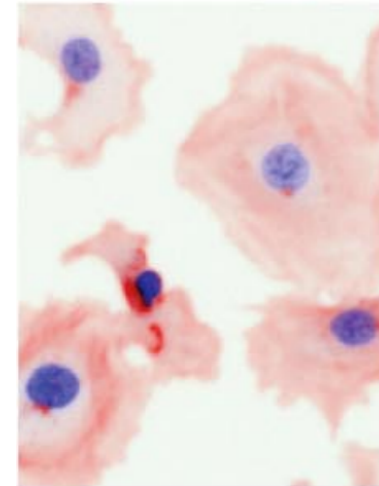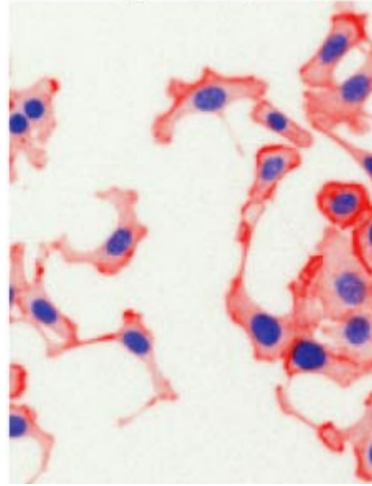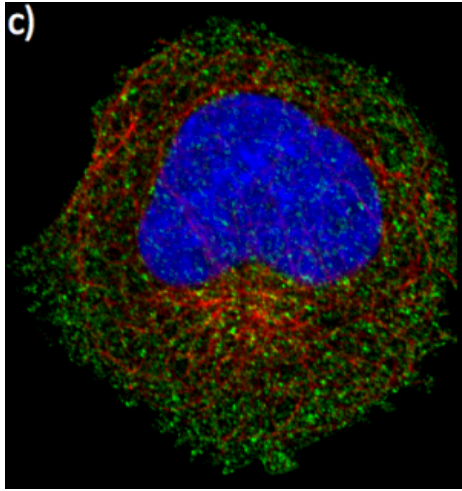# Introduction:
Machine learning on microscopy images

CS/CME/Biophys/BMI 371

March 6, 2018

Ron Dror

Structure of a cell vs. structure of a protein

- The structure of a protein can be specified by listing the spatial coordinates of each atom
  - The atoms move relative to one another, but by a limited amount
- In a cell, many of the molecules can move from one side of the cell to another
  - We could specify the position of every molecule, but that will vary substantially from one point in time to another (for a given cell) or between genetically and functionally identical cells
- Instead, we want a *statistical* description
- "Machine learning" = statistics

Morphological profiling

- Goal: recognize whether or not (or how much) a cell has been perturbed in a certain way
  - For example, distinguish between cells that have been treated by a particular drug and those that have not
  - Or distinguish between diseased and healthy cells
- Traditionally, this is done using certain hand-picked features (e.g., micronucleus count)
- One of Thursday's paper instead does it with convolutional neural networks
  - Without even segmenting out the cells
  - Using pre-training based on "natural" photographs

Repurposing image assays

- High-throughput screening (HTS) is a standard method to find initial hits in drug discovery
- In each discovery campaign, one generally develops a new assay (measurement)
  - Often these assays involve imaging and then extracting a certain measurement from each image
- One of Thursday's papers shows that one can take existing HTS imaging data and use it to replace a screen for an entirely different property
  - To do this, learn how image features relate to the property of interest.

Tracking cells over time

- The third paper presents a method to track cells in videos
- The problem is complicated by the facts that:
    - They want to track cell *lineages*, meaning they have to keep track of which cells divide into which other cells
    - The density of cells in the images can be very high
- They use machine learning (classifiers) in a couple different ways
    - Classifying background objects
    - Predicting cell types

Background

- Fluorescence microscopy

  – Slides from CS 279: http://web.stanford.edu/class/cs279/lectures/lecture11.pdf

  – Review paper: http://www.nature.com/nmeth/journal/v2/n12/pdf/nmeth817.pdf