A Comparative Study of Segmentation Method for Computer-aided Diagnosis (CAD) Leukemia AML Subtype M0, M1, and M2

Wiharto[1]*, Wisnu Widiarto[2], Esti Suryani[3], Nurmajid Hidayatullah[4]
Department of Informatics, Universitas Sebelas Maret
Surakarta, Indonesia

Abstract—A computer-based diagnosis model for Acute Myelogenous Leukemia (AML) is carried out using white blood cell image processing. The stages in computer-aided diagnosis (CAD) are included pre-processing, segmentation, feature extraction, and classification. The segmentation method has many approaches, namely, clustering, region growing, and thresholding. The number of approaches that can be used requires proper selection because it will have an impact on CAD performance. This study aims to conduct a comparative study of the performance of the WBC segmentation method on the AML M0, M1, and M2 subtype leukemia CAD system. The segmentation algorithm used is k-means, fuzzy c-means, SOM, watershed, chan vese (active contour), otsu thresholding, and histogram. The feature extraction method uses GLCM, while the classification algorithms tested are SVM, Random-forest, decision tree, naive Bayesian, and k-NN. The test results show that the histogram segmentation method is able to provide the best average performance when using SVM, namely 90.3% accuracy, 85.9% sensitivity, and 92.7% specificity.

Keywords—Acute myelogenous leukemia; leukemia; segmentation; feature extraction; classification

I. INTRODUCTION

Leukemia is a blood cancer caused by the body producing too many abnormal white blood cells. Leukemia can occur in both adults and children. White blood cells are part of the immune system produced in the bone marrow. When the function of the spinal cord is disturbed, the white blood cells produced will change and no longer perform their role effectively. Leukemia has several types, including Acute myeloblastic leukemia (AML). The AML is acute myeloblastic leukemia that occurs when the bone marrow overproduces immature myeloid cells or myeloblasts. The diagnosis of leukemia AML can be made with a white blood test (WBC) for analysis. WBC analysis can be done computerized, namely by photographing the WBC, so that WBC is obtained in the form of a digital image.

One way to develop a leukemia diagnosis system model is by using a computer-based diagnostic model. This diagnostic system model is done by analyzing WBC. This analysis has a number of advantages compared to existing models, namely the process is faster. Processing is carried out in a number of stages. The stages in computer-based diagnosis include preprocessing, segmentation, feature extraction, and classification. The preprocessing stage is used to improve

image quality so that it is ready to continue with the segmentation process. At the segmentation stage, it is done to separate the object from the background. The segmentation process can use a number of approaches, namely Region growing, edge detection, thresholding, and clustering [1]–[3][4]–[6]. Each approach in segmentation also has many algorithms that have advantages and disadvantages. The next stage after segmentation is feature extraction. Feature extraction has a number of approaches, namely color, texture, statistical, and geometry [7]–[9]. The next stage is the classification process, which can be done using a number of algorithm choices, such as SVM and decision tree [10].

Many developments of AML leukemia CAD models have been carried out [11]–[14], but unfortunately many still produce low related performance. One of the factors causing the low performance is the wrong choice of the segmentation method used. To overcome this, there have been a number of studies that have also compared segmentation performance but only limited to segmentation algorithms on one approach, such as clustering [15] and thresholding [13], [16]. This makes the best algorithm only limited to that approach and has not been carried out for inter-approach. This condition causes the lack of literature that can be used to determine the best segmentation algorithm for WBC segmentation in AML leukemia, so a comparative study of a number of segmentation algorithms from several segmentation approaches is needed, such as thresholding, clustering, and region growing.

II. LITERATURE REVIEW

The segmentation approach for white blood cell (WBC) image analysis has advantages and disadvantages, so the selection of the segmentation method will affect the performance of the AML leukemia diagnosis system. The segmentation method used in the diagnosis system for leukemia AML subtypes M0, M1, and M2 mostly uses thresholding [11], [12], [14]. The weakness of using thresholding is in determining the threshold value, which is done manually so that sometimes it is not suitable for different image conditions. The weakness of thresholding with a static threshold value can be overcome by a dynamic threshold value, namely multi-Otsu thresholding. Multi-Otsu thresholding can provide good performance compared to static thresholding in the case of AML subtype M0 and M1 classification [13]. Another image segmentation method that is also widely used is active contour without edge combined

*Corresponding Author.

with watershed distance transform which is used for segmentation of AML subtypes M2, M3, and M4[17].

Another alternative approach to segmentation is to use a clustering algorithm. The concept of clustering is the same as the concept of segmentation, which separates the background and objects. Clustering will automatically form a cluster center depending on the data, and from the data center, it can be used to threshold the binary image formation process [18], [19]. A study conducted by Dhanachandra et al. [4], showed that the k-mean and subtractive clustering algorithms can be used for segmentation with good results. Good ability in segmentation is also shown in the case of segmentation in the case of AML subtype M4, M5, and M7, which uses the k-mean algorithm [20]. The k-mean algorithm is also used in image segmentation in cases of AML M2, M3, M4, and M5 leukemia diagnosis, and can give good results [21].

Referring to a number of studies that have been carried out to develop leukemia CAD models, at the segmentation stage, in choosing the segmentation method used, it is not explained in detail what the considerations are. Referring to the research conducted by Arumugadevi et al. [22] shows that the segmentation approach using clustering shows that the segmentation performance using FCM and SOM algorithms provides better performance than K-means. When referring to clustering performance using FCM and SOM, FCM is able to provide better clustering performance than SOM [23]. This shows that the selection of the segmentation method is very influential in producing the performance of the CAD system.

The low performance of the AML subtype diagnosis systems M0, M1, and M2 is not only influenced by the lack of accuracy in choosing the segmentation method, but also by the feature extraction method used for diagnosis. The diagnosis of AML M0, M1, and M2 developed in the study of Suryani et al. [24] using the features of WBC diameter, nucleus ratio, and nucleus roundness. These features are only able to provide an accuracy that is still below 80%. The analysis in this study showed that only WBC diameter could be used for features in the diagnosis of AML subtypes M0 and M1. Leukemia AML subtypes M0 and M1 are diagnosed by referring to what blast cells are dominant, each cell has different characteristics in size, shape, and color. A study conducted by Mutlag et al. [9], from feature extraction testing using leaf images, showed that the texture approach was able to provide better accuracy performance than the geometric, color, and statistical approaches. This is as shown in the research of Rawat et al. [25], where the feature extraction Gray level co-occurrence matrices (GLCM) is better than shape-based (geometry). GLCM capability is better than Local binary pattern (LBP) when combined with SVM classification algorithm, with histopathological Specimen image data [26].

Referring to previous research, this study aims to conduct a comparative study of segmentation performance with clustering, region growing, active contour, thresholding, and histogram approaches. The segmentation algorithm was used to segment WBC in cases of CAD leukemia AML subtypes M0, M1 and M2. The AML leukemia CAD model uses the GLCM feature extraction method. The performance of the

CAD system is measured using the parameters of accuracy, sensitivity, and specificity.

III. METHOD

This study uses WBC image data taken from The Hospital of Dr. Moewardi, Surakarta Indonesia. The WBC images analyzed were AML subtypes M0, M1, and M2. The total AML data used are 105 WBC images, consisting of 33 M0, 32 M1, and 40 M2. This research method uses a number of stages as shown in Fig. 1 where the main process is divided into fourparts, namely pre-processing, segmentation, feature extraction, classification, and performance evaluation. At the segmentation stage, three segmentation approaches will be used as shown in Table I. The segmentation algorithm used for each approach is K-means, Fuzzy c-Means, SOM, watershed, chan vese (active contour), Otsu thresholding, and histogram.

TABLE I.   SEGMENTATION APPROACH AND ALGORITHMS

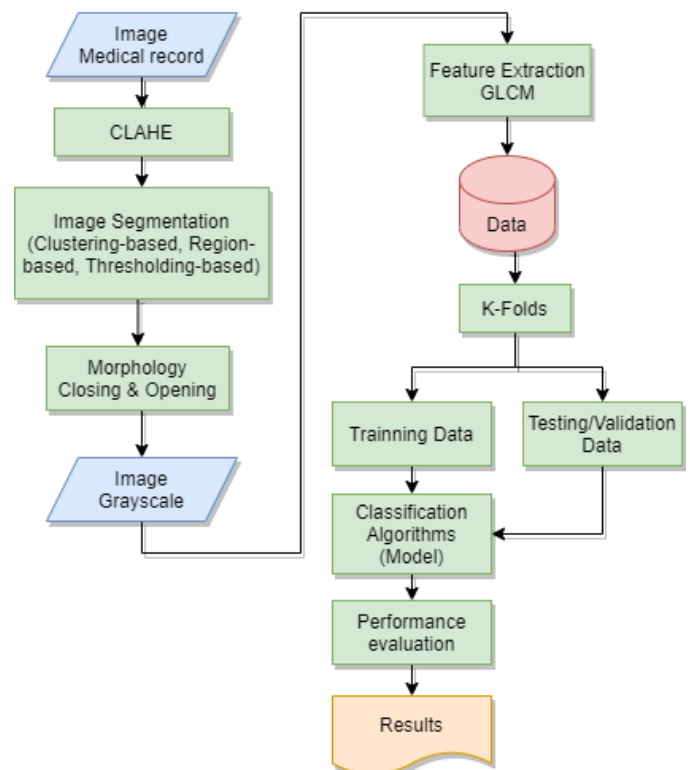| No	Approach	Algorithms	References
1	Clustering-based	K-Means	[27]
		Fuzzy C-Means	[27]
		Self-Organizing Maps	[23]
2	Region-based	Chan Vese (Active Contour)	[28], [29]
		Watershed	[30]
3	Thresholding-based	Otsu Thresholding	[31], [32]
		Histogram	[33]–[35]

Fig. 1.   Research Method.

The next step is feature extraction using Gray Level Co-occurrence Matrix (GLCM). GLCM is a texture analysis technique on the image. GLCM represents the relationship between neighboring 2-pixels that have grayscale intensity, distance, and angle. Distances (d) are expressed in pixels, while angles (θ) are in degrees. The distance between pixels is usually one pixel. The GLCM has 8 angles, including angles 0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315°, in this study using angles 0°, 45°, 90° and 135° [25], [36], [37]. The resulting features for each angle are Contrast, Dissimilarity, Homogeneity, Energy, Correlation, and ASM [38]. These features can be shown in equations (1-7).

$$P(i,j) = NGLCM(i,j) \tag{1}$$

$$P(i,j) = \frac{GLCM(i,j)}{\sum_{i=0}^{N}\sum_{j=0}^{M} GLCM(i,j)}$$

$$Contrast = CT = \sum_{i=0}^{N}\sum_{j=0}^{M}(i-j)^2 \, P(i,j) \tag{2}$$

$$Dissimilarity = DS = \sum_{i=0}^{N}\sum_{j=0}^{M}|i-j| \, P(i,j) \tag{3}$$

$$Homogenity = HG = \sum_{i=0}^{N}\sum_{j=0}^{M}\frac{P(i,j)}{1+(i-j)^2} \tag{4}$$

$$ASM = \sum_{i=0}^{N}\sum_{j=0}^{M}[P(i,j)]^2 \tag{5}$$

$$Energy = EG = \sqrt{ASM} \tag{6}$$

$$Correlation = CR = \sum_{i=0}\sum_{j=0}\frac{(i-\mu i)(j-\mu j)P(i,j)}{\sigma i \, \sigma j} \tag{7}$$

The next step is to classify using a number of algorithms. The classification algorithms tested are SVM, Random Forest, decision tree, k-NN, and Naive Bayesian. The classification algorithm is applied to every angle in the feature extraction process with GLCM, the angles used are 0°, 45°, 90°, and 135°. The test is carried out using k-folds cross-validation, with the performance parameters measured are sensitivity, specificity, and accuracy. Performance calculation is done by referring to Table II, and by equation (8-10).

TABLE II. CONFUSION MATRIX

Actual Class	Predictive Class	
	Positive	Negative
Positive	TP	FN
Negative	FP	TN

$$Sensitivity = SEN = \frac{TP}{TP+FN} \tag{8}$$

$$Specificity = SPE = \frac{TN}{TN+FP} \tag{9}$$

$$Accuracy = ACC = \frac{TP+TN}{TP+FN+FP+TN} \tag{10}$$

IV. RESULTS

A. Result of Preprocessing

In the AML leukemia CAD system, before segmenting the WBC image, the steps taken are to improve the quality of the WBC image. Improvements were made using contrast limited adaptive histogram equalization (CLAHE) [39]. CLAHE is used to enhance the color and appearance of blurry objects in an image. The results of preprocessing using CLAHE can be shown in Fig. 2.
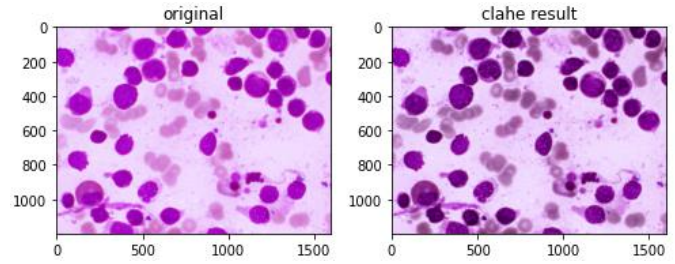

Fig. 2. Result of Preprocessing.

B. Result of Segmentation

This study uses three segmentation approaches, namely, cluster-based, region-based and thresholding-based. The segmentation algorithm for each approach is shown in Table I. In segmentation using the k-mean algorithm, the first step after the pre-processing process is to convert the image to grayscale. Segmentation is carried out with a maximum iteration value of 100 and epsilon 0.2. The clustering process carried out resulted in a total of four clusters. To separate the object with the background from the image resulting from clustering with four clusters, the separation process is carried out using the quartile value of the centroid. The quartile-1 value is used as the threshold value, if the tested pixels value is smaller than Quartile 1 than the centroid value, it is made white and vice versa. The results of the segmentation are then carried out by morphological closing and morphological opening processes [40] to smooth and eliminate noise in the image segmentation results. The complete results of the segmentation process with K-means are shown in Fig. 3. The concept for clustering using fuzzy c-means and SOM is almost the same as the k-means algorithm.
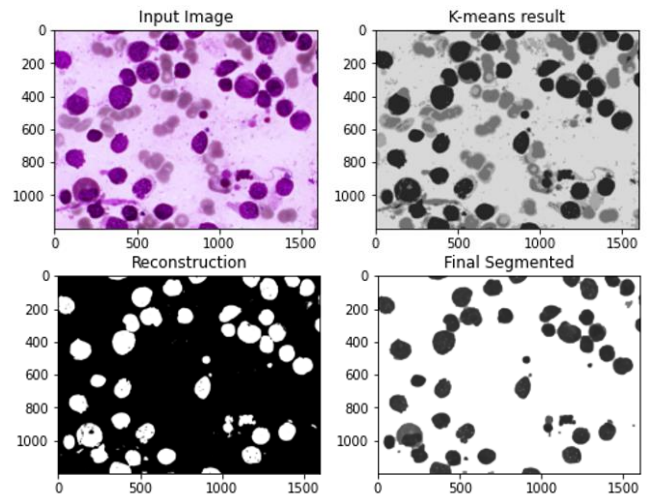

Fig. 3. The Results of the Segmentation Process with k-means.

The next segmentation model is using the thresholding approach. One of the algorithms in this approach is the histogram algorithm. In this algorithm, before the histogram process is carried out, the conversion to grayscale is done first. The grayscale image is then carried out by the histogram process so that it will produce the histogram value and bin

edges or the scale of the histogram boundary. In this study, 10 bin edges were used to form the histogram. From each edge, the middle value is searched for each edge and with this value, the quartile 1 value is sought, then used as the threshold value in separating the object from the background. The results of the segmentation are then carried out by morphological closing and morphological opening processes [40] to smooth out and eliminate noise in the image segmentation results. The results of segmentation using histograms can be shown in Fig. 4.

The next segmentation approach is a region growing. The segmentation algorithm used is Chan-Vese (active contour) [26],[27]. This algorithm is designed to group objects without clear boundaries. This algorithm is based on a set of levels developed iteratively to minimize the energy determined by a weighted value. This value corresponds to the sum of the intensity differences from the mean value outside the segmented region and a term that depends on the length of the segmented region boundary. In chan-vese active contours do not require a cropping process because of the nature of Region-based active contours that find lesions using the globalizing method [30], [41]. This is considered autonomous segmentation because the initial contour placement is the entire image and does not need to be defined [42]. The results of segmentation with the chan-vese algorithm can be shown in Fig. 5.
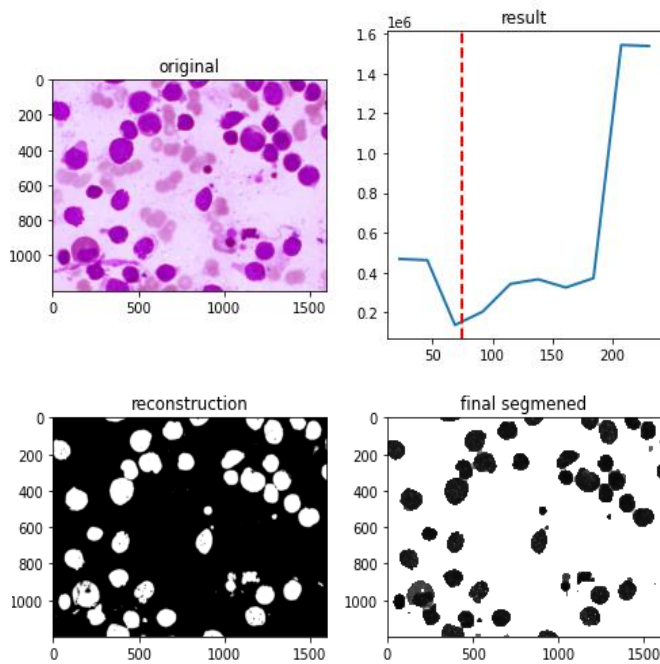

Fig. 4.   Segmentation Results with Histogram.

Another region growing-based segmentation is watershed starting with conversion to grayscale, the grayscale image will then be processed using the otsu thresholding algorithm to separate objects from the background. Furthermore, morphological reconstruction is carried out to remove noise. The next stage is to find the border using the distance transform and label the peak of the object in the image, then the image will be processed by the watershed algorithm [28].

Then the results of the reconstruction are processed using morphology closing and opening morphology to smooth and remove noise in the reconstructed image. The final stage is to use the reconstructed image to take objects in the original image and convert the image to grayscale so that it is ready to be processed in the next stage. The results of this segmentation can be shown in Fig. 6.

C. Result of Feature Extraction

Feature extraction is done using the texture approach, namely by using the GLCM method. Image segmentation results will be carried out a feature extraction process with 6 features, namely Contrast, Dissimilarity, Homogeneity, Energy, Correlation, and ASM (Angular Second Moment). The GLCM method uses angles of 0°, 45°, 90°, and 135°. The results of feature extraction using GLCM at an angle of 0°, and when using the k-mean segmentation algorithm is shown in Table III. In Table III, 5 samples of each AML subtype M0, M1, and M2 are taken.
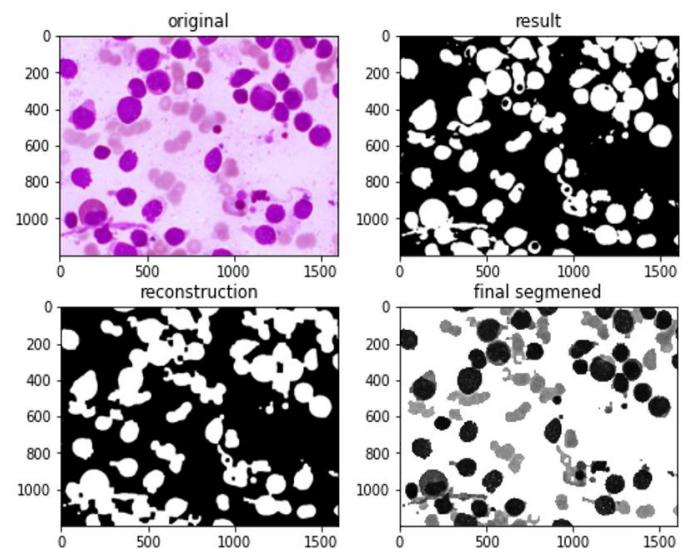

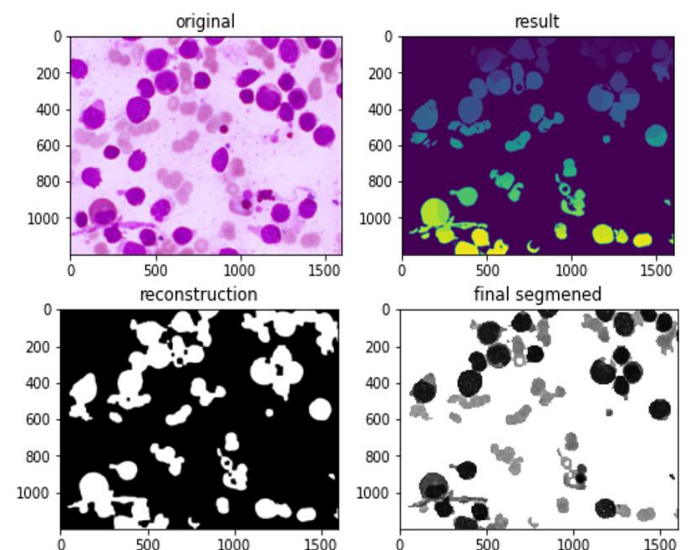Fig. 5.   Segmentation Results with Chan Verse Active Counter.

Fig. 6.   Segmentation Results with Watershed.

TABLE III.   EXAMPLE OF GLCM 0^0 FEATURE EXTRACTION RESULTS WITH K-MEANS SEGMENTATION

Type	CT	DS	HG	EG	CR	ASM
M0	1.3044	0.2308	0.9260	0.5198	0.9842	0.2702
M0	0.7966	0.1612	0.9460	0.6718	0.9882	0.4513
M0	0.5528	0.0961	0.9704	0.8087	0.9889	0.6540
M0	0.7719	0.1321	0.9603	0.7445	0.9871	0.5543
M0	0.5669	0.0987	0.9698	0.8123	0.9887	0.6599
M1	0.3386	0.0893	0.9671	0.8213	0.9911	0.6746
M1	0.4711	0.0944	0.9693	0.8097	0.9888	0.6556
M1	0.6201	0.1306	0.9564	0.7575	0.9870	0.5738
M1	0.3433	0.0727	0.9757	0.8558	0.9900	0.7324
M1	0.6437	0.1127	0.9668	0.8062	0.9839	0.6500
M2	0.5831	0.1225	0.9610	0.7209	0.9870	0.5197
M2	0.8049	0.1443	0.9568	0.6895	0.9847	0.4754
M2	0.8339	0.1592	0.9501	0.6409	0.9871	0.4107
M2	0.9491	0.1615	0.9525	0.6070	0.9865	0.3684
M2	1.0880	0.2081	0.9338	0.5733	0.9826	0.3287

D. Result of Classification

The stage after feature extraction is classification. At the classification stage, a number of classification algorithms were tested, namely support vector machine (SVM), Random Forest (RF), Decision Tree (DT), KNN, and Naïve Bayesian (NB). Performance parameters measured are accuracy, sensitivity, and specificity. Tests were carried out for each angle of the GLCM, namely angle 0°, 45°, 90°, and 135°. The test results for each angle with 5-fold cross-validation validation can be shown in Fig. 7 to Fig. 10.

Fig. 7 to Fig. 10 illustrates the performance of the CAD system when using GLCM feature extraction with the orientation angle, with values 0°, 45°, 90°, 135°. The feature extraction with GLCM will produce a co-occurrence matrix, which is a square matrix with the number of elements as much as the square of the number of pixel intensity levels in the image. Each point (p, q) in the co-occurrence matrix with orientation angle contains the probability of occurrence of a pixel worth p next door to a pixel worth q at a distance d and orientation θ and (180°−θ). If the value is 0°, then the performance of the resulting CAD model is shown in Fig. 7, for the value of 45° is shown in Fig. 8, while the value of 90° and 135° are shown in Fig. 9 and Fig. 10. Fig. 7 to Fig. 10 shows that changing the orientation angle does not affect the performance of the SVM algorithm. The orientation angle of 135° can provide relatively the same performance for all types of classification algorithms, while angles of 0°, 45°, and 90° algorithms other than SVM provide poor performance. Especially for the SVM algorithm, the best performance is given at an orientation angle of 0°, when using the clustering and histogram segmentation algorithm.

Fig. 7 (GLCM with an angle of 0°) shows that the segmentation performance with histogram and k-means gives better performance than the others. Segmentation performance with clustering approach, SOM algorithm gives the lowest performance. Segmentation with thresholding approach, histogram algorithm is better than otsu thresholding, while in region growing approach, the watershed algorithm is better than chan vese. Fig. 7 to Fig. 10 shows that the performance of the classification algorithm that gives the best performance is SVM. At 135° GLCM angle, all tested classification algorithms give a good performance, while at other GLCM angles, the good performance is dominated by the SVM algorithm, while for the others it varies.
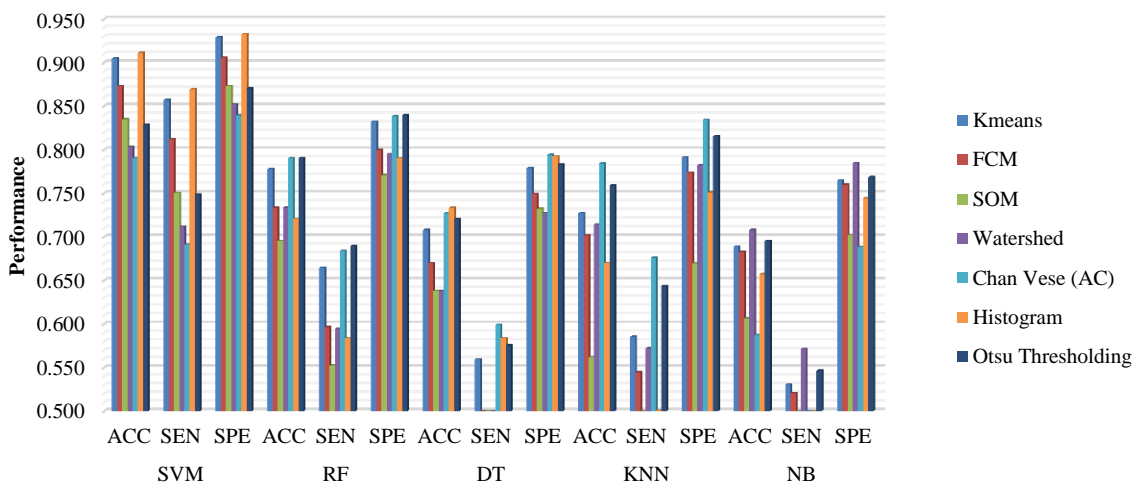

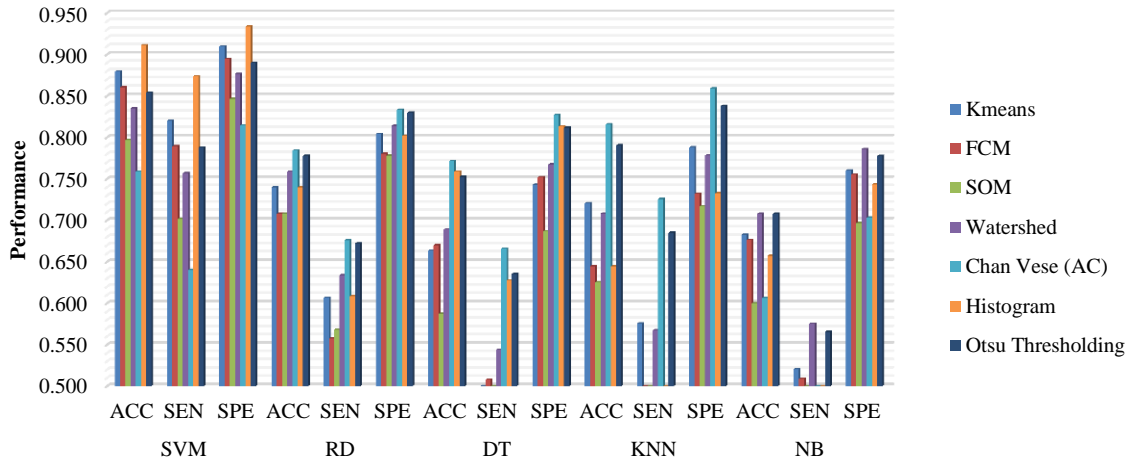Fig. 7.   Performance of Model using GLCM 0°.

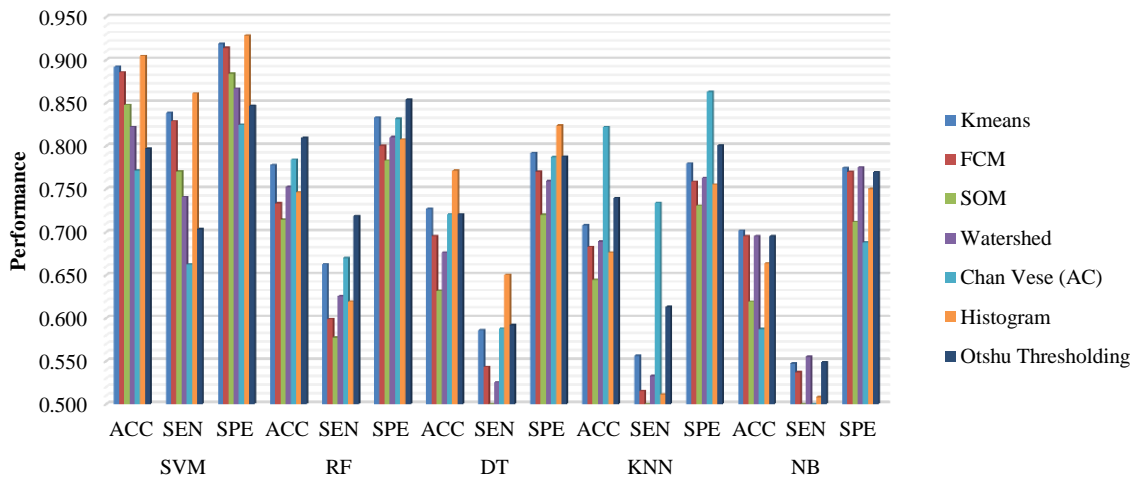Fig. 8. Performance of Model using GLCM 45°.
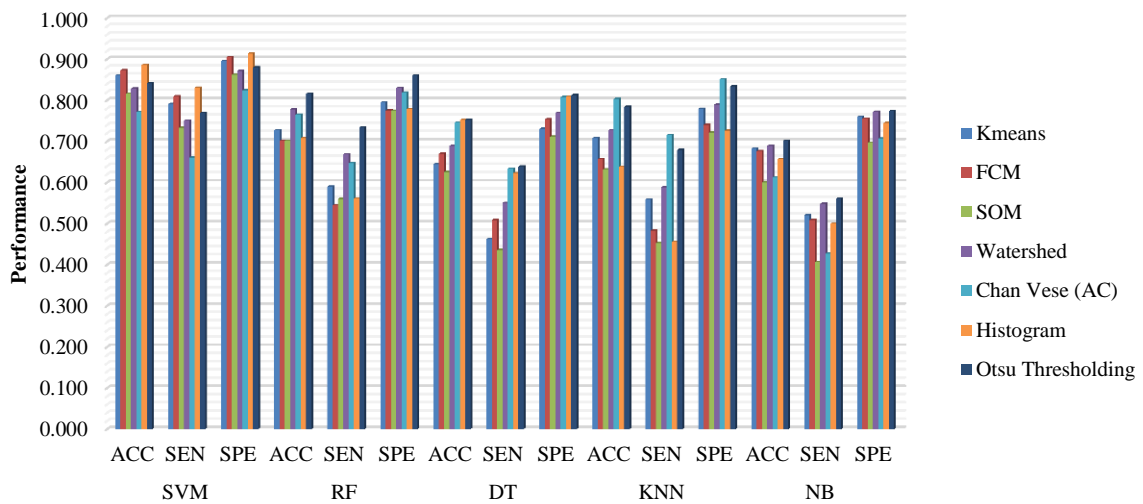

Fig. 9. Performance of Model using GLCM 90°.

Fig. 10. Performance of Model using GLCM 135°.

V. DISCUSSION

The test results of several segmentation algorithms show that the thresholding-based segmentation approach, by using the histogram method, is able to provide the best performance. The best performance is obtained, when using the SVM classification algorithm. Histogram segmentation is able to provide the best performance for all angles in GLCM when using SVM. The segmentation with the best clustering approach is given the k-means algorithm, which gives the best accuracy, for the GLCM angle of 0°, 45°, and 90°, while for the 135° angle the best performance is given by the FCM algorithm. The SOM algorithm gives the lowest performance between FCM and K-means. In the case of image segmentation of AML subtypes M0, M1, and M2, the segmentation approach that is less than optimal is shown in the region growing approach, with the chan vese (active contour) algorithm, where the performance for all GLCM angles is the lowest. Chanvese's algorithm has weaknesses, namely, it is not able to divide non-uniform regions, and is sensitive to noise [43]. The Chan-Vese algorithm also has problems in terms of deviations from the class center in the Chan-Vese model. In particular, followed by inserting the energy function into the level set evolution without a re-initialization framework, the variation formulation can force the level set function to be closed to the object boundary [44].

The segmentation of the region growing approach generally shows less than optimal performance. This is not only shown by the performance of the Chan Vese algorithm, but also by the watershed algorithm. The watershed algorithm has a number of weaknesses, namely over-segmentation, manual intervention is needed, sensitivity to noise, and poor detection of significant areas with low contrast [45]. Weaknesses of the watershed algorithm confirmed the resulting performance in the diagnosis of AML leukemia subtypes M0, M1, and M2. This is different from segmentation with a thresholding approach. Histogram segmentation performance is able to provide the best classification results in the CAD system. This ability is caused by the dynamics of histogram segmentation in determining the threshold value, it's just that the weakness is when there is grayscale whose values overlap, so it becomes inaccurate [45].

The performance of computer-aided diagnosis of AML leukemia is not only influenced by the classification algorithm but also by the feature extraction algorithm used. Feature extraction GLCM is used to obtain features from the segmentation process. Comparison of values for each feature when using the GLCM angle of 0° and using the histogram segmentation method can be shown in Table IV. Table IV shows that the mean ± STD of each feature between M0, M1, and M2 has almost the same value. Referring to Table IV and supported by statistical test results, shows several features that are not significantly different between M0, M1, and M2 (p-value>0.05). The results of statistical tests with a 95% confidence level for GLCM with an angle of 0° in the complete histogram segmentation are shown in Table IV. This condition causes the results of classification, especially using algorithms other than SVM, the results are not optimal. This means that some features cannot be used to distinguish AML subtypes of leukemia M0, M1, and M2.

TABLE IV. COMPARISON OF FEATURES ON GLCM 0^0 (HISTOGRAM)

Feature	M0	M1	M2
	Mean±STD	Mean±STD	Mean±STD
Contrast	0.870±0.248	0.543±0.229	0.797±0.242
Dissimilarity	0.159±0.044	0.114±0.043	0.144±0.043
Homogeneity	0.950±0.015	0.962±0.014	0.956±0.014
Energy	0.681±0.097	0.773±0.093	0.687±0.103
Correlation	0.987±0.002	0.988±0.003	0.986±0.002
ASM	0.472±0.127	0.606±0.138	0.482±0,146

TABLE V. STATISTICS TEST ON GLCM 0^0 (HISTOGRAM)

Feature	P-value		
	M0 x M1	M0 x M2	M1 x M2
Contrast	**0.000**	0.169	**0.000**
Dissimilarity	**0.000**	0.100	**0.007**
Homogeneity	**0.002**	0.066	0.087
Energy	**0.000**	0.598	**0.001**
Correlation	0.110	**0.035**	**0.000**
ASM	**0.000**	0.574	**0.001**

The ability of the histogram segmentation algorithm, when viewed from the results of the significance test with a 95% confidence level, is not the best. This is indicated by the comparison between M0xM1, M0xM2, and M1xM2 features, the number of different features is significantly less compared to segmentation using k-means. Feature extraction, which was preceded by segmentation using K-means, was able to produce a feature correlation to distinguish AML M0 from AML M1, while the histogram could not differentiate. Feature correlation shows the size of the linear relationship of the neighboring pixel gray-level values.

The ability of the histogram and k-means segmentation algorithms, when viewed from statistical tests on features generated from GLCM, shows that the ability to distinguish AML M0 from AML M2 is less than optimal. Referring to Table V, only the correlation features have a significant difference, while the other features are not significantly different. In addition to feature correlation, feature homogeneity is also relatively able to distinguish, compared to other features. This condition shows that there is a close texture between M0 and M2, so it is not optimal when using the six GLCM features. Overall the performance generated when using histogram segmentation with GLCM angle 0°, the area under the curve (AUC) value [46] is above 90%, which is included in the very good category [47].

The chan-vese algorithm produces the lowest classification performance; this is confirmed from the results of the significance test with a confidence level of 96%, where many features generated from GLCM cannot show significant differences over M0xM1, M0xM2, and M1xM2. The test results for Chan-vese segmentation with a GLCM angle of 0° can be shown in Table VI. In Table VI, it is shown that the features generated when using the Chan-vese segmentation algorithm can distinguish significantly between M0, M1, and

M2, only two features, namely contrast, and dissimilarity, while for feature dissimilarity it cannot distinguish between M1xM2.

TABLE VI. STATISTICS TEST ON GLCM 0^0 (CHAN VESE)

Feature	P-value		
	M0xM1	**M0xM2**	**M1xM2**
Contrast	**0.000**	**0.000**	**0.003**
Dissimilarity	**0.003**	**0.005**	0.961
Homogeneity	0.132	0.054	0.544
Energy	0.397	0.928	0.354
Correlation	0.258	0.874	0.218
ASM	0.406	0.883	0.412

VI. CONCLUSION

The results of the comparison of the performance of the segmentation method based on clustering, thresholding, and region growing show that the Histogram algorithm gives the best performance. The best performance is obtained when using the SVM classification algorithm. Performance can also be seen from the features generated from GLCM, the results of the significance test show the performance of K-mean and Histogram capable of producing features that can distinguish AML M0, M1, and M2 leukemia. The conclusion that can be drawn is that histogram and k-means segmentation algorithms can be an alternative segmentation method in cases of CAD leukemia AML.

REFERENCES

[1] N. Sharma and A. Verma, "Segmentation-and-Detection-of-Optic-Disc-Using-Kmeans-Clustering.docx," International Journal of Scientific & Engineering Research, vol. 6, no. 8, pp. 237–240, 2015.

[2] P. Sharma and J. Suji, "A Review on Image Segmentation with its Clustering Techniques," International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 9, no. 5, pp. 209–218, 2016, doi: 10.14257/ijsip.2016.9.5.18.

[3] S. Yuheng and Y. Hao, "Image Segmentation Algorithms Overview," ArXiv, pp. 1–6, 2017.

[4] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image Segmentation Using K-means Clustering Algorithm and Subtractive Clustering Algorithm," Procedia Computer Science, vol. 54, pp. 764–771, 2015, doi: 10.1016/j.procs.2015.06.090.

[5] N. Dhanachandra and Y. J. Chanu, "A Survey on Image Segmentation Methods using Clustering Techniques," European Journal of Engineering Research and Science, vol. 2, no. 1, p. 15, 2017, doi: 10.24018/ejers.2017.2.1.237.

[6] N. Dhanachandra and Y. J. Chanu, "A new approach of image segmentation method using K-means and kernel based subtractive clustering methods," International Journal of Applied Engineering Research, vol. 12, no. 20, pp. 10458–10464, 2017.

[7] A. Humeau-Heurtier, "Texture Feature Extraction Methods: A Survey," IEEE Access, vol. 7, pp. 8975–9000, 2019, doi: 10.1109/ACCESS.2018.2890743.

[8] S. A. Medjahed, "A Comparative Study of Feature Extraction Methods in Images Classification," IJIGSP, vol. 7, no. 3, pp. 16–23, Feb. 2015, doi: 10.5815/ijigsp.2015.03.03.

[9] W. K. Mutlag, S. K. Ali, Z. M. Aydam, and B. H. Taher, "Feature Extraction Methods: A Review," J. Phys.: Conf. Ser., vol. 1591, p. 012028, Jul. 2020, doi: 10.1088/1742-6596/1591/1/012028.

[10] Ahmed, Yigit, Isik, and Alpkocak, "Identification of Leukemia Subtypes from Microscopic Images Using Convolutional Neural Network," Diagnostics, vol. 9, no. 3, pp. 1–11, Aug. 2019, doi: 10.3390/diagnostics9030104.

[11] E. Suryani, Wiharto, S. Palgunadi, and Y. R. Putra, "Cells identification of acute myeloid leukemia AML M0 and AML M1 using K-nearest neighbour based on morphological images," in Proceedings of 2017 International Conference on Data and Software Engineering, ICoDSE 2017, Palembang, Indonesia, 2018, vol. 2018-Janua, pp. 1–6. doi: 10.1109/ICODSE.2017.8285851.

[12] W. Wiharto, E. Suryani, and Y. R. Putra, "Classification of blast cell type on acute myeloid leukemia (AML) based on image morphology of white blood cells," TELKOMNIKA (Telecommunication Computing Electronics and Control), vol. 17, no. 2, p. 645, 2018, doi: 10.12928/telkomnika.v17i2.8666.

[13] E. Suryani, E. I. Asmari, and B. Harjito, "Image Segmentation of Acute Myeloid Leukemia Using Multi Otsu Thresholding," J. Phys.: Conf. Ser., vol. 1803, no. 1, p. 012016, Feb. 2021, doi: 10.1088/1742-6596/1803/1/012016.

[14] E. Suryani, Wiharto, A. P. Putra, and W. Widiarto, "Detection of Acute Myeloid Leukemia Based on White Blood Cell Morphological Imaging Using Naïve Bayesian Algorithm," International Journal of Advanced Computer Science and Applications, vol. 12, no. 10, pp. 244–251, 2021.

[15] T. Z. T. Muda and R. A. Salam, "Comparative Analysis on Blood Cell Image Segmentation:," presented at the 2nd International Symposium on Computer, Communication, Control and Automation, Singapore, 2013. doi: 10.2991/3ca-13.2013.115.

[16] S. Chand and V. P. Vishwakarma, "Comparison of Segmentation Algorithms for Leukemia Classification," presented at the Proceedings of the Fist International Conference on Advanced Scientific Innovation in Science, Engineering and Technology, ICASISET 2020, 16-17 May 2020, Chennai, India, Chennai, India, 2021. doi: 10.4108/eai.16-5-2020.2303967.

[17] A. Harjoko, T. Ratnaningsih, E. Suryani, Wiharto, S. Palgunadi, and N. P. T. Prakisya, "Classification of acute myeloid leukemia subtypes M1, M2 and M3 using active contour without edge segmentation and momentum backpropagation artificial neural network," in MATEC Web of Conferences, Yogyakarta, 2018, vol. 154, p. 01041. doi: 10.1051/matecconf/201815401041.

[18] W. Wiharto and E. Suryani, "The Analysis Effect of Cluster Numbers On Fuzzy C-Means Algorithm for Blood Vessel Segmentation of Retinal Fundus Image," p. 5, 2019.

[19] W. Wiharto, E. Suryani, and M. Susilo, "The Hybrid Method of SOM Artificial Neural Network and Median Thresholding for Segmentation of Blood Vessels in the Retina Image Fundus," IJFIS, vol. 19, no. 4, pp. 323–331, Dec. 2019, doi: 10.5391/IJFIS.2019.19.4.323.

[20] N. Khomairoh, R. Sigit, T. Harsono, Y. Hernaningsih, and A. Anwar, "Segmentation System of Acute Myeloid Leukemia (AML) Subtypes on Microscopic Blood Smear Image," in 2020 International Electronics Symposium (IES), Surabaya, Indonesia, Sep. 2020, pp. 565–570. doi: 10.1109/IES50839.2020.9231651.

[21] F. Kazemi, T. Najafabadi, and B. Araabi, "Automatic recognition of acute myelogenous leukemia in blood microscopic images using K-means clustering and support vector machine," J Med Signals Sens, vol. 6, no. 3, p. 183, 2016, doi: 10.4103/2228-7477.186885.

[22] S. Arumugadevi and V. Seenivasagam, "Comparison of Clustering Methods for Segmenting Color Images," Indian Journal of Science and Technology, vol. 8, no. 7, pp. 670–677, Apr. 2015, doi: 10.17485/ijst/2015/v8i7/62862.

[23] S. A. Mingoti and J. O. Lima, "Comparing SOM neural network with Fuzzy c-means, K-means and traditional hierarchical clustering algorithms," European Journal of Operational Research, vol. 174, no. 3, pp. 1742–1759, 2006, doi: 10.1016/j.ejor.2005.03.039.

[24] E. Suryani, Wiharto, S. Palgunadi, and Y. R. Putra, "Cells identification of acute myeloid leukemia AML M0 and AML M1 using K-nearest neighbour based on morphological images," Proceedings of 2017 International Conference on Data and Software Engineering, ICoDSE 2017, vol. 2018-Janua, pp. 1–6, 2018, doi: 10.1109/ICODSE.2017.8285851.

[25] J. Rawat, A. Singh, H. S. Bhadauria, and J. Virmani, "Computer Aided Diagnostic System for Detection of Leukemia Using Microscopic Images," Procedia Computer Science, vol. 70, pp. 748–756, 2015, doi: 10.1016/j.procs.2015.10.113.

[26] Ş. Öztürk and B. Akdemir, "Application of Feature Extraction and Classification Methods for Histopathological Image using GLCM, LBP, LBGLCM, GLRLM and SFTA," Procedia Computer Science, vol. 132, pp. 40–46, 2018, doi: 10.1016/j.procs.2018.05.057.

[27] S. Ghosh and S. Kumar, "Comparative Analysis of K-Means and Fuzzy C-Means Algorithms," International Journal of Advanced Computer Science and Applications, vol. 4, no. 4, 2013, doi: 10.14569/ijacsa.2013.040406.

[28] A. A. Yahya, J. Tan, and M. Hu, "A Novel Model of Image Segmentation Based on Watershed Algorithm," Advances in Multimedia, vol. 2013, pp. 1–8, 2013, doi: 10.1155/2013/120798.

[29] S. Li and Q. Zhang, "Fast Image Segmentation Based on Efficient Implementation of the Chan-Vese Model with Discrete Gray Level Sets," Mathematical Problems in Engineering, vol. 2013, pp. 1–16, 2013, doi: 10.1155/2013/508543.

[30] X.-F. Wang, D.-S. Huang, and H. Xu, "An efficient local Chan–Vese model for image segmentation," Pattern Recognition, vol. 43, no. 3, pp. 603–618, Mar. 2010, doi: 10.1016/j.patcog.2009.08.002.

[31] M. A. Ansari and S. K. Mahraj, "A Robust Method for Identification of Paper Currency Using Otsu's Thresholding," 2018 International Conference on Smart Computing and Electronic Enterprise, ICSCEE 2018, pp. 1–5, 2018, doi: 10.1109/ICSCEE.2018.8538424.

[32] J. Delon, A. Desolneux, J.-L. Lisani, and A. B. Petro, "A Nonparametric Approach for Histogram Segmentation," IEEE Trans. on Image Process., vol. 16, no. 1, pp. 253–261, Jan. 2007, doi: 10.1109/TIP.2006.884951.

[33] H. T. A. Wahhab, "Classification of Acute Leukemia Using Image Processing and Machine Learning Techniques," Disertation, University of Malaya, Malaysia, 2015.

[34] P. D. R. Raju and G. Neelima, "Image Segmentation by using Histogram Thresholding," International Journal of Computer Science Engineering and Technology, vol. 2, no. 1, pp. 776–779, 2012.

[35] K. Qin, K. Xu, F. Liu, and D. Li, "Image segmentation based on histogram analysis utilizing the cloud model," Computers & Mathematics with Applications, vol. 62, no. 7, pp. 2824–2833, Oct. 2011, doi: 10.1016/j.camwa.2011.07.048.

[36] F. Asadi, F. M. Putra, M. Indah Sakinatunnisa, F. Syafria, Okfalisa, and I. Marzuki, "Implementation of Backpropagation Neural Network and Blood Cells Imagery Extraction for Acute Leukemia Classification," in 2017 5th International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME), Bandung, Nov. 2017, pp. 106–110. doi: 10.1109/ICICI-BME.2017.8537755.

[37] Y. Jusman, L. A. Dewiprabamukti, A. N. N. Chamim, Z. Mohamed, S. N. A. M. Kanafiah, and N. H. A. Halim, "Application of Watershed Algorithm and Gray Level Co-Occurrence Matrix in Leukemia Cells Images," in 2020 3rd International Conference on Mechanical, Electronics, Computer, and Industrial Technology (MECnIT), Medan, Indonesia, Jun. 2020, pp. 9–14. doi: 10.1109/MECnIT48290.2020.9166651.

[38] A. R. Zubair and O. A. Alo, "Grey Level Co-occurrence Matrix (GLCM) Based Second Order Statistics for Image Texture Analysis," International Journal of Science and Engineering Investigations, vol. 8, no. 93, pp. 64–73, 2019.

[39] P. Musa, F. A. Rafi, and M. Lamsani, "A Review: Contrast-Limited Adaptive Histogram Equalization (CLAHE) methods to help the application of face recognition," in 2018 Third International Conference on Informatics and Computing (ICIC), Palembang, Indonesia, Oct. 2018, pp. 1–6. doi: 10.1109/IAC.2018.8780492.

[40] D. Sundararajan, "Morphological Image Processing," in Digital Image Processing, Singapore: Springer Singapore, 2017, pp. 217–256. doi: 10.1007/978-981-10-6113-4_8.

[41] M. Mustafa, H. Najwa Omar Rashid, N. Rul Hasma Abdullah, R. Samad, and D. Pebrianti, "Mammography Image Segmentation: Chan-Vese Active Contour and Localised Active Contour Approach," IJEECS, vol. 5, no. 3, p. 577, Mar. 2017, doi: 10.11591/ijeecs.v5.i3.pp577-583.

[42] T. F. Chan, B. Y. Sandberg, and L. A. Vese, "Active Contours without Edges for Vector-Valued Images," Journal of Visual Communication and Image Representation, vol. 11, no. 2, pp. 130–141, Jun. 2000, doi: 10.1006/jvci.1999.0442.

[43] L. Wang and H. Fan, "Segmentation of Ultrasonic Images Based on Modified Chan-Vese algorithm," in Proceedings of the 6th International Conference on Electronic, Mechanical, Information and Management Society, Shenyang, China, 2016, pp. 710–714. doi: 10.2991/emim-16.2016.147.

[44] D. Zhou, H. Zhou, and Y. Shao, "An improved Chan–Vese model by regional fitting for infrared image segmentation," Infrared Physics & Technology, vol. 74, pp. 81–88, Jan. 2016, doi: 10.1016/j.infrared.2015.12.003.

[45] R. Sarma and Y. K. Gupta, "A comparative study of new and existing segmentation techniques," IOP Conf. Ser.: Mater. Sci. Eng., vol. 1022, p. 012027, Jan. 2021, doi: 10.1088/1757-899X/1022/1/012027.

[46] E. Ramentol, Y. Caballero, R. Bello, and F. Herrera, "SMOTE-RSB *: A hybrid preprocessing approach based on oversampling and undersampling for high imbalanced data-sets using SMOTE and rough sets theory," Knowledge and Information Systems, vol. 33, no. 2, pp. 245–265, 2012, doi: 10.1007/s10115-011-0465-6.

[47] F. Gorunescu, Data Mining: Concepts, Models and Techniques. Berlin, Heidelberg: Springer, 2011.