

A Computer-Aided Diagnosis (CAD) System for Automatic Counting of Ki67 Cells in Meningioma

Fahmi Akmal Dzulkifli^{1*}, Mohd Yusoff Mashor¹, Hasnan Jaafar²

¹Faculty of Electronic Engineering Technology, Universiti Malaysia Perlis (UniMAP), 02600, Arau, Perlis, Malaysia

²Department of Pathology, School of Medical Sciences, Health Campus Universiti Sains Malaysia, 16150, Kubang Kerian, Kelantan, Malaysia

ABSTRACT – Meningioma is a type of primary brain tumour where this tumour arises in the three thin layers of tissues, called meninges. Tumour grading is usually used to describe tumour cells' characteristics and behaviours and how they look under a microscope. There were many techniques used for determining the grade of the tumour. Ki67 was the most common proliferation marker used to measure cell proliferation activity. Currently, pathologists used the manual counting technique to count the Ki67 cells before determining tumour grading. However, this technique was time-consuming, tiring and the counting results are often not accurate. Besides that, manual counting has poor reproducibility and discordant between counting values' among the pathologist. Therefore, this study aimed to develop a Computer-Aided Design (CAD) software that automatically counts the Ki67 cells for determining tumour grading. The purpose of developing this software is to alleviate pathologists' workload associated with counting Ki67 cells and scoring the Ki67 index. The CAD software was developed through seven stages. Based on Pearson Correlation Coefficient results, there was a good positive correlation between the proposed technique with the manual counting technique in counting positive and negative Ki67 cells with a correlation of 0.99 and 0.72 respectively. The proposed CAD system also showed promising results in computing the Ki67 labeling index with a low percentage absolute error of 1.85%.

ARTICLE HISTORY

Received: 10 Feb 2022

Revised: 10 April 2022

Accepted: 9 May 2022

KEYWORDS

Automated Counting

CAD

Ki67

Meningioma

Tumour Grading

INTRODUCTION

A brain tumour can occur at any age. These tumour cells can develop at anybody without considering their ages, gender, and ethnicities. A brain tumour refers to a collection of abnormal cells that typically form a mass of tissue within the brain. In 2017, the American Cancer Society released a report where there were an estimated 16,700 mortalities in the US due to brain and nervous system cancer [1]. Meningioma is a type of primary brain tumour, where the tumour forms in one of the three thin layers of tissue that protects the brain. These tissue layers are called meninges. According to the CBTRUS statistical report, meningioma was the most frequently diagnosed among primary brain tumours from 2011 through 2015, which accounts for 37.1% [2]. The meningioma is usually benign and grows slowly. Due to this fact, these tumour cells are difficult to identify in some locations unless the tumour has an enormous size to be discovered. Thus, this can lead to severely disabling and life-threatening diseases.

Ki67 is a nuclear protein that is present only in growing and dividing phases of the cell (G1, S, G2, and M) but absent during the resting phase (G0) [3]. The Ki67 is usually associated with tumour cell proliferation and growth. Ki67 Labelling Index (LI) is defined as the percentage of immunoreactive tumour among the cell population. The pathologists will use the percentage results to determine the tumour grading. The grading results will help the pathologists determine the type of treatment for a suspected cancer patient. Generally, to obtain the LI value, the pathologists need to count the number of positive and negative Ki67 cells. Currently, the pathologists used manual counting techniques to count and score the Ki67 LI. The 'eye-balling' estimation technique is the most common technique used among pathologists since it is fast, simple, and straightforward. However, this technique produced poor accuracy and reliability in counting the Ki67 cells [4]. Another manual counting technique was by counting each cell individually on the printed image. The cell counting accuracy by using this technique was higher than the 'eye-balling' estimation technique. Nevertheless, this technique was tedious, laborious, and time-consuming [5]. The delays in reviewing the test results can lead a patient to life-threatening situations.

Due to those aforementioned issues, many researchers started developing an automated CAD system for counting and grading tumour disease. In medical imaging, CAD is a system that consists of a few elements such as computer science, pattern recognition, artificial intelligence technologies, and image processing [6]. ImageJ is a well-known open-source software that serves to process and analyse scientific images [7]. Aside from the fact that it is in the public domain, where the source code is openly available and licence-free software, this software also provides various image processing and image manipulation techniques, making it user-friendly [8]. However, this software requires manual adjustments to the

input images for cell counting purposes. It usually occurs when the input images have different image qualities, background intensities, and color intensities between the cells. Therefore, it requires the users to manually adjust the brightness or contrast of the input image, and manually select the thresholding value for cells segmentation and noise removal purposes.

Another preferable software used for cells counting is ImmunoRatio. This software was developed at the University of Tampere, Finland [9]. ImmunoRatio software is commonly used in various studies, especially for computing the Ki67 labelling index. This software is a public domain software, which is free to be used for automated image analysis. This system is designed for the automated evaluation of immunohistochemical (IHC) biomarkers (ER, PR, and Ki67) [10]. However, this software is a web application that requires an Internet connection for the users to access it. Besides that, this web application also had stopped its services starting May 2019. Hence, this study aims to develop a CAD system that automatically counts Ki67 cells and determines the Ki67 labelling index without any adjustments. The main advantage of this CAD system is that it eliminates the subjectivity issue because all image analyses are being done automatically. The following section will discuss the previous works related to the development of CAD systems in the medical field. The next section will present in detail the procedures and methodology to develop the automated CAD system for counting the Ki67 cells and determining the tumour grading. The following section will demonstrate more discussion and the analytical findings and it will be concluded in the last section of this paper.

RELATED WORK

Nowadays, CAD software is more widely used in medical applications. These include medical image analysis, tumour detection, cell counting, designing 3D bioprinting, surgical simulation, and hospital management. In general, the number of studies focusing on CAD systems for medical applications is increasing over the years.

Demir and Camurcu [11] had built a CAD system to detect lung nodules by using outer surface features in computed tomography (CT) images. There were two objectives for developing the CAD system. The first objective was to reduce the time taken by radiologists when examining the CT images. The second objective was to reduce the margin of error when distinguishing the nodules from lung structures. The CAD system consists of three stages: the pre-processing stage, feature extraction, and nodule detection. The pre-processing stage's purpose was to enhance the captured images and identify the region of interest (ROI) or volume of interest (VOI). The next stage was feature extraction. Four features have been extracted from a VOI. These features consist of morphological features, statistical and histogram features, outer surface statistical and histogram features, and outer surface textural features. For nodule detection, the authors used the Support Vector Machine (SVM) to classify the nodules and non-nodules structures. Based on the results, the CAD system was able to identify 597 of 609 nodules correctly. For non-nodules structures, the CAD system was able to detect for 1748 of 1983 correctly. The false positive (FP) result per patient was only 2.45%. Thus, it showed that the outer surface features were helpful for lung nodules detection.

Ştefănescu *et al.* [12] had developed an automatic CAD algorithm for diagnosing colorectal cancer disease. This CAD system comprises three modules. The first module is an algorithm that calculates the fractal dimension and lacunarity for each image. Then, the second module is to compute the gray-level co-occurrence matrix (GLCM) to examine the texture that reflects the spatial relationship of pixels. The third module is focusing on identifying specific anatomical features from the normal colon images. As a result, the proposed CAD system was able to achieve a low diagnosis accuracy error of 15.48%. Aslantas *et al.* [6] developed a CAD system using whole-body bone scintigraphy scans for detecting metastases. The purpose of developing this CAD system was to facilitate physicians' decisions. The proposed system consists of a few steps, which include hot-spot segmentation, feature extraction, and selection and classification. The hot-spot segmentation was done to differentiate the hot-spot area from the skeletal system. A set of active contour segmentation was used to segment the hot-spots area. Ten features have been selected and extracted by using the principal component analysis (PCA) method. Then, the artificial Neural Network (ANN) was used to classify the images for detecting the presence or absence of metastases. Based on 130 images, it showed that the proposed system was able to identify 120 images correctly. The accuracy, sensitivity, and specificity results were also good with 92.30%, 94%, and 86.67%, respectively. Thus, it shows that the CAD system can become a reliable system for detecting metastases.

Karim, Celebi and Mohammed [13] had proposed a CAD system that can determine the type of blood disease in patients. The system was developed using EXSYS CORVID. The development of the CAD system consists of two stages. The first stage was to collect data and medical background of the blood disease. In the second stage, the system will match the blood test results with the medical information before deciding the type of blood disease. The results showed that the system had the potential to be used in the future for blood disease determination. The advantage of this system was that it was easy to use and able to give the results within one minute.

Shaharuddin and Mahmud [14] had designed CAD software for kidney detection based on ultrasound images. This study's objective was to develop CAD software that can detect kidneys from any ultrasound images without considering the image qualities. The proposed system consists of six steps: image acquisition, image pre-processing, image texture analysis, image classification, development of CAD, and performance evaluation of CAD system. In image pre-processing, the colour space of the captured images was converted from RGB to grayscale colour space. Then the images were enhanced using histogram equalization. The speckle noises were removed using the Wiener filter. In texture analysis, five features have been extracted, consisting of coarsening, granulation, regularity, randomness, and smoothness features. Next was the image classification. ANN algorithm was used to classify the images into two categories: a kidney or not. The proposed software had four-button features and two axes for display input and output images. The button features

consist of image analysis, save images, reset and exit function. Four ultrasound machines with different specifications were used in this experiment. Based on 188 ultrasound images, it showed the CAD system able to identify the kidney correctly with a percentage of 94.3%. For non-kidney, the percentage of correctly-identified was 92.2%.

Bahreyni-Tossi, Moghadam, and Nekooie [15] had developed CAD software for the detection and diagnosis of breast lesions in contrast-enhanced magnetic resonance (MR) mammography. The developed software had a few features for image evaluation purposes. The first feature was the colouring. In the colouring process, the system used a specific colour map to highlight all suspected points in the image to provide easy reading. The second feature was the selection of lesion areas. The user can select the suspected region for further evaluation by drawing a free-size ROI around the lesion. The third feature was the margin, shape, and heterogeneity assessments, which are used to display the lesion's margin and internal structures. Next, the CAD software could also read and show the same image cut on other imaging sequences such as TIRM, T2 weighted, and T1 pre-contrast. The final results showed that CAD software's sensitivity for lesion detection was high, with 94.1%, and the specificity was 85%.

Besides the function to diagnose the disease, the CAD system is also developed for cell counting purposes. Li *et al.* [16] had developed an open-source software, called CELLCOUNTER for recognizing and counting the total number of cells in transwell assays. At first, the original assay image was converted into a grayscale image. Then, an adaptive thresholding technique was applied to differentiate between cell areas and background areas. Several image processing procedures such as contrast enhancement, dilation, erosion, and smoothing techniques were applied to remove the noise in the image. Next, the system applied Otsu's method to partition the cell areas into true cells and unwanted objects to improve counting's accuracy. The results showed a high correlation coefficient between the CELLCOUNTER and cell counting results performed by experts with $R^2 = 0.96$.

O'Brien, Hayder and Peng [17] had developed two plugins within ImageJ software for automated hemocytometer and migration/invasion cell counting. These plugins are referring to Cell Concentration Calculator and Migration Assay Counter. For measuring the performance of the proposed plugins, the results obtained from both proposed systems were compared with the manual counts by using the ImageJ plugin Cell Counter. As a result, the correlation between both plugins with the manual counts was high with $R^2 = 0.99$ for the Cell Concentration Calculator and $R^2 = 0.97$ for the migration assay counter.

González-González *et al.* [18] investigated the performance of the ImmunoRatio software by comparing the manual and automated counting results obtained from the software for Ki67 in ameloblastoma. The nuclear area was identified using two methods during the automated counting process. The first method was the basic method, which did not require any adjustments, while the second method was the advanced method, which included a color correction protocol. A blank field image was also used in this study to assess the quality of the brightness and contrast of each ameloblastoma case's image. The results presented showed that the ImmunoRatio software was suitable for ameloblastoma cases since the percentage difference between the conventional labelling indexes and advanced ImmunoRatio was less than 0.5%. However, some adjustments are required to achieve better accuracy results.

Loddo *et al.* [19] proposed a CAD system to automatically detect and identify the red blood cells and white blood cells from blood smear images. At first, the proposed system will segment the input images based on the machine learning approach, which by using the Support Vector Machine (SVM) method. Next, the system will count every single cell in the image. If any clumped cells were identified, the proposed system will apply the Circular Hough transform to locate only the circular objects with the appropriate grey level values. Finally, the system will count the cells in this stage, and the results will be added to the counting results obtained from single cells that were computed previously. Based on the accuracy results acquired by the proposed system, it showed the proposed CAD system was effective in counting the red blood cells and white blood cells with average values of 98% and 99.2% respectively.

Venter and Niesler [20] developed a protocol for identifying and quantifying adherent cell numbers and wound area by using ImageJ software. The cell number was computed and analyzed using three methods. The first two methods were based on ImageJ software, which referred to manual and automated identification. The third method was based on the spectrophotometric assay. For manual identification, the input image was first converted into grayscale. Then, the cells were manually labelled using Microsoft Paint. The next step was to apply a colour thresholding technique for identifying the labelled cells and proceed with the counting process. For automated identification, various image processing techniques were applied such as noise removal, brightness and contrast adjustment, thresholding based on the Phansalkar method, and watershed technique. Another protocol that was developed in this study was automated wound area measurements. Image processing techniques involved in this protocol consisted of colour space conversion, edge detection, and image smoothing. Then, the minimum error thresholding was applied to automatically detect the wound area. The authors concluded that using ImageJ was less laborious and faster compared to manual analysis. The automated analysis using ImageJ was able to accurately define the wound edges more effectively than the manual method.

COMPUTER-AIDED DIAGNOSIS SOFTWARE

This section will explain in detail the design and development of the CAD software built for this study. This CAD software's function was to help the pathologists by providing automated Ki67 counting cells and scoring the Ki67 index. The operation of the CAD software consists of seven stages. Figure 1 shows the block diagram of the CAD software.

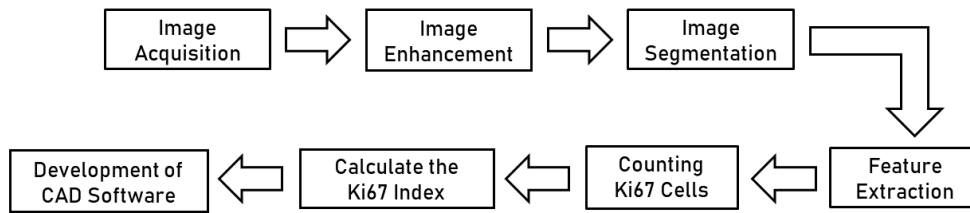


Figure 1. Block Diagram for CAD Software

Image Acquisition

A total of 12 histopathological images were selected and captured at the Department of Pathology, Hospital Universiti Sains Malaysia (HUSM). The meningioma slides were stained using immunohistochemical (IHC) staining. The basic principle of IHC stains is that an immune protein or known as an antibody, will be attached itself to certain substances or called antigens, which is on the cell [21]. Each type of antibody will recognize and attach exactly to the antigen that fits it. In some instances, the normal and cancer cells had unique antigens. A cell with a specific antigen will bind to the antibody that fits the antigen. In order to decide whether the antibodies have been attracted to the cells, chemicals are applied, causing them to change colour when a specific antibody was present [21]. In this study, the sample tissue will be stained with the Diaminobenzidine (DAB) for visualizing positive cells and counterstained with Haematoxylin and Eosin (H&E). As shown in Figure 2, IHC staining makes the positive Ki67 cells appear in granular brown colour, while the negative Ki67 cells appear in blue. The Ki67 images were captured under 40x magnification using an Olympus BX51 microscope and Cell^F software that works as an interface to the digital camera attached to the microscope. The sample images were then saved in (*.jpg) format with a resolution of 1360×1024 pixels and 24-bit RGB. Figure 2 shows an example of the Ki67 image of meningioma using IHC stains.

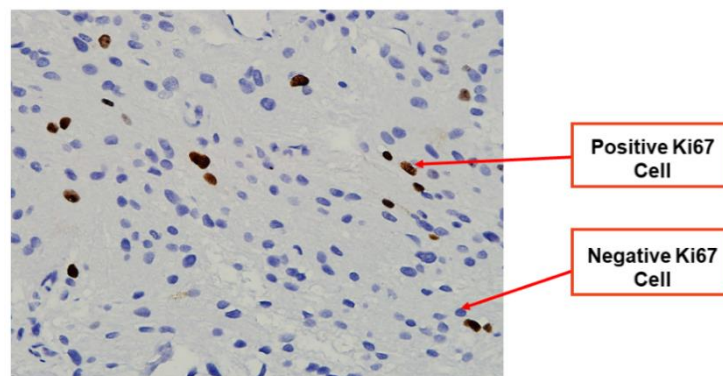


Figure 2. A Ki67 image of meningioma using IHC stains

Image Enhancement

In this study, image enhancement was applied to enhance the contrast and brightness of the captured images. At first, the colour space of the captured images was converted from RGB to $L^*a^*b^*$ colour space. Then, the contrast enhancement was applied to the luminosity, 'L' channel while keeping a^* and b^* channels remain unchanged. For this study, the contrast stretching technique [22] was selected to enhance the images' contrast. After obtaining the output images from the contrast enhancement techniques, the images were converted back to the RGB colour space. The next step was to increase the brightness of an image. For increasing the brightness of an image, a fixed value needs to be selected. The selected value must be greater than 0 to get a brighter image. After performing an analysis using 12 meningioma images, the best value to increase the brightness of the Ki67 image was by adding all the pixels of that image with 20. Based on the observation from the output enhancement images, if the constant value was higher or lower than 20, it will degrade the segmentation results.

Image Segmentation

The next step was image segmentation. In this study, the thresholding method was selected since this method was simple and effective in partitioning the image into the foreground and background. The colour space of the resultant image from the pre-processing image technique was converted into Hue, Saturation, and Intensity (HSI) colour space. Based on the observation, some of the information, especially the negative Ki67 cells in H and S components, were lost, which may degrade the segmentation accuracy. As a result, the intensity component was extracted from the image to perform the segmentation process. The segmentation process was divided into two parts, in which the first part was to segment positive Ki67 cells and followed by the second part to segment negative Ki67 cells.

The process of segmentation was started by removing the image background using Otsu's thresholding. Otsu's technique is a nonparametric and unsupervised technique that automatically selects the threshold region from a grey-level histogram [23]. This method aims to find the threshold value to minimize the weighted within-class variance of the black and white pixels [24]. In this study, the proposed system will remove the background based on the threshold value calculated using Otsu's technique. Then the image was retrieved back to the RGB image. The positive Ki67 cells appeared to be in red-brownish colour. The issue that occurred was determining the RGB values for each channel that described a positive Ki67 cell. For a solution, the proposed system needs to identify the redness in the image first. Equation 1 demonstrates a simple equation to find the redness of a pixel:

$$f(x, y) = (R1) - \max((G1), (B1)) \quad (1)$$

where $f(x, y)$ is the output image after redness thresholding. $R1$, $G1$, and $B1$ indicate the values for each red, green, and blue component from the image that has been through the background removal process. Then, colour thresholding was carried out based on the colour information of the positive Ki67 cell to remove the parts of the image that fell within a specified colour range. The RGB image was separated into three channels. Later, the system will calculate the minimum and maximum values for each colour channel. Colour thresholding for positive Ki67 cell was calculated as:

$$g(x, y) = \{f(x, y) * 0.7, \quad \text{if } f(x, y) == Bmax \ 0, \quad \text{otherwise} \quad (2)$$

where the $g(x, y)$ is the output pixel value from the colour thresholding technique. $f(x, y)$ is the pixel value from the output image after thresholding the redness of a pixel. $Bmax$ represents the maximum values for the blue colour channels after thresholding the redness. The blue component must have a low value since the positive Ki67 cells tend to be in red. Based on the analysis using 12 meningioma images, the most suitable value to segment the positive Ki67 cells was 0.7. This value produces the best segmentation results for positive cells. Next, two morphological operations: region filling and area opening techniques, were applied to the segmented image. The type of region filling technique used in this study is the flood-fill operation, where this technique is used to fill the holes or interior region in the segmented image [25]. The area opening technique is a filter used in the binary image to eliminate the components that have a smaller area than a parameter λ [26]. For this study, the small objects were referred to as all connected components representing noises or irrelevant image objects in the binary image. After performing analysis using 12 meningioma images (which consists of approximately 750 positive Ki67 cells), it has been found that the positive Ki67 cells have an area greater than 130 pixels. Thus any objects smaller than 130 pixels will be eliminated from the image.

The second part of the segmentation process was to segment negative Ki67 cells. The procedure to find the negative Ki67 cells is composed of the following steps:

- Acquire the resultant image after applying the contrast stretching technique.
- Subtract the image in Figure 3(a) with the output image from the colour thresholding technique. If the pixels in Figure 3(a) had the exact pixel values with the image Figure 3(b), the pixels would be replaced by white colour (255). Figure 3 shows the step to obtain the negative K67 image. Figure 3(c) shows that the red arrows pointed to the pixels have been converted into white pixels.

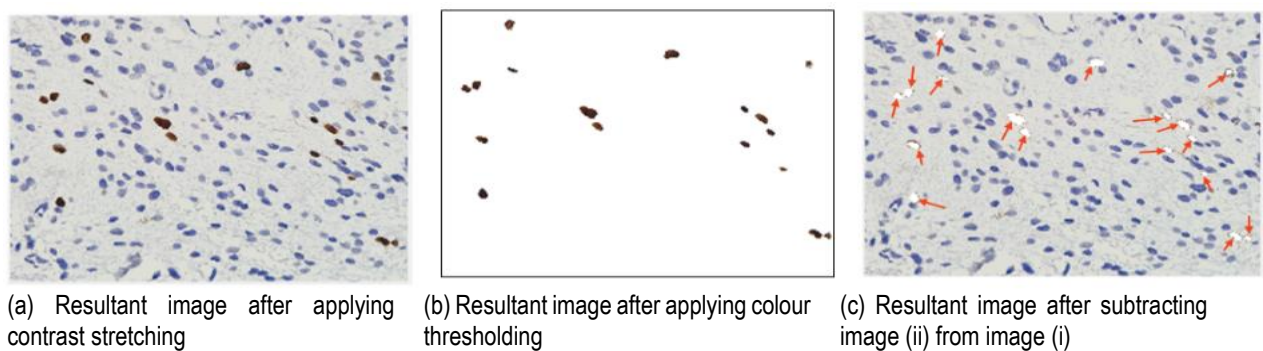


Figure 3. Procedure for acquiring the negative Ki67 image

After obtaining the negative Ki67 image, the following process was to apply the Otsu's thresholding to segment negative Ki67 cells. Similar to segmenting the positive Ki67 cells process, the next step was to apply the region filling and area opening techniques. The proposed system will eliminate the small objects representing noises or irrelevant image objects in the binary image.

Feature Extraction

In image processing, a feature is a piece of information that is usually used to describe an object in an image. Feature extraction is a technique that aims to retrieve the significant information from current data and represent it in a lower

dimensionality space [27]. The selection of a suitable feature needs to be done carefully since the result can seriously affect the system's performance, especially counting the Ki67 cells. The extracted features will provide useful information for identifying the ideal of Ki67 cells and removing unwanted objects like stained artefacts. Three main features have been selected from the Ki67 cell, namely circularity, area, and solidity. Area feature is defined as the number of pixels in a region. Circularity represents the shape of the Ki67 cell. Solidity is used to measure the density of a cell. These three features were used as the input to the proposed system for distinguishing between the Ki67 cells and unwanted objects. The extracted image of positive and negative Ki67 cell was later saved into '*.bmp' file extension. The proposed system will eliminate any objects that do not meet the threshold value of the selected features. Table 1 lists the minimum values for extracting the Ki67 cells based on the selected features.

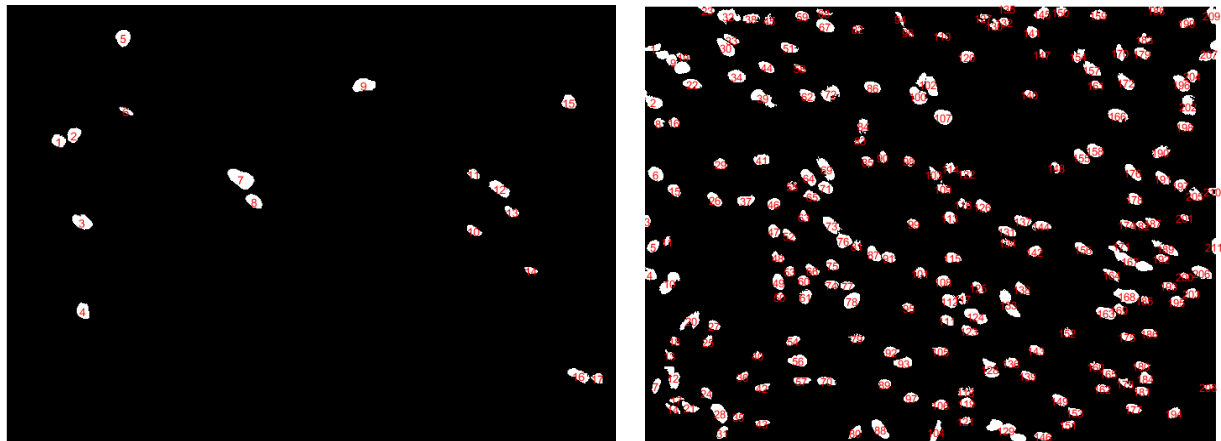
Table 1: Minimum values for extracting circularity, area and solidity features

Ki67 Cells	Circularity	Area	Solidity
Positive Ki67 Cells	1.0	100	0.7
Negative Ki67 Cells	1.0	100	0.7

Table 1 listed the values that would be used to extract the positive and negative Ki67 cells from unwanted objects like noises. These values are the minimum values that identified manually by performing analyses using 12 Ki67 histopathological images (which consists approximately of 750 positive Ki67 cells and 2800 negative Ki67 cells)

Counting Ki67 Cells

The resultant image from the feature extraction process will be used in this step. As shown in Figure 4, the processed image only had two possible values (0 or 1) for each pixel. The '0' represents the background of the image while '1' refers to Ki67 cells. The following step was to count the extracted immunopositive and immunonegative Ki67 cells. The system will label each of the cells with different numbers from 1 to n numbers for counting purposes. Figure 4 illustrates the process of counting the Ki67 cells. After performing the counting process, the processed image was retrieved to the RGB colour space.



(a) Counting Immunopositive Ki67 Cells

(b) Counting Immunonegative Ki67 Cells

Figure 4. The labelling process for counting each of the Ki67 cells

Calculating Ki67 Labelling Index

Equation 3 presents the formula to calculate the Ki67 index. This step can proceed after obtaining the number of immunostained positive and immunostained negative Ki67 cells. According to the 2016 World Health Organisation Classification of Tumours of the Central Nervous System grade, a meningioma is classified into three classes, which are Grade I, Grade II, and Grade III [28]. The tumour was labelled as grade I if the index was 0% to 7.49%, grade II for 7.5% to 19.99%, and grade III for greater than 20%.

$$\begin{aligned}
 & \text{Ki67 Index} \\
 &= \frac{\text{No. of positive Ki67 Cells}}{\text{No. positive Ki67 Cells} + \text{No. of negative Ki67 Cells}} \times 100\% \quad (3)
 \end{aligned}$$

Layout and Functions of the CAD Software

This section will discuss the available features that were used for designing the layout of the CAD software. The reason for using GUIDE to develop the software was because there are many advanced programming tools available for GUI-based operating systems that may help the developer construct a GUI application. GUIDE is a high-level GUI development tool provided in MATLAB, which allows the developer to create their GUI design. The current MATLAB version comes with advanced GUI development and graphics capabilities which includes real-time graphics support, 2D and 3D hardware acceleration using OpenGL, and many of the standard user interface controls common to graphical operating systems [29]. Figure 5 shows the layout of the CAD software.

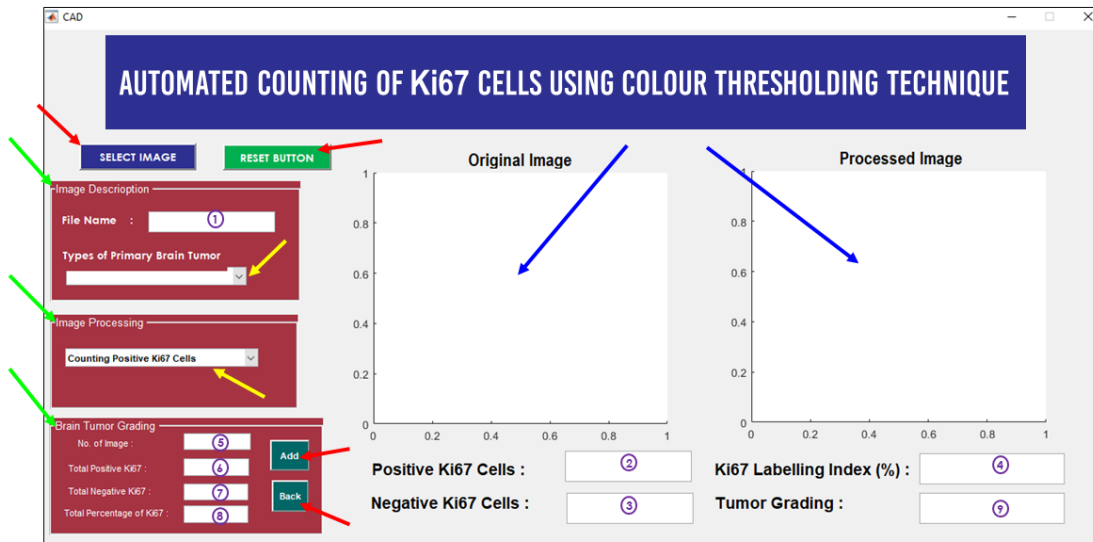


Figure 5. Layout of CAD Software

RESULT AND DISCUSSION

Image Segmentation Results

This section will discuss the performance of the proposed segmentation procedure in segmenting the Ki67 cells from the IHC staining image. The performance of the proposed system was assessed by comparing the segmentation results provided by the proposed system with the manual segmentation. The manual segmentation was performed by using the ImageJ software. At first, the input image was converted into grayscale colour space. Automatic global thresholding was then applied to the images for segmenting the Ki67 cells. The post-processing techniques such as noise removal and region filling were done manually by using the ImageJ software. The resultant segmented image was then validated by the HUSM pathologists. Figure 6 presents the resultant images after applying colour thresholding technique for segmenting positive Ki67 cells.

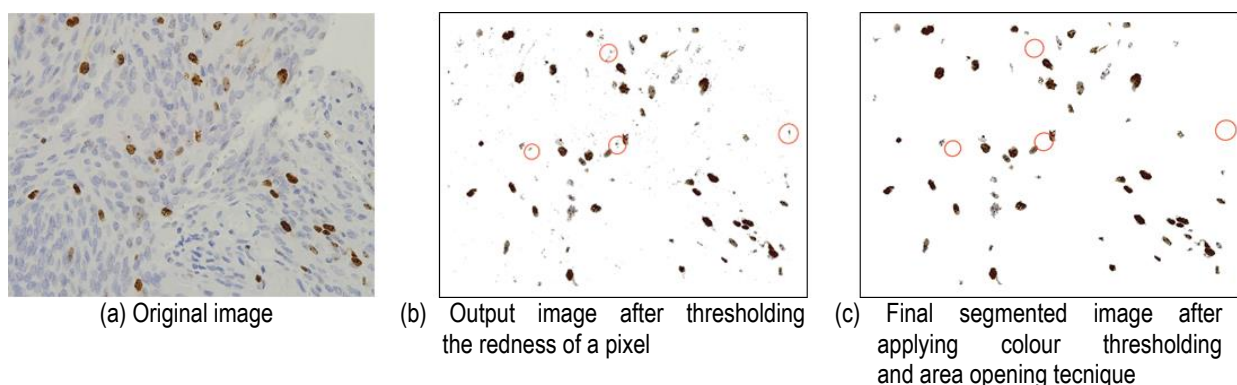


Figure 6. Resultant image for segmenting the positive Ki67 cells

Three analyses have been performed to measure the system performance in segmenting the Ki67 cells. These analyses include calculating the accuracy, sensitivity, and specificity. The accuracy indicates the percentage of how often the system detects and segments the Ki67 cells correctly. The sensitivity demonstrates the percentage of actual pixels that were correctly segmented as positive Ki67 cells. Specificity in this study was referring to the percentage of actual pixels

that were correctly identified as negative Ki67 cells. The accuracy, sensitivity, and specificity can be calculated as expressed in Equations 4 to 6.

$$Accuracy = \left(\frac{TP + TN}{(TP + TN + FP + FN)} \right) \times 100\% \quad (4)$$

$$Sensitivity = \left(\frac{TP}{(TP + FN)} \right) \times 100\% \quad (5)$$

$$Specificity = \left(\frac{TN}{(TN + FP)} \right) \times 100\% \quad (6)$$

where *TP* is true positive, *TN* is true negative, *FP* refers to false positive while *FN* is false negative. In this study, the *TP* defines the number of pixels that are correctly segmented as positive Ki67 cells. *TN* refers to the number of pixels that were correctly segmented as negative Ki67 cells. For *FP*, it specifies the number of pixels of negative Ki67 cells and unwanted objects that were incorrectly segmented as positive Ki67 cells. *FN* signifies the number of pixels of positive Ki67 cells and unwanted objects that were incorrectly segmented as negative Ki67 cells. Table 2 demonstrates the segmentation results for all sample images.

Table 2. Segmentation analysis results in segmenting Ki67 cells using proposed segmentation technique

Input image	Accuracy (%)	Sensitivity (%)	Specificity (%)
Cell 01.bmp	97.90	96.38	99.52
Cell 02.bmp	97.11	95.08	99.33
Cell 03.bmp	98.06	96.91	99.27
Cell 04.bmp	96.27	93.43	99.51
Cell 05.bmp	96.11	93.02	99.68
Cell 06.bmp	96.06	93.10	99.47
Cell 07.bmp	98.44	97.28	99.65
Cell 08.bmp	97.88	96.38	99.47
Cell 09.bmp	98.35	97.24	99.52
Cell 10.bmp	89.96	83.93	98.62
Cell 11.bmp	90.15	84.52	97.99
Cell 12.bmp	83.92	76.56	96.92
Average (%)	95.02	91.99	99.08

Based on the results in Table 2, the proposed system was able to segment the positive Ki67 cells with an average accuracy of 95.02%. The table also showed most of the images have an average accuracy of more than 90%. The average sensitivity and specificity were also high with an average of 91.99% and 99.08% respectively. Therefore it indicates the proposed algorithm had successfully detected most of the positive Ki67 cells in the image.

Results of Feature Extraction for Ki67 Cells

This section showed the results of the meningioma images after extracting the Ki67 features which would be used in the next counting process. As discussed in the Feature Extraction section, three features were used to distinguish the Ki67 cells. These are circularity, area, and solidity features. Each feature had the threshold value which was already listed in Table 1. The proposed system will remove the objects or cells directly if the feature values are lower than the specified range. From the observation, the proposed system was able to extract the true Ki67 cells and eliminated most of the artefacts and noises by using the specified feature values. Figure 7 presents the result after extracting Ki67 cells from a meningioma histopathological image.

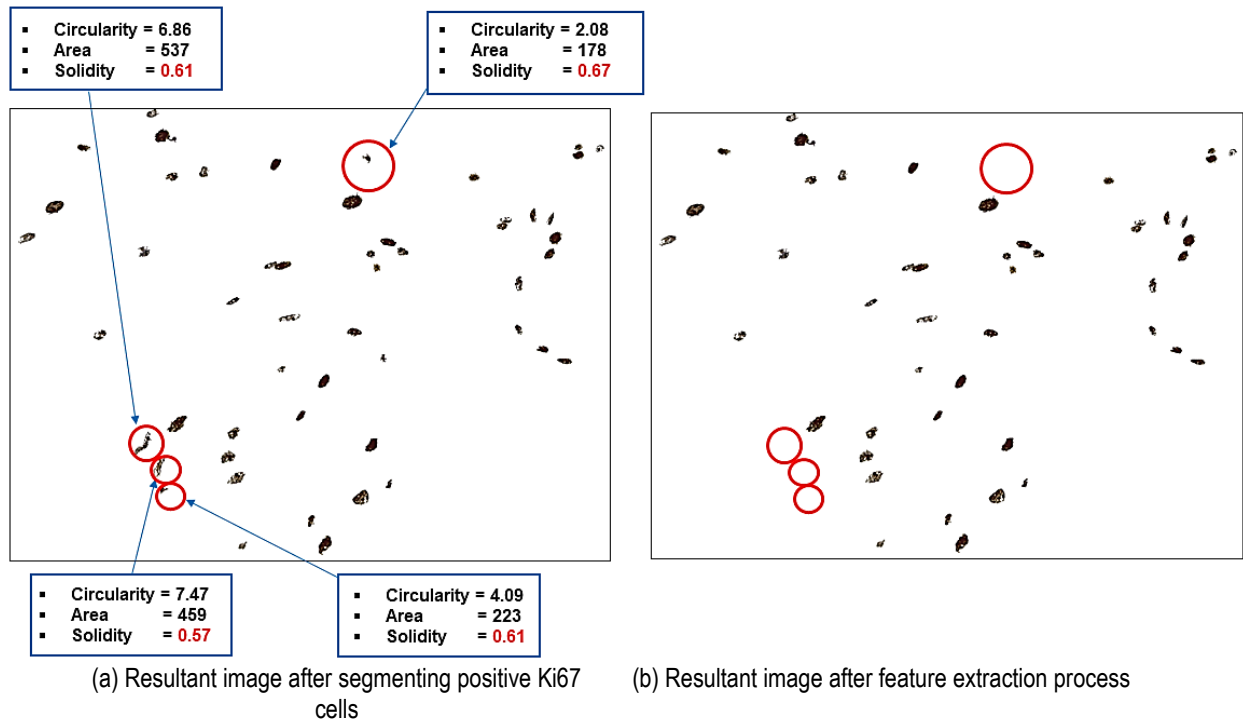


Figure 7. Extracting Ki67 cells based on selected features

Figure 7(b) shows the output image after applying the feature extraction technique. Based on Figure 7(b), the proposed system removed the objects since the solidity values of those objects did not meet the requirements although the circularity and area values exceeded the range.

Comparison between the Automated Counting System and Manual Counting

In this section, the counting results of Ki67 cells for the proposed technique were compared with manual counting. For the manual counting process, the captured meningioma images were printed out using a HP Deskjet 2135 printer. The counting process was done by marking and counting each of the cells at the printed colour image. The HUSM pathologist then validated the cell's marking and the results of the manual counting. The relative error and relative accuracy analysis were calculated to assess the inter-observer agreement between manual counting and proposed automated counting. The purpose of this analysis was to find how close a measured value is to a standard value on relative terms. The relative error can be obtained by using Equation 7 [30]:

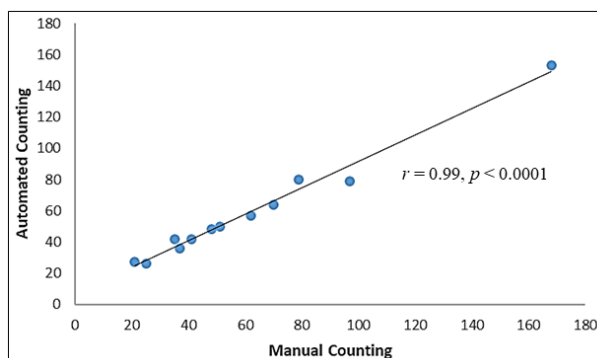
$$\varepsilon_r = \left| \frac{V_A - V_E}{V_E} \right| \quad (7)$$

where ε_r was the relative error. V_A was the result of counting from the proposed system and V_E was the value of manual counting cells. The relative accuracy can be calculated by subtracting the value of 1 with the relative error value (ε_r) [31]. Table 3 presents the analysis result of counting Ki67 cells.

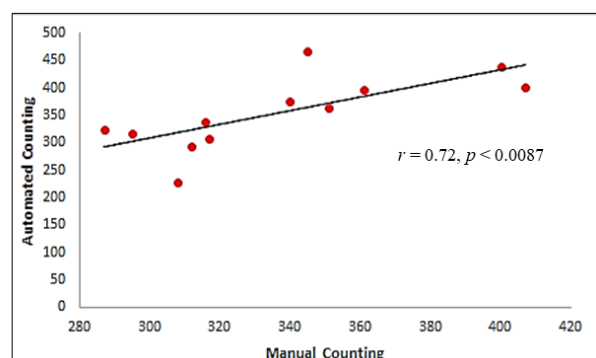
Table 3. Comparison of counting performance for Ki67 cells between manual and automated counting

Image	Manual Counting		Automated Counting		Relative Error		Relative Accuracy	
	Positive Ki67	Negative Ki67	Positive Ki67	Negative Ki67	Positive Ki67	Negative Ki67	Positive Ki67	Negative Ki67
Cell 01.bmp	51	407	50	403	0.02	0.01	0.98	0.99
Cell 02.bmp	62	340	57	375	0.08	0.10	0.92	0.90
Cell 03.bmp	70	351	64	363	0.09	0.03	0.91	0.97
Cell 04.bmp	21	317	27	309	0.29	0.03	0.71	0.97
Cell 05.bmp	25	312	26	293	0.04	0.06	0.96	0.94
Cell 06.bmp	35	400	42	438	0.20	0.10	0.80	0.90
Cell 07.bmp	41	316	42	338	0.02	0.07	0.98	0.93
Cell 08.bmp	37	295	36	318	0.03	0.08	0.97	0.92
Cell 09.bmp	48	287	48	325	0.00	0.13	1.00	0.87
Cell 10.bmp	79	345	80	467	0.01	0.35	0.99	0.65
Cell 11.bmp	97	361	79	396	0.019	0.10	0.81	0.90
Cell 12.bmp	168	308	153	228	0.09	0.26	0.91	0.74
Average					0.09	0.11	0.91	0.89

As shown in Table 3, the proposed system was capable of producing good results in counting positive and negative Ki67 cells. Based on the relative accuracy results, nine of the twelve images obtained an accuracy of more than 0.90 in counting positive and negative Ki67 cells. As shown in Table 3, the proposed system was able to produce good results in counting positive and negative Ki67 cells. Based on the relative accuracy results, nine out of twelve images obtained accuracy with more than 0.90 in counting positive and negative Ki67 cells. The average relative accuracy for counting positive Ki67 cells was 0.91, while 0.89 for counting negative Ki67 cells. The Pearson Correlation Coefficient (PCC) was used to measure how strong the correlation between manual and automated cell counts. The correlation test was applied for the 12 samples. A p -value less than 0.05 is considered statistically significant. Figure 8 shows the correlation of Ki67 cell counting by manual and automated counting.



(a) Correlation result of counting positive Ki67 cells between manual and automated counting based



(b) Correlation result of counting negative Ki67 cells between manual and automated counting based

Figure 8. Correlation of Ki67 cells counting by manual and automated counting

Figure 8(a) showed a positive correlation in counting positive Ki67 cells between manual and automated counting. Most of all data points were lying on the regression line. By using the Pearson Correlation Coefficient (PCC) test, there was a strong correlation between the result of both methods with ($r = 0.99, p < 0.05$). Figure 8(b) showed that there was a good positive correlation between these two methods in counting negative cells with ($r = 0.72, p < 0.05$).

The Operation of the CAD Software

From Figure 5, few design tools were used to develop CAD software. The green arrows showed the 'Panel' box. Two 'Axes' (as shown in the blue arrow) were used to display the original and resultant image. The red arrows present the 'Push Button' features, while the yellow arrows indicate the 'List box' features.

Before executing the system, the user should click the 'RESET BUTTON' to make sure the system will clear all the previous data. At this time, the system will display the '0' value in the box (from number 2 to 8). Then, the user will click the 'SELECT IMAGE' button to choose the image that wants to be analysed. Each software was designed to have three panels, as shown in Figure 5. The first panel was called 'Image Description'. When the user selects the input image, the file's name will be displayed in the 'Edit Box' (number 1). In this panel, there was also a box under the title 'Types of Primary Brain Tumour'. Basically, each type of primary brain tumour had a different range of Ki67 index. For this study, this box only contained one type from the primary brain tumours, which is the meningioma. This box indicated that the Ki67 index and the tumour grading result are based on the classification for meningioma type. The percentage of the Ki67 index and tumour grading for meningioma type have been discussed previously in the *Calculating Ki67 Labelling Index* section. Once the user clicks on this box, the system will display the grading results in the 'Tumour Grading' box (number 9).

The second panel is called Image Processing. This panel is where the process of analysing and counting Ki67 cells will take place. Another 'List box' has been designed in this panel. There were two items in this box, namely 'Counting Positive Ki67 Cells' and 'Counting Negative Ki67 Cells'. When the user clicks for the first item, the system will count the positive Ki67 cells and display the result in the 'Positive Ki67 Cells' box (number 2). The output image for counting the positive Ki67 cells will be displayed at the 'Processed Image'. Figure 9 shows the layout of the software in counting the positive Ki67 cells. At this time, all the boxes will display the '0' value. Then, the user must click at the 'Counting Positive Ki67 Cells', which is in the 'Image Processing' panel for starting the counting process for positive Ki67 cells. The result of counting will display at the 'Positive Ki67 Cells' box.

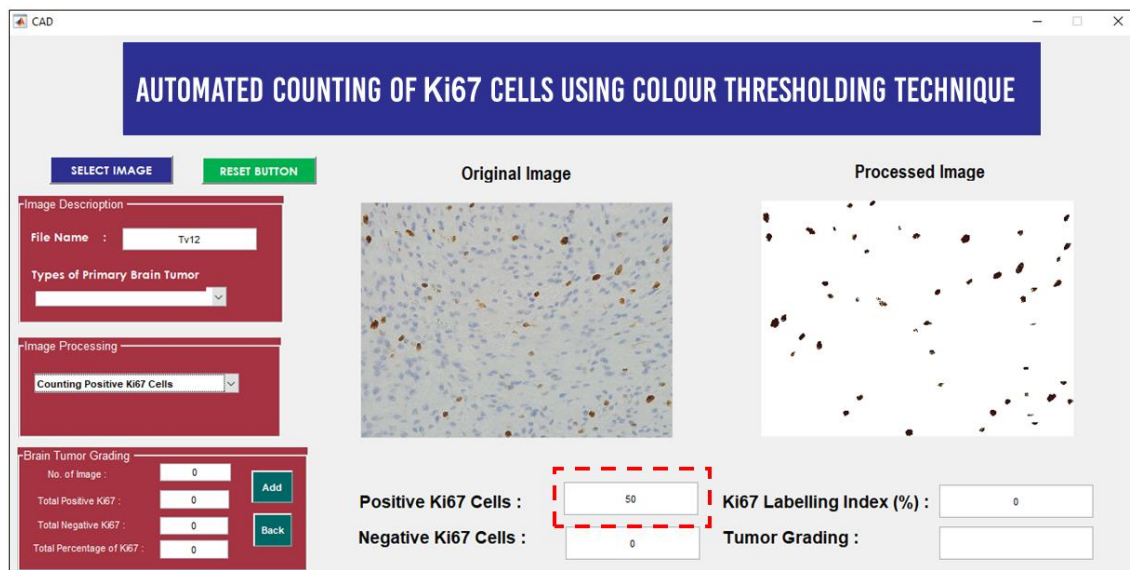


Figure 9. Layout of CAD software after counting positive Ki67 cells

The second item was the process of analysing and counting the negative Ki67 cells. When the user clicks the second item, the system will count the negative cells and display the result at the 'Negative Ki67 Cells' box (number 3). The output image for counting the negative Ki67 cells will be displayed at the 'Processed Image'. At this time, the system will also calculate and display the Ki67 index result at the 'Percentage of Ki67 (%)' box (number 4). Note that this result was the result of the current image being analysed. Figure 10 illustrates the layout of the software in counting negative Ki67 cells and calculating the ki67 index.

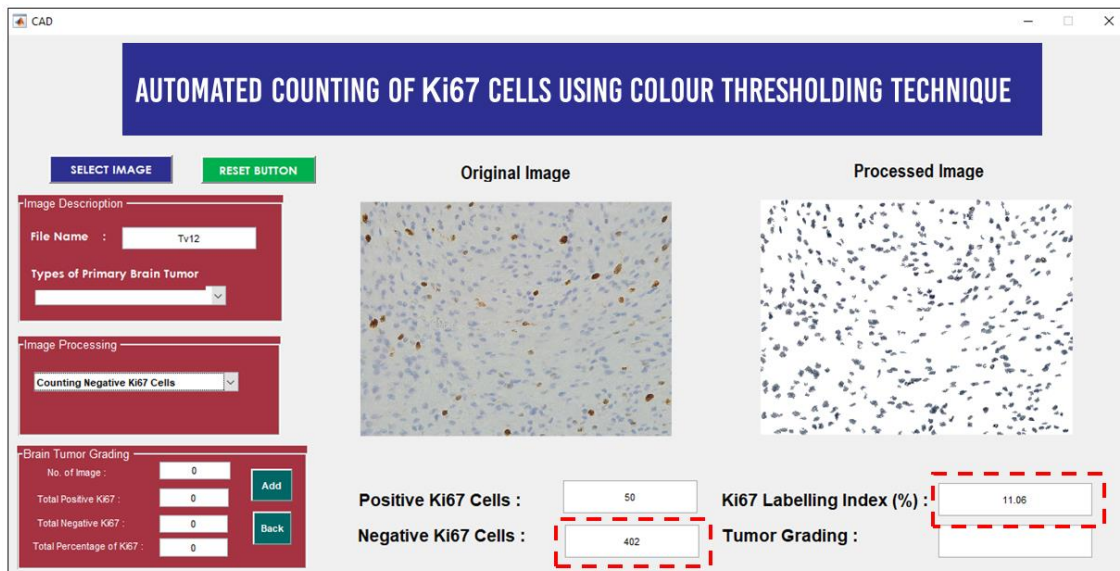


Figure 10. Layout of CAD software after counting negative Ki67 cells

The third panel was named 'Brain Tumour Grading'. Usually, the pathologists will use three or four histopathological images to determine the Ki67 index. The percentage result in box number 4 (Figure 9) was the index result for one image. Due to this factor, the third panel was designed. From Figure 7, there were four boxes (number 5 to 8) designed, which were named as 'No. of Image', 'Total Positive Ki67', 'Total Negative Ki67', and 'Total Percentage of Ki67'. Once the user clicks the 'Add' button, the system will add the current result (includes the number of images, number of tumour cells counting, number of non-tumour cells counting, and the percentage of Ki67 index) with the results that have been saved previously in those four boxes (number 5 to 8). If the user does not feel satisfied with the results or by mistake and inadvertently the user pressed the 'Add' button more than once, the user can click the 'Back' button to subtract the latest data from the old data values. After all the results have been obtained, the user should go to the 'Image Processing' panel and click on 'Meningioma' to obtain the tumour grading. Figure 11 shows an example of using three histopathological images to find the average of the Ki67 index.

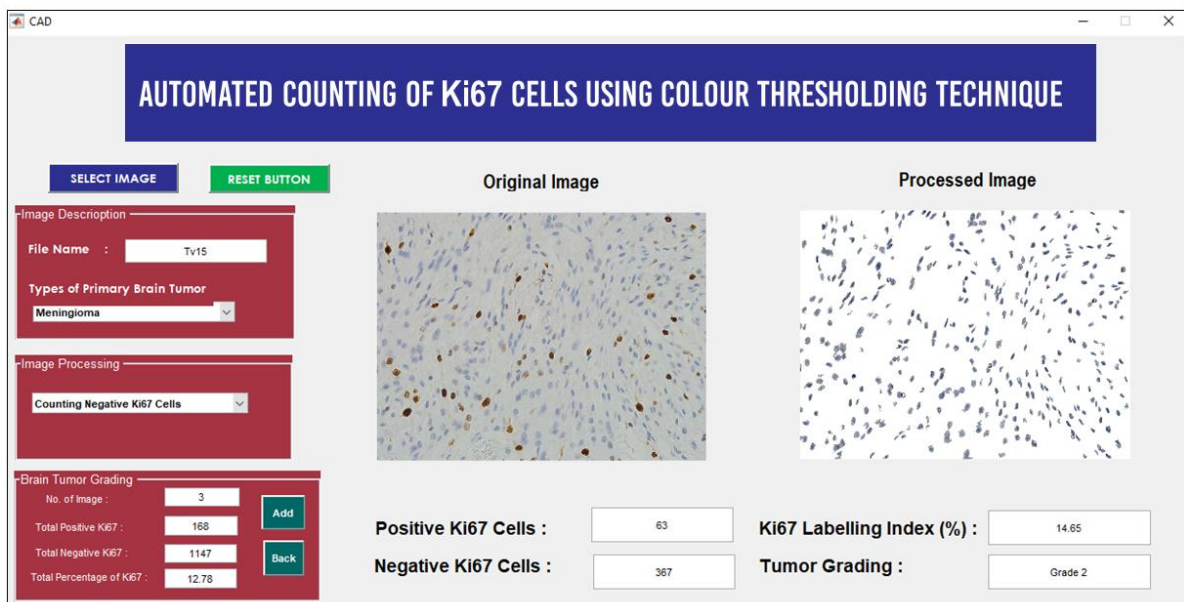


Figure 11. Layout of CAD software after calculating the average Ki67 index for determining the tumour grading.

Calculation of Ki67 Labelling Index and Determination of Tumour Grading

After counting the Ki67 cells, the next step was to calculate the Ki67 Labelling Index (LI). Table 4 compares the results of the Ki67 index between manual counting and automated counting. The absolute error was calculated to identify the amount of error obtained between manual and automated counting. The absolute error can be obtained by finding the differences of Ki67 index results between the automated and manual counting techniques.

Table 4. Comparison results of Ki67 index between the manual and automated counting techniques

Image	Ki67 Labelling Index		Absolute Error (%)
	Manual Counting (%)	Automated Counting (%)	
Cell 01.bmp	11.14	11.06	0.08
Cell 02.bmp	15.42	13.19	2.23
Cell 03.bmp	16.63	14.99	1.64
Cell 04.bmp	6.21	8.04	1.83
Cell 05.bmp	7.42	8.15	0.73
Cell 06.bmp	8.05	8.75	0.70
Cell 07.bmp	11.48	11.05	0.43
Cell 08.bmp	11.14	10.17	0.97
Cell 09.bmp	14.33	12.87	1.46
Cell 10.bmp	18.63	14.63	4.00
Cell 11.bmp	13.32	16.63	3.31
Cell 12.bmp	35.29	40.16	4.87
Average (%)			1.85

Based on Table 4, the average absolute error was obtained at 1.85%. The last three images showed higher results of absolute error. The counting results of Ki67 cells can affect the result of the Ki67 index. Several factors influenced the counting results. Firstly, it was due to the complex nature of the cells, which have a heterogeneous shape. Secondly was the quality of the captured images. The last three images (Cell 10.bmp, Cell 11.bmp, and Cell 12.bmp) were captured from the old slides. The quality of these images was low, where the colour intensity of negative cells was dull. Thus, it makes the proposed system challenging to identify and count the cells. Thirdly, it was due to the uncertainties of objects or noises that present in microscopic images. These uncertainties may occur during surgical removal fixation, staining procedures, processing, embedding, and microtomy [32].

The last step was to determine brain tumour grading. There was a difference between the manual counting technique and the proposed automated counting techniques in determining the grading. For the manual counting technique, the Ki67 index was calculated for each 'hot-spot' area image. Then, the pathologists will calculate the total average by adding all the Ki67 index results and dividing them by the number of 'hot-spots' area images. Thus, the grading will be determined based on the total average value. For the proposed automated counting techniques, the Ki67 index was calculated by finding the total numbers of positive Ki67 cells for all 'hot-spot' area images. Then, the results will be divided by the total number of positive and negative Ki67 cells for all 'hot-spot' area images. The resultant value will be used to determine brain tumour grading. Table 5 compares the grading results between the manual counting and the proposed automated counting techniques.

Table 5. Comparison results of grading between the manual and automated counting technique

Slide	Image	Manual Counting		Automated Counting	
		Average Ki67 Index (%)	Grade	Average Ki67 Index (%)	Grade
1	Cell 01.bmp	14.50	2	13.04	2
	Cell 02.bmp				
	Cell 03.bmp				
2	Cell 04.bmp	7.23	1	8.37	2
	Cell 05.bmp				
	Cell 06.bmp				
3	Cell 07.bmp	12.32	2	11.38	2
	Cell 08.bmp				
	Cell 09.bmp				
4	Cell 10.bmp	22.41	3	23.97	3
	Cell 11.bmp				
	Cell 12.bmp				

Based on Table 5, the automated counting technique was able to give an accurate grading result. From the four slides used in this study, the proposed system can acquire three slides with the same grading result as obtained from the manual counting technique. As for slide 2, the grading results produced by the two techniques are different. Since the result of the average Ki67 index for automated counting does not fall within the range, the grading result became different even though the gap of the average Ki67 index between these two techniques was small. Overall, the proposed system is able to deliver promising results in counting Ki67 cells and determine tumour grading.

CONCLUSION

In this study, a computer-aided diagnosis (CAD) system was designed to reduce pathologists' workload in evaluating the grading of one of the primary brain tumours, which is a meningioma. Previous studies have shown that a

few techniques can be applied to measure tumour cell proliferation activity. For this study, the nuclear antigen, Ki67 has been selected since it is a well-known independent prognostic and predictive indicator for evaluating the tumour proliferation cells. The Ki67 IHC staining was used in this study since pathologists commonly use it for determining the grade of a tumour from various types of cancer and tumour. The gold standard for measuring prognosis and the patient outcome is still using the manual diagnosis by counting each Ki67 cell individually. However, this technique is highly subjective, laborious, time-consuming, and subject to inter- and intra-pathologists' experience. Due to this reason, a semi-automated CAD system was developed to assist the pathologists by providing an automated image analysis to determine the grading of meningioma. Before using this CAD system, the user needs to capture the hot-spot areas manually from the specimen slide. All of the image analysis processes in this system are performed automatically, where no additional pre-setup is needed for the user to use it.

This CAD system consists of two components, which are counting Ki67 cells and scoring the Ki67 index. Various techniques of image processing have been applied in the first component to count the Ki67 cells. The second component was developed to calculate the Ki67 index for the determination of meningioma grading. Overall, the proposed technique was able to achieve good results in detecting and counting Ki67 cells. According to the segmentation results, most of the Ki67 cells were detected by the proposed system with an average accuracy of 95.02%. For counting results, there was a positive correlation between the proposed automated counting system with the manual counting technique performed by the pathologist. The Ki67 index result also demonstrated promising results with a low average absolute error of 1.85%.

This system has several functional features. The CAD system was developed using GUIDE. This software is capable of counting the Ki67 cells automatically, and it has the potential to reduce the cost and time of counting Ki67 cells and determining the meningioma grading. The main advantage of this CAD system was free from subjectivity issues, where this CAD system does not require any adjustments to perform the counting process. Besides, this software contained several GUI functions that may reduce pathologists' work, especially in counting Ki67 cells. Furthermore, this software can also count the Ki67 cells in a short time with an average computational time of fewer than 30 seconds with high accuracy. With this automated analysis system development, the diagnosis process becomes faster, more objective, and less laborious than manual diagnosis examination. Thus, it may reduce the time management of the pathologists in counting Ki67 cells. Based on the performance of the proposed system and the features provided by this CAD software, it has good potential to be used as an automatic counting Ki67 system to evaluate the grade of meningioma. For achieving better accuracy, it is suggested to increase the number of sample images in the future to provide convincing results, especially in counting the Ki67 cells and determining the tumor grading.

ACKNOWLEDGEMENT

The author would like to thank the pathologists from the Department of Pathology, Hospital University Science of Malaysia (HUSM) for helping and contributing to this study.

REFERENCES

- [1] American Cancer Society, "Cancer Facts & Figures 2017 - 2019," *Am. Cancer Soc.*, pp. 1–40, 2017.
- [2] Q. T. Ostrom, H. Gittleman, G. Truitt, A. Boscia, C. Kruchko, and J. S. Barnholtz-Sloan, "CBTRUS statistical report: Primary brain and other central nervous system tumors diagnosed in the United States in 2011-2015," *Neuro. Oncol.*, vol. 20, pp. 1–86, 2018.
- [3] L. T. Li, G. Jiang, Q. Chen, and J. N. Zheng, "Ki67 is a promising molecular target in the diagnosis of cancer (Review)," *Mol. Med. Rep.*, vol. 11, no. 3, pp. 1566–1572, 2015.
- [4] J. H. Neltner, H. Su, and F. Xing, "Comparison of Manual and Automatic Methods of Ki-67 Proliferation Index For Neuroendocrine Tumors: The Development and Validation of a Novel Digital Pathology Tool (Ki67 Counter)," *11th Annu. ENETS Conf.*, vol. 43, no. 3, pp. 250–250, 2014.
- [5] C. G. Loukas, G. D. Wilson, B. Vojnovic, and A. Linney, "An image analysis-based approach for automated counting of cancer cell nuclei in tissue sections," *Cytometry*, vol. 55A, no. 1, pp. 30–42, 2003.
- [6] A. Aslantaş, E. Dandil, S. Sağlam, and M. Çakiroğlu, "CADBOS: A computer-aided diagnosis system for whole-body bone scintigraphy scans," *J. Cancer Res. Ther.*, vol. 12, no. 2, pp. 787–792, 2016.
- [7] C. A. Schneider, W. S. Rasband, and K. W. Eliceiri, "NIH Image to ImageJ: 25 years of image analysis," *Nat. Methods*, vol. 9, no. 7, pp. 671–675, 2012.
- [8] M. D. Abràmoff, P. J. Magalhães, and S. J. Ram, "Image processing with ImageJ Part II," *Biophotonics Int.*, vol. 11, no. 7, pp. 36–43, 2005.
- [9] V. Tuominen, S. Ruotoistenmäki, A. Viitanen, M. Jumppanen, and J. Isola, "ImmunoRatio: a publicly available web application for quantitative image analysis of estrogen receptor (ER), progesterone receptor (PR), and Ki-67," *Breast Cancer Res.*, vol. 12, no. 4, p. R56, 2010.
- [10] S. Robertson, H. Azizpour, K. Smith, and J. Hartman, "Digital image analysis in breast pathology—from image processing techniques to artificial intelligence," *Transl. Res. J. Lab. Clin. Med.*, vol. 194, pp. 19–35, 2018.
- [11] Ö. Demir and A. Y. Çamurcu, "Computer-aided detection of lung nodules using outer surface features," *Biomed. Mater. Eng.*, vol. 26, pp. S1213–S1222, 2015.
- [12] D. Ștefănescu, C. Streba, E. T. Cârțână, A. Săftoiu, G. Gruionu, and L. G. Gruionu, "Computer aided diagnosis for confocal laser endomicroscopy in advanced colorectal adenocarcinoma," *PLoS One*, vol. 11, no. 5, pp. 1–9, 2016.
- [13] A. M. Karim, F. V. Çelebi, and A. S. Mohammed, "Software Development for Blood Disease Expert System," *Lect. Notes Softw. Eng.*, vol. 4, no. 3, pp. 179–183, 2016.

- [14] N. A. Shaharuddin and W. M. H. W. Mahmud, "Development of computer aided diagnosis system (CAD) for detection of kidney ultrasound images," *2017 Int. Conf. Eng. Technol. Technopreneurship, ICE2T 2017*, no. September, pp. 1–4, 2017.
- [15] M. H. Bahreyni-Tossi, M. S. Moghadam, and S. Nekooie, "A NEW CAD SOFTWARE FOR EVALUATION OF BREAST LESIONS IN CONTRAST – ENHANCED MR MAMMOGRAPHY," *Int. J. Life Sci. Pharma Res.*, vol. 8, no. 2, pp. 2–7, 2018.
- [16] X. Li, H. Yang, H. Huang, and T. Zhu, "CELLCOUNTER: Novel open-source software for counting cell migration and invasion in vitro," *Biomed Res. Int.*, vol. 2014, no. Figure 1, 2014.
- [17] J. O'brien, H. Hayder, and C. Peng, "Automated quantification and analysis of cell counting procedures using imagej plugins," *J. Vis. Exp.*, vol. 2016, no. 117, 2016.
- [18] R. Gonzi, J. Gonzi, and J. Gonzi, "Comparison between Manual and Automated Methods for Ki-67 Immunoexpression Quantification in Ameloblastomas," *Anal. Cell. Pathol.*, 2016.
- [19] A. Loddo, L. Putzu, C. DI Ruberto, and G. Fenu, "A Computer-Aided System for Differential Count from Peripheral Blood Cell Images," *Proc. - 12th Int. Conf. Signal Image Technol. Internet-Based Syst. SITIS 2016*, no. November 2017, pp. 112–118, 2017.
- [20] C. Venter and C. U. Niesler, "Rapid quantification of cellular proliferation and migration using ImageJ," *Biotechniques*, vol. 66, no. 2, pp. 99–102, 2019.
- [21] American Cancer Society, "Testing Biopsy and Cytology Specimens for Cancer," *Am. Cancer Soc.*, p. 20, 2015.
- [22] J. Kaur and A. Choudhary, "Comparison of Several Contrast Stretching Techniques on Acute Leukemia Images," *Int. J. Eng. Innov. Technol.*, vol. 2, no. 1, pp. 332–335, 2012.
- [23] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Trans. Syst. Man. Cybern.*, vol. 20, no. 1, pp. 62–66, 1979.
- [24] Khushbu and I. Vats, "Otsu Image Segmentation Algorithm," *Int. J. Innov. Res. Comput. Commun. Eng.*, vol. 5, no. 6, pp. 11945–11948, 2017.
- [25] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd ed. New Jersey, USA: Pearson Education, 2008.
- [26] L. Vincent, "Morphological Area Openings and Closings for Grey-scale Images," in *Shape in Picture*, Berlin, Heidelberg: Springer, 1994.
- [27] G. Kumar and P. K. Bhatia, "A detailed review of feature extraction in image processing systems," in *International Conference on Advanced Computing and Communication Technologies, ACCT*, 2014, pp. 5–12.
- [28] D. N. Louis et al., "The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary," *Acta Neuropathol.*, vol. 131, no. 6, pp. 803–820, 2016.
- [29] S. T. Smith, *MATLAB: Advanced GUI Development*, 1st ed. Indianapolis: Dog Ear Publishing, 2006.
- [30] P. Purkait, B. Biswas, S. Das, and C. Koley, "Measurement of Errors," in *Electrical and Electronics Measurements and Instrumentation*, New Delhi: McGraw-Hill Education, 2013.
- [31] Y. Zhang, H. Wang, Z. Yang, and J. Li, "Relative Accuracy Evaluation," *PLoS One*, vol. 9, no. 8, pp. 1–13, 2014.
- [32] S. Jimson, L. Malathi, G. M. K. Kumar, and N. Balachander, "Artifact in histological section," *Biomed. Pharmacol. J.*, vol. 9, no. 2, pp. 843–845, 2016.