Flow Cytometry Analysis

- Determining cell types using deep learning techniques
- Creating software to allow streamlined process for flow cytometry analysis

Nathan Wong

The problem

FlowJo

The current standard of analyzing flow cytometry data is the use of the paid application FlowJo, which is very tedious and done manually.

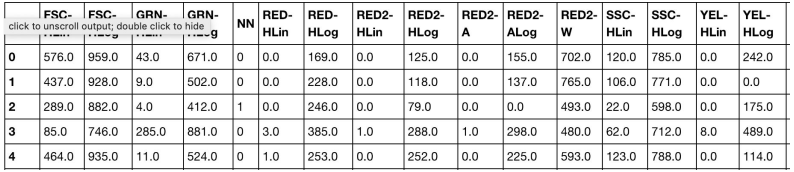
Variability in Data

The method for determining different cell states and types is done by eye, and lacks the capability to determine non-linear patterns. In addition, the multi-channel capabilities of the flow data are not utilized.

Problem statement

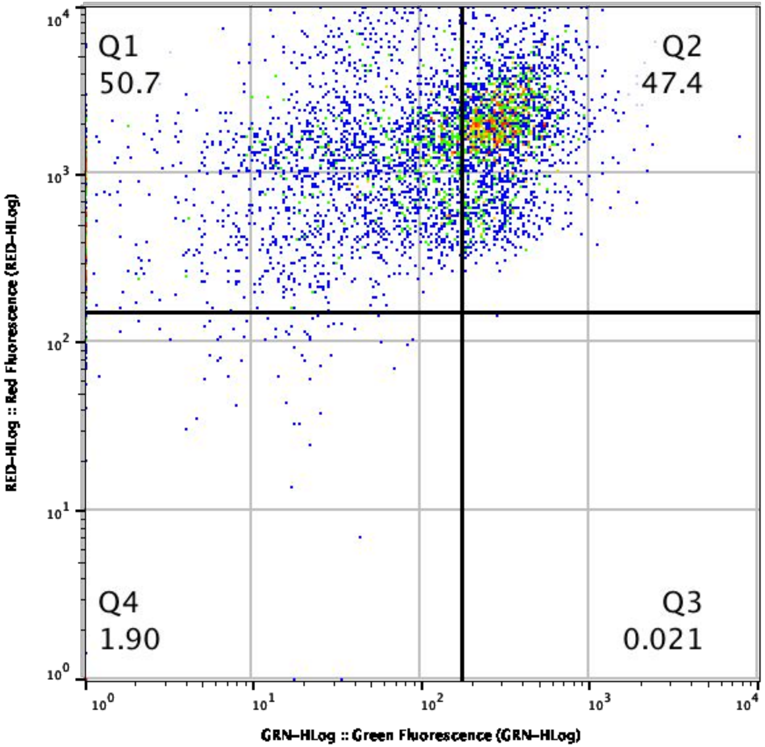
There are significant limitations in the standard ways used to interpret flow cytometry data, and specific learning algorithms can be used to better determine various cell states.

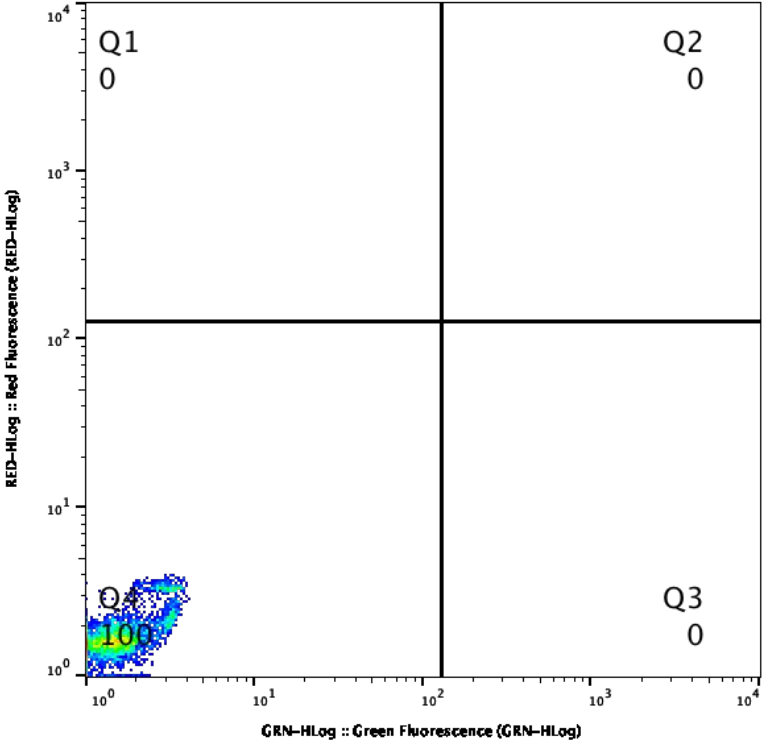
Challenges



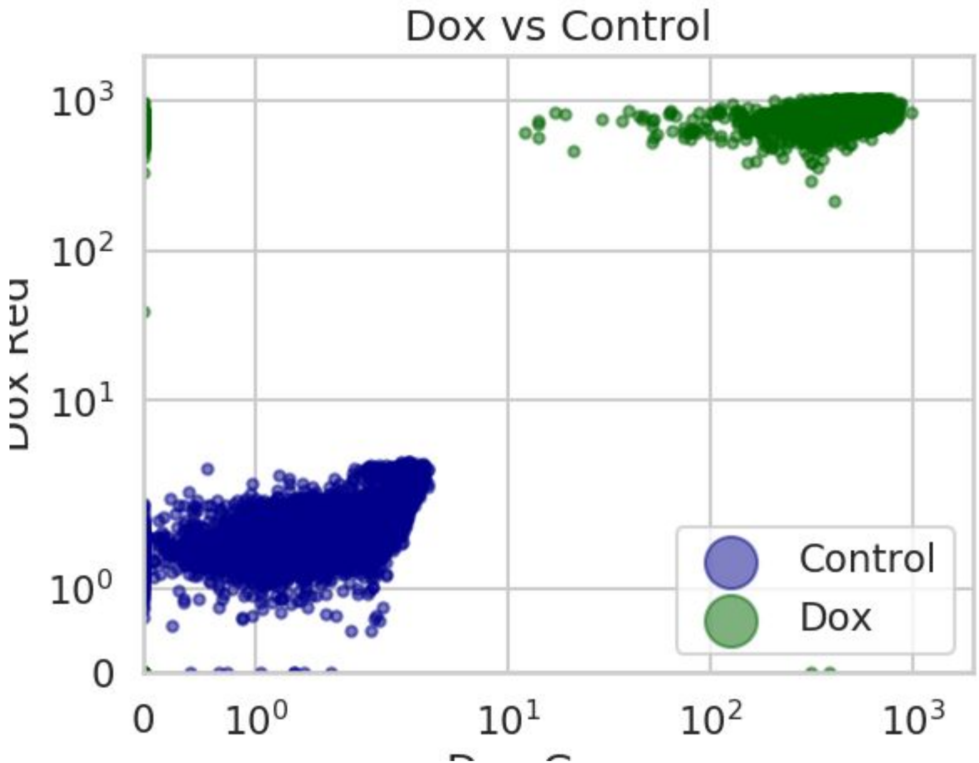
click to unscroll output; double click to hide

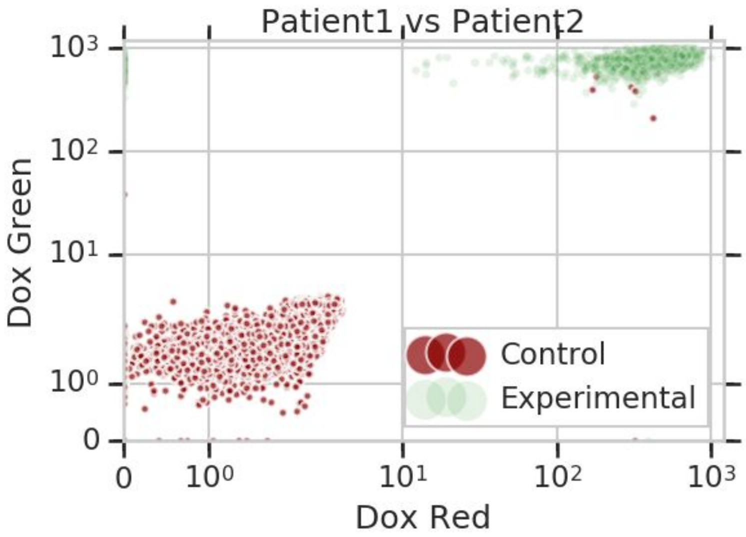
	FSC-HLin	FSC-HLog	GRN-HLin	GRN-HLog	NN	RED-HLin	RED-HLog	RED2-HLin	RED2-HLog	RED2-A	RED2-ALog	RED2-W	SSC-HLin	SSC-HLog	YEL-HLin	YEL-HLog
0	576.0	959.0	43.0	671.0	0	0.0	169.0	0.0	125.0	0.0	155.0	702.0	120.0	785.0	0.0	242.0
1	437.0	928.0	9.0	502.0	0	0.0	228.0	0.0	118.0	0.0	137.0	765.0	106.0	771.0	0.0	0.0
2	289.0	882.0	4.0	412.0	1	0.0	246.0	0.0	79.0	0.0	0.0	493.0	22.0	598.0	0.0	175.0
3	85.0	746.0	285.0	881.0	0	3.0	385.0	1.0	288.0	1.0	298.0	480.0	62.0	712.0	8.0	489.0
4	464.0	935.0	11.0	524.0	0	1.0	253.0	0.0	252.0	0.0	225.0	593.0	123.0	788.0	0.0	114.0



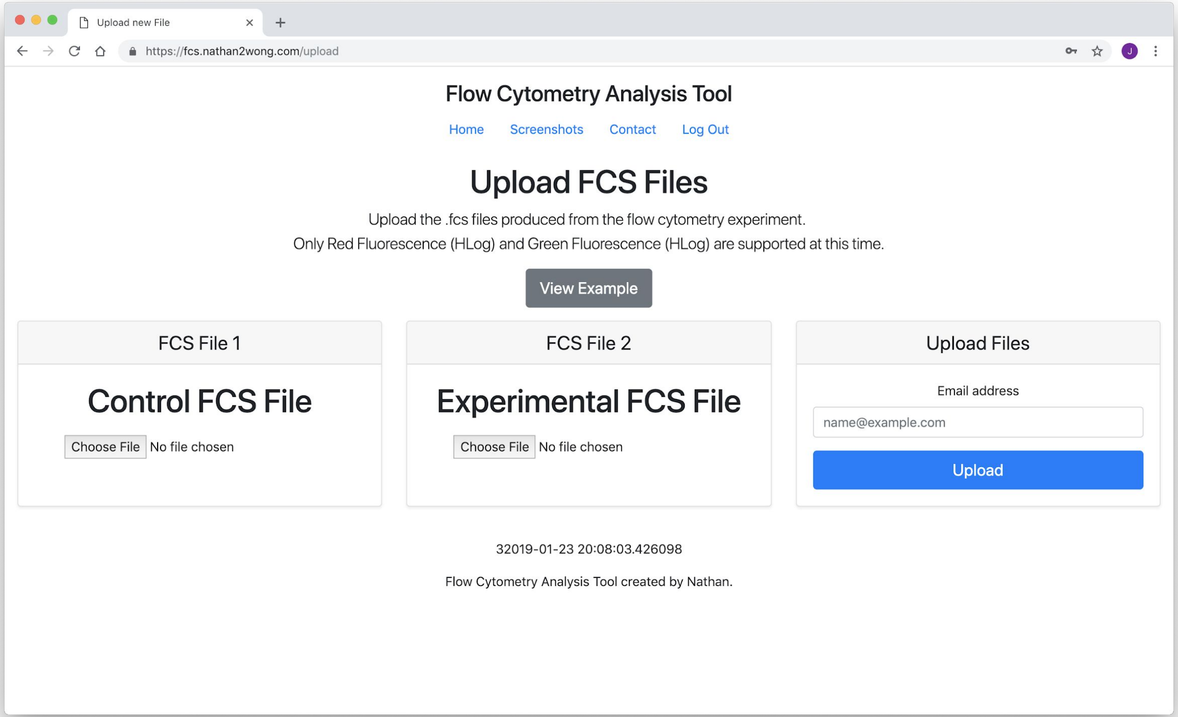


Plotting multiple experiments





Solution: Web application for Flow Cytometry Analysis


Upload new File
https://fcs.nathan2wong.com/upload

Flow Cytometry Analysis Tool

Home
Screenshots
Contact
Log Out

Upload FCS Files

Upload the .fcs files produced from the flow cytometry experiment.
Only Red Fluorescence (HLog) and Green Fluorescence (HLog) are supported at this time.

FCS File 1

Control FCS File

Choose File No file chosen

FCS File 2

Experimental FCS File

Choose File No file chosen

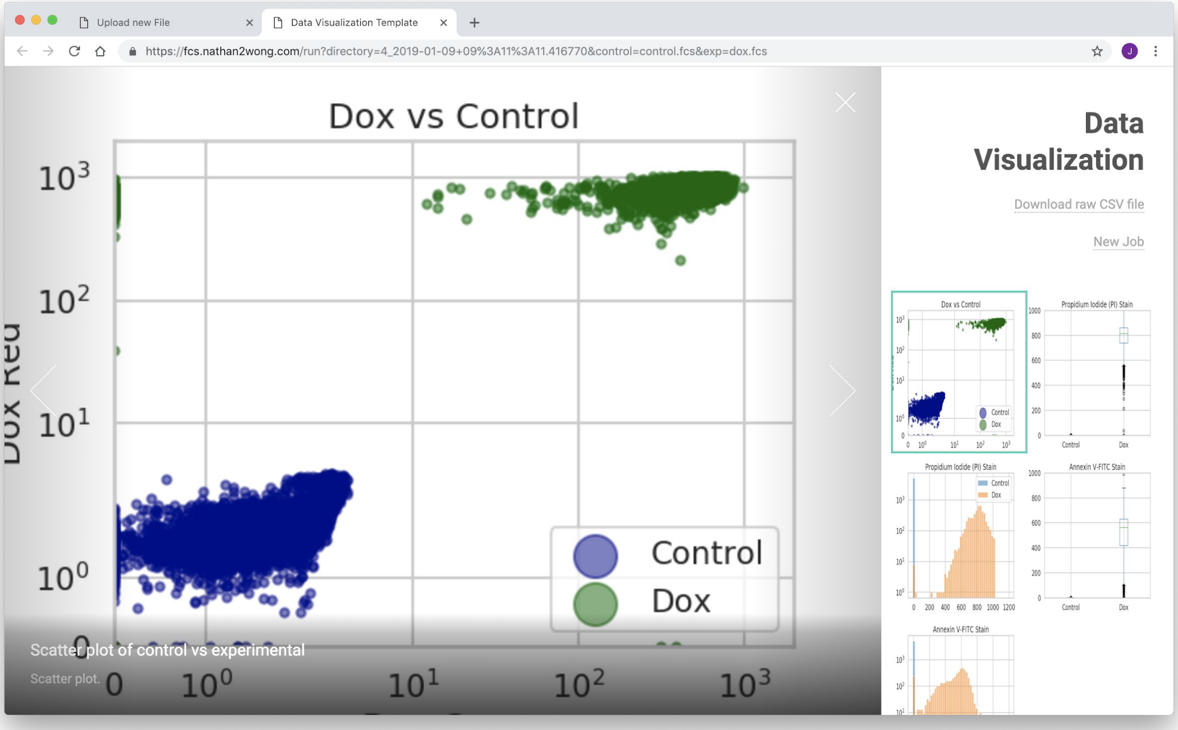
Upload Files

Email address: name@example.com

32019-01-23 20:08:03.426098

Flow Cytometry Analysis Tool created by Nathan.

https://fcs.nathan2wong.com



Upload new File
Data Visualization Template
https://fcs.nathan2wong.com/run?directory=4_2019-01-09+09%3A11%3A11.416770&control=control.fcs&exp=dox.fcs
Dox vs Control
Dox Red
10^3
10^2
10^1
10^0
0
0
10^0
10^1
10^2
10^3
Control
Dox

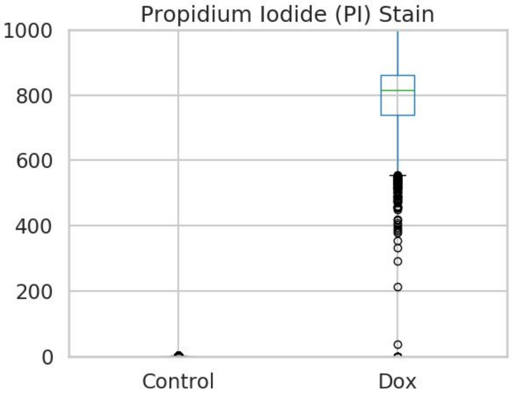
Scatter plot of control vs experimental

Scatter plot.

Data Visualization

Download raw CSV file

New Job

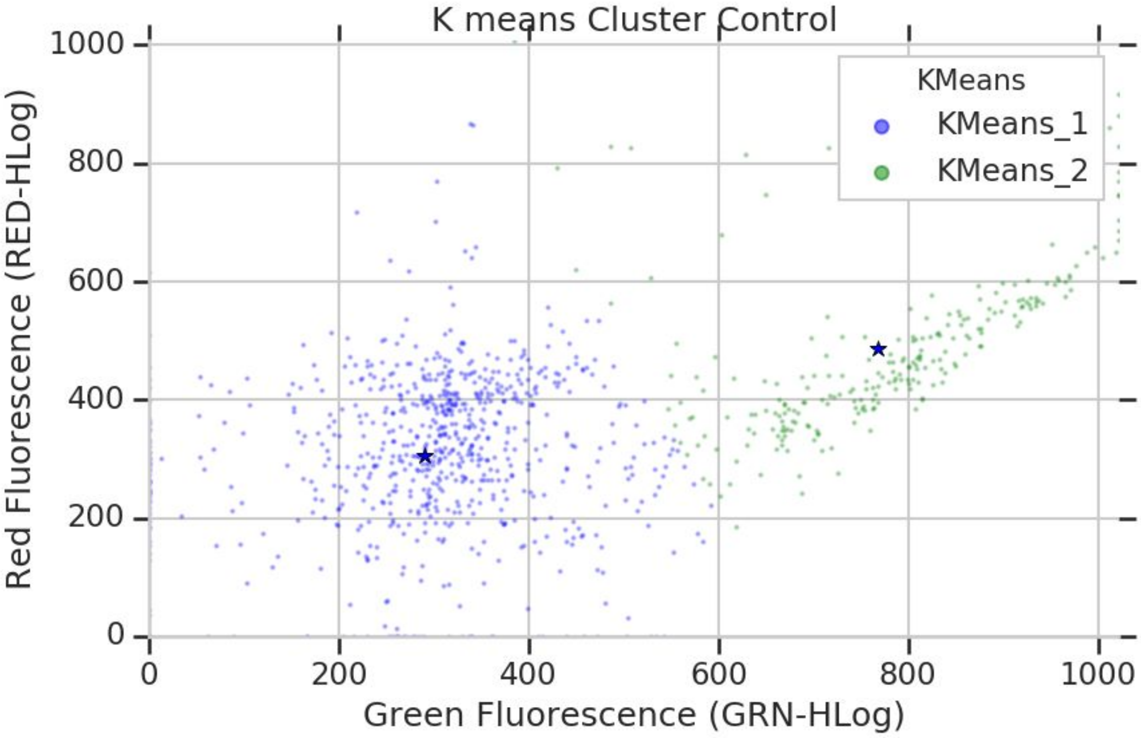


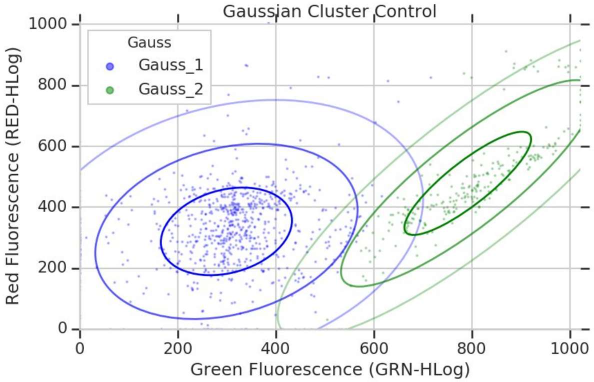
Demonstration Video

Machine Learning and Classification

- Initial goal: determine difference between different samples (ie. Dox vs Nothing)
- Secondary goal: classify into 4 different cell types
- Unsupervised learning, Perceptrons, Decision-trees, Convolutional Neural Networks, Support Vector Machine, K-nearest neighbors, Monte-Carlo Random Walk

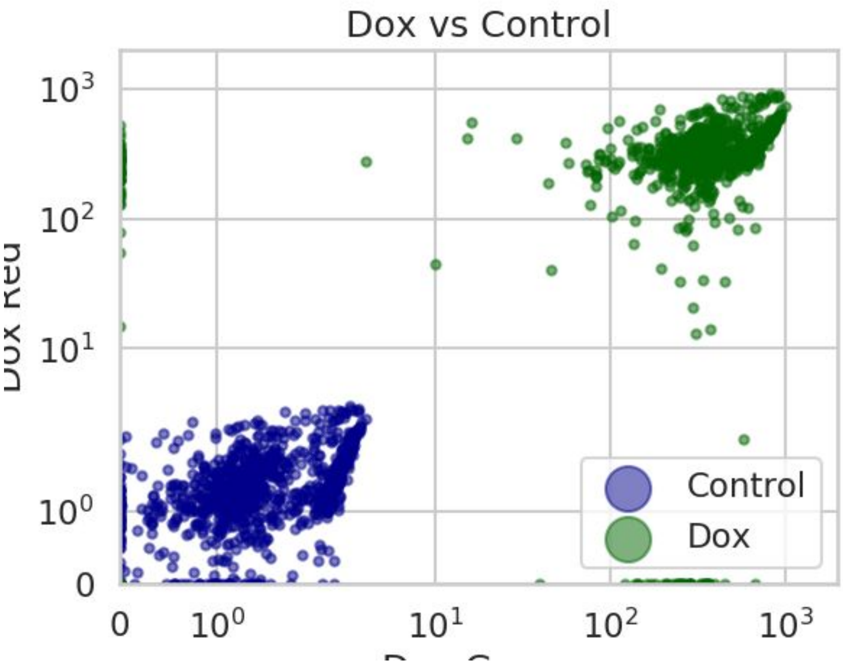
Unsupervised Learning

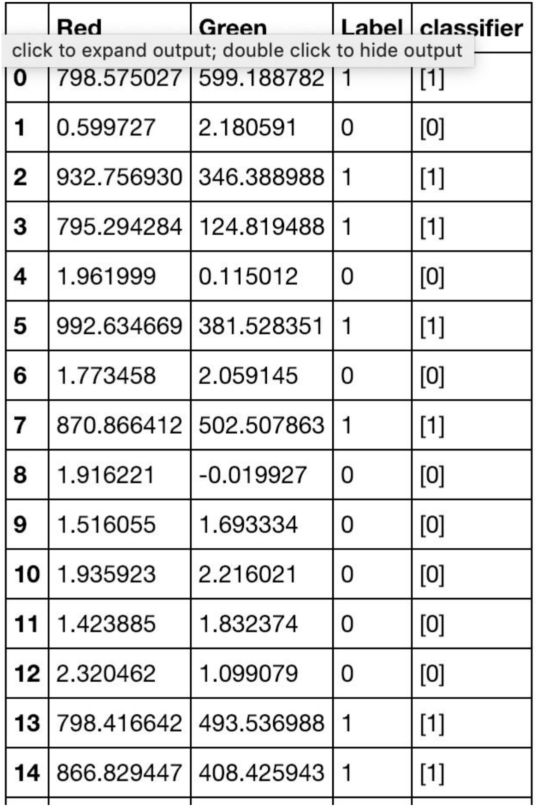




Train to differentiate datasets (ie. Dox vs Control)

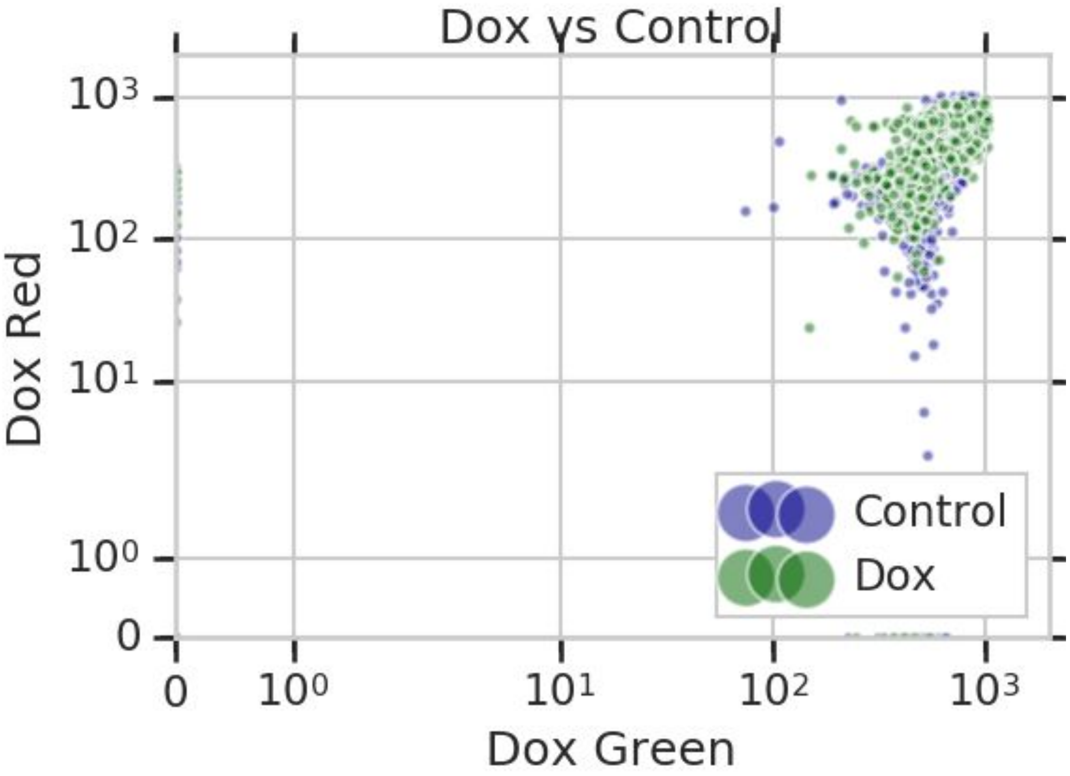
Perceptron: Binary Linear Classifier (~99% accuracy)

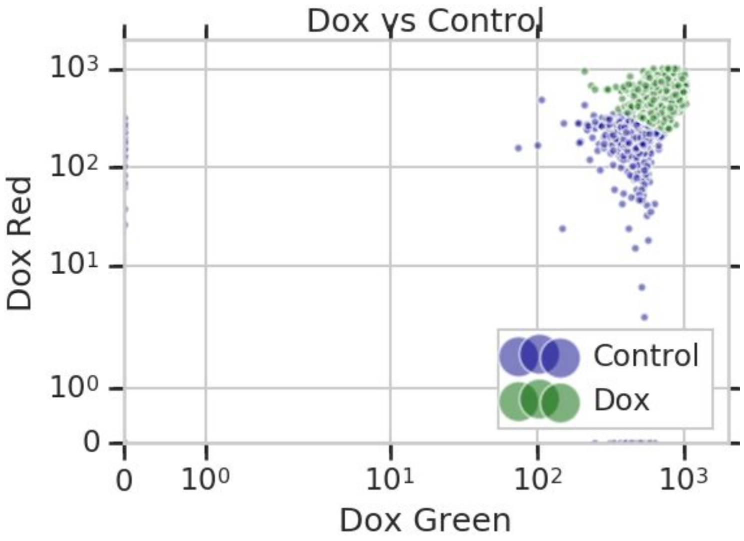



click to expand output; double click to hide output

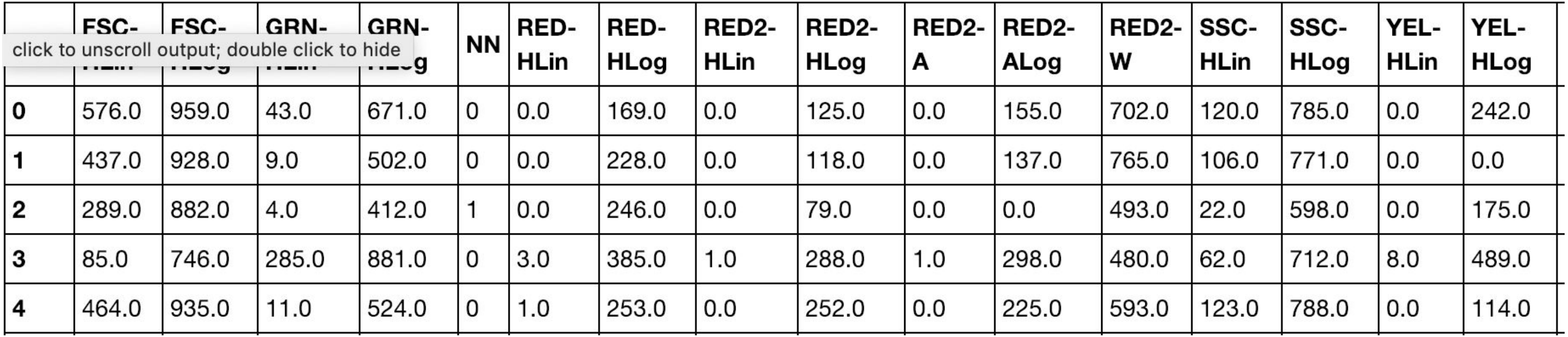
	Red	Green	Label	classifier
0	798.575027	599.188782	1	[1]
1	0.599727	2.180591	0	[0]
2	932.756930	346.388988	1	[1]
3	795.294284	124.819488	1	[1]
4	1.961999	0.115012	0	[0]
5	992.634669	381.528351	1	[1]
6	1.773458	2.059145	0	[0]
7	870.866412	502.507863	1	[1]
8	1.916221	-0.019927	0	[0]
9	1.516055	1.693334	0	[0]
10	1.935923	2.216021	0	[0]
11	1.423885	1.832374	0	[0]
12	2.320462	1.099079	0	[0]
13	798.416642	493.536988	1	[1]
14	866.829447	408.425943	1	[1]

Perceptron: Binary Linear Classifier (~74% accuracy)



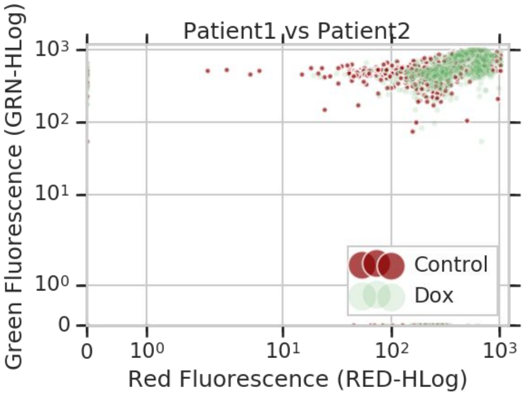


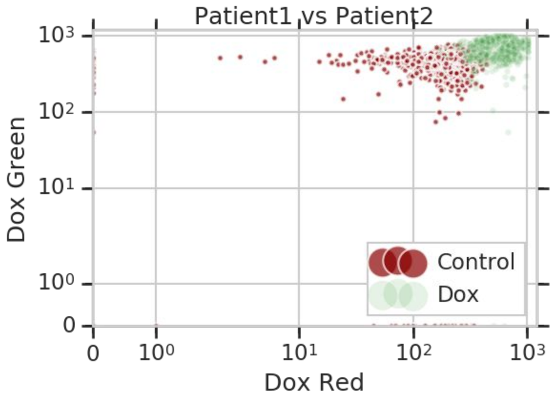
Adding more features (~79% accuracy)



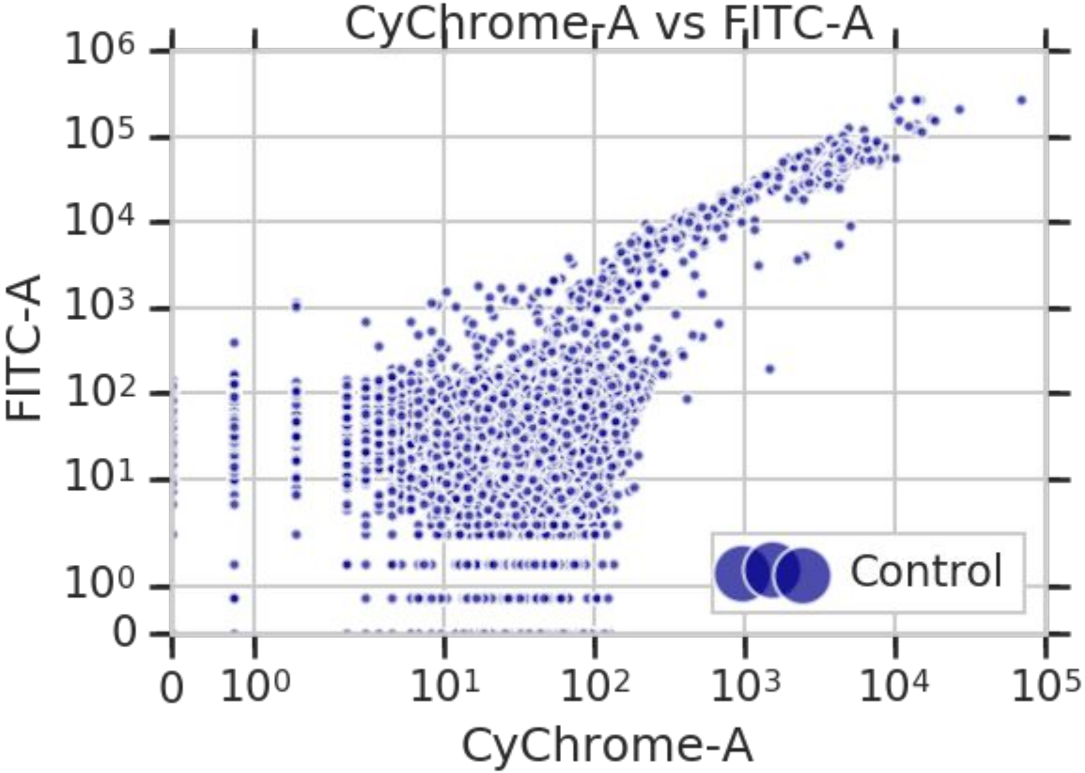
click to unscroll output; double click to hide

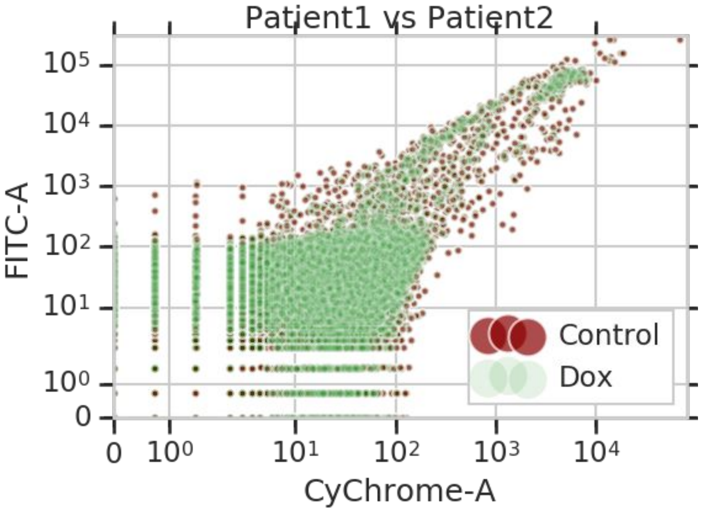
	FSC-HLin	FSC-HLog	GRN-HLin	GRN-HLog	NN	RED-HLin	RED-HLog	RED2-HLin	RED2-HLog	RED2-A	RED2-ALog	RED2-W	SSC-HLin	SSC-HLog	YEL-HLin	YEL-HLog
0	576.0	959.0	43.0	671.0	0	0.0	169.0	0.0	125.0	0.0	155.0	702.0	120.0	785.0	0.0	242.0
1	437.0	928.0	9.0	502.0	0	0.0	228.0	0.0	118.0	0.0	137.0	765.0	106.0	771.0	0.0	0.0
2	289.0	882.0	4.0	412.0	1	0.0	246.0	0.0	79.0	0.0	0.0	493.0	22.0	598.0	0.0	175.0
3	85.0	746.0	285.0	881.0	0	3.0	385.0	1.0	288.0	1.0	298.0	480.0	62.0	712.0	8.0	489.0
4	464.0	935.0	11.0	524.0	0	1.0	253.0	0.0	252.0	0.0	225.0	593.0	123.0	788.0	0.0	114.0





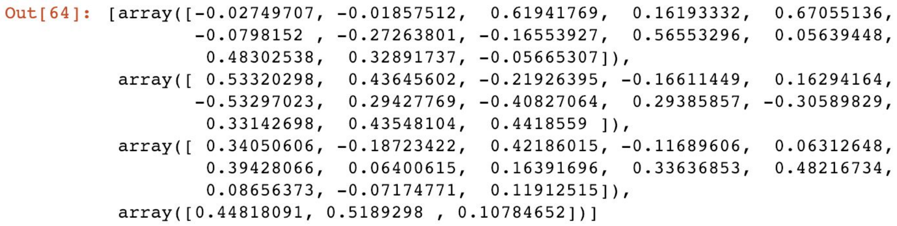
Nature Article Flow Cytometry Data



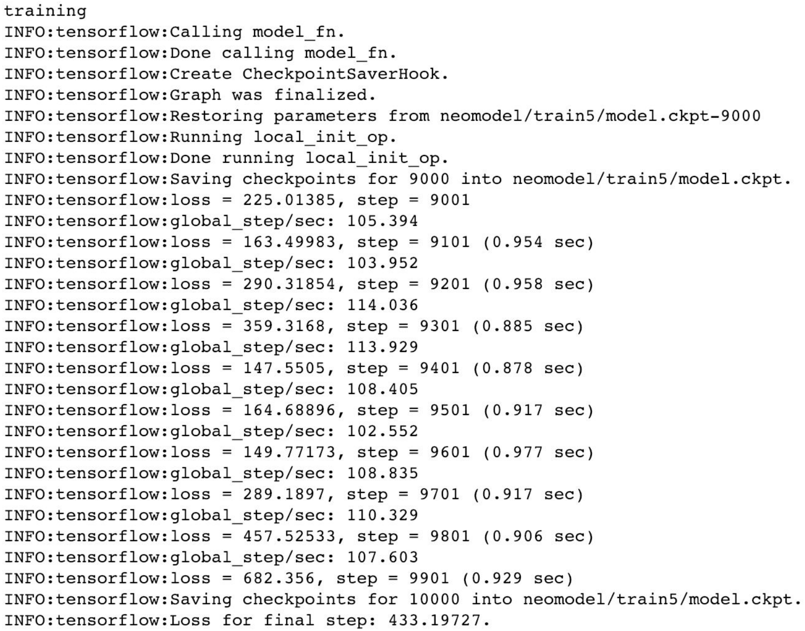


Other Classifiers (~ 70-75% accuracy)

- Random Forest (Decision tree)
- Hyperparameter regularization for perceptrons
- Convolutional Neural network (10 hidden nodes)



```
Out[64]: [array([-0.02749707, -0.01857512,  0.61941769,  0.16193332,  0.67055136,
                -0.0798152 , -0.27263801, -0.16553927,  0.56553296,  0.05639448,
                 0.48302538,  0.32891737, -0.05665307]),
          array([ 0.53320298,  0.43645602, -0.21926395, -0.16611449,  0.16294164,
                -0.53297023,  0.29427769, -0.40827064,  0.29385857, -0.30589829,
                 0.33142698,  0.43548104,  0.4418559 ]),
          array([ 0.34050606, -0.18723422,  0.42186015, -0.11689606,  0.06312648,
                 0.39428066,  0.06400615,  0.16391696,  0.33636853,  0.48216734,
                 0.08656373, -0.07174771,  0.11912515]),
          array([0.44818091, 0.5189298 , 0.10784652])]
```



```
training
INFO:tensorflow:Calling model_fn.
INFO:tensorflow:Done calling model_fn.
INFO:tensorflow:Create CheckpointSaverHook.
INFO:tensorflow:Graph was finalized.
INFO:tensorflow:Restoring parameters from neomodel/train5/model.ckpt-9000
INFO:tensorflow:Running local_init_op.
INFO:tensorflow:Done running local_init_op.
INFO:tensorflow:Saving checkpoints for 9000 into neomodel/train5/model.ckpt.
INFO:tensorflow:loss = 225.01385, step = 9001
INFO:tensorflow:global_step/sec: 105.394
INFO:tensorflow:loss = 163.49983, step = 9101 (0.954 sec)
INFO:tensorflow:global_step/sec: 103.952
INFO:tensorflow:loss = 290.31854, step = 9201 (0.958 sec)
INFO:tensorflow:global_step/sec: 114.036
INFO:tensorflow:loss = 359.3168, step = 9301 (0.885 sec)
INFO:tensorflow:global_step/sec: 113.929
INFO:tensorflow:loss = 147.5505, step = 9401 (0.878 sec)
INFO:tensorflow:global_step/sec: 108.405
INFO:tensorflow:loss = 164.68896, step = 9501 (0.917 sec)
INFO:tensorflow:global_step/sec: 102.552
INFO:tensorflow:loss = 149.77173, step = 9601 (0.977 sec)
INFO:tensorflow:global_step/sec: 108.835
INFO:tensorflow:loss = 289.1897, step = 9701 (0.917 sec)
INFO:tensorflow:global_step/sec: 110.329
INFO:tensorflow:loss = 457.52533, step = 9801 (0.906 sec)
INFO:tensorflow:global_step/sec: 107.603
INFO:tensorflow:loss = 682.356, step = 9901 (0.929 sec)
INFO:tensorflow:Saving checkpoints for 10000 into neomodel/train5/model.ckpt.
INFO:tensorflow:Loss for final step: 433.19727.
```

Challenges and Next Steps

Wet lab side

- Need to re-evaluate goals for this project

Computational

- Develop a reliable method to determine ‘true’ cell states
  - Sequencing is most reliable method, as described in Nature paper
- Gather more data, with more parameters (MTT can be used)
  - Currently looking into full-well imaging
  - Image Cytometry

Web application

- Expand to fit a library of detection parameters
- Add more visualization capabilities