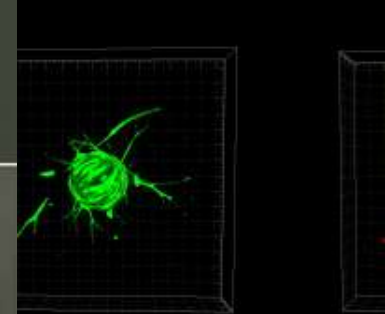
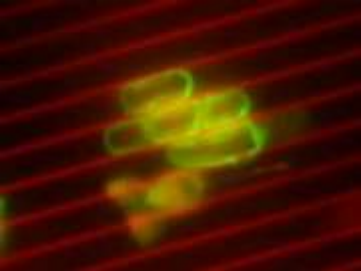




Data analysis

image analysis, machine learning, statistics



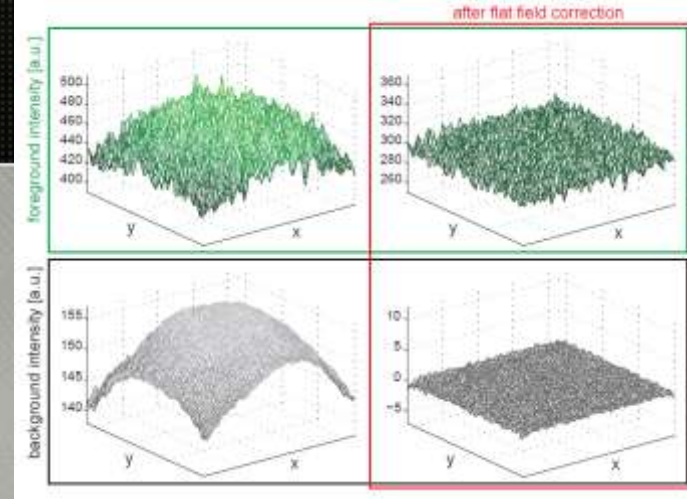
Peter Horvath
ETH Zurich
LMSC

Dublin 2013.

Roadmap

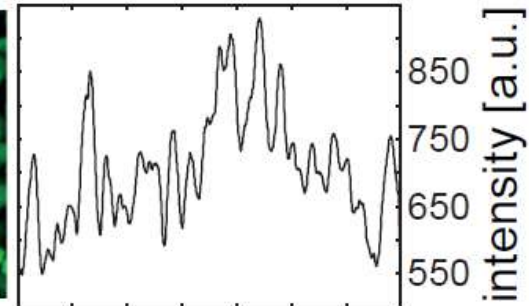
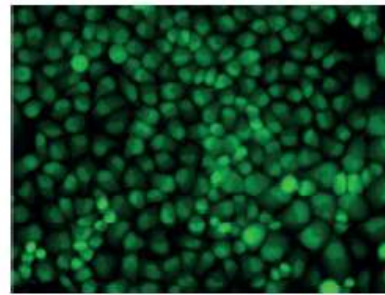
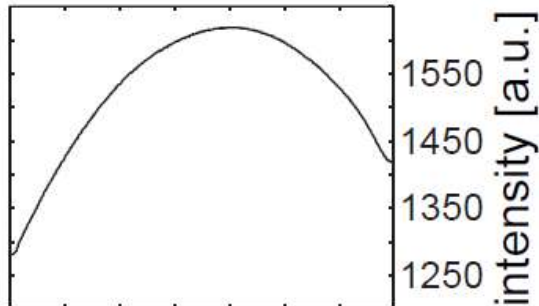
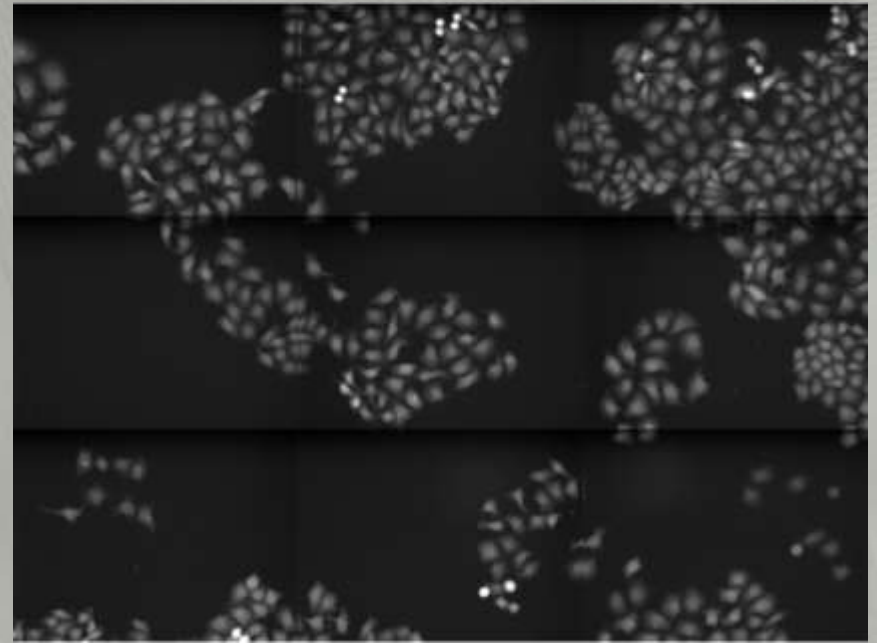
- Research/services
 - Flat field correction
 - Image analysis
 - Machine learning
 - Fingerprint/cluster analysis
- Future work
- Education
- Collaboration
- Usable software

ILLUMINATION CORRECTION IN HIGH CONTENT SCREENING



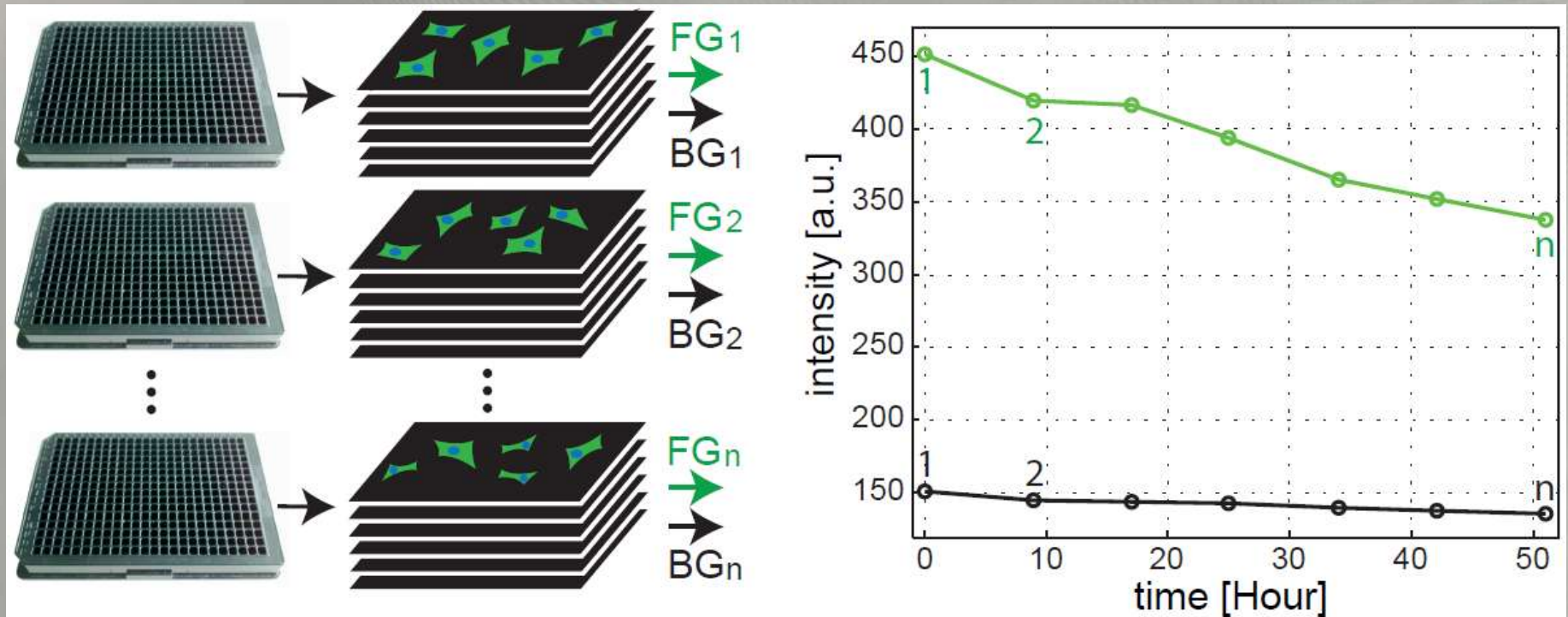
HCS image correction – Problems

1. Vignetting effect



HCS image correction – Problems

2. Light intensity decrease



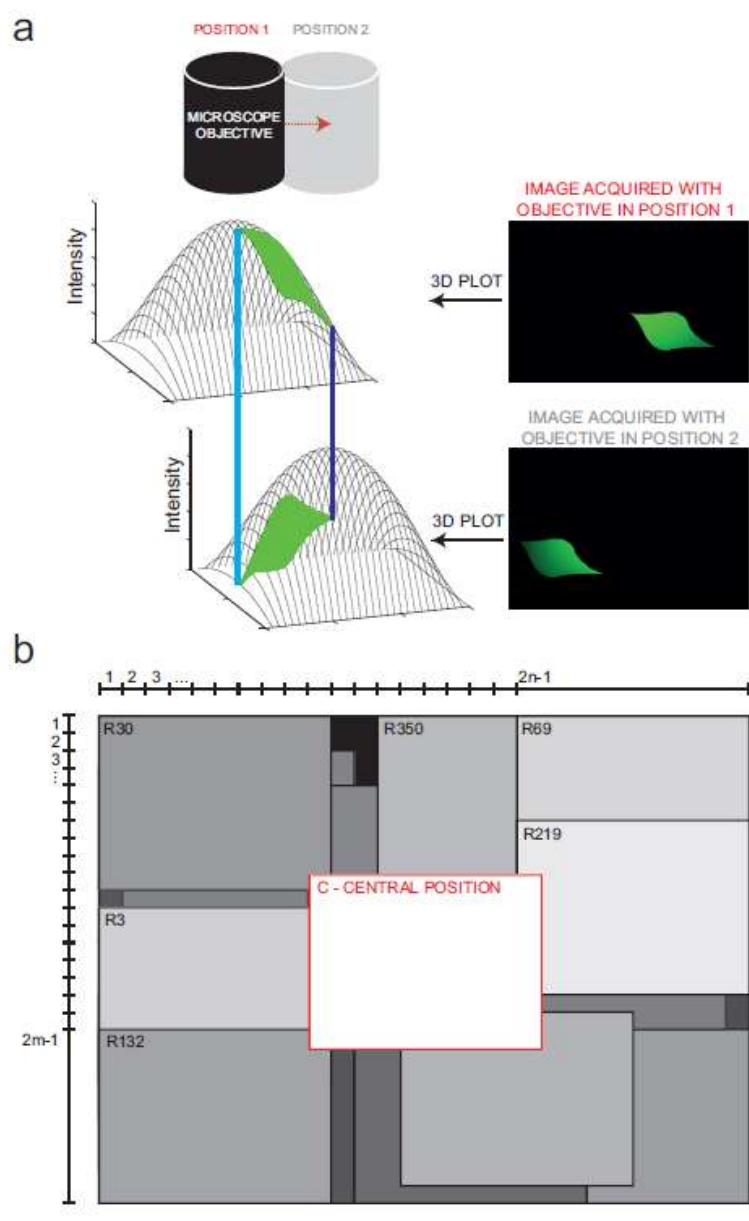
Methods

➤ Pre-acquisition

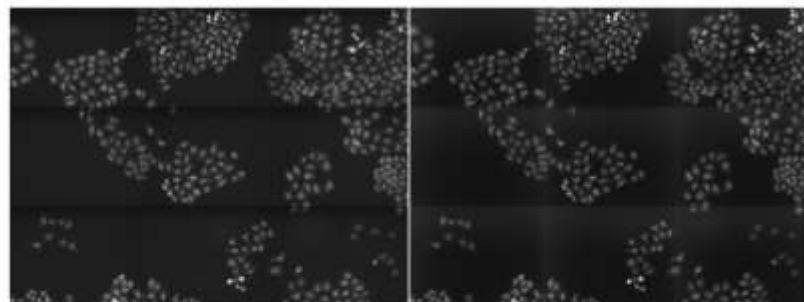
- Based on over
- Models vignette simultaneously
- 3-5 minutes mi scenarios
- Easy to build in

➤ Post acquisition

- State of the ar
- Works for most
- Large test set

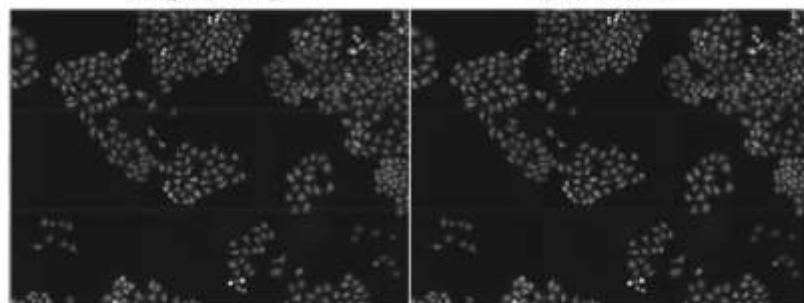


Results



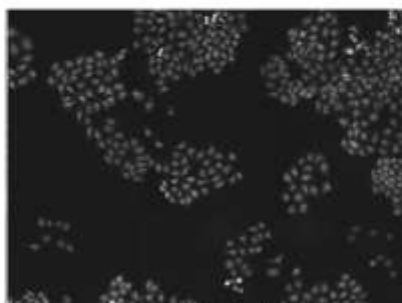
Original images

FCS correction



CP Mean correction

CP Both correction



Our Method

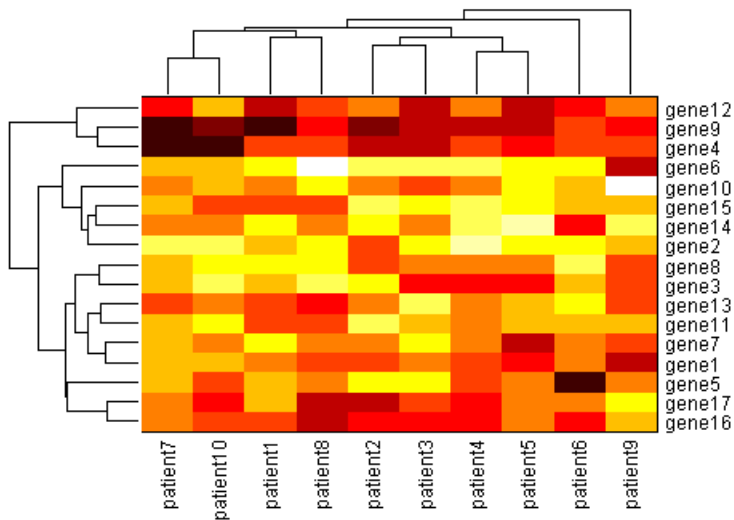


Machine learning

- Phenotypic profiling using:
 - Supervised
 - Semi-supervised
 - Active learning
- Regression models
- Machine learning optimization
 - Suggest a learner toolbox

Fingerprinting clustering

- Similarity studies using image-based features
 - Patient group specific drug discovery
 - Gene fingerprint similarities



Future directions

- Illumination problems
- 3D image analysis
- Unsupervised and semi supervised methods, analyzing the role of field experts
- Better clustering methods

Education

- High-content analysis day
 - 1 full day
 - Image analysis and downstream
- Matlab course for biologists
 - 2-3 times a year
 - 2 full days
- Image analysis courses
 - Imaris, ImageJ
- IPSS Image processing summer school for biologists

Usable software

- CellTracker
- Advanced Cell Classifier
- SALT



Advanced Cell Classifier

[Home](#) | [ACC](#) | [HTS Browser](#) | [Contact Us](#) | [ACC Wiki](#) | [HCS Blog](#)

Advanced Cell Classifier is a data analyzer program to evaluate cell-based high-content screens developed at ETH Zurich. The basic aim is to provide a very accurate analysis with minimal user interaction using advanced machine learning methods.

News

New publication with ACC
[more>>](#)

ACC workshop in Lausanne
[more>>](#)

Third genome-wide

Advanced Cell Classifier is a data analyzer program to evaluate cell-based high-content screens developed at ETH Zurich. The basic aim is to provide a very accurate analysis with minimal user interaction using advanced machine learning methods.



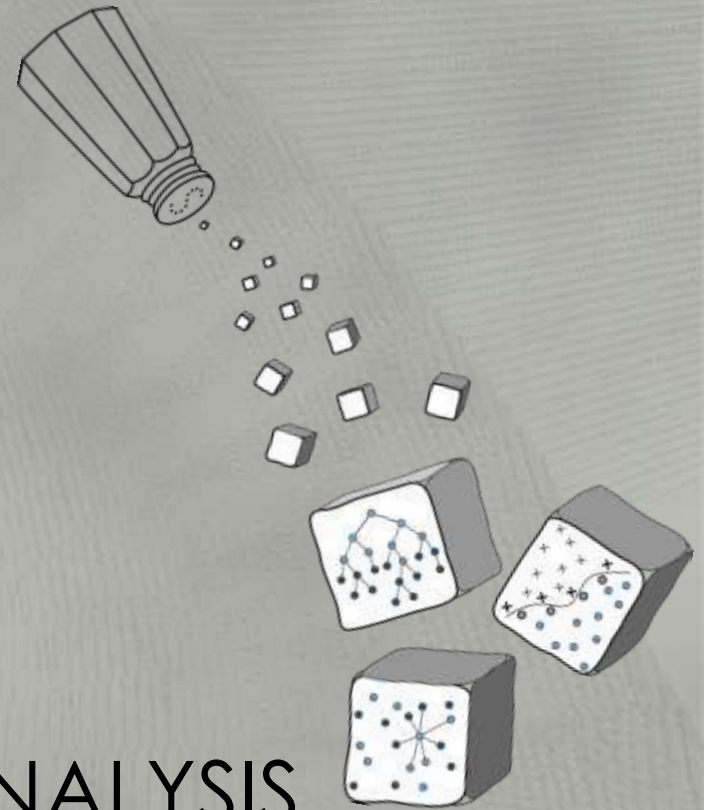
Last update: 08.10.2010



Huotari, Horvath *et al.*; **PNAS** 2012
Yamauchi, Horvath *et al.*; **PLoS Pathogens** 2011
Horvath, Wild *et al.*; **J. of Biomol Screening** 2011
Laurell, Horvath *et al.*; **Cell** 2011
Wild, Horvath *et al.*; **PLoS Biol** 2010
Turgay, Horvath *et al.*; **EMBO journal** 2010

SALT your data

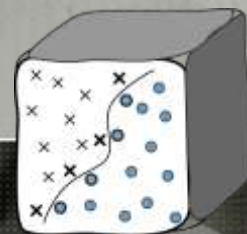
A MACHINE LEARNING ANALYSIS TOOLBOX FOR NON-EXPERTS



Suggest a Learner Toolbox (SALT)



- Matlab-based toolbox to find and optimize machine learning algorithms
- Easy to use (3 commands; train, predict, suggest + time limit)
- Advanced usage (easy to define new methods (XML), metrics, optimizers)
- Currently
 - 20-30 classifiers (Weka, LibSVM, Matlab Neural Networks, Vigna, SVMlight)
 - 10+ metrics (ROC, AUC, BEP, Confusion,...)
- Parallelized
 - Single machines (MAC, UNIX, and Win currently 12 parallel cores)
 - Distributed on clusters



Kevin Smith
Filippo Piccinini
Qasim Bukhari



Thank you for your attention!