Automated nerve identification in histopathology slides enables comprehensive analysis of innervation in cancer and tumor neurobiology

Alison R. Miller[1], Dan Krasnonosenkikh[1], Monica Thanawala[1], Kai-Chih Huang[1], George V. Thomas[2], Alexandra B. Lantermann[1], Hongyue Dai[1], A. Masoud Sadaghiani[1], John A. Wagner[1], Pearl Huang[1]

[1]Cygnal Therapeutics, [2]Department of Pathology & Laboratory Medicine, Knight Cancer Institute, Oregon Health & Science University

ABSTRACT

Linking features of the tumor microenvironment (TME) to patient outcome is ambitious yet crucial to develop new therapies. Histological analysis by a pathologist is the gold standard for diagnosing patients, but does not scale well for discovery biology and population-based studies for drug discovery. Recent studies have shown intriguing relationships between innervation in tumors (tumor exoneural biology) and patient outcomes, but there are no readily-available tools to quantify nerves in H&E images. Here we present an automated tool that can detect and quantify nerve presence in tumors. We manually annotated a set of digital slides from The Cancer Genome Atlas (TCGA) in order to develop a deep learning model to quantify the presence of nerves in head and neck tumors stained with H&E. This tool is generalizable and may be further applied to other structural features, such as vasculature. We used another model to identify tumor-infiltrating lymphocytes (TILs) in and around tumors in digital slides from TCGA. This enables integration of multiple types of image features with multi-omics data to uncover potential biological pathways that are regulated in groups with dense innervation compared to sparse innervation in cancer. This platform-based methodology can be expanded to other disease areas and could ultimately provide valuable insight about exoneural biology and its role in disease physiology to identify new avenues for therapies.

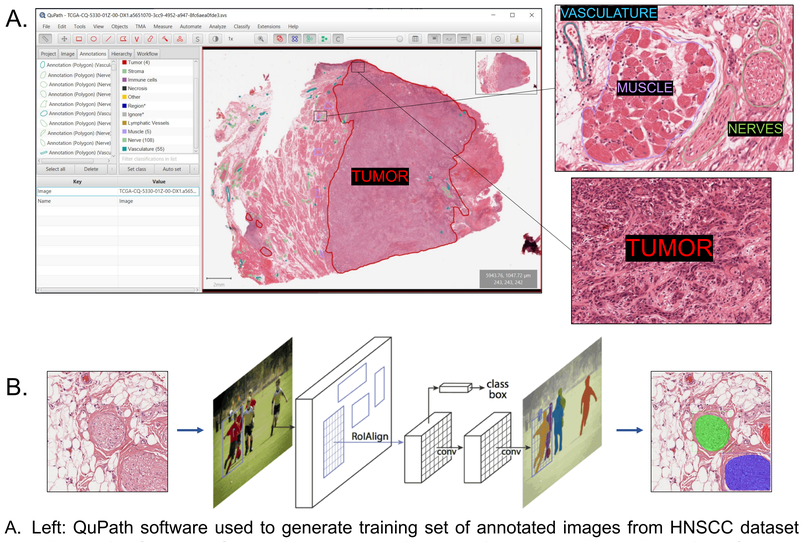
INTRODUCTION

Previous data strongly suggest that nerves in the tumor microenvironment (TME) contribute to cancer proliferation and tumorigenesis, with a greater nerve density often correlating to poorer cancer prognosis[1,2]. Reported investigations into the role of nerves in promoting cancer progression have included methods like nerve ablation and visualization via IHC/IF in prostate cancer, vagotomy suppressing tumorigenesis in pancreatic cancer, and nerve growth factors promoting tumorigenesis in gastric cancer[3,4,5]. While the techniques and cancer indications vary in these studies, the consensus appears to be the same: that across cancer types, nerves play an important role in facilitating tumor growth. Another prior study made use of hundreds of biopsy images stained with hematoxylin and eosin (H&E) for head and neck squamous cell carcinoma (HNSCC) available from The Cancer Genome Atlas (TCGA) and reported that patient survival was inversely correlated with nerve density in their corresponding biopsy image[6]. Despite the many published approaches to visualizing nerves in the TME, this previous study is unique in that it makes use of the thousands of H&E images across cancer types available from TCGA by correlating information within the image data to patient clinical data also available[7]. Because of this promising result, and the fact that H&E images are commonly used and publicly available on multiple databases, the initial training set we used in the current study to develop our machine-learning model was comprised of H&E images from the HNSCC dataset. This dataset served as a good starting point, and we intend to translate the nerve model to other cancer indications. With the ability to rapidly gather data about the presence of features in the TME and their relationships to clinical data across entire datasets, future research can be expanded to understanding the interaction of these features with cancer proliferation and patient outcomes.

OBJECTIVE & HYPOTHESIS

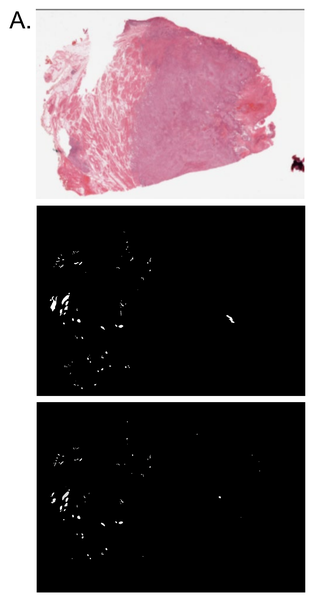
1. **Objective**: To automate feature detection and quantification in H&E images.
2. **Hypothesis**: Human-generated annotations of structures of interest will adequately train a machine-learning model for unbiased detection of these structures in H&E images.

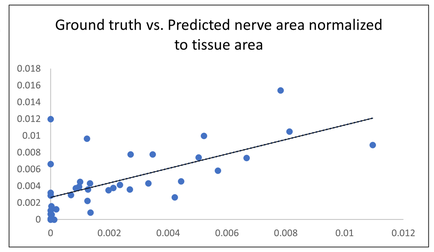
MATERIALS & METHODS



A. Left: QuPath software used to generate training set of annotated images from HNSCC dataset with regular feedback from a licensed pathologist until accurate. Right: Examples of nerve, muscle, vasculature, and tumor annotations.
B. Annotations were used as input for Mask R-CNN model (center) to generate segmented masks (right).

QUANTIFYING THE ACCURACY OF NERVE IDENTIFICATION MODEL





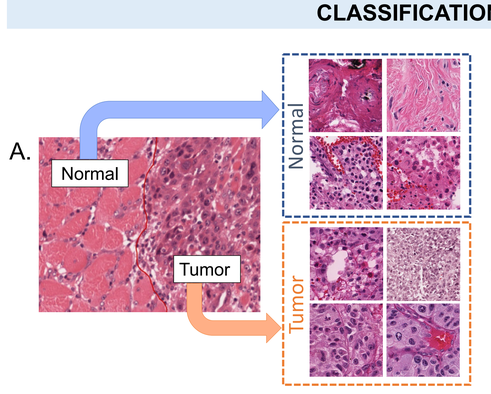
C.

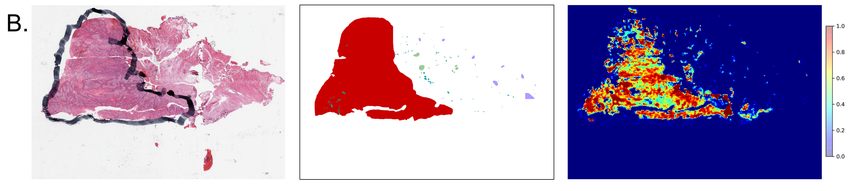
$$Precision = \frac{TP}{TP + FP} \qquad Recall = \frac{TP}{TP + FN}$$

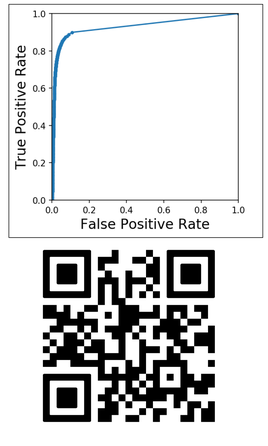
TP = True positive
TN = True negative
FP = False positive
FN = False negative

A. Top: Whole slide example H&E image TCGA-CQ-5330 from HNSCC dataset. Middle: Binary mask showing ground truth nerve annotations from top of panel A. Bottom: Binary mask showing model prediction output from top of panel A.
B. Correlation between ground truth nerve areas and predicted nerve areas from HNSCC training set. Full slide nerve area correlation is 0.65. Full slide pixel-based average precision is 0.341 and recall is 0.256.
C. Equation for calculating precision and recall values to quantify model accuracy[9].

TUMOR IDENTIFICATION IN HNSCC USES PATCH-BASED CLASSIFICATION

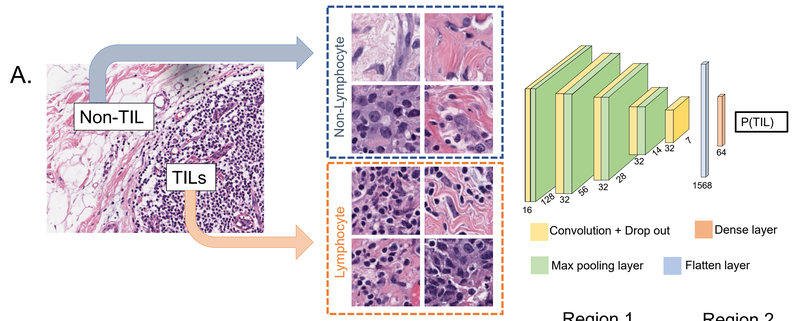


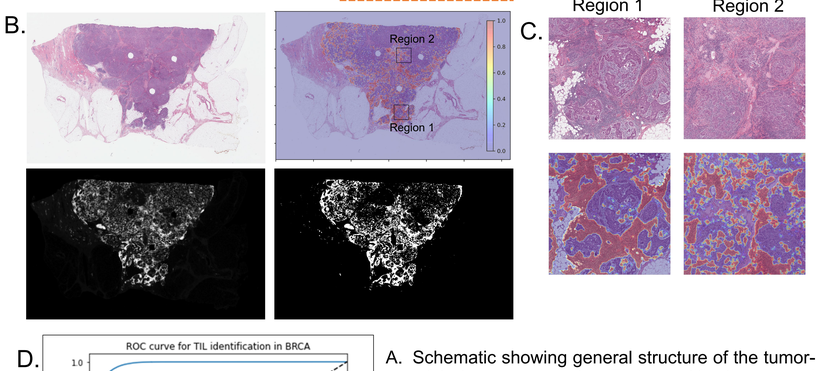


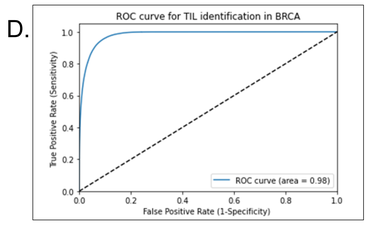


A. Schematic showing general structure of the tumor identification model using patch-based classification. Patch-based classification allows us to extract features that cover increasingly large areas, starting with a fine resolution that detects individual cells and then repeatedly down-sampling to detect multicellular features. This method results in areas that have a certain probability of being the feature of interest, which we manually threshold in order to maximize detection accuracy.
B. Left: whole slide example H&E image TCGA-P3-A6T0 from HNSCC dataset. Center: Ground truth annotations, using red for tumor area. Right: Prediction heatmap where +1.0 is high confidence and 0.0 is low confidence that an area is tumor. Threshold was set to +0.5.
C. ROC curve to quantify tumor prediction accuracy. Area under curve is 0.93. Full slide average precision is 0.846, recall is 0.701.

IMMUNE CELL IDENTIFICATION IN BRCA USES PATCH-BASED CLASSIFICATION







A. Schematic showing general structure of the tumor-infiltrating lymphocyte (TIL) identification model using patch-based classification.
B. Top left: Whole slide example H&E image TCGA-E2-A14P from BRCA dataset. Top right: Probability mask where +1.0 is high confidence and 0.0 is low confidence that an area is TILs. Bottom left: Grayscale ground truth mask from convolutional neural network model presented in Saltz et al. 2018[10]. Bottom right: Grayscale probability map from our model using +0.5 threshold.
C. Insets from top right panel of image B showing probability map overlay (bottom).
D. ROC curve to quantify TIL prediction accuracy. Area under curve is 0.98. Full slide average precision is 0.96, and recall is 0.85.

DISCUSSION

Here we present an automated tool to detect and quantify certain features of the TME. A few limitations remain with this model, including the need to build specialized models for each new cancer indication, but this is easy to overcome with a small number of manual annotations. In the future, this model can be expanded to other features of interest, such as vasculature or non-cancerous glands, in order to independently investigate the presence of these features in the TME and their role in cancer progression. Feature detection can further be used to correlate patient proteomic and genomic data on TCGA and other H&E image databases to the presence, absence, or pattern of features in the TME in order to increase understanding of how these features contribute to tumor proliferation and patient outcome. Because H&E staining is routine, staining and automated analysis can be applied to new patient samples to detect features in tissues related to other diseases of interest where nerves, vasculature, or immune cells may influence disease progression.

CONCLUSIONS

1. We demonstrated highly predictive deep-learning models to detect tumor and TILs in HNSCC and BRCA respectively.
2. We demonstrated a moderately predictive deep-learning model to detect nerves in HNSCC and identified further methods for improvement.
3. Mask R-CNN and patch-based classification can be used to identify diverse structures of interest in tumor biopsies and can be expanded in the future to identify other structures of interest using a training set of hand-annotations.

REFERENCES

1. Monje M, Borniger JC, D'Silva NJ, et al. Roadmap for the emerging field of cancer neuroscience. Cell. 2020;181(2):219-222.
2. Zahalka AH, Frenette PS. Nerves in cancer. Nat Rev Cancer. 2020;20(3):143-157.
3. Magnon C, Hall SJ, Lin J, et al. Autonomic nerve development contributes to prostate cancer progression. Science. 2013;341(6142):1236361.
4. Renz BW, Tanaka T, Sunagawa M, et al. Cholinergic signaling via muscarinic receptors directly and indirectly suppresses pancreatic tumorigenesis and cancer stemness. Cancer Discov. 2018;8(11):1458-1473.
5. Hayakawa Y, Sakitani K, Konishi M, et al. Nerve growth factor promotes gastric tumorigenesis through aberrant cholinergic signaling. Cancer Cell. 2017;31(1):21-34.
6. Amit M, Takahashi H, Dragomir MP, et al. Loss of p53 drives neuron reprogramming in head and neck cancer. Nature. 2020;578(7795):449-454.
7. GDC. Cancer.gov. Accessed 2020-2021. https://portal.gdc.cancer.gov/
8. He K, Gkioxari G, Dollar P, Girshick R. Mask R-CNN. In: 2017 IEEE International Conference on Computer Vision (ICCV). IEEE; 2017.
9. Hui J. MAP (mean average precision) for object detection - Jonathan Hui - medium. Medium. Published March 7, 2018. Accessed September 7, 2021. https://jonathan-hui.medium.com/map-mean-average-precision-for-object-detection-45c121a31173
10. Saltz J, Gupta R, Hou L, et al. Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. Cell Rep. 2018;23(1):181-193.e7.