Acute Lymphoblastic Leukemia Classification with Blood Smear Microscopic Images Using Taylor-MBO based SVM

Nashwan Jasim Hussein
Department of Information Technology, AL-Furat Al-Awsat Technical University, Technical College of Management, Najaf, Iraq.
E-mail: alsalihnashwan@atu.edu.iq, alsalihnashwan@yahoo.com

Received February 16, 2021; Accepted June 14, 2021
ISSN: 1735-188X
DOI: 10.14704/WEB/V18SI02/WEB18104

Abstract

Acute lymphoblastic leukemia (ALL) is a serious hematological neoplasis that is characterized by the development of immature and abnormal growth of lymphoblasts. However, microscopic examination of bone marrow is the only way to achieve leukemia detection. Hence, an effective leukemia detection approach is designed using the proposed Taylor-Monarch Butterfly Optimization based Support Vector Machine (Taylor-MBO based SVM). However, the proposed Taylor-MBO is designed by the integration of Taylor series and Monarch Butterfly Optimization (MBO), respectively. The sparking process is designed to perform automatic segmentation of blood smear image by estimating optimal threshold values. By extracting the features, such as texture features, statistical and grid-based features from the segmented smear image, the performance of classification is increased with less training time However, the proposed Taylor-MBO based SVM obtained better performance using the metrics, such as accuracy, sensitivity, and specificity with the values of 94.5751%, 95.526%, and 94.570%, respectively.

Keywords

Acute Lymphoblastic Leukemia (ALL), Taylor Series, SVM, Sparking Process, Blood Smear Image.

Introduction

Microscopic analysis with blood smear images is the major source of information, which shows a modification of growth of specific disease. However, the blood smear image contains the background, platelets, red blood cells and leukocyte cells (Mohammed and et al 2017). The basic procedure involved in the automatic computer assisted classification (CAC) system is leukocyte segmentation that is mainly used for classifying and identifying the malignant cells with respect to the examination of blood smear image (R.D. Labati and

et al 2011). In the recent haematology effect, blood cancer is the severe lymphoblastic leukemia that is commonly found in the medical practices such that it can be characterized over the abnormal functionality and the activation of blood smears. However, the abnormal functions are examined using physical examination by screening the blood smears. This approach is labour intensive and error prone task for the haematologist (Sara swat, M., Arya and et al 2013). Therefore, the haematologist requires computer aided diagnostic system (CAD) to compact with the issues of this system and enable to have the capability of finding adolescent leukemic cells separately from the adult healthy cells (M. Habibzadeh, A. Krzyzak, T. Fevens and A. Sadr 2011).

The organization of the paper is made as follows: section 2 presents the review of various existing leukemia detection approaches and section 3 elaborates proposed Taylor-MBO based SVM. Section 4 describes the results and discussion of proposed approach and section 5 concludes the paper.

Literature Survey

Various existing leukemia detection methods are surveyed in this section. (Mishra, S *et al 2019).* Introduced an ad boost algorithm for the detection of leukemia. Initially, the images were pre-processed using triangle method. The texture features were extracted using discrete orthonormal s-transform (DOST) and the linear discriminant analysis were used to decrease the dimensionality of features. Here, base classifier was designed by random forest (RF) for classifying leukemia. This method achieved better accuracy, but it failed to perform the sub-classification of ALL. (Laosai, J. and Chamnongthai, K 2018). Introduced a morphological cell-subtype classification model using coarse-to-fine idea for detecting ALL. Initially, the AML, ALL and the healthy cell groups were separated and the cells that were classified into the AML and ALL were grouped into subtypes. However, the subtypes specified the results of morphological classification. Finally, these results were utilized to validate the cell subtypes using CD markers. It increased the accuracy but lacks in the enhancement process. (Jothi, G *et al 2018).* Introduced a backtracking search optimization (BSA) based clustering for segmenting ALL image. Here the features, like wavelet, texture, color, statistical and morphological features were extracted from leukemia image. It increased the classification accuracy but it was not applicable for the multimodal data, like CT, MRI, mammography and PET to predict leukemia. (Moshavash, Z *et al 2018).* Introduced an automatic approach for segmenting leukocyte from the blood microscopic images. This method segments the leukocyte in a robust way under uneven imaging and lighting conditions. It increased the classification accuracy and was highly effective for various levels of analysis. However, it was not applicable for the cells that were clumped

and overlapped. (Rawat, J *et al 2017).* Introduced a genetic algorithm based SVM for classifying myeloblast cell, acute lymphoblast, and the blast subtypes. However, the shape, color and the texture features were obtained and SVM was employed for the classification process. However, this method obtained better accuracy but it was not applicable for large sized database. (Vogado, L.H *et al 2018).* Introduced the convolutional neural network (CNN) for diagnosing the leukemia using blood images. The features were extracted and the robustness was validated by pre-training CNN. IT required a greater number of attributes for classifying the images with number of leukocytes. It does not use segmentation process and it helped the patients and physicians for diagnosing this disease. (Rawat, J *et al 2017).* Introduced an automated hybrid hierarchical classification model for classifying cancerous lymphoblast cells with the hybrid hierarchical classifier. It computed the true edges of cytoplasm and nucleus, and classified the normal and malignant cells. The localization of lymphoblast was achieved and increased the classification accuracy. However, it failed to enhance the exactness and the strength of classification task. Mohapatra, (S *et al 2014).* Introduced an ensemble model for diagnosis of ALL with the blood microscopic images. Authentic and accurate diagnosis of ALL was obtained using the features selected and the segmentation process. It increased the recognition rate and does not consider the sub-classification of ALL. The computation speed was low.

Proposed Taylor-Monarch Butterfly Optimization based Support Vector Machine for Acute Lymphoblastic Leukemia Detection

Leukemia is a disease that contains more amounts of fatalities between children and adolescents. It is variety of cancer that occurs in bone marrow and sources abnormality of WBC. Hence an automatic classification method is designed to detect leukemia with blood smear images by proposed Taylor-MBO based SM. The proposed model accomplishes the detection practice by involving the phases, like pre-processing, segmentation, feature extraction and leukemia classification. At first, the input blood smear image is fed to pre-processing stage, where input image is pre-processed to remove the noise. The pre-processed image is fed to segmentation phase, where the image is segmented by employing sparking process. Thereafter, the features associated with the smear images are extracted such that the features include texture features, statistical features and grid based features. Based on extracted features, leukemia classification is carried out using SVM, which is tuned by proposed Taylor-MBO algorithm. Accordingly, proposed Taylor-MBO is designed by the integration of Taylor series (M. Habibzadeh and et al 2011). Figure 1 represents the schematic view of proposed Taylor-MBO based SVM.

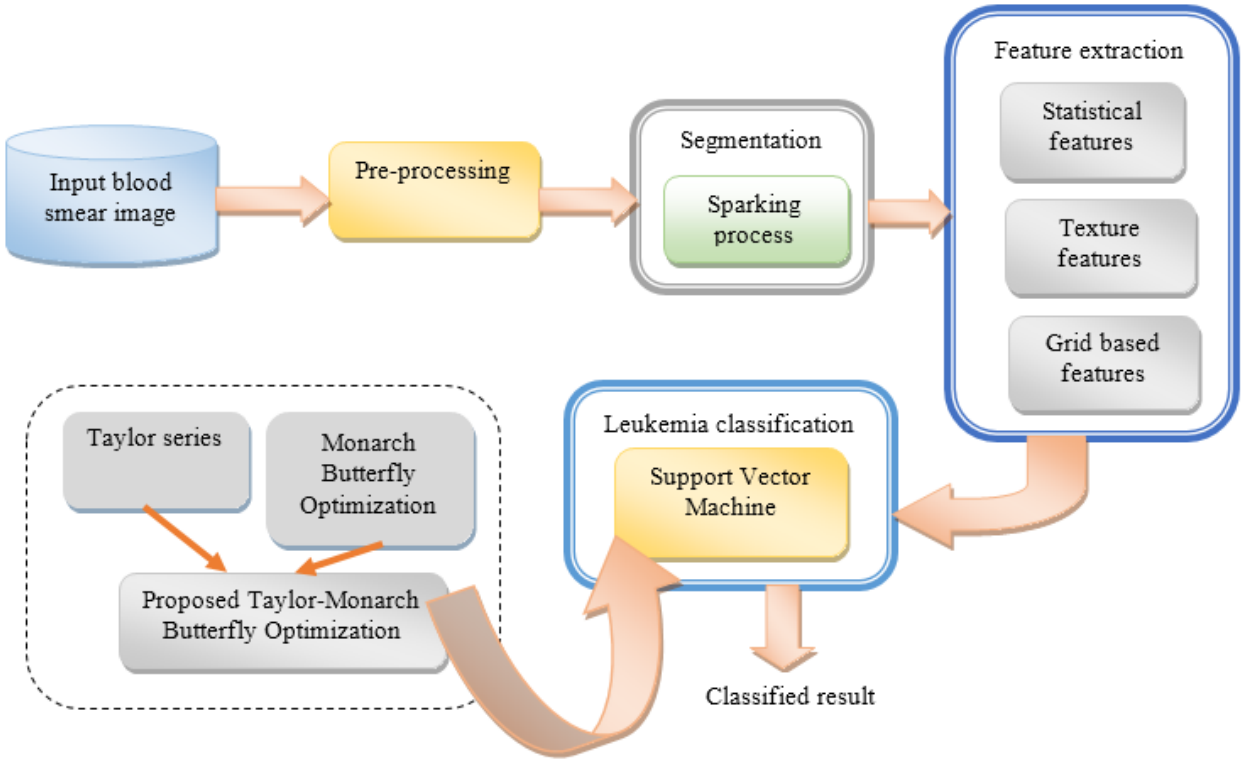


Figure 1 Schematic diagram of proposed Taylor-MBO based SVM

Algorithm 1 Pseudo code of proposed Taylor-MBO based SVM

Sl. No	**Pseudo code of proposed Taylor-MBO based SVM**
1	Input:
2	Output:
3	**Begin**
4	Initialize the population
5	Compute fitness
6	**while** $(r < G_{\max})$ or best solution is not found
7	do
8	Sort all monarch butterflies based on fitness
9	Divide the butterfly individuals into two subpopulations as land-1 and land-2
10	**for** $m = 1$ to T_1
11	do
12	**for** $n = 1$ to G
13	do
14	Randomly generate $rand$ by the uniform distribution
15	$w = rand * h$
16	**if** $w \le l$
17	Randomly select a butterfly in subpopulation 1 and represent as (w_1)
18	Generate n^{th} element of R_m^{r+1} using Eq. (20)

19	**else**
20	Randomly select a monarch butterfly in subpopulation 2 and represent as (w_2)
21	Generate n^{th} element of R_m^{r+1} using Eq. (22)
22	**end if**
23	**end for** n
24	**end for** m
25	**for** $z = 1$ to T_2
26	do
27	Compute dR using Eq. (33)
28	Compute δ using Eq. (34)
29	**for** $n = 1$ to G
30	do
31	Randomly generate $rand$ by the uniform distribution
32	**if** $rand \le l$ then
33	Generate n^{th} element of R_z^{r+1} using Eq. (23)
34	**else**
35	Randomly select a monarch butterfly in subpopulation 2 and represent as (w_3)
36	Generate n^{th} element of R_z^{r+1} using Eq. (24)
37	**if** $rand > t$
38	$R_{zn}^{r+1} = \frac{1}{0.5}\left[\begin{array}{l} 1.3591R_{zn}^{r-1} - 1.359R_{zn}^{r-2} + 0.6795R_{zn}^{r-3} - 0.2259R_{zn}^{r-4} \\ \quad + 0.0555R_{zn}^{r-5} - 0.0104R_{zn}^{r-6} + 1.38e^{-3}R_{zn}^{r-7} - 9.92e^{-5}R_{zn}^{r-8} \end{array}\right] + \delta(dR_n - 0.5)$
39	**end if**
40	**end if**
41	**end for** n
42	**end for** z
43	Combine the two subpopulation into single population
44	Evaluate it based on the updated position
45	$r = r + 1$
46	**end while**
47	Return best solution
48	**end**

Results and Discussion

This section presents discussion of proposed Taylor-MBO based SVM by considering the metrics, like specificity, sensitivity, and accuracy. Figure 2 shows image results of proposed Taylor-MBO based SVM for ALL classification.

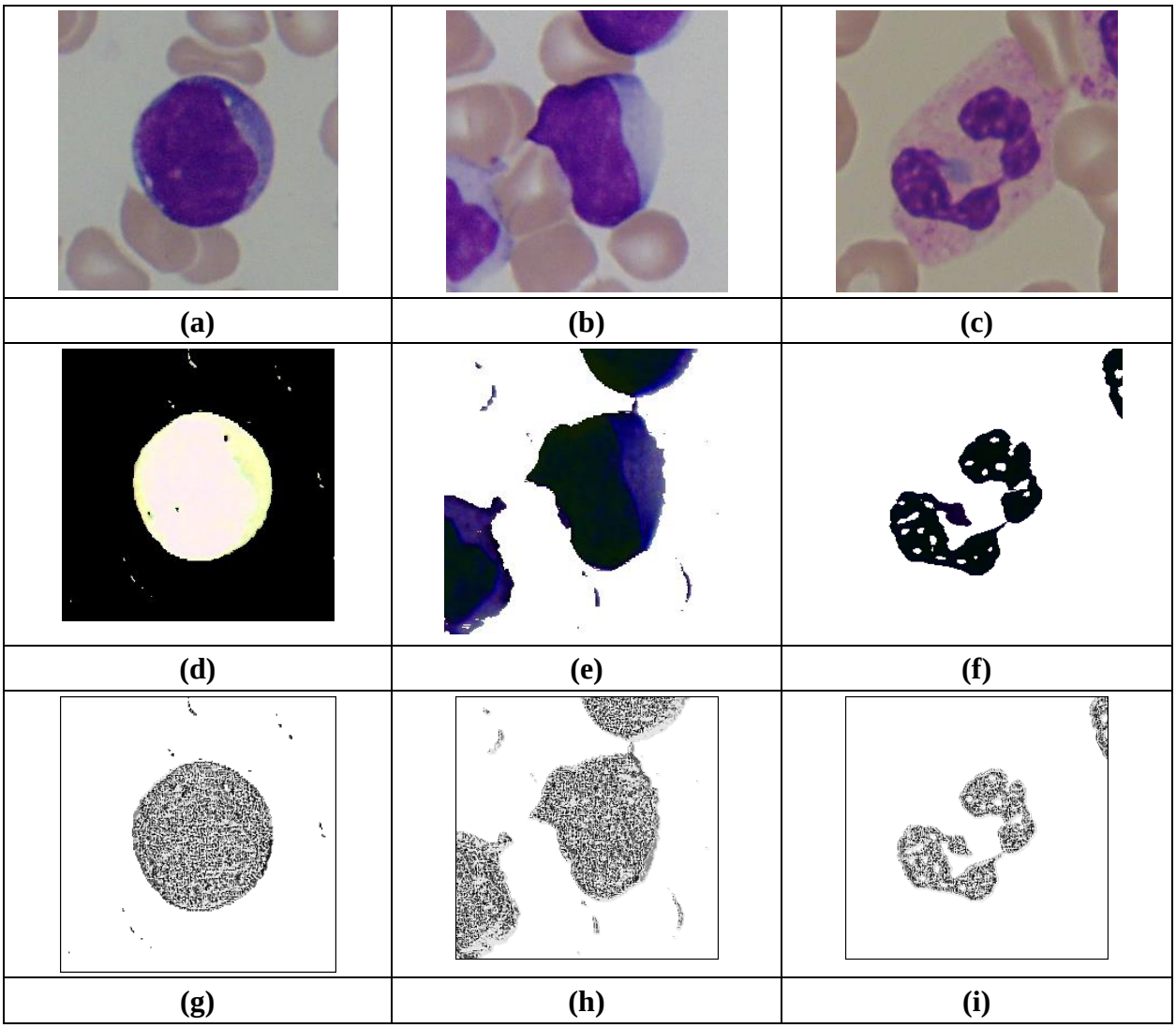


Figure 2 Image results, a) input image-1, b) input image-2, c) input image-3, d) segmented image-1, e) segmented image-2, f) segmented image-3, g) LOOP feature for image-1, h) LOOP feature for image-2, i) LOOP feature for image-3

Analysis by Varying SVM Kernels

The analysis of Taylor-MBO based SVM based on SVM kernels by varying K-fold is shown in figure 3. Figure 3 a) portrays the analysis of accuracy with SVM kernels. For K-fold=7, the accuracy of linear kernel is 92.274%, quadratic kernel is 92.284%, radial kernel is 92.293%, rbf kernel is 92.303%, polynomial kernel is 92.312%, mlp kernel is 92.321%, and proposed Taylor-MBO based SVM is 92.331%, respectively. For K-fold=8, the accuracy of linear kernel is 93.214%, quadratic kernel is 93.224%, radial kernel is 93.233%, rbf kernel is 93.243%, polynomial kernel is 93.252%, mlp kernel is 93.262%, and proposed Taylor-MBO based SVM is 93.271%, respectively. For K-fold=9, the accuracy of linear kernel is 93.511%, quadratic kernel is 93.520%, radial kernel is 93.530%, rbf

kernel is 93.539%, polynomial kernel is 93.548%, mlp kernel is 93.558%, and proposed Taylor-MBO based SVM is 93.567%, respectively.

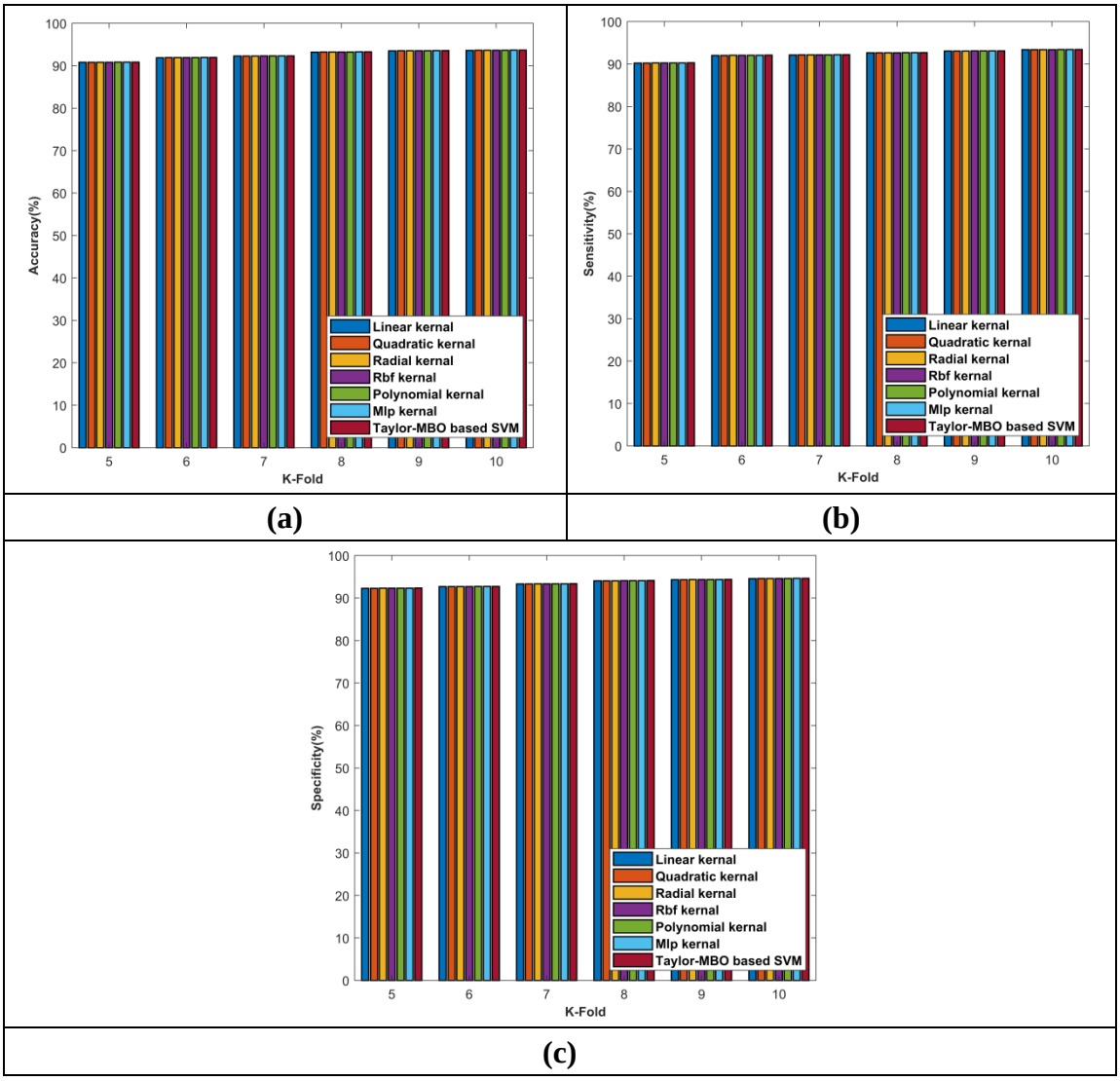


Figure 3 Analysis based on SVM kernels with K-fold, a) accuracy, b) sensitivity, c) specificity, d) TPR

Analysis of ROC Curve

Figure 4 shows the ROC curve of proposed Taylor-MBO based SVM. When FPR=50%, the TPR of existing ADBRF, RST+Jaya algorithm, SVM and the proposed Taylor-MBO based SVM is 78.086%, 95.515%, 96.505%, and 97.495%. When FPR=70%, the TPR of existing ADBRF, RST+Jaya algorithm, SVM and the proposed Taylor-MBO based SVM is 82.121%, 95.674%, 96.664%, and 97.654%.

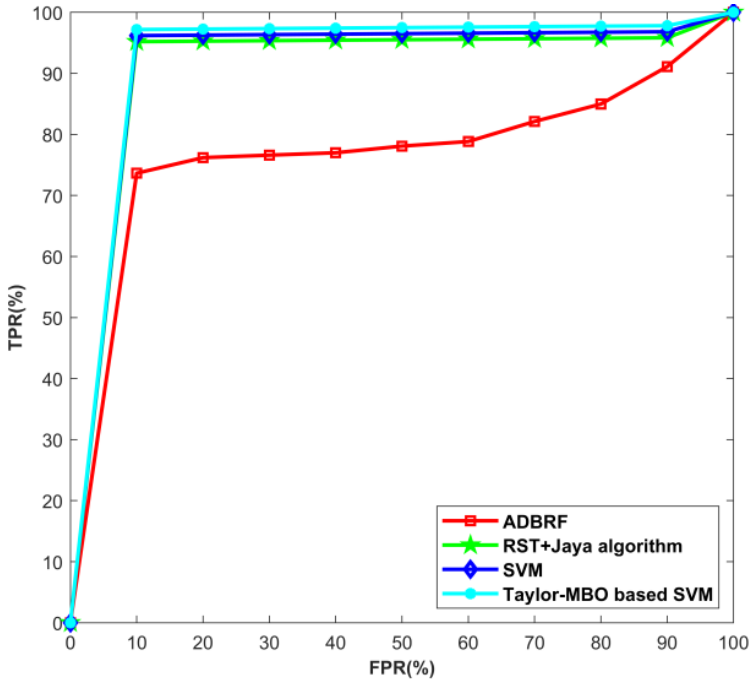


Figure 4 ROC curve of proposed Taylor-MBO based SVM

Figure 5 represents the ROC curve based on SVM kernels. When FPR=50%, TPR of linear kernel is 97.505%, quadratic kernel is 97.515%, radial kernel is 97.525%, rbf kernel is 97.535%, polynomial kernel is 97.545%, mlp kernel is 97.555%, and proposed Taylor-MBO based SVM is 97.565%. When FPR=60%, TPR of linear kernel is 97.584%, quadratic kernel is 97.594%, radial kernel is 97.604%, rbf kernel is 97.614%, polynomial kernel is 97.624%, mlp kernel is 97.634%, and proposed Taylor-MBO based SVM is 97.644%.

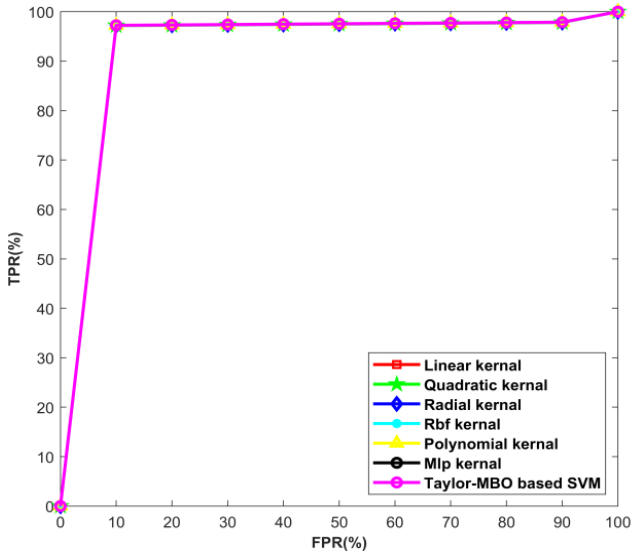


Figure 5 ROC curve based on SVM kernels

Conclusion

The proposed approach accomplishes the detection process by concerning the phases, like pre-processing, segmentation, feature extraction and leukemia classification. The input blood smear image is passed to the pre-processing phase, where the image is pre-processed more effectively to remove the noise. The pre-processed image is fed to segmentation phase, where the image is segmented using sparking process. The proposed Taylor-MBO is designed by the integration of Taylor series and MBO. However, the proposed Taylor-MBO based SVM obtained better performance using the metrics, such as accuracy, sensitivity, and specificity with the values of 94.5751%, 95.526%, and 94.570% by varying training data. The future dimension of research would be the consideration of some other optimization algorithms for training the classifier to enhance the classification performance.

References

Mishra, S., Majhi, B., & Sa, P.K. (2019). Texture feature based classification on microscopic blood smear for acute lymphoblastic leukemia detection. *Biomedical Signal Processing and Control, 47,* 303-311.

Laosai, J., & Chamnongthai, K. (2018). Classification of acute leukemia using medical-knowledge-based morphology and CD marker. *Biomedical Signal Processing and Control, 44,* 127-137.

Jothi, G., Inbarani, H.H., Azar, A.T., & Devi, K.R. (2019). Rough set theory with Jaya optimization for acute lymphoblastic leukemia classification. *Neural Computing and Applications, 31*(9), 5175-5194.

Moshavash, Z., Danyali, H., & Helfroush, M.S. (2018). An automatic and robust decision support system for accurate acute leukemia diagnosis from blood microscopic images. *Journal of digital imaging, 31*(5), 702-717.

Rawat, J., Singh, A., Bhadauria, H.S., Virmani, J., & Devgun, J.S. (2017). Computer assisted classification framework for prediction of acute lymphoblastic and acute myeloblastic leukemia. *Biocybernetics and Biomedical Engineering, 37*(4), 637-654.

Vogado, L.H., Veras, R.M., Araujo, F.H., Silva, R.R., & Aires, K.R. (2018). Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification. *Engineering Applications of Artificial Intelligence, 72,* 415-422.

Rawat, J., Singh, A., Bhadauria, H.S., Virmani, J., & Devgun, J.S. (2017). Classification of acute lymphoblastic leukaemia using hybrid hierarchical classifiers. *Multimedia Tools and Applications, 76*(18), 19057-19085.

Mohapatra, S., Patra, D., & Satpathy, S. (2014). An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images. *Neural Computing and Applications, 24*(7), 1887-1904.

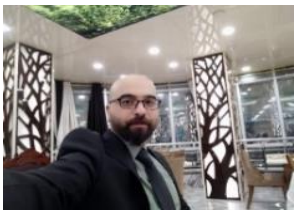
Mohammed, E.A., Mohamed, M.M., Naugler, C., & Far, B.H. (2017). Toward leveraging big value from data: chronic lymphocytic leukemia cell classification. *Network Modeling Analysis in Health Informatics and Bioinformatics, 6*(1), 6.

Labati, R.D., Piuri, V., & Scotti, F. (2011). All-IDB: The acute lymphoblastic leukemia image database for image processing. *In 18th IEEE International Conference on Image Processing*, 2045-2048.

Habibzadeh, M., Krzyzak, A., Fevens, T., & Sadr, A. (2011). Counting of RBCs and WBCs in noisy normal blood smear microscopic images. *In Medical Imaging 2011: Computer-Aided Diagnosis, International Society for Optics and Photonics, 7963*, 79633I.

Saraswat, M., Arya, K. V., & Sharma, H. (2013). Leukocyte segmentation in tissue images using differential evolution algorithm. *Swarm and Evolutionary Computation, 11*, 46-54.

Author Profile



Nashwan Jasim Hussein received B.Sc. degree in Computer and software engineering from university of technology –Baghdad from Iraq, and the MSc Engineering. Degree in computer engineering and information technology from university of PUNE –VIT collage from 2010 to 2012, and he was joined in Ph.D. Research student in 2015 in electronics, communication engineering department, telecommunication in Huazhong University of science and technology, China. His general research interests lie in image processing, signaling analysis, and object detection.