

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/388959013>

# AUTOMATED CERVICAL CELL NUCLEI SEGMENTATION BASED ON MULTILAYER UNSUPERVISED CLUSTERING ALGORITHM AND MORPHOLOGICAL APPROACH

Article in *Mitteilungen Klosterneuburg* · January 2025

DOI: 10.61586/48dz8

CITATIONS

0

READS

9

6 authors, including:



**Khalis Nukman**

University of Malaysia, Perlis

2 PUBLICATIONS 2 CITATIONS

SEE PROFILE



**Wan Azani Mustafa**

University of Malaysia, Perlis

418 PUBLICATIONS 2,354 CITATIONS

SEE PROFILE



**Hiam H Alquran**

Yarmouk University

107 PUBLICATIONS 1,198 CITATIONS

SEE PROFILE



**Syahrul N. Junaini**

University of Malaysia, Sarawak

75 PUBLICATIONS 559 CITATIONS

SEE PROFILE

# AUTOMATED CERVICAL CELL NUCLEI SEGMENTATION BASED ON MULTILAYER UNSUPERVISED CLUSTERING ALGORITHM AND MORPHOLOGICAL APPROACH

**Khalis Danial Nukman Khiruddin<sup>1</sup>, Wan Azani Mustafa<sup>1, 2\*</sup>, Khairur Rijal Jamaludin<sup>3\*</sup>,  
Khairul Shakir Ab Rahman<sup>4</sup>, Hiam Alquran<sup>5</sup> and Syahrul Junaini<sup>6</sup>**

<sup>1</sup> Faculty of Electrical Engineering & Technology, Universiti Malaysia Perlis, Pauh Putra Campus, 02600 Arau, Perlis, Malaysia

<sup>2</sup> Advanced Computing (AdvComp), Centre of Excellence (CoE), Universiti Malaysia Perlis (UniMAP), Campus Pauh Putra, 02600 Arau, Perlis, Malaysia

<sup>3</sup> Faculty of Artificial Intelligence, Universiti Teknologi Malaysia, Jalan Sultan Yahya Petra, 54100 Kuala Lumpur

<sup>4</sup> Department of Pathology, Hospital Tuanku Fauziah, Kangar, Perlis, 02000, Malaysia

<sup>5</sup> Department of Biomedical Engineering, Jordan University of Science and Technology, Irbid 22110, Jordan

<sup>6</sup> Faculty of Computer Science & Information Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

*Corresponding author(s): Wan Azani Mustafa: wanazani@unimap.edu.my, Khairur Rijal Jamaludin: khairur.kl@utm.my*

## ABSTRACT

Cervical cancer, a leading cause of female mortality globally, results from abnormal cell growth in the cervix, making early detection crucial. This study suggests an automated segmentation approach that is more accurate and faster than traditional methods, which face challenges such as contrast problems and noise. The research aims to develop an algorithm for autonomously segmenting the nucleus of cervical cells to aid in diagnosis and future research. The proposed methodology involves extracting and enhancing the brightness (V channel) of input images using a median filter and Pairing Adaptive Gamma Correction and Histogram Equalisation (PAGCHE). A segmentation method based on multiple Fuzzy C-Means Clustering (FCM) layers and flexible morphological approaches is used to segment the nuclei in Pap smear images. The study utilized 917 images from the Herlev dataset to evaluate the method's performance. Image Quality Assessment (IQA) metrics, including accuracy, sensitivity, precision, specificity, and F-measure, demonstrate the method's efficacy. Results show the proposed approach consistently achieves over 90% accuracy. It outperforms other methods like Chan-Vese (CV), Canny edge-based, adaptive threshold, and FCM, with the highest accuracy, F1-measure, and sensitivity at 92.19%, 94.40%, and 93.38%, respectively. It also ranks second in precision and specificity, at 96.41% and 94.25%. These results indicate the approach's high accuracy, sensitivity, and specificity, making it a reliable tool for early detection and diagnosis. The algorithm's successful implementation could improve patient outcomes and support further research in cervical cancer diagnostics. The average segmentation score of the 917 images exceeds 90%, highlighting the method's flexibility.

**Keywords:** cervical cell; contrast enhancement; image segmentation; nucleus; pap smear.

## INTRODUCTION

In the cells of the cervix, cervical cancer originates from the lower portion of the uterus that is anatomically connected to the vagina. It occurs when the cervix cells increase abnormally, as stated by Sausen et al. (Guimarães et al., 2022; Sausen et al., 2023) An estimated 604,000 new cases of cervical cancer and 342,000 deaths occurred globally in 2020, according to Sausen et al. and H. Sung et al. in their papers (Sausen et al., 2023; Sung et al., 2021). Cervical cancer is a common form of the disease that likely every woman will eventually develop; furthermore, it is associated with a high mortality rate, as stated by Mustafa et al. (Mustafa et al., 2021; Mustafa, Halim, & Ab Rahman, 2020). Based on the paper by the same author (Mustafa, Halim, Jamlos, et al., 2020), the Papanicolaou test developed during the 1940s is also known as the Pap test or Pap smear, which enables microscopic analysis of these cells to detect precancerous or cancerous carries. The traditional method of cervical cancer screening is based on the Pap smear test. However, Macios et al. and Schiffman et al. stated in their studies (Macios & Nowakowski, 2022) and (Schiffman & de Sanjose, 2019) that this method has several limitations, including low sensitivity and specificity, which can lead to false-negative or false-positive results.

Cell segmentation has become an essential stage in defining the progression of cervical cancer. Mustafa et al., in their paper (Mustafa et al., 2021), stated that this resulted in the widespread use of computer-aided detection technologies in cervical cancer screening. Traditional manual therapies help to address the difficulty caused by a lack of medical resources. Although technological advancements significantly boost the early detection of cervical cancer, Alias et al. emphasised in their paper (Alias et al., 2022) that correct diagnosis remains problematic for a variety of reasons. Segmenting nuclei in cervical cytology pap smear images is essential in automated cervical cancer screening. Wan et al. noted in their paper (Wan et al., 2019) that cervical cells with false edges, overlapping cells, neutrophils, and artefacts make segmentation more challenging. Nonetheless, the complex technique appears incapable of automatic diagnosis due to the relatively low segmentation accuracy for aberrant cells. Artificial intelligence is gaining popularity in computer-aided diagnosis due to the capacity of various deep learning approaches to automatically extract visual attributes with high accuracy and low error (Fujita, 2020). This breakthrough has been widely used in multiple studies to

segment cell pictures. This study intends to contribute to a more effective nucleus identification approach.

There are many cell or nuclei image segmentation methods, which can be generally divided into several categories. First, cell or nucleus segmentation methods adopt region information to classify each pixel in an image (M. Zhao et al., 2021). Typical nucleus segmentation methods include threshold methods (Hameed et al., 2015; Madhloom et al., 2010; Mayala & Haugsøen, 2022; Yang et al., 2014), region growing (Arya et al., 2020; Nahrawi et al., 2021; Shilpa et al., 2021), clustering (Arya et al., 2020; Bandyopadhyay & Nasipuri, 2020; Haridas & Jayamalar, 2023) and watershed algorithm

Second, segmentation methods using cell or nucleus edge information utilise the discontinuity of grey information to segment the image; they mainly include differential operator methods and active contour models (Alias et al., 2022; Mustafa et al., 2021; M. Zhao et al., 2021).

Third, cell or nucleus segmentation methods are mainly based on related theories such as wavelet analysis (Imtiaz et al., 2023; Wu et al., 2021), mathematical morphology (Nahrawi et al., 2021; Plissiti et al., 2011) genetic algorithm (Li et al., 2010) and neural networks (Prasad Battula & Chandana, 2022). In addition, many state-of-the-art cell or nuclei segmentation algorithms are no longer based on a single algorithm but multiple algorithms together, namely, fusion algorithms, which aim to make up for the shortage of a single algorithm and achieve better segmentation results (Glotsos et al., 2018; Huang & Zhu, 2020)

Rajaroo and Singh employed a Hybrid dual-stage active contour method in the Construction of an Improved Normalized Graph Cut using an Image unfolding process (Rajaroo & Singh, 2020). Nevertheless, Plissiti et al. proposed a new segmentation approach based on the variability of intensity among superpixels (Plissiti et al., 2011). Jinjie Huang proposed a multi-scale fuzzy clustering algorithm to exclude nuclei, and they obtained high accuracy (William et al., 2019). Meanwhile, Meng Zhao et al. exploited the new Selective-Edge-Enhancement-based Nuclei Segmentation method. This method automatically segments the cervical image into small regions after considering the image pre-processing operation techniques (M. Zhao et al., 2021). On the other hand, many researchers exploited the benefits of deep learning to

segment the nuclei automatically, such as Alquran et al., who segmented the cervical images into nuclei and cytoplasm by employing a transfer learning method. They achieved higher accuracy for the nucleus region, reaching 92.0%. Jie Zhao et al. proposed a new method for segmenting nuclei accurately. Their method depends on using the Deformable Multipath Ensemble Model (DMMM). Their results were expressed similarly, reaching 93.3% (J. Zhao et al., 2019).

Deep learning has emerged as a promising solution for improving diagnostic accuracy, with various studies demonstrating its ability to automatically extract visual features with high precision and low error rates. However, deep learning-based segmentation methods come with significant challenges—they require high computational power, large training datasets, and are often ineffective when dealing with low-quality images. These limitations are particularly problematic in cervical cancer screening, where datasets are often small and image quality varies significantly.

Accurate cervical cell nucleus segmentation is crucial for early cervical cancer diagnosis, yet existing methods struggle with three major challenges:

1. Deep learning techniques require extensive processing power and resources, making them impractical for certain medical applications.
2. Supervised learning methods rely on large amounts of labelled data, which are often unavailable for cervical cytology.
3. Variability in cytology images leads to segmentation errors, reducing the reliability of automated methods.

To address these challenges, this study introduces a novel segmentation approach leveraging a multilayer Fuzzy C-Means algorithm, integrating clustering, region properties, and morphological techniques. Unlike deep learning-based methods, this approach is computationally efficient, does not require large training datasets, and remains effective despite low-quality images. Hence, the contributions of this study are as follows:

1. This study introduces a novel approach by integrating a fusion of clustering algorithms, region properties, and a morphological approach for segmentation. To the best of our

knowledge, this is the first study to combine all these methods, aiming to enhance the accuracy and efficiency of cervical cell nucleus segmentation.

2. This study proposed multilayer selection algorithm which only require low computational compared to deep learning techniques. This attribute makes it particularly suitable for cervical datasets, which often have limited available data and resources, thus eliminating the need for an extensive training dataset. Consequently, this method ensures a more accessible and efficient diagnostic tool for cervical cancer screening.

## MATERIALS AND METHODS

This section explained the proposed methodology for segmenting nuclei from Pap smear images. The sample images should consist of a single cervical cell, all colour images. This process was performed using the Matlab R2021b software. The framework involved image acquisition, pre-processing and segmentation. The main idea of the proposed method is based on the combination of Fuzzy C-Means Clustering (FCM) and several mathematical morphologies for segmenting the nucleus in the Pap smear images.

### Image Acquisition

The 917 sample images were collected from the Herlev database (Herlev University Hospital, Denmark). The database was obtained from/developed by NiSIS (EU coordination action, contract 13569), Nature-inspired Smart Information Systems, with particular significance for the Nature-Inspired Data Technology focus group. It is thus accessible (<http://fuzzy.iau.dtu.dk/download /smear2005>) on the World Wide Web. The sample images shown in Figure 1 below display the sample images of each class.

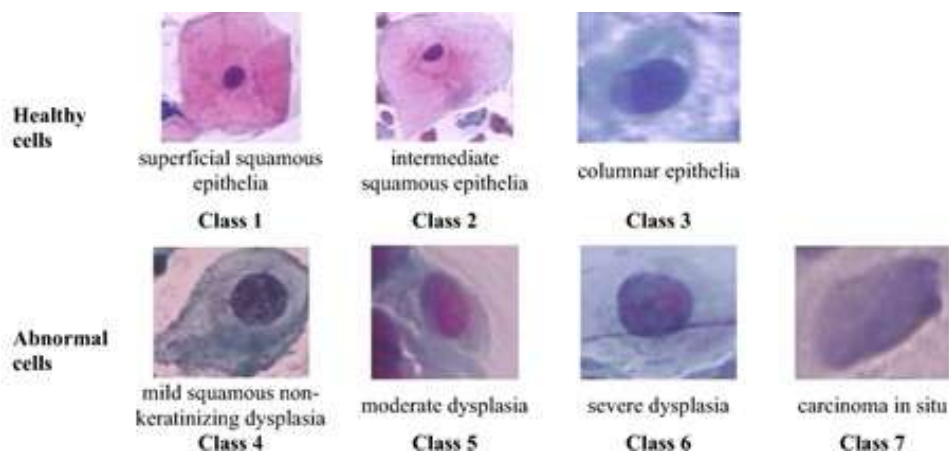


Figure 1: Sample Images of Herlev dataset.

### Pre-Processing

This study proposed two steps: denoise and colour contrast enhancement. The input images were converted to the colour space of the HSV system. HSV is one of the easy-to-use systems to identify colours and illumination and is closest to how humans perceive and compare colours. Extracted V channel will be applied with the median filter with  $15 \times 15$  windowing size to remove salt and pepper noise. As there will be a problem with the contrast after using a median filter in the image, therefore a novel Pairing Adaptive Gamma Correction and Histogram Equalization (PACGHE) proposed by Bataineh (Bataineh, 2023) was used to improve the colour contrast of pixels in the image. Referring to Figure 2, gamma and histogram will be extracted from the V channel of the input images. An adaptive gamma generator calculates gamma parameters ( $\gamma'$ ) based on the image's dark, medium, or bright conditions using the following proposed equation:

$$\gamma' = \frac{Mean}{L \times 0.33} + C \quad (1)$$

Where L is the maximum grey level of the images, here,  $L = 255$ , and it is multiplied by 0.33 to adapt the  $\gamma'$  value to one of the three global states, which are dark, medium or bright, for the images. C is a bias constant to overcome the problem of  $\gamma'$ ; here,  $C = 0.1$ . The computed gamma parameters propose a cumulative distribution function (CDF) that optimises the illumination values. Next, A significant modification of histogram equalisation (HE) is performed to improve image contrast and prepare the illumination for subsequent stages. A

proposed CDF function is used to enhance the illumination greatly. This function aims to improve the illumination and contrast adaptively to unify the visual properties of any image, ranging from considerably light to dark. New values of the brightness matrix are calculated using the processed cdf value as the following equation:

$$V_{MID}(x, y) = cdf'(V(x, y)) \times 255 \quad (2)$$

After calculating the  $V_{MID}$  matrix, a second equalisation process is applied to modify the displacement of values evenly. Instead of using a standard pdf, an equal density probability is used here. The distribution of displaced values in the  $V$ -MID matrix is corrected by rearranging equally to correct the illumination levels and contrast of the processed images. The result is a processed brightness matrix with improved contrast and illumination properties for the original  $V$  channel.

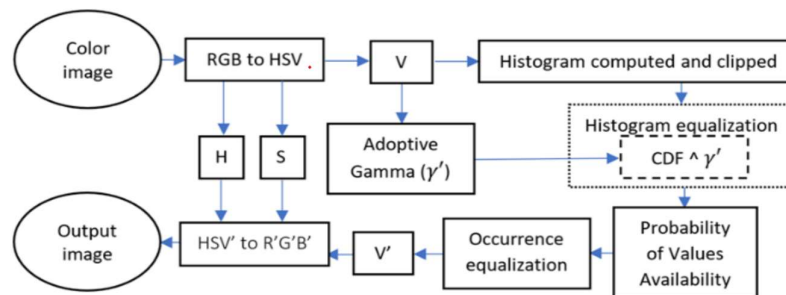


Figure 2: Flowchart of PAGCHE (Bataineh, 2023).

## Post-Processing

The nucleus image needs to be segmented after pre-processing. Figure 3 shows the flowchart for the overall process of the segmentation method. The process begins with an input image, which is first converted into the HSV (Hue, Saturation, Value) color space. The H, S, and V channels are separated, and the V channel undergoes pre-processing, resulting in a modified  $V'$  channel. This processed  $V'$  channel is then recombined with the H and S channels and converted back into the RGB color space. Next, the image is transformed into grayscale.



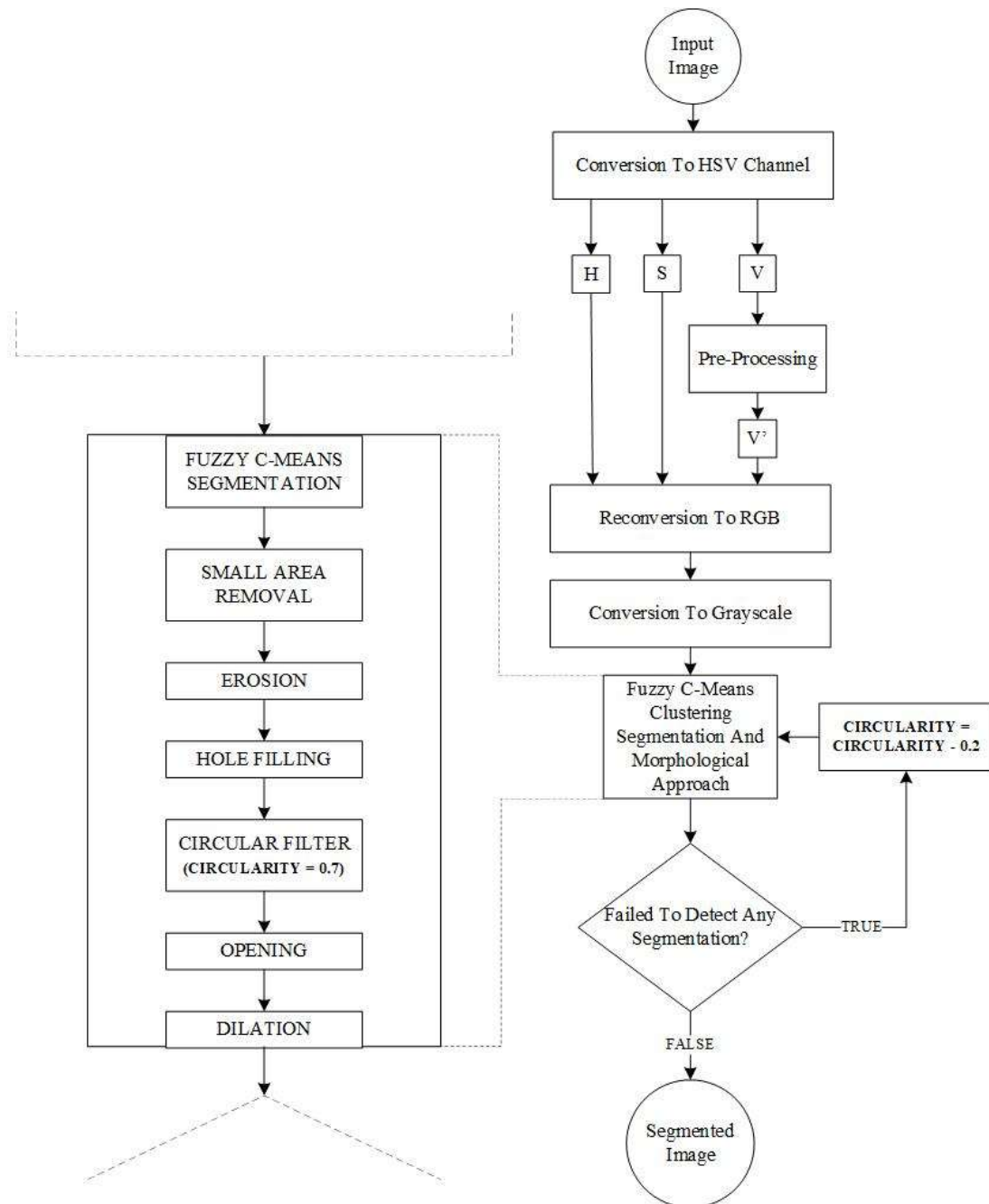


Figure 3: Overall Flowchart of Proposed Segmentation

The core segmentation method employs a multilayer approach combining Fuzzy C-Means clustering, region properties filtering based on circularity, and selected morphological operations. This includes several steps: Fuzzy C-Means segmentation, small area removal, erosion to eliminate noise, hole filling in the segmented regions, region properties filtering based on circularity (initially set to 0.7, and if segmentation fails, the circularity index is reduced by 0.2 and the process repeats until the circularity index becomes zero), morphological opening to smooth the image, and dilation to enhance the segmented regions. Fuzzy C-Means (FCM) clustering algorithm is employed by start to partition the Pap smear image into clusters, with  $k$  set to 5, fuzziness parameter,  $p$  to default value of 2 and threshold criteria,  $\epsilon$ , set to default value of 0.01 as shown in Table 1.  $\epsilon$  refers to a value used to determine when the FCM algorithm should stop iterating. The value of  $k$  is based on five regions: nucleus, cytoplasm, white background, light noise, and dark noise.

Table 1: Selected Original Images

Parameter	Value
Fuzziness parameter, $p$	2
Cluster Number, $k$	5
Threshold criteria, $\epsilon$	0.01

Following the segmentation, circularity is checked using region properties filtering with an initial value of 0.2. The process includes a decision point to determine if any segmentation was detected. If segmentation fails, further steps are performed, such as adjusting circularity to 0.7 and reapplying the segmentation techniques until successful. This iterative process, often referred to as a multilayer fuzzy C-Means on selected images, ensures thorough segmentation of the input image, leveraging various techniques to refine the final output.

The core segmentation method employs a multilayer approach that integrates Fuzzy C-Means (FCM) clustering, geometric property filtering based on circularity, and advanced morphological operations to achieve accurate segmentation. The methodology consists of the following steps: FCM segmentation, small area removal, erosion to eliminate noise, hole filling to ensure region continuity, region properties filtering based on circularity, morphological opening for smoothing, and dilation to enhance the segmented regions.

The FCM clustering algorithm is initially employed to partition the Pap smear image into 5 clusters( $k$ ), representing five distinct regions: nucleus, cytoplasm, white background, light noise, and dark noise. The fuzziness parameter ( $p$ ) is set to the default value of  $p=2$ , while the threshold criteria ( $\epsilon$ ) is set to 0.01. The clustering is governed by the minimization of the objective function:

$$J_m = \sum_{i=1}^N \sum_{j=1}^k u_{ij}^m \|x_i - c_j\|^2 \quad (3)$$

Where  $u_{ij}$  is the degree of membership of data point  $x_i$  in cluster  $j$ ,  $m$  is the fuzziness exponent controlling the degree of cluster overlap.  $c_j$  is the center of cluster  $j$  and  $\|x_i - c_j\|$  denotes the Euclidean distance between  $x_i$  and  $c_j$ . The algorithm iterates until the change in cluster centers falls below  $\epsilon$ .

Following the segmentation, geometric properties are used to filter regions based on circularity. The circularity ( $C$ ) is calculated as:

$$C = \frac{4\pi \cdot A}{p^2} \quad (4)$$

where  $A$  is the area and  $p$  is the perimeter of a region. Initially, Threshold  $C$  is set to 0.7. If segmentation fails to detect regions, the circularity threshold is reduced by 0.2 iteratively until it reaches zero. This adaptive adjustment ensures regions of varying geometries are considered. Several morphological operations are applied to refine the segmented regions:

1. Small Area Removal: Regions with areas below a specified threshold are removed to eliminate noise.
2. Erosion: Morphological erosion is performed to remove spurious pixels along boundaries, defined as:

$$A \ominus B = \{z \in E \mid (B)_z \subseteq A\} \quad (5)$$

where A is the image, B is the structuring element, and  $\ominus$  denotes erosion.

3. Hole Filling: Any holes within segmented regions are filled using connected component labeling to ensure continuity.
4. Morphological Opening: Applied to smooth regions and remove small artifacts, defined as:

$$A \circ B = (A \ominus B) \oplus B \quad (6)$$

where  $\circ$  denotes opening,  $\ominus$  is erosion, and  $\oplus$  is dilation




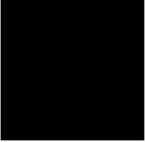









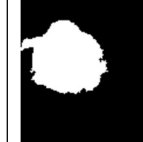






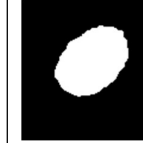
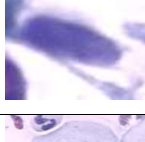






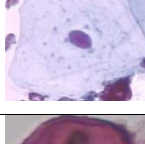





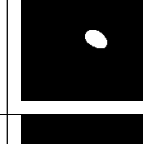
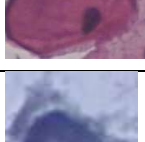



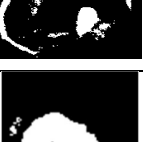









5. Dilation: Finally, dilation is used to enhance the segmented regions, making them more distinct, defined as:

$$A \oplus B = \{z \in E \mid (B)_z \cap A \neq \emptyset\} \quad (7)$$

## RESULTS AND DISCUSSION

This section employs two types of analysis, visual and quantitative, to assess the approaches. In visual testing, the subjective perception of human eyesight is used to quantify image quality. Humans have a decent understanding of image quality because these investigations are limited to evaluating image quality visually. They are applied to images from the previously described data sets. Table 2 features the selection of seven cases for each class image. The first column depicts the original image before applying any of the methods. The second column shows the ground truth of the segmentation image. After using four different segmentation methods, the second to fifth columns are the resulting images. The segmentation process is repeated to 917 image datasets from all the classes. The average of the results will be discussed in the quantitative analysis.

Table 2: Selected Original Images

ORIGINAL	ADAPTIVE THRESHOLD	CANNY EDGE BASED	CHAN-VESE	FUZZY C-MEAN	PROPOSED	GROUND TRUTH
						
						
						
						
						
						
						

The results reveal that all selected images are visually segmented differently depending on the approaches implemented. The result of the segmentation method can be observed in Table 2. In choosing the best method, the method must be able to differentiate the nucleus and the cytoplasm; it also has to be immune to noise. Next, an exemplary method must segment the images with low contrast or low-light features. The proposed methods were applied to 917 images from all classes. Table 2 shows the result of images from each class and compares four different methods: adaptive thresholding, Canny edge-based, Chan-Vese (CV), and FCM

segmentation. Based on the result, adaptive thresholding and Canny edge-based segmentation are very sensitive to the noise for all images, resulting in the noise being included as an object after segmentation. On the other hand, CV and FCM are able to segment the object definitively and better in preventing noise segmentation compared to adaptive thresholding and Canny edge-based. However, by comparing to the ground truth, both methods cannot accurately segment the object of interest. For example, the CV method is unable to segment the nucleus entirely due to the contrast sensitivity. At the same time, this method is also unable to segment any object at all, as observed in the first image; next, by referring to the sixth image, CV and adaptive thresholding are unable to segment the nucleus of interest from the cytoplasm, resulting the segmentation of cytoplasm as an object. Canny edge-based, and FCM segmentation is able to segment the nuclei. However, the method also included another nucleus due to the contrast sensitivity. As observed, the proposed method is able to differentiate between the nucleus and cytoplasm; secondly, it is immune to noise presence; and lastly, it is able to select the nucleus of interest to be segmented while immune to the presence of the other nucleus and debris with similar intensity of pixels. From Table 2, it can be shortly concluded that the proposed method is able to differentiate and segment the nucleus better compared to the other traditional segmentation methods.

An analysis of image quality assessment (IQA) will evaluate the segmented image. Image Quality Assessment (IQA) is a process to determine the value of accuracy, sensitivity, specificity, f-measure and precision. The IQA can be used to monitor the image quality of the segmented image. IQA involves some calculations to obtain the value of the accuracy sensitivity, specificity, f-measure, and precision. The calculation consists of true positive pixels (TP), false positive pixels (FP), false negative pixels (FN), and true negative pixels (TN). Below are the mathematical formulae for accuracy, sensitivity, specificity, f-measure and precision:

$$Accuracy = \frac{\sum TP + \sum TN}{\sum TP + \sum TN + \sum FP + \sum FN} \times 100 \quad (8)$$

$$Sensitivity = \frac{\sum TP}{\sum TP + \sum FN} \times 100 \quad (9)$$

$$Precision = \frac{\sum TP}{\sum TP + \sum FP} \times 100 \quad (10)$$

$$\text{Specificity} = \frac{\sum TN}{\sum TN + \sum FP} \times 100 \quad (11)$$

$$F1 - \text{measure} = \frac{2 \cdot \sum TP}{2 \cdot \sum TP + \sum FP + \sum FN} \times 100 \quad (12)$$

Each measurement parameter result defines a specific meaning, and it is important to correlate all of the results without interpreting each result individually, as the analysis will be less meaningful. For example, a high percentage of sensitivity means the method was able to segment the object successfully (TP) but neglected the falsely segmented pixel (FP), as shown in the equation (Eq. 9), while precision (Eq. 10) only focused on positive segmented neglecting the negative segmented pixel. Accuracy (Eq. 8) defining overall pixel, however, in this study, might suggest that the accuracy will always be high as the nucleus, which is the object of interest, is small and the background, which is easier to segment, will always be dominant.

Quantitative analysis was conducted after applying segmentation methods to the original images. Table 3 shows the results of the evaluation using five methods: adaptive thresholding, Canny edge-based, Chan-Vese, FCM, and the proposed methods. The data on accuracy, sensitivity, precision, specificity, and F1-measure for the five methods were collected and recorded. The study found that the proposed method has the highest accuracy of 92.19% of the other four methods, which is 90.33% for the FCM method, 66.42% for the adaptive thresholding method, 64.92% for the CV method, and the lowest of 60.35% for the Canny edge-based method. Next, the proposed method yields the highest average sensitivity value of 93.38%, which is closely followed by the FCM method at 93.34%, the CV method at 77.88%, the adaptive thresholding method at 61.32%, and the Canny edge-based method at the lowest, at 58.37%. The adaptive thresholding method also has the highest average precision value of 98.37%. On the other hand, the other four methods managed to obtain average precision values of 96.41% for the proposed method, 93.70% for the Canny edge-based method, 93.93% using the FCM, and 85.19% for the CV method. The adaptive thresholding method yields the highest specificity average of 97.60%. The proposed method comes in second with 94.25%, followed by the Canny edge-based method with 85.53%, the FCM method with 86.24%, and the CV method with the lowest specificity average of 62.73%. The proposed method yielded the highest average value of 94.40% for the F1-measure, followed by the FCM method at 93.22%,

the CV method at 76.22%, the adaptive thresholding method at 74.09%, and the Canny edge-based method at 69.41%.

Table 3: Comparison Segment Image between Five Methods

	Accuracy	Sensitivity	Precision	Specificity	F1-measure
Adaptive Thresholding	66.42	61.32	<b>98.37</b>	<b>97.60</b>	74.09
Canny Edge-Based	60.35	58.37	93.70	85.53	69.41
Chan-Vese	64.92	77.88	85.19	62.73	76.22
Fuzzy C-Means (FCM)	90.33	93.34	93.93	86.24	93.22
<b>Proposed Method</b>	<b>92.19</b>	<b>93.38</b>	96.41	94.25	<b>94.40</b>

The results show that the proposed method outperformed the other four methods in accuracy. Adaptive thresholding has the highest precision and specificity scores, with over 95%. However, the method's sensitivity is very low, less than 70%, this suggests it is conservative in predicting nuclei pixels, so misses a lot of true nuclei, but the ones it does predict are likely to be correct. Fuzzy C-Means has good performance, close to the proposed method. However, the proposed method edges it out in all metrics. Canny edge detection has high precision but poor accuracy and sensitivity with lower than 65%. As an edge detector, it is likely marking nucleus boundaries but unable to get full nuclei regions. Overall, the proposed method produced scores greater than 90%, with the highest accuracy, sensitivity, and F1-measure score. Finally, by analyzing the relationship between qualitative and quantitative analysis, CV, FCM, and proposed methods can solve the problem of noise based on the impact on the standard deviation value by using a range of grey level values in the images may cause the high value of sensitivity compared to adaptive thresholding and FCM segmentation. However, the CV method could not distinguish between the cytoplasm and the nucleus due to a slight difference in contrast in the grey-level image cause the lowest accuracy and sensitivity



comparatively to the proposed method and FCM. Though FCM performed well in segmenting the nucleus, there was still some noise in the background that was segmented as an object of interest, which explains why most of the FCM's average score was slightly lower than the proposed method. Noise may reduce the performance of Canny edge-based and adaptive thresholding. In contrast, low contrast in the original image may make it difficult for the FCM method to segment the object accurately. To overcome the noise and contrast issues, the proposed method incorporates pre-processing.

## CONCLUSION

This study examined a variety of automated nucleus segmentation methods. There are three primary stages: image acquisition, pre-processing and post-processing. The proposed methods were evaluated alongside four other traditional segmentation methods. This study used the median filter and PAGCHE in pre-processing to remove noise and improve colour contrast. To maintain colour, the original red, green, and blue (RGB) images were transformed into the HSV colour space, and the V channel is isolated for enhancement before being converted back to RGB and grayscale for post-processing. The proposed pre-processing method combines multilayer FCM, region properties filtering and mathematical morphology segmentation. Each layer has different fine-tuned parameter values. When selecting the best segmentation method, it must be able to differentiate between the nucleus and the cytoplasm and be noise-resistant. Next, a good method must be capable of segmenting images with low contrast distribution. In quantitative analysis, the accuracy, sensitivity, specificity, precision, and F1 score must be high and consistent throughout to ensure the method's robustness. Despite debris and a low-contrast image, the proposed method produces the best visual and quantitative segmentation results. The average score of the 917 images segmented by the proposed method is greater than 90%, demonstrating the proposed method's flexibility.

Nevertheless, the proposed work does have certain limitations. The methods employed may not be universally applicable to all image types or noise levels, which could affect their effectiveness in various contexts. Additionally, the relatively small dataset size constrains the robustness of the findings. Using a larger dataset would provide a clearer understanding of how well the proposed method performs across different conditions. Future research should explore

more advanced techniques and a wider range of datasets to improve the robustness and generalizability of the proposed methods.

### ACKNOWLEDGEMENT

This research was supported by funding from the Ministry of Higher Education (MoHE) Malaysia under the Fundamental Research Grant Scheme (FRGS/1/2021/SKK0/UNIMAP/02/1).

### REFERENCES

- Alias, N. A., Mustafa, W. A., Jamlos, M. A., Nasrudin, M. W., Mansor, M. A. syarafi, & Alquran, H. (2022). Edge Enhancement and Detection Approach on Cervical Cytology Images. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, 28(1). <https://doi.org/10.37934/araset.28.1.4455>
- Arya, M., Mittal, N., & Singh, G. (2020). Three segmentation techniques to predict the dysplasia in cervical cells in the presence of debris. *Multimedia Tools and Applications*, 79(33–34). <https://doi.org/10.1007/s11042-020-09206-9>
- Bandyopadhyay, H., & Nasipuri, M. (2020). Segmentation of Pap Smear Images for Cervical Cancer Detection. *2020 IEEE Calcutta Conference, CALCON 2020 - Proceedings*. <https://doi.org/10.1109/CALCON49167.2020.9106484>
- Bataineh, B. (2023). Brightness and Contrast Enhancement Method for Color Images via Pairing Adaptive Gamma Correction and Histogram Equalization. *International Journal of Advanced Computer Science and Applications*, 14(3), 124–134. <https://doi.org/10.14569/IJACSA.2023.0140314>
- Fujita, H. (2020). AI-based computer-aided diagnosis (AI-CAD): the latest review to read first. In *Radiological Physics and Technology* (Vol. 13, Issue 1). <https://doi.org/10.1007/s12194-019-00552-4>
- Glotsos, D., Kostopoulos, S., Ravazoula, P., & Cavouras, D. (2018). Image quilting and wavelet fusion for creation of synthetic microscopy nuclei images. *Computer Methods and Programs in Biomedicine*, 162. <https://doi.org/10.1016/j.cmpb.2018.05.023>
- Guimarães, Y. M., Godoy, L. R., Longatto-Filho, A., & Dos Reis, R. (2022). Management of Early-Stage Cervical Cancer: A Literature Review. In *Cancers* (Vol. 14, Issue 3). <https://doi.org/10.3390/cancers14030575>
- Hameed, K. A. S., Banumathi, A., & Ulaganathan, G. (2015). Performance evaluation of maximal separation techniques in immunohistochemical scoring of tissue images. *Micron*, 79. <https://doi.org/10.1016/j.micron.2015.07.013>

- Haridas, S., & Jayamalar, T. (2023). A Versatile Detection of Cervical Cancer with i-WFCM and Deep Learning based RBM Classification. *Journal of Machine and Computing*, 3(3). <https://doi.org/10.53759/7669/jmc202303022>
- Huang, Y., & Zhu, H. (2020). Segmentation of Overlapped Cervical Cells Using Asymmetric Mixture Model and Shape Constraint Level Set Method. *Mathematical Problems in Engineering*, 2020. <https://doi.org/10.1155/2020/3728572>
- Imtiaz, T., Fattah, S. A., & Kung, S. Y. (2023). BAWGNet: Boundary aware wavelet guided network for the nuclei segmentation in histopathology images. *Computers in Biology and Medicine*, 165. <https://doi.org/10.1016/j.compbiomed.2023.107378>
- Li, Y., Zeng, X., Han, L., & Wang, P. (2010). Two coding based adaptive parallel co-genetic algorithm with double agents structure. *Engineering Applications of Artificial Intelligence*, 23(4), 526–542. <https://doi.org/10.1016/J.ENGAPPAI.2009.04.004>
- Macios, A., & Nowakowski, A. (2022). False Negative Results in Cervical Cancer Screening— Risks, Reasons and Implications for Clinical Practice and Public Health. In *Diagnostics* (Vol. 12, Issue 6). <https://doi.org/10.3390/diagnostics12061508>
- Madhloom, H. T., Kareem, S. A., Ariffin, H., Zaidan, A. A., Alanazi, H. O., & Zaidan, B. B. (2010). An automated white blood cell nucleus localization and segmentation using image arithmetic and automatic threshold. *Journal of Applied Sciences*, 10(11). <https://doi.org/10.3923/jas.2010.959.966>
- Mayala, S., & Haugsøen, J. B. (2022). Threshold estimation based on local minima for nucleus and cytoplasm segmentation. *BMC Medical Imaging*, 22(1). <https://doi.org/10.1186/s12880-022-00801-w>
- Mustafa, W. A., Halim, A., & Ab Rahman, K. S. (2020). A narrative review: Classification of pap smear cell image for cervical cancer diagnosis. In *Oncologie* (Vol. 22, Issue 2). <https://doi.org/10.32604/ONCOLOGIE.2020.013660>
- Mustafa, W. A., Halim, A., Jamlos, M. A., & Idrus, S. Z. S. (2020). A Review: Pap Smear Analysis Based on Image Processing Approach. *Journal of Physics: Conference Series*, 1529(2). <https://doi.org/10.1088/1742-6596/1529/2/022080>
- Mustafa, W. A., Wei, L. Z., & Ab Rahman, K. S. (2021). Automated cell nuclei segmentation on cervical smear images using structure analysis. *Journal of Biomimetics, Biomaterials and Biomedical Engineering*, 51. <https://doi.org/10.4028/www.scientific.net/JBBBE.51.105>
- Nahrawi, N., Mustafa, W. A., Kanafiah, S. N. A. M., Ahmad, W. K. W., Rohani, M. N. K. H., & Rahim, H. A. (2021). A Novel Nucleus Detection on Pap Smear Image Using Mathematical Morphology Approach. *Journal of Biomimetics, Biomaterials and Biomedical Engineering*, 49, 53–61. <https://doi.org/10.4028/www.scientific.net/JBBBE.49.53>

- Plissiti, M. E., Nikou, C., & Charchanti, A. (2011). Automated detection of cell nuclei in Pap smear images using morphological reconstruction and clustering. *IEEE Transactions on Information Technology in Biomedicine*, 15(2).  
<https://doi.org/10.1109/TITB.2010.2087030>
- Prasad Battula, K., & Chandana, B. S. (2022). Deep Learning based Cervical Cancer Classification and Segmentation from Pap Smears Images using an EfficientNet. In *IJACSA) International Journal of Advanced Computer Science and Applications* (Vol. 13, Issue 9). [www.ijacsa.thesai.org](http://www.ijacsa.thesai.org)
- Rajarao, C., & Singh, R. P. (2020). Improved normalized graph cut with generalized data for enhanced segmentation in cervical cancer detection. *Evolutionary Intelligence*, 13(1).  
<https://doi.org/10.1007/s12065-019-00226-5>
- Sausen, D. G., Shechter, O., Gallo, E. S., Dahari, H., & Borenstein, R. (2023). Herpes Simplex Virus, Human Papillomavirus, and Cervical Cancer: Overview, Relationship, and Treatment Implications. In *Cancers* (Vol. 15, Issue 14).  
<https://doi.org/10.3390/cancers15143692>
- Schiffman, M., & de Sanjose, S. (2019). False positive cervical HPV screening test results. In *Papillomavirus Research* (Vol. 7). <https://doi.org/10.1016/j.pvr.2019.04.012>
- Shilpa, Gopakumar, R., & Acharya, V. (2021). Nucleus Segmentation from Microscopic Bone Marrow Image. *Advances in Intelligent Systems and Computing*, 1198.  
[https://doi.org/10.1007/978-981-15-6584-7\\_5](https://doi.org/10.1007/978-981-15-6584-7_5)
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*, 71(3).  
<https://doi.org/10.3322/caac.21660>
- Wan, T., Xu, S., Sang, C., Jin, Y., & Qin, Z. (2019). Accurate segmentation of overlapping cells in cervical cytology with deep convolutional neural networks. *Neurocomputing*, 365.  
<https://doi.org/10.1016/j.neucom.2019.06.086>
- William, W., Ware, A., Basaza-Ejiri, A. H., & Obungoloch, J. (2019). Cervical cancer classification from Pap-smears using an enhanced fuzzy C-means algorithm. *Informatics in Medicine Unlocked*, 14. <https://doi.org/10.1016/j.imu.2019.02.001>
- Wu, T. C., Wang, X., Li, L., Bu, Y., & Umulis, D. M. (2021). Automatic wavelet-based 3D nuclei segmentation and analysis for multicellular embryo quantification. *Scientific Reports*, 11(1). <https://doi.org/10.1038/s41598-021-88966-2>
- Yang, Y., Cao, Y., & Shi, W. (2014). A method of leukocyte segmentation based on S component and B component images. *Journal of Innovative Optical Health Sciences*, 7(1).  
<https://doi.org/10.1142/S1793545814500072>

Zhao, J., Li, Q., Li, X., Li, H., & Zhangs, L. (2019). Automated segmentation of cervical nuclei in pap smear images using deformable multi-path ensemble model. *Proceedings - International Symposium on Biomedical Imaging, 2019-April*.  
<https://doi.org/10.1109/ISBI.2019.8759262>

Zhao, M., Wang, H., Han, Y., Wang, X., Dai, H. N., Sun, X., Zhang, J., & Pedersen, M. (2021). SEENS: Nuclei segmentation in Pap smear images with selective edge enhancement. *Future Generation Computer Systems, 114*.  
<https://doi.org/10.1016/j.future.2020.07.045>