1st Serbian International Conference on Applied Artificial Intelligence (SICAAI)
Kragujevac, Serbia, May 19-20, 2022

HISTOPATHOLOGICAL H&E-STAINED IMAGE ANALYSIS BASED ON AI

Jelena Musulin[1], Daniel Štifanić[1], Ana Zulijani[2], Sandi Baressi Šegota[1], Nikola Anđelić[1], Ivan Lorencin[1], Matko Glučina[3], Zlatan Car[1]

[1] Faculty of Engineering, University of Rijeka, Vukovarska 58, 51000 Rijeka, Croatia
e-mail: jmusulin@riteh.hr, dstifanic@riteh.hr, sbaressisegota@riteh.hr, ilorencin@riteh.hr, nandelic@riteh.hr, car@riteh.hr
[2] Department of Oral Surgery, Clinical Hospital Center of Rijeka, Krešimirova 40, 51000 Rijeka
e-mail: ana.zulijani@gmail.com
[3] University of Rijeka, Trg braće Mažuranića 10, 51000 Rijeka
e-mail: matko.glucina@uniri.hr

Abstract:

Over the past decade, improvements in image analysis methods and substantial advancements in processing power have allowed the development of powerful computer-aided analytical approaches to medical data. Tissue histology slides can now be scanned and preserved in digital form, thanking to the recent introduction of the entire slide digital scanners. In such form, they can serve as input data for Artificial Intelligence (AI) algorithms that can speed up standard procedures for histology analysis with high accuracy and precision. The aim of this research was to create an automated system based on AI for histopathological image analysis. The first step was to normalize H&E-stain images, and then use them as an input to the convolutional neural network. The best results are achieved using ResNet50 with highest AUC value of 0.98 ($\pm\sigma$=0.02). Such an approach proved to be successful in analyzing histopathological images.

Keywords: artificial intelligence, convolutional neural network, histopathological analysis, oral squamous cell carcinoma

1. Introduction

Image data is extremely important in healthcare. Lately, the massive accumulation of digital images has increased the demand for their analysis, such as computer-aided diagnosis using Artificial Intelligence algorithms. Medical image analysis is one of the areas where histological tissue patterns are combined with computer-aided image analysis to improve detection and classification of disease [1]. There is also the possibility of automating and speeding up processes that take a long time to complete manually. AI-based models may learn to recognize specific traits in these images, making the diagnostic procedure much faster and more accurate.

The dataset used in this research consists of histopathological images of the oral squamous cell carcinoma (OSCC) region, which contains abnormalities. Oral squamous cell carcinoma is the most common histological neoplasm of head and neck cancers, and while it is located in an easily visible area and can be detected early, this does not always eventuate [2]. Despite advances in therapeutic approaches, the morbidity and mortality rates from OSCC have not improved significantly over the last 30 years. The 5-year survival rate for patients with OSCC ranges between 40% and 50% [3]. The most prevalent reasons why

OSCC is detected in advanced stages include an incorrect initial diagnosis and the ignorance from the patient or from the attending physician.

Clinical examination, conventional oral examination (COE), and histological evaluation following biopsy are procedures for detecting oral cancer. These procedures can detect cancer in the stage of established lesions with significant malignant changes. However, the subjective component of the examination, respectively inter- and intra-observer variability, is the fundamental difficulty in employing histopathological examination for tumor differentiation. Moreover, from the pathologist's point of view, providing exact histological identification in the context of multi-class grading is crucial. For this reason, combination of AI-based approaches with clinical prospective could reduce inter- and intra-observer variability as well as assist pathologists in terms of reducing the load of manuals inspection in shorter time [2].

1.1 Related work

Nowadays, there is interest in leveraging AI tools that pathology departments have several commercial digital pathology platforms available for diagnostic work. Several papers have demonstrated the viability of developing AI-based algorithms to analyze histopathology images.

Pantanowitz et al. (2020) developed an AI-based algorithm using H&E-stained slides of prostate core needle biopsies (CNBs). Their paper describes the successful development, validation, and deployment in clinical practice of an AI-based algorithm for accurately detecting, grading, and evaluating clinically relevant findings in digitized slides of prostate CNBs [4]. Song et al. (2020) used a deep learning algorithm trained on H&E-stained whole slide images of gastric cancer in order to create a clinically applicable system. They show that the system could help pathologists improve diagnostic accuracy and avoid misdiagnoses [5]. On the other hand, Chen et al. (2020) demonstrate that a deep learning algorithm could be used to assist pathologists in the classification of histopathology H&E images in liver cancer [6].

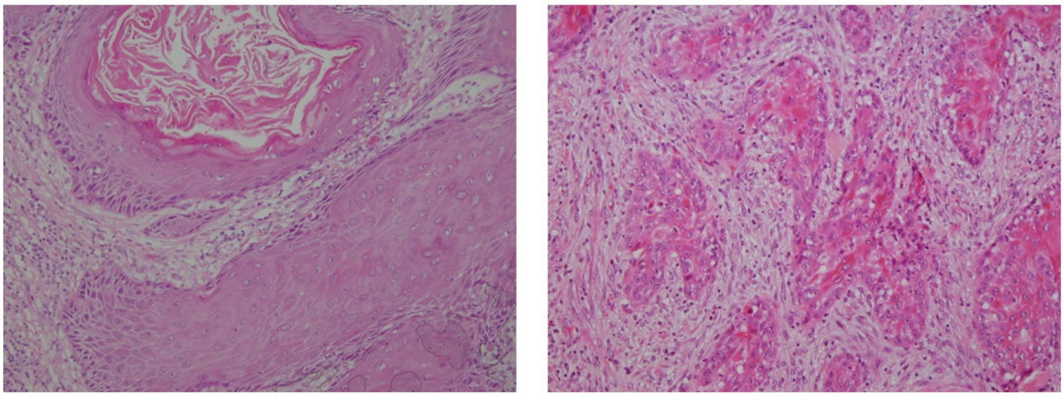
Literature reveals that most researchers have applied AI-based algorithms in order to develop computer-assisted diagnostic tools to evaluate biopsies and improve diagnostic accuracy.

2. Materials and Methods

This section provides a detailed description of the dataset used for classification of oral squamous cell carcinoma, as well as a brief overview of AI-based models.

2.1 Dataset Description

For this research 127 histopathological H&E-stained images with 768x768-pixel size have been used to create the dataset. Haematoxylin and eosin (H&E) staining is one type of tissue staining that is of particular interest to pathologists. This is because the H stain highlights nuclei in blue against a pink cytoplasmic background (and other issue regions). This allows a pathologist to quickly identify and examine tissue, which is a labor-intensive operation. The OSCC samples were retrieved from the archives of the Clinical Department of Pathology and Cytology, Clinical Hospital Center in Rijeka. Two unbiased pathologists analyzed sample slides and classified them according to the 4th edition of the World Health Organization (WHO) classification of Head and Neck malignancies and the 8th edition of the American Joint Committee on Cancer (AJCC) Cancer Staging Manual. In accordance with the aforementioned classification, images have been divided into two classes, well- and moderately differentiated OSCC as shown in Figure 1.



a) Well differentiated b) Moderately differentiated

Fig. 1. OSCC group of well- and moderately differentiated OSCC with magnification x10

Hematoxylin-eosin (HE) staining is most commonly used method in the histopathological examination of tissue sections. For HE-staining, 4 μm thick sections were deparaffinized with xylene and rehydrated in a graded ethanol series and stained with HE according to standard protocol. HE stained sections were captured using the light microscope (Olympus BX51, Olympus, Japan) equipped with a digital camera (DP50, Olympus, Japan) and transmitted to a computer by CellF software (Olympus, Japan). Images were captured at x10 objective lenses.

Since fields like medical image analysis rarely have access to a large number of samples, whilst AI models rely on a large number of samples to achieve good performance and avoid overfitting, it is required to use augmentation techniques to significantly improve the amount and quality of the data [7]. Geometrical transformations used for the augmentation procedure are horizontal flip, horizontal flip combined with 90 degrees anticlockwise rotation, vertical flip, and vertical flip combined with 90 degrees anticlockwise rotation, 90 degrees anticlockwise rotation, 180 degrees anticlockwise rotation and 270 degrees anticlockwise rotation. The augmentation process is used only for the development of training samples, as newly generated data are variants of the original data. Testing samples are not augmented.

2.2 Convolutional Neural Network architectures

Convolutional neural networks (CNN) have emerged as the most prominent strain of neural networks in research in recent years [8]. They have revolutionized computer vision, achieving cutting-edge results in many fundamental tasks while also making significant advances in natural language processing, reinforcement learning, and many other areas. In this research, for classification purposes, we are using three deep CNN architectures.

2.2.1 VGG-16

Simonyan & Zisserman in their paper investigate how the CNN depth affects its accuracy in the large-scale image recording setting. Their main contribution is a thorough evaluation of networks of increasing depth using an architecture with very small convolution filters, which demonstrates that increasing the depth to 16–19 weight layers results in a significant improvement over prior-art configurations [9].

2.2.2 ResNet50

Due to the well-known vanishing gradient problem, deep neural networks become increasingly difficult to train. For that reason, He et al. (2016) propose a residual network (ResNets) to aid in the training of deep neural networks. They refined the residual block as well as the pre-activation variant of the residual block, allowing vanishing gradients to flow unhindered to any previous layer via the shortcut connections. ResNet50 architecture replaces every 2-layer block in the 34-layer network with a 3-layer bottleneck block, which results in 50 layers [10].

2.2.3 InceptionResNetv2

Szegedy et al. proposed several techniques for optimizing the network to loosen the constraints for easier model adaptation in an InceptionV3 architecture, including factorized convolutions, regularization, dimension reduction, and parallelized computations [11]. Furthermore, because the Inception architecture has been demonstrated to be successful at a low computational cost, Szegedy et al. introduce the InceptionResNetv2, which combines the Inception architecture with residual connections. This type of architecture improved both recognition performance and training speed [12].

3. Results and Discussion

Over the past decade, the incidence of oral cancer has increased, especially among young adults. The cause of oral squamous cell carcinoma is multifactorial, and the consumption of tobacco and alcohol have been well established as significant risk factors for the development of oral cancer [13]. As mentioned in the Introduction, the gold standard for diagnosis of oral cancer is tissue biopsy with routine histopathological examination. According to the WHO, OSCC can be graded as well differentiated, moderately differentiated and poorly differentiated, based upon the degree of resemblance to normal squamous epithelium and the amount of keratin production [14]. According to numerous studies, the WHO classification system is not a reliable prognostic and predictive factor for patent outcome. This could be partly due to the fact that grading of OSCC is a subjective process that depends on the area of tumor samples and the evaluation criteria of the individual pathologist. The majority of OSCCs show histological heterogeneity, and in these cases the highest grade should be recorded [15].

An automated H&E-stain histopathological analysis could assist the pathologist in discovering new informative features and in analysing the tumor microenvironment. In order to perform automated image analysis, H&E-stained images need to be normalized as shown in Figure 2.

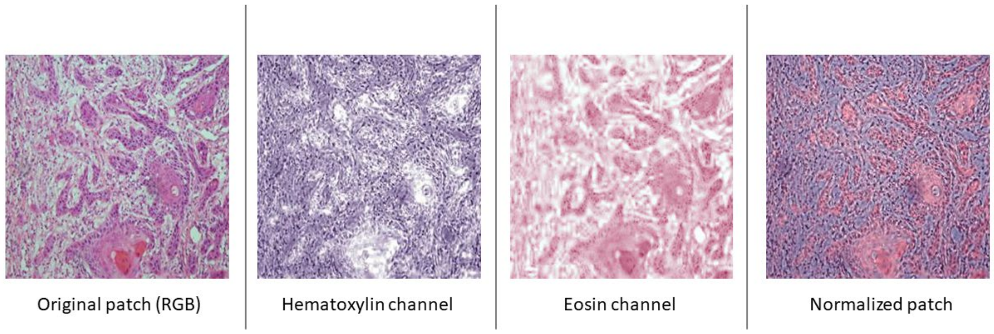


Fig. 2. Visual representation of H&E-stained normalization

This is due to the large color variations in images caused by sample preparation and imaging settings. In our research, we used Macenko approach [10] where the Singular Value Decomposition (SVD) geodesic method is used for obtaining stain vectors. The first step is to convert the RGB color vector to their corresponding optical density (OD) values then remove data with OD intensity less than β. Threshold value of $\beta = 0.15$ was found to provide the most robust results while removing as little data as possible. The next step is to calculate singular value decomposition (SVD) on the OD tuples then create a plane from the SVD directions corresponding to the two largest singular values. After projecting data onto the plane and normalizing to the unit length we calculate the angle of each point with regard to the first

SVD direction. The final step is to convert extreme values back to OD space. Figure 3. shows images before and after normalization.

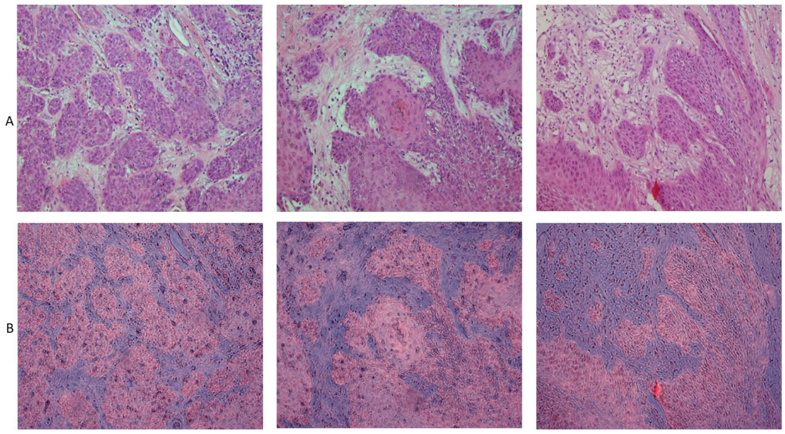


Fig. 3. Visual representation of A) H&E-stained images and B) normalized H&E-stained images

Normalized H&E-stained images are than used as input for deep CNN architectures. The first experimental results are achieved with VGG-16, ResNet50 and InceptionResNetv2 which are pretrained on ImageNet. Stratified 5-fold cross-validation is used to estimate the performance of AI-based model while Area Under the ROC Curve (AUC) is used as evaluation metric.

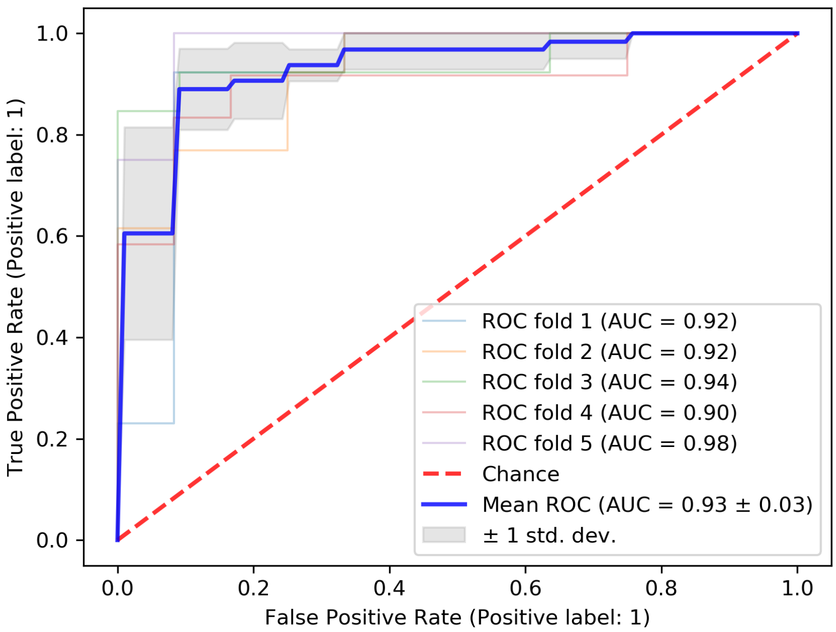


Fig. 4. AUC values of 5-fold cross-validation utilizing VGG-16 architecture

In the case of VGG-16, the highest AUC value of 0.98 is achieved in the fifth fold as shown in Figure 4. However, the same architecture achieves the lowest values in the fourth fold. The mean AUC value of five-fold cross-validation is 0.93 along with a standard deviation (σ) of 0.03.

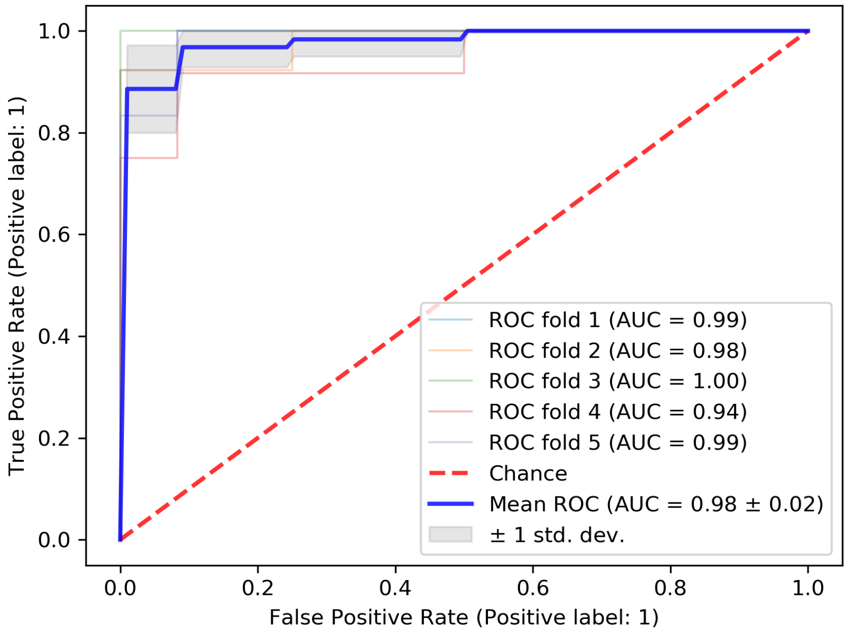


Fig. 5. AUC values of 5-fold cross-validation utilizing ResNet50 architecture

On the other hand, by using ResNet50 architecture the highest AUC value of 1.00 is achieved in the third fold, while the fourth fold achieves the lowest value as shown in Figure 5. The mean AUC and standard deviation values of 5-fold cross-validation are 0.98 ± σ=0.02.

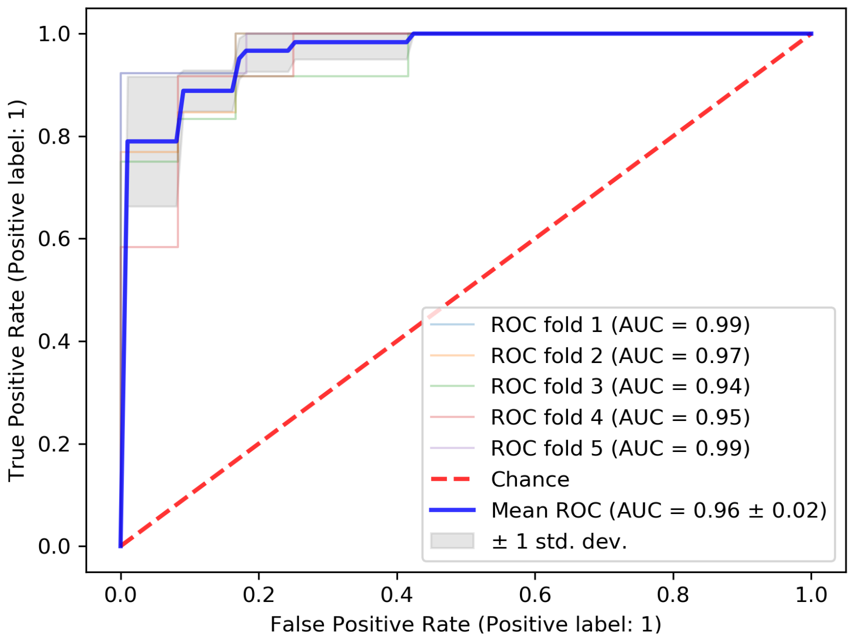


Fig. 6. AUC values of 5-fold cross-validation utilizing InceptionResNetv2 architecture

Figure 6. shows that InceptionResNetv2 architecture achieves the overall highest AUC value of 0.99 in the first and last fold. Moreover, the lowest value is achieved in the third fold. The mean AUC value of five-fold cross-validation is 0.96 along with a standard deviation of 0.02.

After image normalization ResNet50 resulted in the highest classification value of 0.98 (σ ± 0.02) AUC. Figures 4., 5. and 6. represents the ROC metric to evaluate classifier output quality using cross-validation. ROC curves are shown for each of the 5-folds cross-validation and the overall average ROC curve (blue), based on VGG-16, ResNet50 and InceptionResNetv2 predictions. The best results were achieved when two additional layers were added at the end of the base ResNet50 architecture. The first added layer was global average pooling, and the second was the fully connected layer, also known as the output layer.

4. Conclusions

Obtained results reveal that the application of AI-based algorithms along with preprocessing methods, such as image normalization for image analysis, has great potential in the diagnosis of OSCC. Integration of preprocessing method along with the convolutional neural network resulted in 0.98 (σ ± 0.02) AUC. However, data availability was a limitation of the research so the future work should use a dataset with more histopathology images to create a more robust system. The presented approach is the first step in automating histopathological image analysis, therefore, in future work, we plan to integrate more preprocessing methods with other AI classification algorithms.

Acknowledgment: This research has been (partly) supported by the CEEPUS network CIII-HR-0108, European Regional Development Fund under the grant KK.01.1.1.01.0009 (DATACROSS), project CEKOM under the grant KK.01.2.2.03.0004, Erasmus+ project WICT under the grant 2021-1- HR01-KA220-HED-000031177 and University of Rijeka scientific grant uniri-tehnic-18-275- 1447.

References

[1] Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., & Yener, B. (2009). Histopathological image analysis: A review. IEEE reviews in biomedical engineering, 2, 147-171.

[2] Musulin, J., Štifanić, D., Zulijani, A., Ćabov, T., Dekanić, A., & Car, Z. (2021). An enhanced histopathology analysis: An ai-based system for multiclass grading of oral squamous cell carcinoma and segmenting of epithelial and stromal tissue. Cancers, 13(8), 1784.

[3] Zanoni, D. K., Montero, P. H., Migliacci, J. C., Shah, J. P., Wong, R. J., Ganly, I., & Patel, S. G. (2019). Survival outcomes after treatment of cancer of the oral cavity (1985–2015). Oral oncology, 90, 115-121.

[4] Pantanowitz, L., Quiroga-Garza, G. M., Bien, L., Heled, R., Laifenfeld, D., Linhart, C., ... & Dhir, R. (2020). An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: a blinded clinical validation and deployment study. The Lancet Digital Health, 2(8), e407-e416.

[5] Song, Z., Zou, S., Zhou, W., Huang, Y., Shao, L., Yuan, J., ... & Shi, H. (2020). Clinically applicable histopathological diagnosis system for gastric cancer detection using deep learning. Nature communications, 11(1), 1-9.

[6] Chen, M., Zhang, B., Topatana, W., Cao, J., Zhu, H., Juengpanich, S., ... & Cai, X. (2020). Classification and mutation prediction based on histopathology H&E images in liver cancer using deep learning. NPJ precision oncology, 4(1), 1-7.

[7] Musulin, J., Štifanić, D., Zulijani, A., Šegota, S. B., Lorencin, I., Anđelić, N., & Car, Z. (2021, October). Automated Grading of Oral Squamous Cell Carcinoma into Multiple Classes Using Deep Learning Methods. In 2021 IEEE 21st International Conference on Bioinformatics and Bioengineering (BIBE) (pp. 1-6). IEEE.

[8] Štifanić, D., Musulin, J., Car, Z., & Čep, R. (2020). Use of Convolutional Neural Network for Fish Species Classification. Pomorski zbornik, 59(1), 131-142.

[9] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[10] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Proceedings of the IEEE conference on computer vision and pattern recognition.

[11] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826).

[12] Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017, February). Inception-v4, inception-resnet and the impact of residual connections on learning. In Thirty-first AAAI conference on artificial intelligence.

Macenko, M., Niethammer, M., Marron, J. S., Borland, D., Woosley, J. T., Guan, X., ... & Thomas, N. E. (2009, June). A method for normalizing histology slides for quantitative analysis. In 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (pp. 1107-1110). IEEE.

[13] Warnakulasuriya, S. (2009). Global epidemiology of oral and oropharyngeal cancer. Oral oncology, 45(4-5), 309-316.

[14] El-Naggar AK, Chan JKC, Rubin Grandis J, Takata T, Slootweg PJ. International Agency for Research on Cancer. WHO classification of head and neck tumours. World Health Organization classification of tumours. 4. Lyon: International Agency for Research on Cancer; 2017

[15] Rahman, N., MacNeill, M., Wallace, W., & Conn, B. (2021). Reframing histological risk assessment of oral squamous cell carcinoma in the era of UICC 8th edition TNM staging. Head and Neck Pathology, 15(1), 202-211.