models to identify different types of skin lesions (such as melanoma and basal cell carcinoma), researchers can achieve rapid screening of lesions.

B. TILs Detection Using Computer Vision Approach

Tumor-Infiltrating Lymphocytes (TILs) are immune cells that migrate from the blood or lymphatic system into the tumor microenvironment. TILs play an important role in the tumor microenvironment, and studies have shown that their quantity and subtype may be closely related to patient prognosis. A high density of TILs is typically associated with a favorable clinical prognosis, particularly in certain types of cancer, such as melanoma, breast cancer, and non-small cell lung cancer. The density and distribution of TILs can be detected through techniques like pathology slides and immunohistochemistry, and they are often used to evaluate the immune microenvironment of tumors[8]. In digital pathology, automated detection and quantification of TILs is possible with the help of image analysis and deep learning techniques, which not only improves the efficiency of assessment but also helps to identify potential prognostic biomarkers[9-11].

III. METHODOLOGY

A. Image Data Acquisition

The ethics approval for this study was obtained from the medical ethical committee, Faculty of Dentistry, Universiti Malaya [DF OS 2420/0095 (L)]. A total of 11 Whole Slide Images (WSIs) were divided into 512×512 patches at 40× magnification. Then, these patches were classified into tumor and non-tumor regions using a ResNet50 model, with the results visualized as a heat map. Next, tumor patches were further analyzed using segmentation models, such as FCN and K-Net, to detect and quantify TILs at the cellular level. The architecture of the approach is shown in Fig. 1.

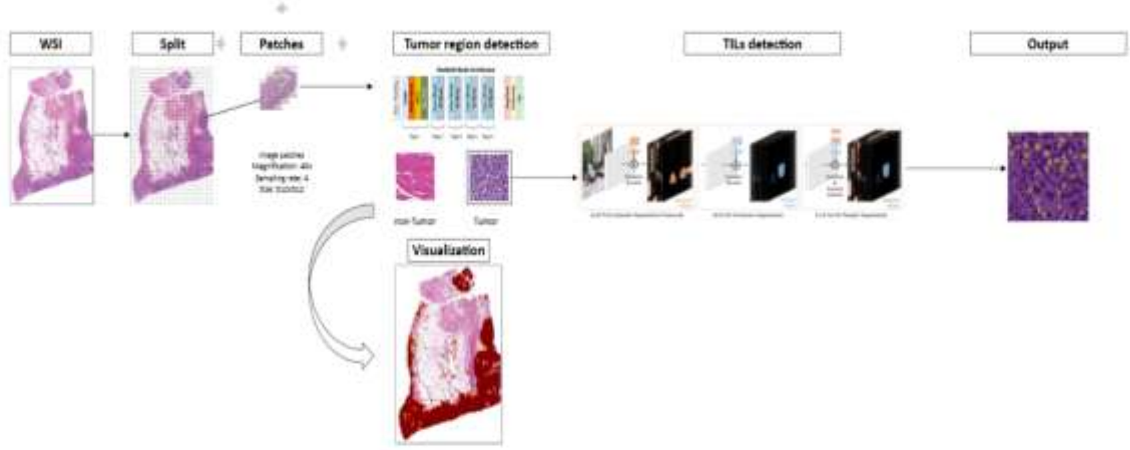


Fig. 1 Overview of the proposed methodology

Due to the extremely high scanning resolution of digital pathology images (e.g., 0.5 microns per pixel), each pathology image file can reach several gigabytes, or even tens of gigabytes in size, making it impractical to use the entire WSI directly for model training. Therefore, it was necessary to segment WSI images into smaller sub-images that were suitable for machine learning and deep learning processing. In practice, one of the most common methods is patch-based processing. This approach divides the WSI image into multiple small patches, typically 256x256 or 512x512 pixels, and processes each patch independently.

By adopting the QuPath software[12], we manually labeled the original WSI as tumor and non-tumor regions, setting the magnification to 40x and the sampling rate to 4. The training set, validation set, and test set were divided in the ratio of 8:1:1. These images were used as the tumour region detection dataset to train the model using the image classification algorithm. The TILs in the images were labelled one by one under the guidance of a professional pathologist (AR), with each image containing from a few to numerous TILs.

B. Tumor Regional Detection

At the macroscopic level, the first task of this study was to distinguish normal tissue regions from tumour tissue regions in WSIs. Due to the ultra-high resolution of WSI images, direct processing of the whole image faced a huge computational burden, so we adopted a classification method based on image chunking (patch-based). For the image classification task, we used Convolutional Neural Networks (CNNs) as the base model. Specifically, we selected ResNet and DenseNet as candidate models, which have shown excellent performance in handling complex image data.

Residual Network (ResNet) is an important neural network architecture in the field of deep learning. ResNet-50 is a classic model within the ResNet architecture, consisting of 50 layers and known for its deep structure and efficient feature extraction capabilities. Compared to other shallow networks, ResNet-50 is better at extracting complex features, making it particularly outstanding in tasks such as image classification and object detection. ResNet-50 employs a bottleneck structure within its residual modules, which maintains the network's depth while significantly reducing the number of parameters, allowing it to train more efficiently on large datasets. On the other hand, DenseNet (Dense Convolutional Network) is a deep convolutional neural network architecture proposed. DenseNet-121 is a relatively lightweight model architecture within the DenseNet family, consisting of 121 layers. Its design retains the core dense connectivity structure of DenseNet while optimizing the balance between parameter count and computational complexity. The architecture of DenseNet-121 includes 4 dense blocks, each composed of multiple convolutional layers, which are directly connected to adjacent layers through dense connections. Additionally, DenseNet-121 employs transition layers between each dense block to reduce the dimensionality of the feature maps, thereby controlling computational complexity.

TABLE I DETECTION PERFORMANCE OF TUMOR USING DEEP LEARNING MODELS

Model	**Epochs**	**Precision(%)**	**Recall(%)**	**F1-Score(%)**	**Accuracy(%)**
ResNet18	30	97.93	**97.93**	**97.93**	**97.93**
	50	97.44	97.41	97.40	97.41
ResNet50	30	97.93	**97.93**	**97.93**	**97.93**
	50	96.89	96.89	96.88	96.89
DenseNet121	30	97.42	97.41	97.40	97.41
	50	97.44	97.41	97.40	97.41
DenseNet169	30	**97.96**	**97.93**	**97.93**	**97.93**
	50	96.94	96.89	96.87	96.89

C. TILs Detection Using Deep Learning

After the detection of tumor regions is completed, the focus shifts to the detection and quantification of TILs within those tumor areas. TILs are an important indicator of immune response, and their quantity and distribution have significant clinical implications for the prognosis of cancer patients. To this end, we have developed an automated detection and quantification method based on deep learning. Unlike the task of detecting tumor regions, the detection and quantification of TILs falls under the category of instance detection tasks, which involves more refined image feature extraction. We selected K-Net[13] and U-Net[14] architectures as the convolutional neural networks for this task.

K-Net is a deep learning-based convolutional neural network architecture designed to improve performance in image segmentation and other visual tasks. The design of K-Net fully considers the diversity of features and the integration of contextual information. By introducing a strategy that fuses local and global features, the model was able to capture both detailed and global information at different levels. This architecture is particularly well-suited for tasks such as medical image analysis and semantic segmentation, effectively handling complex image structures and subtle features. The core idea of K-Net is to enhance feature representation through the fusion of multiple feature layers.

K-Net has significant advantages over traditional convolutional neural networks in the following aspects: (1) Feature Diversity: By combining local and global features, K-Net can capture multi-scale features, enhancing the model's understanding and performance on complex images. (2) Context Awareness: K-Net strengthens the model's context awareness in fine-grained tasks through the integration of global features, allowing it to effectively utilize background information in segmentation and classification tasks. (3) Efficient Training: The structural design of K-Net makes the network more stable during training, enabling faster convergence, which is suitable for training on large-scale datasets.

On the other hand, U-Net is a convolutional neural network (CNN) architecture specifically designed for biomedical image segmentation. Its design aims to address the challenge of image segmentation with limited samples, making it particularly suitable for processing medical images such as cells, tissues, and organs. U-Net achieves efficient feature extraction and reconstruction through its unique U-shaped structure, allowing it to capture contextual information while preserving detailed features. The U-Net architecture consists of two parts: the encoder (contracting path) and the decoder (expanding path).

The U-Net model has the following advantages in medical image segmentation: (1) Efficient feature utilization: Through skip connections, U-Net effectively combines features from different layers, enhancing its ability to capture details. (2) Small sample learning capability: U-Net is designed to tackle the problem of limited samples and performs well on small-scale datasets, making it suitable for medical image analysis. (3) Scalability: U-Net has good scalability in its structure, allowing users to increase the number of layers, adjust convolution kernel sizes, and more to adapt to various image segmentation tasks.

IV. RESULTS AND DISCUSSION

In this study, the aim was to evaluate the performance of the proposed model in the histopathology image classification task. Several evaluation metrics were applied, including accuracy, precision, recall and F1 Score.

A. Performance Evaluation on Tumor Detection

Based on the test results for different models in classifying tumor images, four deep learning models—ResNet18, ResNet50, DenseNet121, and DenseNet169— were evaluated to distinguish between tumor and non-tumor regions. The table presents their performance metrics (precision, recall, F1-score, and accuracy) for 30 and 50 training epochs. The detection performance of tumors by these models is depicted in Table.1.

At 30 epochs, ResNet18 achieved a precision, recall, F1-score, and accuracy of 97.93%. After 50 epochs, the model showed slight decrease in precision and accuracy, reaching 97.44% for these metrics, with a recall and F1-score of 97.41% and 97.4%, respectively. With 30 epochs, ResNet50 achieved 97.93% across all metrics, identical to ResNet18's performance at 30 epochs. However, after 50 epochs, its performance dropped slightly, with all metrics at 96.89%. DenseNet121 demonstrated consistent performance, achieving a precision of 97.42% and accuracy of 97.41% at 30 epochs, and similar scores after 50 epochs, with slight variations in F1-score. DenseNet169 performed best at 30 epochs, with all metrics at 97.93%. At 50 epochs, the performance metrics dropped slightly to 96.94% precision, 96.89% recall, 96.87% F1-score, and 96.89% accuracy.

Overall, DenseNet169 achieved the best performance at 30 epochs, while ResNet18 and DenseNet121 showed consistent results across both epoch settings. ResNet50 showed a decrease in performance at 50 epochs, and DenseNet models generally performed better at 30 epochs compared to 50 epochs. These results suggest that DenseNet architectures, particularly DenseNet169, may be more effective in tumor classification for this dataset at a lower number of training epochs. Fig.2 shows the train loss of these models at 30 and 50 epochs respectively.

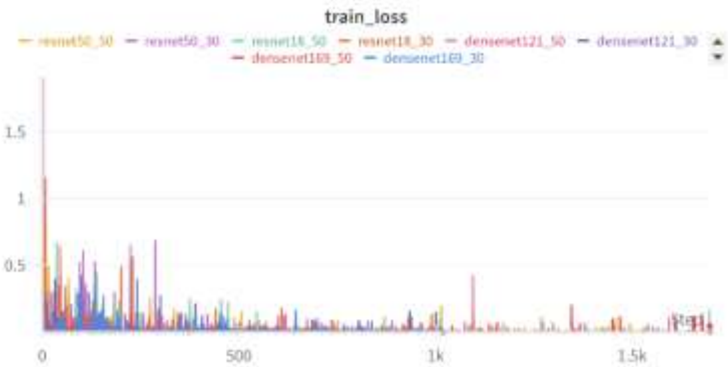
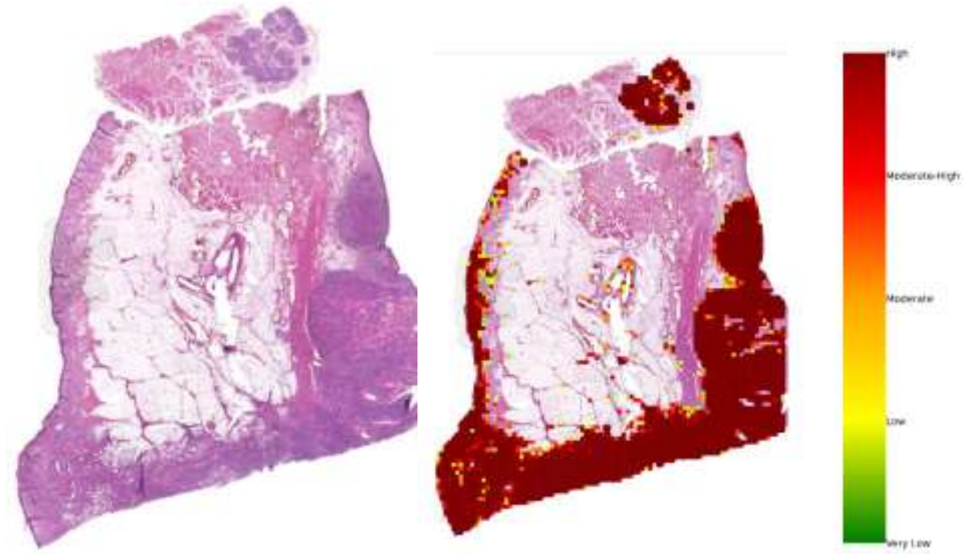


Fig. 2 Train loss for different models at 30 and 50 epochs

The trained model was applied to predict the WSIs, enabling the detection of tumor regions within the WSI. The complete WSI is segmented using an image segmentation algorithm, and the resulting sub-images were then predicted using the model. To facilitate observation, the prediction results are mapped back to the original positions of the sub-images within the WSI, and a heat map was generated to visualize the predictions based on their probabilities. The trained model was applied to predict the WSIs, enabling the detection of tumor regions within the WSI. The complete WSI is segmented using an image segmentation algorithm, and the resulting sub-images were then predicted using the model. To facilitate observation, the prediction results were mapped back to the original positions of the sub-images within the WSI, and a heat map was generated to visualize the predictions based on their probabilities. Fig.3 presents the probabilities of different religions of WSI being detected as tumor area through a heat map.



(a)WSI (b)Tumor Detection

Fig. 3 Tumor detection. (a) Original WSI, (b) Tumor regional detection

B. Performance Evaluation on TILs Detection

We trained the U-Net and K-Net models for a total of 20,000 epochs, to detect TILs. Based on the test results, U-Net showed a precision of 69.73%, recall was 80.12%, F1-score was 74.56%, and accuracy was 80.12%. For the TILs class, K-Net achieved a precision of 78.83%, recall of 77.18%, F1-score of 78.0%, and accuracy of 77.18%. K-Net showed some improvement in TILs detection, especially in precision and F1-score. In summary, for TILs detection, K-Net showed better performance than U-Net, with higher precision and F1-score, suggesting a slight advantage in detecting TILs. Table.2 shows the detection performance of TILs between the two models.

TABLE II

Detection Performance of TILs

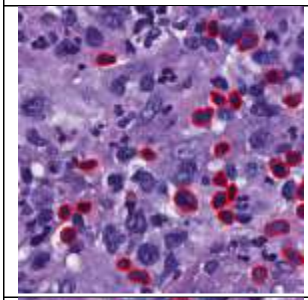
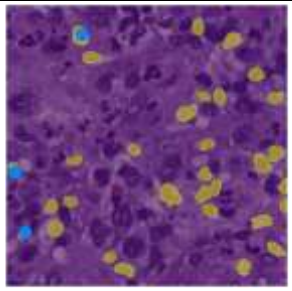
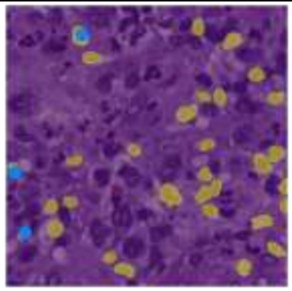
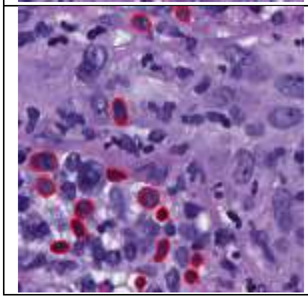
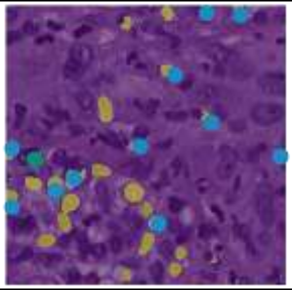
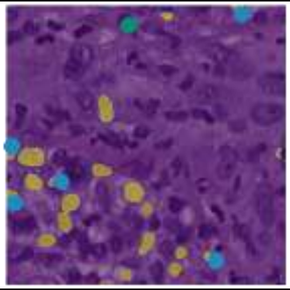
Model	**Precision(%)**	**Recall(%)**	**F1-Score(%)**	**Accuracy(%)**
U-Net	69.73	**80.12**	74.56	**80.12**
K-Net	**78.83**	77.18	**78.00**	77.18

In contrast, the K-Net model demonstrates superior detection performance in terms of precision and F1 score. Its ability to maintain a balanced perspective, considering both the broader context and necessary details, results in more reliable predictions. This distinction highlights the importance of model architecture in achieving effective detection in complex imaging tasks, such as those encountered in medical applications particularly in histopathological images. Overall, the results indicate that the K-Net model was better suited for this analysis, providing a more robust solution for detecting the targeted features within the images.

The detection performance of the two types of models is shown in Table 3. The yellow regions were the correct detected areas, whereas the blue circle represents the incorrect detected areas, and the green circle represents the missed detected areas. The detection performance of U-Net and K-Net still have a certain gap compared with the manual annotation by professional pathologist. However, compared with

U-Net, K-Net has fewer cases of missed detections and false detections.

TABLE III COMPARISON OF TILS DETECTION PERFORMANCE

Manual Annotation	**U-Net**	**K-Net**
		
		

In TILs detection, precision reflects the model's ability to avoid false positives in non-TIL regions. However, due to the complex morphology of TILs, the model still faces challenges in distinguishing between TILs and the background. Low precision can result in many false positives, adversely impacting further pathological analysis. Higher recall ensures that the model captures as many TILs regions as possible, supporting comprehensive analyses of immune responses within the tumor microenvironment.

V. CONCLUSIONS

This study compared the performance of ResNet and DenseNet models in tumor region classification and found that DenseNet169 achieved the best precision, recall, F1-score, and accuracy at 30 training epochs. Compared to other models, DenseNet169 demonstrated a balanced capability to capture details in classification and a high recognition rate for tumor regions. In TILs detection, this study compared the performance of U-Net and K-Net, with K-Net demonstrating significantly higher precision and F1-score in detecting tumor-infiltrating lymphocytes (TILs). K-Net maintained a balanced approach, accurately detecting key features while considering the overall image context, indicating its advantage in complex medical imaging tasks. For future works, multi-scale feature extraction or integrated learning methods can be proposed to better capture TILs of different sizes and morphologies to improve the comprehensiveness and accuracy of detection. More advanced deep learning algorithms, such as Transformer-based models and self-attention mechanisms, may be further explored to improve the detection and classification accuracy of the models. Model compression and acceleration techniques are also worthy of in-depth study to improve computational efficiency in practical applications.

ACKNOWLEDGEMENT

This research work is supported by Universiti Malaya, Malaysia under project number AIM-B05-2025. In addition, this project is also supported by Universiti Teknologi Malaysia under Uni Technologies Sdn Bhd with Vot 2031.

REFERENCES

[1] K. Skok, K. Bräutigam, Tumor infiltrating lymphocytes (TILs)–Pathologia, quo vadis?–A global survey, Pathology-Research and Practice, 266 (2025) 155775.

[2] S. Morales, K. Engan, V. Naranjo, Artificial intelligence in computational pathology–challenges and future directions, Digital Signal Processing, 119 (2021) 103196.

[3] K. Brummel, A.L. Eerkens, M. de Bruyn, H.W. Nijman, Tumour-infiltrating lymphocytes: from prognosis to treatment selection, British Journal of Cancer, 128 (2023) 451-458.

[4] L.C. Schudlo, Y.T. Xie, K. Small, B. Graf, A Novel CNN plus LSTM Approach to Classify Frontal Chest X-Rays for Spine Fractures, in: Conference on Medical Imaging - Computer-Aided Diagnosis, Vol. 12033, Electr Network, 2022.

[5] S. Padmakala, S. Revathy, K. Vijayalakshmi, M. Mathankumar, CNN supported automated recognition of Covid-19 infection in chest X-ray images, in: International Conference on Thermal Analysis and Energy Systems (ICTAES), Vol. 66, Coimbatore, INDIA, 2021, pp. 1201-1210.

[6] S. Gohil, A. Lad, Kidney and Kidney Tumor Segmentation Using Spatial and Channel Attention Enhanced U-Net, in: 2nd International Challenge Conference on Kidney and Kidney Tumor Segmentation (KiTS), Vol. 13168, Electr Network, 2021, pp. 151-157.

[7] G. Renith, A. Senthilselvi, Automated Skin Cancer Diagnosis and Localization Using Deep Reinforcement Learning, Iete Journal of Research, 70 (2024) 3631-3645.

[8] R. Salgado, C. Denkert, S. Demaria, N. Sirtaine, F. Klauschen, G. Pruneri, S. Wienert, G. Van den Eynden, F.L. Baehner, F. Pénault-Llorca, The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014, Annals of oncology, 26 (2015) 259-271.

[9] Y. Schirris, M. Engelaer, A. Panteli, H.M. Horlings, E. Gavves, J. Teuwen, WeakSTIL: Weak whole-slide image level stromal tumor infiltrating lymphocyte scores are all you need, in: Conference on Medical Imaging - Digital and Computational Pathology, Vol. 12039, Electr Network, 2022.

[10] N. Shvetsov, M. Gronnesby, E. Pedersen, K. Mollersen, L.T.R. Busund, R. Schwienbacher, L.A. Bongo, T.K. Kilvaer, A Pragmatic Machine Learning Approach to Quantify Tumor-Infiltrating Lymphocytes in Whole Slide Images, Cancers, 14 (2022).

[11] M. Yosofvand, S.Y. Khan, R. Dhakal, A. Nejat, N. Moustaid-Moussa, R.L. Rahman, H. Moussa, Automated Detection and Scoring of Tumor-Infiltrating Lymphocytes in Breast Cancer Histopathology Slides, Cancers, 15 (2023).

[12] P. Bankhead, M.B. Loughrey, J.A. Fernández, Y. Dombrowski, D.G. McArt, P.D. Dunne, S. McQuaid, R.T. Gray, L.J. Murray, H.G. Coleman, QuPath: Open source software for digital pathology image analysis, Scientific Reports, 7 (2017) 1-7.

[13] W.W. Zhang, J.M. Pang, K. Chen, C.C. Loy, K-Net: Towards Unified Image Segmentation, in: 35th Annual Conference on Neural Information Processing Systems (NeurIPS), Vol. 34, Electr Network, 2021.

[14] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, in: 18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Vol. 9351, Munich, GERMANY, 2015, pp. 234-241.