OFFICIAL ABSTRACT and CERTIFICATION

A Comparison of Machine Learning Methods in the Analysis of Lymphocyte Patterns in Cancer Research

Preethi Krishnamoorthy
Herricks High School, New Hyde Park, NY, United States of America

Cancer is a widespread disease that impacts many people. Tumor infiltrating lymphocytes (TILs) are important biomarkers for cancer that can be used for diagnosis and treatment. Current methods of cancer diagnosis, including manual evaluations of tumors by pathologists, are very time consuming. To make this process more efficient and effective, deep learning methods are being used in cancer diagnosis. In this research, we compare the effectiveness of machine learning techniques in classifying lymphocyte probability maps into immune subtypes. First, the dice coefficient was implemented to compare two sets of probability maps generated from two different networks. Then, several machine learning methods were developed—a fully connected network, a CNN, modified pretrained models, Random Forest, SVM, and K means clustering. These methods classified lymphocyte probability maps and cluster indices. Dice values between the sets of probability maps were high, showing good agreement between them. The machine learning methods implemented in this study had an accuracy ranging from 50-60%. Random Forest classification had the highest accuracy, with and accuracy of 63.5%. This research shows a strong proof of concept for future studies, and with a greater amount of data, this research can be continued and higher accuracies can be obtained to bring machine learning models as a tool in mainstream cancer diagnostics and analysis.

Category
Pick one only—
mark an "X" in box
at right

- Animal Sciences ☐
- Behavioral & Social Sciences ☐
- Biochemistry ☐
- Biomedical & Health Sciences ☐
- Biomedical Engineering ☐
- Cellular & Molecular Biology ☐
- Chemistry ☐
- Computational Biology & Bioinformatics ■
- Earth & Environmental Sciences ☐
- Embedded Systems ☐
- Energy: Chemical ☐
- Energy: Physical ☐
- Engineering Mechanics ☐
- Environmental Engineering ☐
- Materials Science ☐
- Mathematics ☐
- Microbiology ☐
- Physics & Astronomy ☐
- Plant Sciences ☐
- Robotics & Intelligent Machines ☐
- Systems Software ☐
- Translational Medical Sciences ☐

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check ALL that apply):

   ☐ human participants ☐ potentially hazardous biological agents

   ☐ vertebrate animals ☐ microorganisms ☐ rDNA ☐ tissue

2. I/we worked or used equipment in a regulated research institution or industrial setting: ■ Yes ☐ No

3. This project is a continuation of previous research. ☐ Yes ■ No

4. My display board includes non-published photographs/visual depictions of humans (other than myself): ☐ Yes ■ No

5. This abstract describes only procedures performed by me/us, reflects my/our own independent research, and represents one year's work only ■ Yes ☐ No

6. I/we hereby certify that the abstract and responses to the above statements are correct and properly reflect my/our own work. ■ Yes ☐ No

This stamp or embossed seal attests that this project is in compliance with all federal and state laws and regulations and that all appropriate reviews and approvals have been obtained including the final clearance by the Scientific Review Committee.