American Journal of Biomedical Sciences
ISSN: 1937-9080
nwpii.com/ajbms

A Review: Recent Approaches on Improving the Accuracy in Determining the Tumor and Cancer Grading in Histopathological Images Based on Scoring Ki67 Expression

Fahmi Akmal Dzulkifli[1*], Mohd Yusoff Mashor[1], Khairul Shakir AB Rahman[2], Haidi Mohamad[3]

[1] Faculty of Electronic Engineering Technology,Universiti Malaysia Perlis (UniMAP),02600, Arau, Perlis, Malaysia fahmiakmaldzulkifli@gmail.com yusoff@unimap.edu.my
[2] Department of Pathology,Hospital Tuanku Fauziah,Jalan Tun Abdul Razak, Pusat Bandar Kangar 01000, Kangar, Perlis, Malaysia ksyakir@gmail.com
[3] Department of Pathology, Hospital Sultanah Bahiyah,Jln Langgar, Bandar Alor Setar 05460 Alor Setar, Kedah, Malaysia haidimohamad@gmail.com
[*]**Corresponding Author**
Fahmi Akmal Dzulkifli
Faculty of Electronic Engineering Technology,
Universiti Malaysia Perlis (UniMAP)
02600, Arau, Perlis
Malaysia
Email: fahmiakmaldzulkifli@gmail.com

Received:27 May 2021; | Revised:29 May 2021; | Accepted: 09 September 2021

Abstract

Tumor markers are substances formed by cancer cells or other cells in the body, which aims to indicate the presence of cancer and provide information related to cancer growth. The expression of Ki67 is usually associated with tumor cell proliferation and growth. Basically, the pathologists will count the number of Ki67 expressions to determine the growth fraction of tumor cell proliferation. Conventional manual techniques have shown their drawbacks in counting Ki67 expression, such as time-consuming, low accuracy and reliability, poor reproducibility, and highly subjective. Therefore, this paper will focus on the recent studies in improving the quantification of Ki67 expression for determining the tumor or cancer grading. This study divides the approaches into three, which are double or multiple staining methods, image processing or image analysis software, and artificial intelligence.

Keywords: Artificial intelligence, Histopathological Images, Image processing, Ki67 Expression, Tumor Markers

1. Introduction

Tumor markers are biochemical substances, commonly referred to as proteins produced by tumor cells or other body cells in response to cancer or certain benign conditions [1]. These markers can be divided into two types, and each type has different roles in cancer care. First is the circulating tumor markers. These markers are present in the blood, urine, stool, or other cancer patients' bodily

fluids. These markers' function is used to estimate prognosis, examine the responsiveness to treatment, observe whether cancer has been immune to therapy, and identify cancer that remains after treatment or that has returned after treatment [2]. The second type is the tumor tissue markers. These markers are usually found in a biopsy sample, typically presented in formalin-fixed paraffin-embedded (FFPE) tissue samples [3]. These tumor markers are used to diagnose and classify the stage of cancer, estimate prognosis, and to determine the best treatment plan for a patient [2].

Ki67 is one example of prognostic tumor markers that is increasingly gaining attention among researchers and clinicians nowadays. Ki67 is a nuclear antigen that responds to a monoclonal antibody MIB-1. This nuclear protein is generally associated with tumor cell proliferation and growth. Previous studies have shown that the potential of Ki67 as the proliferation marker in diagnosing various types of tumor and cancer diseases such as breast, lung, prostate, pancreatic, and central nervous system. The Ki67 was discovered for the first time in the early 1980s by Gerdes [4], where it was found in an attempt to create cancer-specific monoclonal antibodies. The Ki67 is only present in the cell cycle's growing and dividing phases (G1, S, G2, and M) and absent in the resting phase (G0). The Ki67-antibody did not show any reaction to the cells known to be in a resting phase, such as lymphocytes, monocytes, brain cells, renal cells, and other cells. Due to this fact, it makes the Ki67 appear to be a useful proliferation marker since cancer cells aggressively grow and divide.

The preparation of sample tissues using the histological stains enables the pathologists to measure the percentage of tumor cell proliferation indicated by positive Ki67. This percentage was known as the Ki67 Labelling Index (LI). This LI is represented in percentage form, where this value refers to the percentage of tumor cell proliferation, which is indicated by immunopositive Ki67 among the cell population. The percentage value was obtained by dividing the total number of positive Ki67 cells by the total numbers of positive and negative Ki67 cells and multiplying it by 100.

In order to obtain the number of positive and negative Ki67 cells, pathologists need to count every cell in the slide specimen. Generally, there are several methods available for scoring the Ki67 cells: 'eye-balling' estimation, visual counting using a microscope or any viewer software, and manual counting based on images captured by a digital microscope or on printed images. The 'eye-balling' estimation is the simple, fast, and straightforward technique. For this method, the pathologists only need to scan the whole slide of the tissue specimen through the microscope's eyepiece lenses and estimate the number of Ki67 cells. Although this method is time-saving, the accuracy of the counting is poor with low reliability and reproducibility. The next method is visual counting using a microscope or any viewer software. This method is known as 'real-time counting' since counting is done directly by looking through the microscope. With this method, the pathologists can identify and distinguish between the tumor cells and other cell types. However, when having a large area with large numbers of cells, this method becomes impractical and poorly reproducible. Another method is manual counting based on images captured by a digital microscope or on printed images. This method requires the pathologists to annotate and count each of the cells in the image manually. This method is presently admitted as the 'gold standard' and reliable method since the reproducibility and accuracy are high [5]. Among the researchers, this method becomes the ground truth when they want to measure the performance of the developed system or algorithm. Still, this method also has drawback issues. Since the counting process was done manually, it makes this method look tedious, tiring, and highly subjective [6]. Besides, this method was also reported as a laborious and time-consuming method [7]. Therefore, this paper will review the previous techniques used to solve the addressed issues, especially in counting Ki67 expression.

2. Techniques used in Quantifying Ki67 Expression

An accurate method or developed system is essential for counting Ki67 expression since it will determine the tumors' grades and the future treatment that will be delivered to the patient. Thus, many experiments and studies have been carried out to improve Ki67 scoring.

2.1 Double or Multiple Staining Methods

Staining refers to a set of procedures performed to prepare a tissue sample using dyes or histological stains to visualize cell structures and components under a microscope [8]. In the histopathological study, this staining process can assist the pathologists in identifying tissue abnormalities or any particular indicator in the actual cells. Furthermore, this process can also help locate disease or tumor cells in a sample tissue. The purpose of staining is to highlight the beneficial features of tissues, besides enhancing the tissue contrast.

Ki67 immunohistochemical (IHC) staining is a technique where antibodies are used to identify an antigen in a sectioned tissue. The sample tissue will be stained with the Diaminobenzidine (DAB) and counterstained with Hematoxylin (H). As a result, the immunopositive Ki67 cells will appear in granular brown color, while the immunonegative Ki67 cells appear in blue. However, there are a few aspects that could affect the accuracy results when using this staining. Firstly, there are no clear guidelines or consensus regarding defining the immunopositive Ki67 cell. The variations of cellular level in size, shape, and morphological and textural features may become a crisis, especially for the automated counting system [9]. Secondly is the miscounting issue. This issue arises when some unwanted and related objects have similar color intensities as the Ki67 expression, such as lymphocyte and cytoplasmic brown staining [5].

Matsukuma et al. [10] stated that the common problem in differentiating neuroendocrine tumors in Ki67 staining was the existing background stromal lymphocytes, entrapped non-neoplastic glands, and vascular network that also contained proliferating cells. Next is the negative Ki67 cells, which appear in blue. The negative Ki67 cells consist of stromal and epithelial cells [11]. However, only epithelial cells were considered for scoring purposes. Besides that, the conventional IHC technique only allows labeling a single marker per tissue section [12]. Due to these issues, double or multiplex immunostaining has been introduced to improve the accuracy in counting Ki67 expression. This multiplex staining has increasingly popular in the medical field since it provides precise diagnosis and, eventually, higher treatment success [13]. The multiplex staining refers to a combination of individual antigen detections used to examine the expression of multiple markers on a single tissue section [12]. Many types of experiments related to the application of multiplex staining have been carried out to improve the scoring of Ki67 expression.

Van der Loos et al. [14] had designed a method to improve the Ki67 proliferation index's determination in rabbit liver using a multicolor immunohistochemical approach. This study was performed to improve Ki67 positive cells' quantitation by differentiating between the Ki67-positive hepatocytes and non-parenchymal cells (leucocytes) in the rabbit liver tissue. In this study, three different IHC staining types were conducted to measure the Ki67 positive cells' performance. These stainings consist of single staining (Ki67), double-staining with cytokeratin (Ki67 and CK), and triple-staining (Ki67, CD31, and CK). All of these stainings were then analyzed using the image analysis software. The counting results were further compared with the manual counting results, which was done by two observers. Based on Spearman's Correlation coefficient test, the triple-staining method achieved a high correlation in counting positive nuclei compared to the double-staining technique with 0.91 and 0.76, respectively. Therefore, the results revealed that triple-staining method had significantly improved the accuracy in counting the number of proliferative hepatocytes.

Matsukuma et al. [10] developed a method using the Synaptophysin-Ki67 double immunostaining technique to improve the interobserver agreement in grading the gastrointestinal neuroendocrine tumor. For this study, the authors used Synaptophysin to differentiate between tumor and non-tumor cells to provide consistency in interpreting tumor proliferative index and tumor grade. The proposed technique was prepared for the simultaneous detection of two antigens, where DAB is one chromogen, and Alkaline Phosphate is the second chromogen. The DAB chromogen was selected for detecting Ki67, while the second chromogen was used to visualize Synaptophysin. To measure the proposed method's performance, the counting results for these two staining types (Ki67-only stained slides and Synaptophysin-Ki67 double immnunostaining) were compared with manual counting results. Based on the Intraclass Correlation

Coefficient (ICC) results, the Synaptophysin-Ki67 double immunostaining method had a higher ICC value than the Ki67 immunostaining with 0.79 and 0.51, respectively. The authors summarized that double staining's implementation had increased their confidence in quantifying tumor nuclei in areas that contained a significant number of obscuring non-neoplastic cells.

Koopman, Buikema, Hollema, de Bock, and van der Vegt [15] applied the virtual dual staining (VDS) technique by superimposed two different stained images to assess the Ki67 LI in breast cancer. Two types of antibody markers were used in this study, which consists of Ki67 and cytokeratin. In this study, two digital image analysis platforms were used to perform the image analysis procedures. At first, the VDS technique was applied to locate tumor tissue areas by aligning the corresponding Ki67 and cytokeratin stained sections digitally. During this process, the algorithms automatically performed the distortion and rotation modifications to remove the small difference due to tissue and section processing. Later, the whole Ki67-stained section will be annotated using the cytokeratin-stained, which acts as the reference and tumor classifier. At this moment, the existing large areas of carcinoma in situ, pre-existent epithelium, and artifacts were excluded. Next, both image analysis platforms will proceed to the counting process, and the results of the quantification were compared with the manual counting. Based on the Spearman's correlation coefficient results, Platform A obtained higher inter-observer agreement than Platform B with 0.94 ($p < 0.001$) and 0.93 ($p < 0.001$) respectively. Hence, it showed that the VDS technique's application could be an alternative method to determine the Ki67 proliferation index on a whole section of invasive breast carcinomas.

Mejías-Badillo et al. [17] introduced a protocol that can differentiate between tumoral and non-tumoral cells for accurate grading of pancreatic neuroendocrine tumors. The purpose of this study was to develop a protocol that can distinguish between the tumoral cells and the reactive infiltrating lymphocytes in biopsies of pancreatic neuroendocrine tumors. This protocol involved dual-color immunostaining, which are Ki67 and Leucocyte Common Antigen (LCA). The simultaneous immunostaining allowed the tumoral cells to appear in brown color (Ki67), and the lymphocytes appeared in both nuclear brown and red cytoplasmic color (LCA). The dual-color immunostaining will be compared to the Ki67 single-color immunostaining to test the proposed protocol's efficiency. Based on the Ki67 labeling index results, the dual-color immunostaining achieved higher precision for Grade I than the single-color immunostaining with an average of 1% and 3%, respectively. For Grade II results, the average Ki67 labeling index obtained from the dual-color immunostaining was 6%, while for single-color immunostaining was 7%. The higher percentage values obtained from the single-color immunostaining indicated some miscategorization in defining the tumor cells. Thus, it showed that the proposed protocol had simplified the interpretation and improved precision counting of Ki67 cells.

2.2 Image Processing Techniques and Image Analysis Software

Image processing is a technique applied to a digital image to enhance image quality or extract valuable information from the image. Nowadays, image processing applications are wide-ranging in all fields, such as industrial robotics, astronomy, medical, pattern recognition, video processing, security, traffic control, and object detection. Recently, image processing also had provided enormous contributions to clinical pathology. The development of an automated image analysis system reduces the pathologists' time in diagnosing the tissue specimens and increases the reproducibility of the measurements [18].

Alomari, Abdullah, Zin, and Omar [19] had designed an iterative randomized circular detection algorithm to determine the proliferation rate estimation (PRE) in brain tumor using Ki67 histopathological images. The main objective was to develop an algorithm that can detect and quantify the Ki67 expression with irregular circular shapes. The proposed algorithm is composed of six steps: color space transformation, customized color modification, nuclei segmentation, pre-processing the extracted cells, counting the positive and negative normal cells, and calculate the PRE value. First, the input image's color space was converted from RGB to L*a*b* color space. Second, the pixels with dark brown color intensity will be

modified into a light brown color for segmentation purposes. Third, using the K-means clustering technique to classify the image into three clusters (brown nuclei, blue nuclei, and background). Next, the algorithm applied the Canny operator to detect the edge of the cell nuclei. Later, the algorithm used Otsu's thresholding to obtain the binary image. The erosion operation was then used to reduce the overlapped cells. The counting process was done using the circularity features, which used the iterative structured nuclei detection (IRIC) algorithm. Based on the cells detection results, the proposed algorithm achieved high accuracy in detecting the positive and negative Ki67 expression, with an average of 98.0% and 98.3%, respectively. The counting results obtained by the proposed algorithm were also high, with a precision of 97.8% for negative Ki67 cells and 98.7% for positive Ki67 cells.

Razavi et al. [20] proposed an automated and integrated framework for scoring Ki67 expression in breast cancer tissue images. The proposed framework was constructed based on eight steps. These include image acquisition, tile-based smoothing and enhancement, color decomposition, post-processing, morphological and shape operations, counting positive and negative Ki67 expression, merging tiles, and presenting the visual and analytical pathological results. At first, the whole-slide image (WSI) Ki67 staining image will be divided into several tiles, where each tile had a size of 256 × 256 pixels. The tiled images' quality was enhanced using the edge preserved smoothing, median filter, noise reduction, and bilateral filter. Next was the identification of positive and negative Ki67 expression. Several techniques were used to identify the negative Ki67 cells: extracting hematoxylin channel, application of median and bilateral filters, segmentation using adaptive Huang thresholding method, the watershed technique for removing the connected cells, extracting the morphological and shape features, and implementation of K-means cluster. For detecting positive Ki67 cells, the process was similar as detecting the negative cells. After finishing all the processes, all tiled images' will be merged to get the output result. Based on 30 datasets, the proposed method achieved 91% classification accuracy, 0.93 for precision, 0.89 for recall, and 0.91 for F-score. The authors concluded that the proposed method provides a better scoring performance and faster in delivering the results due to automatically process of discriminating between cancerous and non-cancerous cells.

Sugita et al. [21] investigated the role and effectiveness of the image analysis software in quantifying Ki67 cells and calculating the recurrence-free survival (RFS) in gastrointestinal stromal tumor (GIST) patients. In this study, the authors used two different types of images with two different image analysis software. The sample images consist of WSI and manually-captured images. For the first Ki67 quantitation using WSI, the image was scanned using the slide scanner to obtain the WSI of Ki67. Then, the authors used the image analysis software (Tissue Studio) to perform automatic quantitation. For the second method, the "hot-spot" images were captured using a microscope attached with a digital camera. Later, the authors used the Patholoscope software to count the Ki67 cells. The counting results for both methods were compared with the National Institutes of Health (NIH) risk classification results. According to the Ki67 quantification results, the manually-captured images method produced more accurate results compared to WSI. This result also showed that the manually-captured images method could differentiate between low, intermediate, and high-risk GIST patients.

Saadeh, Abdullah, Erashdi, Sughayer, and Al-Kadi [22] had designed a semi-automated framework for calculating the Ki67 index in gastropancreatic neuroendocrine tumors. The process began using the Gaussian low-pass filtering to the input image for smoothing the inhomogeneity of the nuclear staining. The next step was selecting the red channel, which had maximum nuclei intensity homogeneity, to separate the objects from the background. Then, the authors used the histogram equalization technique to enhance the contrast of the image. Following this step was applying Otsu's thresholding technique to differentiate between the desired objects and the image's background. Several image post-processing techniques, such as image inversion, dilation, and fill image regions, were then applied to the segmented image. Afterwards, the authors employed the watershed algorithm to solve nuclei's overlapping, touching

cells, and occlusion issues. For feature selection, the authors chose the circularity and shape features to separate the Ki67 cells from unwanted objects. Based on the correlation test results, a strong agreement was obtained between the proposed framework and the manual counting method in calculating the Ki67 index with an Intraclass Correlation Coefficient (ICC) of 0.993 and Lin's Concordance Correlation Coefficient (CCC) of 0.986.

Geread et al. [11] developed an automatic and unsupervised color separation method for improving the Ki67 index calculation in breast cancer histopathological images. The workflow involved four phases: pre-processing, color separation, nuclei detection, and proliferation index calculation. In the pre-processing phase, the noises of the image were removed using the Vector Median Filter. After applying the de-noising technique, the proposed method performed the background subtraction to obtain the tissue and cell regions. The next phase was the color separation. Before moving to the next phase, the input image's color space was converted from RGB to L*a*b* color space. For this study, the authors create two different thresholds from the b* channel to separate the image's brown and blue color. The first threshold was to obtain the hematoxylin staining pixels, while the second threshold was to find the DAB pixels. Moving to the next phase was the nuclei detection phase. A voting image was applied to determine the center of a cell with the supplied cell radius. The cells' radius was determined based on the adaptive cell radius estimator, which the authors implemented to automatically obtain the radius parameter values. A scatter plot was constructed to find the agreement between the manual and automated approaches in measuring the Ki67 index. The correlation results revealed a strong relationship between the proposed method with the manual counting method, where the Spearman's correlation coefficient was obtained at 0.944 with an R2 = 0.8667. Based on 80 image datasets, the proposed method could classify 74 images correctly, resulting in 92.5% accuracy.

2.3 Artificial Intelligence (AI)

Digital pathology refers to a collection of digital workflow and imaging solutions that aims to create a digital image-based practice environment for diagnostic, prognostic, and specific content purposes [23]. As the demand for digital pathology is gradually increasing in the market, the technology used in this sector is becoming more advanced. Hence, AI is increasingly and continuously being used to support digital pathology, allowing scientists in developing new capabilities for the technology [24].

Briefly, AI is the capability of a machine or system to imitate intelligent human behavior to solve complex problems in a 'human way' [25]. In digital pathology, AI's application enables processing and analyzing vast datasets and performing more comprehensive and accurate analyses [24]. Generally, the pathologists will examine the histopathologic features of tumor cells before deciding the disease's grading. The diagnostic method can take a long time since it depends on the staining preparation and the quality of the images captured from the digital microscope. Besides, the results' accuracy may be less due to issues of subjectivity and inter-observer variability [26]. Considering the problems mentioned above, many researchers and scientists proposed their techniques based on AI applications to improve the accuracy in determining a tumor disease's grading.

Swiderska, Markiewicz, Grala, and Kozłowski [27] had presented a technique based on mathematical morphology and texture analysis to determine the tumor proliferation index in meningiomas and oligodendrogliomas. This study focused on developing an algorithm to detect the immunopositive Ki67 cells in the tumor area as well as identify and remove the hemorrhage areas from the tissue specimen. The proposed technique consists of three stages: defining a map of the sample, eliminating the hemorrhage areas, and localizing "hot-spot" areas. At first, thresholding and morphological filtering were used to obtain the specimen map. Then, using the Support Vector Machine (SVM) classification and texture analysis, the proposed technique will classify between the tumor and hemorrhage areas. Two different approaches have been proposed to recognize the tumor areas: based on color representation (Algorithm 1) and mathematical morphology operations (Algorithm 2). Next, the proposed technique will locate the "hot-spot" areas by finding

the local maxima with the highest immunopositive cell nuclei density. According to the Wilcoxon matched-pairs test, a significant difference was found when compared between Algorithm 2 with experts results with the correlation value of ($Z = 2.44$, $p = 0.014$). However, the authors concluded that Algorithm 2 had an advantage compared to Algorithm 1 in detecting the "hot-spot" areas in the presence of staining artifacts.

Wang et al. [26] demonstrated an automated grading platform with quantitative results interpretation based on machine learning models for multiparametric glioma. The automated system comprises five major components: automated region-of-interest (ROI) detection, feature extraction, essential feature selections, automated grading, and result interpretation. The first step is to partition the WSI of the input image into several tiles, where each tile will have a resolution size of 5120×5120 pixels. The second step was to apply the watershed nuclei segmentation algorithm in each tile. Five tiles with the highest densities of nuclei proliferation were selected based on the energy index map. The next step was to select the most representative and informative features by using the Random Forest classifier. The extracted features consist of 11 visual features, 15 sub-visual features, and one immunohistochemical feature. Then, the SVM classifier was used to grade the samples automatically. For further understanding, the authors use the Local Interpretable Model-Agnostic Explanations (LIME) algorithm to identify the selected features' contributions in each grading result. As a result, the proposed system achieved a good result with a validation performance of 0.88 ± 0.14. The accuracy in grading the glioma cases was also high, with more than 90%. The result also showed the visual features were the main contribution features in classifying the glioma cases with the highest accuracy of 0.76.

Feng et al. [28] had applied the deep learning algorithm for scoring Ki67 expression in breast invasive ductal carcinoma (IDC). The workflow of the study is consists of three important stages. The first stage is the identification of the IDC area. The pathologists will label the IDC areas first on the WSI specimens. Then, an image processing algorithm was used to segment and extract all the labelled areas into several patches with a size of 128 $\times$ 128 pixels. Next, these patches will be fed up with the training algorithm for extracting features and classification purposes. Afterward, the proposed system will identify the IDC areas based on the testing dataset using the GoogleNet Incubation V1. The second stage is applying image registration to H&E and IHC stained section to locate the immunopositive and immunonegative Ki67 cells. After obtaining the position of those cells, the authors will extract features related to immunopositive and immunonegative Ki67 cells by using the toolbox from ImageJ software. These features later will be trained by using a random forest classifier for the classification task. The classification results will be used for counting purposes and followed by comparing with the calculated results performed by the pathologists. As a result, the proposed system had an accuracy rate of 99.4% for scoring Ki67 and 89.44% for identifying breast IDC regions.

Liu et al. [29] introduced a method that can predict the Ki67 positive cells in H&E stained slides of the neuroendocrine tumor using the deep-learning model. The proposed system will start by identifying and selecting the ROI of positive and negative tumor regions in H&E stained slides by using the IHC stained slides as the reference image. After obtaining the ROI, the authors will annotate the positive cells, negative cells, and the background using a point label. Following this step was the extraction of these annotated images for training the model. Each of these images will have a size of 64×64 pixels. Afterwards was the classification process, where the authors used a modified ResNet18 as the classifier and separated it into three classes that consist of positive Ki67, negative Ki67, and background. The modification part has been made by first removing the last average pooling layer, and the second was changing the number of output nodes. The authors also applied a transformation method at the training network by converting all the fully connected layers into convolutional layers. The purpose of this transformation was to enable the network to perform the classification maps, even though the size of the ROI image was different. According to the classification results, the proposed model obtained high accuracies on the training set and validation set with values of 0.9780 and 0.9371,

respectively. Table 1 summarizes the comparison of three approaches techniques in improving the Ki67 cells' counting and calculation of the Ki67 proliferation index.

Table 1: A summary table of different approaches techniques in improving the Ki67 cells' counting

Authors and Year	Sample Specimens	Approached Techniques	Remarks
Loos *et al.* (2013)	Rabbit Liver Resection	1. Single staining with Ki67. 2. Double-staining with CK and Ki67. 3. Triple-staining with CK, CD31, and Ki67. 4. Multispectral un-mixing analysis.	1. The integration of IHC triple-staining with the tissue segmentation software significantly improved the accurate estimation of the actual number of proliferating hepatocytes. 2. The classification is based on multiple staining and texture analysis. 3. Image pre-processing does not mention in the study.
Swiderska *et al.* (2015)	Meningiomas and Oligodendrogliomas	1. Texture Analysis and Mathematical Morphology Operations. 2. SVM Classification. 3. Mathematical morphology operation. 4. Color Representation.	1. The proposed method can identify and eliminate the hemorrhage areas and staining artifacts from the tumor specimen map. 2. It is required to identify the features first before the classification process. 3. Classification based on texture analysis.
Alomari *et al.* (2016)	Brain Tumor	1. Color Space Transformation 2. Nuclei Segmentation 3. Edge Detection 4. Counting based on circularity features	1. The proposed method can detect irregular circle shapes and overlapping cells. 2. Image pre-processing is based on the color modification. 3. This study focuses on analyzing circle shapes and overlapping cells, which will be useful during the counting process. However, there are no explanations mentioned in differentiating between tumor cells and other cell types.
Matsukuma *et al.* (2017)	Gastrointestinal Neuroendocrine Tumor	1. Double-Staining (Synapthophysin and Ki67)	1. The proposed system can distinguish between gastrointestinal tumors from non-tumor cells. 2. Classification based on multiple staining.
Razavi *et al.* (2018)	Breast Cancer	1. Image enhancement 2. Image Segmentation 3. Patch-based image processing and tiling 4. Morphological operations	1. The proposed method can exclude stromal and epithelial tissues based on morphological and shape features. 2. Image pre-processing focusing on noise reduction, edge-preserved smoothing.
Koopman *et al.* (2018)	Breast Cancer	1. Virtual Dual Staining (CK and Ki67)	1. Able to exclude large areas of carcinoma in situ, pre-existent epithelium, and tissue or staining artifacts. 2. The VDS technique requires calibration, especially in aligning between the two different stained slides. 3. Image pre-processing is not mentioned in this study.

Sugita *et al.* (2018)	Gastrointestinal Stromal Tumor	1. Image Analysis	1. This study uses two different image analysis software with two different quantification methods (based on WSI and manually captured images). 2. The selection of 'hot-spot' areas was made manually. 3. Some parameters need to be set in advance for handling the different staining methods.
Saadeh *et al.* (2019)	Gastrointestinal Neuroendocrine Tumor	1. Image pre-processing 2. Image Segmentation 3. Image post-processing 4. Feature extraction	1. The proposed method shows promising results in quantifying the Ki67 expression with the semi-automated cell counting software. 2. Classification between the Ki67 expression and other unwanted cells is based on the shape and size features. 3. The selection of 'hot-spot' areas was made manually.
Wang *et al.* (2019)	Glioma Tumor	1. Watershed Algorithm. 2. Feature Extraction and Selection. 3. Classification based on a machine learning model	1. Since this study uses the conventional machine learning methods for classification, selecting the important features is necessary to feed up in the classification algorithm. 2. Grade I gliomas are not included in this study.
Geread *et al.* (2019)	Breast Cancer	1. Color Separation 2. Nuclei Detection	1. This study introduces two novel techniques, which are focusing on color separation and nuclei segmentation. 2. Classification between Ki67 expression and other types of cells are not specified in this study.
Feng *et al.* (2020)	Breast Cancer	1. GoogLeNet Inception V1 2. Random Forest classifier 3. Image registration	1. The proposed method able to identify the IDC areas by using the deep learning technique. 2. The details about image processing techniques used are not mentioned in this study.
Liu *et al.* (2020)	Neuroendocrine Tumor	1. ROI selection and background samples extraction. 2. Deep Learning Classification (modified ResNet18)	1. This study can predict Ki67 positive cells from H&E stained images by using a deep convolutional network model. 2. The selection of 'hot-spot' areas are not included in this study. 3. Image pre-processing is not mentioned in this study.
Badillo *et al.* (2020)	Pancreatic Neuroendocrine Tumor	1. Dual-color immunostaining protocol (a combination of Ki67 and leukocyte common antigen).	1. The dual-color staining increased the precision of quantifying Ki67 LI due to the ability to exclude LCA immunoreactive lymphocytes. 2. The authors used two types of counting methods: manual and automated counting. However, the detailed process of the automated counting was not discussed in this study.

3. Conclusion

An accurate system in counting and scoring the Ki67 expression is vital since it will determine the future results and outcome treatment for treating the patient. This paper discusses three different methods used to improve the system's performance in quantifying the Ki67 expression. The first method is double or multiple staining methods. In terms of accuracy, this method had shown significant and promising results in counting the tumor cells. The combination of multicolor and different immunostainings is beneficial to the pathologists since it can differentiate between two (or more) different antigens in the same sample of the tissue specimen. Besides, this method can also solve the inter-observer variability between the pathologists in counting and grading a tumor case. However, this method is more time-consuming and costly in terms of slide preparation than the conventional staining method.

The second method is about using image processing techniques and image analysis software. Various methods include semi-automated and automated systems based on image processing techniques, were discussed in this paper. Overall, this method can reduce the pathologists' workload, especially in counting each of the Ki67 cells in the digital image. The automated counting system development can save the pathologists' time and energy in scoring the cells. According to selected studies in this paper, most researchers who apply the image pre-processing technique focus only on eliminating the existing noises for better segmentation purposes. However, the slide preparation output also plays an important role, especially for the automated system. Geread et al. [11] stated the variability in slide preparation and staining protocols can create stain inconsistency, overstaining, other image artifacts. Thus, it is required to have another algorithm that will standardize the staining intensity for every input image in the system.

The third approach is the use of artificial intelligence. The use of AI in the field of histopathology is gaining attention nowadays. Based on the selected studies that are reviewed in this paper, it showed that the application of artificial intelligence had improved the accuracy of the counting performance. Moreover, this method also allows researchers to solve specific tasks, specifically in areas where humans have limited abilities (such as differentiating between the lymphocytes and positive Ki67 cells, distinguishing between stromal cells and negative Ki67 cells, and identifying the hemorrhages areas). Additionally, this method also showed promising results in classifying the grading of a tumor. Nevertheless, most of the studies still used conventional or supervised machine learning techniques to classify each cell in the specimen. The disadvantage of this technique is it requires pre-defined and selection features so that the algorithm can learn from the input features [30]. The shortcoming of these pre-defined features is the need of finding the most informative features in the classification task. Usually, these informative features are difficult to know, especially when there are numerous data are available [30]. Compared with unsupervised learning like the deep learning technique, the model will work independently to find and discover useful information that can be used for categorization. For this reason, this method is advantageous and preferred compared to the supervised machine learning techniques. It is expected that the combination of image processing methods and deep learning techniques would be effective in the future, notably in boosting pathologists' decision-making for better patient care and management.

Conflicts of Interest

The authors declare that they no conflicts of interest.

Authors' Contributions

F.A.D. prepared the original draft; M.Y.M. reviewed and edited related to engineering field; K.S.A.B. and H.M. reviewed and edited related to the medical and pathological field. All authors read and approved the final manuscript.

References

1 A.Wahed and A.Dasgupta, Pitfalls in Testing for Common Tumor Markers,in Accurate Results in the Clinical Laboratory, 2nd ed., A.

Dasgupta and J. L. Sepulveda, Eds. ***Cambridge: Elsevier***, 2019; 191-211 [DOI: 10.1016/C2016-0-04205-5]

2 N.C.Institute, Tumor Markers, National Cancer Institue, 2019. [Online]. Available: https://www.cancer.gov/about-cancer/diagnosis-staging/diagnosis/tumor-markers-fact-sheet. [Accessed: 27-Dec-2020].

3 Kokkat TJ, Patel MS, McGarvey D, LiVolsi VA, Baloch ZW. Archived formalin-fixed paraffin-embedded (FFPE) blocks: A valuable underexploited resource for extraction of DNA, RNA, and protein. ***Biopreserv Biobank*** 2013; 11(2): 101-106 DOI: 10.1089/bio.2012.0052

4 Gerdes J, Schwab U, Lemke H, Stein H. Production of a mouse monoclonal antibody reactive with a human nuclear antigen associated with cell proliferation. ***Int J Cancer*** 1983; 31(1): 13-20 DOI: 10.1002/ijc.2910310104

5 Cottenden J, Filter ER, Cottreau J, Moore D, Bullock M, Huang WY, Arnason T. Validation of a Cytotechnologist Manual Counting Service for the Ki67 Index in Neuroendocrine Tumors of the Pancreas and Gastrointestinal Tract. ***Arch Pathol Lab Med*** 2018; 142(3): 402-407 DOI: 10.5858/arpa.2017-0203-OA

6 Joseph J, Roudier MP, Narayanan PL, Augulis R, Ros VR, Pritchard A, Gerrard J, Laurinavicius A, Harrington EA, Barrett JC, Howat WJ. Proliferation Tumour Marker Network (PTM-NET) for the identification of tumour region in Ki67 stained breast cancer whole slide images. ***Sci Rep*** 2019; 9(1): 12845 DOI: 10.1038/s41598-019-49139-4

7 Loukas CG, Wilson GD, Vojnovic B, Linney A. An image analysis-based approach for automated counting of cancer cell nuclei in tissue sections. ***Cytometry A*** 2003; 55(1): 30-42 DOI: 10.1002/cyto.a.10060

8 Alturkistani HA, Tashkandi FM, Mohammedsaleh ZM. Histological Stains: A Literature Review and Case Study. ***Glob J Health Sci*** 2015; 8(3): 72-79 DOI: 10.5539/gjhs.v8n3p72

9 Puri M, Hoover SB, Hewitt SM, Wei BR, Adissu HA, Halsey CHC, Beck J, Bradley C, Cramer SD, Durham AC, Esplin DG, Frank C, Lyle LT, McGill LD, Sanchez MD, Schaffer PA, Traslavina RP, Buza E, Yang HH, Lee MP, Dwyer JE, Simpson RM. Automated Computational Detection, Quantitation, and Mapping of Mitosis in Whole-Slide Images for Clinically Actionable Surgical Pathology Decision Support. ***J Pathol Inform*** 2019; 10: 4 DOI: 10.4103/jpi.jpi_59_18

10 Matsukuma K, Olson KA, Gui D, Gandour-Edwards R, Li Y, Beckett L. Synaptophysin-Ki67 double stain: a novel technique that improves interobserver agreement in the grading of well-differentiated gastrointestinal neuroendocrine tumors. ***Mod Pathol*** 2017; 30(4): 620-629 DOI: 10.1038/modpathol.2016.225

11 Geread RS, Morreale P, Dony RD, Brouwer E, Wood GA, Androutsos D, Khademi A. IHC Color Histograms for Unsupervised Ki67 Proliferation Index Calculation. ***Front Bioeng Biotechnol*** 2019; 7: 226 DOI: 10.3389/fbioe.2019.00226

12 Tan WCC, Nerurkar SN, Cai HY, Ng HHM, Wu D, Wee YTF, Lim JCT, Yeong J, Lim TKH. Overview of multiplex immunohistochemistry/immunofluorescence techniques in the era of cancer immunotherapy. ***Cancer Commun (Lond)*** 2020; 40(4): 135-153 DOI: 10.1002/cac2.12023

13 Jurgensmeier JM, Eder JP, Herbst RS. New strategies in personalized medicine for solid tumors: molecular markers and clinical trial designs. ***Clin Cancer Res*** 2014; 20(17): 4425-4435 DOI: 10.1158/1078-0432.CCR-13-0753

14 van der Loos CM, de Boer OJ, Mackaaij C, Hoekstra LT, van Gulik TM, Verheij J. Accurate quantitation of Ki67-positive proliferating hepatocytes in rabbit liver by a multicolor immunohistochemical (IHC) approach analyzed with automated tissue and cell segmentation software. ***J Histochem Cytochem*** 2013; 61(1): 11-18 DOI: 10.1369/0022155412461154

15 Koopman T, Buikema HJ, Hollema H, de Bock GH, van der Vegt B. Digital image analysis of Ki67 proliferation index in breast cancer using virtual dual staining on whole tissue sections: clinical validation and inter-platform agreement. ***Breast Cancer Res Treat*** 2018;

169(1): 33-42 DOI: 10.1007/s10549-018-4669-2
16 Mejias-Badillo L, Jeanty J, Khalid K, Bhalla A, Salem N, Thomas S, Shidham VB. Dual-color immunocytochemistry (Ki-67 with LCA) for precise grading of pancreatic neuroendocrine tumors with applicability to small biopsies and cell blocks. ***Cytojournal*** 2020; 17: 6 DOI: 10.25259/Cytojournal_92_2019
17 Laurinavicius A, Laurinaviciene A, Dasevicius D, Elie N, Plancoulaine B, Bor C, Herlin P. Digital image analysis in pathology: benefits and obligation. ***Anal Cell Pathol (Amst)*** 2012; 35(2): 75-78 DOI: 10.3233/ACP-2011-0033
18 Y. M. Alomari, S. N. H. S. Abdullah, R. R. M. Zin, and K. Omar, Iterative randomized irregular circular algorithm for proliferation rate estimation in brain tumor Ki-67 histology images, ***Expert Syst. Appl.***, 2016; 48: 111 – 129 [DOI: 10.1016/j.eswa.2015.11.012]
19 S.Razavi et al., An Automated and Accurate Methodology to Assess Ki-67 Labeling Index of Immunohistochemical Staining Images of Breast Cancer Tissues, in International Conference on Systems, Signals, and Image Processing, 2018; 1-5 [DOI: 10.1109/IWSSIP.2018.8439184]
20 Sugita S, Hirano H, Hatanaka Y, Fujita H, Kubo T, Kikuchi N, Ito Y, Sugawara T, Segawa K, Hisai H, Yamashita K, Nobuoka T, Matsuno Y, Hasegawa T. Image analysis is an excellent tool for quantifying Ki-67 to predict the prognosis of gastrointestinal stromal tumor patients. ***Pathol Int*** 2018; 68(1): 7-11 [PMID: 29131458 DOI: 10.1111/pin.12611]
21 Saadeh H, Abdullah N, Erashdi M, Sughayer M, Al-Kadi O. Histopathologist-level quantification of Ki-67 immunoexpression in gastroenteropancreatic neuroendocrine tumors using semiautomated method. ***J Med Imaging (Bellingham)*** 2020; 7(1): 012704 DOI: 10.1117/1.JMI.7.1.012704
22 Barisoni L, Lafata KJ, Hewitt SM, Madabhushi A, Balis UGJ. Digital pathology and computational image analysis in nephropathology. ***Nat Rev Nephrol*** 2020; 16(11): 669-685 DOI: 10.1038/s41581-020-0321-6
23 S.Moore, Digital Pathology and Artificial Intelligence,News Medical-Life Sciences, 2020. [Online]. Available: https://www.news-medical.net/life-sciences/Digital-Pathology-and-Artificial-Intelligence.aspx. [Accessed: 19-Dec-2020].
24 M.Hild and B. Stemmer,Can Evolution Produce Robots?, in Consciousness and Cognition-Fragments of Mind and Brain, H. Cohen and B. Stemmer, Eds. ***San Diego, California: CRC Press***, 2007; 53 – 67 [DOI: 10.1016/B978-0-12-373734-2.X5000-6]
25 Wang X, Wang D, Yao Z, Xin B, Wang B, Lan C, Qin Y, Xu S, He D, Liu Y. Machine Learning Models for Multiparametric Glioma Grading With Quantitative Result Interpretations. ***Front Neurosci*** 2018; 12: 1046 DOI: 10.3389/fnins.2018.01046
26 Z. Swiderska, T. Markiewicz, B. Grala, and W. Kozłowski, Texture and Mathematical Morphology for Hot-Spot Detection in Whole Slide Images of Meningiomas and Oligodendrogliomas, in Conference: Computer Analysis of Images and Patterns 16th International Conference, ***CAIP*** 2015, 2015; 9257: 1 – 12 [DOI: 10.1007/978-3-319-23117-4]
27 Feng M, Deng Y, Yang L, Jing Q, Zhang Z, Xu L, Wei X, Zhou Y, Wu D, Xiang F, Wang Y, Bao J, Bu H. Automated quantitative analysis of Ki-67 staining and HE images recognition and registration based on whole tissue sections in breast carcinoma. ***Diagn Pathol*** 2020; 15(1): 65 DOI: 10.1186/s13000-020-00957-5
28 Liu Y, Li X, Zheng A, Zhu X, Liu S, Hu M, Luo Q, Liao H, Liu M, He Y, Chen Y. Predict Ki-67 Positive Cells in H&E-Stained Images Using Deep Learning Independently From IHC-Stained Images. ***Front Mol Biosci*** 2020; 7: 183 DOI: 10.3389/fmolb.2020.00183
29 Ertosun MG, Rubin DL. Automated Grading of Gliomas using Deep Learning in Digital Pathology Images: A modular approach with ensemble of convolutional neural networks. ***AMIA Annu Symp Proc*** 2015; 2015: 1899-1908 [PMID: 26958289 PMCID: PMC4765616]

30 F.Sheikhzadeh, M. Guillaud, Niekerk, D. van, and R. K. Ward, Automatic labeling of molecular biomarkers of immunohistochemistry images using fully convolutional networks, ***PLoS One***, 2016; 13(1): [PMID: 29351281 DOI: 10.17605/OSF.IO/C74WU]

