AI4ALL Project Proposal

Authors: Aaron Kim, Ethan Carr, and Michael Tran

October 15, 2024

Project Title

NeuroVision: Automated Neuron Counting Tool

Research Question

Can computer vision techniques be used to automate neuron counting in neuroscience, achieving accuracy and efficiency that match or surpass those of manual methods?

Topic of Interest & Summary

The project focuses on developing an automated neuron counting tool using computer vision to analyze microscope images of brain tissue. This topic is important because manual neuron counting is a time-consuming and error-prone process that can impede neuroscience research. By automating this task, we aim to enhance efficiency and accuracy, enabling researchers to process larger datasets rapidly. This is essential for understanding brain development, studying neurological diseases, and evaluating treatment effectiveness, ultimately contributing to significant progress in neuroscience.

Machine Learning Algorithms

We intend to use convolutional neural networks (CNNs) for image processing and neuron detection, given their proven effectiveness in handling complex image recognition tasks. The dataset we've selected supports both supervised and unsupervised approaches. Initially, we will focus on a supervised approach to train the model with labeled neuron images, and if time permits, we will explore an unsupervised approach to compare and contrast.

Dataset of Interest

We plan to use the Fluorescent Neuronal Cells v2 dataset, which contains a set of microscope images of fluorescently stained mice neurons and corresponding ground-truth labels. Since we are working with CNNs, image resizing and normalization may be necessary.

Bias

1. **Sources:**

   (a) **Species Bias:** Since the dataset includes images predominantly from one species (mice), the model may not generalize well to neurons from other species (e.g., humans).

   (b) **Staining Technique Variability:** Different staining techniques (e.g., Nissl staining vs. immunohistochemistry) produce images with varying contrast and features, potentially confusing the model if not properly accounted for. This can also be synonymous with noise.

   (c) **Neuron Type Overrepresentation:** If the dataset includes more images of a particular neuron type, the model might be biased toward detecting that type over others, reducing its effectiveness in diverse applications.

2. **Mitigation:**

   (a) To mitigate **Species Bias**, ideally, we would be able to mix in other species into the dataset. However, finding a similar dataset of a different species would be challenging.

   (b) To mitigate **Staining Technique Variability**, we can apply consistent image preprocessing (normalization) across all images to reduce differences in contrast and color, which can help the model focus on neuron structure rather than staining variations.

   (c) To mitigate **Neuron Type Overrepresentation**, we can apply class weights to penalize mis-classification of underrepresented neuron types, encouraging the model to learn from all types more equitably.

Citations

1. Best practices for convolutional neural networks applied to visual document analysis

2. Brain tumor segmentation with Deep Neural Networks

3. Convolutional Networks for Biomedical Image Segmentation

4. Dataset: Fluorescent Neuronal Cells v2

5. Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique

6. The Search for True Numbers of Neurons and Glial Cells in the Human Brain: A Review of 150 Years of Cell Counting

7. You Only Look Once: Unified, Real-Time Object Detection