Evaluating DeLTA 2.0 for Growth-Based Drug Susceptibility Analysis in Microfluidic Time-Lapse Microscopy

Joel Sundin
Image Analysis & Machine Learning Masters Programme,
Uppsala University, Uppsala, Sweden
joel.sundin.9471@student.uu.se

***Abstract*—This work evaluates whether the DeLTA 2.0 single-cell analysis pipeline can be adapted to microfluidic time-lapse phase-contrast microscopy for growth-based phenotypic drug susceptibility testing (pDST). Using a *Mycobacterium smegmatis* dataset from Tran et al., Omnipose-generated masks were refined using distance-transform–based marker-controlled watershed splitting and used to retrain the DeLTA 2.0 segmentation U-Net. Growth rates based on cell area and length were extracted from DeLTA 2.0 outputs and summarized at the population level using 10-frame (20 min) temporal binning and median aggregation to mitigate tracking noise. Drug response was quantified as the percent change in area under the population growth-rate curve (AUC) relative to paired controls, yielding reductions between -7.9% and -81.8% across seven positions. While population-level trends were consistent across positions, tracking instability and imperfect annotations precluded reliable single-cell susceptibility analysis. Overall, DeLTA 2.0 supports robust population-level growth-based pDST analysis on this dataset, but improved ground truth data are required to evaluate or enable single-cell readouts.**

***Index Terms*—phenotypic drug susceptibility testing, microfluidics, time-lapse microscopy, bacterial segmentation, DeLTA 2.0**

I. Introduction

Tuberculosis (TB) continues to pose a major global health challenge, not only because of its disease burden but also due to delays in diagnosis and treatment initiation. According to the World Health Organization, more than ten million people developed TB in 2023 and over one million deaths were reported [1]. The increasing prevalence of drug-resistant forms of TB further underscores the need for rapid and reliable methods to determine antibiotic susceptibility early in the course of infection.

Phenotypic drug susceptibility testing (pDST) directly measures bacterial growth in the presence of antibiotics and remains the clinical gold standard. However, for mycobacteria, conventional pDST methods are slow and often require several weeks due to long doubling times. Recent advances in microfluidics and time-lapse phase-contrast microscopy enable monitoring of bacterial growth at the single-cell level, reducing the time to susceptibility readouts to hours rather than weeks [2].

Despite these advances, automated analysis of microscopy data remains a major bottleneck. Rapid cell motion, overlap, and imaging variability lead to segmentation and tracking errors, which in turn affect growth-rate estimation and downstream susceptibility analysis.

In the work by Tran et al. [2], Omnipose is used to perform frame-wise segmentation of phase-contrast time-lapse images acquired from microfluidic microchambers. Rather than reconstructing long single-cell trajectories, growth is quantified at the population level by monitoring the temporal expansion of the total segmented cell area within each microchamber and estimating growth rates using sliding-window exponential fits. This strategy enables rapid phenotypic drug susceptibility testing without requiring stable cell identity tracking, but it also means that growth estimates remain sensitive to segmentation variability, particularly in dense regions with overlapping cells. Cell overlap can lead to apparent decreases in measured cell area that do not reflect true biological growth inhibition or drug susceptibility.

DeLTA 2.0 represents a complementary approach, as it integrates segmentation, tracking, and growth-rate estimation into a unified single-cell analysis pipeline [3]. By explicitly maintaining cell identities across frames and estimating growth rates at the single-cell level, DeLTA 2.0 has the potential to provide more stable temporal growth measurements. This motivates evaluating whether such a lineage-aware pipeline can be adapted to microfluidic pDST data and whether it can improve discrimination between drug-susceptible and non-susceptible populations compared to frame-based approaches.

The aim of this work is to evaluate whether the DeLTA 2.0 single-cell analysis pipeline can be applied to time-lapse microscopy data from microfluidic pDST experiments and produce reliable distinctions between drug-susceptible and non-susceptible bacterial populations. This evaluation is carried out by:

- assessing the robustness of DeLTA 2.0 growth-rate estimates under realistic imaging conditions,
- analyzing population-level growth differences between control and drug-treated samples,
- identifying limitations related to tracking accuracy and data quality that affect early susceptibility detection.

II. Related Work

Deep-learning–based segmentation methods have substantially improved bacterial cell detection in microscopy images. Omnipose, an extension of the Cellpose framework, provides morphology-independent segmentation and was used by Tran et al. [2] to segment phase-contrast time-lapse images from microfluidic microchambers. In their analysis, growth is quantified at the population level by tracking temporal changes in total segmented cell area using sliding-window exponential fits. Because this readout is based on frame-wise segmentations rather than lineage-resolved trajectories, growth estimates remain sensitive to segmentation variability, particularly in dense regions with overlapping cells.

In contrast, lineage-aware approaches explicitly preserve cell identities across time, which has been shown to be important for accurate quantification of single-cell growth dynamics in microfluidic environments. Lugagne et al. [4] demonstrated that reliable tracking enables stable estimation of growth rates and division events by following individual cells over extended time periods. DeLTA 2.0 builds on this lineage-aware paradigm by integrating segmentation, tracking, and growth-rate estimation into a unified deep-learning pipeline. This motivates evaluating whether such an approach can be adapted to microfluidic pDST data and whether lineage-aware growth summaries can complement or improve upon frame-wise segmentation-based analyses.

III. Materials and Methods

The data used in this study consist of time-lapse phase-contrast microscopy images of *Mycobacterium smegmatis* growing in microfluidic chips, originally published by Tran et al. [2] [5]. The dataset was generated for rapid phenotypic drug susceptibility testing (pDST) experiments using microfluidic cell traps and high-frequency imaging. Images were acquired at two-minute intervals over a total duration of four hours, corresponding to 120 consecutive frames for the *M. smegmatis* experiments.

In addition to the phase-contrast images, the dataset includes corresponding segmentation masks generated automatically using Omnipose. These Omnipose-generated masks were used throughout the analysis as the available annotation source for segmentation. However, the masks frequently contain overlapping and merged cells, limiting their suitability as accurate single-cell ground truth and introducing uncertainty in downstream tracking and growth estimates.

To reduce the prevalence of merged and overlapping cells in the Omnipose-generated segmentation masks, additional binary morphological processing was applied. Each labeled object (cell/cells) in the Omnipose mask was processed independently. First, objects were binarized and a Euclidean distance transform was computed to estimate the distance of each foreground pixel to the nearest background. Local maxima in the distance map were detected to approximate individual cell centers. These maxima were then used as markers for marker-controlled watershed segmentation applied within each connected component, enabling separation of merged objects into multiple candidate cells. Boundary pixels between watershed regions were removed, and the resulting regions were combined into a refined binary mask. This procedure aimed primarily to increase object separation and improve downstream tracking stability, rather than to produce fully accurate single-cell ground truth.

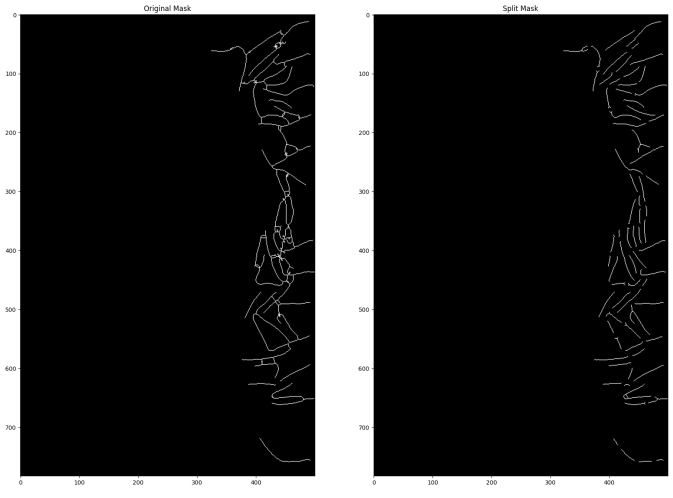

Fig. 1. Comparison between the skeleton of an original Omnipose-generated mask (left) and the skeleton of the corresponding binary-refined split mask (right).

Figure 1 compares the skeletons of an original Omnipose-generated mask and the corresponding binary-refined split mask, illustrating increased separation of individual objects after refinement.

Cell segmentation, tracking, and feature extraction were performed using the DeLTA 2.0 single-cell analysis pipeline. DeLTA 2.0 combines a U-Net–based segmentation model with automated tracking to link detected cells across time-lapse frames and reconstruct cell trajectories. To improve performance under microfluidic pDST imaging conditions, the segmentation U-Net component of DeLTA 2.0 was retrained on 2 420 image–mask pairs. Training was performed using the binary-refined variants of the Omnipose-generated masks. As no manually annotated ground truth was available, these automated masks were used as supervisory signals during training. The aim of retraining was to improve segmentation consistency and reduce merged objects rather than to obtain fully accurate single-cell annotations.

A. DeLTA 2.0 Pipeline Overview

The DeLTA 2.0 workflow used in this project consists of three stages. First, a segmentation U-Net predicts a binary mask for each frame. Second, the tracking component links segmented objects across consecutive frames to form cell trajectories by assigning consistent identities over time. Third, growth features are extracted from each tracked cell, including cell area and cell length, enabling construction of growth-rate time series. In this work, the segmentation model was retrained to better match the imaging conditions and mask characteristics of the microfluidic pDST dataset, while the

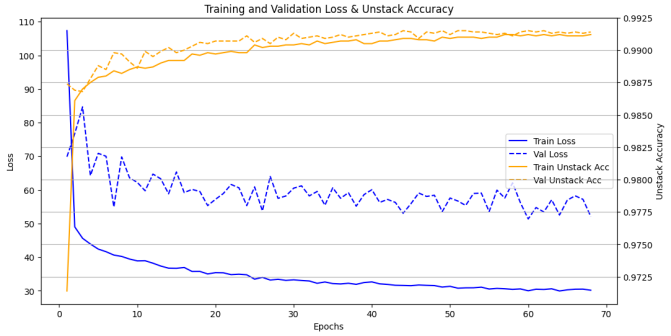

Fig. 2. Training and validation loss together with the unstack accuracy reported by DeLTA 2.0 during retraining of the segmentation U-Net.

downstream tracking and feature extraction modules were applied using the standard DeLTA 2.0 pipeline.

B. Loss Function and Training Monitoring

Accurate separation of individual cells was particularly important because the DeLTA 2.0 segmentation U-Net is optimized using a pixelwise-weighted binary cross-entropy loss with a skeleton-based weighting scheme. In this formulation, a spatial weight map is provided as an additional input during training, assigning the highest weights to the skeletons (medial axes) of individual cells and elevated weights to cell borders, while minimizing the contribution of background pixels. This design emphasizes correct classification along the cell interior and reduces sensitivity to ambiguous boundary regions. Consequently, segmentation errors caused by merged or overlapping cells disproportionately affect the loss and can degrade downstream tracking performance.

Training and validation loss, together with the unstack accuracy reported by DeLTA 2.0, were monitored throughout training. Unstack accuracy is an internal framework metric reflecting the model's ability to separate touching or overlapping cells during segmentation. Figure 2 shows the evolution of these metrics across epochs for both the training and validation partitions.

C. Experimental Design and Evaluation

Evaluation was performed on seven paired control–drug positions, each consisting of 120 time-lapse frames acquired at two-minute intervals. The analysis was carried out per position, and drug response was assessed by comparing each drug-treated position to its paired control recorded under the same experimental setup. Because no manually annotated ground truth was available and the Omnipose-derived labels contain merged objects, object-level segmentation and tracking metrics (e.g., Dice/IoU, MOTA/MOTP) were not computed, as such metrics would not be reliable under label inconsistencies and one-to-many (split) correspondences.

Instead, growth-rate curves were constructed from per-cell measurements extracted from the DeLTA 2.0 outputs. For each frame, cell area and length were computed from segmented masks, and frame-to-frame growth rates were estimated from

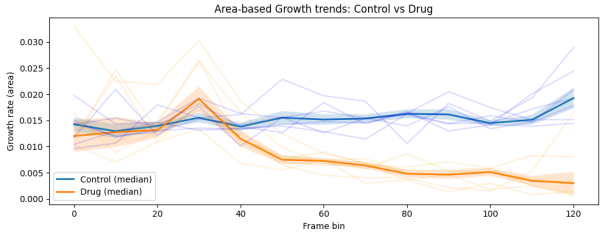

Fig. 3. Area-based growth-rate trajectories for drug-treated (red) and control (orange) cells aggregated at the population level.

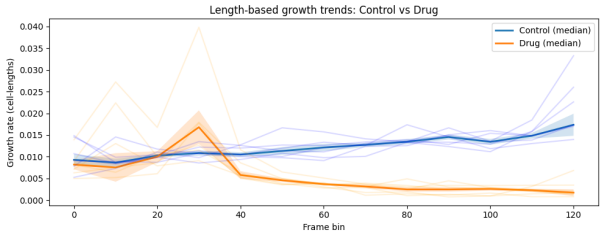

Fig. 4. Length-based growth-rate trajectories for drug-treated (red) and control (orange) cells aggregated at the population level.

changes in these features along tracked trajectories. Because tracking was affected by label switching, fragmentation, and occasional spurious measurements, growth-rate values were aggregated into non-overlapping 10-frame bins (corresponding to 20 minutes) and summarized using the median within each bin. The median was chosen to reduce the influence of outliers introduced by noisy or short-lived tracks. Population-level trajectories were computed by aggregating across cells within each position. Drug response was quantified as the percent change in the area under the smoothed population growth-rate curve (AUC) relative to the paired control.

IV. Results

Figure 2 shows the evolution of training and validation loss together with the unstack accuracy during retraining of the DeLTA 2.0 segmentation U-Net. Both loss curves decrease and stabilize over epochs, indicating convergence of the training process.

Population-level growth-rate trajectories differed between drug-treated and control conditions. Figure 3 shows area-based growth-rate trajectories aggregated across cells for drug-treated and control samples. Figure 4 shows the corresponding length-based growth-rate trajectories.

Drug response was further quantified using the area under the growth-rate curve (AUC). Table I reports AUC values for paired control and drug-treated positions, together with the absolute and percent change relative to the paired control. Percent changes ranged from -7.9% to -81.8% across positions.

V. Discussion

This study evaluated the applicability of DeLTA 2.0 to microfluidic pDST time-lapse microscopy data in the absence of manually annotated ground truth. While the segmentation

Position	AUC Control	AUC Drug	ΔAUC	Percent Change (%)
115	4.451433	0.809448	-3.641986	-81.816
117	4.301386	1.885018	-2.416368	-56.177
114	4.123103	1.974695	-2.148408	-52.107
112	4.343609	2.654542	-1.689067	-38.886
111	3.776717	2.789170	-0.987547	-26.148
113	4.098492	3.582340	-0.516152	-12.594
116	3.453173	3.181973	-0.271200	-7.854

TABLE I

AREA UNDER THE GROWTH-RATE CURVE (AUC) FOR CONTROL AND DRUG-TREATED CONDITIONS, TOGETHER WITH ABSOLUTE AND PERCENT CHANGE RELATIVE TO THE PAIRED CONTROL.

U-Net was successfully retrained and achieved high internal unstack accuracy, tracking performance remained unstable in several sequences. Label switching and identity fragmentation were observed, primarily due to inconsistencies and merged objects in the Omnipose-derived supervisory labels. Because these labels do not always accurately reflect individual cells in the underlying images, maintaining consistent cell identities across frames is challenging and limits reliable single-cell trajectory reconstruction.

Despite these tracking instabilities, population-level growth summaries showed consistent differences between drug-treated and control conditions. Both area-based and length-based growth-rate analyses exhibited reduced growth under drug exposure, and all evaluated positions showed negative percent changes in AUC relative to their paired controls. Aggregating growth metrics using temporal binning and median statistics reduced sensitivity to failed tracking and outlier trajectories, allowing robust population-level comparisons even when individual cell tracks were unreliable.

A key difference compared to Tran et al. [2] is that their susceptibility readout relies primarily on frame-wise segmentation and population-level morphology trends derived from Omnipose [6], without requiring stable long trajectories for individual cells. In contrast, DeLTA 2.0 aims to maintain cell identities over time to enable single-cell growth-rate estimation. In the present dataset, this lineage-aware approach provides an interpretable framework for growth analysis, but its benefits are limited by tracking instability and by the quality of available supervisory labels. These findings suggest that frame-wise approaches may be more robust when annotations are weak, whereas lineage-aware tracking becomes more informative when reliable ground truth and stable imaging conditions support consistent identity assignment.

Notably, the population-level growth inhibition observed in this study is consistent with findings reported by Tran et al. [2]. In our Figures 3 and 4, separation between drug-treated and control growth-rate trajectories becomes apparent during the early portion of the time series. Tran et al. report in their study that susceptible strains can be identified within approximately one hour after antibiotic exposure, corresponding to roughly 60 frames following drug introduction at frame 30. The early divergence observed in our population-level growth summaries therefore aligns with the timescale reported for phenotypic susceptibility detection, suggesting that DeLTA 2.0–based population analyses can support early pDST readouts comparable to frame-wise approaches.

These results suggest that DeLTA 2.0 can be used to extract meaningful population-level growth trends from microfluidic pDST data. However, in the present setting, reliable single-cell susceptibility analysis is not feasible. The lack of high-quality ground truth annotations, together with frequent segmentation inconsistencies and tracking failures, prevents stable reconstruction of individual cell trajectories and precludes robust single-cell growth-rate estimation. As a result, both quantitative evaluation of tracking performance and early susceptibility decisions at the single-cell level are not achievable with the available data. Future work should therefore focus on generating high-quality manual annotations for a representative subset of frames to enable objective benchmarking of segmentation and tracking, as well as on improving tracking robustness under conditions of rapid motion and cell overlap.

Several methodological improvements could strengthen the applicability of lineage-aware pipelines such as DeLTA 2.0 to microfluidic pDST data. First, the availability of manually annotated ground truth for a limited subset of frames would enable objective benchmarking of both segmentation and tracking performance and allow targeted retraining of the tracking component. Second, growth analysis strategies that do not rely on long, uninterrupted single-cell trajectories could be explored, for example by estimating growth from short track fragments or frame-to-frame changes and aggregating these measurements at the population level. Such hybrid approaches could provide a compromise between lineage-aware analysis and purely frame-wise methods when tracking stability is limited. Finally, improved imaging conditions, such as higher frame rates or reduced cell density per microchamber, would likely alleviate several of the observed tracking failures and improve downstream growth-rate estimation.

ACKNOWLEDGMENT

The author thanks the course staff for feedback and Tran et al. for providing the dataset used in this project. All code used in this work is publicly available at this repository https://github.com/joelsundin/forskningsmetodik_2025_Delta2.0

REFERENCES

[1] World Health Organization, "Global tuberculosis report 2024," Geneva, Switzerland, 2024, accessed 2025. [Online]. Available: https://www.who.int/teams/global-tuberculosis-programme/tb-reports

[2] T. Tran *et al.*, "Phenotypic drug susceptibility testing for mycobacterium tuberculosis variant bovis bcg in 12 hours," *bioRxiv*, 2025.

[3] D. Bannon *et al.*, "Delta 2.0: a deep learning pipeline for quantifying single-cell growth dynamics in time-lapse microscopy," *bioRxiv*, 2021.

[4] J.-B. Lugagne, S. S. Carrillo, M. Kirch, H. Klinker, D. Camsund, G. Batt, and P. Hersen, "Balancing a genetic toggle switch by real-time feedback control and periodic forcing," *Nature Communications*, vol. 11, no. 1, pp. 1–11, 2020.

[5] B. M. Tran, J. Larsson, A. Grip, P. Karempudi, and J. Elf, "Microscopy data, analysis code, and segmentation models for phenotypic drug susceptibility testing for *Mycobacterium tuberculosis* variant *bovis* bcg in 12-hour experiments," 2025.

[6] K. J. Cutler, C. Stringer, P. A. Wiggins, and J. D. Mougous, "Omnipose: a high-precision morphology-independent solution for bacterial cell segmentation," *Nature Methods*, 2022.