Introduction

Cancer claimed to be one of the leading causes of death in the world, and it is just second to ischemic heart disease. While unintentional injury such as road accident is accounted for third place of leading causes of death globally [1]. According to the World Health Organization, approximately 9.60 million and 1.35 million deaths each year for cancer and road accidents respectively [2, 3]. While the risk of getting cancer or injuries can be greatly reduced by adopting a healthy lifestyle or paying more attention to the environment respectively, however, accidents still always happen. Once it does, the key point to save lives is to provide dedicated treatment at the earliest possible moment. Thus, the development of those treatments becomes essential in medical fields. To develop medicines or other treatments, various technology has been utilized. Compare to the past, today’s medical technology has reached a point where we can observe individual cells or even further smaller scale under various microscopy technology. With the study of the behavior of cells, many diseases and injuries including cancer that was determined to be impossible to be treated in the past became possible today. Among such studies, one of them is collective cell migration, which described as the movement of a group of cells with collective behavior towards the environment and among each other. It is the main pattern of locomotion of collective cells that contribute to cancer spreading and wound healing [4]. Due to that reason, many kinds of researches have been carried out to investigate the method to control the speed of the movement. If the movement can be slow down, cancer spreading will be slow down. In contrast, if the movement can be speed up, the wound healing effect will also speed up accordingly. Moreover, it is observed that special kinds of cells formed at the leading edges of the cells during collective cell migration, they are called leader cells. They act like leaders to the group of cells and are able to regulate the migration [5]. Thus, scientists are carrying out experiments revolving leader cells to investigate if its existence is the main reason for collective cell migration. Solving this enigma allowing us to take one step forward to develop the medicines that can slow down cancer spreading or speed up would healing, which helps to reduce deaths.

To study leader cells in collective cell migration, an affirmative issue is the identification of leader cells itself. Thus, the characteristics that can distinguish leader cells with other cells, also called follower cells need to be determined. One of those determinants is the fact that it has been observed that leader cells possess distinctively large amounts of proteins from Rho family of GTPases such as Rac1 and its downstream to compare to the follower cells [6]. By protein staining, the signal of those proteins is stronger at the leading edges. While identifying leader cells by protein staining provide a clear result, the staining process itself is considered invasive and might already have disturbed the behavior of collective cell migration.

In this research, a more direct, non-invasive method is proposed to identify leader cells more efficiently, which is utilizing computer vision with deep learning to identify leader cells through filopodia, a piece of basic information from leader cells’ morphology in a faster, more organized manner.

Materials and Methods

Fluorescent Staining

One vial of Alexa Fluor 488 Phalloidin (Invitrogen, USA) is dissolved in 1.5 mL of methanol and divided into 300 units with each unit contains 5 μL methanolic stock solution. The stock solutions are stored in -20°C fridge and protected from light until use. It can last for more than one year by storing it in this manner. First, the spent medium is removed, and the cells are rinsed using prewarmed PBS thrice. Then, fixation is done by soaking the cells in 4% Paraformaldelhyde Phosphate Buffer Solution (Wako Pure Chemical, Japan) for 30 minutes. After that, the cells are rinsed using PBS thrice and permeabilization is done by soaking the cells in 1% of Triton X-100 (Merck Millipore, Germany) diluted in PBS for 15 minutes. Next, the cells are rinsed using PBS thrice and blocking is done by soaking the cells in 1% of albumin from bovine serum (BSA, Wako Pure Chemical, Japan) diluted in PBS for 1 hour and 30 minutes. Last, the cells are rinsed using PBS thrice again and stained by soaking in 3 units (= 15 μL) of Alexa Fluor 488 Phalloidin methanolic stock solution diluted in 0.6 ml of PBS for one hour followed by 2 drops of Hoechst 33342 (Invitrogen, USA) diluted in 2 ml of PBS for another one hour.

Cell Segmentation Training

Cell segmentation training is done by using a deep learning program, UNet++ that written in Python. First, debug is required before running the training and it is done by using PyCharm (JetBrains, Czech), an Integrated Development Environment (IDE) used for
coding. Other than that, some basic requirements have to be fulfilled. For hardware requirements, a computer installed or comes with a CUDA enabled Nvidia GPU needs to be prepared. While for software requirements, Anaconda, a package that included Python and all basic supporting libraries and PyTorch, a library that used in computer vision needs to be installed. The training is done in Command Prompt that comes with Windows. The directory is navigated to the python files’ folder and a conda environment, a virtual environment that consists of all the required libraries is created. Then, the code for cell segmentation training is run using python in Command Prompt. The program will stop itself when the training is finished, and it usually takes one day for the dataset size (one hundred images) used in this research.

**Identification of Filopodia**

To identify filopodia, a plugin of Fiji, Filopodyan is used. Filopodyan can be used to identify and measure filopodia as long as the parameters and criteria are set. The threshold method Huang is chosen, and the ED iterations and LoG sigma are set at 5 and 2.6 respectively. The software uses these values to identify the cell’s body and filopodia. Due to there is no information or guideline for these values from the software at the moment, the values are obtained by trial and error. After that, small particles or noises that are not filopodia are filtered using a batch filter. The particles are filtered by setting Min max length and Max mean waviness at 0.15 and 1.0 respectively.

**Results**

**Cell Segmentation**

One hundred fluorescence images of size 256 x 256 are tested with the model trained by UNet++ of 9.16M parameters. The results produced show an IoU of 99% which is very high accuracy. As the results are depending on the quality of the masks, good quality masks made manually using Photoshop has greatly improved the chance of obtaining high accuracy. In Table 1, the filopodia in the first image appear very long and are well segmented. Even so, the filopodia in the second image are too small to be segmented and the segmentation appears more like the shape of lamellipodia. In the third image, although the cells are overlapping each other, it is still segmented according to the edges. For the last image, since the edges are very clear, the cells are all well segmented. Although the cell nuclei are also showing strong edges, the computer can even learn to ignore that through deep learning. As a result, only cells’ borders are segmented. The main dataset used for this training is fluorescently stained MDCK cell images (as in Table 1).

**Results of Filopodia**

The filopodia identification is only done on the main dataset, the one using MDCK cells as the training model. The reason is due to the parameters required for different cell types can be very different. The cell size, cell shape, filopodia length, filopodia shape, and others are the factors affecting the parameters. As the MDCK cells are the ones we want to assess in this research, the evaluation is focused on MDCK cells.

Firstly, Filopodyan segmented filopodia from the cell body. As in Table 2 first column, along with filopodia, there is also a huge amount of noise. After that, filters are applied to filter out the noises to identify where the filopodia are. With little error, the software identified 17 filopodia. It can identify large numbers of filopodia rapidly. However, accurately segment all filopodia is still difficult due to the variability of the images. While the filopodia that are unable to detect are impossible to add, the filopodia that are falsely detected can be erased manually. Filopodyan shows very good results in this research as the images are of high quality with high magnification. However, its disadvantages are it cannot be used on cells that have filopodia that are too short or too fine and images with cells overlapping each other where the filopodia are not under a high contrast background.

<table>
<thead>
<tr>
<th>Table 1. Four out of One Hundred Cell Segmentation</th>
<th>Cell Images</th>
<th>Segmentation Results</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Cell Image 1" /></td>
<td><img src="segmentation1.png" alt="Segmentation Result 1" /></td>
<td></td>
</tr>
<tr>
<td><img src="image2.png" alt="Cell Image 2" /></td>
<td><img src="segmentation2.png" alt="Segmentation Result 2" /></td>
<td></td>
</tr>
<tr>
<td><img src="image3.png" alt="Cell Image 3" /></td>
<td><img src="segmentation3.png" alt="Segmentation Result 3" /></td>
<td></td>
</tr>
<tr>
<td><img src="image4.png" alt="Cell Image 4" /></td>
<td><img src="segmentation4.png" alt="Segmentation Result 4" /></td>
<td></td>
</tr>
</tbody>
</table>
**Table 2. Segmentation and Identification of Filopodia**

<table>
<thead>
<tr>
<th>Segmentation of filopodia (before filter noise)</th>
<th>Identification of filopodia (after filter noise)</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image1.png" alt="Image" /></td>
<td><img src="image2.png" alt="Image" /></td>
</tr>
<tr>
<td><img src="image3.png" alt="Image" /></td>
<td><img src="image4.png" alt="Image" /></td>
</tr>
</tbody>
</table>

**Discussion**

This research basically divided into two parts. The first part is the culture of cells where we allow collective cell migration to occur in microchannels and promote the formation of leader cells with filopodia.

Collective cell migration is a process by which cells relocate from one part of the body to another by taking various motions but happens to a group of cells that retain cell-cell contacts with coordinated movement [4]. It possessed unique behavior by which the cell-environment interaction and cell-cell communication occur in a collective manner and the movement of the group of cells occurs simultaneously [7]. The migration can occur either in a cohesive group of cells or cells connected by cell-cell adhesion [8, 9]. During the migration, the cells are going through various changes in their cytoskeleton structure [10].

During collective cell migration, it is observed that there is an existence of special kinds of cells possessed unique characteristic, they are leader cells and they seem to be the ones that regulating the movement. They acquire distinctive morphology and motility. Although there are extensive researches done on cell migration, there are not many do on collective cell migration, especially on leader cells. Thus, many characteristics of leader cells have remained unclear. It is proved that if the leader cells are manually killed or removed from their colonies, the migration is disturbed, the cells no longer move in the same direction and the speed is significantly dropped [6]. From here we can see the importance of leader cells to lead the colonies to migrate directionally [11]. It helps the collective cell migration to move more effectively [12]. Leader cells do possess a decisively characteristic, filopodia. Leader cells have larger lamellipodia compare to follower cells. From its lamellipodia, there are finger-like structures protrusion, which are filopodia.

Filopodia are slender finger-like structures that extended from lamellipodium of cells in the leading edges of colonies of migrating cells [13]. They are protrusive actin cytoskeleton of cells that consist of bundles of actin filaments cross-linked structures [14]. Filopodia’s main functions are sensing the environment, direct migration, and communicate with other cells. From their function, we can understand filopodia’s role in collective cell migration, they directly affected the movement of cell colonies in cancer metastasis and wound healing [15, 16]. During sensation, filopodia sweep up and down and from side to side to search through the situation of the environment in front of the cells [17]. While during migration, they extend and retract to shift the cells forward.

For culturing cells in a microchannel, cells usually die within 4 or 5 days due to lack of nutrients as the cell to medium ratio is much smaller than culturing cells in a dish. As changing medium too often will also affect cell growth and has the risk to push out cells from the microchannel which will affect cell migration, it is better to use a smaller channel (around 50 μm) and fix the cells within 3 days after obtaining a good migration. There was formation of leader cells as we can clearly see their filopodia. However, due to there is a shift of the cells’ body to the front during the migration, the filopodia are not always visible even though it is a leader cell.

All three fluorescence dye used emit good fluorescence and do not fade away easily. Although the final combination used was Alexa Fluor 488 Phalloidin and Hoechst 33342, the combination of Rhodamine Phalloidin and Hoechst is recommended as their wavelength is further from each other and the overlapping of fluorescence signal occur less frequent.

In the second part of the research, computer vision with deep learning is proposed in response to the need for faster leader cell identification in collective cell migration. Firstly, masks are created using Photoshop and UNet++ is used to train a model for future segmentation. After the segmentation, Filopodyan is used to identify filopodia.

As Photoshop is the best available photo editing software today, the results are very satisfying where the cells in white pixels are separated from the background in black pixels by the edges of the cells and the cells have very detailed shapes. It greatly helps in cell segmentation training. However, some cells do not have clear edges. Thus, for those cells, the area surrounding the nucleus is drawn instead of the edge.

Both UNet++ and Filopodyan can produce good results providing the cell images are clear. As the images obtained using epifluorescence microscopy are not good enough, TIRFM is recommended as TIRFM allows us to obtain a high contrast image that can show a better difference between cells and their background.

Due to other cell types and other image modalities are not assessed, the performance remained unknown. Even for the same cell type, it is also very difficult to segment as the is variety such as the fluorescence is overexposed, cells are overlapping each other, and
other conditions. In conclusion, more cell images with different conditions are required to train the network, and it can greatly improve the success rate.

As other cells do not show filopodia as in Table 2, this method can be utilized to identify leader cells. The identification for one hundred images is done in just a few seconds and around 90% of the filopodia are accurately detected, 10% miss is due to filopodia that are undetected or falsely detected. However, since in different conditions the level of noise is different, sometimes the parameters for the segmentation and filters need to be fine-tuned.

As cells with filopodia are very few and the model trained is only using MDCK cells, the same experiment needs to be repeated a few more times. Furthermore, using the same experiment set up (microchannels) to culture many different types of cells and obtain different cell images are also required. In order to obtain more data other than the method mentioned above, a collaborator in the medical field can also provide us more data for the dataset required in deep learning training. In this research, other architectures are not assessed. Although many shreds of evidence have shown that the U-Net family is best for biomedical image segmentation, there might still have some other architecture that is more suitable for cell segmentation. To evaluate the outcome of this research, time-lapse fluorescence images are recommended. It can be achieved by culture cells with green fluorescent protein (GFP). Therefore, the cells are no longer needed to be fixed.

References