

## Article type: Brief Report

### Title: A deep learning algorithm to detect cutaneous squamous cell carcinoma on frozen sections in Mohs micrographic surgery: a retrospective assessment

Matthew J. Davis, MD<sup>1,\*</sup>; Gokul Srinivasan<sup>2,\*</sup>; Rachael Chacko<sup>3</sup>, BS, Sophie Chen<sup>4</sup>; Anish Suvarna<sup>5</sup>; Louis J. Vaickus, MD PhD<sup>6</sup>; Veronica C. Torres PhD<sup>7</sup>; Sassan Hodge<sup>7</sup>; Eunice Y. Chen MD<sup>3,8</sup>; Sarah Preum PhD<sup>9</sup>; Kimberley S. Samkoe PhD<sup>3,7</sup>; Brock C. Christensen PhD<sup>10,11,12</sup>; Matthew LeBoeuf, MD PhD<sup>1,\*\*</sup>; Joshua J. Levy, PhD<sup>1,2,3,6,10,13,\*\*</sup>

- 1) Department of Dermatology, Dartmouth-Hitchcock Medical Center, Lebanon, NH
- 2) Dartmouth College, Hanover, NH
- 3) Geisel School of Medicine, Hanover, NH
- 4) Caddo Parish Magnet High School, Shreveport, LA
- 5) Thomas Jefferson School for Science and Technology, Alexandria, VA
- 6) Department of Pathology and Laboratory Medicine, Dartmouth-Hitchcock Medical Center, Lebanon, NH
- 7) Thayer School of Engineering, Dartmouth College, Hanover, NH
- 8) Department of Surgery, Dartmouth Hitchcock Medical Center, Lebanon, NH
- 9) Department of Computer Science, Dartmouth College, Hanover, NH
- 10) Department of Epidemiology, Dartmouth College Geisel School of Medicine, Hanover, NH
- 11) Department of Molecular and Systems Biology, Geisel School of Medicine at Dartmouth, Lebanon, NH 03756
- 12) Department of Community and Family Medicine, Geisel School of Medicine at Dartmouth, Lebanon, NH 03756
- 13) Program in Quantitative Biomedical Science, Dartmouth College Geisel School of Medicine, Hanover, NH

\* Authors contributed equally as first author on the manuscript

\*\* Authors contributed equally as senior author on the manuscript

#### Corresponding authors:

Matthew LeBoeuf, MD, PhD  
Department of Dermatology  
Dartmouth Hitchcock Medical Center  
1 Medical Center Drive, Lebanon, NH 03756  
Email: [Matthew.LeBoeuf@hitchcock.org](mailto:Matthew.LeBoeuf@hitchcock.org)

Joshua J. Levy, PhD  
Department of Dermatology  
Department of Epidemiology  
Dartmouth Hitchcock Medical Center  
Geisel School of Medicine  
1 Medical Center Drive, Lebanon, NH 03756  
Email: [Joshua.J.Levy@hitchcock.org](mailto:Joshua.J.Levy@hitchcock.org)

**Manuscript word count:** 1200 words [maximum 1200]

## Key Points

**Question:** How can the efficiency and accuracy of real-time intraoperative margin analysis for the removal of cutaneous squamous cell carcinoma (cSCC) be improved, and how can tumor differentiation be incorporated into this approach?

**Findings:** A proof-of-concept deep learning algorithm was trained, validated, and tested on frozen section whole slide images (WSI) for a retrospective cohort of cSCC cases, demonstrating high accuracy in identifying cSCC and related pathologies. Histomorphology alone was found to be insufficient to delineate tumor from epidermis in histologic identification of well-differentiated cSCC. Incorporation of surrounding tissue architecture and shape improved the ability to delineate tumor from normal tissue.

**Meaning:** Integrating artificial intelligence into surgical procedures has the potential to enhance the thoroughness and efficiency of intraoperative margin analysis for cSCC removal. However, accurately accounting for the epidermal tissue based on the tumor's differentiation status requires specialized algorithms that consider the surrounding tissue context. To meaningfully integrate AI algorithms into clinical practice, further algorithmic refinement is needed, as well as the mapping of tumors to their original surgical site, and evaluation of the cost and efficacy of these approaches to address existing bottlenecks.

## Abstract

**Importance:** Intraoperative margin analysis is crucial for the successful removal of cutaneous squamous cell carcinomas (cSCC). Artificial intelligence technologies (AI) have previously demonstrated potential for facilitating rapid and complete tumor removal using intraoperative margin assessment for basal cell carcinoma. However, the varied morphologies of cSCC present challenges for AI margin assessment.

**Objective:** To develop and evaluate the accuracy of an AI algorithm for real-time histologic margin analysis of cSCC.

**Design:** A retrospective cohort study was conducted using frozen cSCC section slides and adjacent tissues.

**Setting:** This study was conducted in a tertiary care academic center.

**Participants:** Patients undergoing Mohs micrographic surgery for cSCC between January and March 2020.

**Exposures:** Frozen section slides were scanned and annotated, delineating benign tissue structures, inflammation, and tumor to develop an AI algorithm for real-time margin analysis. Patients were stratified by tumor differentiation status. Epithelial tissues including epidermis and hair follicles were annotated for moderate-well to well differentiated cSCC tumors. A convolutional neural network workflow was used to extract histomorphological features predictive of cSCC at 50-micron resolution.

**Main Outcomes and Measures:** The performance of the AI algorithm in identifying cSCC at 50-micron resolution was reported using the area under the receiver operating characteristic curve. Accuracy was also reported by tumor differentiation status and by delineation of cSCC from epidermis. Model performance using histomorphological features alone was compared to architectural features (i.e., tissue context) for well-differentiated tumors.

**Results:** The AI algorithm demonstrated proof of concept for identifying cSCC with high accuracy. Accuracy differed by differentiation status, driven by challenges in separating cSCC from epidermis using histomorphological features alone for well-differentiated tumors. Consideration of broader tissue context through architectural features improved the ability to delineate tumor from epidermis.

**Conclusions and Relevance:** Incorporating AI into the surgical workflow may improve efficiency and completeness of real-time margin assessment for cSCC removal, particularly in cases of moderately and poorly differentiated tumors/neoplasms. Further algorithmic improvement is necessary to remain sensitive to the unique epidermal landscape of well-differentiated tumors, and to map tumors to their original anatomical position/orientation. Future studies should assess the efficiency improvements and cost benefits and address other confounding pathologies such as inflammation and nuclei.

**Funding sources:** JL is supported by NIH grants R24GM141194, P20GM104416 and P20GM130454. Support for this work was also provided by the Prouty Dartmouth Cancer Center development funds.

**Keywords:** general dermatology; medical dermatology; oncology; Mohs micrographic surgery; clinical research; artificial intelligence

## Introduction

Cutaneous squamous cell carcinoma (cSCC) is the second most common form of skin cancer, with more than one million cases diagnosed in the United States each year <sup>1,2</sup>. While many tumors are isolated to the skin, advanced disease is not uncommon with a metastasis rate of 4% and a disease-specific death estimate of 2.8% <sup>3</sup>. When tumors occur on the head and neck or other high-risk sites, Mohs micrographic surgery (MMS) is the treatment of choice. MMS allows for real-time margin analysis resulting in low rates of recurrence. A recent study by Motley et al., found recurrence rates of cSCC of 3% when treated with MMS and 8% when treated with standard excision, despite a higher proportion of high risk tumors in the MMS group <sup>4</sup>.

Early diagnosis and treatment of cutaneous tumors is essential. Currently, patient demand far outweighs the capacity of the dermatology workforce (Association of American Medical Colleges; AAMC), making early treatment more difficult <sup>5</sup>. Machine learning models exist to detect basal cell carcinoma (BCC) <sup>5-7</sup>, but given the complexities and variable morphologies of cSCC, similar algorithms are yet to be developed for this tumor type <sup>8</sup>. Our study presents an algorithm to detect cSCC on whole slide images (WSI) of frozen sections obtained in MMS <sup>9</sup>. Developing such algorithms may ameliorate the physician deficit by improving access to expert histologic assessment with the potential for broad application across numerous surgical specialties that treat various forms of cSCC.

## Methods

A retrospective study was designed, with WSI scanning (20X resolution using the Aperio AT2 image scanner) of 95 frozen section slides, each containing 3-5 tissue sections, followed by manual annotation of benign tissue structures, inflammation, and tumor by three dermatologists.

WSIs were then split into 256 x 256-pixel image patches (i.e., 50-micron resolution). Patches were randomly distributed into training, testing, and validation sets in an 80:10:10 arrangement, ensuring patches from the same patient were partitioned to the same set (e.g., restricting all patches across all tissue sections for one patient to the validation set). To classify tumors at the patch level, a convolutional neural network (CNN) workflow was implemented, using a ResNet101 model that was pre-trained and selected after comparing multiple neural network architectures (e.g., SWIN-Transformer, EfficientNet) <sup>10</sup>. The CNN workflow dynamically extracts histomorphological features at each 50-micron location, generating a probability score for cSCC between 0 and 1 <sup>11</sup>. After the model was trained and validated, its performance characteristics were evaluated across the validation and testing sets using the Area Under the Receiver Operating Characteristic Curve (AUC), a performance metric that summarizes algorithmic sensitivity and specificity across a range of decision thresholds, with 95% confidence intervals reported using 1000-sample non-parametric bootstrapping.

Distinguishing cSCC from epithelial tissue based on histomorphology alone (i.e., what can be learned by a CNN) can be challenging, particularly in moderate-well to well differentiated squamous cell tumors, we hypothesized that the algorithm would not perform as well in these cases. To test this hypothesis, we annotated the epithelial tissue within the well-differentiated tumors in our cohort and compared the sensitivity and specificity of cSCC detection at 50-micron locations containing either cSCC or epithelium alone. To improve algorithmic performance in distinguishing cSCC from epithelium in well-differentiated tumors, we incorporated larger-scale

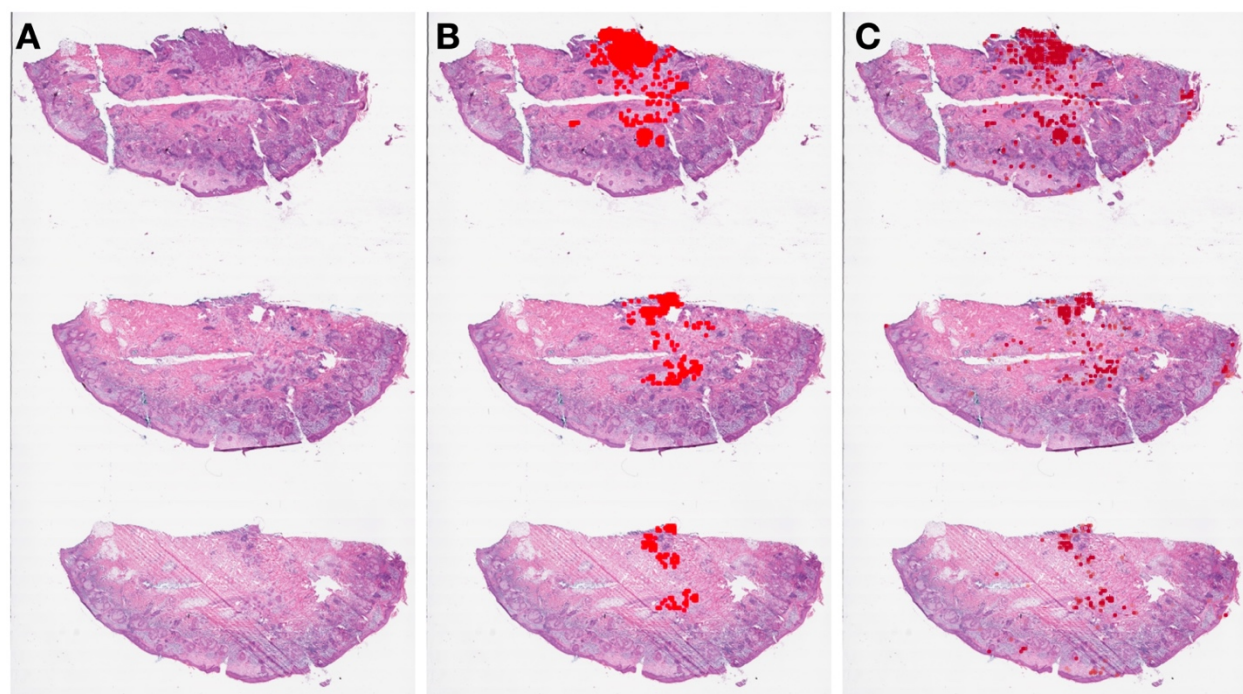
architectural features beyond histomorphology. Specifically, we examined topological and shape descriptors, referred to as "architectural features", of cSCC and epithelial tissue across the training, validation, and test sets<sup>12,13</sup>. Topological and shape (i.e., architectural) features capture the relationships between tissue architectures and their shape properties. For instance, when viewed under a microscope, the epidermis typically appears flat or slightly curved, and may also have ridge-like features in certain areas; in contrast, cSCC is often characterized by a more discohesive and infiltrative growth pattern. The architectural features are numerical descriptors which encapsulate topological and shape differences and were used to train a random forest (RF) model for the purpose of distinguishing between SCC and epithelium. In addition, we incorporated a graph neural network (GNN) to consider contextual information from adjacent image patches<sup>14,15</sup>. A graph neural network (GNN) is a type of neural network designed to operate on graphs and capture complex relationships and interactions between the nodes and edges of a graph. Unlike traditional neural networks that operate on vectorized inputs, GNNs can process structured data, which is useful for a variety of tasks such as node classification, link prediction, and graph clustering. For example, GNN increases the probability of classifying an image patch as epithelium if the surrounding patches were also classified as epithelial. We compared the performance of the architectural and GNN models to that of the CNN workflow to show how using the surrounding tissue architecture improves the accuracy of distinguishing SCC from epithelium in well differentiated tumors.

**Table 1: Performance characteristics for SCC algorithm**, considering histomorphological (CNN), architectural (topology/shape) and contextual (GNN) features across the validation and test sets, broken down by overall performance, tumor differentiation status, and restricting to SCC/epithelium within well-differentiated test set tumors; 95% confidence intervals reported using 1000-sample non-parametric bootstrapping

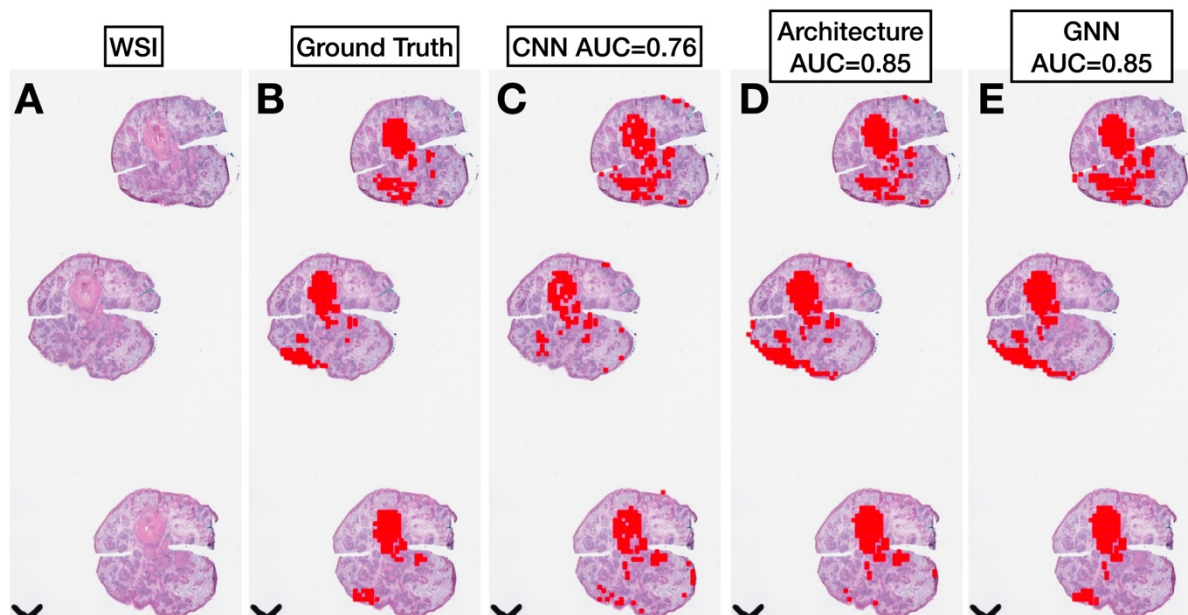
Dataset	Algorithm	AUC	2.5% CI	97.5% CI
---------	-----------	-----	---------	----------



<b>Validation Set: Overall</b>	<b>CNN/Morphology</b>	0.981	0.980	0.982
<b>Test Set: Overall</b>	<b>CNN/Morphology</b>	0.935	0.934	0.936
<b>Poor/Mod-Poor/Mod-Diff</b>	<b>CNN/Morphology</b>	0.968	0.953	0.980
<b>Mod-Well/Well-Diff</b>	<b>CNN/Morphology</b>	0.895	0.837	0.943
<b>cSCC versus Epidermis within Mod-Well/Well-Diff</b>	<b>CNN/Morphology</b>	0.626	0.594	0.658
	<b>RF/Architecture</b>	0.760	0.728	0.792
	<b>GNN/Context</b>	0.764	0.729	0.796



**Figure 1: Example display output of cSCC prediction probabilities at 50-micron resolution for example test-set WSI: A) Original WSI; B) Ground truth cSCC; C) cSCC algorithm predictions.**



**Figure 2: Example display output of cSCC prediction probabilities at 50-micron resolution for cSCC/Epithelium predictions across example test-set WSI for well-differentiated tumor: A) Original WSI; B) Ground truth cSCC; C) cSCC CNN algorithm predictions (histomorphology); D) Topological and shape features (architecture); E) GNN predictions (contextual)**

## Results

The algorithm achieved an AUC of 0.981 (95% CI [0.980-0.982]) and 0.935 (95% CI [0.934-0.936]) for predicting cSCC when applied to the validation and test sets respectively. As expected, the model performed better on poorly to moderately differentiated tumors (AUC=0.968, 95% CI [0.953-0.980]) than on well differentiated tumors (AUC=0.895, 95% CI [0.837-0.943]) (Table 1, Figure 1, Supplementary Figures 1-3). The difficulty in distinguishing normal epidermis from cSCC contributed to these deficiencies, yielding an AUC of 0.626 (95% CI [0.594-0.658]) when distinguishing cSCC from epithelium alone in well differentiated tumors (Figure 1). However, incorporating architectural (AUC=0.760; 95% CI



[0.728-0.792]) and contextual (GNN; AUC=0.764; 95% CI [0.729-0.796]) features significantly improved the algorithm's performance in delineating cSCC from epidermis (**Table 1, Figure 2**).

## **Discussion**

Our study provides an example of a deep learning algorithm used to identify cutaneous squamous cell carcinoma on frozen section slides in MMS. Reducing rate limiting steps to intraoperative margin assessment of cSCC tumors can improve the efficiency and completeness of tumor removal, reducing the burden on laboratory staff while reducing tumor recurrence and repeat procedures<sup>16,17</sup>. When evaluating this study, it should be acknowledged that all slides were obtained from a single MMS clinic and scanned images, not slides, were used for training, which may limit generalizability and real-world implementation. Application of this algorithm requires complete, high-quality tissue sections devoid of tears, holes, and other artifacts which may preclude histological margin assessment. Our data provide evidence supporting the identification of cSCC on frozen section slides, which has historically proven challenging. The algorithm's successful performance in this study suggests its potential for broader use in providing real-time complete margin analysis of cSCC in various body parts. In the future, the focus will be on refining and improving the algorithm's accuracy to enable more detailed identification of various associated pathologies, including single cell analysis, follicles, actinic keratosis, and incidental diagnoses. Additionally, efforts will be made to map tumors to their original anatomical position/orientation and evaluate the efficiency improvements and cost benefits of this algorithmic approach.

## **Conclusion**

This study not only established the general feasibility of histomorphological cSCC detection, but also demonstrated challenges in effectively distinguishing epithelium tissue from cSCC in well-differentiated cases. Therefore, algorithms that consider the surrounding tissue architecture could be useful for these tumors, although further research is necessary to improve the ability to utilize spatial cues. Furthermore, different tumor types may necessitate different algorithms. Future research will also address other confounding pathologies, such as inflammation, nuclei, follicles, architecture, and keratinocyte differentiation, by considering nuclei and large-scale architectural features.

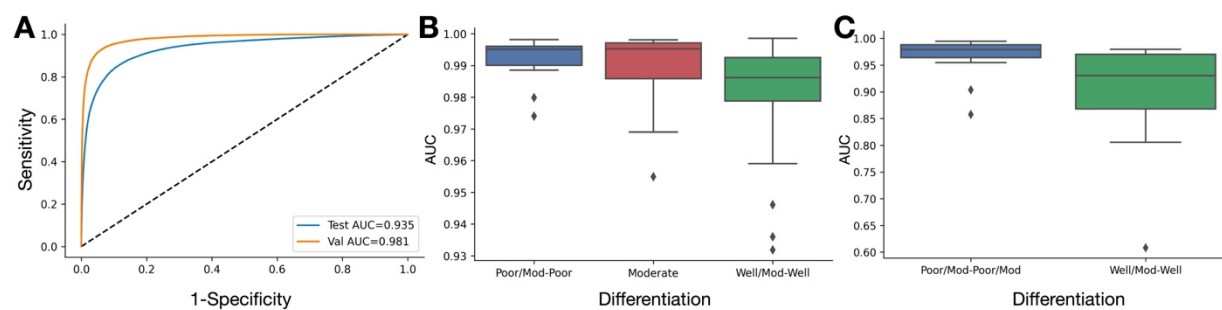
**Acknowledgements:** We would like to congratulate Sophie Chen for presenting an earlier iteration of this work at the national Junior Science and Humanities Symposium, placing 3<sup>rd</sup> in the Computer Science competition.

## References

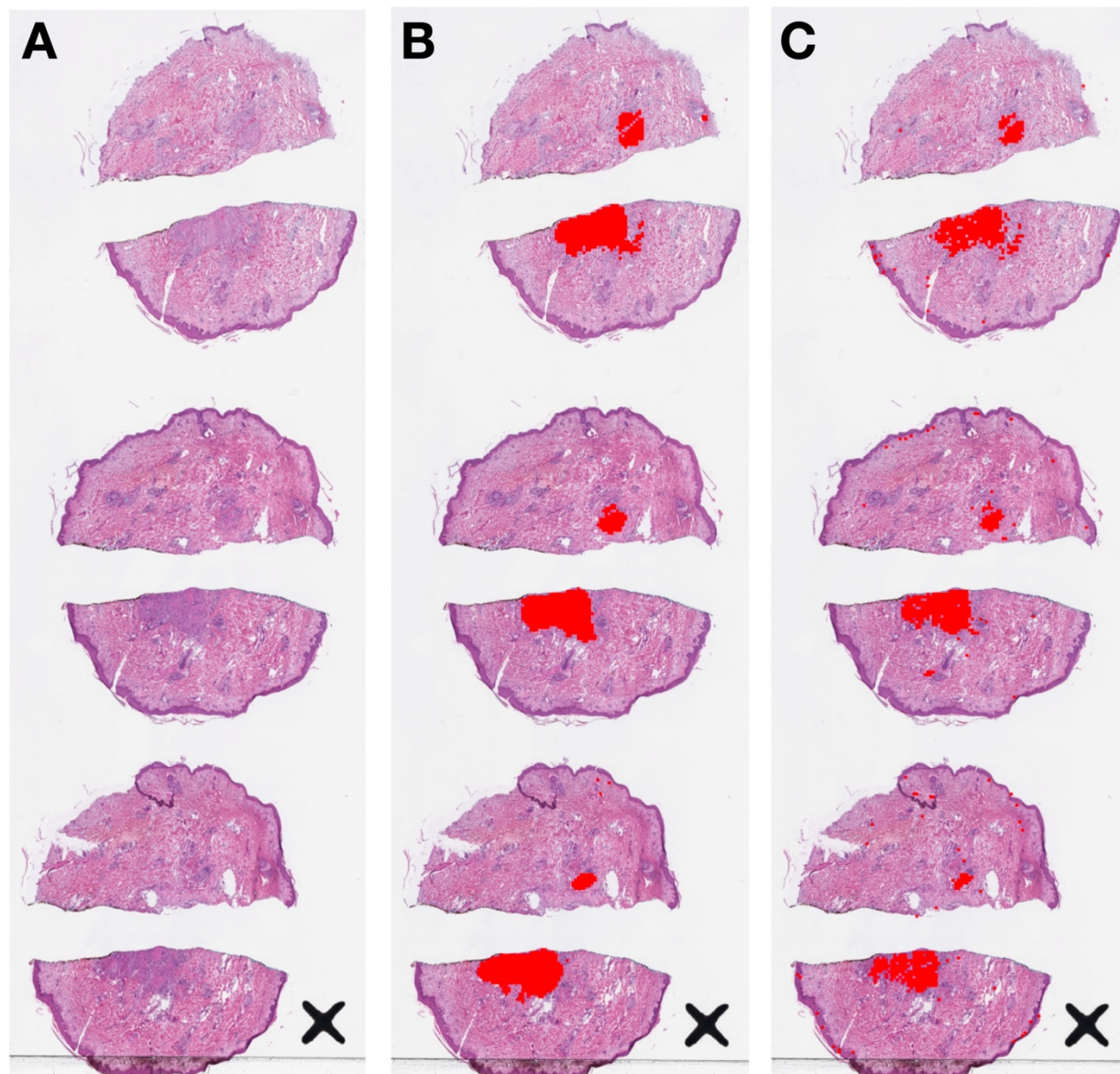
1. Leitenberger JJ, Rogers H, Chapman JC, et al. Defining recurrence of nonmelanoma skin cancer after Mohs micrographic surgery: Report of the American College of Mohs Surgery Registry and Outcomes Committee. *J Am Acad Dermatol*. 2016;75(5):1022-1031. doi:10.1016/j.jaad.2016.06.047
2. Rogers HW, Weinstock MA, Harris AR, et al. Incidence Estimate of Nonmelanoma Skin Cancer in the United States, 2006. *Arch Dermatol*. 2010;146(3):283-287. doi:10.1001/archdermatol.2010.19
3. Eigentler TK, Leiter U, Häfner HM, Garbe C, Röcken M, Breuninger H. Survival of Patients with Cutaneous Squamous Cell Carcinoma: Results of a Prospective Cohort Study. *J Invest Dermatol*. 2017;137(11):2309-2315. doi:10.1016/j.jid.2017.06.025
4. Motley R, Arron S. Mohs micrographic surgery for cutaneous squamous cell carcinoma. *Br J Dermatol*. 2019;181(2):233-234.
5. Dill MJ, Salsberg ES. *The Complexities of Physician Supply and Demand: Projections through 2025*. Association of American Medical Colleges; 2008.
6. Levy J, Davis M, Chacko R, et al. ArcticAI: A Deep Learning Platform for Rapid and Accurate Histological Assessment of Intraoperative Tumor Margins. Published online May 10, 2022:2022.05.06.22274781. doi:10.1101/2022.05.06.22274781
7. Campanella G, Nehal KS, Lee EH, et al. A deep learning algorithm with high sensitivity for the detection of basal cell carcinoma in Mohs micrographic surgery frozen sections. *J Am Acad Dermatol*. 2021;85(5):1285-1286. doi:10.1016/j.jaad.2020.09.012
8. Petter G, Haustein UF. Histologic Subtyping and Malignancy Assessment of Cutaneous Squamous Cell Carcinoma. *Dermatol Surg*. 2000;26(6):521-530. doi:10.1046/j.1524-4725.2000.99181.x
9. Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. *Lancet Oncol*. 2019;20(5):e253-e261. doi:10.1016/S1470-2045(19)30154-8
10. Paszke A, Gross S, Massa F, et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. *ArXiv191201703 Cs Stat*. Published online December 3, 2019. Accessed May 15, 2021. <http://arxiv.org/abs/1912.01703>
11. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521(7553):436-444. doi:10.1038/nature14539
12. Tauzin G, Lupo U, Tunstall L, et al. giotto-tda: A topological data analysis toolkit for machine learning and data exploration. *J Mach Learn Res*. 2021;22(1):1834-1839.
13. Edelsbrunner H. Alpha shapes-a survey. In: *Tessellations in the Sciences: Virtues, Techniques and Applications of Geometric Tilings*. ; 2011.
14. Li MM, Huang K, Zitnik M. Graph representation learning in biomedicine and healthcare. *Nat Biomed Eng*. Published online 2022:1-17.

15. Levy J, Haudenschild C, Barwick C, Christensen B, Vaickus L. Topological Feature Extraction and Visualization of Whole Slide Images using Graph Neural Networks. *Pac Symp Biocomput Pac Symp Biocomput*. 2021;26:285-296.
16. Bialy TL, Whalen J, Veledar E, et al. Mohs Micrographic Surgery vs Traditional Surgical Excision: A Cost Comparison Analysis. *Arch Dermatol*. 2004;140(6):736-742. doi:10.1001/archderm.140.6.736
17. Otley CC, Salasche SJ. Mohs surgery: efficient and effective. *Br J Ophthalmol*. 2004;88(9):1228-1228.

## Supplementary Material

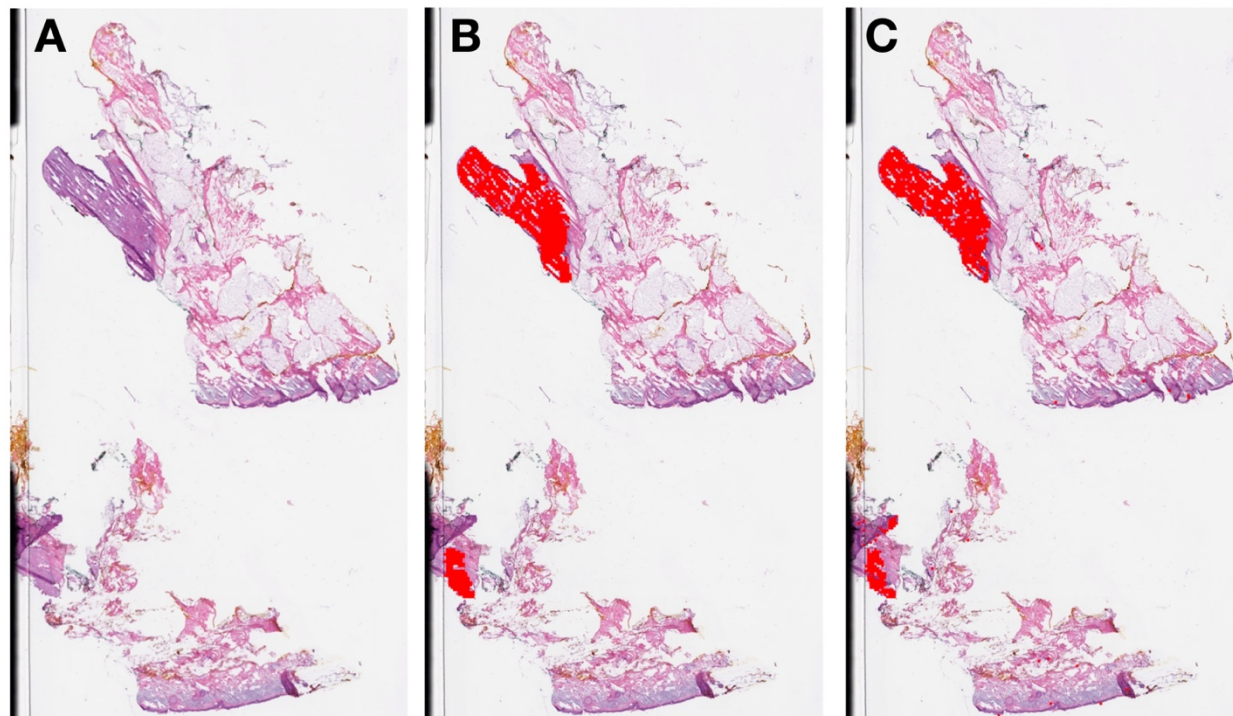


**Supplementary Figure 1: Performance characteristics of cSCC CNN algorithm: A)** Receiver operating characteristic (ROC) curves for validation and test set cases at 50-micron resolution; **B)** Boxplots demonstrating distribution of slide-level AUC values from training set slides, broken down by differentiation status; **C)** Boxplots demonstrating distribution of slide-level AUC values from validation/test set slides, broken down by differentiation status



**Supplementary Figure 2: Example display output of sSCC prediction probabilities at 50-micron resolution for example validation-set WSI (poor-moderately differentiated): A) Original WSI; B) Ground truth cSCC; C) cSCC algorithm predictions**





**Supplementary Figure 3: Example display output of cSCC prediction probabilities at 50-micron resolution for example validation-set WSI (poor-moderately differentiated): A) Original WSI; B) Ground truth cSCC; C) cSCC algorithm predictions**