A Fully Automated Multi-Scale Pipeline for Oral Epithelial Dysplasia Grading and Outcome Prediction

Adam J Shephard[1] ADAM.SHEPHARD@WARWICK.AC.UK
Neda Azarmehr[2] N.AZARMEHR@SHEFFIELD.AC.UK
Raja Muhammad Saad Bashir[1] SAAD.BASHIR@WARWICK.AC.UK
Shan E Ahmed Raza[1] SHAN.RAZA@WARWICK.AC.UK
Hanya Mahmood[2] H.MAHMOOD@SHEFFIELD.AC.UK
Syed Ali Khurram[2] S.A.KHURRAM@SHEFFIELD.AC.UK
Nasir M Rajpoot[1] N.M.RAJPOOT@WARWICK.AC.UK
[1] *Tissue Image Analytics Centre, University of Warwick, Coventry, UK*
[2] *School of Clinical Dentistry, University of Sheffield, Sheffield, UK*

Editors: Under Review for MIDL 2022

Abstract

Oral epithelial dysplasia (OED) is a premalignant histopathological diagnosis given to lesions of the oral cavity, characterised by changes to the nuclear morphometry and the epithelial layers. In this work, we have finetuned HoVer-Net+ for the simultaneous segmentation of nuclei and the epithelial layers in heamatoxylin and eosin (H&E) stained whole slide images (WSIs). We then employed a multi-scale attention-based multiple instance learning architecture for the prediction of OED status, grade, recurrence and malignant transformation. The impressive results have demonstrated the potential of such methods.

Keywords: Oral Epithelial Dysplasia, Histopathology, HoVer-Net+, Multiple Instance Learning, Deep Learning, Instance Segmentation, Segmentation.

1. Introduction

Head and neck cancer has an estimated incidence of 150,000 new cases every year in Europe alone, with oral squamous cell carcinoma (OSCC) being one of the most common types. These cancers are often detected late, with advanced stage cases having a five-year survival rate of just 40%. OSCCs are typically preceded by a pre-cancerous state, lesions termed oral potentially malignant disorders. Following a biopsy and microscopic examination, these lesions may be given a histopathological diagnosis of OED. OED is characterised by a number of architectural and cytological changes to the epithelium and is often graded as mild, moderate, severe dysplasia or high-risk versus low-risk lesions (Takata and Slootweg, 2017). Changes typically manifest in the basal layer and progress upwards through the epithelial layers (e.g. basal, epithelium, keratin) with severity.

In this work, we aim to optimise HoVer-Net+ (Shephard et al., 2021) to perform simultaneous segmentation of nuclear instances and semantic segmentation of the intra-epithelial layers in H&E stained WSIs. We suggest that these steps may enable us to replicate the important architectural/cytological features used by pathologists to diagnose and grade OED. We used both patch-level and global features in a multi-scale attention-based multiple instance learning (MIL) framework to predict patient OED status, grade, recurrence and transformation.

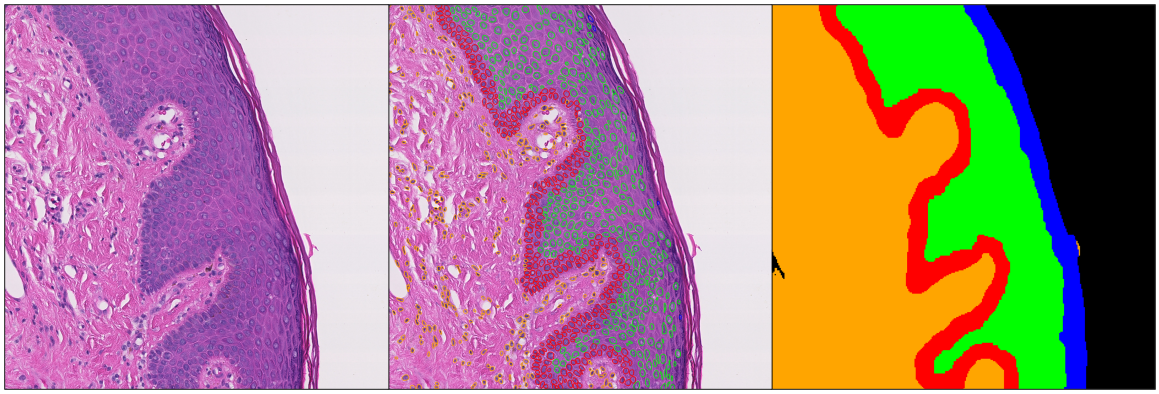

Figure 1: Example output from HoVer-Net+. Left: raw image. Centre: nuclear instance segmentations. Right: layer segmentations. The orange label is for other tissue/nuclei, red for basal epithelium, green for epithelium and blue for keratin.

2. Methods and Materials

134 H&E stained WSIs were scanned at 40× magnification on a Hamamatsu XR scanner at the University of Sheffield (n = 26 controls). Of the 108 OED cases, 27 were graded as high-risk (n = 81 low-risk), 38 had OED recurrence (n = 70 no recurrence), and 20 transformed to malignancy (n = 88 did not transform). A pathologist segmented the intra-epithelial layers (basal, epithelium, keratin) in 20 cases and added point annotations to each nucleus (labelled as "epithelial" or "other") in ten large regions of interest (ROIs).

We used NuClick (Alemi Koohbanani et al., 2020), with manual refinement, to generate nuclear boundaries for the point annotations. Next, we finetuned HoVer-Net+ (Shephard et al., 2021) for the segmentation/classification of nuclear instances and intra-epithelial layers on a subset of our cohort (n = 20), using patches of size 256×256 extracted at 20× magnification. We first tuned the layer segmentation branch of HoVer-Net+ for 20 epochs on all layer patches (trained/tested on 16/4 WSIs). The layer and NuClick segmentations were overlapped to obtain exhaustively annotated ROIs. Patches from these ROIs were used to tune HoVer-Net+ across all branches for 20 epochs with and without the encoder frozen (trained/tested on 6/4 ROIs). HoVer-Net+ was then used for inference on all WSIs.

Each WSI was tessellated into non-overlapping patches of size 1024×1024 (at 20×). 13 features were generated for each nucleus (eccentricity, convex area, contour area, extent, perimeter, solidity, orientation, radius, major/minor axis, equivalent diameter, bounding box area/aspect ratio) and patch-level statistics (mean, minimum, maximum, variance, kurtosis, skew, 25/50/75th percentiles) were calculated for each nucleus type, giving 234 features per patch. We also generated intra-epithelial layer mean widths as global features.

Finally, we used CLAM (Lu et al., 2021), with our patch-level features, to generate WSI-level predictions for subject status, grade, recurrence and transformation. We further modified this architecture to include WSI-level layer features at the final fully-connected layer (multi-scale CLAM). These models were trained using stratified 3-fold cross-validation.

Table 1: Mean (standard deviation) macro F1-score and AUROC for predicting outcomes.

	CLAM		Multi-Scale CLAM	
Outcome	F1M	AUROC	F1M	AUROC
Status	0.84 (0.08)	**0.94 (0.05)**	**0.85 (0.10)**	0.92 (0.06)
Binary grade	0.63 (0.10)	0.75 (0.05)	**0.72 (0.06)**	**0.81 (0.03)**
Recurrence	**0.73 (0.11)**	**0.77 (0.12)**	0.64 (0.05)	0.67 (0.02)
Transformation	0.68 (0.11)	0.72 (0.08)	**0.71 (0.09)**	**0.79 (0.05)**

3. Results

For the segmentation of the basal, epithelial and keratin layers, HoVer-Net+ obtained an F1-score of 0.73, 0.88 and 0.82, respectively. We achieved an F1-score of 0.84 for epithelial nuclei and 0.78 for other nuclei segmentation (see Figure 1). Our multi-scale CLAM achieved high scores for case/control prediction (F1M = 0.85, AUROC = 0.92). Results were lower for the other outcomes, but improved upon CLAM for all outcomes except recurrence (Table 1).

4. Discussion and Conclusion

We have shown the success and generalisability of HoVer-Net+ to be applied to oral dysplasia analysis with finetuning. We also demonstrate the use of multi-scale MIL approaches for histopathological diagnosis and prognosis using an end-to-end fully-automated pipeline.

Acknowledgments

This work was supported by a Cancer Research UK Early Detection Project Grant, as part of the ANTICIPATE study (grant no. C63489/A29674).

References

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.

Ming Y Lu, Drew FK Williamson, Tiffany Y Chen, Richard J Chen, Matteo Barbieri, and Faisal Mahmood. Data-efficient and weakly supervised computational pathology on whole-slide images. *Nature biomedical engineering*, 5(6):555–570, 2021.

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.