Received: 22 July 2023 | Revised: 13 September 2023 | Accepted: 30 September 2023

DOI: 10.1111/exd.14949

CONCISE COMMUNICATION

Experimental Dermatology WILEY

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

Matthew J. Davis[1] | Gokul Srinivasan[2] | Rachael Chacko[3] | Sophie Chen[4] | Anish Suvarna[5] | Louis J. Vaickus[6] | Veronica C. Torres[7] | Sassan Hodge[7] | Eunice Y. Chen[3,8] | Sarah Preum[9] | Kimberley S. Samkoe[3,7] | Brock C. Christensen[10,11,12] | Matthew R. LeBoeuf[1] | Joshua J. Levy[1,2,3,6,10,13]

[1]Department of Dermatology, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, USA

[2]Dartmouth College, Hanover, New Hampshire, USA

[3]Geisel School of Medicine, Hanover, New Hampshire, USA

[4]Caddo Parish Magnet High School, Shreveport, Louisiana, USA

[5]Thomas Jefferson School for Science and Technology, Alexandria, Virginia, USA

[6]Department of Pathology and Laboratory Medicine, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, USA

[7]Thayer School of Engineering, Dartmouth College, Hanover, New Hampshire, USA

[8]Department of Surgery, Dartmouth Hitchcock Medical Center, Lebanon, New Hampshire, USA

[9]Department of Computer Science, Dartmouth College, Hanover, New Hampshire, USA

[10]Department of Epidemiology, Dartmouth College Geisel School of Medicine, Hanover, New Hampshire, USA

[11]Department of Molecular and Systems Biology, Geisel School of Medicine at Dartmouth, Lebanon, New Hampshire, USA

[12]Department of Community and Family Medicine, Geisel School of Medicine at Dartmouth, Lebanon, New Hampshire, USA

[13]Program in Quantitative Biomedical Science, Dartmouth College Geisel School of Medicine, Hanover, New Hampshire, USA

Correspondence
Matthew R. LeBoeuf, Department of Dermatology, Dartmouth Hitchcock Medical Center, 1 Medical Center Drive, Lebanon, NH 03756, USA.
Email: matthew.leboeuf@hitchcock.org

Joshua J. Levy, Department of Dermatology, Department of Epidemiology, Dartmouth Hitchcock Medical Center, Geisel School of Medicine, 1 Medical Center Drive, Lebanon, NH 03756, USA.
Email: joshua.j.levy@hitchcock.org

Funding information
National Institutes of Health, Grant/Award Number: P20GM104416 and P20GM130454

Abstract

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 tumour removal using intraoperative margin assessment for basal cell carcinoma. However, the varied morphologies of cSCC present challenges for AI margin assessment. The aim of this study was to develop and evaluate the accuracy of an AI algorithm for real-time histologic margin analysis of cSCC. To do this, a retrospective cohort study was conducted using frozen cSCC section slides. These slides were scanned and annotated, delineating benign tissue structures, inflammation and tumour to develop an AI algorithm for real-time margin analysis. A convolutional neural network workflow was used to extract histomorphological features predictive of cSCC. This algorithm demonstrated proof of concept for identifying cSCC with high accuracy, highlighting

Matthew J. Davis and Gokul Srinivasan are contributed equally as first author on the manuscript.

Matthew R. LeBoeuf and Joshua J. Levy are contributed equally as senior author on the manuscript.

the potential for integration of AI into the surgical workflow. Incorporation of AI algorithms may improve efficiency and completeness of real-time margin assessment for cSCC removal, particularly in cases of moderately and poorly differentiated tumours/neoplasms. Further algorithmic improvement incorporating surrounding tissue context is necessary to remain sensitive to the unique epidermal landscape of well-differentiated tumours, and to map tumours to their original anatomical position/orientation.

KEYWORDS
artificial intelligence, clinical research, general dermatology, medical dermatology, Mohs micrographic surgery, oncology

1 | BACKGROUND

Cutaneous squamous cell carcinoma (cSCC) is the second most common form of skin cancer, with more than 1 million cases diagnosed in the United States each year.[1,2] While many tumours 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%.[3] When tumours 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 and Arron 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 tumours in the MMS group.[4]

Early diagnosis and treatment of cutaneous tumours is essential. Currently, patient demand far outweighs the capacity of the dermatology workforce (Association of American Medical Colleges, AAMC), making early treatment more difficult.[5] Machine learning models exist to detect basal cell carcinoma (BCC),[5–7] but given the complexities and variable morphologies of cSCC, similar algorithms are yet to be developed for this tumour type.[8] Our study presents an algorithm to detect cSCC on whole slide images (WSI) of frozen sections obtained in MMS.[9] Developing such algorithms can improve the accuracy and expand the applicability of the already tremendously efficient practice of Mohs surgery. In the more distant future, FDA approval and clinical implementation of this technology may address challenges related to clinical capacity in intraoperative margin assessment by enhancing access to rapid and reliable histologic evaluation. This holds potential for broad application across numerous surgical specialties that treat various forms of cSCC.

2 | QUESTIONS ADDRESSED

The aim of the study was to develop and evaluate the accuracy of an artificial intelligence (AI) algorithm for real-time histologic margin analysis of cSCC. More specifically, an algorithm was trained to delineate benign tissue structures, inflammation and tumour on frozen section slides obtained from a MMS clinic. A convolutional neural network (CNN) workflow was used to extract histomorphological features predictive of cSCC at 50-micron resolution, and area under the receiver operating curve (AUC) was used to assess model performance.

3 | EXPERIMENTAL DESIGN

Tissue specimens were collected at a local MMS clinic in Lebanon, NH. Given the retrospective nature of this work, Human Research Protection Program of Dartmouth Health (Institutional Review Board) gave ethical approval for this work. Tissue was grossed, sectioned and stained during MMS, 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 tumour by three dermatologists and dermatology residents. The frozen section slides were obtained from a single site MMS clinic from 1 January to 1 March 2020. Every slide that was deemed appropriate for final interpretation was scanned to generate a WSI. ASAP annotation software (Computational Pathology Group, Nijmegen, Netherlands) was used to generate all annotations. Many of the annotations were performed by dermatology residents, but all were confirmed by a board certified Mohs surgeon. Upon further review, the residents and surgeon also noted a few instances ($n=3$) where slides had significant tissue freezing artefacts and quality issues that impacted the ability to annotate tissue slides—these slides were excluded from the study.

WSIs were then split into 256×256-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/slides 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 only). The random assignment of patients maintained a similar distribution of tumours based on their differentiation status. To classify tumours at the patch level, a 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).[10] The CNN workflow dynamically extracts histomorphological features at each 50-micron location, generating a probability score for cSCC between 0 and 1.[11] After the model was trained and validated, its performance characteristics were evaluated across the validation and testing sets using the 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.

As 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 tumours, 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 tumours 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 tumours, 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.[12,13] 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 dyscohesive and infiltrative growth pattern. The architectural features are numerical descriptors which encapsulate topological and shape differences and were used to train a random forest 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.[14,15] A 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 tumours.

While our study used relatively high-quality slides deemed appropriate for real-world complete margin analysis by a Mohs surgeon, recent studies have shown that similar algorithms are able to perform well even on fragmented low-quality frozen specimens.[16] Despite this, our team believes that identifying holes/fragmentation is perhaps more important in this clinical context, as additional tissue sections are required if any epidermis is missing or if any significant holes or processing artefact is present. In a previous work, we developed an algorithm to detect holes on frozen section tissue with the intent of flagging slides that require further reprocessing.[6]

4 | 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 tumours (AUC=0.968, 95% CI [0.953–0.980]) than on well-differentiated tumours (AUC=0.895, 95% CI [0.837–0.943]) (Table 1; Figure 1; Figures S1–S3). 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 tumours (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).

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, tumour differentiation status and restricting to SCC/epithelium within well-differentiated test set tumours; 95% confidence intervals reported using 1000-sample non-parametric bootstrapping.

Dataset	Algorithm	AUC	2.5% CI	97.5% CI
Validation Set: Overall	CNN/Morphology	0.981	0.980	0.982
Test Set: Overall	CNN/Morphology	0.935	0.934	0.936
Poor/Mod-Poor/Mod-Diff	CNN/Morphology	0.968	0.953	0.980
Mod-Well/Well-Diff	CNN/Morphology	0.895	0.837	0.943
cSCC versus Epidermis within Mod-Well/Well-Diff	CNN/Morphology	0.626	0.594	0.658
	RF/Architecture	0.760	0.728	0.792
	GNN/Context	0.764	0.729	0.796

Abbreviations: CNN, convolutional neural network; cSCC, cutaneous squamous cell carcinomas; GNN, graph neural network; RF, random forest.

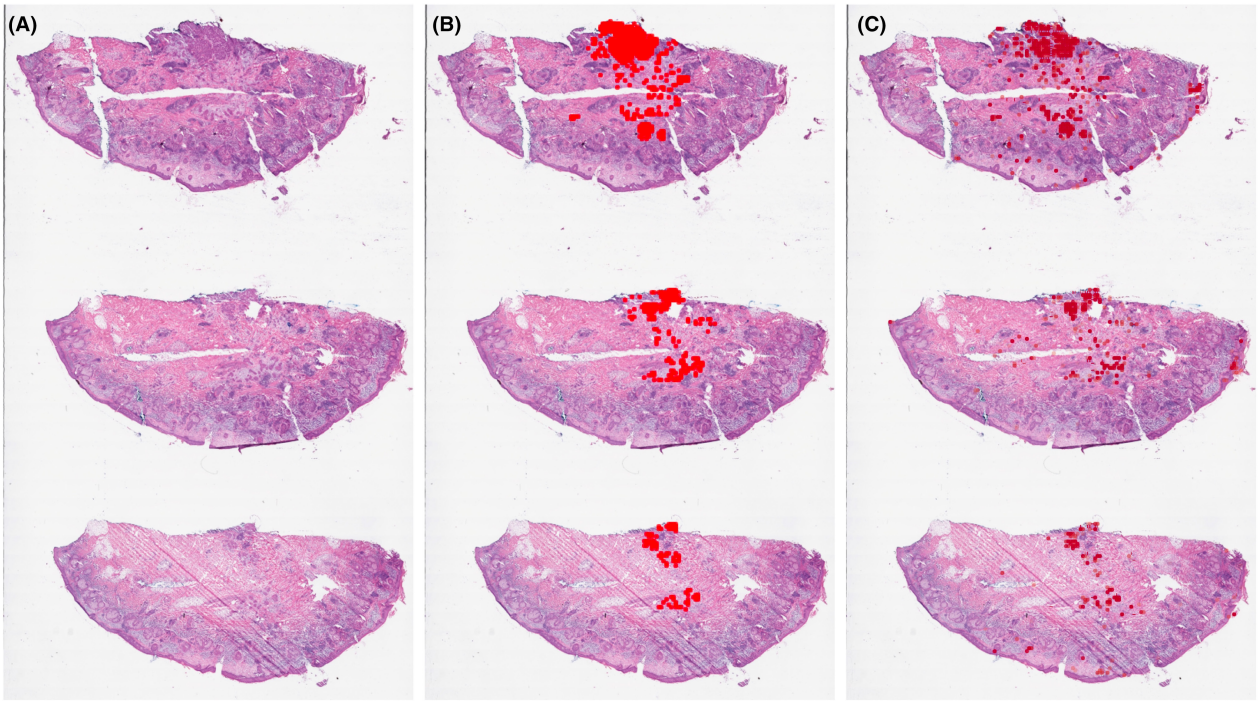

FIGURE 1 Example display output of cutaneous squamous cell carcinomas (cSCC) prediction probabilities at 50-micron resolution for example test-set WSI: (A) original WSI; (B) ground truth cSCC; (C) cSCC algorithm predictions.

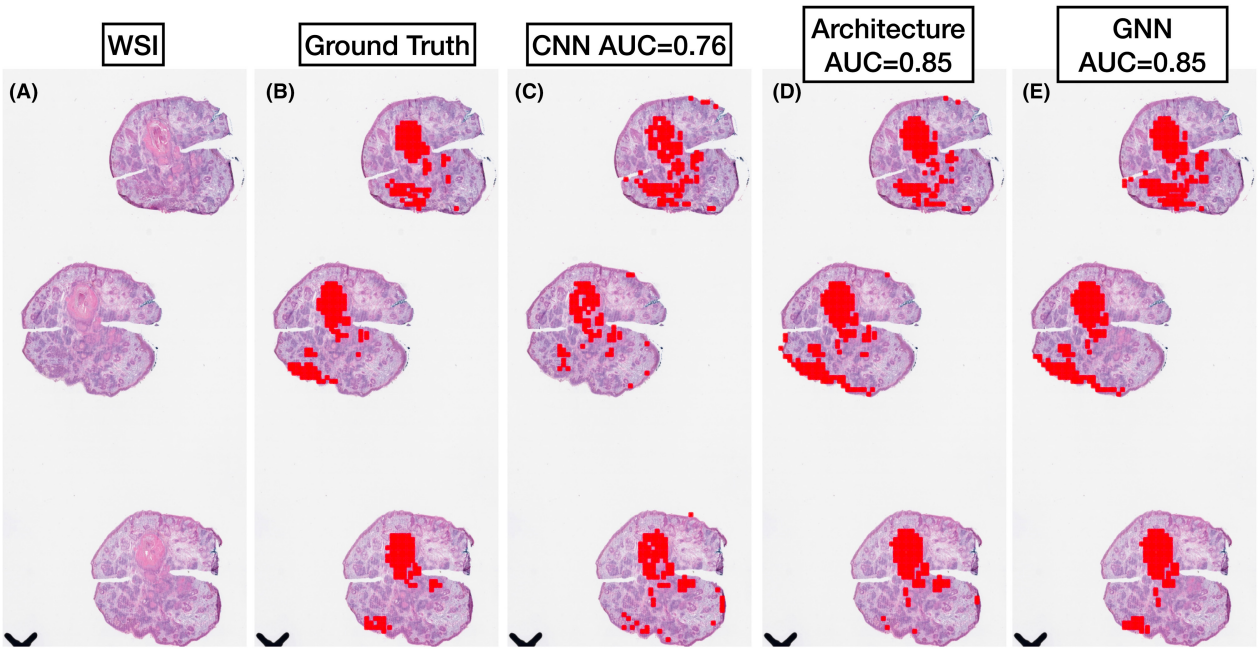

FIGURE 2 Example display output of cutaneous squamous cell carcinomas (cSCC) prediction probabilities at 50-micron resolution for cSCC/Epithelium predictions across example test-set WSI for well-differentiated tumour: (A) original WSI; (B) ground truth cSCC; (C) cSCC convolutional neural network (CNN) algorithm predictions (histomorphology); (D) topological and shape features (architecture); (E) graph neural network (GNN) predictions (contextual).

5 | CONCLUSION AND PERSPECTIVES

Our study provides an example of a deep learning algorithm used to identify cSCC on frozen section slides in MMS. Reducing rate limiting steps to intraoperative margin assessment of cSCC tumours can improve the efficiency and completeness of tumour removal, reducing the burden on laboratory staff while reducing tumour recurrence and repeat procedures.[17–19] 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 artefacts 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.

While slides containing incidental diagnoses such as actinic keratosis and SCC in situ were not excluded from the analysis, our algorithms were not explicitly trained to identify these diagnoses. In the future, the focus will be on refining and improving the algorithm's accuracy to enable more detailed identification of various associated histopathologic features, including single cell analysis, follicles, actinic keratosis and incidental diagnoses (e.g. SCC in situ). Additionally, efforts will be made to map tumours to their original anatomical position/orientation and evaluate the efficiency improvements and cost benefits of this algorithmic approach. Thorough evaluation of implementation barriers and indicators of cost-effectiveness will motivate integration of these algorithms into the clinical workflow.[20] Such studies necessitate a detailed analysis of existing bottlenecks, guiding us to the most suitable places for integrating AI technologies. For example, previous studies have shown that suboptimal specimen quality, like fragmented tissue samples, has limited impact on identifying positive margins. This is partly because it signals the Mohs surgeon to possibly extract additional sections from the tissue block or resected tissue for a more comprehensive evaluation.[6,16] However, cutting additional tissue sections could introduce further bottlenecks when using AI tools for intraoperative assessment. Yet, some prior research has suggested that AI might expedite the decision on when to section blocks.[6,21,22] The prevalence of low-quality sections can differ among institution, thus influencing their impact. This variability warrants further investigation.

Moving beyond frozen section slides, future research may also be focused on the use of machine learning algorithms in non-invasive diagnostic techniques, such as reflectance confocal microscopy (RCM). Such advanced techniques require significant training to attain sufficient diagnostic accuracy to be practically useful, but with advances in machine learning applications, RCM may be more widely accessible. Nascent AI techniques are capable of generating H&E-like digital images from RCM images—these synthetic images could be trained and assessed using the algorithms featured in this work, potentially obviating the need for frozen sections. A recent review article assessing the advancements in this space and potential future research highlights the accuracy and promise of machine learning algorithms applied to RCM.[23] However, similar to the use of machine learning algorithms for assessment of frozen section slides, additional research and randomized controlled trials using the technology are required before widespread application is possible.

This study not only established the general feasibility of histomorphological cSCC detection incorporating AI algorithms, 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 tumours, although further research is necessary to improve the ability to utilize spatial cues. Furthermore, different tumour types may necessitate different algorithms. Future research will also address other confounding histopathologic features, such as inflammation, nuclei, follicles, architecture and keratinocyte differentiation, by considering nuclei and large-scale architectural features.

AUTHOR CONTRIBUTIONS

Matthew R. LeBoeuf and Joshua J. Levy designed the research study. Victoria C. Torres, Sassan Hodge, Eunice Y. Chen, Kimberley S. Samkoe, Rachael Chacko, Matthew J. Davis, Louis J. Vaickus, and Matthew R. LeBoeuf collected the data. Matthew J. Davis, Gokul Srinivasan, Rachael Chacko, Sophie Chen, and Anish Suvarna performed the research and analysed the data. Matthew R. LeBoeuf, Joshua J. Levy, Brock C. Christensen, Sara Preum, and Louis J. Vaickus provided project oversight and mentorship. All authors have read and approved the final manuscript.

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 third in the Computer Science competition.

FUNDING INFORMATION

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

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Anish Suvarna https://orcid.org/0000-0002-6918-7454
Joshua J. Levy https://orcid.org/0000-0001-8050-1291

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*. 2022. 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*. 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: van de Weygaert R, Vegter G, Ritzerveld J, Icke V, eds. *Tessellations in the Sciences: Virtues, Techniques and Applications of Geometric Tilings*. Springer; 2011.
14. Li MM, Huang K, Zitnik M. Graph representation learning in biomedicine and healthcare. *Nat Biomed Eng*. 2022;6:1-17.
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*. 2021;26:285-296.
16. Murphree DH, Kim Y, Sidey KA, Comfere NI, Vidal NY. Artificial intelligence for automatic detection of basal cell carcinoma from frozen tissue tangential biopsies. *Clin Exp Dermatol*. 2023;llad209. Advance online publication. doi:10.1093/ced/llad209
17. 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
18. Otley CC, Salasche SJ. Mohs surgery: efficient and effective. *Br J Ophthalmol*. 2004;88(9):1228.
19. Chacko R, Davis MJ, Levy J, LeBoeuf M. Integration of a deep learning basal cell carcinoma detection and tumor mapping algorithm into the Mohs micrographic surgery workflow and effects on clinical staffing: a simulated, retrospective study. *medRxiv*. 2023. doi:10.1101/2023.05.13.23289947
20. Li RC, Asch SM, Shah NH. Developing a delivery science for artificial intelligence in healthcare. *Npj Digit Med*. 2020;3(1):1-3. doi:10.1038/s41746-020-00318-y
21. Gorman BG, Lifson MA, Vidal NY. Artificial intelligence and frozen section histopathology: a systematic review. *J Cutan Pathol*. 2023;50(9):852-859. doi:10.1111/cup.14481
22. Sohn GK, Sohn JH, Yeh J, Chen Y, Jiang SIB. A deep learning algorithm to detect the presence of basal cell carcinoma on Mohs micrographic surgery frozen sections. *J Am Acad Dermatol*. 2021;84(5):1437-1438. doi:10.1016/j.jaad.2020.06.080
23. Malciu AM, Lupu M, Voiculescu VM. Artificial intelligence-based approaches to reflectance confocal microscopy image analysis in dermatology. *J Clin Med*. 2022;11(2):429. doi:10.3390/jcm11020429

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Data S1.

How to cite this article: Davis MJ, Srinivasan G, Chacko R, et al. A deep learning algorithm to detect cutaneous squamous cell carcinoma on frozen sections in Mohs micrographic surgery: A retrospective assessment. *Exp Dermatol*. 2023;00:1-6. doi:10.1111/exd.14949