Using Deep-Learning Algorithm to Detect Basal Cell Carcinoma and Squamous Cell Carcinoma on Frozen Mohs Micrographic Surgery Sections

Zhuoqing S. Gellatly MS[1], Bichchau Michelle Nguyen MD, MPH[2]
Tufts University School of Medicine, Boston, MA, USA[1], Tufts Medical Center, Boston, MA, USA[2]

Abstract

Mohs micrographic surgery (MMS) is a surgical technique used to treat specific types of skin cancers that ensures maximum tissue sparing and high cure rates. In order to increase accuracy and efficiency of identifying basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) on MMS frozen sections, we aim to develop a deep-learning algorithm using Vision Transformer (ViT), a computer vision model, to classify normal skin tissue, BCC, and SCC on digitized MMS slides. High-quality digital images of MMS slides from 20 cases (10 cases of BCC and 10 cases of SCC) were taken and patched before randomly split into train, test, and validation sets to fine-tune a ViT-L32 model. Model performance evaluation showed an overall accuracy of 95.8% and F1-scores of 0.97, 0.90, and 0.96 for normal skin tissue, BCC, and SCC, respectively. This research holds significant implications for MMS practice, streamlining the review process of MMS slides, and improving quality assurance for Mohs laboratories. The potential of artificial intelligence (AI) in dermatopathology and skin cancer detection is promising, and this project will contribute to the advancement of AI-assisted diagnosis and treatment in dermatology.

Introduction

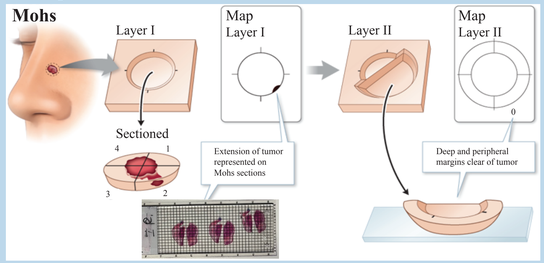
- Mohs micrographic surgery (MMS) is a surgical technique that maximizes tissue-sparing as well as cancer cure rate.
- Basal cell carcinoma (BCC) and squamous cell carcinoma (SCC) are the two most common types of skin cancers treated with MMS. Although recurrence rates are low, there is a need for continuing quality assurance and improvement.
- To comply with CLIA regulations, Mohs surgeons must send out randomly selected Mohs slides to another Mohs surgeon or dermatopathologists for blinded review.
- The application of deep learning algorithms has the potential of not only increasing the efficiency of testing for CLIA purposes but also standardization of the review process.
- The application of AI in dermatology has mainly revolved around dermoscopy, however, studies have shown the potential of deep learning algorithm use to pre-screen histopathology slides.
- Currently published studies employed deep-learning algorithms to specifically identify one type of skin cancer, BCC or SCC, and the models used were convolutional neural network (CNN) and graph neural network (GNN).
- A deep learning algorithm using ViT model that can be used for detection of both BCC and SCC may prove to be more accurate and useful.

Citations

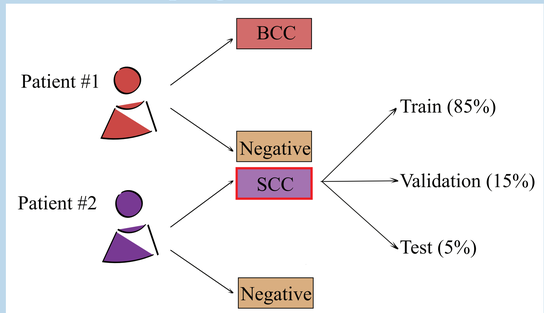
- Bonnefille, H., Abbas, M., Roger, P., Habib, F., Masset, F., Chaumont, M., Subsol, G., & Stoebner, P. E. (2023). A deep-learning algorithm to localize basal cell carcinoma foci on Mohs surgery frozen sections. Journal of the European Academy of Dermatology and Venereology: JEADV. https://doi.org/10.1111/jdv.18859
- Campanella, G., Nehal, K. S., Lee, E. H., Rossi, A., Possum, B., Manuel, G., Fuchs, T. J., & Busam, K. J. (2021). A deep learning algorithm with high sensitivity for the detection of basal cell carcinoma in Mohs micrographic surgery frozen sections. Journal of the American Academy of Dermatology, 85(5), 1285–1286.
- Davis, M. J., Srinivasan, G., Chacko, R., Chen, S., Suvarna, A., Vaickus, L. J., Torres, V. C., Hodge, S., Chen, E. Y., Preum, S., Samkoe, K. S., Christensen, B. C., LeBoeuf, M., & Levy, J. J. (2023). A deep learning algorithm to detect cutaneous squamous cell carcinoma on frozen sections in Mohs micrographic surgery: a retrospective assessment. medRxiv : The Preprint Server for Health Sciences. https://doi.org/10.1101/2023.05.14.23289960
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. https://openreview.net/pdf?id=YicbFdNTTy
- Sohn, G. K., Sohn, J. H., Yeh, J., Chen, Y., & Brian Jiang, S. I. (2021). A deep learning algorithm to detect the presence of basal cell carcinoma on Mohs micrographic surgery frozen sections. Journal of the American Academy of Dermatology, 84(5), 1437–1438.
- Behshad R (2018). Mohs micrographic surgery. Kantor J(Ed.), Dermatologic Surgery. McGraw Hill. https://accessmedicine-mhmedical-com.ezproxy.library.tufts.edu/content.aspx?bookid=2811&sectionid=245221672

Methods

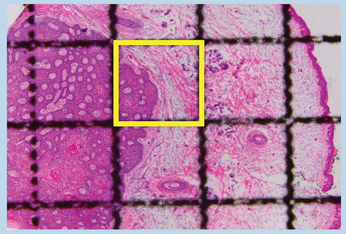
Sample collection



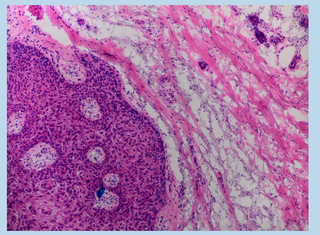
Stratified sampling



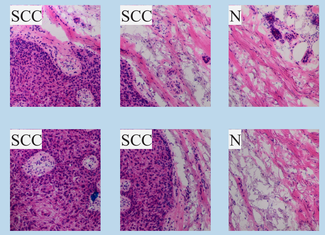
Data collection



2x magification

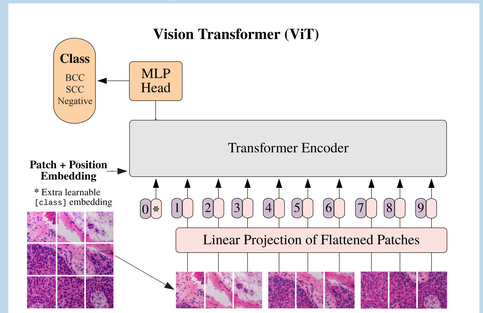


10x magnification

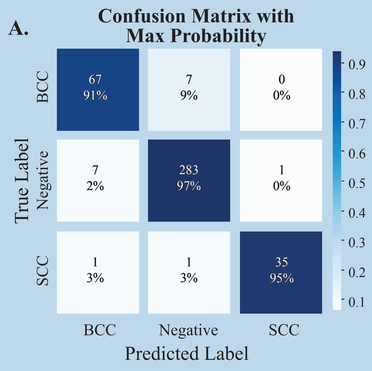


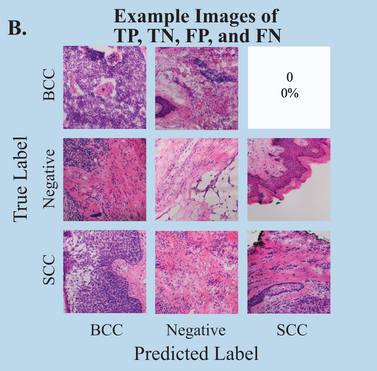
Patching and labeling

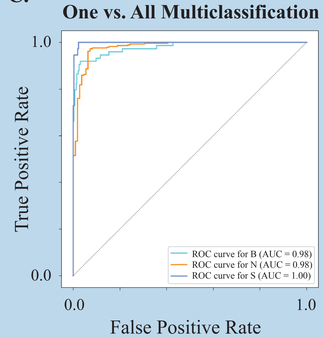
Model training

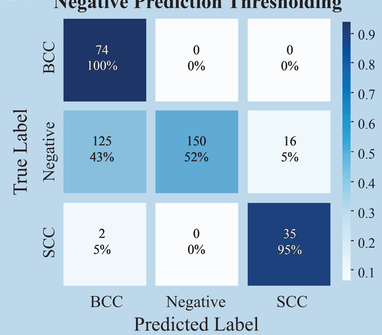


Results









A. Confusion matrix showing model predictions on the test data using the highest probability of each class.
- Overall accuracy: **96%**
- Precision: 95%
- Recall: 94%
- F1 score: 94%

B. Confusion matrix with example images of true positives, false positives, true negatives, and false negatives.

C. Receiver operating characteristic (ROC) curve showing the model's performance for each class in an one vs. all multiclassification setting.

D. Confusion matrix showing the model predications using a threshold on negative predictions (99%) so that there were no false negatives.
- Overall accuracy: **64%**
- Precision: 98%
- Recall: 64%
- F1 score: 66%

Discussion

- Model was trained to classify input images into either SCC, BCC, or normal tissue.
- Model used (ViT) operates on the whole image utilizing spatial context that a traditional CNN does not take into account.
- Model was trained on a dataset collected from 20 cases (n = 8139); however, a large portion of the dataset is of normal tissue (72%). Increasing the number of cases could improve accuracy for both BCC and SCC.
- The current model uses classification to label patches within the input image, however segmentation could also prove to be a powerful method.
- Instead of individual images, Whole Slide Imaging could be used to create high throughput data annotation labels.
- Third-party validation could be used to further assess model accuracy.
- This analysis could be expanded to include mixed tumor types.