

# Automated Cell Counting using Image Processing

**DEWI KARTINI HASSAN, HAZWANI SUHAIMI, MUHAMMAD ROIL BILAD, PG EMEROYLARIFFION ABAS**

Faculty of Integrated Technologies, Universiti Brunei Darussalam, Gadong BE1410, Brunei

Corresponding authors: Hazwani Suhaimi (e-mail: [hazwani.suhaimi@ubd.edu.bn](mailto:hazwani.suhaimi@ubd.edu.bn)); Pg Emeroylariffion Abas (e-mail: [emeroylariffion.abas@ubd.edu.bn](mailto:emeroylariffion.abas@ubd.edu.bn)).

**ABSTRACT** Manual cell counting using Hemocytometer is commonly used to quantify cells, as it is an inexpensive and versatile method. However, it is labour-intensive, tedious, and time-consuming. On the other hand, most automated cell counting methods are expensive and require experts to operate. Thus, the use of image analysis software allows one to access low-cost but robust automated cell counting. This study explores the advanced setting of image processing software to obtain routes with the highest counting accuracy. The results show the effectiveness of advanced settings in CellProfiler for counting cells from synthetic images. Two routes were found to give the highest performance, with average image and cell accuracies of 85% and 99.8%, respectively, and the highest F1 score of 0.83. However, the two routes were unable to correctly determine the exact number of cells on the histology images, albeit giving a respectable cell accuracy of 79.6%. Further investigation has shown that CellProfiler is able to correctly identify the bulk of the cells within the histology images. Good image quality with high focus and less blur was identified as the key to successful image-based cell counting. To further enhance the accuracy, other modules can be included to further segment an object hence improving the number of objects identified. Future work can focus on evaluating the robustness of the routes by comparing them with other methods and validating with the manual cell counting method.

**KEYWORDS** Image Processing; Automated Cell Counting; CellProfiler; Synthetic Cell Images; Histology Cell Images.

## I. INTRODUCTION

Cell counting is a method of quantifying cells to monitor cell proliferation and viability, enhance and optimise the cell culture condition, and prepare for cell-based assays [1, 2]. It is a standard laboratory procedure to investigate cell density and confluency and has been used for various purposes, including monitoring cell proliferation rates and seeding cells for downstream investigations [3]. Furthermore, it is also commonly employed in medical diagnostics and life sciences research [4].

The commonly used method of counting cells is via manual counting using Hemocytometer. This method is very versatile and inexpensive as it only uses essential tools commonly available in the laboratory; however, it is very time-consuming, error-prone, and labour-intensive, especially when dealing with medium to high throughput cell counting [5, 6]. Additionally, human interpretation is required in manual cell counting, which makes the method heavily dependent on the operator's expertise. This potentially leads to variability in results between different operators and even by the same operator but at different times, thereby reducing the reliability of the results. Subsequently, different automated methods of counting cells have been developed to surmount

the existing challenges of manual cell counting.

Many researchers have attempted several methods to automate the cell counting process, including using Hough Transformation [7], gray thresholding [8, 9], artificial neural networks via classification of cell shapes [10], automatic segmentation [11, 12], and mathematical morphology [13], [14]. While these methods are more time-efficient and able to reduce errors, they are not easy to follow, expensive and mostly depend on the operator's capability to understand the algorithms behind the automated programs.

Therefore, the aim of the present study is to explore a low-cost method of automating the cell counting process by optimising the freely available image analyst software for image-based cell counting. It is envisaged that the proposed method of automated cell counting from the present work not only eliminates the current problems associated with manual cell counting but also makes automated cell counting conveniently available to be used by operators. At the initial stage, a preliminary screening of the freely available image processing software of ImageJ and Cell Profiler was conducted. Advanced settings of the selected software were then explored to construct image-processing methods for cell counting. At this stage, 64 routes were identified, and all were

tested for analysing 500 images. The performance of each route was then evaluated using the average cell and image accuracy to find the best routes for image analysis. The significance of the research is that users may use the proposed method to automate their cell counting tasks and utilise these best routes as the basis number of cells in an image. This has the effect of reducing the burden on the users, on what is otherwise a very troublesome and error-prone manual cell-counting task. Finally, the attributes of the selected routes were tested on the blurry-image effect and actual microscopy images.

## II. METHODOLOGY

### A. IMAGE DATASET

Two different image sets; BBBC005V1 from the Broad Bioimage Benchmark Collection [15] and breast cells from a cohort of Triple Negative Breast Cancer (TNBC) patients [16], were used in this paper. A summary of the dataset is given in Table 1, in Fig. 1(a) and Fig. 1(b) depicting sample images from the former and later datasets, respectively. The red box in Fig. 1(a) illustrates one possible situation whereby two cells are considered as one cell, due to clumping together of multiple cells, resulting in overall counting error.

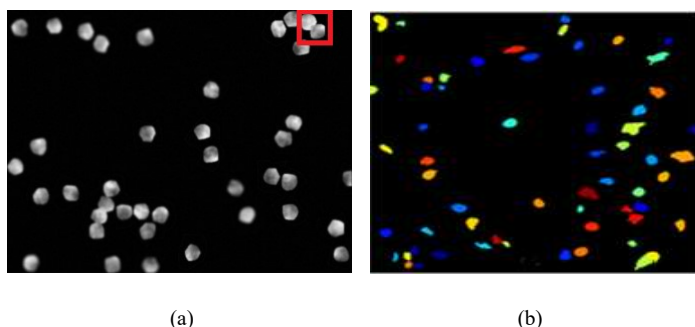


Figure 1. Sample images from (a) the BBCC005v1 dataset [15], and (b) breast cells from a cohort of Triple Negative Breast Cancer (TNBC) patients [16].

The BBBC005V1 simulated high-content screen image of Human Osteosarcoma Cells (U2OS) for a given cell count and was created by the SIMCEP simulating platform for fluorescent cell population images. They were created using a 25% clustering probability and a Charge Coupled Device (CCD) noise variance of 0.0001. Each image has a 696 x 520 pixels resolution. This dataset constitutes the Synthetic HSC image dataset and consists of 500 images. Each image in the dataset contains between 18 and 100 cells.

The image set also includes focus blur, which has been created by applying a Gaussian filter, as well as 16 focus ranges (1-48). However, only five focus ranges of the 500 images were used in this paper to give 2500 images. The image set is provided with labelled nuclei with a given cell count in binary 8-bit TIF images (black and white indicating background and foreground, respectively).

The second image set includes annotated Hematoxylin & Eosin (H&E)-stained histology images at a magnification of 40, containing breast cell samples from 11 TNBC cohort patients. Images in the set are in PNG format and have a resolution of 512 x 512. Furthermore, the image set is also

provided with annotated nuclei, although without the cell counts. The annotated cell nuclei are normal epithelium and myoepithelial breast cells, invasive carcinomatous cells, fibroblasts, endothelial cells, adipocytes, macrophages, and inflammatory cells. To create a cell counting ground truth based on the annotated nuclei, the IdentifyPrimaryObjects of the CellProfiler software was employed. The image set contained 50 images, but only five were utilised in this paper.

Table 1. Summary of image dataset used

Dataset	No. of images	No. of cells in images	Remarks
Synthetic HSC images	500	Between 18 and 100 cells	BBBC005V1 from the Broad Bioimage Benchmark Collection [15]
	2500		BBBC005V1 from the Broad Bioimage Benchmark Collection [15], with 5 focus ranges: F1, F14, F23, F35 and F48
Histology images	5	Between 20 and 103 cells	breast cells from a cohort of Triple Negative Breast Cancer (TNBC) patients [16]

### B. PRELIMINARY ASSESSMENT OF IMAGE SOFTWARE

Initially, two image processing software were considered: CellProfiler and ImageJ. Both CellProfiler [17] and ImageJ [18] software have been used to process several types of microscopy images, including brightfield, histology, fluorescence, differential interference contrast (DIC), and phase imaging. In this paper, they were explored and tested using the same image set: the Synthetic Hematopoietic Stem Cell (HSC) image dataset, to determine the suitability of software in automating the cell counting process. Particularly, the following selection criteria were considered: (i) the accuracy of the results using the software's default setting, and (ii) ease of use of the software. After software selection, the selected software was then used for the counting of cells under different settings.

### C. IDENTIFYING AND COUNTING CELL NUCLEI USING THE CELLPROFILER SOFTWARE

Assessments of the two pieces of software show that the CellProfiler software is more suitable software according to the selection criteria, and as such, it has been chosen as the platform to be used for automated cell counting. A CellProfiler pipeline has been developed, containing fully configurable modules that need to be performed sequentially to fulfil the counting task. The crucial module responsible for identifying and counting objects is the IdentifyPrimaryObjects module. Table 2 gives the list of modules used for the task and their functions, categorised according to their overall usage.

Table 2. The descriptions of CellProfiler modules

Category	Module	Function
Image Acquisition	UnmixColors	Create separated grayscale images from a colour image stained such as Histology. The modules sorted the images based on two or more stains in the background.
	ImageMath	Performs simple mathematics calculations on image intensities.
Segmentation	GaussianFilter	It helps to blur the images to eliminate noise. This module is tremendously

		beneficial to images with noisy foreground signals.
	Smooth	Using this module, one may blur or smooth an image to remove minor artefacts.
	Threshold	Convert the grayscale images into binary using pre-selected threshold methods.
	MaskImage	Masks an image to highlight the identified objects while hiding other areas of the image. This module helps in the pipeline's downstream process.
Object Identification and Counting	IdentifyPrimary Objects	Recognize the biological objects of interest. This module is used in this project to recognize nuclei and count the number of identified nuclei in an image.
Data Export	ExportToSpread sheet	It helps to collect the data from the preceding module and export it to usable files such as Excel.

methods were also explored. Common cells are brighter in the middle and dimmer towards the edge. Consequently, smoothed intensity image obtained via the intensity declumping option can be used to identify local maxima to provide a basis for declumping the foreground into multiple cells. For round-shaped cells, the shape declumping option can be used, whilst none declumping option is suitable for well-dispersed cells image, where no declumping is performed.

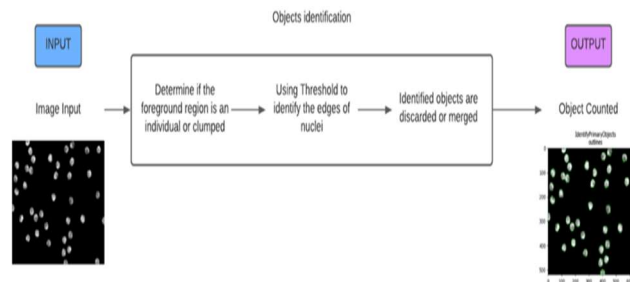


Figure 2. The workflow of the automated cell counting method for the synthetic HSC image dataset (Adapted from [17])

The choice of modules to use is partly dependent on the utilized image set. Two methods were created: 1) the Automated Cell-Counting method for the Synthetic HSC Image dataset, containing the IdentifyPrimaryObjects module, and 2) the Automated Cell-Counting method for the Histology Image dataset, containing other pre-processing modules (Image Acquisition and Segmentation categories) as well as the IdentifyPrimaryObjects module. The first method was developed for the Synthetic HSC Image Set. This is subsequently validated using the Histology Image Set, which represents the actual image analysis in cell counting.

### 6.1 AUTOMATED CELL COUNTING METHOD FOR THE SYNTHETIC HSC IMAGE DATASET

The IdentifyPrimaryObjects module in the CellProfiler software is able to identify and count objects in an image. Fig. 2 depicts the process of this module. In order for the module to function, the input image must be a grayscale image and contains bright foreground cells on a dark background. Similar to other modules in CellProfiler, the IdentifyPrimaryObjects module has configurable options, allowing users to customize the settings according to their needs. To identify objects, the module follows a three-step method: 1) determining if the foreground region is an individual or clumped, 2) using a threshold to identify the edges of nuclei, and 3) making decisions on discarding or merging identified objects. Segmentation of the object or nuclei utilizes a threshold method [19] to recognize the edges of nuclei. This recognition of the edges of nuclei assists in differentiating between the foreground and background, in addition to using the intensity of the objects for differentiation. The success of cell identification and consequently counting is dependent on the contrast of the foreground cells with respect to the background, shapes, and intensity of the cells. Obviously, the identification problem becomes more difficult for images containing multiple cells closely clumped together with irregular shapes and random intensity. Nevertheless, the IdentifyPrimaryObjects module provides multiple options for each of the three-step methods.

Threshold strategy and methods, as well as the declumping (intensity, shape, or none) and dividing lines of clustered objects (intensity, shape, propagate or none)

For cells that do not touch one another, different threshold methods can be used to determine the edges of the cells. In this regard, the module provides various options for threshold strategy (global and local) and methods (minimum cross-entropy, otsu, robust background and sauvola), with default and advanced settings. On the other hand, for images with cells that are not touching one another, either intensity, shape, propagate dividing line options or none at all, can be used. Where the dividing lines between clumped cells are dimmer than the remainder of the foreground, the intensity dividing line option has been shown to work well [20]. With the shape dividing line options, the dividing lines between clumped cells are made based on the shape of the clump, whilst a propagation algorithm is used to draw dividing lines with the propagate dividing line options. Similar to the declumping options, none dividing line options may be suitable for well-dispersed cell images, in which case no dividing line is drawn.

Table 3 illustrates 64 possible combinations/routes of threshold strategy and methods, including the declumping and dividing the line of clustered objects methods.

**Table 3. Overview of 64 routes combination for cell counting derived from the advanced setting (with yellow highlights routes representing some of the best performing routes)**

Route	Threshold		Declumping clustered objects Method	Dividing lines of clustered objects method
	Strategy	Method		
1	Default			
2	Global	Minimum-Cross Entropy	Intensity	Intensity
3				Shape
4				Propagate
5				None
6			Shape	Intensity
7				Shape
8				Propagate

9	Adaptive		None	None	
10			None	N/A	
11			Otsu	Intensity	Intensity
12					Shape
13					Propagate
14				None	
15		Shape		Intensity	
16				Shape	
17			Propagate		
18		None			
19		None	N/A		
20		Robust Background	Intensity	Intensity	
21				Shape	
22				Propagate	
23			None		
24			Shape	Intensity	
25				Shape	
26				Propagate	
27				None	
28			None	N/A	
29			Minimum Cross Entropy	Intensity	Intensity
30		Shape			
31		Propagate			
32		None			
33		Shape		Intensity	
34				Shape	
35				Propagate	
36		None			
37		None		N/A	
38		Otsu		Intensity	Intensity
39			Shape		
40			Propagate		
41			None		
42			Shape	Intensity	
43				Shape	
44				Propagate	
45			None		
46		None	N/A		
47		Robust Background	Intensity	Intensity	
48				Shape	
49				Propagate	
50				None	
51			Shape	Intensity	
52				Shape	
53				Propagate	
54				None	
55		None	N/A		
56		Sauvola	Intensity	Intensity	
57				Shape	
58				Propagate	
59				None	
60			Shape	Intensity	
61				Shape	
62				Propagate	
63				None	
64		None	N/A		

### C.II INFLUENCE OF FOCUS

The occurrence of blur images in a microscope is known as an aberration. This phenomenon occurs when rays fail to converge at one point due to a flaw in a lens or mirror. The influence of blurriness on automated cell counting was also investigated by using different focus ranges. 16 focus ranges between 1 and 48 are available and can be used on an image, with 5 focus ranges: F1, F14, F23, F35 and F48, considered in this paper. These focus ranges represent the amount of focus blur applied to the images. F48 is the most blurriness that can be applied to an image, whilst focus range 1 has no blurriness. Only two of the best routes were tested using five different focus ranges on the 500 images from the Synthetic HSC dataset.

### C.III AUTOMATED CELL COUNTING OF HISTOLOGY IMAGES

The CellProfiler software can also be used to automatically count the number of cells in images extracted from the microscope, including brightfield, fluorescence and histology images. However, some pre-processing needs to be performed on the images before the IdentifyPrimaryObjects module can be used on the images. Fig. 3 depicts the pre-processing required to automate cell counting of raw microscopy images. Image acquisition pre-processing, using the UnmixColors and ImageMath modules, convert the raw microscopy images into grayscale images. Naturally, conversion of the raw microscopy images into grayscale images will result in the loss of information in terms of actual colours [20, 21], however, the IdentifyPrimaryObject module requires its input to be in grayscale. This suggests that the model utilizes only geometrical information and intensities in identifying cells. The image segmentation pre-processing, using the Gaussianfilter, Smooth, Threshold and MaskImage modules, further segment the grayscale images into bright foreground cells on a dark background. Pre-processed images can then be used directly for Object Identification using the IdentifyPrimaryObjects module.

5 histology images from the Histology image dataset were used with the best route used to automatically determine the number of cells in the images. This was performed to explore the applicability of the CellProfiler software to automatically count the number of actual cells from microscopy images, particularly histology images.

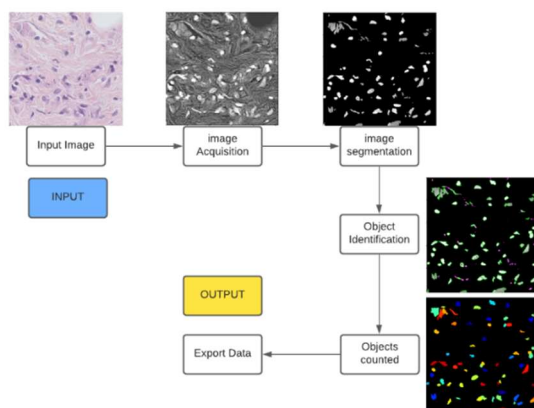


Figure 3. The workflow of the automated cell counting method for the histology image dataset

It is one of the aims of this work to find the best-suited route, i.e., combinations of threshold strategy and method, declumping and dividing lines methods, which is able to accurately count the cell nuclei. Other variables were set to default while exploring the different routes. Each of the 64 routes is used to automatically count the number of cells in the 500 mages from the Synthetic HSC image dataset. This method specifically considered the counting of nuclei only while disregarding other parameters, such as the identified object size.

#### D. PERFORMANCE MEASURES

Different modules in the Cell Profiler software are used to automatically determine the number of cells in an input image. Due to the limitation of the software, the number of counted cells using the software may not correspond to the actual number of cells in an image. A confusion matrix can be constructed, with  $f_{i,j}$  representing the frequency of images with  $i$  actual number of cells but determined as  $j$  number of counted cells using the software.  $\sum_{\forall i,j} f_{i,j}$  then gives the total number of input images  $N$  to the software, i.e.,  $\sum_{\forall i,j} f_{i,j} = N$ .

Average cell accuracy  $Acc_{cell}$  quantifies the accuracy of the software in determining the number of cells in the input image, with high average cell accuracy  $Acc_{cell}$  indicating less deviation from the actual number of cells. This can be calculated from

$$Acc_{cell} = \left[ \frac{\sum_{\forall i,j} (1 - \frac{|i-j|}{i}) \times f_{i,j}}{N} \right] \times 100. \quad (1)$$

Overall image accuracy  $Acc_{image}$  can be defined as the percentage of images whose number of cells within the images have been correctly determined by the software in proportion to the total number of input images. Whilst average cell accuracy  $Acc_{cell}$  considers the size of variations from the actual number of cells, image accuracy  $Acc_{image}$  only considers whether a correct number of cells has been determined. Image accuracy  $Acc_{image}$  is given by

$$Acc_{image} = \frac{\sum_{\forall i} f_{i,i}}{N} \times 100. \quad (2)$$

Image precision can be defined as the average number of images which have been correctly determined as having  $i$  number of cells in proportion to the total number of images which have been determined as having  $i$  number of cells. On the other hand, image recall can be defined as an average number of images which have been correctly determined as having  $i$  number of cells in proportion to the total number of images with  $i$  actual number of cells. Image precision and recall may be determined using

$$Pr_{image} = \frac{\sum_{\forall i} f_{i,i}}{\sum_{\forall j} f_{j,i}}. \quad (3)$$

$$Rec_{image} = \frac{\sum_{\forall i} f_{i,i}}{\sum_{\forall j} f_{i,j}}. \quad (4)$$

Commonly, a single metric, the F1 score, which combines image precision and recall [23, 24], is used as a metric. F1 score is defined as the harmonic mean between precision and recall and may be determined using

$$F1 - Score = \frac{2 \times Pr_{image} \times Rec_{image}}{Pr_{image} + Rec_{image}}. \quad (5)$$

### III. RESULTS AND DISCUSSION

#### A. SOFTWARE PLATFORM PRELIMINARY ASSESSMENT

Two pieces of software: ImageJ and CellProfiler, were identified as possible platforms to develop the image processing method for cell counting applications. Both

software can count cells using their default settings. Nonetheless, based on initial exploration and trials, CellProfiler outperforms ImageJ in terms of user experience.

Although ImageJ is a powerful image analysis tool capable of delivering the desired results in image analysis, it is not user-friendly, particularly for first-time users. It requires the user to seek help from other sources to operate the software. Additionally, it does not give any standards for each setting, making it more challenging to generate good results. Consequently, more time is required to explore the software application. Even so, ImageJ employs a more flexible approach to its settings, offering experienced users more excellent options.

On the other hand, CellProfiler guides the user throughout the interface. There are 'help' buttons to assist new users in using the different settings, and these enable users to select suitable options based on the CellProfiler's recommendations. The modules also offer user-friendly settings that are mostly self-explanatory. Moreover, the CellProfiler's pipeline provides an overview of the modules, making it easy to track and make any necessary adjustments. As a result of these assessments, CellProfiler has been selected as the most suitable platform for automated cell counting in this paper.

#### B. CELLPROFILER SOFTWARE ON THE SYNTHETIC HSC IMAGE DATASET

Cell nucleus identification and segmentation on images from the HSC image dataset uses only one module: the IdentifyPrimaryObjects module on the Cell Profiler software. 500 images containing between 18 and 100 cells in a single image, were used to ascertain the best route from the 64 possible routes. Sample of input HSC image, cell identification results and the output of HSC images are shown in Fig. 4.

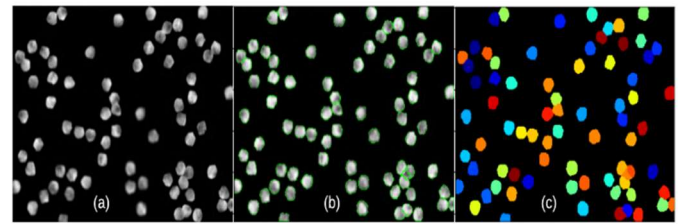


Figure 4. (a) Input of HSC image, (b) Cell identification results, and (c) Output of HSC image

Fig. 5 shows the accuracy of the 64 possible routes: in terms of average image (blue dots) and cell accuracy (orange dots). They were a total of 52 routes with average image accuracies of below 20%, i.e., they were able to correctly determine the number of cells in less than 100 images from the total 500 images. 22 of the routes were unable to correctly determine the number of cells in any of the images. The highest average image accuracy was 85% on routes 34 and 35. Routes 61 and 62 gave a slightly lower accuracy of 84%, whilst routes 7 and 8 gave an accuracy of 81%. All 6 of the best routes in terms of average image accuracy employ 'Shape' and 'Propagate' as the method to divide lines of clustered objects. Similar observations take place in terms of average cell accuracy, with routes 34 and 35 giving the highest average cell accuracy of 99.8%. Routes 7, 8, 61 and

62 gave a slightly lower accuracy of 99.7%. As such, routes 34 and 35 represent the two best routes, with average image and cell accuracies of 85% and 99.8%, respectively; they were able to correctly predict 425 images from the total of 500 images and failed to consider only 0.2% of cells from the 500 images only.

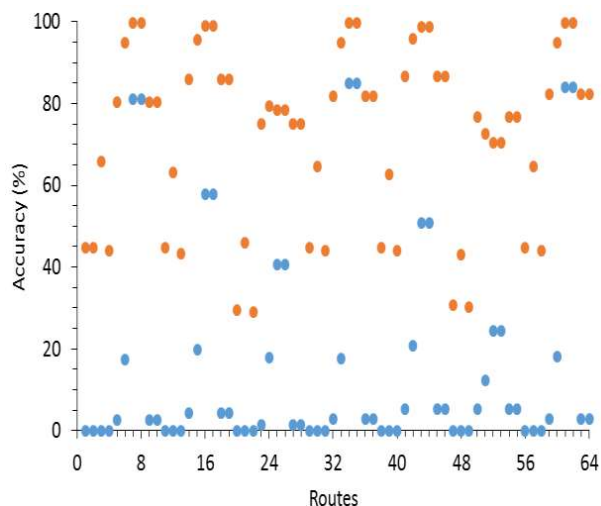


Figure 5. Accuracies of the 64 possible routes for automated cell counting

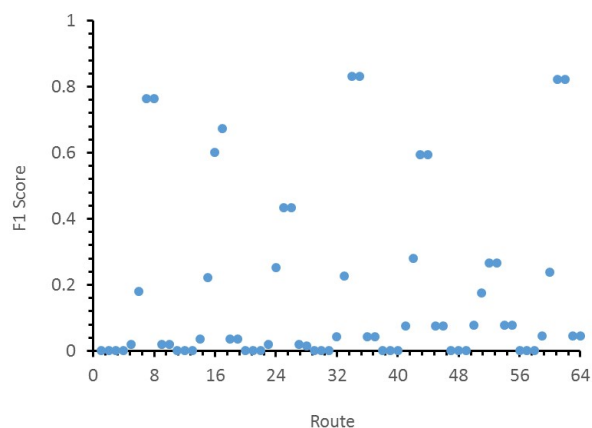


Figure 6. F1 Score of the 64 possible routes for cell counting method on the synthetic HSC dataset

F1 scores for the different routes were calculated and shown in Fig. 6. It can be seen that most of the routes (44 routes) gave an F1 score of less than 0.1. There were four routes with F1 scores greater than 0.8 routes 34, 35, 61 and 62. Routes 61 and 62 gave F1 scores of 0.82, whilst the highest F1 scores of 0.83 were obtained from routes 34 and 35. Despite giving F1 scores which were lower than 0.8, routes 7 and 8 demonstrated respectable F1 scores of 0.76.

It is evident from the above that routes 7, 8, 34, 35, 61 and 62 are some of the best routes that can be used to determine the number of cells in images in terms of average cell accuracy, image accuracy and F1 scores. Routes 7, 34 and 61 were further analyzed in order to distinguish their ability to determine the number of cells in images with a small and large number of cells. Images with 18~57 nuclei counts are considered to have a small number of cells, whilst images with 61~100 nuclei counts are considered to have a large

number of cells. It was found that routes 7, 34 and 61 had small/large nuclei average image accuracies of 92.42%/69.6%, 94.8%/75.2% and 94.0%/74.0%, respectively. These are approximately in the ratio 4:3. Their small/large nuclei counting F1 Score segregation of 0.474/0.289, 0.484/0.347, and 0.482/0.341, respectively. From these results, it can be derived that the software is more able to accurately ascertain the number of cells in images with a small number of cells. Images with a large number of cells may contain closely packed and tightly arranged nuclei, which makes the identification more difficult.

### B.I INFLUENCE OF FOCUS RANGE

The influence of focus range on the ability of the CellProfiler software to determine the number of cells in images was investigated. Only two of the best routes were considered for this purpose, routes 34 and 61. Fig. 7 depicts image accuracies obtained on different focus ranges, with visual presentations of focus range F1, F14, F23, F35 and F48 shown on the insets. It can be seen that as the amount of focus blur applied increases, the accuracy of the routes decreases significantly. This trend illustrates that image with low quality (one of the factors is blurriness) would tremendously affect the accuracy of the image analysis, thus directly influencing the accuracy of the cell counting.

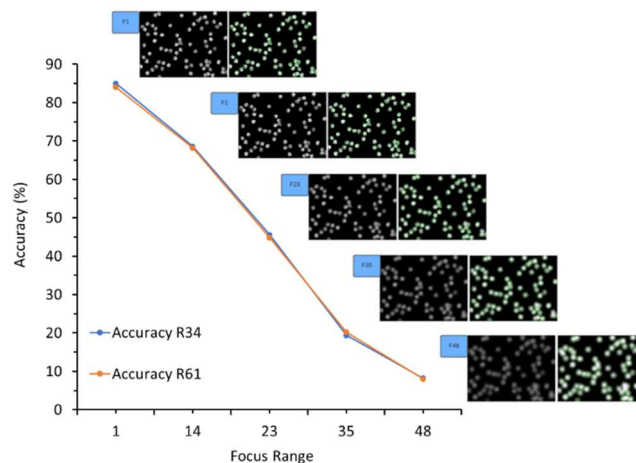


Figure 7. The accuracy of selected routes when blurriness is applied to the images, with insets showing five focus ranges before and after applied to the algorithm under different blurry levels

### C. HISTOLOGY IMAGE SET

The CellProfiler software was also tested by utilizing 5 histology images, with histology images 1, 2, 3, 4 and 5 containing 66, 103, 29, 20 and 68 cells, respectively. Using the best route, i.e., route 34, the number of cells given by the software were 61, 101, 44, 13 and 72, respectively, giving 0% image accuracy but average cell accuracy of 79.6%. The fact that the software was unable to accurately determine the number of cells in any of the histology images is worrying; however, the fact that the deviation from the actual number of cells is small, as given by the almost 80% cell accuracy, necessitates further investigation.

Each of the 5 histology images was manually analyzed and shown in Fig. 8. Three types of identification were obtained: a) correctly identified cell (the coloured objects are

outlined), b) unidentified cell (the coloured object without the red outline), and c) incorrectly identified cell (the red outline without the object). This information is tabulated in Table 4. It can be seen that the software is able to correctly identify a bulk of the cells within the images, with 57, 83, 27, 13 and 51 cells correctly identified from the 66, 103, 29, 20 and 68 cells in images 1, 2, 3, 4 and 5, respectively. Some of the cells cannot be identified by the software, with false identification of non-cells as cells.

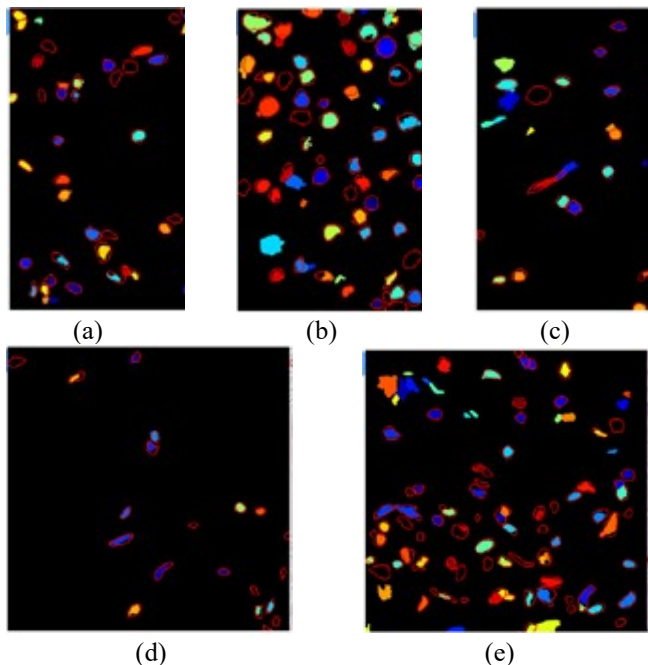


Figure 8. Output from the CellProfiler software for the automated cell counting for images (a) 1, (b) 2, (c) 3, (d) 4, and (e) 5

Table 4. Analysis on the automated cell counting method on the histology images

Image	Actual no. of cell	Correctly identified cell (a)	unidentified cell (b)	incorrectly identified cell (c)
1	66	57	9	4
2	103	83	20	18
3	29	27	2	17
4	20	13	7	0
5	68	51	17	21

**D. DISCUSSION**

Analysis of the results suggests that routes 34 and 35, followed by routes 61 and 62, are able to automate cell counting of synthetic HSC images satisfactorily. Particularly, routes 34 and 35 gave average image and cell accuracies of 85% and 99.8%, respectively, and an F1 score of 0.83. The routes comprise the threshold method, declumping clustered objects and dividing the lines of clustered objects. Cell counting highly relies on thresholding, as also mentioned by others [25–27], with the threshold helping in the conversion of images into binary form, thereby making it easier to distinguish the object’s edges.

All routes 34, 35, 61 and 62 utilized adaptive as their threshold strategy and ‘Shape’ and ‘Propagate’ as methods of declumping the clustered objects and dividing the clustered objects, respectively. As each pixel’s threshold is calculated, the adaptive threshold responds to variations in background

and foreground intensities across the image [17]. This strategy is best suited for the Synthetic HSC Image Set because High Content Screening images are highly contrasted such that the foreground can be detected almost immediately. In image processing, cell counting relies on cell identification and segmentation [28, 29], but this is only partially true for Synthetic Image Sets. The reason is that the HSC images are already ideal for the IdentifyPrimaryObjects to function, perfectly segmenting the objects, as shown in Fig. 4.

Analysis of the ability of the software to determine the number of cells in images with a small and large number of cells was also performed. It has been shown that the software is better able to determine the number of cells in images with a small number of cells. This may be due to the overlapping and congested nature of images with a large number of cells, making it difficult for the software to determine the cell numbers. In order to increase the accuracy of high nuclei counting, the Watershed module may be utilized after the IdentifyPrimaryObjects to segment the objects further. It has also been found that the accuracy of cell counting is highly dependent on the blurriness of the image, with accuracy decreasing with an increase in blurriness. This highlights the importance of getting highly focused images. Indeed, modern microscopes are commonly fitted with auto-focus as standard and there exists different algorithms that may be used to detect blurriness in microscopic images [30]. These algorithms can be combined with the proposed method to give a warning to users of possible errors in cell counting due to the blurring of images. However, this is outside the scope of this current research but provides new direction on this topic.

The applicability of the software was also investigated using some Histology images; however, the result was not that encouraging, as the software was unable to accurately determine the number of cells in any of the histology images. However, in terms of average cell accuracy, it gave 79.6% cell accuracy. Further investigation into the output from the individual histology images shows that the software is able to correctly identify the bulk of the cells within the histology images, albeit some cells were left unidentified and some misidentified. The misidentification may be due to the faint colour of the stain, making it difficult to distinguish the foreground and background. Therefore, the software is unable to pick up the object’s signal intensity.

**IV. CONCLUSION**

Manual cell counting using Hemocytometer, although inexpensive, can be troublesome and prone to error. Current methods of automating the cell-counting are expensive and require expert users. This study explored the use of a low-cost method for automating the cell counting process by using the freely available image analyst software for image-based cell counting. The use of the advanced setting in CellProfiler was proven effective for cell counting applications from input cell images. Of the 64 possible combinations of image processing routes, routes 34 and 35 have been shown to have the highest cell and image accuracies of 85% and 99.8%, respectively. They also resulted in the highest F1 score of 0.83. Analysis of the results has shown that higher accuracy is obtained when processing images with a small number of cells. Routes 34 and 35 utilize adaptive as their threshold strategy and ‘Shape’ and ‘Propagate’ as methods of declumping the clustered objects and dividing the clustered objects, respectively. Cell

counting via image processing was also found to be highly influenced by the focus blur.

As the CellProfiler software is also used on microscopy images, the applicability of the software on histology images was also explored by utilizing additional pre-processing before the images can be fed into the automatic counting module. Unfortunately, the software performs poorly in terms of image accuracy, and it is unable to accurately count the exact number of cells in any of the 5 histology images. Cell accuracy shows a better prospect, with a cell accuracy of 79.6%. These encouraging results prompt us to perform a deeper investigation, whereby it has been demonstrated that the software is able to correctly identify the bulk of the cells within the histology images, albeit some cells were left unidentified and some misidentified.

Good image quality is the key to successful image-based cell counting. To further enhance the accuracy, the Watershed module can be included in the image processing steps, whereby the module could ensure that the objects are further segmented, improving the number of objects identified. Moreover, the robustness of the results can be measured by comparing with other methods and validated with the manual cell counting method and images taken from primary data. The outcome of the present work suggests that the proposed method to automate cell counting can be effectively used by operators.

## ACKNOWLEDGEMENT

This work was supported by UBD/RSCH/1.3/FICBF(b)/2020/012.

## References

- [1] M. C. Phelan and G. Lawler, "Cell counting," *Curr. Protoc. Cytom.*, vol. 00, no. 1, pp. 1–4, 1997. <https://doi.org/10.1002/0471142956.cya03as00>.
- [2] S. Nakada, J. Kohara, and K. Makita, "Estimation of circulating bovine leukemia virus levels using conventional blood cell counts," *J. Dairy Sci.*, vol. 101, no. 12, pp. 11229–11236, 2018. <https://doi.org/10.3168/jds.2018-14609>.
- [3] K. Ongena, C. Das, J. L. Smith, S. Gil, and G. Johnston, "Determining cell number during cell culture using the sceptor cell counter," *J. Vis. Exp.*, no. 45, pp. 1–5, 2010. <https://doi.org/10.3791/2204-v>.
- [4] J. Lojk, L. Šajin, U. Cibej, and M. Pavlin, "Automatic cell counter for cell viability estimation," *Proceedings of the 2014 37th Int. Conv. Inf. Commun. Technol. Electron. Microelectron. MIPRO'2014*, no. May, pp. 239–244, 2014. <https://doi.org/10.1109/MIPRO.2014.6859568>.
- [5] R. Green and S. Wachsmann-Hogiu, "Development, history, and future of automated cell counters," *Clin. Lab. Med.*, vol. 35, no. 1, pp. 1–10, 2015. <https://doi.org/10.1016/j.cll.2014.11.003>.
- [6] J. B. Dixon, M. Faulkner, and J. R. Green, "Electronic counting of dog leucocytes. Discrepancies arising from calibration with Coulter standard 4C and with the haemocytometer," *Res. Vet. Sci.*, vol. 31, no. 2, pp. 249–252, 1981. [https://doi.org/10.1016/S0034-5288\(18\)32503-7](https://doi.org/10.1016/S0034-5288(18)32503-7).
- [7] Z. Xu, "A high accurate static contact angle algorithm based on Hough transformation," *Measurement*, vol. 55, pp. 97–100, 2014. <https://doi.org/10.1016/j.measurement.2014.04.030>.
- [8] P. Patil, "Counting of WBCS and RBCS from blood images using gray thresholding," *Int. J. Res. Eng. Technol.*, vol. 03, pp. 391–395, 2014. <https://doi.org/10.15623/ijret.2014.0304071>.
- [9] L. C. Muskat, Y. Kerkhoff, P. Humbert, T. W. Nattkemper, J. Eilenberg, and A. V. Patel, "Image analysis-based quantification of fungal sporulation by automatic conidia counting and gray value correlation," *MethodsX*, vol. 8, p. 101218, 2021. <https://doi.org/10.1016/j.mex.2021.101218>.
- [10] G. K. Chadha, A. Srivastava, A. Singh, R. Gupta, and D. Singla, "An automated method for counting red blood cells using image processing," *Procedia Comput. Sci.*, vol. 167, no. 2019, pp. 769–778, 2020. <https://doi.org/10.1016/j.procs.2020.03.408>.
- [11] Z. Zhang, Q. Li, W. Song, P. Wei, and J. Guo, "A novel concavity based method for automatic segmentation of touching cells in microfluidic chips," *Expert Syst. Appl.*, vol. 202, p. 117432, 2022. <https://doi.org/10.1016/j.eswa.2022.117432>.
- [12] D. S. Depto *et al.*, "Automatic segmentation of blood cells from microscopic slides: A comparative analysis," *Tissue Cell*, vol. 73, p. 101653, 2021. <https://doi.org/10.1016/j.tice.2021.101653>.
- [13] P. Maji, A. Mandal, M. Ganguly, and S. Saha, "An automated method for counting and characterizing red blood cells using mathematical morphology," *Proceedings of the 2015 Eighth International Conference on Advances in Pattern Recognition (ICAPR)*, 2015, pp. 1–6. <https://doi.org/10.1109/ICAPR.2015.7050674>.
- [14] W. F. Martins, D. A. Longhi, G. M. F. de Aragão, B. Melero, J. Rovira, and A. M. Diez, "A mathematical modeling approach to the quantification of lactic acid bacteria in vacuum-packaged samples of cooked meat: Combining the TaqMan-based quantitative PCR method with the plate-count method," *Int. J. Food Microbiol.*, vol. 318, p. 108466, 2020. <https://doi.org/10.1016/j.ijfoodmicro.2019.108466>.
- [15] V. Ljosa, K. L. Sokolnicki, and A. E. Carpenter, "Annotated high-throughput microscopy image sets for validation," *Nat. Methods*, vol. 9, no. 7, p. 637, 2012. <https://doi.org/10.1038/nmeth.2083>.
- [16] P. Naylor, M. Laé, F. Reyat, and T. Walter, "Segmentation of nuclei in histopathology images by deep regression of the distance map," *IEEE Trans. Med. Imaging*, vol. 38, no. 2, pp. 448–459, 2019. <https://doi.org/10.1109/TMI.2018.2865709>.
- [17] A. E. Carpenter *et al.*, "CellProfiler: Image analysis software for identifying and quantifying cell phenotypes," *Genome Biol.*, vol. 7, no. 10, 2006.
- [18] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," *Nat. Methods* 2012 97, vol. 9, no. 7, pp. 671–675, 2012. <https://doi.org/10.1038/nmeth.2089>.
- [19] L. Sultana and A. G. Sanchis, "Establishing the lower bacterial concentration threshold for different optical counting techniques," *J. Microbiol. Methods*, vol. 203, p. 106620, 2022. <https://doi.org/10.1016/j.mimet.2022.106620>.
- [20] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 6, pp. 583–598, 1991. <https://doi.org/10.1109/34.87344>.
- [21] J. Raihan A, P. E. Abas, and L. C. De Silva, "Depth estimation for underwater images from single view image," *IET Image Process.*, vol. 14, no. 16, pp. 4188–4197, 2020. <https://doi.org/10.1049/iet-ipt.2019.1533>.
- [22] S. Aneja, N. Aneja, E. Abas, and A. G. Naim, "Transfer learning for cancer diagnosis in histopathological images," *LAES Int. J. Artif. Intell.*, vol. 11, no. 1, pp. 129–136, 2022. <https://doi.org/10.11591/ijai.v11.i1.pp129-136>.
- [23] R. Murugaiya, P. E. Abas, K. Mohanchandra, and D. S. Liyanage, "Robust cepstral feature for bird sound classification," *Int. J. Electr. Comput. Eng.*, vol. 12, pp. 1477–1487, 2022. <https://doi.org/10.11591/ijece.v12i2.pp1477-1487>.
- [24] M. G. M. Milani, P. E. Abas, and L. C. De Silva, "Identification of normal and abnormal heart sounds by prominent peak analysis," *Proceedings of the ACM International Conference Preceding Series*, 2019, pp. 31–35. <https://doi.org/10.1145/3364908.3364924>.
- [25] G. Cincotti, L. Nichele, M. Lucidi, and V. Persichetti, "Quantitative evaluation of ImageJ thresholding algorithms for microbial cell counting," *OSA Contin.*, vol. 3, no. 6, pp. 1417–1427, 2020. <https://doi.org/10.1364/OSAC.393971>.
- [26] T. Rasal, T. Veerakumar, B. N. Subudhi, and S. Esakkirajan, "Segmentation and counting of multiple myeloma cells using IEMD based deep neural network," *Leuk. Res.*, vol. 122, p. 106950, 2022. <https://doi.org/10.1016/j.leukres.2022.106950>.
- [27] A. Ossinger *et al.*, "A rapid and accurate method to quantify neurite outgrowth from cell and tissue cultures: Two image analytic approaches using adaptive thresholds or machine learning," *J. Neurosci. Methods*, vol. 331, p. 108522, 2020. <https://doi.org/10.1016/j.jneumeth.2019.108522>.
- [28] S. Wienert *et al.*, "Detection and segmentation of cell nuclei in virtual microscopy images: A minimum-model approach," *Sci. Rep.*, vol. 2, pp. 1–7, 2012. <https://doi.org/10.1038/srep00503>.
- [29] A. LaTorre, L. Alonso-Nanclares, J. M. Peña, and J. DeFelipe, "3D segmentation of neuronal nuclei and cell-type identification using multi-channel information," *Expert Syst. Appl.*, vol. 183, p. 115443, 2021. <https://doi.org/10.1016/j.eswa.2021.115443>.
- [30] H. Zhang and J. Yao, "Automatic Focusing Method of Microscopes Based on Image Processing," *Math. Probl. Eng.*, vol. 2021, p. 8243072, 2021. <https://doi.org/10.1155/2021/8243072>.





**Dewi Kartini Hassan** graduated from Universiti Brunei Darussalam with a BEng in Chemical and Process Engineering (2022), which included a 16-month joint Training Programme at Zhejiang University. After finishing her studies, she is now employed at Hengyi Industries Sdn Bhd. Her research interest focuses on exploring methods to automate cell counting with accessible software using microscopic cell images.



**Hazwani Suhalmi** obtained her BEng in Chemical Engineering (2009), MSc in Advanced Chemical Engineering with IT and Management (2010), and PhD in Chemical Engineering (2015) from Loughborough University. In January 2016, Hazwani joined Universiti Brunei Darussalam as a lecturer in Chemical and Process Engineering, and was promoted to Assistant Professor in December 2021. She has been the Programme Leader of

the Chemical and Process Engineering since August 2016. Her major research is on the nutrient diffusion in tissue engineering membranes and scaffolds. Over the years, she has expanded her research activities and interests including in the area of ankle injury, rehabilitation, sustainability, computer modeling and simulation, and membrane science.



**Muhammad Roll Bilad** is an Associate Professor at the Chemical and Process Engineering Department, Faculty of Integrated Technologies, Universiti Brunei Darussalam. He co-invented 3 EPO patents and co-authored >220 ISI/Scopus journals. Prior to his academic faculty career, he worked as a postdoctoral researcher at KU Leuven (Belgium), Masdar

Institute (United Arab Emirates), and NTU (Singapore). He completed his Ph.D. from KU Leuven (Belgium) on membrane fouling control. His current research focuses on membrane technology for water and wastewater treatment, CO<sub>2</sub> separation, point-of-use treatment, and bioresource management. He is also interested in exploiting "internet-of-thing" via "big data" analysis for membrane plant optimization.



**Pg Emeroylariffion Abas** received his B.Eng. Information Systems Engineering from Imperial College, London in 2001, before obtaining his PhD Communication Systems in 2005 from the same institution. He is now working as a Senior Assistant Professor in System Engineering, Faculty of Integrated Technologies, Universiti Brunei Darussalam. His present research interest are data analysis, IoT and Sensor Network, photonics, and energy system.

...