SRTP - Project Description Form #213

PART I:

Name of Schulich faculty member who will supervise the project	Matthew Cecchini
Supervisor's Schulich, Western, Hospital or Lawson Email	matthew.cecchini@lhsc.on.ca
Schulich Department	Pathology & Laboratory Medicine

PART II - Project Description

Title of Project	Developing a machine learning approach to detect malignant lymph nodes by ultrasound at the time of gross examination.

Background

In the macroscopic examination of resected pathology specimens, meticulous dissection of pericolic soft tissue is undertaken for the identification of lymph nodes (LNs). Lymph nodes serve as highly organized, encapsulated conglomerates of immune cells, and function as the primary filtration points within the lymphatic system, responsible for the clearance of toxins and other antigens. Given their pivotal role as the sentinel stations for cancer metastasis, the status of LNs remains instrumental in cancer staging and informs the subsequent therapeutic management. Further in some tumor types such as colorectal cancer the location of involved lymph nodes in relation to the surgical margin is critical, however, this can be challenging in the current systems which utilize manual palpation and dissection to identify lymph nodes. The macroscopic identification of all LNs poses a considerable challenge and utilizes significant pathology resources. Ultrasound technology offers a potential rapid and non-invasive modality for the identification of LNs within pericolic soft tissues in cancer specimens, the novel lymph node device developed in my lab in collaboration with Tenomix offers a unique technologic solution to this problem.
Emerging innovations in machine learning, particularly the deployment of transformer-based neural network models, have demonstrated notable efficacy in the field of pathology. Accordingly, we hypothesize that the application of such advanced neural network architectures on ultrasound imaging of cancer tissue would allow not only for the identification of lymph nodes but identifying lymph nodes with a high probably for involvement by cancer. These lymph nodes can then be appropriately sampled in relationship with relevant surgical margins.

Hypothesis

A machine learning system will be able to identify lymph nodes involved by cancer in colorectal cancer specimens.

Proposed Methodology

Colorectal cancer specimens from patients receiving treatment at the London Health Sciences Centre will be utilized for the study. The colorectal soft tissue will be imaged using the non-invasive ultrarouns device in my lab. The specimen will then be processed according to standard lab protocols. The blocks corresponding to the identified lymph nodes we be recorded. The histology slides of the cases will be reviewed after sign out to map the involved lymph nodes to the imaging data. This will be used as the training dataset. The model will be developed using the TensorFlow library in Python. Specifically, we will utilize a vision transformer model using TensorFlow and Keras. This model will be pre-trained on publicly available ultrasound data such as the Learn2Reg dataset. Subsequently, this pre-trained model will be applied to the ultrasound scans generated by this study. Model training will be done using a high-performance computing cluster (>16GB of GPU). A subset of cases will be used held back for the validation datasets to test the performance of the model.

Expected Outcomes

We expect to scan 50 patients over the next two years with between 12 and 20 lymph nodes in each case. Based on performance metrics from other studies that have applied transformer-based neural network models to ultrasound imaging, we expect our model to attain an accuracy between 80-95% for the detection of cancer in lymph nodes. We expect that the model will perform best with larger metastatic lesions. We will also correlate with foci of tumor deposits and integrate these into the training model. An expansion of the project may be to distinguish between tumor deposits and lymph nodes on the ultrasound imaging.

Research Environment – Description of the number of research personnel, primary location of research, size of lab, etc

This is a collaborative project between pathology medical biophysics and an industrial collaborator.

Names and titles of other individuals who will be involved with the research project?

N/A

Can this project be done remotely?	Yes
Duration of Project	Two Summers

Expected Objectives/Accomplishments for Student for Year 1?

Collect the specimens and data in year 1
Start to build the pipeline and infrastructure for the machine learning model
Learn the basics of to deploy these tools

Expected Objectives/Accomplishments for Student for Year 2?

Develop and refine the machine learning model to detect tumor in lymph nodes.
Finalize data collection
Write up research study for publication.

PART III - Certifications

If the project will require any certification approvals from one or more of the following offices, please check the appropriate box below.	- Human Ethics
Human Ethics: If you have the protocol information, please enter it below (or enter the status of the approval).	119579

Note: certification approval should be obtained prior to the start of the summer. Projects without this approval will not be a priority for funding.