Career Properties

I didn't write much code at the beginning of the project, but I tried to get familiar with using OpenCV code (the use of this package is forbidden in the 373 courses). The class was still beneficial in that it gave me an understanding of images and how the various algorithms for image processing work, opening my mind to conceptualize better how the problem should be solved. The self-learning skills I developed in school, and the ability to meet deadlines were very helpful because I always met the time limit, and for a project that involved much knowledge, the ability to learn on my own helped me get into the project as quickly as possible.

I had some innovative ideas when researching how to find contigs' peaks properly. For example, if a contig finds a certain contig that is the peak, then go to that contig and look for it. If that contig is also found, add 1 to the score, and the final relationship with the highest score will be identified as interconnected.

About network skills, I have to express my ideas and algorithms in python code, and I've practiced this skill very well during this internship。

Development process

During the first two weeks of my internship, my main task was to familiarize myself with the principles and requirements of the project; my supervisor Chen gave me lots of help with the project because this is a biological project, and I needed that information. During the third and fourth weeks, the IT supervisor and I had regular meetings, and on Mondays, we would review the previous week's progress and discuss what we should do next.

During the third week, the project focused on how to generate a specific heatmap using code rather than using online tools and taking screenshots. The project started in week 4 with the task of partitioning a heatmap into 566 2D matrices based on the number of contigs in order to detect special blocks better.

The project needed to detect peaks on each 2D matrix in the fifth week. Due to the project's specificity, we thought it would be more appropriate to convert the 2D array into 1D and perform peak detection. This week we focused on how to convert 2D into 1D and give a suitable smoothing algorithm and then perform detection.

In week 6, the peek detection was completed, and depending on the contig detected, the code would return the position of the detected peek on the x-axis. Another function I wrote would determine which contig the position was in.

In the seventh week, at first, we intended to do the sorting. However, we found that due to biological peculiarities, peak detection is impossible to be precise, so we used another method to convert 2D to 1D, which previously averaged the values of each column to generate a sequence of length 5503 before processing (the original image size is 5503*) 5503), this time we plan to average the 2D array of each contig to find a value, which will generate a sequence of length 566 (since there are a total of 566 contigs) and operate afterward. Finally, it was found that since some contigs have a threshold set to a specific value that is appropriate but not feasible for others, a particular operation is needed to filter the results after the first round of detection.