Artificial Intelligence Predicts Cell Proliferation from DAPI images of (hISC)-derived colorectal cancer model

Maurycy Chronowski1, Konrad Wojtowicz1, Krzysztof Rataj1, Agata Stachowicz-Walaszek2, Marcin Duleba2, Eliza Zimolag2, Joanna Drozdowska2 , Karolina Pyziak2, Andrzej Mazan2, Magdalena Otrocka1, Michał Warchoł1

1 Ardigen, Kraków, Poland; 2 Ryvu Therapeutics, Kraków, Poland Contact: maurycy.chronowski@ardigen.com

INTRODUCTION

In recent years, Artificial Intelligence has paved its way into the world of Drug Discovery, proving to be a game-changing tool in the search for novel medicines. At the dawn of the first AI-designed drugs being approved by FDA, deep neural networks are finding a multitude of applications, allowing to reduce the cost and complexity of the experiments.

In this work, we demonstrate how an AI-based pipeline can significantly enhance and simplify cell proliferation detection in human intestinal stem cell (hISC)-derived cancer cell lines cultured on a monolayer of supporting mouse feeder cells. While this culture method brings the research closer to real biological conditions, it also generates image analysis challenges that significantly increase experimental costs. We propose the application of a deep segmentation model (U-Net) which we use to robustly classify parts of the images that contain proliferating cancer cells stained with DAPI, effectively removing the need to perform a costly EdU cell proliferation assay on the whole screening. We demonstrate that our model, after being trained on a preliminary batch of measurements, transfers to new experiments performed on a similar cell line achieving results comparable to those obtained from EdU positive nuclei count.

DATA

- We used the dataset generated during phenotypic screening campaign performed on human intestinal stem cell (hISC)-derived cancer cell lines cultured on a monolayer of supporting mouse feeder cells.
- In the primary screening, approximately 4000 compounds were tested in triplicates on AKTS (AKT and SMAD4 knockout) cell line.
- This was followed by eighth concentrations dose response study for 300 selected primary hits on AKTS cells and counter screening on WT (wild type) cell line.
- The cells were co-stained with DAPI to visualize all nuclei and EdU reagents to detect nuclei of proliferating (cancer) cells.
- Imaging was done using Nikon Ti2-E fluorescent microscope in DAPI and FITC (for EdU) fluorescence channels at 4X magnification (1 image per well).

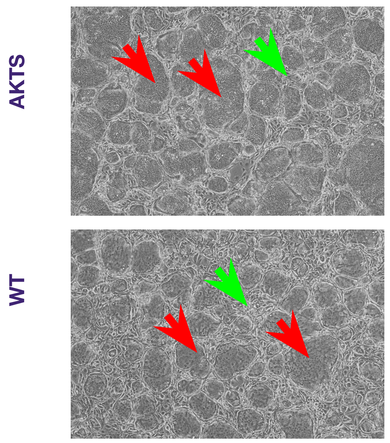
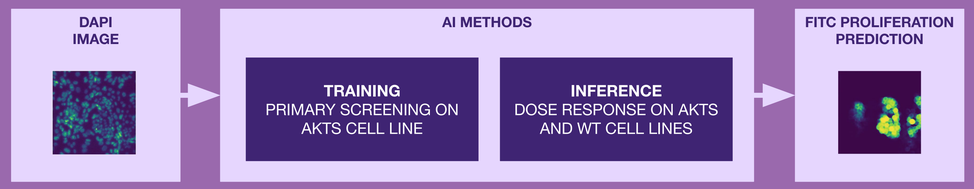


Figure 1: Bright field images of AKTS (uper) and WT (down) human intestinal stem cell (hISC)-derived cancer cell lines (red arrows) cultured on a monolayer of supporting mouse feeder cells (green arrows). Co-culture models are more biologically relevant, however, the analysis of proliferation of different cell types is challenging.

GRAPHICAL ABSTRACT



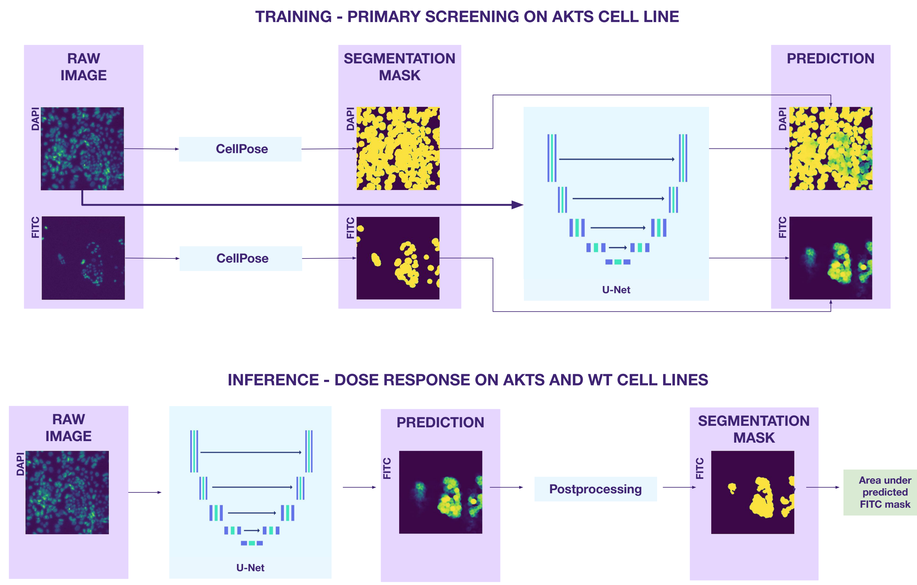


Figure 2: The overview of developed method. Top panel depicts the training procedure performed on images generated during primary screening on AKTS cell line. Bottom panel shows the inference procedure performed on images from dose response study performed on AKTS and WT cell lines.. Both CellPose and U-Net models are open-source, available online ([2], [3]). We keep CellPose frozen, while U-Net is being finetuned.

METHODS

Training on Primary screening dataset:

- We generated segmentation masks from DAPI and FITC images using an open-source, pretrained Cellpose model [1] and use them as ground true in the next step.
- We finetune open-source U-Net model pretrained on instance segmentation task [2]:
  - Dataset consisting of 14 plates was split into train and test folds of sizes 11 and 3, respectively.
  - The images and masks were randomly cropped to square patches of size 256px during training. Augmentation consists of random rotation and flips. The model was trained for 23 epochs.
  - The model predicts DAPI and FITC segmentation simultaneously, from raw DAPI input images only.

Inference on Dose-Response dataset:

- We inferred our trained U-Net model on imaged KO cells (same as in training) and WT cells (different from training).
- To clean up the predicted masks, we thresholded U-Net outputs (sigmoid-activated) using Otsu thresholding clipped to a minimum value of 0.1 (to avoid segmenting noise from empty images). We then applied top-hat transformation to remove small segmentation debris.
- We then calculated the total area of the predicted binary FITC mask and compared it to the actual FITC cell count measured in laboratory experiments.

IC50 estimation:

- For each tested compound, we normalized the predicted area to: max = negative control, min = positive control and fitted a sigmoid curve using non-linear regression, 4-parameter logistic model

REFERENCES

[1] Stringer, C., Wang, T., Michaelos, M., & Pachitariu, M. (2021). Cellpose: a generalist algorithm for cellular segmentation. Nature methods, 18(1), 100-106.
[2] UNet for Instance Cell Segmentation on Pytorch, PARMAGroup, https://github.com/PARMAGroup/UNet-Instance-Cell-Segmentation/tree/master

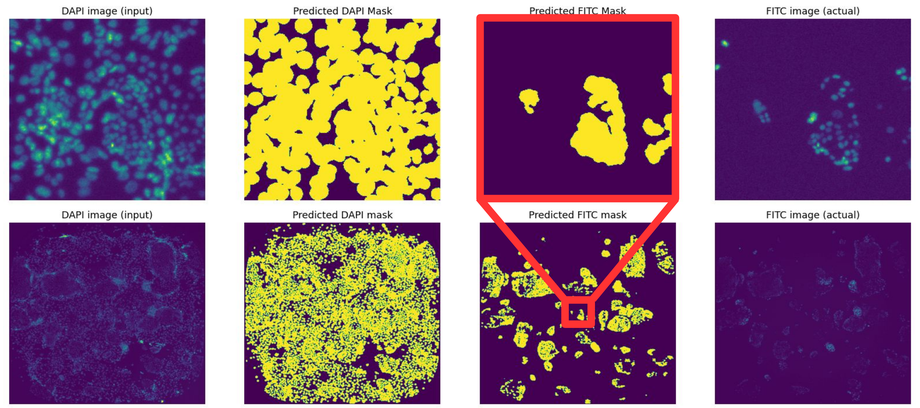


Figure 3: Example of processed image. DAPI image shown in the left panel is used to predict segmentation masks not only for DAPI staining (left inner) but pressly predicts FITC signal signal visualising EdU staining (right inner). The original raw FITC image is not used for the prediction and presented in the left panel for the reference. Top row shows the single patch, as processed within a model, while bottom shows entire well.

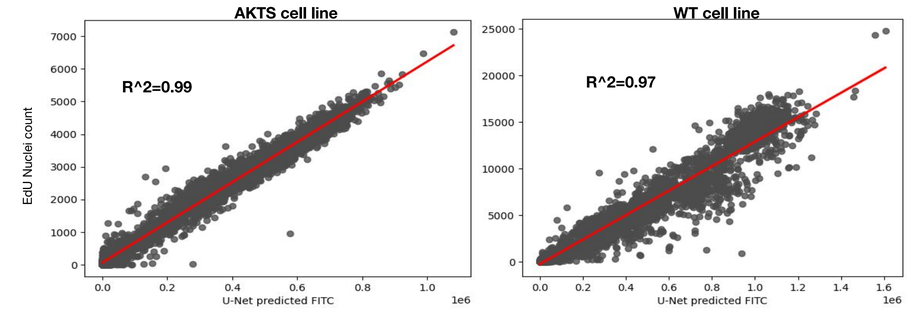


Figure 4: The area measured under FITC mask predicted from DAPI images correlates well with the EdU (FITC) stained proliferating nuclei for both AKTS (left) for which the model was trained, as well as for WT (right) cell line. This shows that our model can be used to asses the cell proliferation from DAPI images in different types of cell co-culture models.

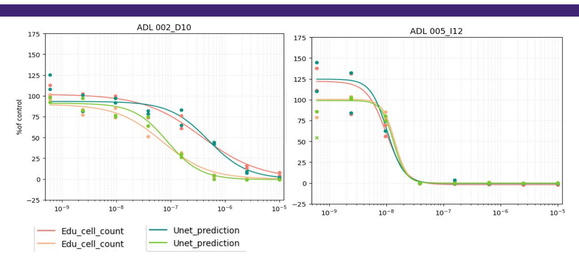


Figure 5: Examples of dose-response curves for AKTS and WT cell lines generated with the outputs from different cell proliferation assessment methods. The curves obtained based on the prediction of cell proliferation from DAPI images are overlapping with those obtained based EdU based cell nuclei count.

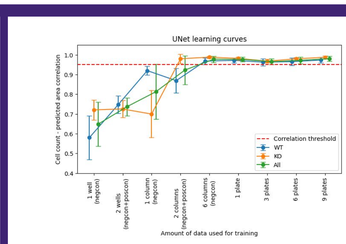


Figure 6: Learning curves for the model. Our method is capable of learning to predict the cell count with as little data as the negative controls in 3 replicates (2 columns/plate) or a single, non-replicated plate.

KEY CONCLUSIONS

- **Our method can replace the costly and time-consuming EdU staining procedure allowing for increasing throughput of compound testing on complex cancer models.**
- **U-Net model accurately predicts the level of cancer cell proliferation from images of DAPI-stained nuclei in complex co-culture models.**
- **Quantification of EdU signal predicted from DAPI staining can be used for IC50 estimation of tested compounds.**
- **Our method requires minimal amount of data to train while retaining the capability to transfer between cell lines.**