Deep Learning Algorithm for Biomarker Classification on Multiplexed Immunofluorescence Images using Repel Coding

Satarupa Mukherjee[1], Yao Nie[1], Jim Martin[1]

[1]Imaging and Algorithms, Digital Pathology, R&D, RTD, Santa Clara, California

SCAN ME

1 – Introduction

Study of the heterogeneous tumor microenvironment has significant clinical implications. New assays to guide immunotherapy require accurate characterization of interactions between multiple phenotypes (e.g. tumor cells and different types of immune cells) in the tumor microenvironment. Multiplexed immunofluorescence (IF) enables us to understand the complexity and heterogeneity of the immune context of tumor microenvironments and its influence on response to immunotherapies. In this project, we report the development of an automated digital pathology algorithm for biomarker classification using a deep learning approach. The biomarker used in this approach is PD1, which is primarily expressed on lymphocytes. We developed an end-to-end deep learning algorithm to perform both cell detection as well as classification using only point labels. In order to perform better localization of cells, we used repel coding which enhances the center locations of the cells. We applied our algorithm on multiple IF stained tissue types, which includes Gastric, Pancreas, Lung, Breast, Colon, and Bladder.

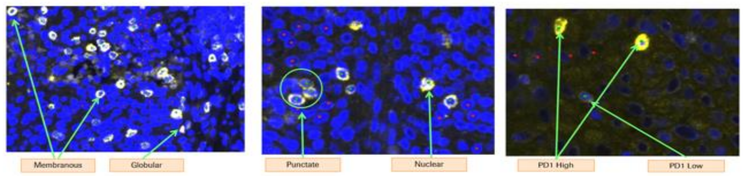
2 – Dataset

- **Pleiades Panel 2 was used as the dataset; it is an immunofluorescent multiplex assay developed by Roche for detecting colocalization of cells expressing one or more of the following: Gzmb, PanCK, PD-1, PD-L1, CD3.**
- **100 field of views (FOVs) of variable sizes were acquired from 15 slides.**
- **The FOVs contained 9800 PD1+ cells and 13000 PD1- cells.**
- **The following table illustrates the distribution of slides over tissue types.**

Tissue Type	No of slides Used Per Tissue Type
Gastric	4
Pancreas	1
Lung	1
Breast	3
Colon	2
Bladder	4

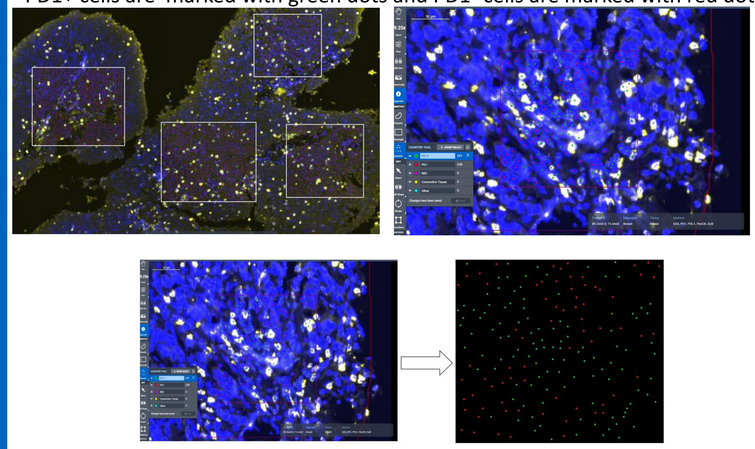
3 – Marker Patterns

The biomarker that has been used for classification is PD1, which is an immune cell marker. This biomarker is primarily expressed on lymphocytes and manifests in numerous patterns including membranous (partial or complete), punctate, nuclear, globular and also combination of these patterns. The biggest challenge in identifying PD1 patterns is that it expresses at a broad range of intensities. The following figure shows different expression patterns of PD1 (figures and patterns approved by pathologist).



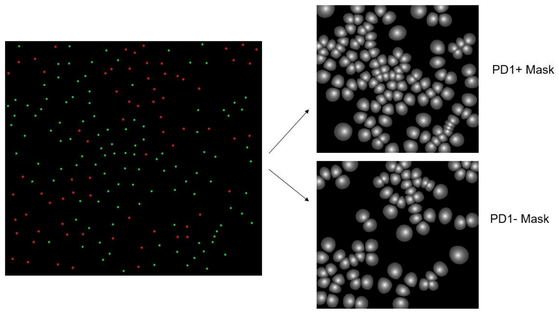
4.1 – Methodology : Groundtruth Collection

A Pathologist has manually annotated each individual cell in selected FOVs through the internal platform dPATH.
The following figures represent screenshots of the annotation procedure.
PD1+ cells are marked with green dots and PD1- cells are marked with red dots.



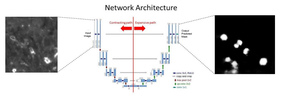
4.2 – Methodology : Groundtruth Label Mask – Repel Coding

- Repel encoding is an enhanced center encoding for cells which defines a 2D decaying function with peak located at the cell center point label.
- Compared to the commonly used Gaussian and proximity encoding, the repel code decays faster for cells which have shorter distance to the neighboring cells.
- Therefore, we exploited the repel code in our algorithm to promote better cell separation and better center localization for the cell identification task.



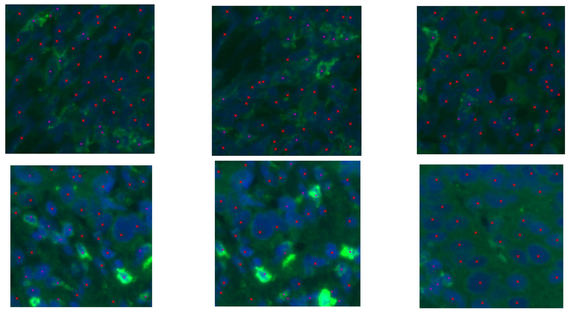
4.3 – Methodology : Model Architecture

A Unet architecture was used with the encoder layers replaced by the convolutional layers of ResNet34.
The input to the network (as shown in figure) was a 6 channel tif image and output was a probability mask.
Among the 100 FOVs, 80 were used for training and 20 for validation.
Recall and Precision were calculated for both training and validation sets.

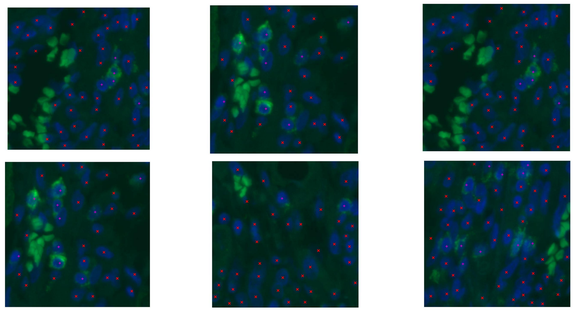


5 - Results

Visual Results (Training Set) : '+' indicates PD1+ and 'x' indicates PD1- cells



Visual Results (Validation Set) : '+' indicates PD1+ and 'x' indicates PD1- cells



	Recall (%)	Precision (%)
PD1+ (Training Set)	94.37	98.74
PD1- (Training Set)	98.12	98.5
PD1+ (Validation Set)	87.45	90.12
PD1- (Validation Set)	86.5	88.37