CALIFORNIA SCIENCE & ENGINEERING FAIR
2018 PROJECT SUMMARY

Name(s)

James L. Wang

Project Number

38752

Project Title

Improving Mitosis Detection and Localization in Breast Cancer Histology Using Deep Convolutional Neural Networks

Abstract

Objectives/Goals

To (1) create an enhanced machine learning model for classification of mitosis in breast histology through transfer learning on pre-trained deep convolutional neural networks and (2) to develop a more effective and accurate algorithm for localizing instances of mitosis in breast histology instead of using recurrent convolutional neural networks.

Methods/Materials

Open-source histology slides were first obtained from the MITOS-ATYPIA 14 Grand Challenge These slides contained 391 instances of mitosis and 741 instances of false-positives, which were extracted through image segmentation techniques. Data augmentation techniques were then used to artificially augment the dataset to 3,000 images, with 1,500 in each class. These images were then fed through a convolutional neural network for training using MATLAB and an NVIDIA GTX 1060 graphical processing unit. Using an optimal trained model, a selection-search based algorithm was developed to more accurately localize mitosis instances in breast histology.

Results

After training this dataset on different transfer learning implementations on well-known pre-trained convolutional neural networks, it was found that retraining the final fully-connected layer and softmax classifier of the VGG16 model yields the best performance, with a 91.33% classification accuracy on a reserved test set. Our selection-search based algorithm was able to localize most instances of mitosis on histology images; however, some false positive instances were also detected by our algorithm.

Conclusions/Discussion

Our histology analysis pipeline is able to localize and accurately classify instances of mitosis in whole-slide images. While we did find limitations in our localization algorithm, they can be improved by allowing the reading kernel size to be adaptive to the relative sizes of mitotic instances in each whole slide image. Our algorithm design examines the problem of localization with emphasis on accuracy rather than efficiency, which has been explored less in this field of research. The methodology used in this research can be easily generalized and applied to other medical imaging tasks.

Summary Statement

A breast histology analysis pipeline was developed to more accurately detect instances of mitosis using transfer learning on convolutional neural networks and by developing a selection-search based localization algorithm.

Help Received

Breast histology slide images were acquired from the MITOS-ATYPIA-14 Grand Challenge Image Dataset. Neural network transfer learning implementations and algorithm design were done independently without external guidance.