Mathematical Biosciences and Engineering

MBE, 19(2): 1721–1745.
DOI: 10.3934/mbe.2022081
Received: 15 October 2021
Accepted: 08 December 2021
Published: 15 December 2021

http://www.aimspress.com/journal/MBE

Research article

Automated knowledge-assisted mitosis cells detection framework in breast histopathology images

Xiao Jian Tan[1,*], Nazahah Mustafa[2,*], Mohd Yusoff Mashor[2] and Khairul Shakir Ab Rahman[3]

[1] Centre for Multimodal Signal Processing, Department of Electrical and Electronic Engineering, Faculty of Engineering and Technology, Tunku Abdul Rahman University College (TARUC), Jalan Genting Kelang, Setapak 53300, Kuala Lumpur, Malaysia
[2] Biomedical Electronic Engineering Programme, Faculty of Electronic Engineering Technology, Universiti Malaysia Perlis (UniMAP) 02600 Arau, Perlis, Malaysia
[3] Department of Pathology, Hospital Tuanku Fauziah 01000 Jalan Tun Abdul Razak Kangar Perlis, Malaysia

* **Correspondence:** Email: tanxj@tarc.edu.my, nazahah@unimap.edu.my; Tel: +(6)0185701012, +(6)049885281.

Abstract: Based on the Nottingham Histopathology Grading (NHG) system, mitosis cells detection is one of the important criteria to determine the grade of breast carcinoma. Mitosis cells detection is a challenging task due to the heterogeneous microenvironment of breast histopathology images. Recognition of complex and inconsistent objects in the medical images could be achieved by incorporating domain knowledge in the field of interest. In this study, the strategies of the histopathologist and domain knowledge approach were used to guide the development of the image processing framework for automated mitosis cells detection in breast histopathology images. The detection framework starts with color normalization and hyperchromatic nucleus segmentation. Then, a knowledge-assisted false positive reduction method is proposed to eliminate the false positive (i.e., non-mitosis cells). This stage aims to minimize the percentage of false positive and thus increase the F1-score. Next, features extraction was performed. The mitosis candidates were classified using a Support Vector Machine (SVM) classifier. For evaluation purposes, the knowledge-assisted detection framework was tested using two datasets: a custom dataset and a publicly available dataset (i.e., MITOS dataset). The proposed knowledge-assisted false positive reduction method was found promising by eliminating at least 87.1% of false positive in both the dataset producing promising results in the F1-score. Experimental results demonstrate that the knowledge-assisted detection framework can achieve promising results in F1-score (custom dataset: 89.1%; MITOS dataset: 88.9%)

and outperforms the recent works.

Keywords: breast carcinoma; mitosis cells detection; segmentation; handcrafted feature-based approach; knowledge-assisted; domain knowledge; false positive reduction; histopathology image

1. Introduction

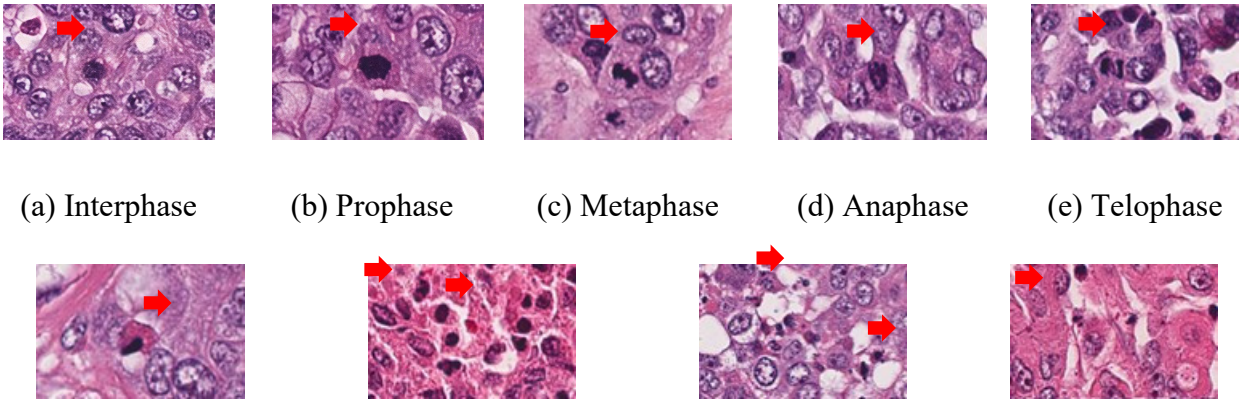
Mitosis counting on Hematoxylin and Eosin (H & E) [1,2] stained breast histopathology slides is one of the important prognostic markers that indicate the proliferative activity of breast carcinoma. The Nottingham Histopathology Grading (NHG) is a scoring system to assess the grade of breast carcinoma [2]. The NHG system was developed based on the following three morphological criteria: mitosis count, degree of tubule formation, and nuclear pleomorphism [2]. Each criterion of the NHG system is scored from 1 to 3 and the summation of these scores provides the grade. The score of breast carcinoma is determined manually by the histopathologist. There are three grades for breast carcinoma: scores 3–5 for Grade 1; scores 6 and 7 for Grade 2; scores 8 and 9 for Grade 3. Distinct prognosis and treatment planning are developed according to the obtained grade.

Conventional assessment of mitosis count is done through visual examination method under a light microscope. For each histopathology slide, multiple images/frame (approximate 2000 frames) is created corresponding to different areas on a slide [3]. The slide is examined at a high-power field (HPF) under 40x magnifications. As the histopathology slides are complex and heterogeneous [4–6], visual examination requires the slide to be examined frame by frame. During the slide examination, the histopathologist needs to identify a defined area express in mm^2 with the highest mitosis activity to provide scores based on the NHG system [2]. In many hospitals, histopathologist typically handles around 100 grading cases per day [7]. Thus, conventional visual examination of breast carcinoma tissues is cumbersome and tedious [1,8]. During the past few decades, the emergence of Whole-Slide-Imaging (WSI) scanner has made a giant leap in the pathology laboratory [9–11]. This provides a significant benefit over the conventional histopathology slides. WSI scanner takes the analogue histopathology slides as input and produces digital images as output.

Mitosis detection is very challenging [12]. Most of the challenges arise from the variability of the mitosis cell appearance and the imperfection of H & E staining [13,14]. Mitosis cell is formed through a process known as mitosis [15,16]. Mitosis is a process where the chromosome is replicated and separated into two new nuclei. This process involves five phases [16]: interphase, prophase, metaphase, anaphase, and telophase (Figure 1, top row). The appearance of the mitosis cell varies from a round shape with rough texture (interphase and prophase) into a narrow oval shape with hyperchromatic properties (metaphase). It then becomes a separate-like nuclei with rough texture and hyperchromatic properties (anaphase and telophase) [16]. Apart from this, there are few artefacts such as karyorrhectic and apoptotic nuclei that have similar appearance with the true mitosis cells (Figure 1, bottom row). Imperfection in the H & E staining, variability of the mitosis cells and artefacts could influence the mitosis count and breast carcinoma grading.

In recent years, the research interest in mitosis detection is mainly driven by the open challenges and availability of public access datasets. These datasets allow the researchers to test and validate the respective methods and benchmark the results with other re-searchers across the globe. To date, the MITOS [17], AMIDA [18], MITOS-ATYPIA [19], and TUPAC [20] are four public access datasets,

specifically for mitosis detection in breast carcinoma. Table 1 summarizes the public access datasets for mitosis cells detection.



(a) Interphase (b) Prophase (c) Metaphase (d) Anaphase (e) Telophase

(f) Karyorrhectic (g) Apoptotic nuclei

Figure 1. The top row shows samples of mitosis cells in different phases, whereas the bottom row shows the artefacts with similar appearances to the mitosis cells.

Table 1. Public access datasets for mitosis cells detection.

Datasets	Year	Resolutions	Magnification	Number of images	Number of mitosis cells
MITOS [17]	2012	Aperio scanner: 2084 × 2084 Hamamatsu scanner: 2250 × 2252	40 x	50	327
AMIDA [18]	2013	Multispectral images: 1360 × 1360 2000 × 2000	40 x	311	550
MITOS-ATYPIA [19]	2014	Aperio scanner: 1539 × 1376 Hamamatsu scanner: 1663 × 1485	40 x	1200	749
TUPAC [20]	2016	2000 × 2000	40 x	321	Unlabeled

The recent works in mitosis cells detection can be characterized into three different approaches: handcrafted feature-based approach, deep learning approach, and fusion approach. In mitosis cells detection, the unbalanced data between the mitosis cells and non-mitosis cells are significant. Therefore, to ensure a fair comparison, results benchmarking is usually performed by observing the percentage of F1-score obtained from a proposed method. Tables 2–4 show the recent works in mitosis detection using the public access datasets with handcrafted feature-based, fusion, and deep learning approaches, respectively.

In handcrafted feature-based approaches, Irshad et al. proposed a mitosis cells detection framework that implements multispectral band selection and spatial characterization [21]. The selected spectral bands are meant to reduce the computation load and storage complexity, whereas the multi-features are meant to discriminate between the mitosis cells and non-mitosis cells. Lu and Mandal proposed a multi-stage mitosis cells detection framework that involved: discriminative image generation stage, mitosis candidate detection, and segmentation stage, followed by the classification stage [22]. Tashk et al. performed mitosis cells detection by using a combination of different features: Completed Local Binary Pattern, Statistical Moment Entropy, and Stiffness Matrix which includes

textural, statistical, and innovative mathematical features [23]. Paul and Mukherjee proposed a mitosis cells detection framework that is driven by domain knowledge in mitosis cells [24]. Based on the fact that the intensity pattern of the mitosis remains similar in breast histopathology images, mitosis cells detection can be done by implementing the Relative Entropy Maximized Scale Space using area morphological opening and closing approaches. Nateghi et al. proposed a mitosis cells detection framework that implemented a novel Maximized Inter-class Weighted Mean method to reduce the false positive in mitosis cells candidates selection stage [25]. Color, textural, and shapes features are then extracted from the mitosis cells candidates and are used as input for the classifier.

In deep learning approaches, Li et al. proposed a novel multi-stage deep learning framework [8]. The deep segmentation framework is able to detect the mitosis cells regions by using a weak label such as the centroid pixel of the mitosis cells. Also, the deep learning framework is facilitated using contextual region information and a deep verification network. The proposed method is termed as DeepMitosis. Cai et al. proposed an effective mitosis cells detection framework using a modified regional convolutional neural network (RCNN) [26]. The proposed RCNN implemented the Resnet-101 network as the backbone feature extractor and is pre-trained on the ImageNet dataset. The output channels are then reduced using a large separable convolution. This proposed method is found promising in reducing the computation time in both training and inference stages. Alom et al. proposed an end-to-end mitosis cells detection framework, namely MitosisNet [27]. The proposed method constitutes of segmentation, detection, and classification stages. The output from segmentation and detection stages are used as mitosis cells reference regions while further confirmation is performed using the classification stage. Sebai et al. proposed a multi-task deep learning framework for mitosis cells detection and instance segmentation using Mask RCNN which is termed as MaskMitosis [10]. The MaskMitosis is meant to perform multiple tasks such as mitosis mask estimation, bounding box labelling on weakly annotated and unannotated datasets, mitosis cells localization, and mitosis cells classification using the estimated pixel-level annotations.

In fusion approaches, Wang et al. proposed a mitosis detection framework that combines the convolution neural network (CNN) and handcrafted features such as morphological, textural, and color features [28]. The proposed fusion approach was found promising in reducing the computation load while achieving high accuracy. Beevi et al. proposed a multi-classifier system for mitosis cells detection using the Deep Belief Network [29]. The proposed system combined methods such as the Localized Active Contour Model and Krill Herd Algorithm for segmentation purposes and a multi classifier system to improve the prediction accuracy. Das and Dutta proposed a deep CNN with wavelet decomposed images for mitosis detection [30]. The input images are decomposed into smaller images using Haar wavelet. These images are then used to develop a deep CNN for automatic mitosis cells detection. Mahmood et al. proposed a multi-stage mitosis cells detection framework based on a faster region convolutional neural network (Faster R-CNN) and deep CNNs [31]. The proposed framework included a post-processing stage that aims to reduce the false positive using handcrafted features.

Overall, the recent works aforementioned benefit from applying similar image processing methodologies for mitosis detection. These methodologies included image pre-processing, mitosis cells candidate estimation and segmentation, feature extraction, feature selection, and classification. There are two known challenges in mitosis detection: the number of mitosis cells is very low as compared to non-mitosis cells and the mitosis can present in various shapes and morphologies (refer to the top row in Figure 1). A plethora of non-mitosis cells are usually found after the segmentation stage. This contributes to a high number of false positives and impinges the overall detection accuracy.

In addition, a large number of non-mitosis cells can significantly increase the computation time of the proposed framework. Several recent works highlight and propose solutions to this issue. Khan et al. proposed a gamma-gaussian mixture model to reduce the number of non-mitosis objects after the segmentation stage [36]. A more recent work by Nateghi et al. proposed a false positive mitosis candidate reduction using a novel Maximized Inter-Class Weighted Mean method [25].

Table 2. Recent works in mitosis cell detection using handcrafted feature-based approach.

Authors	Year	Datasets	Methods	F1-score (%)
Sommer et al. [32]	2012	MITOS [17]	Mitosis detection using hierarchical learning workflow	62.9
Irshad [33]	2013		Mitosis detection using the multi-channel features computation framework where the nuclei feature extraction is performed in selected channels of color spaces at a fixed image scale	63.0
Tek [34]	2013		Mitosis detection via application of generic features and an ensemble of cascade adaboosts	58.0
Irshad et al. [35]	2013		Mitosis detection using texture, sift features, and HMAX biologically inspired approach	76.0
Tashk et al. [25]	2013		Mitosis detection using objective and pixel-wise textural features	70.1
Khan et al. [36]	2013		Mitosis detection using Gamma-Gaussian mixture model	75.4
Irshad et al. [21]	2014		Mitosis detection via the selection of spectral bands and focal plane	74.0
Nateghi et al. [37]	2014		Mitosis detection using genetic algorithm	78.4
Irshad et al. [21]	2014		Mitosis detection using multispectral band selection and spatial characterization	57.0
Lu and Mandal [22]	2014		Mitosis detection using three main stages: discriminative image generation, mitosis cell candidate detection and segmentation, and followed by a classification	47.9
Tashk et al. [23]	2015		Mitosis detection using a combination of textural, statistical, and innovative mathematical features	82.7
Nateghi et al. [25]	2017		Mitosis detection using Maximized Inter-Class Weighted Mean method	88.4
Nateghi et al. [25]	2017	AMIDA [18]	Mitosis detection using Maximized Inter-Class Weighted Mean method	75.3
Paul and Mukherjee [24]	2015	MITOS-ATYPIA [19]	Mitosis detection by implementing the Relative Entropy Maximized Scale Space using area morphological opening and closing	73.4
Nateghi et al. [25]	2017		Mitosis detection using Maximized Inter-Class Weighted Mean method	83.7

Recognition of complex and inconsistent objects specifically in the medical histopathology images could be achieved by incorporating domain knowledge in the field of interest [24,47,48]. There are studies obtained encouraging results when incorporating domain knowledge, image processing, and classification methods [24,48–53], but the number is very few. During manual vision inspection, histopathologists do not check the presence of mitosis on every cell in each frame under a light

microscope. Histopathologists scan through the frames using the ball rolling method and scrutinize cells with characteristics (e.g., hyperchromatic properties) that are specifically demonstrated by mitosis cells. In this study, a more promising handcrafted feature-based approach mitosis detection framework is proposed. The proposed framework considered both the challenges highlighted herein. The proposed mitosis detection framework is driven by domain knowledge in mitosis such that the domain knowledge and strategies of the histopathologist (human expert) were utilized to guide the development of the image processing framework for mitosis detection. Additionally, the proposed framework clusters the segmented cells into different groups based on the morphological structures. This aims to triage and eliminate false positives (i.e., non-mitosis cells) in different phases (i.e., interphase, prophase, metaphase, anaphase, and telophase). The main novelty of the proposed detection framework lies within the simple yet powerful detection capability of the framework with knowledge transferred from human experts, producing promising detection results, reflected with high percentages in F1-score. To the best of our knowledge, mitosis cells detection using a knowledge-assisted detection framework by incorporating domain knowledge as well as knowledge transfer from human experts is yet to be available in the literature.

The objectives of this study are: 1) to remove and minimize the segmented non-mitosis cells (false positive); 2) to provide a mitosis cells detection framework that can serve as the second opinion for histopathologists to facilitate the mitosis cells assessment procedure; 3) to validate the applicability of the proposed detection framework by benchmarking the output of the proposed detection framework with ground truth and other recent works.

Table 3. Recent works in mitosis cell detection using fusion approach.

Authors	Year	Datasets	Methods	F1-score (%)
Malon and Cosatto [46]	2013	MITOS [17]	Mitosis detection using the convolutional neural networks and seeded blob features	55.7
Wang et al. [28]	2014		Mitosis detection using a combination of handcrafted and convolutional neural network features	73.4
Mahmood et al. [31]	2020		Mitosis detection using a multistage mitosis-cell-detection method based on faster region convolutional neural network and deep convolutional neural network	85.8
Beevi et al. [29]	2017	MITOS-ATYPIA [19]	Mitosis detection using a multi-classifier system	76.7
Das and Dutta [30]	2019		Mitosis detection using the deep convolution neural network with wavelet decomposed patches	55.9
Saha et al. [12]	2018		Mitosis detection using a supervised model of deep learning architecture with handcrafted features.	90.0
Mahmood et al. [31]	2020		Mitosis detection using a multistage mitosis-cell-detection method based on faster region convolutional neural network and deep convolutional neural network	69.1

This paper is organized as follows: Section 2 details the methodology of the proposed mitosis detection framework; Section 3 describes the datasets used in this study; Section 4 presents the experimental results and analysis obtained from the proposed framework. In this section, the proposed

framework is benchmarked and compared with ground truth annotated by the histopathologist (using the custom dataset) and recent works (using the MITOS dataset [17]) for validation purposes. In Section 5, the conclusion of this study is presented.

Table 4. Recent works in mitosis cell detection using deep learning approach.

Authors	Year	Datasets	Methods	F1-score (%)
Ciresan et al. [38]	2013	MITOS [17]	Mitosis detection using the deep max-pooling convolutional neural networks	78.2
Chen et al. [39]	2016		Mitosis detection using the deep regression network	79.0
Wahab et al. [40]	2017		Mitosis detection using the two-phase deep convolutional neural network	79.0
Alom et al. [27]	2020		Mitosis detection via an end-to-end multi-task learning system from pathological images which is termed as "MitosisNet"	87.8
Sebai et al. [10]	2020		Mitosis detection using multitask deep learning framework for object detection and instance segmentation Mask RCNN	86.3
Xi et al. [41]	2020		Mitosis detection using the cascaded convolutional neural network (CNN) based on UNet.	83.6
Albarqouni et al. [42]	2016	AMIDA [18]	Mitosis detection using the convolutional neural network (CNN) via additional crowdsourcing layer (AggNet)	74.2
Wollmann and Rohr [43]	2017		Mitosis detection using deep residual Hough voting	60.9
Romo-Bucheli et al. [44]	2019		Mitosis detection using deep learning strategy with gene expression derived risk categories in estrogen receptor-positive breast carcinomas	55.6
Chen et al. [45]	2016	MITOS-ATYPIA [19]	Mitosis detection using the deep cascaded convolutional neural network	48.2
Li et al. [8]	2018		Mitosis detection using the multi-stage deep learning framework	43.7
Cai et al. [26]	2019		Mitosis detection using the modified regional convolutional neural network	58.5
Alom et al. [27]	2020		Mitosis detection via an end-to-end multi-task learning system from pathological images which is termed as "MitosisNet"	75.9
Sebai et al. [10]	2020		Mitosis detection using multitask deep learning framework for object detection and instance segmentation Mask RCNN	60.8
Xi et al. [41]	2020		Mitosis detection using the cascaded convolutional neural network (CNN) based on UNet.	57.1
Wahab et al. [40]	2017	TUPAC [20]	Mitosis detection using the two-phase deep convolutional neural network	55.0
Cai et al. [26]	2019		Mitosis detection using the modified regional convolutional neural network	73.6

2. Methods

The proposed mitosis cells detection framework consists of five main stages: 1) color normalization using the optimal stain vector approach [54]; 2) hyperchromatic nucleus segmentation performed using the knowledge-assisted K-Mean; 3) the proposed knowledge-assisted false positive reduction stage based on the domain knowledge in mitosis cells; 4) features extraction on morphological and textural features; 5) classification using the Support Vector Machine (SVM) with radial-based function (RBF) kernel. Figure 2 shows the block diagram of the proposed knowledge-assisted mitosis cells detection framework. The following subsections detail each stage in the proposed detection framework.

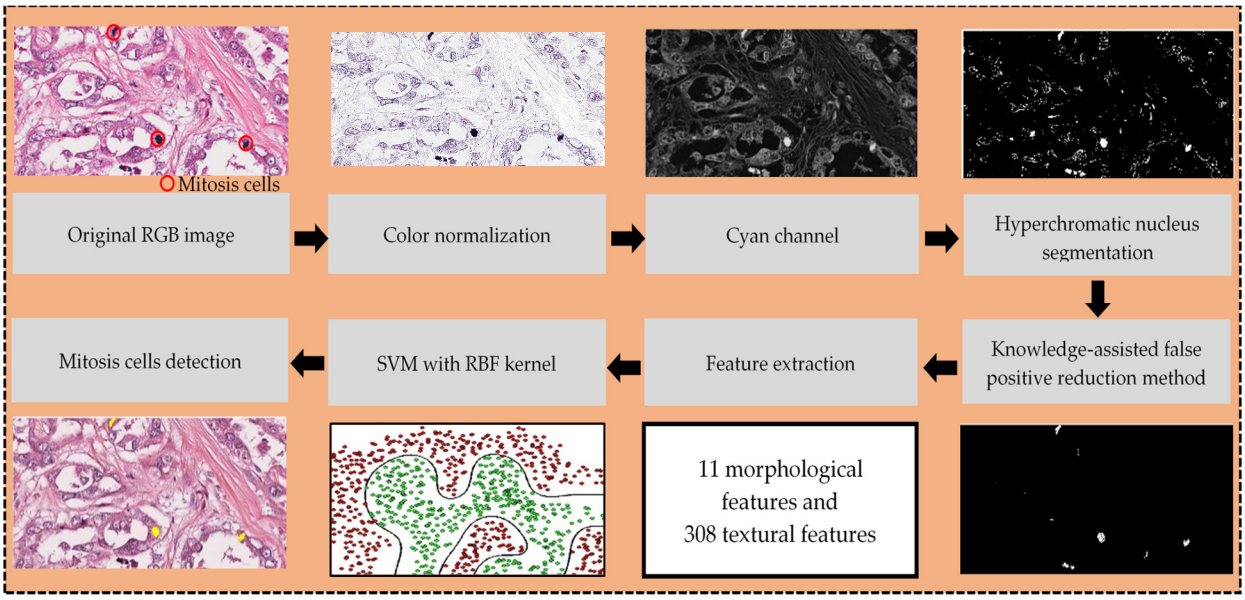
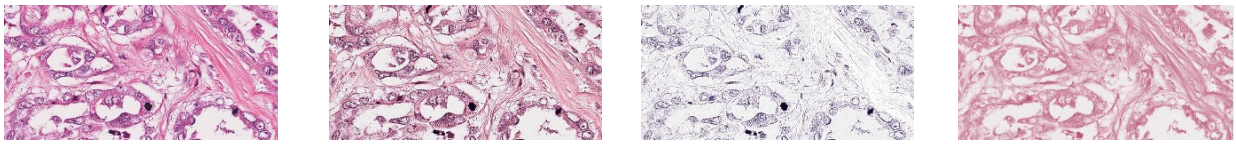


Figure 2. Block diagram of the proposed knowledge-assisted mitosis cells detection framework.

2.1. Color normalization

Color variation in H & E-stained breast histopathology images is a common issue that arises from the use of different stains/dyes reactivity and coloring from a different manufacturer, use of different WSI scanner, and use of different equipment. To ensure excellent stability in the performance of the proposed detection framework, color normalization shall be performed to minimize color variation amongst the input images in any datasets. In this study, the optimal stain vector approach [54] was implemented for color normalization purposes. Briefly, the input images were first transformed to the optical density domain. The singular value decomposition was then calculated from the optical density tuples. The optical density was then transformed and projected onto the plane and was normalized to the unit length. Next, calculate the angle of each point with respect to the first singular value decomposition direction. Mapping of direction was then performed. All the extreme values were then converted back to the optical density domain. The optimal stain vector approach formalized three outputs: a color normalized image, a Hematoxylin-stained image, and an Eosin-stained image. In this study, only the Hematoxylin-stained image is used as the mitosis cells captured only the Hematoxylin stain in the standard staining procedure.

Figure 3 shows a sample of outputs obtained from the color normalization stage.



(a) Input image (b) Color normalized image (c) Hematoxylin-stained image (d) Eosin-stained image

Figure 3. A sample of outputs obtained from the color normalization stage.

2.2. Hyperchromatic nucleus segmentation

By nature, the chromosomal condensation that occurs during mitosis resulted in a higher degree of Hematoxylin stain absorption during the staining procedure. This causes the mitosis cells to become remarkable (i.e., darker in color) than the normal cells and is prominent in the Hematoxylin-stained image (Figure 3(c)). Thus, it is reasonable to conclude that the nucleus with low intensity is closely related to the mitosis cells. In this study, the nucleus with remarkable intensity values (i.e., darker in color) as compared to the normal nucleus in the Hematoxylin-stained image is termed as the hyperchromatic nucleus.

To perform hyperchromatic nucleus segmentation, the Hematoxylin-stained image was first converted to the cyan channel (determined heuristically). The number of clusters for K-Mean was set as three, such that cluster 1 for hyperchromatic nucleus (white objects), cluster 2 for tumor regions (grey regions), and cluster 3 for background (black regions). In hyperchromatic nucleus segmentation, only information from cluster 1 (i.e., hyperchromatic nucleus) is relevant to this study. Nevertheless, it is important to assign two different clusters for tumor regions and background. If only two clusters were assigned in this stage (i.e., cluster 1 for hyperchromatic nucleus and cluster 2 for tumor regions and back-ground), the background (i.e., unstained areas that appeared in black color in the cyan channel) would act as an outlier and impinge the clustering accuracy for both the clusters. For knowledge-assisted K-Mean, the initial centroids are not generated randomly but based on domain knowledge. Since the hyperchromatic nucleus, tumor regions, and background respectively appeared in white, grey, and black pixel ranges, the initial centroids for each cluster can be selected by simply partitioning the gray level (i.e., 255) into three regions, such that region 1 ranges from 171 to 255, region 2 ranges from 86 to 170, and region 3 ranges from 0 to 85. Then, the median intensity of each region is selected as the initial centroid.

2.3. Knowledge-assisted false positive reduction method

The segmented hyperchromatic nucleus images contain mitosis and non-mitosis cells. Mitosis cells in different phases demonstrate different morphological features specifically in shape (see Figure 1, top row). The knowledge-assisted false positive reduction method was mainly based on domain knowledge in mitosis cells and is formalized through empirical analysis. To support the empirical analysis, Monte Carlo simulation [55,56] was implemented to determine the number of mitosis cells to be used in this stage. Parameters such as area, mean intensity, aspect ratio (AR), form factor (FF), and filled area (FA) are observed as random variables. The Cochran formula (refer to Eq (1)) is implemented to calculate the

sampling data at a confidence level of 95.0%, provided the z value is at 1.96 (obtained from the z table).

$$n_0 = \frac{z^2 pq}{e^2} \tag{1}$$

where z denotes the z value, p (set as 0.5 in this study) denotes the estimated probability a mitosis cell is found; q denotes the probability of 1-p; and e denotes the margin of error. From Eq (1), sampling data of 385 is required in order to obtain a confidence level of 95.0%. Thus, in this stage, the empirical analysis was performed using 385 mitosis cells. It is important to remark that all the 385 mitosis cells used for empirical analysis are first undergone color normalization as described in Section 2.1. As the mitosis cells demonstrate different morphological features, typically in shape, the mitosis cells can be grouped into four shapes (Figure 4): Large Circular-Shaped (LCS) nucleus, Ellipse-Shaped (ES) nucleus, Hollow-Shaped (HS) nucleus, and Curl-Shaped (CS) nucleus. Domain knowledge of mitosis cells and strategies used by the histopathologist were incorporated in this stage (i.e., Checking procedures 1 and 2, as in Figure 5) to reduce the number of false positive. The steps in the knowledge-assisted false positive reduction stage are shown in Figure 5.

(a) LCS nucleus (b) ES nucleus (c) HS nucleus (d) CS nucleus

Figure 4. Four different shapes of nucleus. In images (a) to (d), the left image shows the RGB image of nucleus and the right image shows the segmentation results obtained using the knowledge-assisted K-Mean.

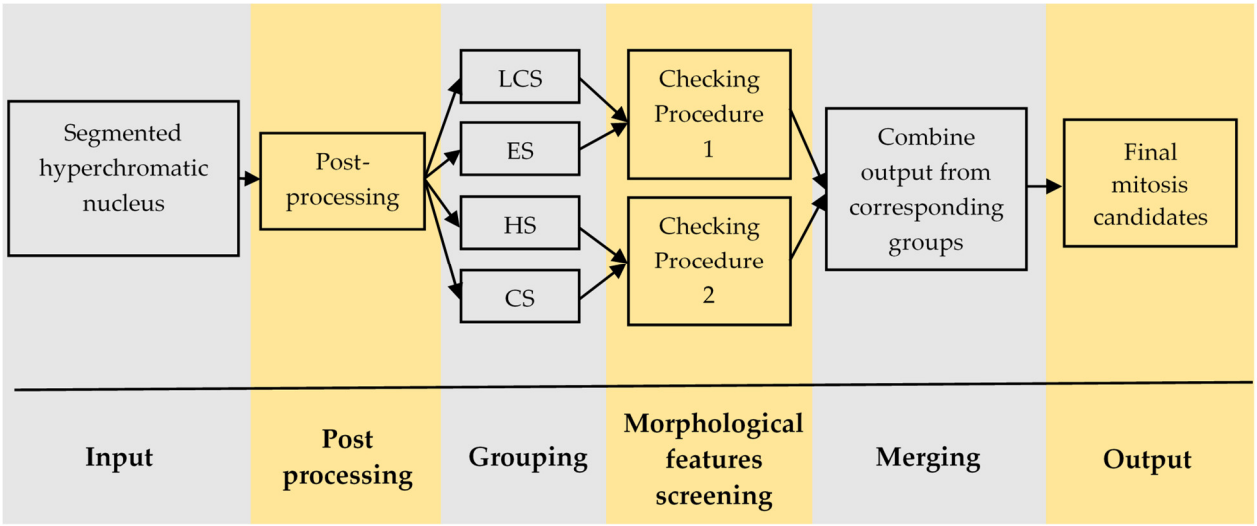


Figure 5. Block diagram showing processes in the knowledge-assisted false positive reduction stage.

2.3.1 Post-processing

Post-processing was first implemented to eliminate the small objects, typically noise and/ or non-mitosis cells. Manual segmentation was performed on the 385 mitosis cells to obtain the area in pixel. The mean area (± standard deviation (SD)) for the mitosis cells was 312 (± 76) pixels. The smallest

area of the mitosis cell is 118 pixels. To avoid over-elimination, in this step, only objects with areas less than half of this smallest size (i.e., 59) were eliminated from the images and considered as artefacts. Table 5 shows the results of the empirical analysis for the area.

Table 5. Results of empirical analysis for the area in 385 mitosis cells.

Feature	Parameters		
	Mean (± SD)	Maximum	Minimum
Area	312 (± 76)	564	118

2.3.2 Grouping

As aforementioned, the knowledge-assisted false positive reduction stage is meant to triage and eliminate false positive in different phases (i.e., interphase, prophase, metaphase, anaphase, and telophase). Thus, the outputs from the previous step (i.e., post-processing step) are characterized into four different groups to facilitate the morphological features screening step in Section 2.3.3. The four groups were LCS, ES, HS, and CS groups. These groups are derived from the morphological features of the mitosis cells from different phases (i.e., interphase, prophase, metaphase, anaphase, and telophase) as in Figure 1, top row, such that the LCS group aims to remove false positive in interphase and prophase, ES group in metaphase; HS group in anaphase, and CS group in telophase.

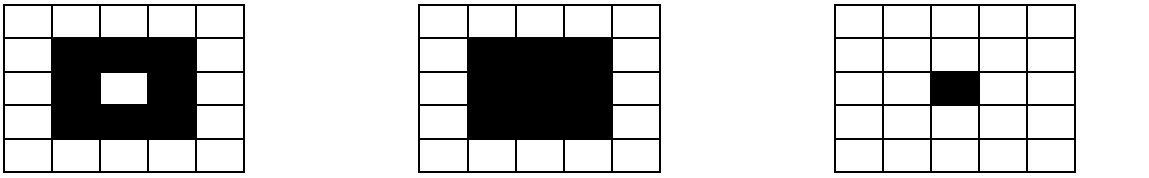
Grouping was performed based on the following shape parameters: AR, FF, and FA. AR is a ratio of the major axis length to the minor axis length (Eq (2)). A high AR value (e.g., > 2.0) relates to an ellipse shape. When AR approaching 1.0, it indicates a circular shape. FF is a shape descriptor (Eq (3)). It is used to indicate the degree of similarity to a perfect circle [57]. FF value of 1.0 refers to a perfect circle and the object becomes less circular when approaching 0.0 [57]. FA is used to detect a hollow shape (Eq (4)). A high value of FA indicates an object with a cavity inside whereas a small value of FA refers to a solid object. Figure 6 illustrates how FA can be obtained. It is important to remark that in any dataset, the number of mitosis cells is very limited. Hence, it is very challenging to perform empirical analysis using 385 mitosis cells (at a confidence level of 95.0%) for each group. Therefore, the 385 mitosis cells were first manually examined and characterized into LCS, ES, HS, and CS groups. Based on the manual examination, from the 385 mitosis cells, there are 134, 91, 82, and 78 mitosis cells in LCS, ES, HS, and CS groups, respectively. The shape parameters (i.e., AR, FF, and FA) were then extracted from the mitosis cells in each group. Table 6 shows the parameter values of the mitosis cells in each group.

$$AR = \frac{Major\ axis\ length\ in\ pixel}{Minor\ axis\ length\ in\ pixel} \tag{2}$$

$$FF = \frac{4\pi * area}{perimeter^2} \tag{3}$$

$$FA = x - y \tag{4}$$

where x denotes the enclosed area of the nucleus and y denotes exact area of the nucleus.



(a) The exact area of the nucleus (y) (b) The enclosed area of the nucleus (x) (c) FA obtained from $x - y$

Figure 6. Illustration on how to obtain FA.

Table 6. Results of empirical analysis for shape parameters on the mitosis cells in each group.

Groups	Shape parameters								
	AR			FF			FA		
	Mean	Max	Min	Mean	Max	Min	Mean	Max	Min
LCS	1.5	1.9	1.2	0.6	0.9	0.4	-	-	-
ES	3.1	4.5	2.6	-	-	-	-	-	-
HS	-	-	-	-	-	-	41	57	28
CS	1.6	1.8	1.4	0.2	0.3	0.1	-	-	-

2.3.3 Morphological features screening

Based on the parameter values in Table 6, conditions for each group were set as shown in Table 7. The segmented hyperchromatic nuclei were put into any of these groups based on the conditions. The shape parameters of the segmented hyperchromatic nuclei were first checked to see if they fit the condition for the HS group. This is then followed by checking the conditions for the ES group, LCS group, and CS group. The mitosis candidates in each group then underwent two checking procedures: Checking Procedures 1 and 2. These procedures were used to eliminate the false positive in each group. Checking Procedure 1 was implemented on the LCS and ES groups, whereas Checking Procedure 2 was implemented on the HS and CS groups. In Checking Procedure 1, mitosis candidates with an area smaller than 118 pixels (obtained in Table 5) were recognized as false positive and eliminated from the image. In Checking Procedure 2 (Figure 7), the minor axis length (in pixel) ($minor_{axis}$) of each nucleus was obtained and was used to develop a square window, such that window = [$minor_{axis}$ $minor_{axis}$]. The window was then superimposed and placed on the centroid of the respective nucleus. The mean intensity within the window patch was calculated. Checking Procedure 2 eliminates candidates with mean intensity lower than 0.23 (obtained in Table 8). This value is obtained from the empirical analysis of 385 mitosis cells. Table 8 shows the results of empirical analysis in terms of mean intensity calculated using the aforementioned square window.

Table 7. Setting conditions of shape parameters for four different shape groups.

Groups	Shape parameters		
	AR	FF	FA
LCS	≤1.9	≥0.4	-
ES	>1.9	-	-
HS	-	-	>28 pixels
CS	≤1.9	<0.4	-

Table 8. Results of empirical analysis for mean intensity in 385 mitosis cells.

Feature	Parameters Mean (± SD)	Maximum	Minimum
Mean intensity (calculated using square window, where $window = [minor_{axis}\ minor_{axis}]$)	0.53 (± 0.113)	0.76	0.23

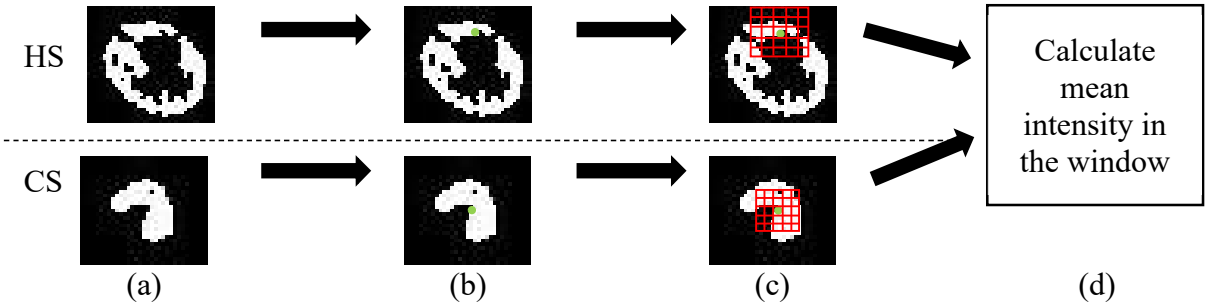


Figure 7. Illustration on Checking Procedure 2 for HS and CS groups. The top row images show the nucleus from HS group and the bottom row images show the nucleus from CS group. (a) The sample nucleus, (b) the centroid of the corresponding nucleus is shown by the green dot, (c) a window patch with the size of minor axis length (pixel) of the corresponding nucleus was developed and (d) the mean intensity in the corresponding window is calculated.

2.4 Features extraction

Morphological feature describes the form, shape and the structure of the mitosis cell. Morphological features such as eccentricity, extent, perimeter, convex area, area, equivalent diameter, AR, FF, FA, solidity, and ratio of perimeter to area were extracted from the mitosis candidates. Texture features hold a significant characteristic in describing objects or region of interest in an image. The condensation of chromosomes during mitosis causes the cell to show random but consistent texture properties throughout the five phases. Gray-Level Co-occurrence Matrix (GLCM) was applied to examine properties of the mitosis candidates. GLCM with radius, δ = 1 pixel and orientations (θ) of 0°, 45°, 90°, 135° were used. δ of 1 pixel was used as it could provide the best result in examining the texture properties [58]. As the mitosis cell is not an isotropic object (i.e., the directional information of mitosis cell is important), thus, it is essential to obtain GLCM at four different orientations [59]. 11 texture features were computed from the GLCM at four distinct orientations in different channels: entropy, energy, inertia, inverse different moment, correlation, sum average, sum variance, difference average, difference variance, skewness and kurtosis. Seven color channels were used to generate the co-occurrence matrix: red, green, blue, hue, Cr, Cb, and C channels. A total of 319 features (11 morphological features and 308 texture features) were extracted from the mitosis candidates and used as inputs to the SVM classifier.

2.5 Classification

SVM was selected to classify the extracted features [60]. SVM is optimized based on risk minimization [60]. The optimization process is defined by a function in order to find the parameters of a hyperplane that minimizes the risk. For a non-linear case, a kernel-based decision function can be applied [61]. By introducing kernel, SVM will gain its flexibility during the choice in the form of threshold separating the classes. As the kernel implicitly contains a non-linear transformation, no assumption about the functional form of transformation, which makes the data linearly separable, is needed. This study used RBF as the kernel-based decision function [61]. In this study, the parameters used in the SVM classifier were the default settings. Some of the default settings for the SVM classifier can be found in Appendix A (Table A).

3 Datasets

To evaluate the proposed knowledge-assisted detection framework, two datasets were used: a custom dataset and a publicly available dataset (i.e., MITOS dataset [17]). The custom dataset is a self-collected dataset locally in Malaysia. This dataset consists of 48 breast histopathology slides and was obtained from the Pathology Department, Hospital Tuanku Fauziah, Kangar, Perlis, Malaysia. These slides were prepared under a standard staining procedure known as H & E stains. An Aperio CS2 WSI scanner was used to convert the slides into digital form. 20 images were captured at 40x magnification from different dominant areas in each digital slide. The captured image is in 8-bit RGB color with dimensions of 614 × 1264 pixels (size of pixel: 0.2521 μm per pixel) and is in tiff file format. The custom dataset consists of 960 images with a total of 832 mitosis cells. Another dataset, namely MITOS dataset [17] was used in this study. MITOS dataset [17] is a publicly available dataset that is widely used to evaluate the superiority of a mitosis detection framework. Images in MITOS dataset [17] were prepared using the standard H & E stains and were captured at 40 x magnification using Aperio XT WSI scanner, Hamamatsu NanoZoomer WSI scanner, and multi-spectral microscope. In this study, the dataset from the Aperio XT WSI scanner was used for evaluation purposes. The captured image is in 8-bit RGB color with dimensions of 2084 × 2084 pixels (size of pixel: 0.2456 μm per pixel). This dataset consists of 50 images with a total of 327 mitosis cells. Table 9 summarizes the dataset used in this study.

For the custom dataset, the ground truth (i.e., mitosis cells) annotation was performed by an experienced histopathologist from the Hospital Tuanku Fauziah, Kangar, Perlis, Malaysia. Each image was visually examined and the mitosis cells were annotated care-fully in accordance with the NHG system. The main purpose of using the custom dataset is to evaluate the applicability of the proposed knowledge-assisted detection framework and to determine the degree of agreement with respect to the ground truth. The main purpose of using the publicly available dataset (i.e., MITOS dataset [17]) is to perform benchmarking with recent works in order to justify the superiority of the proposed knowledge-assisted detection framework.

Table 9. The summary of dataset used in this study.

Datasets	Number of images	Number of mitosis cells	Magnification	Resolution
Custom	960	832	40 x	614 × 1264
MITOS [17]	50	327	40 x	2084 × 2084

4 Results and analysis

To evaluate the applicability of the proposed knowledge-assisted detection framework, performance evaluation was done in four main aspects: initialization method, false positive reduction method, capability in mitosis cells segmentation (after implementation of false positive reduction method), and overall performance in mitosis cells detection. The subsequent sub-sections show the performance evaluation in detail.

4.1 Initialization method performance evaluation

To evaluate the applicability of the improved initialization method, the knowledge-assisted K-Mean method is compared to the baseline initialization method (i.e., random method [62]) and some other popular methods (i.e., Forgy method [63], K-Mean++ [64], and Greedy K-Mean++ [65]). The main purposes of this comparison are to: 1) justify if the knowledge-assisted initialization method is effective to avoid limitations in conventional K-Mean (using random initialization method) by minimizing the possibility of center redundancy, dead center, and the possibility of initial centroid to trap in local minima; 2) justify if the knowledge-assisted initialization method could reduce the iteration numbers (i.e., by reducing the search space) of the K-Mean algorithm. The boxplots in Figure 8 shows the comparison in terms of iteration numbers for different initialization methods using both the dataset. Based on the boxplot, it is found that the knowledge-assisted K-Mean has a smaller search space reflected by lower iteration numbers to compute the final centroids. As compared to other initialization methods, the boxplot of the knowledge-assisted K-Mean has a compact distribution and outperformed the Random method [62], Forgy method [63], and K-Mean++ [64]. When compare to Greedy K-Mean++ [65], the iteration numbers obtained from the knowledge-assisted K-Mean are highly comparative and slightly better by demonstrating lower iteration numbers with lower median value (Greedy K-Mean++ [65] (median): 9 and knowledge-assisted K-Mean (median): 7).

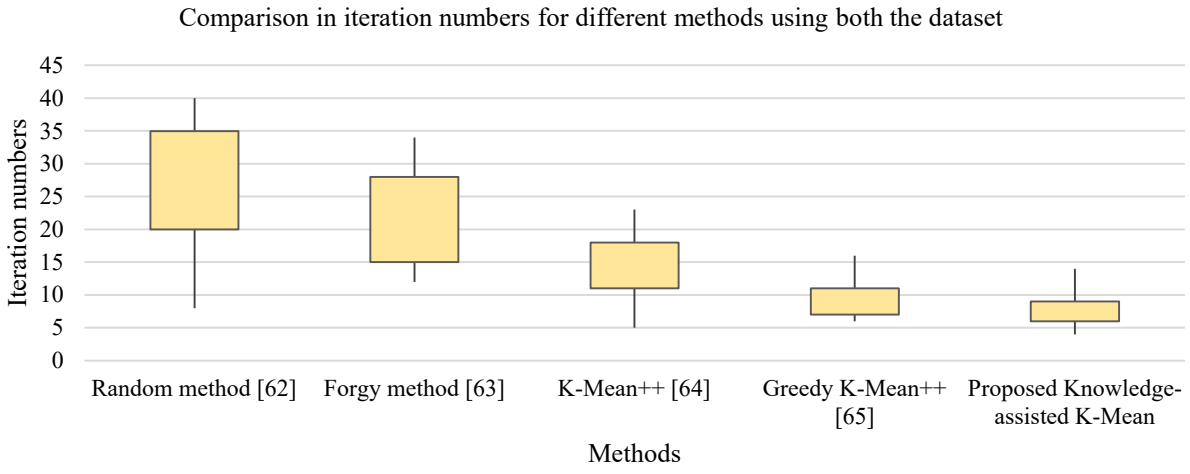


Figure 8. Comparison in terms of iteration numbers for different methods using both the datasets.

4.2 Mitosis cells segmentation performance evaluation

The main purpose of the mitosis cells segmentation performance evaluation is to determine the capability of retaining the mitosis cells after implementation of segmentation and false positive reduction stages in a detection framework. For ideal performance, in the segmentation stage, the detection framework aims to segment all the mitosis cells as foreground, whereas, in the false positive reduction stage, the detection framework aims to remove most of the false positive (i.e., non-mitosis cells) and retain all the true positive (i.e., mitosis cells). Therefore, in this study, a high total number of segmented mitosis cells is preferable. Tables 10 and 11 show the comparison in terms of the total number of segmented mitosis cells with different methods after implementation of the respective segmentation and false positive reduction stages using the custom and MITOS [17] datasets, respectively.

Based on the table, it is found that the proposed knowledge-assisted detection framework demonstrates a better performance as the total number of segmented mitosis cells for both the datasets are the highest (i.e., custom dataset: 825/ 832 and MITOIS dataset [17]: 321/ 327). In the proposed knowledge-assisted detection framework, all the mitosis cells were segmented as foreground in the segmentation stage using the knowledge-assisted K-Mean method with a trade-off of a high number of false positive. Nevertheless, it is important to remark that most of the false positive were then removed via the knowledge-assisted false positive reduction stage as in Section 2.3 with minimal loss in mitosis cells.

Table 10. Comparison in terms of the total number of segmented mitosis cells for different methods using custom dataset.

Methods	Total number of segmented mitosis cells	Total number of mitosis cells
Gamma-gaussian mixture model [36]	808	832
Maximized Inter-Class Weighted Mean [25]	817	
Proposed	825	

Table 11. Comparison in terms of the total number of segmented mitosis cells for different methods using MITOS dataset [17].

Methods	Total number of segmented mitosis cells	Total number of mitosis cells
Blue ratio [35]	314	327
Anisotropic Diffusion [66]	314	
Gradient minimization [67]	315	
Gamma-gaussian mixture model [36]	317	
Pixel features [32]	318	
Maximized Inter-Class Weighted Mean [25]	318	
Proposed	321	

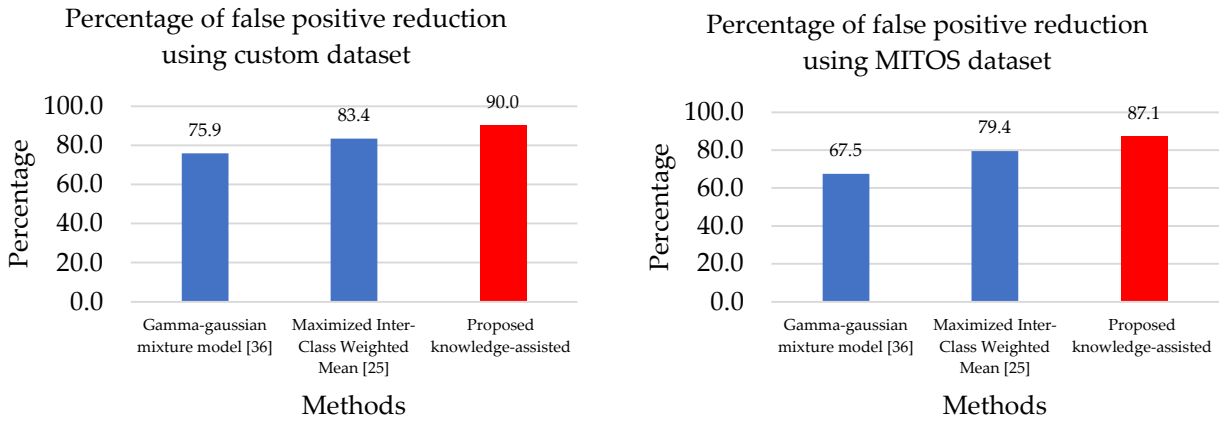
4.3 False positive reduction method performance evaluation

Based on Tables 10 and 11 in the previous sub-section, generally, the majority of the established works can retain most of the mitosis cells after the segmentation and false positive reduction stages. Nevertheless, as aforementioned, the number of false positive (i.e., non-mitosis cells) versus the number of true positive (i.e., mitosis cells) are commonly in great unbalance, where a high number of false positive would impinge the overall performance of the detection framework (low F1-score). Thus, removing the false positive is essential to improve the performance of mitosis cells detection. To establish a fair com-parison, in this section, the performance evaluation is done by benchmarking with two established works that incorporate different false positive reduction methods (i.e., Gamma-gaussian mixture model [36] and Maximized Inter-Class Weighted Mean [25]) after the respective segmentation stage. The percentage of false positive reduction can be calculated using Eq (5).

$$Percentage\ of\ false\ positive\ candidates\ reduction = \frac{i-o}{i} x100\% \tag{5}$$

where i denotes the number of mitosis candidates before the false positive reduction stage and o denotes the number of mitosis candidates after the false positive reduction stage. Figure 9 shows the comparison in terms of percentages in false positive reduction for different methods, where Figure 9 (a) and (b) show the outputs evaluated using the custom and the MITOS [17] datasets, respectively. A high percentage in false positive reduction reflects that the false positive reduction method is capable to remove more false positive (i.e., non-mitosis cells) and is preferable in this study.

Based on Figure 9, it is evident that the proposed knowledge-assisted false positive reduction method can remove most of the false positive (i.e., 90.0 and 87.1% of the false positive, respectively, for custom and MITOS [17] datasets). For the custom dataset, the percentage obtained by the proposed knowledge-assisted false positive reduction method is 14.1 and 6.6% higher than that of the Gamma-gaussian mixture model [36] and Maximized Inter-Class Weighted Mean [25], respectively, whereas, for the MITOS dataset [17], 19.6 and 7.7% higher than that of the Gamma-gaussian mixture model [36] and Maximized Inter-Class Weighted Mean [25], respectively.



(a) Outputs obtained using custom dataset. (b) Outputs obtained using MITOS dataset [17].

Figure 9. Comparison in terms of percentage in false positive reduction using (a) custom and (b) MITOS [17] datasets.

4.4 Mitosis cells detection performance evaluation

As aforementioned, in mitosis cells detection, the number of mitosis cells (true positive) versus the number of non-mitosis cells (true negative) is usually in great unbalance, where mitosis cells are very few as compared to non-mitosis cells in any dataset. Therefore, evaluation parameter such as F1-score is useful where the detection of non-mitosis cells (true negative) is excluded in the calculation. To calculate the F1-score, one must first compute the precision and recall. Equations 6 to 8 show the formula for F1-score, precision, and recall, respectively.

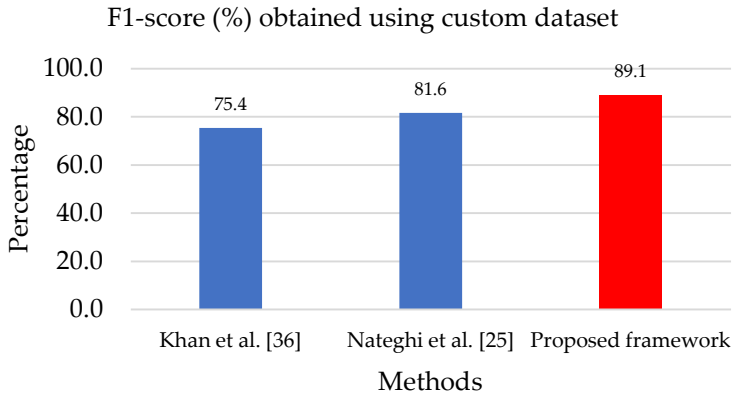
$$F1\text{-}score = \frac{2x\ Precision\ x\ Recall}{Precision+Recall}\ x\ 100\% \tag{6}$$

$$Precision = \frac{TP}{TP + FP}\ x\ 100\ \% \tag{7}$$

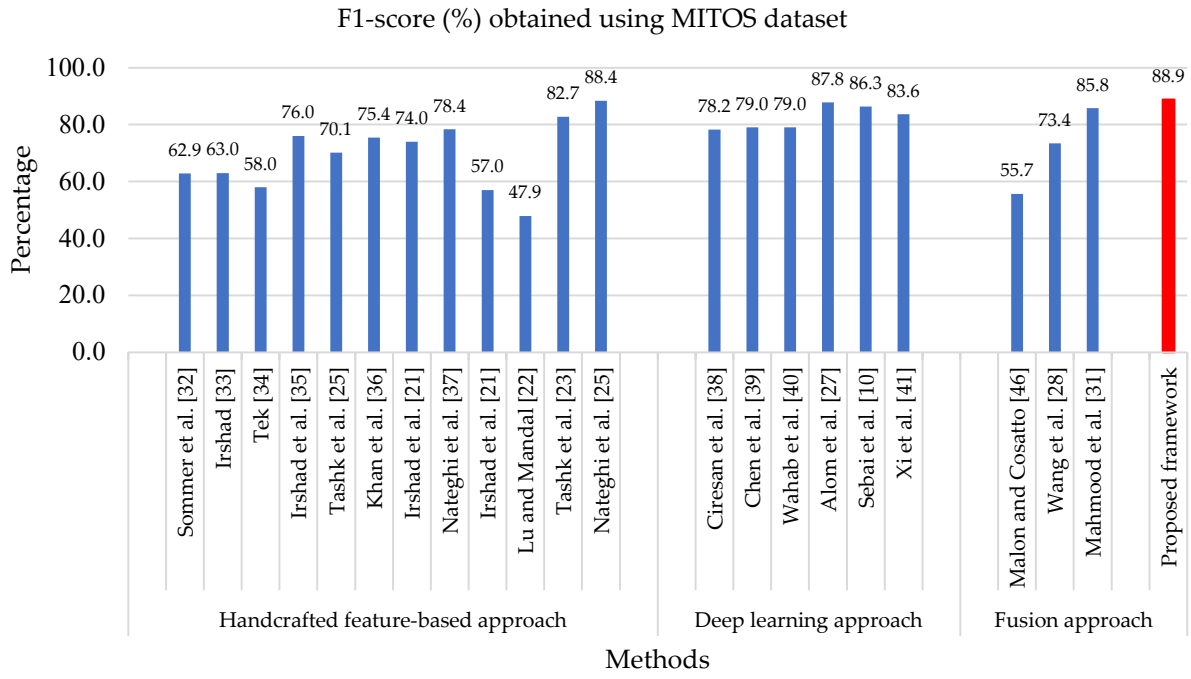
$$Recall = \frac{TP}{TP + FN}\ x\ 100\% \tag{8}$$

where true positive (TP) denotes the mitosis cells that are correctly labeled as mitosis cells, false positive (FP) denotes the non-mitosis cells that are wrongly labeled as mitosis cells; and false negative (FN) denotes the mitosis cells that are wrongly labeled as non-mitosis cells. In this study, F1-score would be used as the primary evaluation parameter. A high percentage in F1-score reflects an accurate detection of mitosis cells and is preferable in this study.

The mitosis cells classification was done using a 10-fold cross-validation SVM classifier with an RBF kernel. The data were randomly divided into 10 equal pieces. Each selected piece was chosen as a test set with training done on the remaining data. The cross-validation was then repeated 10 times (folds) with each subsample used exactly once as the validation data. All the observations were used for both training and validation, and each observation was used for validation exactly. To demonstrate the capability of the proposed detection framework, the outputs of the proposed detection framework using both datasets are benchmarked to the recent works. For the custom dataset, the comparison of the outputs was done on recent works which incorporate different false positive reduction methods: Gamma-gaussian mixture model [36] and Maximized Inter-Class Weighted Mean methods [25]. The proposed knowledge-assisted detection framework outperformed these recent works by achieving 89.1% in F1-score. For the MITOS dataset [17], the performance comparison is mainly focuses on recent works using handcrafted features-based approach. Nevertheless, considering the growing body of literature in deep learning and fusion (i.e., handcrafted features-based and deep learning) approaches, several recent works in respective approaches are included for benchmarking purposes to justify the applicability of the proposed detection framework. Based on the outputs, the proposed knowledge-assisted detection framework outperformed all the recent works using handcrafted feature-based approach by achieving 88.9% in F1-score. The F1-score of the proposed knowledge-assisted detection method is 0.5% higher than that of the best performing method amongst the handcrafted feature-based method (i.e., Nateghi et al. [25]). When comparing to recent works using deep learning and fusion approaches, the F1-score of the proposed knowledge-assisted detection framework is found promising and outperformed the best-performing methods using deep learning and fusion approaches by 1.1% (i.e., Alom et al. [27]) and 3.1% (i.e., Mahmood et al. [31]), respectively. Figure 10 shows the comparison of the overall outputs in terms of F1-score for different detection frameworks using custom (Figure 10 (a)) and MITOS [17] datasets (Figure 10 (b)).



(a) Outputs obtained using custom dataset.



(b) Outputs obtained using MITOS dataset [17].

Figure 10. Outputs comparison in terms of F1-score for different detection frameworks using (a) custom and (b) MITOS [17] datasets.

For both datasets, the classification obtained high TP values. The high TP is reflected in the high FI-score percentages. The high TP perhaps caused by the good performance in the hyperchromatic nucleus segmentation using the improved K-Mean and the knowledge-assisted false positive reduction methods. In the hyperchromatic nucleus segmentation, all the mitosis cells were retained on the images. This demonstrates the ability of the knowledge-assisted K-Mean method in providing a promising segmentation without removing the region of interest or ground truth (i.e., mitosis cells). In the knowledge-assisted false positive reduction stage, for custom and MITOS [17] datasets, only 0.8 (7 mitosis cells) and 2.0% (6 mitosis cells), respectively, of the ground truth were lost. In terms of false positives, this stage successfully removed 90.0 and 87.1% false positives in the custom and MITOS [17] datasets, respectively. The overall classification result demonstrates robustness with an optimal F1-

score of 89.1 and 88.9% for custom and MITOS [17] datasets, respectively. Overall, the obtained F1-score was found promising and demonstrated superiority when compared to recent works using handcrafted feature-based, deep learning, and fusion approaches for both datasets. This shows integration between domain knowledge (from human experts) and image processing procedure can help to improve the performance of the mitosis cells detection producing complementary results.

The low FP and high TP have enhanced the overall F1-score of the proposed knowledge-assisted detection framework. The combination of morphological and texture features has provided better feature discrimination between the mitosis and non-mitosis cells. The computation complexity of the proposed detection framework remains low as the proposed detection framework did not involve complex mathematic equations (as in the deep learning and fusion methods). Domain knowledge was incorporated in the detection framework and was used to guide the image analysis algorithms. The domain knowledge and strategies of the histopathologist were effectively incorporated in the mitosis cells detection framework and were used to formalize the basic building blocks in mitosis cells detection. Based on the results obtained from each stage and the overall performance, the domain knowledge is found prominent and potentially holds a significant role in medical image analysis, especially for mitosis cells detection.

5 Conclusions

In this study, an automated knowledge-assisted mitosis cells detection framework is proposed. The proposed detection framework is developed by incorporating the strategies of the histopathologist and domain knowledge approach which aims to perform recognition of complex and inconsistent mitosis cells in breast histopathology images. The main novelty of the detection framework lies within the simple yet powerful detection capability with knowledge transferred from human experts, producing promising detection results. The contributions of the detection framework are: 1) a knowledge-assisted K-Mean that is able to segment all the mitosis cells from the background and reduce the search space in computing the final centroids, which is reflected with low iteration numbers; 2) a knowledge-assisted false positive reduction method which capable to remove most of the false positive (i.e., non-mitosis cells). For both the datasets, the knowledge-assisted K-Mean is found able to segment all the mitosis cells with a trade-off of high false positive numbers. The knowledge-assisted false positive reduction method is then applied to remove most of the false positive (i.e., 90.0 and 87.1% in the custom and MITOS [17] datasets, respectively); 3) the proposed framework achieves a high percentage in F1-score (i.e., 89.1%) when tested using the custom dataset which reflects a good agreement with respect to ground truth demarcated by the histopathologist, 4) the proposed framework achieves a high percentage in F1-score (i.e., 88.9%) when tested using the MITOS dataset [17] and outperformed the recent works with handcrafted feature-based, deep learning, and fusion approaches; 5) to the best of our knowledge, mitosis cells detection using a knowledge-assisted detection framework by incorporating domain knowledge as well as knowledge transfer from human experts is yet to be available in the literature. The proposed knowledge-assisted K-Mean and false positive reduction method are found capable to improve the performance of the mitosis cells detection producing complementary results. This is justifiable as the proposed knowledge-assisted detection framework obtains high F1-scores when the proposed detection framework is evaluated using two different datasets. Additionally, when benchmarking with recent works using handcrafted feature-based, deep learning, and fusion approaches, the F1-scores of the proposed detection framework are found promising and outperformed the recent works. For future work, the proposed framework is intended to test using other publicly available datasets (e.g., AMIDA [18], MITOS-ATYPIA [19], and TUPAC [20]) to

further justify the applicability and capability of the proposed framework.

Acknowledgments

The authors gratefully acknowledge the financial support from the Fundamental Research Grant Scheme (FRGS) under a grant number of FRGS/1/2016/SKK06/UNIMAP/02/3 from the Ministry of Higher Education Malaysia. The protocol of this study had been approved by Medical Research and Committee of National Medical Research Register (NMRR) Malaysia referring to the protocol number: NMRR-17-281-34236.

Conflict of interest

The authors declare no conflict of interest.

References

1. T. Mathew, J. R. Kini, J. Rajan, Computational methods for automated mitosis detection in histopathology images: A review, *Biocybern. Biomed. Eng.*, **41** (2021), 64–82. doi: 10.1016/j.bbe.2020.11.005.
2. P. H. Tan, I. Ellis, K. Allison, E. Brogi, S. B. Fox, S. Lakhani, et al., The 2019 World Health Organization classification of tumours of the breast, *Histopathology*, **77** (2020), 181–185. doi: 10.1111/his.14091.
3. E. A. Rakha, J. S. Reis-Filho, F. Baehner, D. J. Dabbs, T. Decker, V. Eusebi, et al., Breast cancer prognostic classification in the molecular era: the role of histological grade, *Breast Cancer Res*., **12** (2010), 207. doi: 10.1186/bcr2607.
4. S. M. Samuel, E. Varghese, S. Varghese, D. Büsselberg, Challenges and perspectives in the treatment of diabetes associated breast cancer, *Cancer Treat. Rev.*, **70** (2018), 98–111. doi: 10.1016/j.ctrv.2018.08.004.
5. A. D. Shah, A. K. Mehta, N. Talati, R. Brem, L. R. Margolies, Breast tissue markers: Why? What's out there? How do I choose? *Clin. Imaging*, **52** (2018), 123–136. doi: 10.1016/j.clinimag.2018.07.003.
6. A. Ramírez-torres, R. Rodríguez-Ramos, F. J. Sabina, C. García-Reimbert, R. Penta, J. Merodio, et al., The role of malignant tissue on the thermal distribution of cancerous breast, *J. Theor. Biol.*, **426** (2017), 152–161. doi: 10.1016/j.jtbi.2017.05.031.
7. J. Dalle, W. K. Leow, D. Racoceanu, A. E. Tutac, T. C. Putti, Automatic breast cancer grading of histopathological images, in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, **2** (2008), 3052–3055. doi: 10.1109/IEMBS.2008.4649847.
8. C. Li, X. Wang, W. Liu, L. J. Latecki, DeepMitosis: Mitosis detection via deep detection, verification and segmentation networks, *Med. Image Anal.*, **45** (2018), 121–133. doi: 10.1016/j.media.2017.12.002.
9. A. Duggento, A. Conti, A. Mauriello, M. Guerrisi, N. Toschi, Deep computational pathology in breast cancer, *Semin. Cancer Biol.*, **72** (2021), 226–237. doi: 10.1016/j.semcancer.2020.08.006.
10. M. Sebai, X. Wang, T. Wang, MaskMitosis: a deep learning framework for fully supervised, weakly supervised, and unsupervised mitosis detection in histopathology images, *Med. Biol. Eng. Comput.*, **58** (2020), 1603–1623. doi: 10.1007/s11517-020-02175-z.

11. M. Peikari, M. J. Gangeh, J. Zubovits, G. Clarke, A. L. Martel, Triaging diagnostically relevant regions from pathology whole slides of breast cancer: A texture based approach, *IEEE Trans. Med. Imaging*, **35** (2016), 307–315. doi: 10.1109/TMI.2015.2470529.
12. M. Saha, C. Chakraborty, D. Racoceanu, Efficient deep learning model for mitosis detection using breast histopathology images, *Comput. Med. Imaging Graph.*, **64** (2018), 29–40. doi: 10.1016/j.compmedimag.2017.12.001.
13. S. Roy, A. k. Jain, S. Lal, J. Kini, A study about color normalization methods for histopathology images, *Micron*, **114** (2018), 42–61. doi: 10.1016/j.micron.2018.07.005.
14. X. Li, K. N. Plataniotis, A complete color normalization approach to histopathology images using color cues computed from saturation-weighted statistics, *IEEE Trans. Biomed. Eng.*, **62** (2015), 1862–1873. doi: 10.1109/TBME.2015.2405791.
15. Ş. Öztürk, B. Akdemir, Phase classification of mitotic events using selective dictionary learning for stem cell populations, *Comput. Electr. Eng.*, **67** (2018), 25–37. doi: 10.1016/j.compeleceng.2018.03.025.
16. B. Plasma, C. Mott, Enigmatic morpho insight: Mitosis at a glance, *J. Oral Maxillofac. Pathol.*, **18** (2014), 2–5. doi: 10.4103/0973-029X.141175.
17. MITOS–dataset, 2012. Available from: http://ludo17.free.fr/mitos_2012/dataset.html.
18. AMIDA–dataset, 2013. Available from: http://amida13.isi.uu.nl/?q=node/62.
19. MITOS-ATYPIA–dataset, 2014. Available from: https://mitos-atypia-14.grand-challenge.org/Dataset/.
20. TUPAC–dataset, 2016. Available from: http://tupac.tue-image.nl/node/3.
21. H. Irshad, A. Gouaillard, L. Roux, D. Racoceanu, Multispectral band selection and spatial characterization: Application to mitosis detection in breast cancer histopathology, *Comput. Med. Imaging Graph.*, **38** (2014), 390–402. doi: 10.1016/j.compmedimag.2014.04.003.
22. C. Lu, M. Mandal, Toward automatic mitotic cell detection and segmentation in multispectral histopathological images, *IEEE J. Biomed. Heal. Informatics*, **18** (2014), 594–605. doi: 10.1109/JBHI.2013.2277837.
23. A. Tashk, M. S. Helfroush, H. Danyali, M. Akbarzadeh-jahromi, Automatic detection of breast cancer mitotic cells based on the combination of textural, statistical and innovative mathematical features, *Appl. Math. Model.*, **39** (2015), 6165–6182. doi: 10.1016/j.apm.2015.01.051.
24. A. Paul, D. P. Mukherjee, Mitosis detection for invasive breast cancer grading in histopathological images, *IEEE Trans. Image Process.*, **24** (2015), 4041–4054. doi: 10.1109/TIP.2015.2460455.
25. R. Nateghi, H. Danyali, M. S. Helfroush, Maximized inter-class weighted mean for fast and accurate mitosis cells detection in breast cancer histopathology images, *J. Med. Syst.*, **41** (2017), 146. doi: 10.1007/s10916-017-0773-9.
26. D. Cai, X. Sun, N. Zhou, X. Han, J. Yao, Efficient mitosis detection in breast cancer histology images by RCNN, in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, (2019), 919–922. doi: 10.1109/ISBI.2019.8759461.
27. M. Z. Alom, T. Aspiras, T. M. Taha, T. Bowen, V. K. Asari, MitosisNet: End-to-end mitotic cell detection by multi-task learning, *IEEE Access*, **8** (2020) 68695–68710. doi: 10.1109/ACCESS.2020.2983995.
28. H. Wang, A. C. Roa, A. N. Basavanhally, H. L. Gilmore, N. Shih, M. Feldman, et al., Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features, *J. Med. Imaging*, **1** (2014), 034003. doi: 10.1117/1.JMI.1.3.034003.

29. K. S. Beevi, M. S. Nair, G. R. Bindu, A multi-classifier system for automatic mitosis detection in breast histopathology images using deep belief networks, *IEEE J. Transl. Eng. Heal. Med.*, **5** (2017), 1–11. doi: 10.1109/JTEHM.2017.2694004.
30. D. K. Das, P. K. Dutta, Efficient automated detection of mitotic cells from breast histological images using deep convolution neutral network with wavelet decomposed patches, *Comput. Biol. Med.*, **104** (2019), 29–42. doi: 10.1016/j.compbiomed.2018.11.001.
31. T. Mahmood, M. Arsalan, M. Owais, M. B. Lee, K. R. Park, Artificial intelligence-based mitosis detection in breast cancer histopathology images using faster R-CNN and deep CNNs, *J. Clin. Med.*, **9** (2020), 749. doi: 10.3390/jcm9030749.
32. C. Sommer, L. Fiaschi, F. A. Hamprecht, D. W. Gerlich, Learning-based mitotic cell detection in histopathological images, in *Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012)*, (2012), 2306–2309.
33. H. Irshad, Automated mitosis detection in histopathology using morphological and multi-channel statistics features, *J. Pathol. Inform.*, **4** (2013), 10. doi: 10.4103/2153-3539.112695.
34. F. B. Tek, Mitosis detection using generic features and an ensemble of cascade adaboosts, *J. Pathol. Inform.*, **4** (2013), 12. doi: 10.4103/2153-3539.112697.
35. H. Irshad, S. Jalali, L. Roux, D. Racoceanu, L. J. Hwee, G. L. Naour, et al., Automated mitosis detection using texture, SIFT features and HMAX biologically inspired approach, *J. Pathol. Inform.*, **4** (2013), 12. doi: 10.4103/2153-3539.109870.
36. A. M. Khan, H. ElDaly, N. M. Rajpoot, A gamma-gaussian mixture model for detection of mitotic cells in breast cancer histopathology images, *J. Pathol. Inform.*, (2013), 11. doi: 10.4103/2153-3539.112696.
37. R. Nateghi, H. Danyali, M. S. Helfroush,F. P. Pour, Automatic detection of mitosis cell in breast cancer histopathology images using genetic algorithm, in *2014 21th Iranian Conference on Biomedical Engineering (ICBME)*, (2014), 1–6. doi: 10.1109/ICBME.2014.7043883.
38. D. C. Cireşan, A. Giusti, L. M. Gambardella, J. Schmidhuber, Mitosis detection in breast cancer histology images with deep neural networks, *Lect. Notes Comput. Sci.*, **8150** (2013), 411–418. doi: 10.1007/978-3-642-40763-5_51.
39. H. Chen, X. Wang, P. A. Heng, Automated mitosis detection with deep regression networks, in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, (2016), 1204–1207. doi: 10.1109/ISBI.2016.7493482.
40. N. Wahab, A. Khan, Y. S. Lee, Two-phase deep convolutional neural network for reducing class skewness in histopathological images based breast cancer detection, *Comput. Biol. Med.*, **85** (2017), 86–97. doi: 10.1016/j.compbiomed.2017.04.012.
41. X. Lu, Z. You, M. Sun, J. Wu, Z. Zhang, Breast cancer mitotic cell detection using cascade convolutional neural network with U-Net, *Math. Biosci. Eng.*, **18** (2020), 673–695. doi: 10.3934/MBE.2021036.
42. S. Albarqouni, C. Baur, F. Achilles, V. Belagiannis, S. Demirci, N. Navab, AggNet: Deep learning from crowds for mitosis detection in breast cancer histology images, *IEEE Trans. Med. Imaging*, **35** (2016), 1313–1321. doi: 10.1109/TMI.2016.2528120.
43. T. Wollmann, K. Rohr, Deep residual hough voting for mitotic cell detection in histopathology images, in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, (2017), 341–344. doi: 10.1109/ISBI.2017.7950533.

44. D. Romo-Bucheli, A. Janowczyk, H. Gilmore, E. Romero, A. Madabhushi, A deep learning based strategy for identifying and associating mitotic activity with gene expression derived risk categories in estrogen receptor positive breast cancers, *Cytometry Part A*, 91 (2017), 566–573.doi: 10.1002/cyto.a.23065.
45. H. Chen, J. Qin, Q. Dou, X. Wang, P. A. Heng, Mitosis detection in breast cancer histology images via deep cascaded networks, in *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, (2016), 1160–1166.
46. C. D. Malon, E. Cosatto, Classification of mitotic figures with convolutional neural networks and seeded blob features, *J. Pathol. Inform.*, **4** (2013), 9. doi: 10.4103/2153-3539.112694.
47. X. J. Tan, N. Mustafa, M. Y. Mashor, K. S. Ab Rahman, A novel quantitative measurement method for irregular tubules in breast carcinoma, *Eng. Sci. Technol. Int. J.*, (2021). doi: 10.1016/j.jestch.2021.08.008.
48. A. Basavanhally, E. Yu, J. Xu, S. Ganesan, M. Feldman, J. Tomaszewski, et al., Incorporating domain knowledge for tubule detection in breast histopathology using O'Callaghan neighborhoods, *SPIE Med. Imaging*, **7963** (2011), 796310–796315. doi: 10.1117/12.878092.
49. D. Racoceanu, F. Capron, Towards semantic-driven high-content image analysis: An operational instantiation for mitosis detection in digital histopathology, *Comput. Med. Imaging Graph.*, **42** (2015), 2–15. doi: 10.1016/j.compmedimag.2014.09.004.
50. M. C. Clark, L. O. Hall, D. B. Goldgof, R. Velthuizen, F. R. Murtagh, M. S. Silbiger, Automatic tumor segmentation using knowledge-based techniques, *IEEE Trans. Med. Imaging*, **17** (1998), 187–201. doi: 10.1109/42.700731.
51. A. Madabhushi, D. N. Metaxas, Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions, *IEEE Trans. Med. Imaging*, **22** (2003), 155–169. doi: 10.1109/TMI.2002.808364.
52. H. Pan, J. Li, W. Zhang, Incorporating domain knowledge into medical image clustering, *Appl. Math. Comput.*, **185** (2007), 844–856. doi: 10.1016/j.amc.2006.06.083.
53. S. Naik, S. Doyle, S. Agner, A. Madabhushi, M. Feldman, J. Tomaszewski, Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology, in *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, (2008), 284–287. doi: 10.1109/ISBI.2008.4540988.
54. M. Macenko, M. Niethammer, J. S. Marron, D. Borland, J. T. Woosley, X. Guan, et al., A method for normalizing histology slides for quantitative analysis, in *2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, (2009), 1107–1110. doi: 10.1109/ISBI.2009.5193250.
55. J. Wang, M. J. Slattery, M. H. Hoskins, S. Liang, C. Dong, Q. Du, Monte carlo simulation of heterotypic cell aggregation in nonlinear shear flow, *Math. Biosci. Eng.*, **3** (2006), 683–696. doi: 10.3934/mbe.2006.3.683.
56. J. C. Han, F. Shang, P. Li, B. Li, Y. Zhou, Y. Huang, Coupling bayesian-monte carlo simulations with substance flow analysis for efficient pollutant management: A case study of phosphorus flows in China, *Resour. Conserv. Recycl.*, **169** (2021), 105550. doi: 10.1016/j.resconrec.2021.105550.
57. Q. Li, Y. Wang, H. Liu, X. He, D. Xu, J. Wang, et al., Leukocyte cells identification and quantitative morphometry based on molecular hyperspectral imaging technology, *Comput. Med. Imaging Graph.*, **38** (2014), 171–178. doi: 10.1016/j.compmedimag.2013.12.008.

58. M. M. Saleck, A. Elmoutaouakkil, M. Moucouf, Tumor detection in mammography images using fuzzy C-means and GLCM texture features, in *2017 14th International Conference on Computer Graphics, Imaging and Visualization*, (2017), 122–125. doi: 10.1109/CGiV.2017.22.
59. B. Abraham, M. S. Nair, Computer-aided classification of prostate cancer grade groups from MRI images using texture features and stacked sparse autoencoder, *Comput. Med. Imaging Graph.*, **69** (2018), 60–68. doi: 10.1016/j.compmedimag.2018.08.006.
60. I. Qabajeh, F. Thabtah, F. Chiclana, A recent review of conventional vs. automated cybersecurity anti-phishing techniques, *Comput. Sci. Rev.*, **29** (2018), 44–55. doi: 10.1016/j.cosrev.2018.05.003.
61. H. Amitha, I. Selvamani, D. A. S. Dhas, Developement of computer aided system for detection and classification of mitosis using SVM, in *2017 International Conference on Inventive Computing and Informatics (ICICI)*, (2017), 954–958. doi: 10.1109/ICICI.2017.8365278.
62. J. MacQueen, Some Methods for classification and snalysis of multivariate observations, in *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, (1967), 281–297.
63. E. Forgy, Cluster analysis of multivariate data: Efficiency vs. interpretability of classification, *Biometrics*, **21** (1965), 768.
64. D. Arthur, S. Vassilvitskii, K-means++: The advantages of careful seeding, in *Proceedings of the eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, (2006), 1027–1035.
65. A. Bhattacharya, J. Eube, H. Röglin, M. Schmidt, Noisy, greedy and not so greedy k-means++, preprint, arXiv:1912.00653v1.
66. P. Perona, J. Malik, Scale-space and edge detection using anisotropic diffusion, *IEEE Trans. Pattern Anal. Mach. Intell.*, **12** (1990), 629–639. doi: 10.1109/34.56205.
67. L. Xu, C. Lu, Y. Xu, J. Jia, Image smoothing via L0 gradient minimization, *ACM Trans. Graph.*, **30** (2011), 1–12. doi: 10.1145/2024156.2024208.

Appendix A

Table A shows some of the default settings for the SVM classifier used in this study.

Table A. Settings of the SVM classifier (default settings).

Parameters	Settings
Box constraint	1
Kernel function	RBF kernel
Kernel scale	1
Flag to standardize predictor data	False
Solver (optimization routine)	Iterative Single Data Algorithm (ISDA)
Flag to clip alpha coefficients	True