Journal of Telecommunication, Electronic and Computer Engineering
ISSN: 2180 – 1843 e-ISSN: 2289-8131 Vol. 14 No. 3
jtec.utem.edu.my

Efficient Classification Techniques in Classifying Human Intestinal Parasite Ova

N. A. A. Khairudin[1,2], C. C. Lim [2], A. S. A. Nasir[1] and Z. Mohamed [3]
[1]*Faculty of Electrical Engineering Technology, Universiti Malaysia Perlis, UniCITI Alam Campus, Sungai Chuchuh, 02100 Padang Besar, Perlis, Malaysia.*
[2]*Faculty of Electronic Engineering Technology, University Malaysia Perlis, Pauh Putra Campus, 02600 Arau, Perlis, Malaysia.*
[3]*Department of Microbiology & Parasitology, School of Medical Sciences, Health Campus, Universiti Sains Malaysia, 16150 Kubang Kerian, Kelantan, Malaysia.*
norhanisayunie@studentmail.unimap.edu.my

Article Info

Article history:
Received Aug 20th, 2022
Revised Sep 20th, 2022
Accepted Sep 26th, 2022

Index Terms:
Helminth
Parasite ova
k-NN
SVM
Ensemble

Abstract

Helminth parasites live in the human body and can cause serious health problems that will lead to cancer and may cause death in patients. These parasitic helminths congregate in the intestines to mate and produce ova. Therefore, early identification screening is necessary to prevent the spread of helminth parasites throughout the body. A manual microscopic feces test is still the most often used approach for helminth detection. As a result, the purpose of this research is to investigate the effectiveness of three classifiers in classifying four types of human intestinal parasite ova. Three classifier techniques used are k-Nearest Neighbourhood (k-NN), Support Vector Machine (SVM), and Ensemble classifier. There are four types of helminth ova which are *Ascaris Lumbricoides Ova* (ALO), *Enterobius Vermicularis Ova* (EVO), *Hookworm Ova* (HWO), and *Trichuris Trichiura Ova* (TTO). A total of 664 helminth parasite ova images were analyzed, consisting of 166 images from each helminth species. The Linear kernel function from the SVM classifier has obtained the highest accuracy performance reaching 92.23%. Followed by Cityblock distance from the k-NN classifier with an accuracy of 91.16% and AdaBoostM2 from the Ensemble classifier with an accuracy of 89.94%.

I. INTRODUCTION

Parasitic disorders are distinguished through the identification of parasitic organisms as well as parasite ova in blood, urine, feces, and tissues using appropriate recognition methods [1]. Helminths have been discovered as infectious agents for parasitic worm-type diseases worldwide known as helminthiasis. Helminth ova can be transmitted to humans by direct contact with polluted sludge or feces, through ingestion of contaminated food or water, and through contact with dust from animal hair and bodies. [2-3]. Inhaling helminth ova in the air dust is also one of the causes of helminth penetration into the human body [4]. Figure 1 shows the images of four helminth species.

The helminth ova are microscopic in size and each species has a different ovum size and shape. Their durability varies as well [5]. Global statistic indicates that helminthiasis has infected more than 1.5 billion individuals and children aged 5 to 15 are especially vulnerable to infections that can harm their physical, emotional, and mental health. By 2020, upwards of 436 million children worldwide have been given helminthiasis treatment. [6].

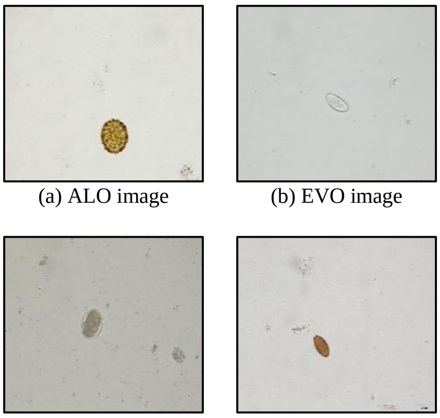

(a) ALO image (b) EVO image
(c) HWO image (d) TTO image.

Figure 1. Helminth ova images

Severe infections can result in a variety of symptoms, such as impaired growth, malnutrition, intestinal manifestations, and physical development. Meanwhile, high-concentration infections might cause intestinal blockage, which is usually treated surgically. Early helminthiasis identification is crucial for patient recovery, especially in young patients. Helminth parasites ova can be identified through feces samples, tissue samples, mucus samples, and blood samples from patients.

It becomes an issue when the samples must be diagnosed by a parasitologist as soon as possible since they must be in fresh condition. The morphology of the helminth ova must be carefully studied in order to correctly identify and classify the helminth species. This method necessitates the observer to have a sharp focus and pay close attention to the helminth ovum samples. Therefore, it takes a significant amount of time [7-8]. Due to these drawbacks, digital image processing for helminth ova detection and classification has been implemented. Thus, this paper aims to identify the classification performance of helminth ova by using three different classification techniques: k-NN, SVM, and Ensemble classifier to find out the appropriate technique for helminth ova classification.

In the prior study, various researchers have put forward digital image processing techniques to detect and classify the helminth ova in the human body. Hadi et al. [9] proposed a Threshold with Logical Classification Method (TLCM) classifier for detecting helminth-based intestinal parasites, namely ALO and TTO. Three image processing methods have been presented; method I has used contrast enhancement with the gray threshold, method II has used canny edge detection, whereas, method III has used contrast enhancement and canny edge detection. Based on the analysis obtained, method III has obtained the highest accuracy results with a value of 93% in identifying ALO species and a value of 94% in identifying TTO species.

Sengul [10] presented a procedure for distinguishing helminth parasites based on ovum structure. The Gray-Level Co-occurrence Matrix (GLCM), a texture-based statistical approach was used for feature extraction. The k-NN classifier was then used to automatically classify the helminth parasites. The results demonstrated a 99% accuracy rate in classifying 14 different types of helminth parasites. Next, Jimenez *et al.* [11] suggested a technique for distinguishing and identifying helminth species in wastewater. Helminth contours were obtained using grayscale profile segmentation. As a result, the genera and species may be distinguished more easily. The system can discriminate seven kinds of helminth species with 99% and around 80% to 90% for specificity and sensitivity, respectively.

Avci and Varol [12] proposed an expert diagnosis system based on a multi-class support vector machine (MCSVM) classifier to identify sixteen types of human parasite ova in microscopic images. In the pre-processing stage, noise reduction, contrast enhancement, thresholding, and post-processing have been used. Meanwhile, an invariance moment was used for the feature extraction stage. This method has provided an accuracy of 97.70% as an overall success rate.

Previous studies focused on helminths that inhabit the human body. Therefore, many systems have been proposed to identify these helminths, but erroneous results are likely to occur even if high accuracy results are obtained. This is due to some species of helminths might have a similar size, shape, or character as each other, causing them to be mistakenly identified. Therefore, this paper proposed an algorithm that specifically identifies and classifies the ova of human intestinal parasitic helminths based on feces samples to achieve more accurate results.

.

II. MATERIAL AND METHODS

A. Image Acquisition

A computerized microscope is used to capture images of the targeted helminth ova from feces sample slides. The Department of Microbiology and Parasitology at Hospital Universiti Sains Malaysia prepared these feces samples (HUSM), which are freshly collected from patients. The samples images collected are from four different species; *Ascaris Lumbricoides Ova* (ALO), *Enterobius Vermicularis Ova* (EVO), *Hookworm Ova* (HWO), and *Trichuris Trichiura Ova* (TTO). An Olympus Digital Microscope is used to examine these samples at a 40X magnification. To achieve clear vision, normal saline is used for staining. Each species got 166 images, with 33 for under-exposed illumination, 33 for over-exposed illumination, and 100 for normal illumination. The total number of helminth ova images is 664 and stored in.jpg format. The collected images have a resolution of 1294 × 980 pixels.

B. Image Pre-processing

Image pre-processing is a procedure used to boost the quality of the original image sample and is a crucial step before the segmentation procedure. It includes image enhancement and color conversion, which allows for more efficient data analysis [13]. Pre-processing step is necessary for designing a reliable system that can be used in different conditions such as different illuminations, human error and different sample staining techniques, and so on [14].

1) Image Enhancement

The collected data of the helminth ova images appear in three different illuminations that need to be standardized into a similar illumination without disrupting the original image color structure. A good enhancement technique helps in enhancing the ROI as well as improving the overall quality of the helminth image. The modified global contrast stretching (MGCS) technique is applied to standardize the illumination in the helminth ova images. This technique uses the particular minimum and maximum values that lie in a specified percentage of pixels from the total number of pixels in the RGB image [15]. The image enhancement of helminth ova depends directly on the minimum and maximum values that will be used during the contrast stretching process. The percentage of minimum value (min_p) is obtained from the lowest value among the R, G, and B color components out of the total numbers of pixels, likewise for maximum (max_p). The lowest and highest values obtained must satisfy the conditions in Equation (1) and Equation (2) [16].

$$\frac{Tmin_{(RGB)}}{total\ number\ of\ pixels\ in\ image} * 100 \geq min_p \qquad (1)$$

$$\frac{Tmax_{(RGB)}}{total\ number\ of\ pixels\ in\ image} * 100 \geq max_p \qquad (2)$$

where: T_{min} = total amounts of pixels that fall between a particular minimum percentage
T_{max} = total amounts of pixels that fall between a particular maximum percentage
min_p = minimum percentages
max_p = maximum percentages

2) Color Model

A color model is used to identify, simplify, extract and edit the particular color needed. It is used to distinguish the desired region and reduce undesired artifacts on the helminth image. In this paper, the K component from CMYK color space (K_CMYK) was applied to the enhanced helminth ova images. The calculation to obtain the K_CMYK [17] is shown in Equation (3) and Equation (4).

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3)$$

$$K = min\,(C, M, Y) \quad (4)$$

C. Image Segmentation

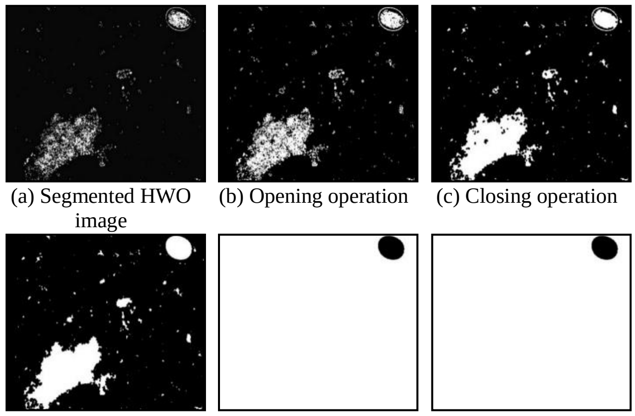
The primary goal of helminth segmentation is to distinguish the regions in the helminth image by splitting them into ROI and background regions. Thus, the Fuzzy c-means (FCM) clustering algorithm is an iterative distributing strategy that produces optimal c-partitions. This method calculates the cluster centers and generates the class membership matrix. It also is designed to minimize the goal function, J_m (U,v) as in Equation (5):

$$J_m(U, v) = \sum_{k=1}^{N} \sum_{i=1}^{C} (U_{ki})^m \,\|Y_k - V_i\|^2\, A \quad (5)$$

The $Y = \{y_1, y_2, \dots y_n\} \subset R^n$ is the data set. C is the number of clusters in $Y; 2 \leq c < n$, and m is the weighting exponent. Then, U is the fuzzy c-partition of Y. While $\|Y_k - V_i\|A$ is an induced a-norm of R^n. The FCM technique generates the initial random membership matrix. This random membership matrix is used to determine how well each sample fits into each cluster. The membership matrix is then updated using the newly produced cluster centers. Then, the new membership matrix is compared to the prior membership matrix. If the difference is greater than a specified threshold, another iteration is generated. Otherwise, the algorithm is terminated [18].

D. Image Post-processing and Masking

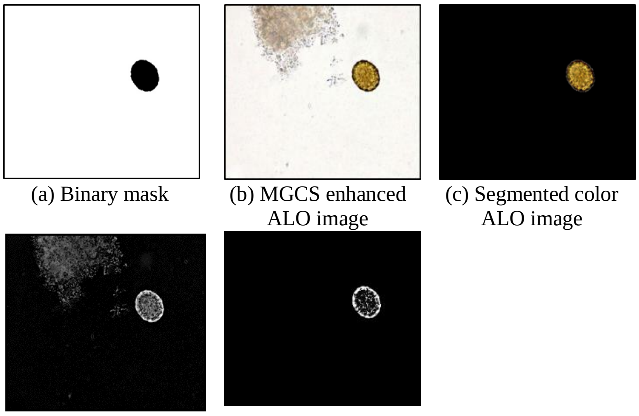
After image segmentation, the segmented images are subjected to post-processing methods that included several morphological operations. This post-processing procedure is used to help to gain a better-segmented image. First, morphological opening and closing are used to restructure and recover the removed information in the ROI. After that, filling image regions and holes are applied to fill the holes in the target region. The object remover approach is then used to eliminate image pixels that are smaller than the size of the intended segmented image in order to minimize misidentifying the helminth species. As a result, any segmented regions with fewer than 6000 pixels and more than 38000 pixels are deemed undesirable and deleted during the object-removing operation. Figure 2 shows the sequence of the post-processing operations applied to achieve a clean helminth ova image.



(a) Segmented HWO image (b) Opening operation (c) Closing operation (d) Fill holes operation (e) Object removing operation (f) Final image.

Figure 2. The post-processing operations on the HWO image

Using the layer masking approach, image masking is performed to retrieve the ROI from the helminth image. When a binary mask was applied to an image to perform a selective adjustment, the ROI appeared in the output image. Figure 3 shows the example of the layer masking technique on an image to obtain the colored and grayscale ROI without any noise or background in the resultant image.



(a) Binary mask (b) MGCS enhanced ALO image (c) Segmented color ALO image (d) K_CMYK on ALO image (e) Segmented grayscale ALO image.

Figure 3. The layer masking technique on ALO image: ALO image

E. Feature Extraction

Feature extraction transfers the input data obtained from the segmented ROI into different sets of features. In this paper, the features of the segmented ROI are extracted to classify the types ofhelminth ova. Five types of feature extractions are Hu's invariant moment [19], Affine Moment Invariants (AMI) [20], color feature [21], Gray Level Co-occurrence Matrix (GLCM) [22], and simple shape [23] are applied. Then, the data collected in feature extraction has been analyzed and tabulated as preparation for the classification process.

Three out of five feature extractions are used to extract the features from the binary image which are Hu's invariant moment, Affine Moment Invariants, and simple shape. Seven features have been obtained from Hu invariance and AMI methods, respectively. Simultaneously, six features are obtained from the simple shape technique. A total of twenty features are extracted by these three feature extraction methods from the binary image.

Next, the GLCM feature extraction technique was applied to the grayscale image. A total of twenty-two features have been obtained through the grayscale image of the helminth ovum. In the meantime, there are six color features are obtained from the segmented color helminth image which is mean, and standard deviation for each red, green, and blue. These five feature extraction methods yielded a total of 48 features, as shown in Table 1.

Table 1
Features obtained from the masking and feature extraction technique

Feature Extraction Method	Hu	AMI	Simple Shape	GLCM	Color
Resultant segmented image	Binary	Binary	Binary	Grayscale = Binary + CMYK_K	Color = Binary + MGCS
Features	Hu 1, Hu 2, Hu 3, Hu 4, Hu 5, Hu 6, Hu 7	AMI 1, AMI 2, AMI 3, AMI 4, AMI 5, AMI 6, AMI 7	Area, perimeter, diameter, eccentricity, centroid 1, centroid 2,	Autocorrelation, contrast, correlation 1, correlation 2, cluster prominent, cluster shade, dissimilarity, energy, entropy, homogeneity 1, homogeneity 2, maximum probability, variance, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation 1, information measure of correlation 2, inverse difference normalized, inverse difference moment normalized	Mean R, mean G, mean B, std R, std G, std B,

F. Classification

Three types of classifiers, k-Nearest Neighbourhood (k-NN), Support Vector Machine (SVM), and Ensemble were used and compared to identify the most appropriate classifier for classifying the types of helminth ovum. Comparisons were made through the performance measures obtained when a 10-fold cross-validation method was applied to each classifier.

The k-NN is a classification technique based on the non-parametric procedure that is usually used in machine learning and pattern recognition applications. It predicts the new data input based on how close the distance matches the training dataset [24]. Then, four different types of distance metrics; Euclidean, Cityblock, Minkowski, and Chebychev, were trained and tested in the k-NN classifier. The calculations of these four distance metrics are shown in Equations (6) until (9).

$$Euclidean = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \tag{6}$$

$$Cityblock = \sum_{i=1}^{k}|x_i - y_i| \tag{7}$$

$$Minkowski = \left(\sum_{i=1}^{k}(|x_i - y_i|)^q\right)^{1/q} \tag{8}$$

$$Chebychev = max\left(\sum_{i=1}^{k}|x_i - y_i|\right) \tag{9}$$

where: k = number of variables
x_i = variables of vector x
y_i = variables of vector y

The SVM classifier's basic principle is to discover the optimal linear hyperplane equation, which maximizes the distance between it and the nearest data point [25]. One-versus-one coding design is applied to four classes of helminth ova to produce four binary learners. Through this coding design, the SVM classifies one class of helminth ovum samples as positive and the rest of the samples as negative samples. Three types of kernel functions applied in this paper are Gaussian, Linear, and Polynomial. The calculations for these three kernel functions are shown in Equations (10) until (12), where x^T is the unit vector, y is a constant, and p is the polynomial degree.

$$Gaussian, k(x,y) = exp\left(\frac{\|x-y\|^2}{-2*\sigma}\right) \tag{10}$$

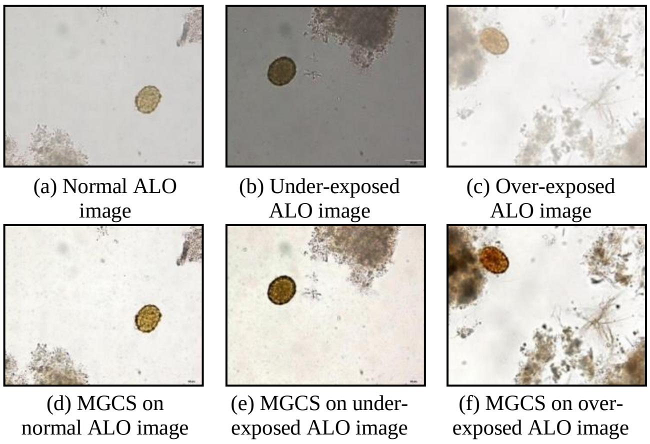
$$Linear, k(x,y) = x^T.y \tag{11}$$

$$Polynomial, k(x,y) = (x^T, y)^p \tag{12}$$

Meanwhile, the Ensemble classifier is a popular classifier because it can produce diversity in the classifier [26]. In this paper, three types of ensemble aggregation methods were used which are AdaBoostM2, Bootstrap aggregation (Bag), and RUSBoost. The decision learner tree used for these three ensemble aggregations is fixed through the best split value. The value for the best split is set to 5 in the split criterion.

III. RESULTS AND DISCUSSION

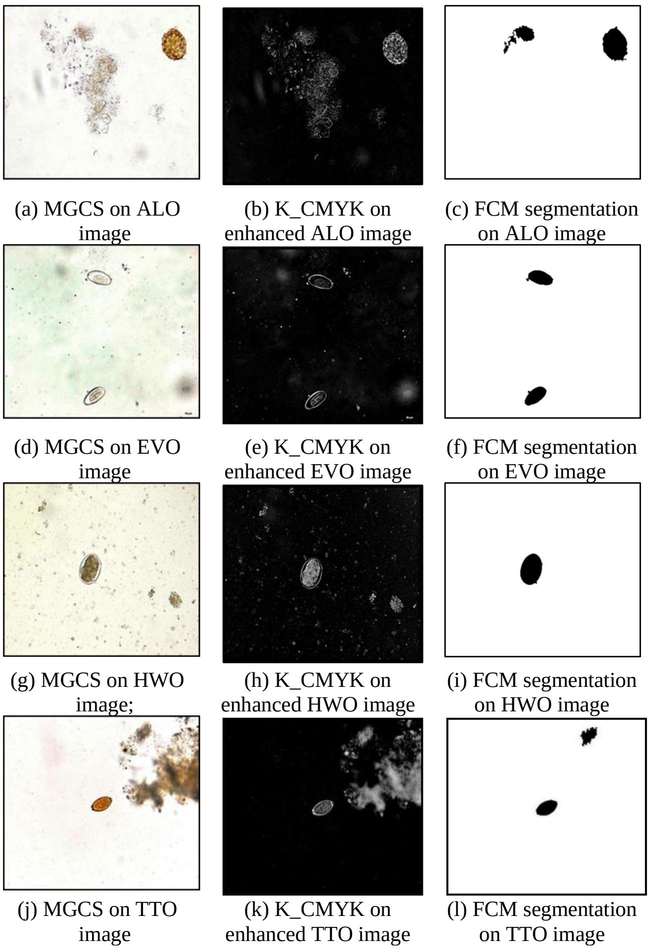
A total of 664 images from the ALO, EVO, HWO, and TTO species were obtained. Despite the fact that it comes with three different illuminations: normal illumination, under-exposed illumination, and over-exposed illumination, the MGCS approach can standardize the illumination for all images produced. Figure 4 shows the example of the ALO species captured in three different illuminations with the results obtained when the MGCS technique was applied. The ROI is well enhanced but the artifacts are enhanced as well.



(a) Normal ALO image
(b) Under-exposed ALO image
(c) Over-exposed ALO image
(d) MGCS on normal ALO image
(e) MGCS on under-exposed ALO image
(f) MGCS on over-exposed ALO image

Figure 4. The results of the MGCS technique on ALO images with three different illuminations

After the MGCS technique was applied, K_CMYK was applied to the MGCS-enhanced image to distinguish the ROI, artifact, and background. Then, the FCM segmentation technique was applied to separate the ROI region from the unwanted region. Followed by a post-processing procedure to obtain a clean and clear ROI. Figure 5 shows the resultant images obtained from the color model and final segmentation on each of helminth ovum species.



(a) MGCS on ALO image
(b) K_CMYK on enhanced ALO image
(c) FCM segmentation on ALO image
(d) MGCS on EVO image
(e) K_CMYK on enhanced EVO image
(f) FCM segmentation on EVO image
(g) MGCS on HWO image;
(h) K_CMYK on enhanced HWO image
(i) FCM segmentation on HWO image
(j) MGCS on TTO image
(k) K_CMYK on enhanced TTO image
(l) FCM segmentation on TTO image

Figure 5. The results of a color model and segmentation technique on helminth ova images

Through the results obtained, most of the species are able to achieve a clean and clear ROI. However, some of the artifacts that appeared in the same size as the ROI are unable to be removed and segmented together with the ROI as shown in Fig. 5 (c) and (l). Quantitative measures have been done to analyze the FCM segmentation performance of the segmented helminth ova images. Table 2 tabulated the FCM segmentation analysis obtained from each species of segmented helminth ovum images.

Table 2
Quantitative measure on segmentation performance of helminth ova images

Species	Accuracy (%)	Sensitivity (%)	Specificity (%)
ALO	98.54	78.98	98.94
EVO	97.94	84.12	98.15
HWO	99.48	81.13	99.81
TTO	98.75	85.81	98.85

Through the quantitative analysis obtained, the results show that each of the species achieves high accuracy and specificity results which are higher than 97%. However, the results obtained for the sensitivity are in the range of 78% to 86% for each species which is quite low. This happens because of the appearance of the artifacts in the final segmented image that cannot be removed because it has the same size as the helminth ova.

After the segmentation procedure is done, image masking takes place to acquire binary, grayscale, and color images of segmented ROI. These images are used for feature extraction and a total of 48 features are collected and tabulated as input data for the classification procedure. For classification, three types of classification techniques are used, and ten-fold cross-validation is applied for 10% of training data and 90% of testing data for each of the classification techniques.

Figure 6 shows the training and testing performances obtained from four different types of distance from k-NN: Euclidean, Cityblock, Minkowski, and Chebychev when using a k-value equal to 1 [27]. The results obtained by training performance are the same for each distance which is 99.86%. Meanwhile, Cityblock distance has the best testing performance with a score of 91.16%. Minkowski distance came in second with 90.24%, Euclidean distance came in third with 89.94% and Chebychev distance came in fourth with 85.67%. Cityblock distance is chosen as the best technique to be implemented in k-NN classifier compared to the Euclidean, Minkowski, and Chebychev distance.

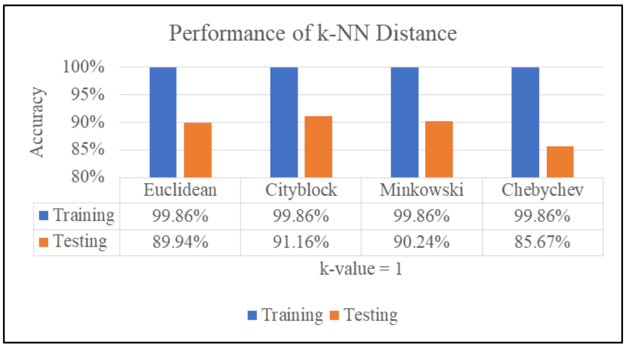


Figure 6. The highest performance for each distance in the k-NN classifier.

Then, three different kernels which are Polynomial, Gaussian, and Linear were used to discover the best kernel function for the SVM classifier. Figure 7 shows the training and testing performance for accuracy obtained from each kernel function. Since all of the outcomes are greater than

90%, these kernel functions exhibit good training and testing performance. For the training accuracy, Polynomial and Gaussian kernel has achieved the highest accuracy analysis with a value of 99.86%, while Linear has achieved 96.05%. In the meantime, the Linear kernel has achieved the highest accuracy for testing performance with a value of 92.23%, followed by Gaussian and Polynomial kernel with 92.07% and 90.55%, respectively. Nevertheless, the Linear kernel function has been chosen as the best kernel function because it achieves the highest testing performance compared to the other kernel functions.

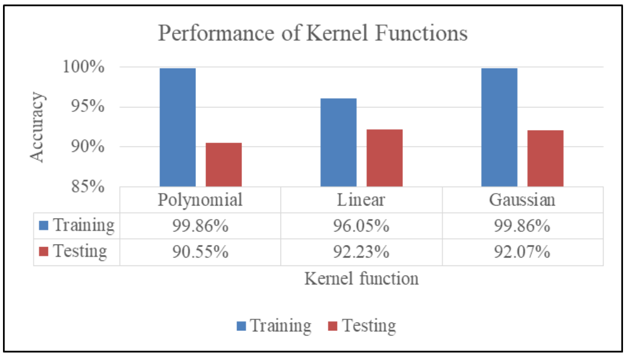


Figure 7. Performance of the kernel functions used in the SVM classifier.

Next, Fig. 8 presents the training and testing performance when three ensemble aggregations are applied. AdaBoostM2 has conquered the highest accuracy for training and testing performance with 98.37% for training performance and 89.94% for testing performance. Then, the Bag aggregation achieved the second highest results for accuracy with 88.11% for training performance and 83.84% for testing performance. RUSBoost aggregation is the lowest with a value of 83.40% for training performance and 80.49% for testing performance.

Based on the overall results obtained from the aggregation methods in the Ensemble classifier, AdaBoostM2 has shown a good analysis performance compared to the other aggregation methods when it achieved the highest accuracy analysis. With an accuracy of 89.94%, AdaBoostM2 is selected as the best aggregation method to be applied in the Ensemble classifier compared to the Bag and RUSBoost three different kernels which are Polynomial, Gaussian, and Linear were used to discover the best kernel function for the SVM classifier.

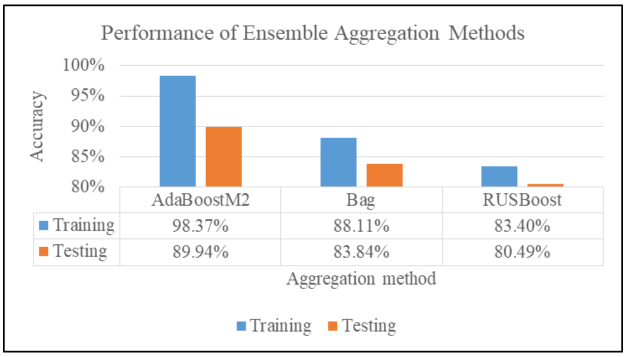


Figure 8. Performance of the aggregation methods used in the Ensemble classifier.

Then, Fig. 9 shows the clustered graph and data collected from the best methods for the three classifiers used in this of research. Through the results presented, the Linear kernel function from the SVM classifier has acquired the highest accuracy for testing analysis with 92.23%. Followed by Cityblock distance from the k-NN classifier with 91.16% and AdaBoostM2 from the Ensemble classifier with an accuracy of 89.94%. This indicates that the Linear kernel function has a better performance in classifying the helminth ova compared to AdaBoostM2 and Cityblock.

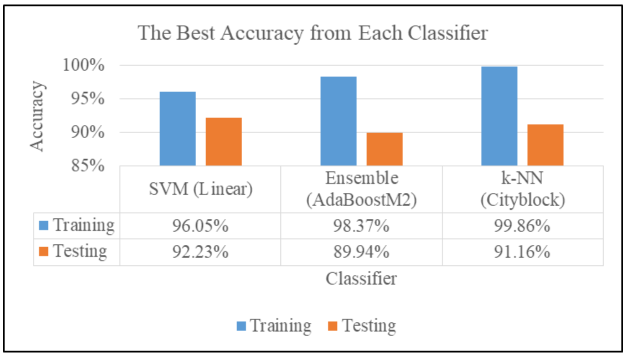


Figure 9. Performance of the kernel functions used in the SVM classifier.

IV. CONCLUSION

In this paper, the overall classification results for the four types of helminths ova have been obtained. The MGCS enhancement technique applied is capable to standardize the illumination in the original helminth ovum images. Then, the combination of color model, segmentation, and post-processing which are K components from the CMYK color model, FCM segmentation, and post-processing operations shows excellent results in segmenting the ROI. However, some of the artifacts have also been segmented which reduces the segmentation performance and also classification performance. Thus, it is recommended to compare the results obtained by the K component from the CMYK color model with several color models to identify which color component is the best to assist the segmentation procedure.

Based on the three types of classifiers used for helminth ova images, the Linear kernel function from the SVM classifier has the highest accuracy for testing accuracy of 92.23%, followed by the Cityblock distance from the k-NN classifier with 91.16% and AdaBoostM2 from Ensemble classifier with an accuracy of 89.94%. As a result, the Linear kernel function from the SVM classifier is discovered as the most appropriate technique to be used in classifying the human intestinal parasites based on helminth ova compared to the other classifier used in this paper.

As a recommendation, it is better to use a staining solution that is able to provide color or reveal distinctions between the region of interest (helminth ovum) and the artifacts in the data collection stage. This is due to the fact that HWO and EVO species used in this study are colorless. This circumstance complicates obtaining a clear and clean target image without the appearance of artifacts. Then, it is also suggested that the size of the helminth ova images be calculated based on their actual size (µm) for better detection during the morphological operation compared to the pixel size.

ACKNOWLEDGMENT

The authors gratefully acknowledge team members and thank Hospital Universiti Sains Malaysia (HUSM) for providing helminths ova samples.

REFERENCES

[1] Y. S. Yang, K. P. Duck, C. K. Hee, M. H. Choi, J. Y. Chai, "Automatic identification of human helminth eggs on microscopic fecal specimens using digital image processing and an artificial neural network," *IEEE Trans. Biomed. Eng.* 2001, *vol. 48(6)*, pp. 718–730.

[2] I. D. Amoah, G. Singh, T. A. Stenström, P. Reddy, "Detection and quantification of soil-transmitted helminths in environmental samples: A review of current state-of-the-art and future perspectives," *Acta Trop.* 2017, *vol. 169, no. February*, pp. 187–201.

[3] N. A. A. Khairudin, N. S. Rohaizad, A. S. A. Nasir, L. C. Chin, H. Jaafar, Z. Mohamed, "Image segmentation using k-means clustering and otsu's thresholding with classification method for human intestinal parasites," *IOP Conf. Ser. Mater. Sci. Eng.* 2020, *vol. 864, no. 1*.

[4] L. B. Huat, A. K. Mitra, J. N. I. Noor, P. C. Dam, M. H. J. Jan, M. W. A. M. Wan, "Prevalence and risk factors of intestinal helminth infection among rural Malay children," *J. Glob. Infect. Dis.* 2012, *vol. 4, no. 1*, pp. 10–14.

[5] K. H. Ghazali, R. S. Hadi, M. Zeehaida, "Microscopy image processing analaysis for automatic detection of human intestinal parasites ALO and TTO," *Int. Conf. Electron. Comput. Comput. ICECCO 2013*, 2013, pp. 40–43.

[6] WHO, "Training manual on diagnosis of intestinal parasites : tutor's guide [electronic resource]," 2004, pp. 1 CD-ROM.

[7] N. Mohd-Shaharuddin, Y. A. L. Lim, N. A. Hassan, S. Nathan, R. Ngui, "Soil-transmitted helminthiasis among indigenous communities in Malaysia: Is this the endless malady with no solution?," *Trop. Biomed.*, 2018, *vol. 35, no. 1*, pp. 168–180.

[8] N. A. A. Khairudin, A. S. A. Nasir, L. C. Chin, Z. Mohamed, C. Y. Fook, "An Improvement for Human Intestinal Parasites Detection Methodology using k-Means and Fast k-Means Clustering," *Proc. - 2020 IEEE EMBS Conf. Biomed. Eng. Sci. IECBES 2020*, 2021, *no. March*, pp. 378–383.

[9] R. S. Hadi, K. H. Ghazali, I. Z. Khalidin, M. Zeehaida, "Human parasitic worm detection using image processing technique," *ISCAIE 2012 - 2012 IEEE Symp. Comput. Appl. Ind. Electron.*, 2012, pp. 196–201.

[10] S. Gokhan, "Biomedical research," *Aging Res. - Methodol. Issues*, 2015, *vol. 27, no. 3*, pp. 27–38.

[11] B. Jimenez, C. Maya, G. Velasquez, F. Torner, F. Arambula, J.A. Barrios, "Identification and quantification of pathogenic helminth eggs using a digital image system," *Exp. Parasitol.*, 2016, *vol. 166*, pp. 164–172.

[12] D. Avci, A. Varol, "An expert diagnosis system for classification of human parasite eggs based on multi-class SVM," *Expert Syst. Appl.*, 2009, *vol. 36, no. 1*, pp. 43–48.

[13] M. H. Motlagh, "Automatic Segmentation and Classification of Red and White Blood Cells in Thin Blood," Concordia University, 2015, no. August.

[14] M. Habibzadeh, A. Krzyzak, T. Fevens, A. Sadr, "Counting of RBCs and WBCs in noisy normal blood smear microscopic images," *Med. Imaging 2011 Comput. Diagnosis*, 2011, *vol. 7963, no. September*, p. 79633I.

[15] A. S. Abdul-Nasir, M. Y. Mashor, Z. Mohamed, "Modified global and modified linear contrast stretching algorithms: New colour contrast enhancement techniques for microscopic analysis of malaria slide images," *Comput. Math. Methods Med.*, 2012, *vol. 2012, no. June*.

[16] L. C. Chin, N. A. A. Khairudin, S. W. Loke, A. S. A. Nasir, C. Y. Fook, Z. Mohamed, "Comparison of Human Intestinal Parasite Ova Segmentation Using Machine Learning and Deep Learning Techniques" *Applied Sciences*, 2022, vol 12, no 15, pp 7542. https://doi.org/10.3390/app12157542

[17] M. E. Latoschik, "Realtime 3D Computer Graphics / Virtual Reality-WS: Color Models (p. 13)," Retrieved from https://www.techfak.unibielefeld.de/ags/wbski/lehre/digiSA/WS0607/3DVRCG/Vorlesung/8a.RT3DCGVR-color.pdf

[18] A. E. Hassanien, "Fuzzy rough sets hybrid scheme for breast cancer detection," *Image Vis. Comput.*, 2007, *vol. 25, no. 2*, pp. 172–183.

[19] M. Arafah, Q. A. Moghli, "Efficient Image Recognition Technique Using Invariant Moments and Principle Component Analysis," *J. Data Anal. Inf. Process.*, 2017, *vol. 05, no. 01*, pp. 1–10.

[20] E. G. Karakasis, A. Amanatiadis, A. Gasteratos, S. A. Chatzichristofis, "Image moment invariants as local features for content based image retrieval using the Bag-of-Visual-Words model," *Pattern Recognit. Lett.*, 2015, vol. 55, pp. 22–27.

[21] J. M.; Patel, N. C. A. Gamit, "review on feature extraction techniques in Content Based Image Retrieval," *Proc. 2016 IEEE Int. Conf. Wirel. Commun. Signal Process. Networking, WiSPNET 2016*, 2016, pp. 2259–2263.

[22] P. Shan, "Image segmentation method based on K-mean algorithm," *Eurasip J. Image Video Process.*, 2018, *vol. 2018, no. 1*.

[23] S.; Shafique, S. Tehsin, "Computer-Aided Diagnosis of Acute Lymphoblastic Leukaemia," *Comput. Math. Methods Med.*, 2018, *vol. 2018*.

[24] M. M. D. Joshi, P. A. H. Karode, P. S. R. Suralkar, "White Blood Cells Segmentation and Classification to Detect Acute Leukemia," 2013, *vol. 2, no. 3*, pp. 147–151.

[25] C. Cortes, V. Vapnik, "Support-Vector Networks," *Machine Learning*, 1995, *vol. 20*, pp. 273-297.

[26] M. R. Sumathi, B. Poorna, "Design and development of ensemble of naïve bayes classifiers to predict social and communication deficiency among children," *Int. J. Appl. Eng. Res.*, 2017, *vol. 12, no. 24*, pp. 14190–14198.

[27] L. C. Chin, C. Y. Fook, A. S. A. Nasir, S. N. Basah, M. Y.; Din, Z. Zainuddin, "Classification of the Severity Level for Lower Limb Joint Injuries," *2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, 2021, pp. 267-270.