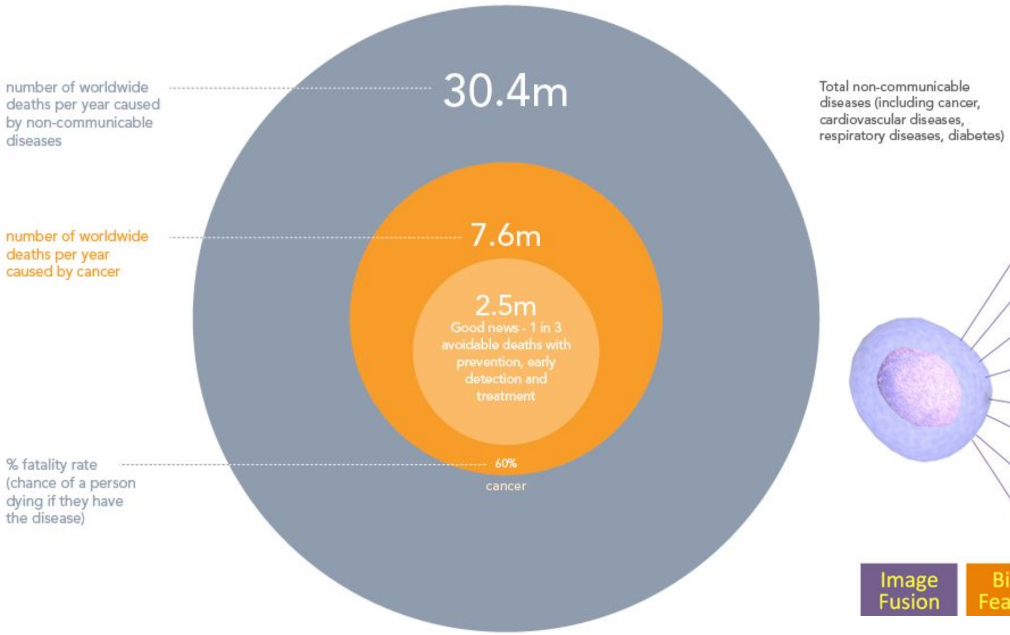
Scalable Informatics Tools for Investigating Intra-Tumor Heterogeneity in Breast Cancer

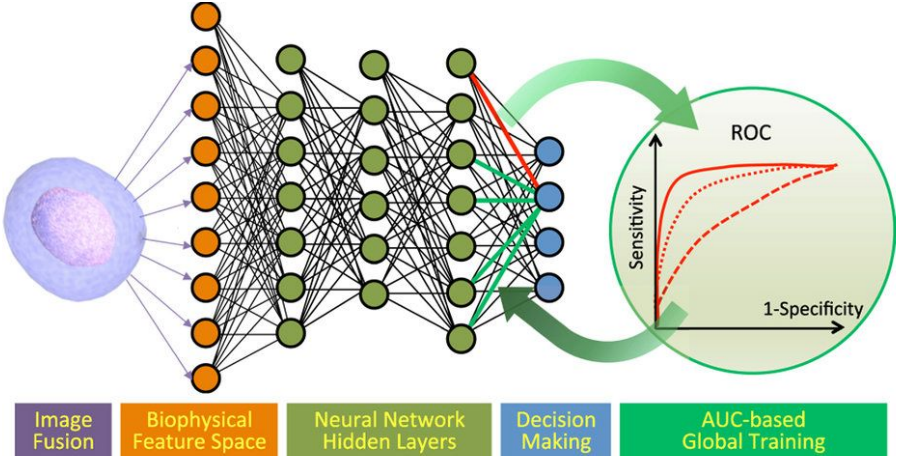
Alex Andonian

[1]TECBio REU @ Pitt, Department of Computational and Systems Biology University of Pittsburgh,
[2]Bates College, Lewiston, Maine

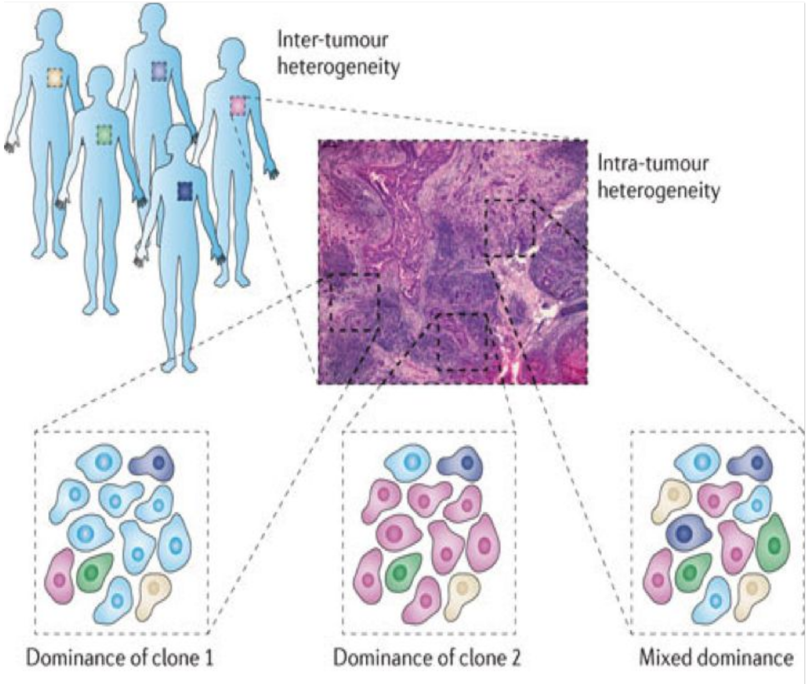
Motivation

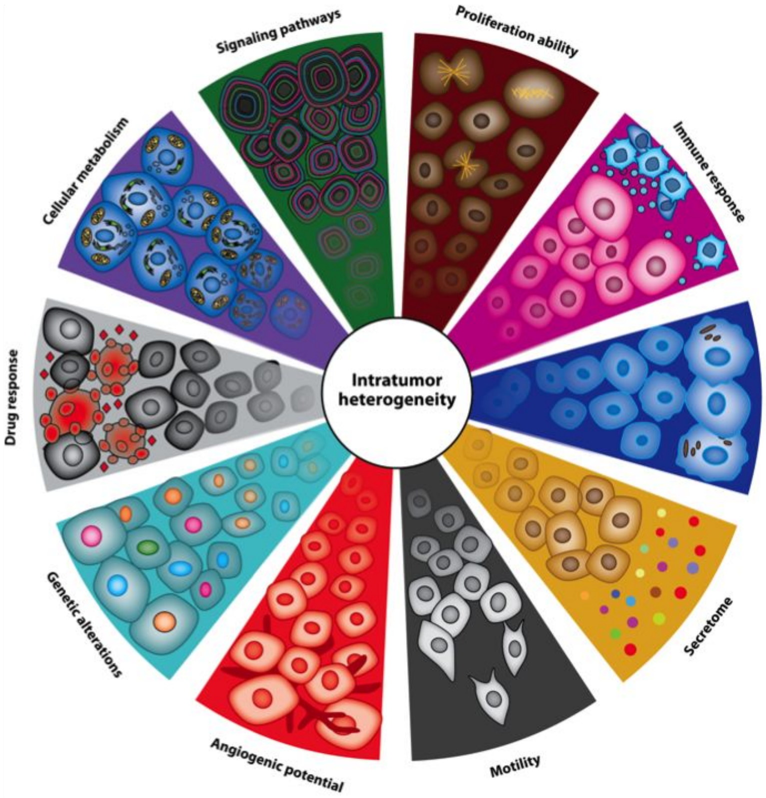
Cancer mortality compared to other diseases





Tumor Heterogeneity





Data

Tissue microarray of 99 samples consisting of triplicate, 1mm diameter cores from 24 invasive breast tumor tissues.

Immunohistochemical staining revealed 4 cohorts:

ER(+) IDC

ER(+) ILC

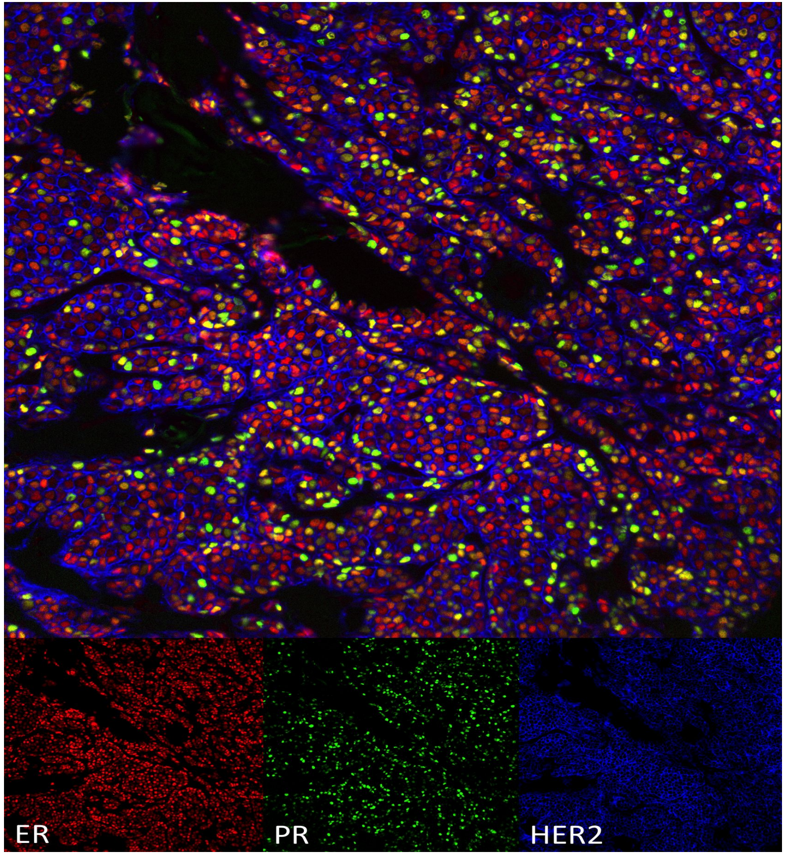
ER(-) IDC

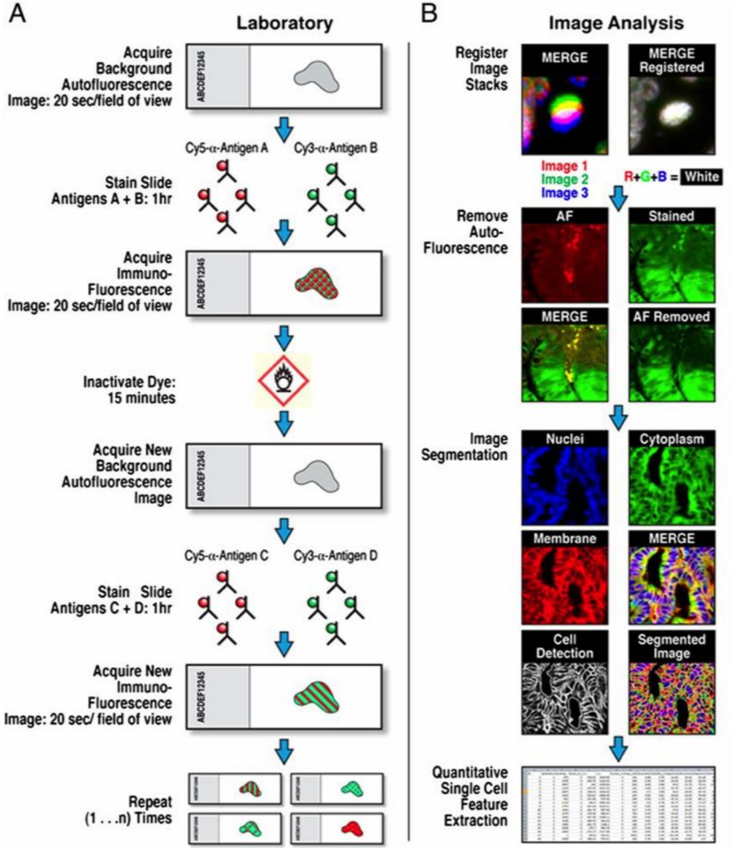
HER2(+) IDC

IDC - invasive ductal carcinoma

ILC - invasive lobular carcinoma

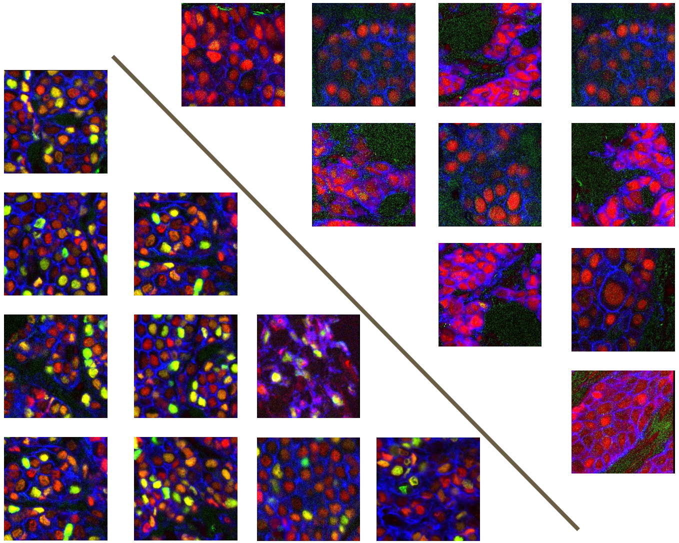
Multiplexed Immunofluorescence Imaging



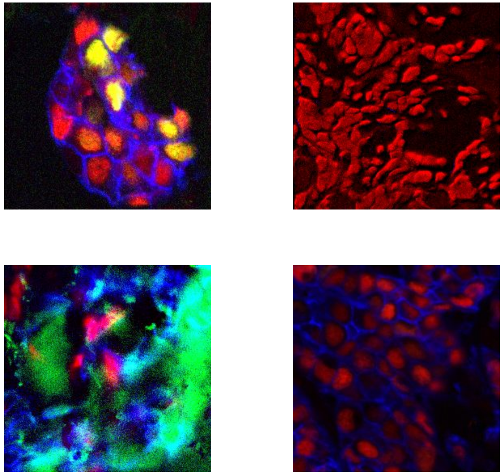


Research Goals

Develop high throughput informatics tools for integrating and analyzing cancer data obtained from a variety of imaging modalities

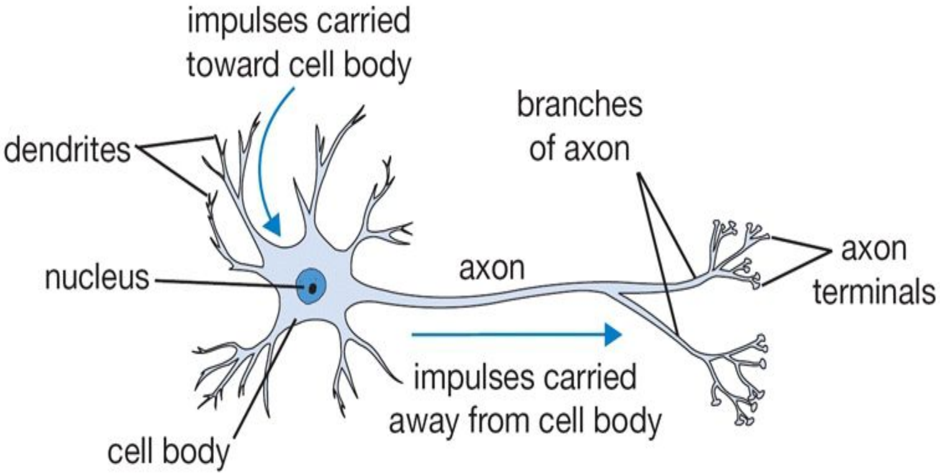


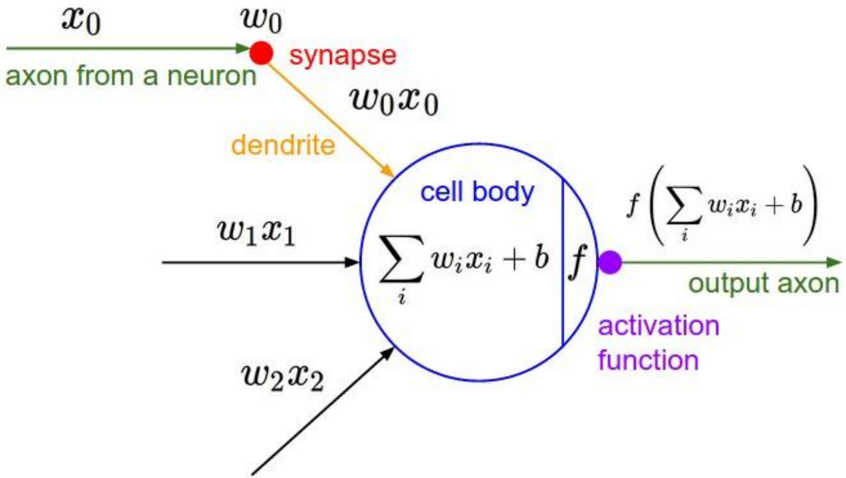
1. Cancer Classification



2. IF Signatures

Neural Networks - Biological Motivation





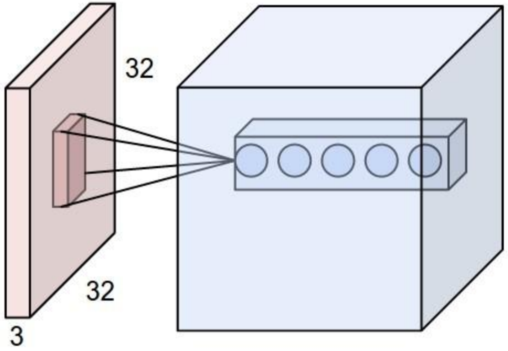
http://cs231n.github.io//neural-networks-1/

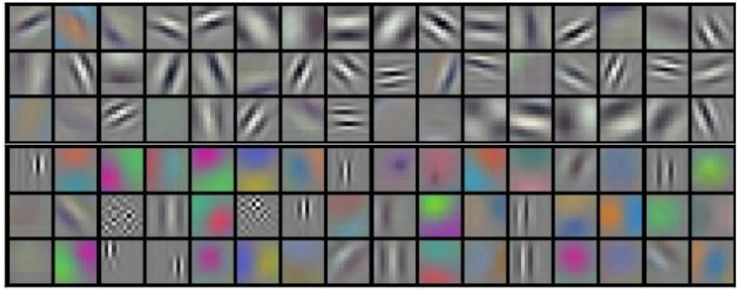
Convolutional Neural Networks

CNNs are very similar to ordinary neural networks, but...

- Now, we make the explicit assumption that input are **images**

- Since fully connected layers don't scale well - take advantage of the fact that portions of images are correlated

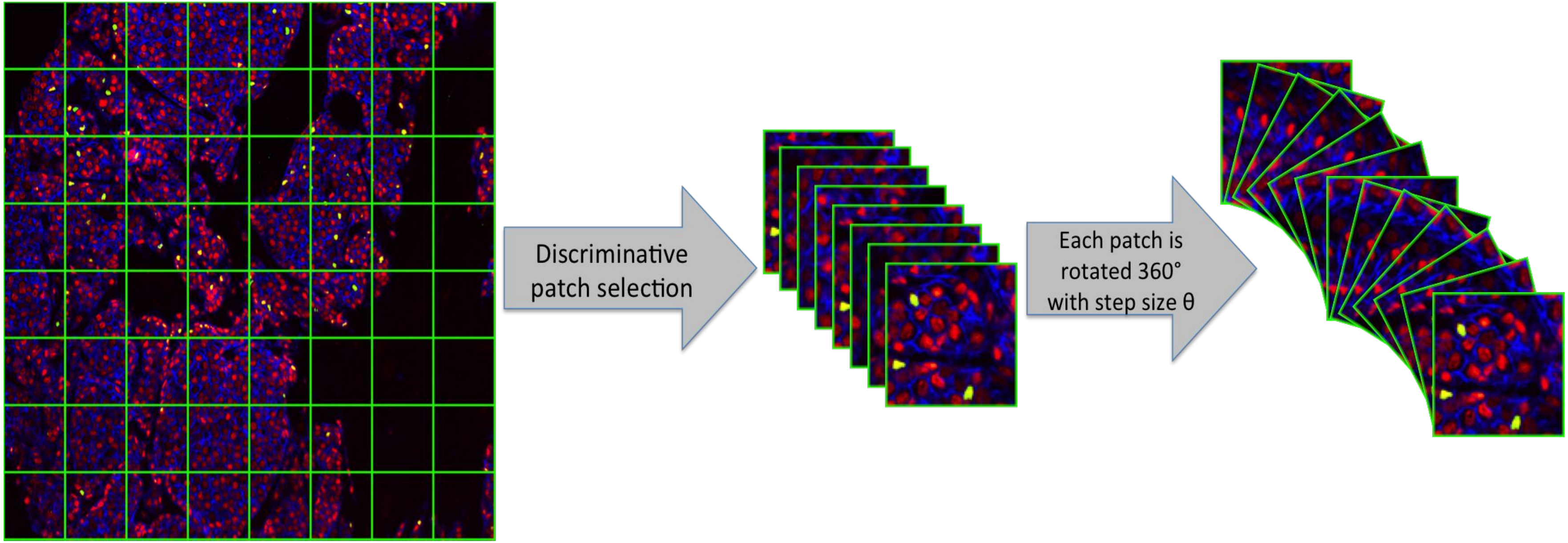




Example filters learned by AlexNet

http://cs231n.github.io/convolutional-networks/

Patch Selection and Augmentation



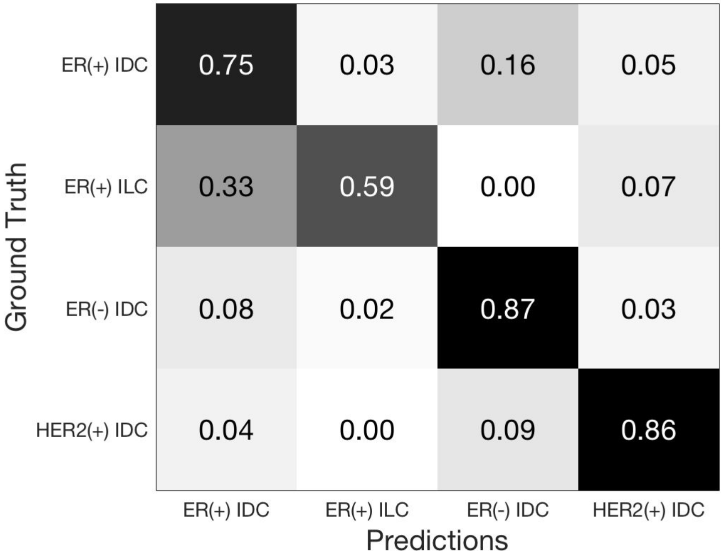
Results

Classification accuracy by configuration:

	256×256					512×512
θ step size	3	6	6	18	36	9
patch overlap	-	-	64	-	-	-
CNN Acc	0.79	0.77	**0.81**	0.76	0.75	0.78

Classification accuracy:

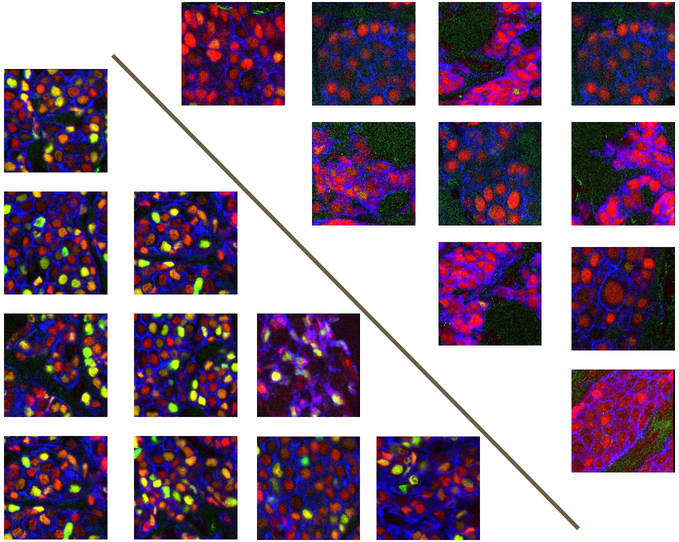
increases with patch overlap
decreases with rotation step size



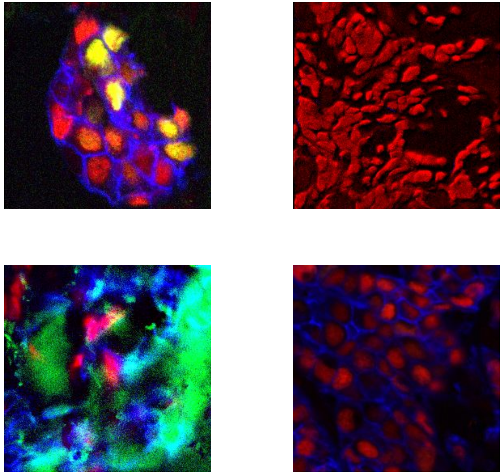
Confusion matrix for best performing configuration

Research Goals

Develop high throughput informatics tools for integrating and analyzing cancer data obtained from a variety of imaging modalities

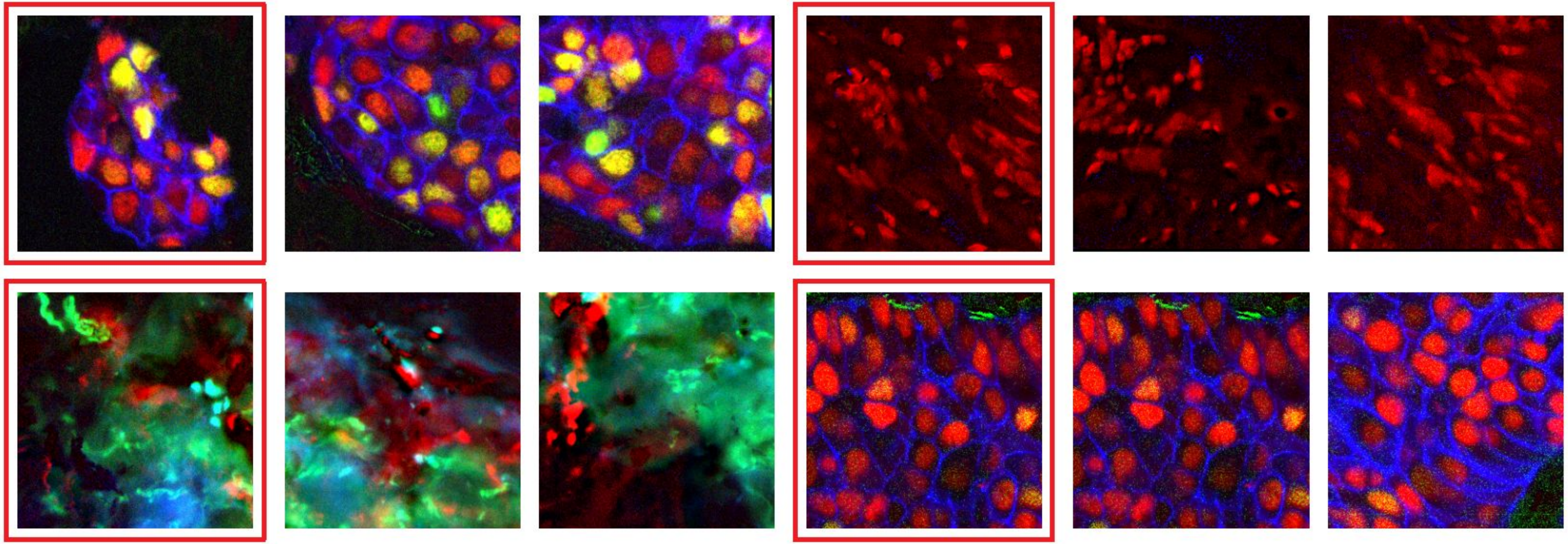


1. Cancer Classification



2. IF Signatures

Nearest Neighbor Visualization

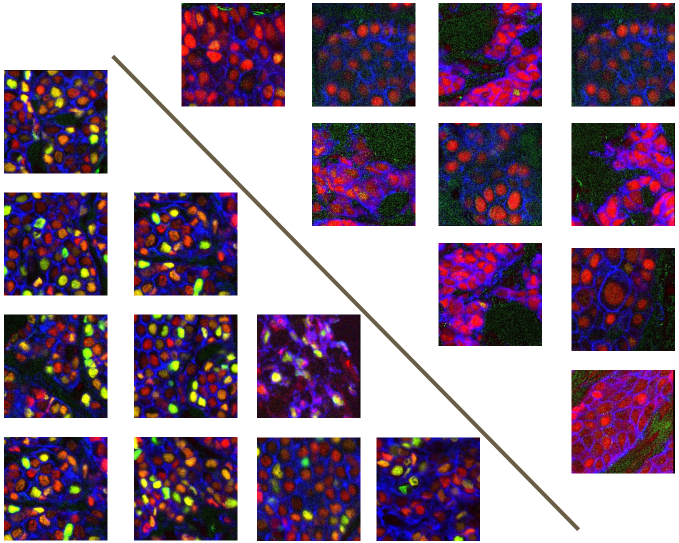


Could form the basis for a powerful and interactive visualization tool for clinicians

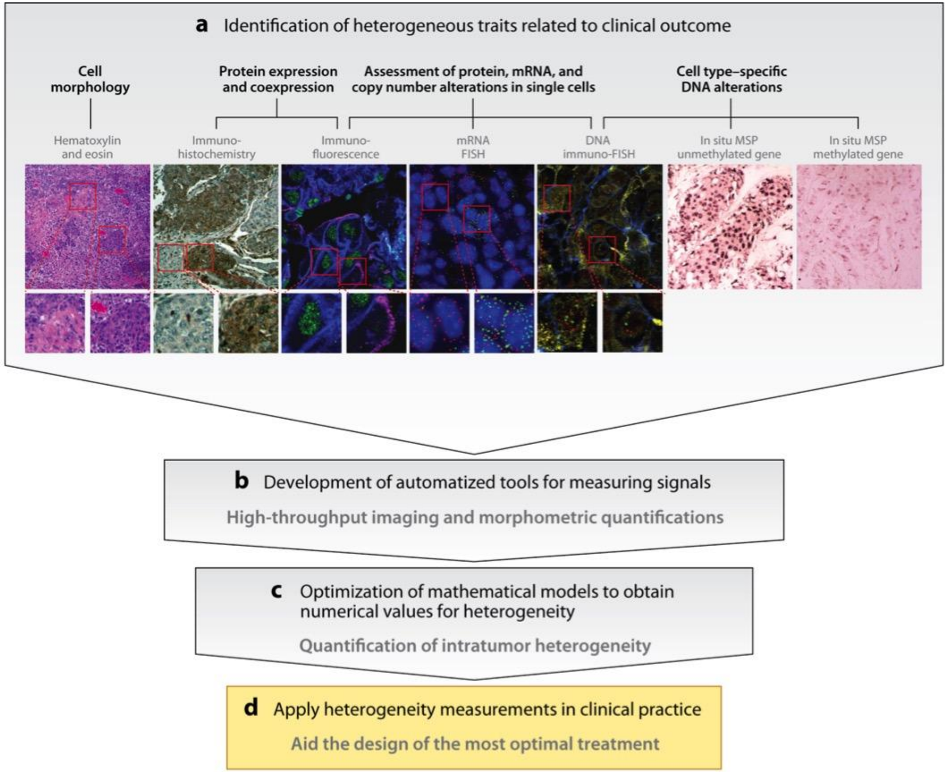
Conclusions

The main contributions of this work:

- classify cancer subtypes with respectable accuracy
- Identify immunofluorescent signatures associated with a cancer subtype



Future Work



a Identification of heterogeneous traits related to clinical outcome
Cell morphology
Protein expression and coexpression
Assessment of protein, mRNA, and copy number alterations in single cells
Cell type–specific DNA alterations
Hematoxylin and eosin
Immuno-histochemistry
Immuno-fluorescence
mRNA FISH
DNA immuno-FISH
In situ MSP unmethylated gene
In situ MSP methylated gene
b Development of automatized tools for measuring signals
High-throughput imaging and morphometric quantifications
c Optimization of mathematical models to obtain numerical values for heterogeneity
Quantification of intratumor heterogeneity
d Apply heterogeneity measurements in clinical practice
Aid the design of the most optimal treatment

Acknowledgements

TECBio REU @ Pitt is supported by the National Science Foundation under Grant DBI-1263020 and is co-funded by the Department of Defense in partnership with the NSF REU program. Many thanks to Dr. Chennubhotla and his lab group.

References

1. Siegel, R. L., Miller, K. D. and Jemal, A. (2015), Cancer statistics, 2015. CA: A Cancer Journal for Clinicians, 65: 5–29. doi: 10.3322/caac.21254
2. Marusyk, A., et al. “Intra-tumor heterogeneity: a looking glass for cancer?.” Nature Reviews Cancer 12.5 (2012): 323-334.
3. Almendro, V., et al. ”Cellular heterogeneity and molecular evolution in cancer.” Annual Review of Pathology: Mechanisms of Disease 8 (2013): 277-302.
4. Gerdes, Michael J., et al. “Highly multiplexed single-cell analysis of formalin-fixed, paraffin-embedded cancer tissue.” PNAS 119.29 (2013): 11982-11987.
5. Jia, Y., et al. “Caffe: Convolutional architecture for fast feature embedding.” arXiv preprint:1408.5093 (2014).

Questions?