Machine learning supported image analysis of microfluidic droplets

Carl-Magnus Svensson[1], Arjun Sarkar[1,2], Sundar Hengoju[3], Miriam Agler-Rosenbaum[3,4], Marc Thilo Figge[1,4]

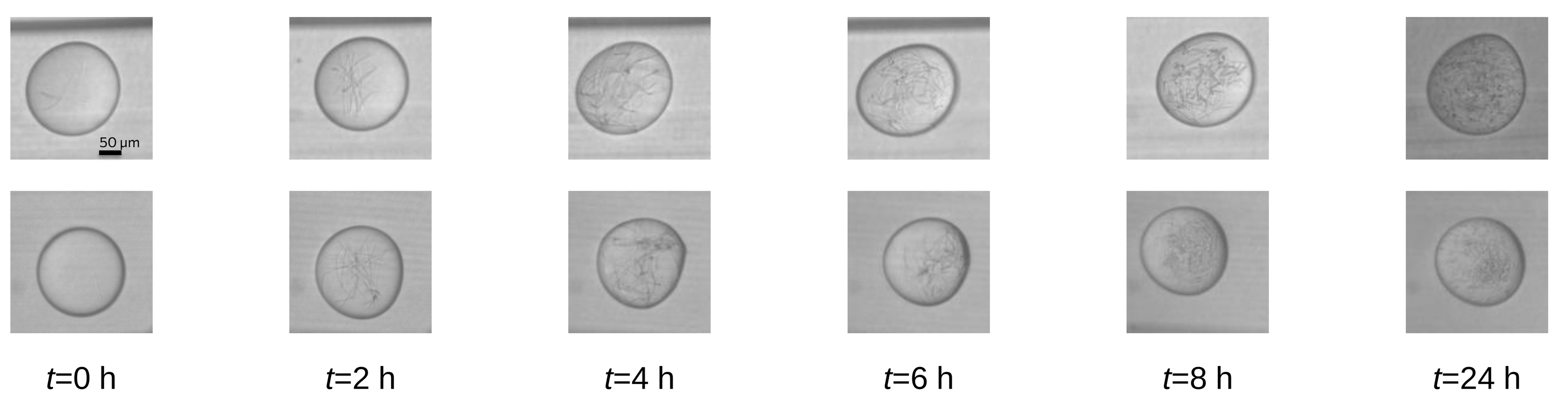
[1] Applied Systems Biology, Leibniz Institute for Natural Product Research and Infection Biology - Hans Knöll Institute, Jena, Germany
[2] Faculty of Biological Sciences, Friedrich Schiller University Jena, Jena, Germany
[3] Bio Pilot Plant, Leibniz Institute for Natural Product Research and Infection Biology - Hans Knöll Institute, Jena, Germany
[4] Institute of Microbiology, Faculty of Biological Sciences, Friedrich Schiller University Jena, Jena, Germany

Microfluidic droplets

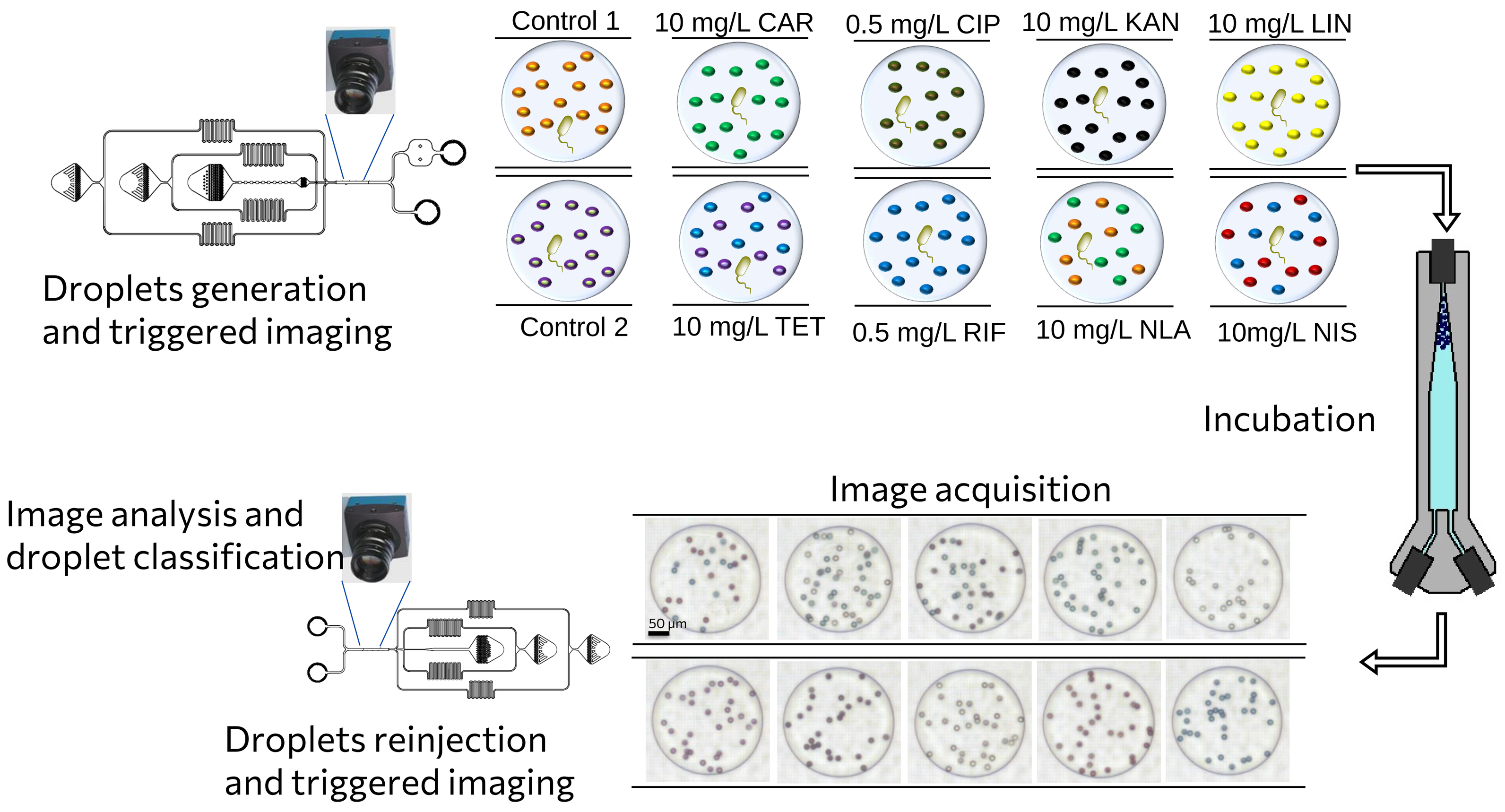
- Collaboration with Bio-Pilot Plant at Leibniz-HIK and Fraunhofer IOF
- Encapsulation of single cells in droplets
- Droplet generation and screening in the range of 100-1000 Hz
- Monitor microbial growth
- Demonstration of encoding of experimental conditions using color beads [1]
- Application to antibiotic susceptibility testing of eight different antibiotics validated by Microtiterplate (MTP) experiments

- Using Deep Learning to determine growth stages in droplets
- Angle-resolved scattering spectroscopy to determine single cell division events

E. Coli growth calibration curves

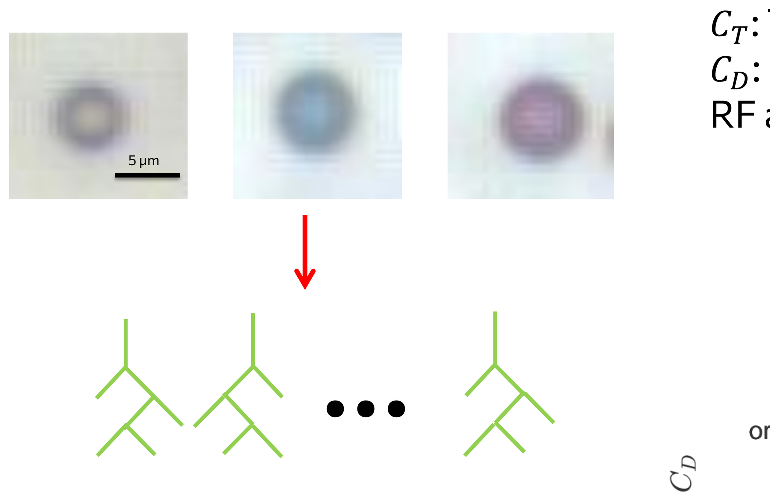


Droplet encoding

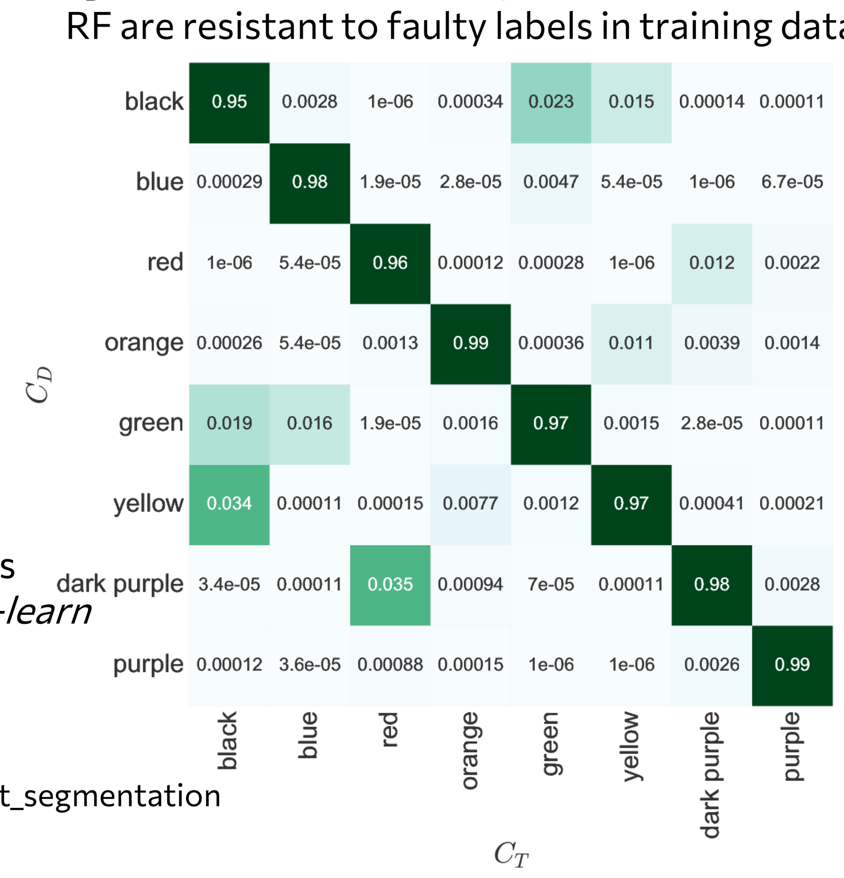


Droplet content determination

Bead classification



C_T: True bead color
C_D: Bead color detected by RF
RF are resistant to faulty labels in training data [2]

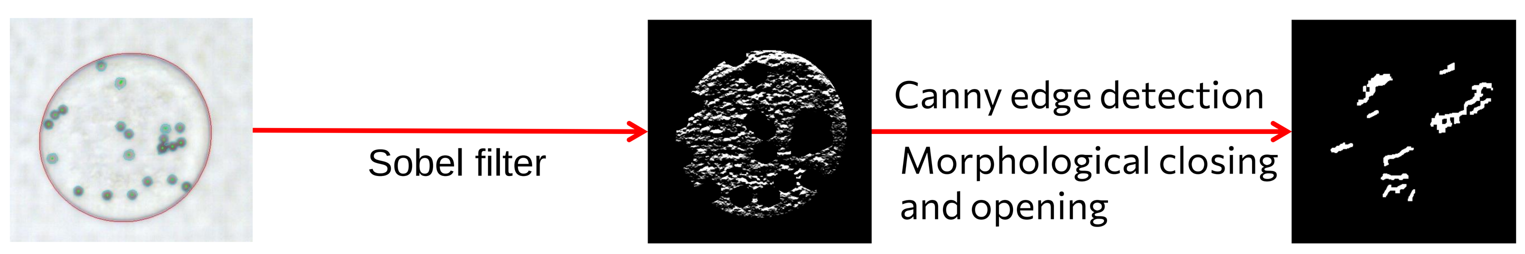


- 8 bead colors
- 11x11 pixels
- Lab colour space
- Random forest (RF) with 5000 trees
- Implemented in Python using *scikit-learn*

GitHub

github.com/applied-systems-biology/Droplet_segmentation

Growth detection



3000 droplets: $Acc = 0.925$, $Rec = 0.788$ and $Pre = 0.992$

Droplet decoding

- Each droplet has a true code (K_T) and the RF gives a detected code (K_D)
  $K_D = \{C_{D,1}, \dots, C_{D,N_D}\}$
- We are looking for the most likely code given:
  - Bayes theorem
    $p(K_T|K_D) \sim p(K_D|K_T)p(K_T)$
  - The misclassification rate of the RF
    $$p(K_D|K_T) = \prod_{j=1}^{N_D} p(C_{D,j}|C_{T,j})^{\boxed{k_j}}$$
    Number of beads of color j
  - A multinomial distribution with equal probability for each color
    $$p(K_T) = p_{MN}(\boldsymbol{v}_{K_T}) = \frac{\sum_{j=1}^{N_T} k_j}{\prod_{j=1}^{N_T}(k_j!)}\left(\frac{1}{N_T}\right)^{\sum_{j=1}^{N_T} k_j}$$

$K_D = \{g, o\}$
$\boldsymbol{v}_{K_D} = [k_1, k_2] = [22,1]$
$K_T = \{g, o\}, \{g\}$ or $\{o\}$
N_D: Number of detected colors in a droplet

$p(\{g\}) = p(g|g)^{22}p(o|g)^{1}p_{MN}([23]) \approx 7.98 \cdot 10^{-5}$ Most probable code

$p(\{o\}) = p(g|o)^{22}p(o|o)^{1}p_{MN}([23]) \approx 1.82 \cdot 10^{-76}$

$p(\{g,o\}) = p(g|g)^{22}p(o|o)^{1}p_{MN}([22,1]) \approx 1.39 \cdot 10^{-6}$

- 4259 droplets manually validated
- 18 were wrongly decoded (<0.5%)

No microbial growth
Diminished microbial growth
No effect on microbial growth

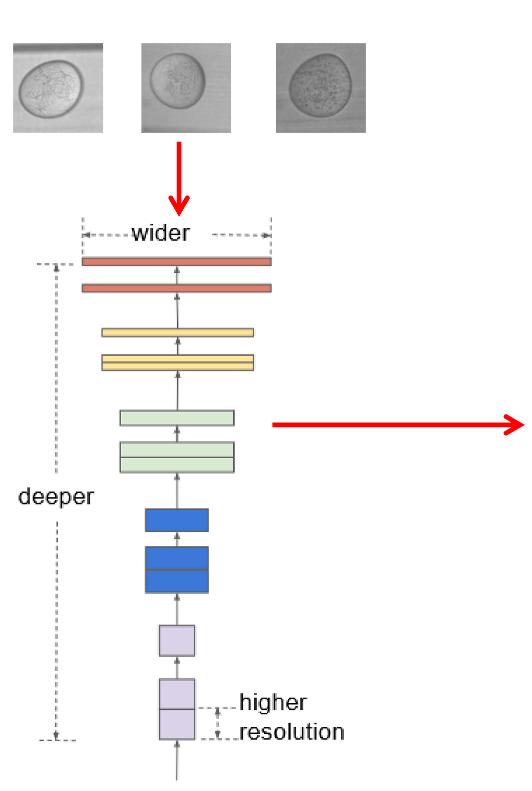
Antibiotic susceptibility testing

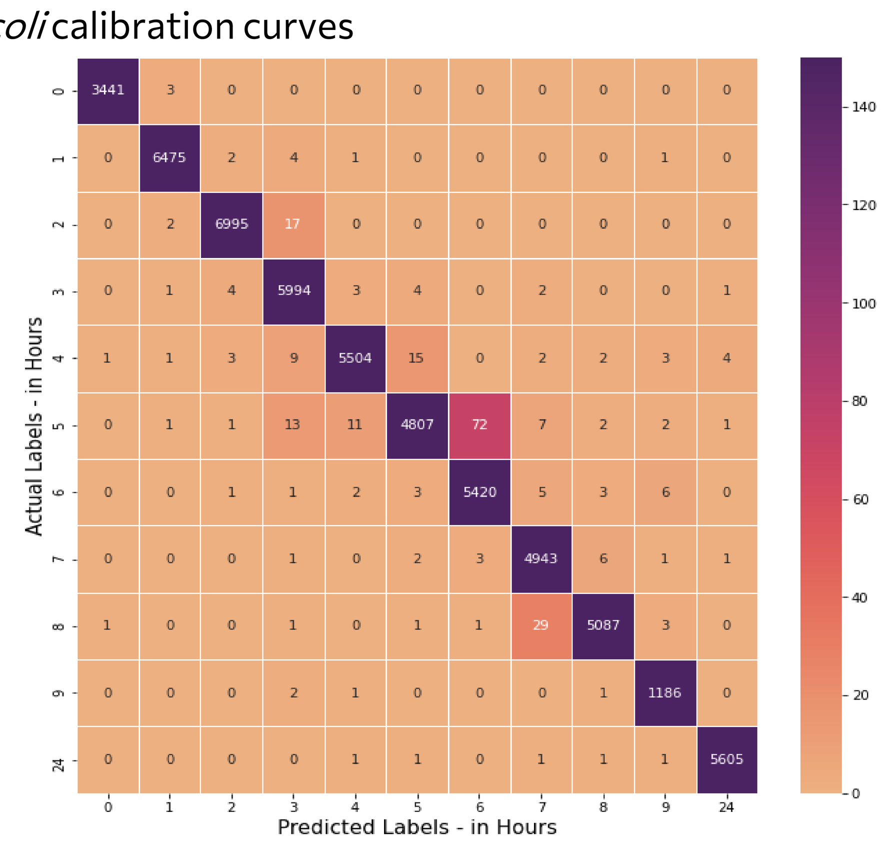
Experimental condition	Control 0 mg/ml	CAR 10 mg/ml	CIP 0.5 mg/ml	GEN 10 mg/ml	KAN 10 mg/ml	LIN 10 mg/ml	NIS 10 mg/ml	NLA 10 mg/ml	RIF 0.5 mg/ml	TET 10 mg/ml
Droplets	1	0.466	0.090	0.073	0.592	0.661	0.526	0.482	0.565	0.055
MTP	1	0.998	0.023	0.005	1.003	0.961	0.989	0.998	0.950	0.001

Future developments

Deep learning

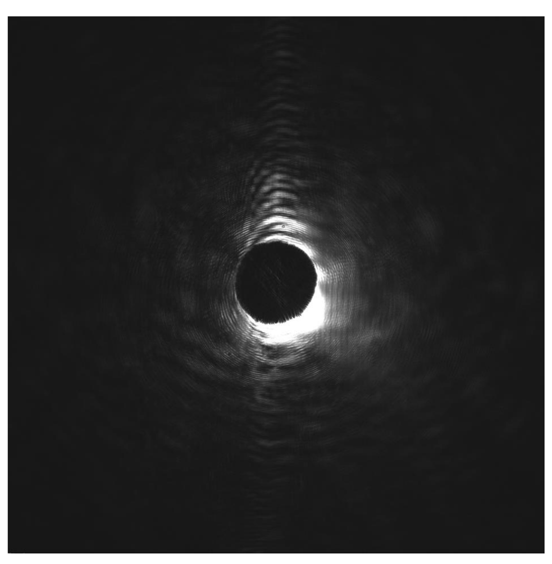
- Use Convolutional Neural Networks to learn growth at different time points
- Tuned version of the EfficientNet-B4 was used [3]
- Trained and evaluated on *E. coli* calibration curves



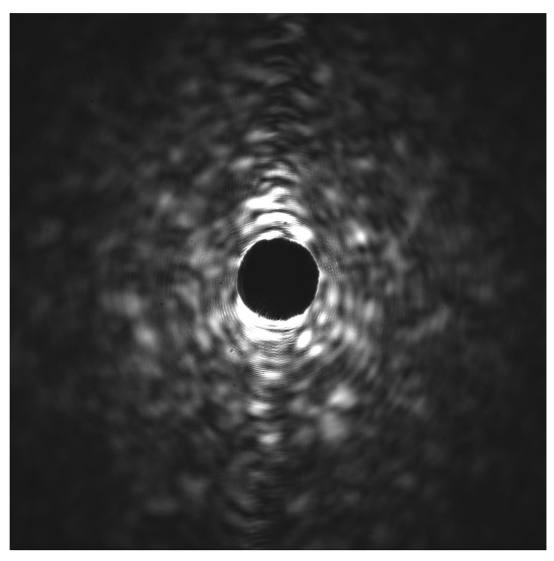


Angle resolved spectra

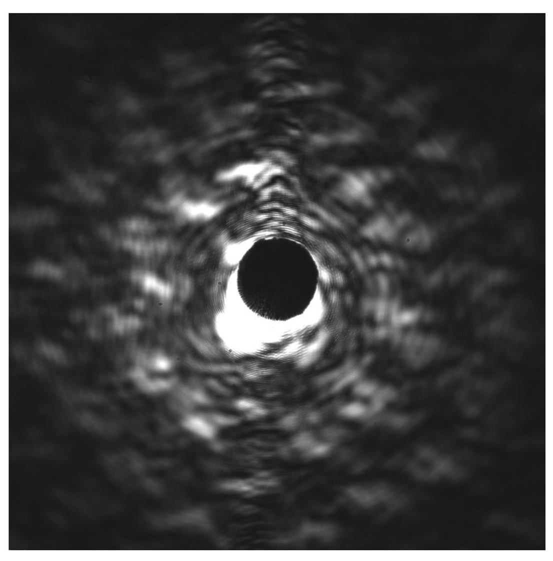
- Droplets in flow imaged using angle resolved scattering techniques [4]
- Will use traditional image analysis and Deep Learning technique to quantify small changes in spectra
- Goal is to detect single cell division events for rapid antibiotic susceptibility testing (<1 h after lag phase)



Empty droplet



Droplet with *E. coli*



Droplet with beads

carl-magnus.svensson@leibniz-hki.de

www.leibniz-hki.de

References
[1] Svensson *et al.* 2019. *Small*. 15(4):e1802384
[2] Svensson *et al.* 2015. *J Immunol Res*. 2015:573165
[3] Tan and Le. 2020. *ArXiv*. 1905.11946v5
[4] Schröder *et al.* 2011. *Appl Opt*. 50(9):C164-C171

Funded by Thüringer Aufbaubank through the projects Dropcode (2014FE9037) and Autoscreen (2017FE9071). Angle resolved scattering work is funded by BMBF through InfectoGnostics 2 (ADA Nr. 13GW0456B)

Thüringer Aufbaubank

InfectoGnostics Forschungscampus Jena

Bundesministerium für Bildung und Forschung