Classifying T Cell Activity with Convolutional Neural Networks

Zijie J. Wang[1,2], Alex J. Walsh[2], Melissa C. Skala[1,2], Anthony Gitter[1,2]

[1]University of Wisconsin–Madison, [2]Morgridge Institute for Research

Introduction

- T cell activity state is important for immunotherapy
- Autofluorescence imaging is label-free and non-destructive
- Use machine learning to classify activated and quiescent T cells with only autofluorescence intensity images
- Results of five classifiers show that advanced models can accurately classify T cell activity

Image Processing

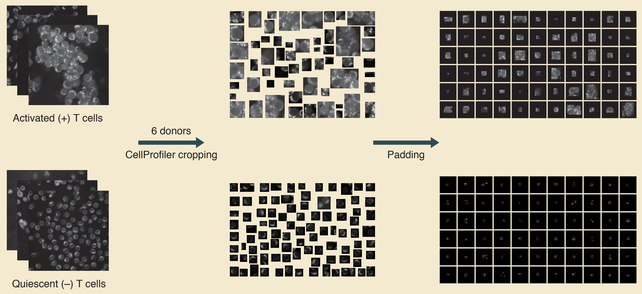
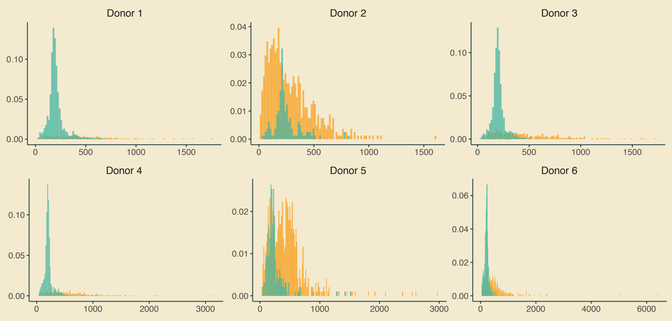


Image Size (Mask Pixel Count) Histogram



Bad Image Filtering

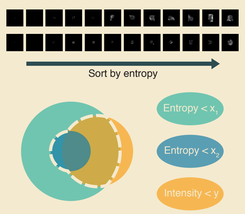


Image Augmentation

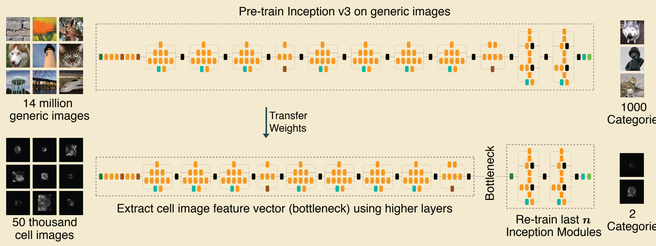
Final Image Count

Donor	Activated	Quiescent
1	1410	9306
2	3882	846
3	2676	7428
4	2892	9414
5	4098	1476
6	2652	3480
Total	17610	31950

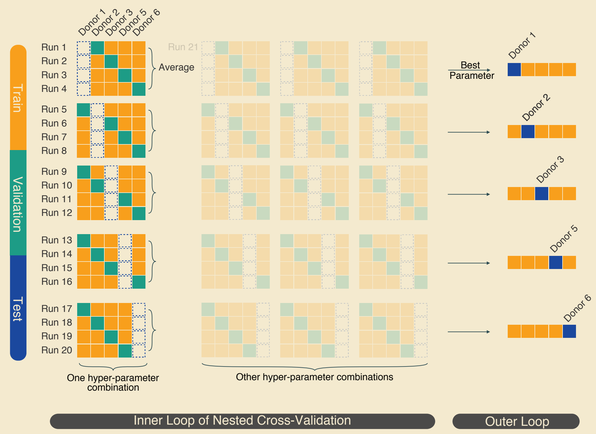
Machine Learning Models

- Binary classification: activated and quiescent
- Various machine learning models of increasing complexity
- Understand when and why deep learning is needed

Model	Description	Hyper-parameter
Frequency Classifier	Predict class probability using the positive frequency in training set	None
Logistic Regression with Pixel Matrix	Regularized logistic regression model fitted with image pixel matrix	L1 penalty power λ
Logistic Regression with Image Intensity and Size	Regularized logistic regression model fitted with two numerical values: total intensity and mask size	L1 penalty power λ
Logistic Regression with CellProfiler Features	Regularized logistic regression model fitted with 123 of Intensity, Texture, and Area features extracted using CellProfiler	L1 penalty power λ
Simple Fully Connected Neural Network	Fully connected one-layer neural network with pixel intensity matrix as input	Num of neurons, learning rate, batch size
Simple Convolutional Network	LeNet network with pixel intensity as input	Learning rate, batch size
Pre-trained CNN with retraining the Last Layer	Retrain the last layer of Inception v3 model	Learning rate, batch size
Pre-trained CNN with retraining the Best Layer	Retrain the last n layers of Inception v3 model	Learning rate, batch size, n

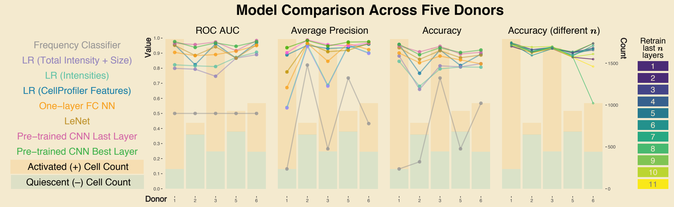


Nested Cross-validation

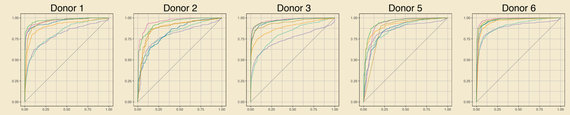


Results and Interpretations

Model Comparison Across Five Donors



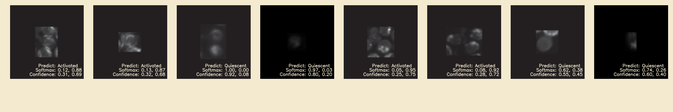
ROC Curves Across Five Donors



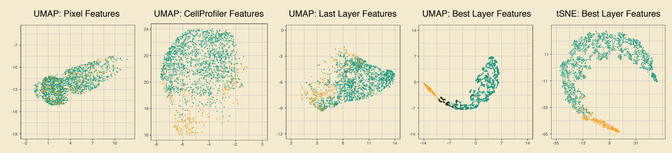
Donor 4 Result

Metric	Value
Accuracy	98.59%
AP	99.67%
AUC	99.89%

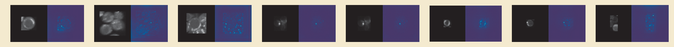
Donor 1: Some Misclassified Images with Temperature Scaled Confidence Score



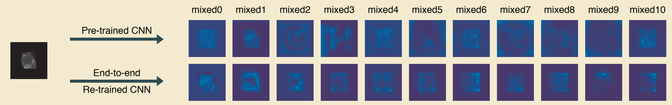
Donor 1: Feature Visualization with Dimension Reduction



Donor 1: Saliency Map of Output Layer

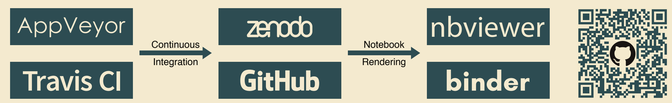


Donor 2: Saliency Map of Inception Layers



Takeaways and Reproducibility

- Advanced models can accurately classify T cell activity state
- Transfer learning outperforms other models
- Performance boost of retraining more layers is not significant
- Release code as Jupyter notebooks: reproducibility and tutorial (https://github.com/gitter-lab/t-cell-classification)



WISCONSIN
UNIVERSITY OF WISCONSIN–MADISON

MORGRIDGE
INSTITUTE FOR RESEARCH