Automated Oral Epithelial Dysplasia Grading Using Neural Networks and Feature Analysis

Neda Azarmehr[1] n.azarmehr@sheffield.ac.uk
Adam Shephard[2] Adam.Shephard@warwick.ac.uk
Hanya Mahmood[1] h.mahmood@sheffield.ac.uk
Nasir Rajpoot[2] N.M.Rajpoot@warwick.ac.uk
Syed Ali Khurram[1] s.a.khurram@sheffield.ac.uk
[1] *School of Clinical Dentistry, University of Sheffield, UK, Sheffield, UK*
[2] *Department of Computer Science, University of Warwick, UK, Coventry, UK*

Abstract

Oral epithelial dysplasia (OED) is a precancerous lesion, histologically graded as mild, moderate or severe. The manual histological diagnosis of OED is time-consuming and subjective. We explore a customised Neural Architecture Search (NAS) technique to optimise an efficient architecture for full epithelium and individual nuclei segmentation in pathology whole slide images (WSIs). Results show the NAS-derived model outperforms all state-of-the-art networks. Accurate nuclear segmentation allows us to extract morphometric features. We propose a random forest model, using these features, to differentiate between OED grades.

Keywords: Oral Epithelial Dysplasia, Histopathological Images, Neural Architecture Search, Deep Learning, Machine Learning.

1. Introduction

OED is a spectrum of architectural and cytological epithelial changes, with increased risk of progression to squamous cell carcinoma. Deep learning (DL) has been applied to automatically segment the epithelium and individual nuclei within WSIs. However, head and neck cancer research is limited. Previous studies propose segmenting the epithelium into three sub-layers (Shephard et al., 2021). Since layers are manually annotated, to reduce subjectivity, our study focuses on full epithelium segmentation. The study aims are three-fold: **1.** adopt NAS-based (i.e. Auto-Deeplab) method for a customised networks to segment the epithelium, the individual nuclei and stroma. We believe this is the first study applying NAS techniques on nuclear segmentation in OED. **2.** compare performance with well-established architectures. **3.** extract patch-based nuclear morphological and non-morphological features for automated OED grading.

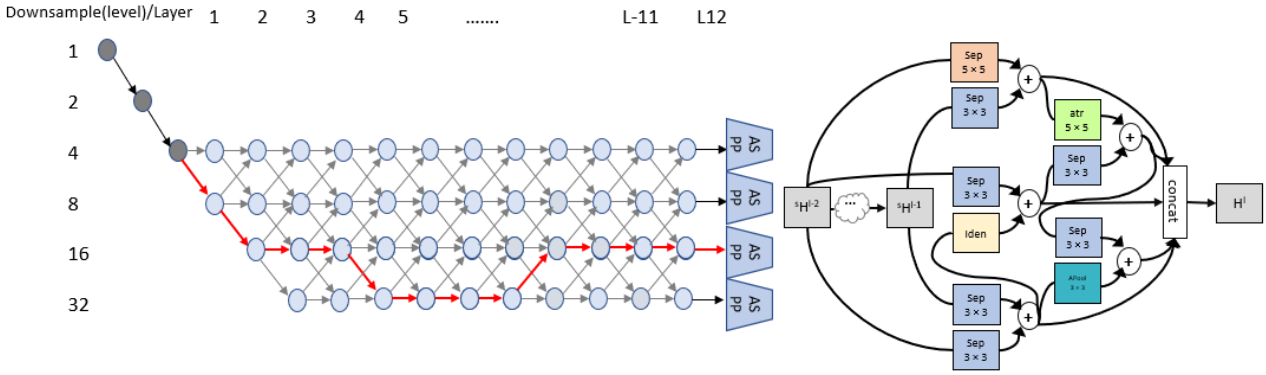
2. Dataset

43 WSIs were collected at the University of Sheffield at 20× magnification. OED cases were labelled by a pathologist and categorised as either mild, moderate, or severe and low- or high-risk based on the world health organisation (WHO) and binary grading systems respectively(Takata and Slootweg, 2017; Barnes et al., 2005), then a contour was drawn around the epithelial boundaries. For each WSI, tissue masks were produced using Otsu thresholding to identify tissue and further exclusion of small holes. Lastly, a single point was assigned to each nucleus and NuClick (Koohbanani et al., 2020) was used to generate the boundaries of individual nuclei for algorithm training. The dataset was split 60/20/20 for training, validation and testing. WSIs were divided into patches of 256×256 pixels.

3. Method

Auto-DeepLab (Liu et al., 2019) applies a two-stage differential NAS: cell and architecture search and architecture evaluation. First, fixed, two-layer "stem" structures are added. Next, localised cells are selected to find a good path in the trellis. The Atrous Spatial Pyramid Pooling (ASPP) module is applied to each layer. A weight (noted as beta) is associated with each grey arrow to run differentiable NAS on this search space; connections are strengthened or weakened after optimisation. The Viterbi algorithm is used to decode the path with the greatest beta. This network-level architecture is searched in addition to the cell-level one. Information about the candidate layer types in every cell and operation can be found in (Liu et al., 2019).

Figure 1: Left: Proposed architecture. Right: The best cell. atr: atrous convolution. Sep:depthwise-separable convolution, Iden: skip connection, AP: average pooling, MP: max pooling. Horizontal axis= number of layers/cells, vertical axis= downsampling rate.



4. Experiment Design and Results

Models were trained separately using the following parameters: cross-entropy loss, learning rate of 10-4, weight decay of 0.0001, momentum of 0.9, and max. 100 epoch. The search took ≈40 hours using Distributed Data-Parallel techniques. The model with the highest F1-score is used to extract six nuclear shape features and two non-morphological features. A Random Forest model is trained using these features to predict OED grade. The highest F1-score (0.935) for full epithelium segmentation is achieved by the proposed model (Auto-Deeplab) using the NAS solution. Also, the proposed model achieves the highest F1-score (0.945) for nuclear segmentation when compared with standard architectures (see details in Table 1). A Random Forest model was trained using the extracted features mentioned earlier for OED grade prediction. The model on the test set for binary grading achieves a higher F1-score than the WHO system (0.787 compared to 0.627- details table 2).

5. Discussion and Conclusion

Initial results using the NAS approach outperform state-of-the-art DL models for epithelium and nuclear segmentation and highlight the potential role of NAS architectures in computational pathology. Further testing on a larger, multicentric cohort is needed to improve accuracy and determine clinical applicability.

Table 1: Comparison of results for full epithelium (top) and nuclear (bottom) segmentation

Model/task	F1-score	Precision	Recall	Accuracy
Full epithelium segmentation				
U-Net	0.744	0.740	0.765	0.751
SegNet	0.739	0.731	0.723	0.724
DeeplabV3ResNet101	0.808	0.853	0.786	0.820
Auto-Deeplab	**0.935**	**0.962**	**0.933**	**0.930**
Nuclear segmentation				
U-Net	0.757	0.742	0.740	0.7611
SegNet	0.660	0.653	0.652	0.681
DeeplabV3ResNet101	0.895	0.882	0.895	0.910
Auto-Deeplab	**0.945**	**0.944**	**0.945**	**0.952**

Table 2: Comparative experiments for grade prediction using two grading systems

Grade system	F1-score	Precision	Recall	AUC
WHO	0.627	0.634	0.615	0.797
Binary	0.787	0.786	0.793	0.853

Acknowledgments

This work was supported by a Cancer Research UK (Grant no. C63489/A29674).

References

Leon Barnes, John Eveson, Peter Reichart, David Sidransky, et al. World health organization classification of tumours: pathology and genetics of head and neck tumours. 2005.

Navid Alemi Koohbanani, Mostafa Jahanifar, Neda Zamani Tajadin, and Nasir Rajpoot. Nuclick: a deep learning framework for interactive segmentation of microscopic images. *Medical Image Analysis*, 65:101771, 2020.

Chenxi Liu, Liang-Chieh Chen, Florian Schroff, Hartwig Adam, Wei Hua, Alan L Yuille, and Li Fei-Fei. Auto-deeplab: Hierarchical neural architecture search for semantic image segmentation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 82–92, 2019.

Adam J Shephard, Simon Graham, Saad Bashir, Mostafa Jahanifar, Hanya Mahmood, Ali Khurram, and Nasir M Rajpoot. Simultaneous nuclear instance and layer segmentation in oral epithelial dysplasia. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 552–561, 2021.

T Takata and P Slootweg. Tumours of the oral cavity and mobile tongue: epithelial precursor lesions. In AK El-Naggar, JKC Chan, JR Grandis, T Takata, and PJ Slootweg, editors, *WHO Classification of Head and Neck Tumours*. 4th edition, 2017.