*2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)*
*Galgotias University, Greater Noida, UP, India. Oct 30-31, 2020*

# Effective Performance of Bins Approach for Classification of Malaria Parasite using Machine Learning

Hrishikesh Telang
B.E. Student, Department of Computer Engineering
St. Francis Institute of Technology
Mumbai, India
telang.hrishikesh@gmail.com

Kavita Sonawane
Professor, Department of Computer Engineering
St. Francis Institute of Technology
Mumbai, India
kavitasonawane@sfit.ac.in

***Abstract*—Malaria is a severe infectious disease transferred through the bite of an infected mosquito caused by a blood parasite of the genus Plasmodium. In the past, conventional microscopy techniques have proven to be time-consuming and had observed a lack of differentiation due to poor accuracy and a few algorithms used. In this paper, our approach primarily focuses on image processing techniques to process and enhance stained thin blood smear images for feature extraction, as well as machine learning techniques for the final classification of feature space. Our emphasis is also to address the drawbacks, as mentioned earlier, by taking input cell images and performing classification using CNN. As an alternate approach, we have also computed color features using a novel Bins Approach Algorithm, statistical features using color moments, and texture features using GLCM, which also equally played a pivotal role in feature extraction for classification. Further, these images are classified into parasitized and uninfected cells by applying machine learning classifiers such as Linear SVM, Random Forest Algorithm (RNN), and KNN over feature space. The proposed algorithms have been experimented using the subset of Lister Hill National Center for Biomedical Communication (LHNCBC) dataset, a division of the National Library of Medicine (NLM). The performance of the algorithms is evaluated and compared using different parameters like accuracy, precision, recall, and F1-score. The proposed application of Bins Approach in malaria parasite detection has proved better in terms of all parameters as compared to the other existing algorithms.**

***Keywords—Malaria; parasites; image processing; CNN; feature extraction; Bins Approach; statistical features; color moments; texture features; GLCM; classification; SVM; Random Forest; KNN.***

## I. INTRODUCTION

Malaria is a mosquito-borne communicable disease that grows from person to person or other animals due to infectious bites of a female Anopheles mosquito [1]. In a UNICEF/WHO report dated 2015, every year, about 3.2 billion people are at risk of being infected with malaria, which is nearly half of the world's population [2]. According to the WHO World Malaria Report 2019, an estimated 228 million cases of malaria occurred worldwide in 2018, compared with 251 million cases in 2010 and 231 million cases in 2017 [3]. Moreover, in 2018, children aged under five years are the most vulnerable group, as they accounted for 67% (272,000) of all malaria deaths worldwide [3]. Malaria is the primary cause of death across many countries because, in 2018, there were 94% of malaria deaths reported in the WHO African Region [3]. The report claims that one of the main reasons for febrile children not receiving medical attention may largely owe to inadequate access to health-care providers, with an increased effort of availing malaria diagnosis and treatment across community level [2]. Moreover, malaria is challenging to diagnose owing to common symptoms of normal fever and fevers due to malaria amongst these febrile children [2]. In the past analysis performed, some systems are not standardized or rigorous enough to ensure an accurate diagnosis [4]. As a result, they largely depend on the skills, expertise of the "microscopist," and the environment where they work [4]. Such claims suggest that innovative ways may be required to expand access to better healthcare systems for malaria elimination.

Malaria is a severe health problem, and for quicker results and diagnosis, a more computationally efficient mode is of foremost importance. Our primary focus is on building a system in which a model can process blood smear images in a lesser amount of time. This helps in the deployment of such systems on gadgets with minimal functioning across remote areas where malaria is one of the major causes of high mortality rates.

## II. RELATED WORK

The concept of a Malaria Parasite Classification system came into inception over the past decade due to the growing demand for machine learning across various applications in medical science. Olugboja *et al.* invented a malaria parasite detection using different classifiers with a dataset of 130 images used, over which Watershed segmentation was employed [5]. Similarly, Pattanaik *et al.* performed a three-stage object detection model for malaria parasite using input images from the CPC-DPDx1 data set of Laboratory Identification of Parasitic Diseases [6]. It incorporated Kalman filter to remove the noise, Reinhard method used for color normalization, followed by stain normalization to improve the accuracy of the mining algorithm [6]. Its features were extracted using Angular distance and Softmax Classifier [6].

Bashir *et al.* extracted intensity and threshold features classified using Neural Networks and Back Propagation Algorithm and taking 77 input images, from which a total of 1120 erythrocytes sub-images were cropped [7]. Bairagi *et al.* primarily focused on texture and statistical-based features that were segmented using Otsu's method, watershed transform using distance metrics, and classified using Support Vector Machine [8]. The images were taken from the Center for Disease Control (CDC) and Prevention and Public Health Image Library [8].

Silamut *et al.* have used a myriad of preprocessing techniques along with morphologic operations in different

**978-1-7281-6324-6/20/$31.00 ©2020 IEEE**

combinations [9]. Applications of PCA such as F-statistic, 1-way analysis of variance, information gain, and support vector machine-based recursive feature elimination were employed using Decision Trees as one of the classifiers [9]. Somasekar *et al.* also attempted to perform malaria parasite detection in giesma blood samples using image processing using a square median filter to remove the noise, followed by HSV and Watershed segmentation [10].

While we were motivated by the different techniques used in detecting malaria parasites, we also researched through various Content-Based Image Research (CBIR) techniques. One of such methods that piqued our interest was the novel Bins Approach Technique for feature extraction. Kekre *et al.* well substantiates this process as they extracted feature vectors using color contents of the image. These approaches are considering the significance of the relationship of R, G, B intensities in a color image. Different size bins like 8, 27, and 64 were formed based on different partitioning techniques (Linear and Centre of gravity-based) to partition the histograms of R, G, and B planes of the image into two, three and four parts respectively [11,12]. They have also used histogram modification functions for enhancing the image contents in order to improve the quality of the features to be extracted. Based on the performance analysis, CG partitioning has been recommended by the authors [13].

## III. Challenges Identified

Although past innovations in this domain have produced plausible solutions in classifying malaria parasites, the methods used are traditional and have certain drawbacks in terms of inefficiency. The drawbacks are as shown below:

### *A. Time Consuming*

Conventional microscopy has been an incompetent mode of analysis since it is a laborious task to procure relevant results owing to a lack of accuracy and shreds of evidence provided using appropriate manual classifications.

### *B. Lack of differentiation and classification*

One of the previous systems cannot distinguish between the species of plasmodium parasite [5]. Another approach also failed in providing comparisons as only one algorithm was used [8].

### *C. Uneven intensity ranges*

For images with varying intensities of light among the parasites, this system tends to generate erroneous results and renders them less accurate [10].

## IV. Workflow Of the System

Our primary focus in this work is to prepare an automated system of detecting malaria parasites using effective image processing and classifying them using machine learning techniques. The fundamental aim of this system is to detect the Plasmodium parasite in a stained blood sample image. This system will help to detect malaria parasites from stained blood sample images.

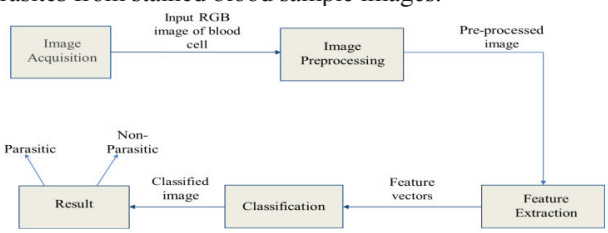



Fig. 1. The Proposed solution for Malaria Classification

The workflow of our system is shown in Fig. 1 that illustrates the process of image acquisition, image preprocessing, feature extraction, and classification. The system has been designed and implemented using two different approaches. One of the machine learning techniques used is CNN [7]. The second method applied is the novel Bins Technique [11,12,13], focusing on color contents of the image to improve the quality of features with the hope of improving the accuracy in detection of the malaria parasite. We have also procured other characteristics such as color moments and texture features in order to improve the classification [8,11,12,13]. We pass these feature databases to three machine learning classifiers, namely SVM, RNN, and KNN, respectively. Then, we optimize the classifier that yields the best result to further enhance the model accuracy in making predictions.

## V. Algorithmic View with Implementation Details

We have experimented with our proposed system with three different methods, as mentioned in Section IV. They have been explained in detail:

### *A. Classification of Malaria Parasites using CNN:*

A convolutional neural network (CNN) is a deep learning architecture that has been suited to perform operations such as pattern recognition and its classification [14]. A CNN model takes an input image and passes it through multiple layers; however, it achieves in reducing a large number of parameters in ANN [14]. It mainly revolves around critical features such as shared weights, pooling, local connections, and multi-layered architectures [15]. CNNs have been used in a variety of fields, including speech recognition [16], text recognition [17], and can also be extended to other applications such as handwriting recognition [18]. All of these justifications make CNN an ideal learning model for malaria parasite detection too.

In this work, we have used the Sequential model by applying a 12-layer for parasite classification. Our model only deals with the feed-forward neural networks in which at every layer, an activation function such as the rectified linear unit (ReLU) is applied to the input weights of the other layers and passes it forward to the next. The pool size used in this model is 2x2 at regular intervals within the layers. Since the output to be procured is binary, we used sigmoid as the final activation function for the last layer. In order to train our CNN model, we used 90% of the images for training and 10% of the images for testing.

### *B. Classification of Malaria Parasites using Bins Approach:*

Bins Approach is a novel technique that can extract essential image contents and represent in the form of simple feature vectors with a significant reduction in the feature dimensions. It is one of the Content-Based Image Retrieval (CBIR) processes that attempts at retrieving appropriate features from large digital image databases. It has proved its best performance in the CBIR domain. In order to check its performance in the medical field, we are proposing to apply the same for our malaria parasite detection system. The Bins Approach applied for feature extraction in malaria parasite detection is given in Table I as follows:

TABLE I. BINS FORMATION PROCESS: 8 BINS [11]

***Step1:*** *Spilt the image into R, G and B planes.*
***Step2:*** *Obtain the histogram for each plane.*
***Step3:*** *Partition the R, G and B histograms in two parts (part 0 and part 1) by computing the Center of gravity (CG).*

***Step4:*** *Let's process each pixel* $P_i$ $(\boldsymbol{r_i},\boldsymbol{g_i},\boldsymbol{b_i})$. *Compare its* ***r, g*** *and* ***b*** *values with respective* ***CG*** *values of three planes and assign a unique flag to each pixel based on the partition (either '0' OR '1') of respective histograms in which it falls.*

*For e.g.- Pi (1,0,0)- flag assigned for pixel Pi- is (100- Bin no 4). This leads to the generation of eight-Bin address for each pixel that ranges from 000 to 111.*

***Step5****:* ***Count of pixels in the bin****: All image pixels will get segregated into 8 bins based on the r, g, b intensities it owns.*

The bins formation process is demonstrated with the sample cell image shown in Fig. 2

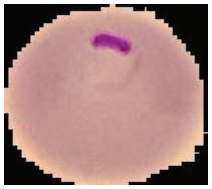


Fig. 2. Sample cell image

Image noise is removed by preprocessing the cell image, as shown in Fig. 3

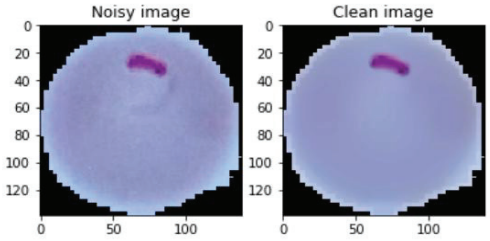



Fig. 3. Original noisy image and preprocessed Image

Once preprocessed, Image gets split into R, G, and B planes as

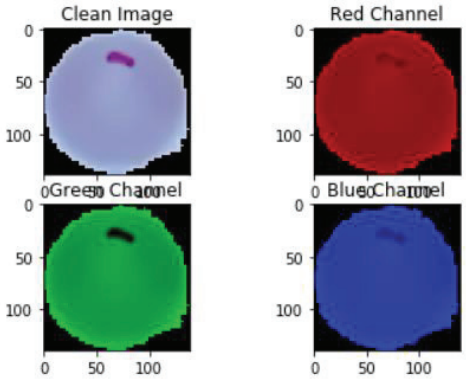
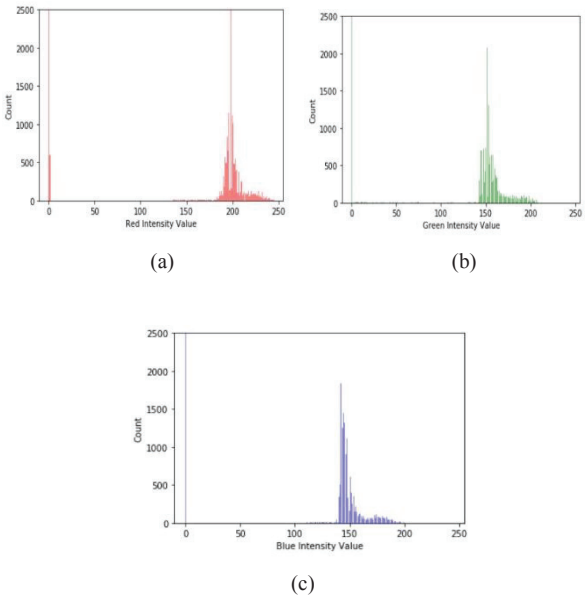



Fig. 4. R, G and B planes of Sample cell_clean image

Later, Compute histograms for each of these planes.




(a) (b)

(c)

Fig. 5. (a)(b)(c). Histograms for R, G and B planes shown in Fig.4

Further, we compute the *Centre of Gravity (CG)* to get balanced partitioning of a histogram in two parts [13]. In this step, we are giving equal importance to the pixels as well as their intensity values. Each of these intensities is considered as the weight of the pixels so that intensities get divided into two partitions equally. This idea is precisely what led us to label these partitions and identify them as '0' and '1', such that both of these partitions will have the same intensity. Every pixel of the image is processed and checked with the respective R, G, B intensities of its channels with respective CG to decide which partition it falls on (0 or 1). For example, suppose we consider a pixel p1(r, g, b) of an image, of which the value of r falls in '1', g in '0', and b in '0'. In such a case, the binary value is processed as '100' that corresponds to bin number- 4, and the pixel gets dumped into it. In this way, all the remaining pixels are segregated as per their computed bin addresses. CG computed for sample image is mentioned as an example in Table II.

TABLE II. TOTAL CG VALUES FOR SAMPLE CELL IMAGES

| | |
|---|---|
| **CG (R-plane)** | 152.7129 |
| **CG (G-plane)** | 119.4505 |
| **CG (B-plane)** | 114.4074 |

This process extracts all color contents of the image and represents each cell image as bins feature vector of 8 components only. Since we already calculated the bin values of the cell images for a parasitic as well as a non-parasitic sample, we pass a loop function wherein we acquire 12,000 clean image samples each of parasitic and non-parasitic cell images. These images are then passed through the entire bins process to produce the values that are written on a .csv file to produce a dataset of 24,000 sample cell images. We maintained separate programs to acquire datasets for parasitic and non-parasitic cell images, labeled all of them as Class '1' (for parasitic cell images) or Class '0' (for non-parasitic cell images). We merged both of them in a standard dataset for the classification process.

*C. Preparation of Fusion Feature vectors using Statistical and Texture Features:*

Although our earlier predictions with bins have deduced better results than CNN, we still wished to improve our accuracy even further. To accomplish this goal, we decided to extract other features such as statistical and texture features from blood smear images. After calculating their features individually using classification, we found out that they were not performing well. Thus, we merged them with bins feature vector to produce a fusion feature vector.

*1) Calculation of Statistical Features:*

For statistical features, we decided to use color moments such as mean, standard deviation, skewness, and kurtosis for each of the R, G, B channels of the corresponding image. This produces twelve features from such measures.

Suppose an image function $I(x, y)$ contains two-pixel variables $x$ and $y$, where $x = 0, 1, 2, \ldots, X-1$ and $y = 0, 1, 2, \ldots, Y-1$. In this case, the discrete values obtained from $I(x, y)$ can be transformed into an intensity value "a" in such a way that $a = 0, 1, 2, \ldots, M-1$, where $M$ represents the total number of intensity levels present in the image. The features are computed using the histogram $h(a)$ and probability density function $p(a)$ of the respective cell image. In this case, (a) is the ratio of all the intensity values computed using $h(a)$ and the total number of pixels in the image. We obtain the features such as mean, standard deviation (SD), skewness, and kurtosis as mentioned in equations (1), (2), (3), (4).

$$\text{Mean: } \mu - \sum_{k=0}^{G-1} ap(a) \tag{1}$$

$$\text{SD: } \sqrt{\sigma^2 - \sum_{k=0}^{G-1}(a-1)^2 p(a)} \tag{2}$$

$$\text{Skewness: } \mu 3 - \sum_{k=0}^{G-1}(a-\mu)^3 p(a) \tag{3}$$

$$\text{Kurtosis: } \mu 4 - \sum_{k=0}^{G-1}(a-\mu)^4 p(a) - 3 \tag{4}$$

*2) Calculation of Texture Features:*

The technique of texture analysis has been used widely in the analysis of malaria parasites. The grey-level co-occurrence matrix (GLCM) performs computations in gathering the texture information by analyzing how often individual pairs of pixels (i, j) with fixed values and having uniform spatial relationships occur together in an image.

The following sample Fig. 6 shows the construction of a sample 8x8 GLCM, which encodes the frequency of occurrence of two neighborhood pixels. In every process of GLCM, an image is converted into grayscale, and the process of feature extraction is immediately followed [19, 20].

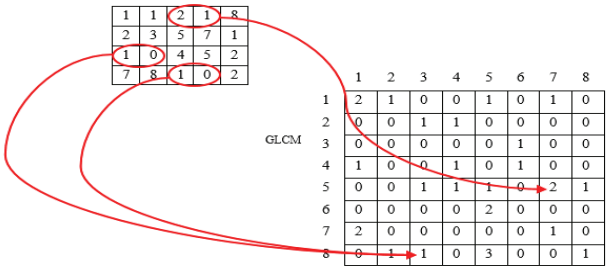



Fig. 6. Working of GLCM for the Gray-image of a sample image

From the matrix, we computed six values such as Variance, Angular Second Moment (ASM), Contrast, Inverse Difference Moment (IDM), Entropy, Homogeneity, Energy, and Dissimilarity [19, 20].

Now that we have finally calculated Statistical and Texture features using the parameters as mentioned earlier, we calculated individual performances using these feature vectors through the classification algorithms mentioned earlier in Section IV.

*D. Classification using SVM, RNN, and KNN:*

Classification is the final process of predicting the class of given sets. The main goal of a classification problem is to identify which new data will fall under what category/class. The input for this stage will be the feature vectors, and output is the classified label, corresponding to a relevant result, namely parasitic and non-parasitic. In our final step, we pass this dataset to three supervised learning classifiers SVM, RNN, and KNN. With its robust features, SVM uses a property known as a kernel trick that performs multiple data transformations and computes an optimal boundary between the classifying outputs. RNN has always been known for providing and maintaining higher accuracy, as it handles missing values and avoids overfitting of many trees in the same model. KNN has one of the most straightforward implementations with little cost and high robustness that requires tuning of only two parameters, i.e., the distance metric used and the number of neighbors.

## VI. Performance Evaluation Parameters

The various performance evaluation parameters that need to be considered while evaluating a malaria parasite classification system are as follows:

*1) Accuracy*: Accuracy determines how often the classifier shows this correct output. The accuracy of a system must be higher. This is shown in Equation (5).

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{5}$$

*2) Precision:* Precision is defined as the ratio of the number of accurate positive results generated to the total positive results predicted by the classifier. A higher precision denotes that the results are more relevant than the irrelevant ones. This is shown in Equation (6).

$$\text{Precision} = \frac{TP}{TP+FP} \tag{6}$$

*3) Recall:* Recall is the ratio of the correct positive results generated by the classifier to the number of all relevant samples. A higher recall denotes that the classifier returns more actual results. This is shown in Equation (7).

$$\text{Recall} = \frac{TP}{TP+FP} \tag{7}$$

*4) F1 measure:* F1-measure is the harmonic mean between precision and recall. It tells how accurate and robust the classifier is. This is shown in Equation (8).

$$\text{F1 measure} = \frac{2*recall*precision}{recall+precision} \tag{8}$$

## VII. Experimental Setup

The dataset was obtained from the Lister Hill National Center for Biomedical Communication (LHNCBC), a part of the National Library of Medicine (NLM). The total dataset

contains 27,558 segmented cell images extracted from thin blood smear slide images, containing 13,779 samples with equal instances each of malaria parasitized as well as uninfected cells. Since the cells have already been segmented, we did not require to perform grave image pre-processing techniques throughout the system. The cell images are in .png format. The following Fig. 7 and 8 are the sample cell images taken from the dataset. Python 3 is the platform over which the system was created using the Windows 10 operating system.

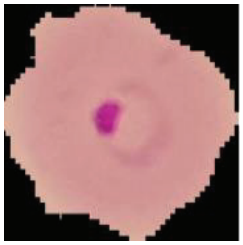


Fig. 7. Sample parasitized cell

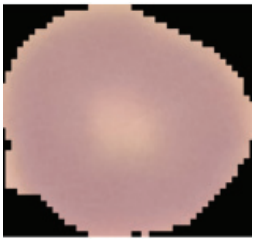


Fig. 8. Sample uninfected cell

## VIII. Results and Discussion

Concerning Equations (5), (6), (7), (8) from Section VII, we have evaluated the parameters of Accuracy, Precision, Recall, and F1-score, and we have accordingly found the results of the outputs as follows:

TABLE III. Result obtained for CNN and Bins Approach with ML Classifiers

| Approach Parameter | CNN | Bins with SVM | Bins with RNN | Bins with KNN |
|---|---|---|---|---|
| Accuracy | 94% | 93.1% | **96.3%** | 86.5% |
| Precision | 93% | 98.2% | **97.1%** | 94.8% |
| Recall | 96% | 88.1% | **95.6%** | 77.6% |
| F1 Score | 94% | 92.9% | **96.3%** | 85.3% |

From Table III, we can observe that Bins Technique works considerably as good as CNN, with accuracy scaling at 93.1%. Thus, we can conclude that SVM works reasonably well with our approach during classification, although slightly lesser than the CNN approach. The above table also delineates that RNN is the best classifier for this approach as it achieves an accuracy of 96.3%, which is much better than the rest of the methods mentioned so far (the results are highlighted in yellow). Moreover, KNN classifier hails the least in terms of accuracy; however, owing to its drawbacks in misclassification of labels, it holds a relatively more substantial accuracy. It should be noted that we employed KNN classification using neighbors n=2 in Table III.

Although models like CNN achieved a great accuracy result, there are a few drawbacks to consider using this approach. Firstly, CNNs take much time during the training of the model. During this process, we took approximately 30 to 45 minutes to train the model on a simple Windows 10 PC. Although the processing time may differ in other PCs, loading even a modest model and performing necessary computations is more expensive.

With the results generated above, we still wished to improve our accuracy further. As mentioned in Section VI, we decided to extract statistical and texture features from blood smear images. The following Table IV describes the results obtained. We have also employed a nomenclature of the respective headings using the following abbreviations:

***Stat_SVM = Statistical feature vector with SVM***
***Stat_RNN = Statistical feature vector with RNN***
***Stat_KNN = Statistical feature vector with KNN***
***Text_SVM = Texture feature vector with SVM***
***Text_RNN = Texture feature vector with RNN***
***Text_KNN = Texture feature vector with KNN***

TABLE IV. Result obtained from SVM, RNN and KNN classifiers using statistical and texture features

| Approach parameter | Stat_ SVM | Stat_ RNN | Stat_ KNN | Text_ SVM | Text_ RNN | Text_ KNN |
|---|---|---|---|---|---|---|
| Accuracy | 81.4% | 81.4% | 68.2% | 79.6% | 88.2% | 70.3% |
| Precision | 86.4% | 82.5% | 74.2% | 83.0% | 88.6% | 79.1% |
| Recall | 75.6% | 80.7% | 56.4% | 75.7% | 88.2% | 55.9% |
| F1-score | 80.6% | 81.6% | 64.1% | 79.2% | 88.4% | 65.5% |

We infer from the table that the results obtained provide much lesser and inadequate accuracy than Bins Approach itself. Thus, we conclude that statistical and texture features alone are not enough to perform substantial classification. This was when we decided to prepare a ***fusion feature vector*** that combines the features of ***bins, color moments, and GLCM parameters all in one feature dataset***. Since this fusion dataset is a conglomeration of over 27 features, we cannot fit all of them during classification. Thus, we employed feature selection using Recursive Feature Elimination (RFE) with Logistic Regression. We wanted to test whether these features are enough to see an improvement in the classification that we saw in Tables IV and V, and the results have met our expectations. Note that the nomenclature used as **"Fusion"** in the following table is the same as fusion feature vectors described in the paragraph earlier. It is only for simplicity that we reduced it to this term.

***Fusion Feature Vector = (Bins + Color Moments + GLCM parameters)***

TABLE V. Result obtained for Fusion Dataset with ML Classifiers

| Approach Parameter | Fusion Features with SVM | Fusion Features with RNN | Fusion Features with KNN |
|---|---|---|---|
| Accuracy | 94.6% | **96.1%** | 95.2% |
| Precision | 98.6% | **97.8%** | 97.4% |
| Recall | 90.7% | **94.4%** | 92.8% |
| F1 Score | 94.5% | **96.1%** | 95.1% |

If we compare the results of Bins Technique implemented here in our system with a previous approach [21] to parasite classification, one of the computer scientists performed different preprocessing techniques. The workflow describes the process of resizing, grayscaling, histogram normalizing, and extracting contrast features of black and white colors by dividing image histograms into ten bins (from 0.1 to 10). The accuracy shown was barely 83%. However, from the table above, our approach works much better, with accuracy scaling at 94.6% with SVM, and 96% with RNN classifier. Note that during the classification of the fusion dataset, the KNN classifier was predicted with neighbors n=4.

With this respect, when we infer Table V, we can confirm that Bins Technique indeed improves the accuracy of the classification as it dramatically reduces the dimensionality to provide numerical integers as against text-based features of traditional classification approaches. It plays a significant role in terms of all the performance evaluation parameters with better values.

## IX. Conclusion and Future Scope

Through the course of our research work, we primarily focused on image processing techniques to process stained thin blood smear images for feature extraction, as well as machine learning techniques for the final classification. To achieve this, we employed the concepts of CNN and Bins Approach, respectively.

The drawbacks of CNN mentioned in Section VIII outweigh the advantage of high accuracy of neural networks and impact the quality of feature engineering because of a lack of insights about the actual features being extracted. We addressed them using the Bins Technique that believed to be a simple but effective technique deals with the actual color image contents and represents them in the form of a compact feature as part of the feature engineering process [12,13]. Further, these features were fused with statistical moments and texture features using GLCM.

After comparing the performances of all approaches, it can be deduced that Bins Approach indeed plays a significant role in feature engineering. It produces better accuracy and classification in terms of precision as much as or better than CNN (values: 93% and 96%), shown in Table III. Also, if it is compared with the results of the GLCM and Texture approach shown in Table IV, it has proved far better in terms of all parameters. Thus, a further fusion of GLCM, Texture with Bins has also have deduced better results. Based on overall experimentation, along with the results and discussion, we can delineate the following points:

*1) Speed:* The classification system is faster than most other traditional techniques employed, if not the best of them all. It does not require any expert to perform tedious procedures to obtain the required results.

*2) User-friendly:* The proposed system is efficient, straightforward, and simple to implement.

Thus, Bins Technique can prove to be an alternative approach in feature extraction techniques as it also reduces the computational complexity to a great extent with the feature vector of size eight components. We can deploy such an approach in the medical domain, where proper and prompt diagnosis and treatment are of paramount importance.

With an extension of other features, changing color space, improving classifier performance using Grid Search, or using ensemble ML techniques such as Bagging or Boosting classifiers like AdaBoost, we may further try to improve the accuracy and performance of the model that can be of significant contribution to medical science.

## References


[1] Caraballo, Hector, and Kevin King. "Emergency department management of mosquito borne illness: malaria, dengue, and West Nile virus." Emergency medicine practice 16, no. 5 (2014): p. 1-23.

[2] 2015. *Achieving The Malaria MDG Target: Reversing The Incidence Of Malaria 2000-2015*. Geneva: World Health Organization and the United Nations Children's Fund, p. 6-18.

[3] World Health Organization, World Malaria Report 2019. Geneva, Global Malaria Programme, WHO Global, 2019, p. xii-xiii.

[4] Organization, W., 2016. Malaria Microscopy Quality Assurance Manual - Version 2. Geneva: World Health Organization.

[5] Olugboja, and Z. Wang, "Malaria parasite detection using different machine learning classifiers," 2017 International Conference on Machine Learning and Cybernetics (ICMLC), Ningbo, 2017, pp. 246-250.

[6] P. A. Pattanaik, T. Swarnkar and D. Sheet, "Object detection technique for malaria parasite in thin blood smear images," 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, 2017, pp. 2120-2123.

[7] A. Bashir, Z. A. Mustafa, I. Abdelhameid and R. Ibrahem, "Detection of malaria parasites using digital image processing," 2017 International Conference on Communication, Control, Computing and Electronics Engineering (ICCCCEE), Khartoum, 2017, pp. 1-5.

[8] Vinayak K. Bairagi and Kshipra C. Charpe, "Comparison of Texture Features Used for Classification of Life Stages of Malaria Parasite," International Journal of Biomedical Imaging, vol. 2016, Article ID 7214156, 9 pages, 2016.

[9] K. Silamut, N.H. Phu, C. Whitty, *et al.* "Image analysis and machine learning for detecting malaria "Am J Pathol, 155 (2017), pp. 395-410

[10] Somasekar J, Reddy B, Reddy E, Lai C, (2018) "Detection of malaria parasite in giemsa blood sample using image processing", International Journal on Computer Science and Engineering 3: 2251-2256.

[11] H. B. Kekre, Kavita Sonawane, " Image Retrieval Using Histogram Based Bins of Pixel Counts and Average of Intensities" (pp. 74-79, (IJCSIS) International Journal of Computer Science and Information Security, Vol. 10, No. 1, 2012

[12] H. B. Kekre, Kavita Sonawane, " Image Retrieval Using Histogram Based Bins of Pixel Counts and Average of Intensities" (pp. 74-79, (IJCSIS) International Journal of Computer Science and Information Security, Vol. 10, No. 1, 2012

[13] H. B. Kekre, Kavita Sonawane, "Comparative Performance of Linear and CG Based Partitioning Of Histogram for Bins Formation in CBIR", International Journal of Engineering Research and Development e-ISSN: 2278-067X, p-ISSN : 2278-800X, www.ijerd.com Volume 5, Issue 7 (January 2013), PP. 75-83

[14] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," *2017 International Conference on Engineering and Technology (ICET)*, Antalya, 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.

[15] Y. LeCun, Y. Bengio, G. Hinton, "Deep learning." Nature. 2015 May28;521(7553), pp. 436-44.

[16] D. Palaz, M. Magimai.-Doss and R. Collobert, "Convolutional Neural Networks-based continuous speech recognition using raw speech signal," 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brisbane, QLD, 2015, pp. 4295-4299, doi: 10.1109/ICASSP.2015.7178781.

[17] X. Liu, T. Kawanishi, X. Wu and K. Kashino, "Scene text recognition with high performance CNN classifier and efficient word inference," 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, 2016, pp. 1322-1326, doi: 10.1109/ICASSP.2016.7471891.

[18] Darmatasia and M. I. Fanany, "Handwriting recognition on form document using convolutional neural network and support vector machines (CNN-SVM)," 2017 5th International Conference on Information and Communication Technology (ICoIC7), Malacca City, 2017, pp. 1-6, doi: 10.1109/ICoICT.2017.8074699.

[19] Suresh, Annamalai & Shunmuganathan, K.. (2012). Image Texture Classification using Gray Level Co-Occurrence Matrix Based Statistical Features. European Journal of Scientific Research. 75.

[20] Zulpe, Nitish & Pawar, V. (2012). GLCM textural features for Brain Tumor Classification. IJ CSI. 9. 354-359.

[21] Fiott, S., 2018. Malaria Parasite Detection Using An SVM. [online] Stefanfiott.com. Available at: <https://www.stefanfiott.com/machine-learning/malaria-parasite-detection-using-an-svm/> [Accessed 23 June 2020].