Semester report

by **Hamid Cheraghi** (hamidchr@student.elte.hu)

PhD Program: Biological physics, Statistical physics, and Quantum physics

Supervisor: Bálint Szabó

Ph.D. Thesis title: Image-based single cell isolation with deep learning

Introduction

My research at the Eötvös Loránd University focuses on single-cell isolation of yeast and unicellular parasites. In this project, one of the main objectives is to develop hardware and software tools to isolate *T brucei* cells for subsequent DNA/RNA sequencing. Further goal of the project is to establish different strategies to isolate cells by an automated micropipette from agar plates and, if possible, from microfluidic devices with the aim to maintain single-cell lineages separated for further downstream analysis (e.g., next generation sequencing).

Cell Cycle

Cell cycle is the series of events that takes place in a cell as it grows and divides. A cell spends most of its time in what is called interphase, and during this time it grows, replicates its chromosomes, and prepares for cell division. The cell then leaves interphase, undergoes mitosis, and completes its division. The resulting cells, known as daughter cells, each enter their own interphase and begin a new round of the cell cycle [1]. Figure.1 shows the illustration of cell cycle.

FUCCI Cell Cycle reporter construct

The FUCCI (Fluorescent Ubiquitination-based Cell Cycle Indicator) system was developed by Sakaue-Sawano et al. as a tool to visualize the dynamics of cell cycle progression. HeLa-Fucci cells periodically express two fluorescent protein tagged molecules namely Geminin fused with monomeric Azami green (mAG-hGem) and Cdt1 fused with monomeric Kusabira Orange (mKO2-hCdt1) the expression levels of which oscillate periodically during the cycle: mKO2-hCdt1 levels are high during G1, while mAG-hGem levels are high during the S/G2/M phases.

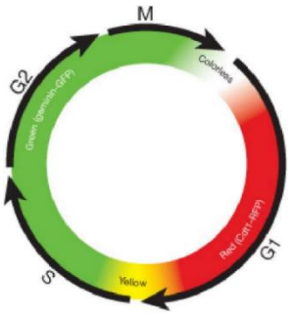

Figure.1 Two-color FUCCI cell cycle.

In this project, I am working on a dataset of fluorescent and phase contrast HeLa cell images. HeLa cells are the most widely used human cell line in biological research [2]. My dataset contains 324

Phase contrast images as well as the corresponding Green Fluorescent Protein (GFP) and the Red Fluorescent Protein images to the Phase contrast images. Images were acquired on an automated microscope by me.

Description of my current research

Our project includes two major stages. The first step is to generate a ground truth dataset using Artificial Intelligence (AI) algorithms. To do so, we utilized a generalist algorithm called Cellpose [3] for cell segmentation. So, we could segment cells and generate mask images and outlines for each cell in the unlabeled phase-contrast images. Figure 2 demonstrates an example of segmentation of the phase contrast images.

When we are able to obtain the mask and outline for each cell in a given phase contrast image, we segment the fluorescent red and green images using the mask determined in the phase contrast images. The next part of this stage is to crop the image of cells in all 3 channels as well as the mask image. To do so, we need to achieve coordinates of each cell to be able to draw a rectangle around them. Also, we can annotate each rectangle and put the number of cells next to the rectangle. Figure 2 shows an example of the mask, the red, and the green images along with rectangles and the annotations.

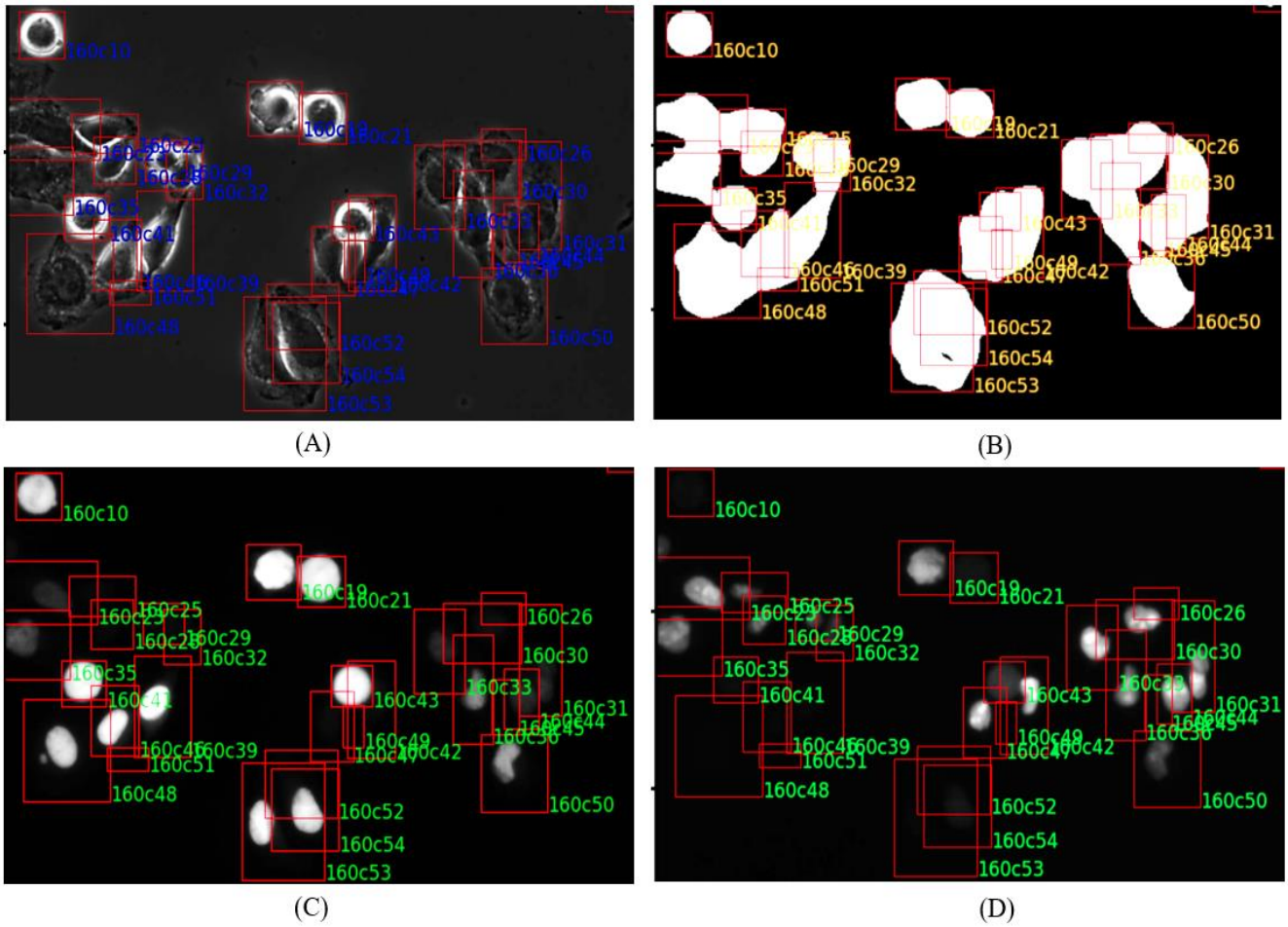

Figure.2 An example of phase contrast, mask, red, and the green images along with rectangles and the annotations. (A) a phase contrast image, (B) a generated gray mask, (C) a fluorescent red image, (D) a fluorescent green image.

The next step of this part is to calculate the sum and the spatially averaged fluorescent intensity of each cell in both red and green channels. By this, we can see the fluorescent signal of each cell.

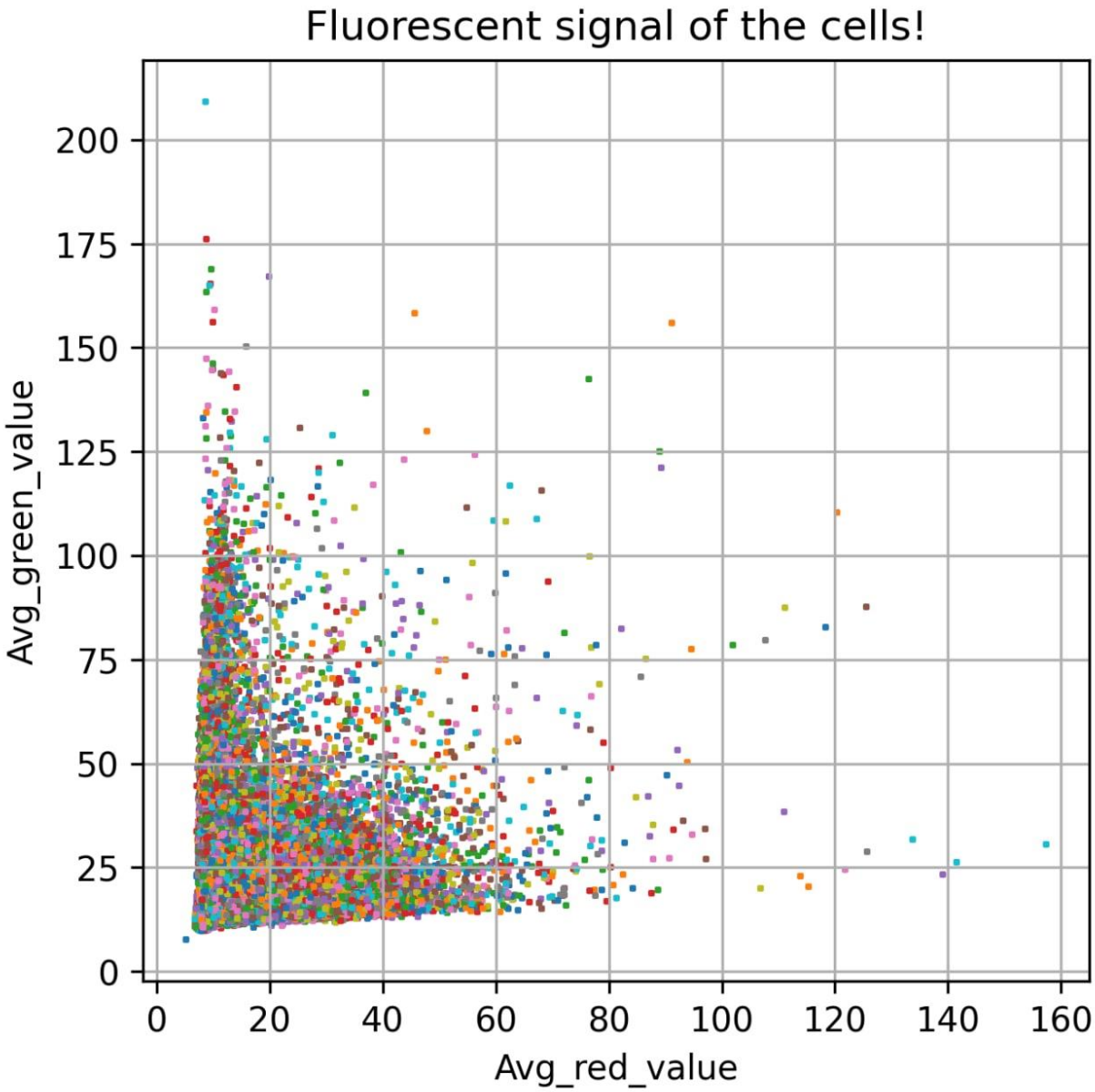

Figure.3 Average intensity of red fluorescent protein and green fluorescent protein for 42,000 cells. (Average means the integral pixel intensity of cells divided by cell area.)

The second step is implementing a deep Convolution Neural Network (CNN) for cell cycle estimation (prediction). In this step, we use pre-trained CNN networks such as: ResNet50, VGG16, and the Xception model. On the other hand, we could implement our own network for cell cycle prediction. Figure 4. demonstrates the pipeline of our work.

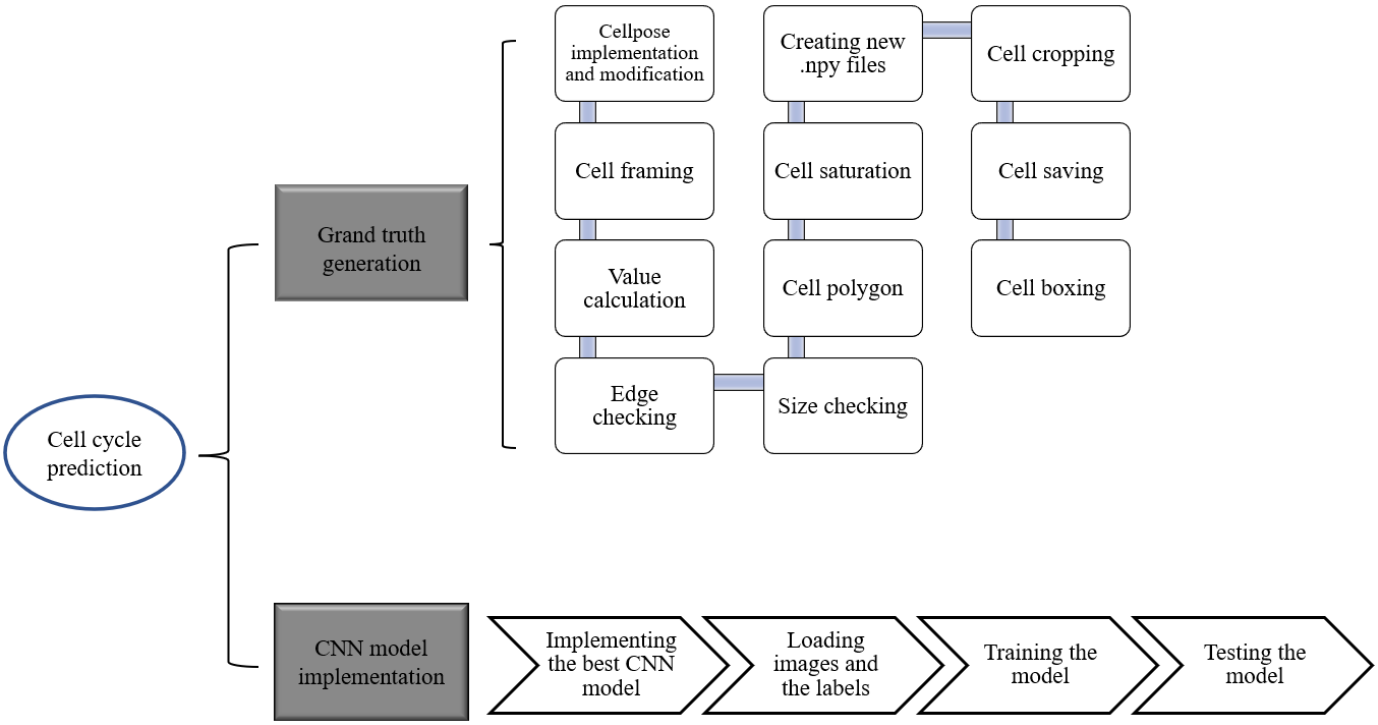

Figure.4 Cell cycle prediction pipeline. I built an automated computational workflow for image processing and subsequent cell phase prediction.

The current performance of our neural network is 9.6 percentage error in both red and green channels.

We are almost ready to start writing our paper with the title of: Deep neural network predicts cell cycle phase based on a standard phase contrast image.

Authors of the paper are: Hamid Cheraghi, Kinga Dóra Kovács Inna Székács, Robert Horvath, Bálint Szabó

Study activities

In this semester, I continued my studies by taking courses entitled “Statistical physics of polymers and membranes” (Neptun code: FIZ/3/021E). Also, I completed the course with the title “Sensory biophysics” (Neptun code: FIZ/3/010E).

During the semester, I participated in weekly lectures organized by the Ludwig Maximilian University (LMU) of Munich in the Cell2Cell ITN: https://cell2cell.eu

Speaker: Dr. Ritwick Sawarkar, MPI Immunology and Epigenetics, Freiburg

Title: Cellular response to stress: a chromatin perspective

Speaker: Ritwick Sawarkar, MPI Immunology and Epigenetics, Freiburg

Title: Chaperones@chromatin.nucleus: an emerging address of an ancient family

Speaker: Claudia Keller Valsecchi, Institute of Molecular Biology Mainz (IMB)

Title: will be indicated later

Speaker: Claudia Keller Valsecchi, Institute of Molecular Biology Mainz (IMB)

Title: Cis-acting non-coding RNAs and X chromosome dosage compensation

I participated in Advanced FISH technologies and related analysis workshop offered by Cell2Cell ITN: https://cell2cell.eu.

Organizers of this workshop were Karolinska Institute (https://ki.se/en) and Scilifelab (https://www.scilifelab.se/), Stockholm, Sweden.

The Imaging workshop overview is as follows:

Day1: probe hybridization

- Prepare the hybridization mixture 1:

  99 µl of the RNA HYB 25% buffer + 1 µl of a probe
- Immerse coverslips from the ethanol-containing dish in the RNA WASH buffer
- Incubate for 5 min at RT
- place the coverslip with cells facing down on 100 µl hybridization mix1 in a humidity chamber
- Seal the chamber with parafilm, and incubate for 24 h at 30 C

Day2: fluorescent oligo hybridization

- Prepare the hybridization mixture 2:

  99 µl of the RNA HYB 25% buffer + 1 µl of a fluorescently-labelled oligo
- Transfer each coverslip into 6 cm dishes filled with 2 ml RNA WASH 25% buffer
- Incubate 30 min at 30C
- place the coverslip with cells facing down on 100 µl hybridization mix 2 in a humidity chamber
- Seal the dish with parafilm and incubate for 24 h at 30 C

Day3: Mount slides

- Wash with pre-warmed 0.2xSSC/0.2% Tween in a 6-well plate floating on the water bath surface at 45°C for 7’ 2hours Lectures+2 Hours Practicals
- Repeat the wash once
- Rinse with 2xSSC, RT
- Incubate with 1 ng/ul Hoechst 33342/2xSSC, 10’, RT
- Wash briefly a couple of times with 2xSSC
- Mount with the GLOX buffer & image

References

[1]. National Human Genome Research Institute: https://www.genome.gov/genetics-glossary/Cell-Cycle

[2]. British Society for Immunology: https://www.immunology.org/hela-cells-1951

[3]. Stringer, Carsen, et al. "Cellpose: a generalist algorithm for cellular segmentation." Nature Methods 18.1 (2021): 100-106.