

Image Processing for Automatic Cell Nucleus Segmentation Using Super pixel and Clustering Methods on Histopathological Images

Firhouse Hassan Ahamed Shibly¹

¹Senior Lecturer in Information Technology, South Eastern University of Sri Lanka.
Email: shiblymis@gmail.com

Lakshmana Kumar. R²

²Assistant Professor, Department of CSE, SNS College of Technology, Coimbatore-641035, India
E-mail: veerandrakumarr9@gmail.com

Abstract

On a daily basis, it appears that the number of cancer cases and cancer-related deaths are increasing. Early detection and treatment of the malignant region are critical for successful treatment. Early detection of sick cells is made possible with the use of computer-assisted programmes, which are then diagnosed by experienced pathologists due to their efforts. Using computer-aided programmes, this research found that cell nuclei could be automatically detected in high-resolution histopathological images using global segmentation methods such as k-Means and Fuzzy C Means and algorithms from superpixel segmentation methods such as SLIC, Quick-shift, Felzenszwalb, Watershed, and ERS. Using high-quality histopathology pictures, the researchers discovered that the k-means and FCM algorithms performed significantly better than the baseline techniques in the study. In terms of precision, the Quickshift and SLIC approaches produced superior outcomes. The K-means and FCM algorithms perform best in the F-M test, and the true negative ratio is more successful in the Quickshift and SLIC methods than in the F-M test.

Keywords: Image Processing, Histopathological Image, Machine Learning, Fuzzy C Means algorithm, k-Means algorithm

1. Introduction

Cancer is still the leading cause of death globally. Cancer affects people of all ages and backgrounds. In Iraq, thousands of people get cancer each year, including men and women. In Iraq country has a lower cancer incidence. Women get breast cancer at a rate of one in four, compared to one in ten men [1]. This allows for early cancer detection and treatment. Cancer awareness grows as countries progress. CT, MR, and ultrasound can now detect suspicious cells. High-resolution histopathology images aid in early cancer detection. Pathology is vital in these assessments. In

Greek, pathology means disease. Pathology studies disease mechanisms and organ morphology. It studies the cellular changes that occur in infected organs. They examine cells under a microscope to diagnose diseases. A pathologist takes a biopsy of the suspect tissue. Microscope examines it after fixation and tracking. It uses expertise and consultation to diagnose illness. Sharp imaging has made skilled pathologists' jobs easier. Like the microscope, digital image processing is useful. Only a few people can use the microscope at once, preventing distant pathologists from inspecting the tissue [2].

Hundreds of pathologists can examine the tissue simultaneously [5]. These images are made by magnifying and digitizing human tissue. Medical images with composite structures and information are histopathological images. Expert pathologists examine organs and tissues in the cell to diagnose the disease: expertise influences tissue or organ size, patient risk, and diagnosis. Pathologists can automate tedious tasks using image processing software. Pathologists have to examine cells one by one. Examination of cells one by one can be difficult and tiring for expert pathologists. By using computer-aided diagnosis (CAD) systems [3], this process, which is difficult and tiring by expert pathologists, can be carried out in software with technological devices. The computer-aided diagnosis system aims to analyze the digitized image in the computer environment and create decision support for diagnosis.

Histopathology is a branch of science that examines the microscopy of lesions occurring in the natural structure of organs, tissues and cells [4]. With the development of computer technology, fast processing and saving time have increased the importance of computer technology. Digital cameras are used to transfer medical images to digital media. Images taken with a microscope are called histopathological images.

1.1 Segmentation of Histopathological Images:

Effects of Color Spaces on Cell Segmentation: Color spaces in the literature are a model used to express colours mathematically. Colour spaces are formed to represent all colours found in nature. Colour diversity is provided by changing the three main colours in certain proportions in colour spaces [6-7].

- RGB (Red, Green, Blue) colour space: The colour space is widely used to express colours in devices such as scanners and monitors. Mixing the three basic colours in a certain ratio allows the formation of different colours. Mixing these basic colours at 100% creates white, and mixing them at 0% creates black [7].
- HSV (Hue, Saturation, Value) colour space: HSV colour space consists of Hue, Saturation and Value variables. Hue is the quality that distinguishes a colour from other colours according to its type or tone, and Saturation indicates how saturated the colour is used. On the other hand, value is used to distinguish whether the colour is light or dark
- XYZ colour space: XYZ colour space was designed by CIE in 1931 and constituted the basic structure of all colours. XYZ colour space is obtained by transforming the RGB colour space into a coefficient matrix. The sum rate of the impulses of the nerves transmitting the RGB colours indicates X, Y, Z colours.

- LAB colour space: It was created by CIE in 1976, and it is aimed to create an easier expression of colours. The LAB colour spaces components stand for L (Luminance), a (Gradation) and b (Saturation). The most important feature of LAB colour space is its uniform distribution in human perception. For this reason, LAB colour space is used in histopathological image analysis studies.
- Image Processing: Image processing is a technique for converting a physical image into a digital format that can then be processed. As a form of input, images such as photographs or video clips. A portion of the image you specify appears in the output [6]. Images are typically treated as two-dimensional (2D) signals and processed using predetermined signal processing methods by the image processing system. After obtaining the digital image, the next step is preprocessing. In these processes, the acquired image is made clearer to be processed to the next stage easily and without errors. Some of these stages are;
 - Making the image clear
 - Filtering impurities in the image
 - Eliminating or minimizing the structural defects on the image In the field of image processing medicine, certain techniques are applied to the images obtained from devices such as CT, MR or microscope to obtain a certain region.

1.2 Review of Literature

According to Mittal, et al. (2018) [9], the segmentation technique is a multi-step procedure that includes image enhancement, core region extraction, core centroid marking, core area enhancement, and complex core separation. The proposed system also included a new performance measure as a cumulative factor.

Albayrak, et al. (2018) [8] compared various machine learning and deep learning models on the CRC colon cancer dataset in their study. Because traditional machine learning algorithms, such as convolutional neural networks, do not take a two-dimensional input, they have benefited from local binary images. It is stated in the conclusion section that the K-nearest neighbour and RF algorithms produce excellent results, while the CNN algorithm produces better results without feature extraction.

Saturi, et al. (2020) [10] used deep learning methods during the treatment phase to conduct a study and analyze it in order to predict side effects in patients diagnosed with breast cancer over the age of 65. As a result of the research, doctors will be able to decide whether or not to treat breast cancer patients over the age of 65 with chemotherapy. As a result, they were able to achieve a 74 per cent accuracy rate.

Information security is mentioned in websites and applications in Cuadros Linares et al.,(2020) [11] studies.

The datasets they obtained and the original sized datasets were compared using k nearest neighbour, support vector machines, and extreme vector machines, as well as various metrics like the success rate of the models with the best performance for the test data, sensitivity false alarm rate, and so on. They claim that extreme learning machines can be easily integrated into online intrusion detection systems and used as a backup method.

2. Materials and Method

The Beck Laboratory at Harvard University provided the data for our investigation. High-resolution histopathological images of kidney cancer from TCGA (The Cancer Genome Atlas) data are included in this collection. Large-scale TCGA is supported by the NHGRI and the American National Cancer Institute of the National Human Genome Research Institute. TCGA has about ten thousand comprehensive molecular studies in 25 common cancer types. With this situation, it collects the whole image from its participants within the scope of the study. It is seen that TCGA constitutes the main source for projects in computational pathology by examining the molecular, morphological and clinical features of cancer disease together [12]. Histopathological images were cropped image sections of 400×400 dimensions. A total of 810 kidney cell carcinomas were obtained, 81 of which were present on each slide. Cellular structures in the image were marked by expert pathologists, and 64 histopathological images were used. Performance comparison of images tagged by pathologists and images obtained by segmentation methods was made. Three histopathological images were obtained from the Harvard University Beck Laboratory renal cell carcinoma data set, and the reference images of these images are given in Figure 2.1.

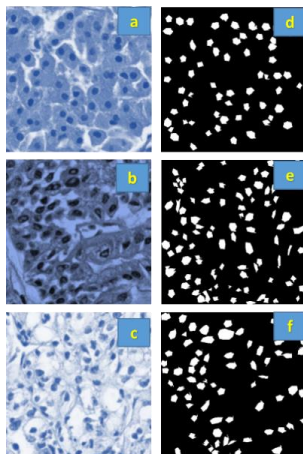


Figure 2.1 (a), (b) and (c) high-resolution images obtained in the dataset. Reference images of (d), (e) and (f) images marked by expert pathologists

2.2.1 Segmentation Methods Used: Segmentation is the process of identifying the relevant data region. During segmentation, some clustering algorithms are used. Clustering groups similar objects. Hard and fuzzy clustering are the two basic clustering methods. Hard clustering assigns data to a separate cluster for processing. This means that each data point is assigned to only one cluster when it clusters the image. Overlapping structures, noise, and low contrast reduce the method's effectiveness [11]. Soft clustering is a method used successfully in image segmentation. Fuzzy clustering algorithms are more resilient to uncertainty. The fuzzy c means method is the most widely used fuzzy clustering method [13].

2.2.2 k-Means Method: The K-means method (kMM) divides N data into c clusters by minimizing the distance between the data and the cluster. MacQueen proposed it in 1967. The unsupervised algorithm is the simplest. It has two stages. First, 'c' cluster centres are chosen at random. Next, each data centre is assigned to the nearest cluster. The cluster distance is calculated using Euclidean and Manhattan calculations. Repetition of the process until new cluster centres are found and reach the minimum level [14].

K – Steps performed in the means algorithm;

- Determination of the number of clusters (c)
- Determination of the maximum value (ϵ) for the stopping condition
- Cluster centre is randomly determined
- Euclidean distance is calculated separately for each data
- The calculated data is assigned to the cluster closest to it.
- If the condition is not met, the process continues. If so, the clustering process is terminated.

In the k-means clustering algorithm, the computation time increases as the number of data and clusters increases [14].

$$b_n = \sum_{j=1}^N \max(d_{k-1}^j - \|x_n - x_j\|^2, 0) \dots (Eq.1)$$

$$i = \arg \max_n b_n \dots (Eq.2)$$

d_{k-1}^j in Equation 1 is the square distance between the nearest cluster centres located between the cluster centres. In other words, it is the cluster centre of x_j . The b_n value aims to reduce the error by placing a new cluster centre at the x_n location. x_j is the cluster centres obtained so far (Likas et al., 2003). In the k-means clustering algorithm, the computation time increases as the number of data and clusters increase.

2.2 Methods used in the Current Study

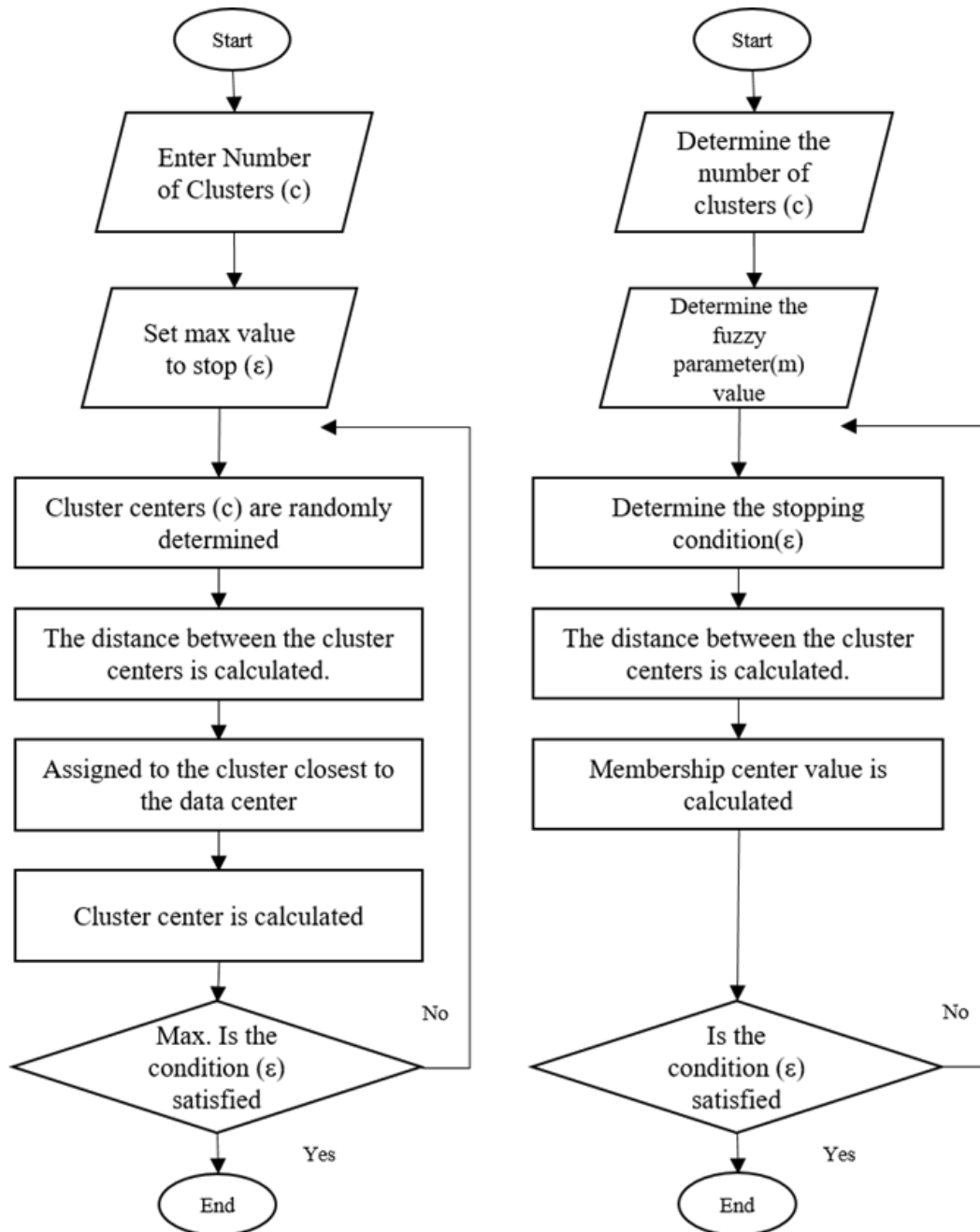


Figure 4.2 K-Means Flow Chart Figure 4.3 Fuzzy C-Means Flow Chart

2.2.3. Fuzzy C-Means Method (FCM)

Dunn proposed the FCM algorithm in 1972, and Bezdek developed it. Granath has reported several FCM algorithm variants [15]. Used in many image processing algorithms. It is widely used because it is not sensitive to noise. The Fuzzy C-Means algorithm assigns data to multiple clusters. A membership degree allows more than one cluster to be owned. The degree of membership is determined by taking values between 0 and 1. The algorithm creates a membership matrix (U0) with random values. The stopping conditions are the threshold value (ϵ) and the number of clusters (c). The fuzzy parameter m is determined. In the fuzzy c Means method, the cluster centre is calculated as in equation 3. u_{ij} is the membership value of the pixel in the set i .

$$C_j = \frac{\sum_{i=1}^n (u_{ij}^k)^m \cdot j_i}{\sum_{i=1}^n (u_{ij}^k)^m} \dots (Eq .3)$$

The membership matrix is calculated as in equation 4 using U^{k+1}

$$u_{ij}^{k+1} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij}}{d_{kj}} \right)^{\frac{2}{m-1}}} \dots (Eq .4)$$

$$d_j = \|y_j - c_i\|^2 \dots (Eq .5)$$

The process continues until the condition $\max_{ij} \|u_{ij}^k - u_{ij}^{k+1}\| < \epsilon$ is satisfied. After the transaction is fulfilled, the centre returns to the matrix.

The steps performed in the Fuzzy C-Means algorithm;

- The number of clusters (c) and the fuzzy parameter value (m) are determined
- The threshold value (ϵ), which is the stopping condition, is determined
- Calculate the distance between cluster centres and pixels
- Membership centre value is calculated
- Suppose the difference between the membership matrix and the calculated membership matrix as a result of the calculations is higher than the first determined value. In that case, it is repeated from the 3rd step.

2.2.4. Superpixel Method for Segmentation of Histopathological Images

Detecting cancerous cells and structures in histopathological images requires accurate segmentation. Experts need to segment correctly to detect and diagnose. This study evaluated the superpixel method's success in segmenting high-resolution histopathological images [16].

In high-resolution image segmentation, each pixel is semantically labelled. Each pixel's texture is grouped by properties like colour and density. Superpixels are these groups. Superpixels are meaningful superpixels in

histopathological images. Thus, complex processes are avoided and reduced processing time.

2.2.5. Simple Linear Iterative Clustering (SLIC) Method

SLIC Achanta et al. SLIC's superpixel segmentation algorithm uses k-means to cluster pixels based on colour and coordinate information [17]. Superpixels provide computation speed, memory savings, and ease of use when preprocessed with SLIC.

The SLIC-Segnet algorithm has two parts: encoder and decoder (decoder). Encoder classification uses the vgg16 algorithm. The decoder increases the image size to match the input image. The operation performed by the segmentation method for each pixel is given in equation 6.

$$d_c = \sqrt{(I_j - I_i)^2 + (a_j - a_i)^2 + (b_j - b_i)^2} \dots (Eq .6)$$

In Equation 7, it is stated how the coordinate information of each pixel is calculated. The j value specified in the equation is the pixel accepted as the center, and the i value is the value to be clustered.

$$d_s = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \dots (Eq .7)$$

It is the realization of the assignment of the relevant pixel calculated in Equation 8 to a centre.

$$D' = \sqrt{\left(\frac{d_c}{N_c} \right)^2 + \left(\frac{d_s}{N_s} \right)^2} \dots (Eq .8)$$

The maximum expected spatial distance within a given cluster should correspond to the sampling interval (N_s). SLIC superpixels correspond to clusters on the lab colour image plane. Determining the maximum colour distance (N_c) will not be easy, as colour distances will vary significantly from image to image and from cluster to cluster. This situation can be avoided by fixing N_c to a fixed value of m .

$$D = \sqrt{(d_c)^2 + \left(\frac{d_s}{S} \right)^2 (m)^2} \dots (Eq .9)$$

The m value in Equation 9 also enables the determination of the relative importance between colour similarity and spatial proximity [17].

2.2.6 Quick Shift Segmentation Method

Quickshift is a well-known superpixel segmentation method. It is an iterative algorithm that identifies data in a set of data points. Any field can use Quickshift. It distinguishes modes when all data points are connected. Various data clusters are separated [18]. This method creates superpixels by changing the fast transition parameters based on recognition. Neighbors are considered within a spatial distance with a Gaussian kernel. Then, a tree is formed

connecting each image pixel (x, y) to the nearest neighbor (x', y') with a higher intensity value, as shown in the equation.

$$P(X_i) = \frac{1}{N} \sum_{j=1}^N k(D(x_i, x_j)), x_i, x_j \in R^d \dots (Eq. 10)$$

N is data, and $k(x)$ is the kernel function with a Gaussian window. The expression $D(x_i, x_j)$ is the distance between x_i and x_j . Using the Euclidean distance in the spatial and colour space, the distances are calculated by the equation in 11.

$$d_q(i, j) = d_{xy}(i, j) + w_q \cdot d_{lab}(i, j) \dots (Eq. 11)$$

w_q is the weighting parameter in the equation, and the smaller the weighting parameter, the more important the spatial area [17-18].

2.2.7. Felzenszwalb Segmentation Method (FSM)

Felzenszwalb is a local variation graph algorithm. Because the algorithm has no density constraints, it typically produces skewed regions [19]. The algorithm is fast and works almost linearly with time. The segmentation method can ignore detail in high-variability images while retaining detail in low-variability images. The input of a graph-based algorithm; It is a graph $G=(V, E)$ with n vertices and m sides.

Its output is the division of V into $S=(C_1, C_2, \dots, C_r)$ components. $C \subseteq V$ is expressed as the internal difference of the component, the maximum weight that covers the component minimum, $MST(C, E)$. Well;

$$Int(C) = \max_{e \in MST(C, E)} W(e) \dots (Eq. 12)$$

C_1 , the difference between the components C_2 and V are defined as the minimum weight edge connecting the two components.

$$dif(C_1, C_2) = \min_{vi \in C_1, vj \in C_2, (vi, vj) \in E} W(vi, vj) \dots (Eq. 13)$$

2.2.8 Watershed Segmentation Method

With Watershed, the image gradient is calculated with good adherence to the object boundaries and allows control over the number and spatial arrangement of the resulting image (Machairas et al., 2014). The algorithm requires a grayscale gradient image to express the borders. It interprets the image's bright pixels as a hilly landscape. A new image segment is created from the high hills and floodplains. Lalitha et al.(2016) [20] A grayscale image is $G=(D, E, F)$. Only critical points in a connected domain D form functions on D if F is a continuous space $C(D)$. The distance between the points is calculated as in the following equation

$$Tf(p, q) = \inf_{\lambda} f_{\lambda} \parallel \nabla f(\lambda(s)) \parallel ds \dots (Eq. 14)$$

2.2.9 ERS Segmentation Method

It uses a graphics-based approach to find compact and homogeneous superpixels, unlike other well-known superpixels like SLIC. The ERS has two parts: entropy and balancing. These are homogeneous clusters, and the balancing term provides similar-sized clusters [21]. The number of clusters (k) and weight factor (w_{ij}) are adjusted in the ERS algorithm. The original subset (A) has superpixel segmentation [14]. Entropy measures the uncertainty of a random variable (H). The probability mass function $p(x)$ and the entropy of a discrete random variable X are calculated as in the equation

$$H(X) = - \sum_{x \in X} p(x) \log p(x) \dots (Eq. 15)$$

2.3 Metrics Used to Evaluate Segmentation Performance

In the experiments, some measurements were used to determine operational performance. These; True Positive Ratio, True Negative Ratio, Precision, precision, Overlap Ratio, False Positive Ratio, False Negative Ratio are the measurements used to evaluate FM segmentation performance [13]

Table 4.1 Metrics and Equations

Metric	Equation
Precision	$\frac{TP}{TP + FP}$
True Positive Ratio (TPR)	$\frac{TP}{TP + FN}$
F-Measure (F-M)	$\frac{Two \times precision \times TPR}{precision + TPR}$
True Negative Rate (TNR)	$\frac{TN}{TN + FP}$
Overlap Rate (OR)	$\frac{A(S) \cap A(G)}{A(S) \cup A(G)}$

$A(S)$ is the area of the segmented image, and $A(G)$ is the area of the real image in the table above. True positive (TP), true negative (TN), false positive (FP), false negative (FN), precision, and recall are the parameters specified in the

formulas. True positive refers to the number of pixels in the cellular structure area that are correctly affected as positive samples, whereas true negative refers to the number of pixels outside of the cellular structures that are negatively affected.

True Positive Ratio- (TPR) is given in equation 2.1, which expresses the ratio of positively labelled pixels to all pixels of the kernel, according to their belonging to the kernel.

$$TPR = \frac{TP}{TP + FN} \dots (Eq. 2.1)$$

True Negative Ratio- TNR (True Negative Ratio- TNR) expresses the ratio of pixels that do not belong to the core but that is marked as negative to all non-core pixels. It is given in Equation 2.2.

$$TNR = \frac{TN}{TN + FP} \dots (Eq. 2.2)$$

Precise Hit (KI) is the ratio of true positives to positive predictions. It is given in Equation 2.3.

$$KI = \frac{TP}{TP + FP} \dots (Eq. 2.3)$$

F-Measure is calculated as a joint function of Precise Accuracy and True Positive Rate.

$$F - M = \frac{2 \times KI \times TNR}{KI + TNR} \dots (Eq. 2.4)$$

Overlap-O (Overlap) is calculated as in equation 5.5.

$$O = \frac{TP}{TP + FP + FN} \dots (Eq. 2.5)$$

3.0 Results and Discussion

3.1. Segmentation Results

In the study, 64 high-resolution histopathological images of renal cell carcinoma obtained at Harvard University Beck Laboratory were performed using seven different segmentation methods. The images obtained as a result of segmentation were evaluated with the performance criteria using the reference image. Some preprocessing steps were applied to the histopathological images and then segmentation processes were applied. After preprocessing, global methods such as kMeans, Fuzzy C Means, superpixel methods, SLIC, Quick Shift, Felzenswalb, Watershed and ERS segmentation methods were used [22]. In the structure of histopathological images, there are sections such as cytoplasm, cell wall, adipose tissue and nucleus. For this reason, cluster centres such as 3,4 and 5 were selected in the k Means, and Fuzzy C Means methods, which are global segmentation methods.

3.1.1. Segmentation with K Means method

The K Means segmentation method used clusters 3,4,5 and 6 for segmentation. When the number of clusters was 2, 3, or 4, the nuclear structure could be segmented as nuclei, but it could not when the number of clusters was 4. It has been found that 5 clusters perform better segmentation. Figure 3.1 shows the performance with 5 clusters.

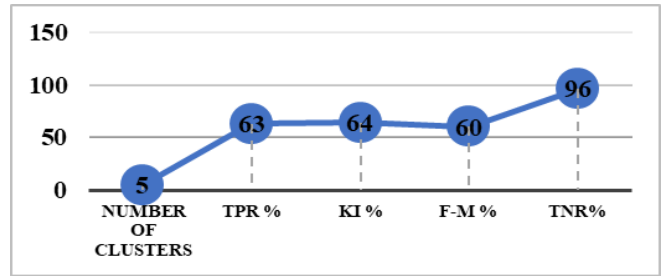


Figure 3.1 Performance Criteria Values of Images Segmented by K-Means Method

3.1.2. Segmentation with Fuzzy C Means method

The Fuzzy C Means segmentation method used clusters 3,4,5,6 for segmentation. When the number of clusters was 2 or 3, cell walls were perceived as nuclei, but when the number of clusters was 5, the nuclear structure could not be segmented as nuclei. It has been found that 4 clusters perform better segmentation. Figure 3.2 shows the performance with 4 clusters.

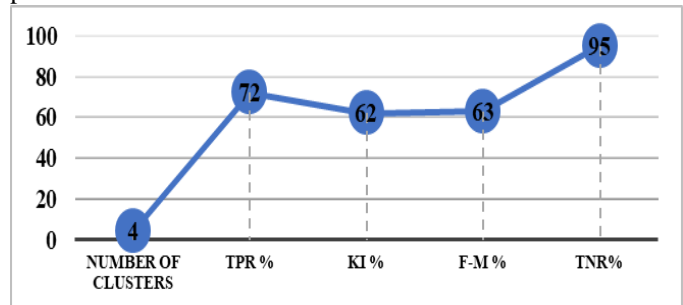


Figure 3.2 Performance Criteria Values of Images Segmented by Fuzzy C Means Method

3.2.2. segmentation with the SLIC method

In the SLIC segmentation method, it is seen that it gives better results when divided into 3000-pixel groups. It has been seen that as the number of pixels and compact value increase in the parameters, it gives better results. It has been observed that the performance remains constant after a certain value. The parameters used and their results are shown in Figure 3.1.

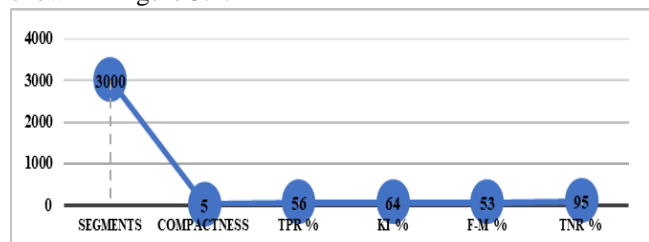


Figure 3.3 Performance Criteria Values of Images Segmented by SLIC Method

The quick shift segmentation method principle is the algorithm that identifies the data in a series of data points and does it iteratively. The segmentation method has two main parameters. The performance value according to these two parameters is given in Figure 3.3.

3.1.4. segmentation with the Quick Shift method

The quick shift segmentation method principle is the algorithm that identifies the data in a series of data points and does it iteratively. The segmentation method has two main parameters. The performance value according to these two parameters is given in Figure 3.4.

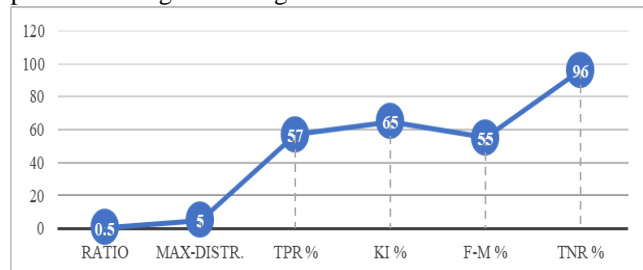


Figure 3.4: Performance Criteria Values of Images Segmented by Quick Shift Method

3.1.5. Segmentation with the Felzenszwalb method

The actual size and number of segments in the Felzenszwalb algorithm can vary greatly depending on local contrast. The algorithm scale parameter has a single parameter that affects the segment size. The performances obtained with the parameters used are given in figure.3.5.

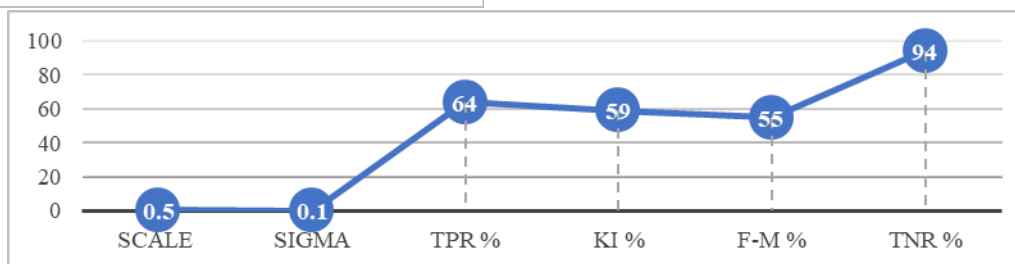


Figure 3.5 Performance Criteria Values of Images Segmented by Felzenszwalb Method

3.1.6. segmentation with the Watershed method

Two different parameters, markers and compactness, were used in the Watershed segmentation algorithm. It is seen that

the performance value increases as the markers value increases. However, it seems to remain constant after a certain value.

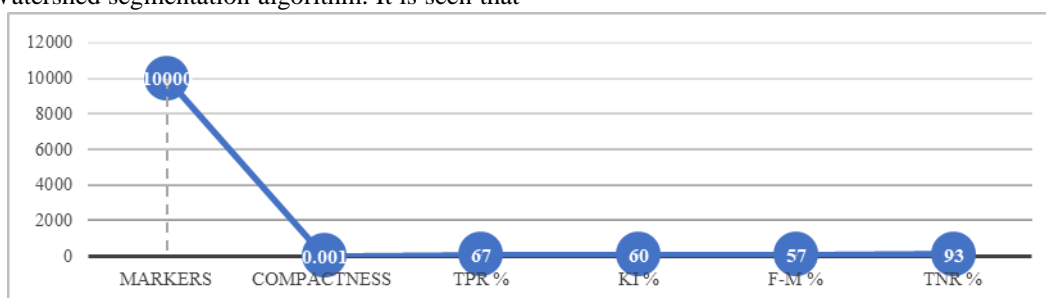


Figure 3.6 Performance Criteria Values of Images Segmented by Watershed Method

In this study, apart from the methods used in the literature, Felzenszwalb and Watershed methods were used. It is seen that the true positive rate gives better results in K-Means and Fuzzy C-Means segmentation methods.

4. Conclusions and Recommendations

Conclusions: Automatic detection of unhealthy cellular structures on high-resolution histopathological images is of great importance for cancer detection. Different segmentation methods of 64 histopathological images were used. Some parameters were used to evaluate the results

obtained. This study evaluated the performance performances of segmentation methods for automatic cell detection on high-resolution histopathological images. Performances of k-Means, Fuzzy C Means and superpixel segmentation algorithms (SLIC, Watershed, Quickshift, Felzenszwalb, ERS) were obtained on 64 high-resolution histopathological images.

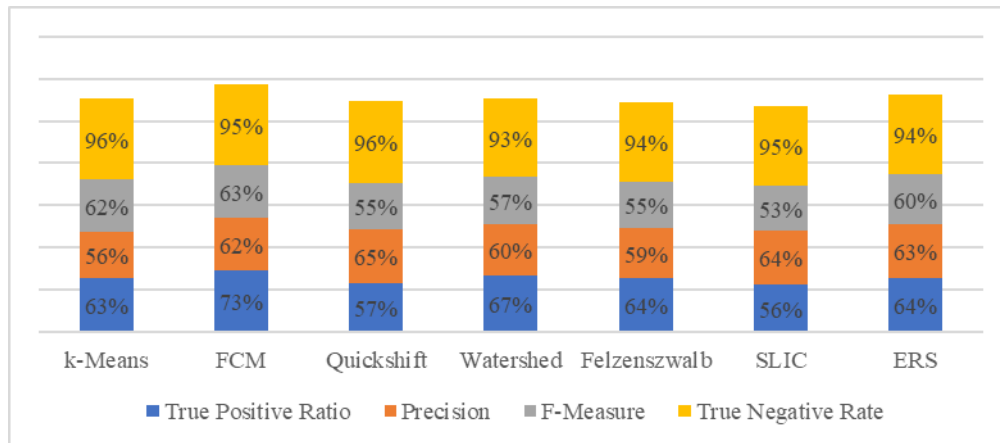


Figure 4.1 Performance Performances of the Segmentation Algorithms Used

Figure 4.1 shows that seven segmentation methods were employed in the study. According to this study, the k-Means and FCM algorithms perform better in high-resolution histopathological images. In terms of precision, the Quickshift and SLIC methods performed better. Algorithms using k-Means and FCM provide the best results for the F-M, while Quickshift and SLIC methods do better with the true negative ratio (TNR). Since the cells marked by the pathologists are not completely marked, their performance decreases. The performance graph of the segmentation algorithms used in the graph is shown in Figure 4.1.

Recommendations: Cancer is one of the most common diseases today. Segmentation methods increase the chance of early detection of cancerous cells day by day. With the use of segmentation methods, the accuracy of the diagnostic method and the determination of the appropriate treatment method can be ensured. It is thought that this study will lead to further research for the diagnosis, diagnosis and treatment of diseases with segmentation methods. By improving the methods and techniques used, the detection of diseased cells can be concluded more successfully.

References

- [1]. Al-Akeedi, Janan & Noor, Nawar & Alhiti, Mustafa. (2020). Epidemiological study of cancers among iraqi people in oncology teaching hospital in medical city in baghdad for 2018. *European Journal of Pharmaceutical Sciences*. 7. 45-51.
- [2]. Cooper, Lee & Carter, Alexis & Farris, Alton & Wang, Fusheng & Kong, Jun & Gutman, David & Widener, Patrick & Pan, Tony & Cholleti, Sharath & Sharma, Ashish & Kurc, Tahsin & Brat, Daniel & Saltz, Joel. (2012). Digital Pathology: Data-Intensive Frontier in Medical Imaging. *Proceedings of the IEEE*. 100. 991-1003. 10.1109/JPROC.2011.2182074.
- [3]. Land, Walker & Wong, L. & McKee, D.W. & Masters, T. & Anderson, F.R.. (2003). Breast cancer computer aided diagnosis (CAD) using a recently developed SVM/GRNN Oracle hybrid. 5. 4705 - 4711 vol.5. 10.1109/ICSMC.2003.1245727.
- [4]. Musumeci, Giuseppe. (2014). Past, present and future: overview on Histology and histopathology. *Journal of Histology and Histopathology*. 1. 5. 10.7243/2055-091X-1-5.
- [5]. Fischer, Andrew & Jacobson, Kenneth & Rose, Jack & Zeller, Rolf. (2008). Hematoxylin and Eosin Staining of Tissue and Cell Sections. *CSH protocols*. 2008. pdb.prot4986. 10.1101/pdb.prot4986.
- [6]. Taqi, SyedAhmed & Sami, SyedAbdus & Sami, LateefBegum & Zaki, SyedAhmed. (2018). A review of artifacts in histopathology. *Journal of Oral and Maxillofacial Pathology*. 22. 279. 10.4103/jomfp.JOMFP_125_15.
- [7]. Xie, Juanying & Liu, Ran & Luttrell, Joseph & Zhang, Chaoyang. (2019). Deep Learning Based Analysis of Histopathological Images of Breast Cancer. *Frontiers in Genetics*. 10. 10.3389/fgene.2019.00080.
- [8]. Albayrak, Abdulkadir & Bilgin, Gökhan. (2018). Automatic cell segmentation in histopathological images via two-staged superpixel-based algorithms. *Medical & Biological Engineering & Computing*. 57. 10.1007/s11517-018-1906-0.
- [9]. Mittal, Himanshu & Saraswat, Mukesh. (2018). An automatic nuclei segmentation method using intelligent gravitational search algorithm based superpixel clustering. *Swarm and Evolutionary Computation*. 45. 10.1016/j.swevo.2018.12.005.
- [10]. Satri, Rajesh & Chand, P.. (2020). Implementation of Efficient Segmentation Method for Histopathological Images. 419-423. 10.1109/ICICT48043.2020.9112386.
- [11]. Cuadros Linares, Oscar & Soriano Vargas, Aurea & Faical, Bruno & Hamann, Bernd & Fabro, Alexandre & Traina, Agma. (2020). Efficient Segmentation of

- Cell Nuclei in Histopathological Images. 47-52. 10.1109/CBMS49503.2020.00017.
- [12]. Caicedo, Juan & Roth, Jonathan & Goodman, Allen & Becker, Tim & Karhohs, Kyle & Broisin, Matthieu & Molnar, Csaba & McQuin, Claire & Singh, Shantanu & Theis, Fabian & Carpenter, Anne. (2019). Evaluation of Deep Learning Strategies for Nucleus Segmentation in Fluorescence Images. *Cytometry Part A*. 95. 10.1002/cyto.a.23863.
- [13]. Mandloi, Ayush & Daripa, Ushnesha & Sharma, Mukta & Bhattacharya, Mahua. (2019). An Automatic Cell Nuclei Segmentation based on Deep Learning Strategies. 1-6. 10.1109/CICT48419.2019.9066259.
- [14]. Tariq, Humera & Burney, S.M.Aqil. (2014). K-Means Cluster Analysis for Image Segmentation. *International Journal of Computer Applications*. 96. 10.5120/16779-6360.
- [15]. Yang, Yong & Huang, Shuying. (2007). Image Segmentation by Fuzzy C-Means Clustering Algorithm with a Novel Penalty Term.. *Computers and Artificial Intelligence*. 26. 17-31.
- [16]. Albayrak, Abdulkadir & Bilgin, Gökhan. (2018). A Hybrid Method of Superpixel Segmentation Algorithm and Deep Learning Method in Histopathological Image Segmentation. 1-5. 10.1109/INISTA.2018.8466281.
- [17]. Kim, Kwang-Shik & Zhang, Dongni & Kang, Mun-Cheon & Ko, Sung-Jea. (2013). Improved simple linear iterative clustering superpixels. *Digest of Technical Papers - IEEE International Conference on Consumer Electronics*. 259-260. 10.1109/ISCE.2013.6570216.
- [18]. Ibrahim, Abdelhameed & Salem, Muhammed & Ali, Hesham. (2014). Automatic Quick-Shift Segmentation for Color Images. *IJCSI International Journal of Computer Science Issues*. 11. 122-127.
- [19]. Guimarães, Silvio & Kenmochi, Yukiko & Cousty, Jean & Patrocínio Jr, Zenilton & Najman, Laurent. (2017). Hierarchizing graph-based image segmentation algorithms relying on region dissimilarity: the case of the Felzenszwalb-Huttenlocher method. *Mathematical Morphology - Theory and Applications*. 2. 10.1515/mathm-2017-0004.
- [20]. K V, Lalitha & .R, Amrutha & Michahial, Stafford & M, Dr. (2016). Implementation of Watershed Segmentation. *IJARCCCE*. 5. 196-199. 10.17148/IJARCCCE.2016.51243.
- [21]. Contreras, Cecilia & Khodadadzadeh, Mahdi & Tolosana-Delgado, Raimon & Gloaguen, Richard. (2020). Drill-Core Hyperspectral and Geochemical Data Integration in a Superpixel-Based Machine Learning Framework. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. PP. 1-1. 10.1109/JSTARS.2020.3011221.
- [22]. Khan, Am & Srisha, Ravi. (2013). Image Segmentation Methods: A Comparative Study. *International Journal of Soft Computing and Engineering*. 2231-230