

Improved Classification of Histopathological UImages using the Feature Fusion of Thepade Sorted Block Truncation Code and Niblack Thresholding

Sudeep D. Thepade[#], Abhijeet V. Bhushari^{*}

^{#, *}*Dept. of Computer Engineering, Pimpri Chinchwad College of Engineering, Savitribai Phule Pune University, Pune, India*

[#]*sudeepthepade@gmail.com*, ^{*}*abhijeetvijaybhushari242002@gmail.com*

Received 9th of January, 2022; accepted 10th of May 2023

Abstract

Histopathology is the study of disease-affected tissues, and it is particularly helpful in diagnosis and figuring out how severe and rapidly a disease is spreading. It also demonstrates how to recognize a variety of human tissues and analyze the alterations brought on by sickness. Only through histopathological pictures can a specific collection of disease characteristics, such as lymphocytic infiltration of malignancy, be determined. The "gold standard" for diagnosing practically all cancer forms is a histopathological picture. Diagnosis and prognosis of cancer at an early stage are essential for treatment, which has become a requirement in cancer research. The importance and advantages of classification of cancer patients into more-risk or less-risk divisions have motivated many researchers to study and improve the application of machine learning (ML) methods. It would be interesting to explore the performance of multiple ML algorithms in classifying these histopathological images. Something crucial in this field of ML for differentiating images is feature extraction. Features are the distinctive identifiers of an image that provide a brief about it. Features are drawn out for discrimination between the images using a variety of handcrafted algorithms. This paper presents a fusion of features formed with Thepade sorted block truncation code (TSBTC) and the Niblack thresholding algorithm for classifying histopathological images. The experimental validation is done using 960 images present in the Kimiapath-960 dataset of histopathological images with the help of performance metrics like sensitivity, specificity, and accuracy. Better performance is observed by an ensemble of TSBTC N-ary and Niblack's thresholding features as 97.92% of accuracy in 10-fold cross-validation.

Key Words: Thepade Sorted Block Truncation Code, Niblack Thresholding, Feature Fusion, Ensemble learning.

Correspondence to: sudeepthepade@gmail.com[#]

Recommended for acceptance by Angel D. Sappa

<https://doi.org/10.5565/rev/elcvia.1644>

ELCVIA ISSN: 1577-5097

Published by Computer Vision Center / Universitat Autònoma de Barcelona, Barcelona, Spain

1 Introduction

Histopathology is the study of illness symptoms utilizing a microscopic inspection of a prepared and fixed surgical specimen on glass slides. This quantitative examination of digital pathology is crucial for understanding the underlying causes of a given disease from a diagnostic standpoint. There need to be more qualified practitioners to visually diagnose these photos. Early diagnosis considerably increases the likelihood of receiving the proper care and survival of the patient. However, this procedure is time-consuming and frequently causes conflict between pathologists. Systems for computer-aided diagnosis (CAD) have the potential to increase both the diagnostic rate and accuracy.

Traditional machine-learning techniques are applied to classify histopathology pictures for cancer detection. Histopathological images are subject to complex models or models with significant bias and high variance. The model may perform exceptionally well on training datasets, i.e., it produces low bias, but it may fail on test datasets and give significant variance. This happens when the model treats even random oddities of data as patterns. Therefore, ensemble learning techniques are developed to increase the model's accuracy (estimate). Multiple models are trained using machine-learning algorithms in an ensemble. It combines individual model predictions with low-performing classifiers (also known as weak learners or base learners) to produce the final prediction.

The contributions of the method proposed here are:

- Feature Fusion of TSBTC and Niblack Thresholding for improving the histopathological image classification.
- Exploring the 12 ML algorithms and nine ensembles to classify histopathological images.
- The experimental exploration of the proposed method with the 960 images of the histopathological dataset (KIMIA-PATH 960) is done using performance metrics like Accuracy, Sensitivity, and Specificity.

The paper's organization has section 2, which holds a literature review, and section 3 represents the proposed method. Results and Discussion of the explorations done for the proposed method are put forth in section 4. The concluding summarisation is in section 5.

2 Related Work

Many previous works regarding histopathology and tissue image classification use various techniques. These techniques aim to produce better results than the previous one. Some of these techniques analyze the spatial structure of histopathology images and have been followed in some of the previous research. Other techniques include classification using segmentation, using global features and window-based features. Other classification methods use deep features, which require much training with large datasets of balanced images.

A comparison analysis explaining the prospective future of deep features, BoVW, and the local binary patterns (LBP) methods of histopathological image classification is performed by Meghna et al. [1]. They have attempted classification using the LBP histograms, deep features, and BoVW through leave-one-out validation. The classification accuracy derived using LBP was 90.62%, whereas the best accuracy was 94.72%, derived through deep features. The strength of the approach presented is the automated feature extraction of deep learning after data training, the requirement of a large amount of data for training of deep learning models is the limitation of the approach.

A. Ganguly et al. [2] performed histopathological image classification on the dataset used in this paper, i.e., KIMIA Path 960. They have used a customized version of pre-trained Resnet50 for this

dataset. They have also used a five-layer CNN for classification and explored various optimization algorithms like AdaMax, Adam, Radam, and AdamW with the histopathological dataset. Their best accuracy reached 99.9%. That makes them better than the accuracies achieved by IRRCNN, i.e., 98.79%, and BOVW, which is 96.5%. The disadvantage of the method is the need for more time and effort to train the residual neural networks used.

A pre-trained convolutional neural network for histopathological image classification using the KIMIA Path 960 dataset with various combinations of GLCM and MobileNetV2 is proposed by Anish et al. [3]. The best results were observed for the technique GLCM + Mean of Sorted Gray Values for the classification using neural networks, which provided an AUC of 0.999, Precision of 0.951, F1 Score of 0.951, and a Recall of 0.951. This approach is able to perform more complex activities resulting in robust features as compared to other algorithms.

A study comparing three classification models: support vector machines, artificial neural networks, and decision trees, is presented by Taha et al. [4]. The dataset used for this study is KIMIA Path960, where the features are extracted using a pre-trained deep network, local binary patterns (LBP), and the histogram of gradients. The highest accuracy received in this study is 90.52% for SVM classification using local binary patterns, whereas the accuracy received for deep features is 81.14%. The use of LBP restricts the captured structural information of the images is the drawback of this method.

Alexander et al. [5] utilized CNN (Convolutional Neural Networks) for image classification of a histopathological dataset of breast cancer. They have used the microscopic histopathological stained images presented in the ICIAR 2018 Grand Challenge. Their approach uses neural networks along with gradient-boosted trees for classification. The advantage of this approach is the automatic detection of the discriminating features by CNN without any human supervision. Their results for 2-class and 4-class classification were 93.8% accuracy and 87.2% accuracy, respectively.

Overall, feature generalization is found to be a challenge in most of the existing approaches being used in histopathological image classification.

Table 1 gives us a Brief about the techniques used by previous researchers and the accuracies obtained while classifying histopathological images in the KIMIA Path960 dataset.

Authors	Year of Publication	Dataset used	Technique used	Advantages	Limitations
Ambarish et al. [2]	2020	KIMIA Path960	Customized Resnet50 with various optimization algorithms	The model has over 23 million trainable parameters, which makes it better for image recognition.	Using residual neural networks requires more time and effort to train.
Meghna et al. [1]	2017	KIMIA Path960	Deep features	Deep learning automatically performs feature extraction after data training.	Deep learning requires time-intensive training.
Taha et al. [4]	2018	KIMIA Path960	SVM + features from LBP	LBP provides high discrimination power and computational simplicity.	The structural information captured is limited.

Anish et al. [3]	2021	KIMIA Path960	NN + GLCM + Mean of sorted grey values	Able to perform more complex activities as compared to other algorithms.	Needs a high amount of data and is computationally expensive.
Alexander et al. [5]	2018	Histopathological images of ICIAR 2018 Grand Challenge	CNN + Gradient Boosted Trees	CNN can detect discriminating features automatically without any human supervision.	CNN requires a huge amount of training data for it to be effective.

Table 1: Comparison of existing methods of classification of histopathological images

Few of them have attempted to use the feature fusion of various feature extractors to provide vivid data to the classification algorithms, which might help them increase their performance. Also, few classifications have been done using the ensembles of classification algorithms that combine the best of the multiple algorithms to improve results. This paper tries to achieve performance appraisal over the classification using unsupervised algorithms by applying feature fusion and ensemble classification on the histopathological dataset.

3 Proposed Method

The proposed method of classifying histopathological images uses feature fusion of the features extracted from TSBTC and Niblack thresholding, which are used with the ensemble of ML algorithms as depicted in figure 1.

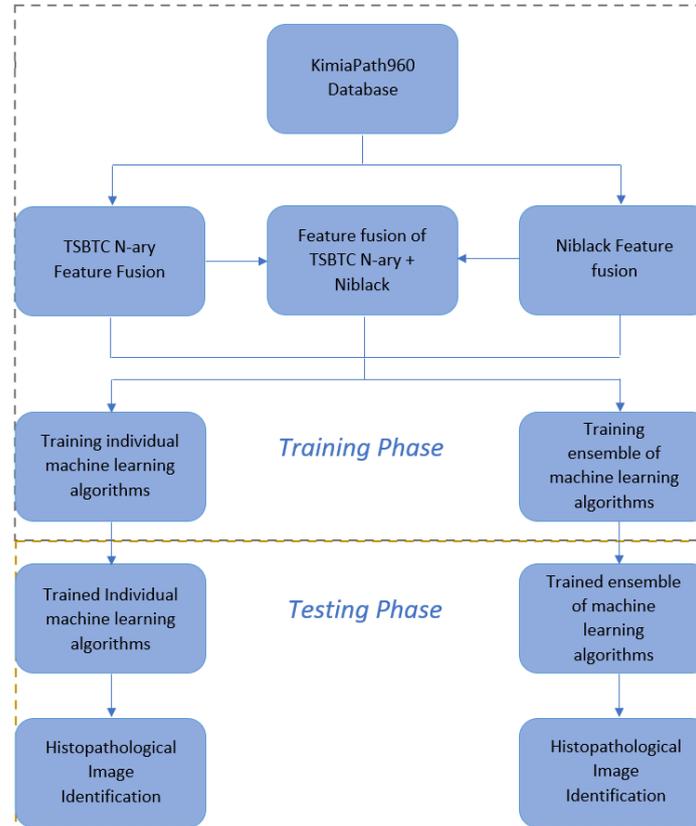


Figure. 1. Representation of the working of the proposed method, divided into two phases: the operation is performed in the training phase, and the result is tested in the testing phase.

3.1 Thepade Sorted Block Truncation Code (TSBTC):

For any image with size $(a \times b)$ sq. pixels, one blue, green, and red value exists for each pixel. Firstly, the RGB planes for the image are separated, where each pixel contains a single intensity value for that respective colour. After this, the values of RGB matrices are stored in an array of size $(a \times b)$ in a sorted manner.

For each sorted array, it is divided into 'n' clusters. Then the values in each cluster are replaced by the Mean of those values. In this method, Thepade sorted block truncation is performed. For example, TSBTC 3-ary is formed, as shown in figure 2.

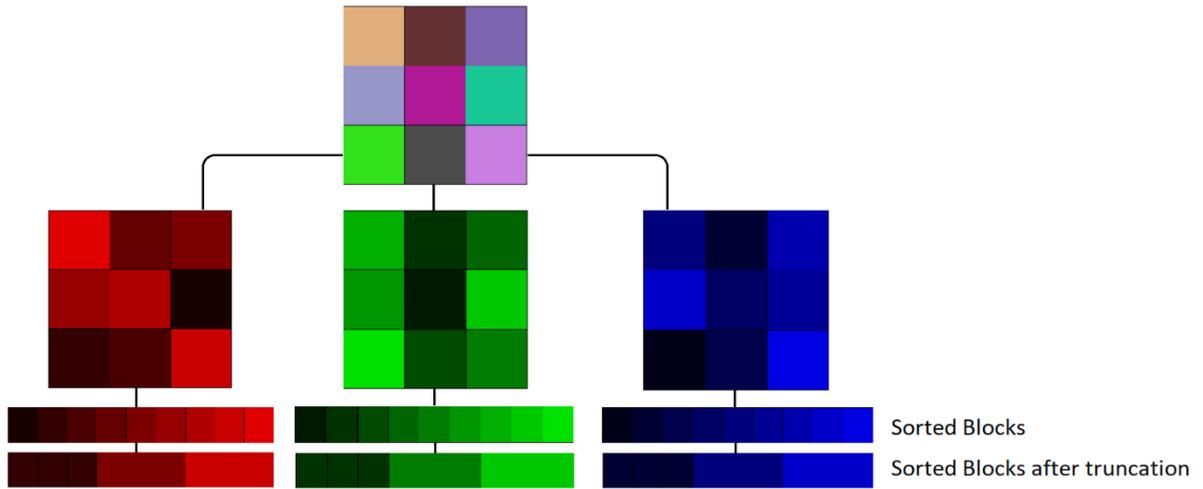


Figure. 2. Pictorial depiction of TSBTC for a sample 3x3 RGB image

For an image of size (a×b) with three colour planes, i.e., Rplane, Gplane and Bplane, the TSBTC N-ary feature vector of TSBTC N-ary becomes [R₁...R_n, G₁...G_n, B₁...B_n]. Here the R_i, G_i and B_i values represent the centroid of respective planes of the ith cluster.

The colour planes of the images are transformed into a sorted array of sizes 'a × b'. Each sorted array is stored as sortedR, sortedG, and sortedB. Using the sorted array stored previously, the TSBTC feature vector is calculated as made known in equations 1,2 and 3.

$$R_i = \frac{n}{a \times b} \sum_{k=1 + \frac{a \times b}{n} \times (i-1)}^{\frac{a \times b}{n} \times i} \text{sortedR}[k] \quad [1]$$

$$G_i = \frac{n}{a \times b} \sum_{k=1 + \frac{a \times b}{n} \times (i-1)}^{\frac{a \times b}{n} \times i} \text{sortedG}[k] \quad [2]$$

$$B_i = \frac{n}{a \times b} \sum_{k=1 + \frac{a \times b}{n} \times (i-1)}^{\frac{a \times b}{n} \times i} \text{sortedB}[k] \quad [3]$$

3.2 Niblack Thresholding Algorithm

Based on a goal-directed evaluation, it is demonstrated that the Niblack thresholding method outperforms the other eleven local thresholding techniques empirically. Numerous image thresholding applications, including display inspection and OCR, benefit from the algorithm's efficiency. In regions of the image with a big, uniform backdrop, the Niblack algorithm is perceptive to the size of the window and gives noisy segmentation results.

By moving a rectangular window over the picture data, Niblack thresholding determines the threshold for each pixel (Let m x n be the size of image IMG). This approach makes use of the threshold determined by the Blue, Red, and Green means calculated locally are represented by the values mb(i, j), mr(i, j), mg(i, j) and b(i, j), r(i, j) and g(i, j) respectively, determined across a specified size of the window (k×k). The Niblack threshold values for the Blue, Green and Red elements are considered NTB, NTG, and NTR, respectively, as shown in the following equations:

$$NTR(i, j) = k \times \sigma r(i, j) + mr(i, j) \quad [4]$$

$$NTG(i, j) = k \times \sigma g(i, j) + mg(i, j) \quad [5]$$

$$NTB(i, j) = k \times \sigma b(i, j) + mb(i, j) \quad [6]$$

Constructed on these threshold values, equations 7, 8 and 9 give the binary maps of the Red, Green, and Blue planes. The following criteria are used to choose the binary bitmaps that were produced via Niblack thresholding:

$$BM_R(i, j) = \begin{cases} 1, & \text{if } IMG(i, j) > NTR(i, j) \\ 0, & \text{if } IMG(i, j) \leq NTR(i, j) \end{cases} \quad [7]$$

$$BM_G(i, j) = \begin{cases} 1, & \text{if } IMG(i, j) > NTG(i, j) \\ 0, & \text{if } IMG(i, j) \leq NTG(i, j) \end{cases} \quad [8]$$

$$BM_B(i, j) = \begin{cases} 1, & \text{if } IMG(i, j) > NTB(i, j) \\ 0, & \text{if } IMG(i, j) \leq NTB(i, j) \end{cases} \quad [9]$$

These binary maps are then taken to create the features of the feature vector (or image's "signature") that is utilized for classification. The equations 10 to 16 are used to compute the feature vector of Niblack thresholding of image IMG are [uNR, INR, uNG, ING, uNB, INB].

$$uNR = \frac{1}{\sum_{i=1}^m \sum_{j=1}^n BM_R(i, j)} \sum_{i=1}^m \sum_{j=1}^n IMG(i, j) * BM_R(i, j) \quad [10]$$

$$INR = \frac{1}{\sum_{i=1}^m \sum_{j=1}^n (1 - BM_R(i, j))} \sum_{i=1}^m \sum_{j=1}^n IMG(i, j) * (1 - BM_R(i, j)) \quad [11]$$

$$uNG = \frac{1}{\sum_{i=1}^m \sum_{j=1}^n BM_G(i, j)} \sum_{i=1}^m \sum_{j=1}^n IMG(i, j) * BM_G(i, j) \quad [13]$$

$$ING = \frac{1}{\sum_{i=1}^m \sum_{j=1}^n (1 - BM_G(i, j))} \sum_{i=1}^m \sum_{j=1}^n IMG(i, j) * (1 - BM_G(i, j)) \quad [14]$$

$$uNB = \frac{1}{\sum_{i=1}^m \sum_{j=1}^n BM_B(i, j)} \sum_{i=1}^m \sum_{j=1}^n IMG(i, j) * BM_B(i, j) \quad [15]$$

$$INB = \frac{1}{\sum_{i=1}^m \sum_{j=1}^n (1 - BM_B(i, j))} \sum_{i=1}^m \sum_{j=1}^n IMG(i, j) * (1 - BM_B(i, j)) \quad [16]$$

To train various classifiers, feature vectors of Niblack thresholding (signatures) are created for every dataset image and saved in Niblack's feature vector database. These classifiers using Niblack feature training are used to categorize the requested photos.

3.3 Fusion of Thepade SBTC and Niblack Thresholding

The TSBTC N-ary provides the dataset's global features, and its local features are provided via Niblack thresholding. Combining the advantages of these two approaches, the proposed method aims to improve the desired outcomes. The accuracy is seen to increase over individual classification. When given a range of features, classification algorithms perform better. By using this reasoning, feature fusion is developed to improve classification accuracy.

3.4 Ensemble of ML algorithms

The ensemble of various classification algorithms is considered with majority voting. This machine learning ensemble model, called a majority vote ensemble, integrates the predictions from various other models [19]. It is a method for enhancing model performance that, in theory, produces results superior to those of any individual model utilized in the ensemble.

A voting ensemble operates by aggregating the results from other models' predictions, and the same can be applied to regression or classification [20]. Computing the average of the model predictions is necessary in the case of regression. While classifying data, predictions of individual labels are added together, and the most voted label is predicted.

3.5 Performance Metrics used

Several other variables are measured that are necessary to give a high value for a test to prove its effectiveness. Some such variables are sensitivity and specificity. Sensitivity and specificity represent how well a test performs in relation to a reliable referent.

		<i>Truth</i>		
		<i>Disease (number)</i>	<i>Non Disease (number)</i>	<i>Total (number)</i>
<i>Test Result</i>	<i>Positive (number)</i>	A <i>(True Positive)</i>	B <i>(False Positive)</i>	$T_{\text{Test Positive}}$
	<i>Negative (number)</i>	C <i>(False Negative)</i>	D <i>(True Negative)</i>	$T_{\text{Test Negative}}$
		T_{Disease}	$T_{\text{Non Disease}}$	Total

Figure. 3. Tabular depiction of the various determining variables such as TP, FP, FN and TN

Sensitivity gives us the probability of a test result indicating 'disease' among those with the disease:

$$\text{Sensitivity} = \frac{TP}{TP+FN} = TP \text{ rate} = \text{Recall} \quad [17]$$

Specificity is the probability that those without the disease indicate a negative test result:

$$Specificity = \frac{TN}{TN+FP} = 1 - FP \text{ rate} \quad [18]$$

Accuracy depicts the correctly identified instances and is represented as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100 \quad [19]$$

Where (as shown in figure 3),

TP (True Positive): Conclusion where histopathological image identification is correct.

FP (False Positive): Conclusion where histopathological image identification is incorrect.

TN (True Negative): Conclusion where the procedure correctly recognizes the histopathological image fitting some other histopathological image category type as an image of the current class.

FN (False Negative): Conclusion where the procedure incorrectly recognizes the image not fitting in its original category.

TP rate: Proportion of correct identifications in the positive predictions class.

Recall: Fraction of the relevant predictions that are successfully made.

FP rate: Proportions of incorrect identifications in the positive predictions class.

4 Results and Discussion

The proposed model for the histopathological image classification experiments on a dataset named 'Kimia-path960' consists of 960 total images, each of which falls under one of the 20 categories present. The images are equally distributed within categories, each having exactly 48 images.

The experimental validation is done on histopathological images from the dataset 'Kimia-Path960' [1], which contains 960 images. The dataset is equally distributed into 20 classes. Thus, it includes 48 images per class. A few images from the Kimia-path960 dataset are shown in figure 4.

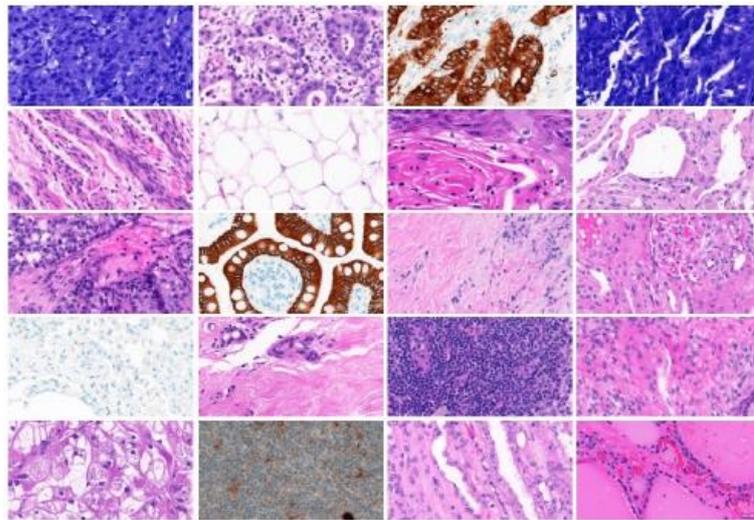


Figure. 4. Illustrative Images from the Kimia Path960 dataset [1], each from one class present in the dataset

The dataset experiments with 12 classification algorithms and nine distinct combinations of ensembles, local feature extraction with Niblack thresholding, global feature extraction with ten variations of TSBTC n-ary followed by the feature fusion of TSBTC N-ary algorithm and Niblack's algorithm. This feature fusion results in combining local and global aspects of the image. The result section shows the results obtained from experimentation and comparisons between results.

TSBTC is performed for each image in the dataset. The performance graph is calculated using 12 classification algorithms (BayesNet, NaiveBayes, SMO, Simple Logistic, MultilayerPerceptron, IBk, KStar, J48, RandomForest, LMT, RandomTree & REPTree). The various ensembles used here are alias 'SimpleLogistic + Multilayer Perceptron + LMT + KStar + RandomForest (SL + MP + LMT + KS + RF)', 'SimpleLogistic + Multilayer Perceptron + LMT + RandomForest (SL + MP + LMT + RF)', 'SimpleLogistic + LMT + Multilayer Perceptron (SL + LMT + MP)', 'SimpleLogistic + LMT + RandomForest (SL + LMT + RF)', 'SimpleLogistic + Multilayer Perceptron + RandomForest (SL + MP + RF)', 'SimpleLogistic + LMT (SL + LMT)', 'SimpleLogistic + Multilayer Perceptron (SL + MP)', 'SimpleLogistic + RandomForest (SL + RF)', 'Multilayer Perceptron + RandomForest (MP + RF)'.

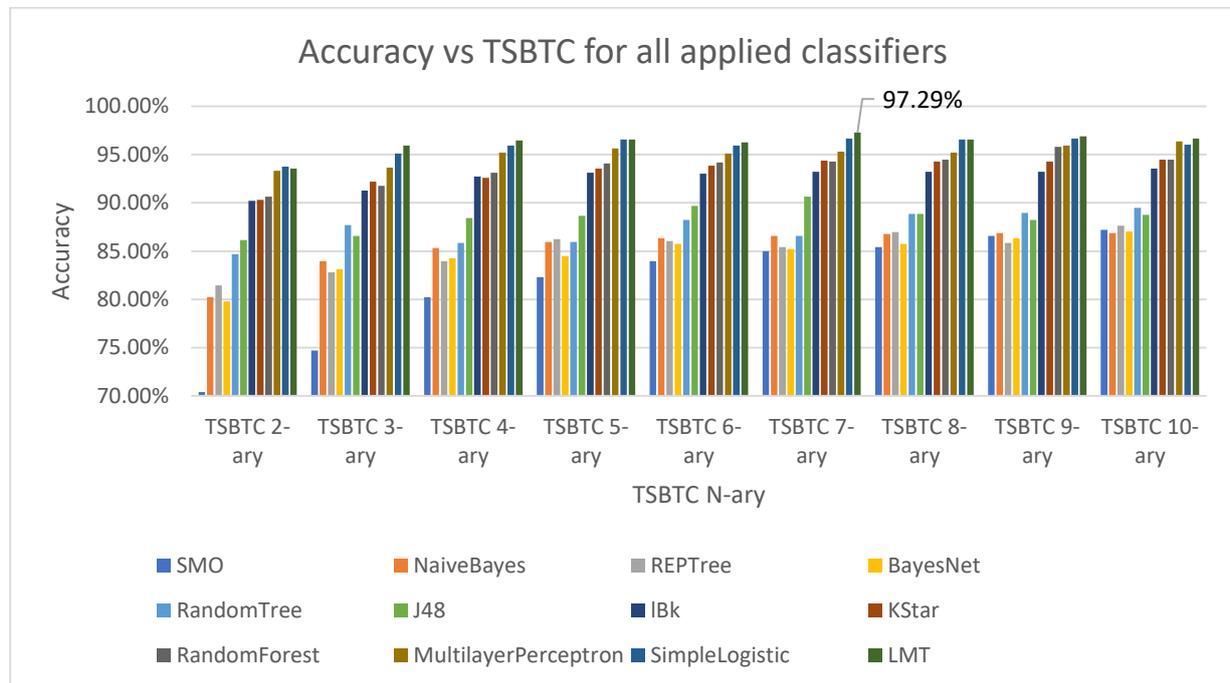


Figure. 5. Chart representing the accuracy for each Thepade SBTC N-ary using various Machine Learning algorithms for classification.

The percentage accuracy shows excellent improvement at the beginning from TSBTC 2-ary to TSBTC 6-ary but shows minimal improvement from TSBTC 6-ary to TSBTC 10-ary, as shown in Figure 5. Overall, the accuracy for all algorithms is best for TSBTC 7-ary. The five best-performing algorithms giving the best results are LMT, Simple Logistic, MultilayerPerceptron, RandomForest and KStar, the accuracies for which are shown in figure 6.

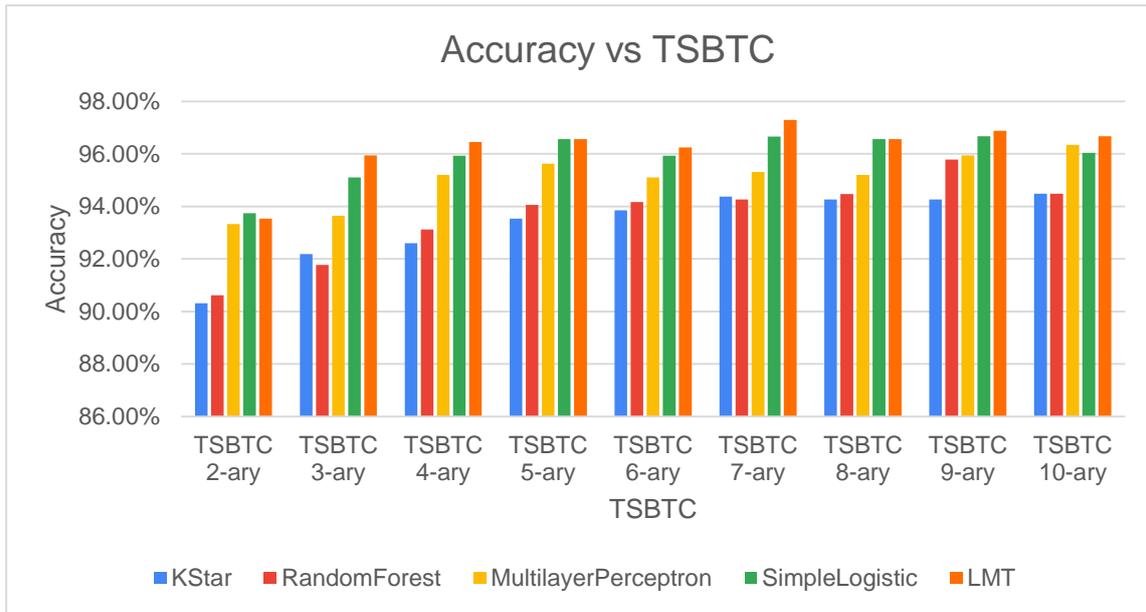


Figure. 6. chart representing the accuracy vs Thepade SBTC for the best five machine learning algorithms used for classification

Niblack thresholding is then performed on each image, and then accuracy for the same is calculated using the above-stated Machine Learning algorithms. For this feature, too, the best-performing algorithms remain the same. Concatenating the feature elements of TSBTC N-ary with Niblack thresholding, the proposed method creates a feature fusion where TSBTC N-ary gives us the global features, and Niblack thresholding adds the local features. Figure 7 shows the best accuracies for Niblack thresholding for this dataset.

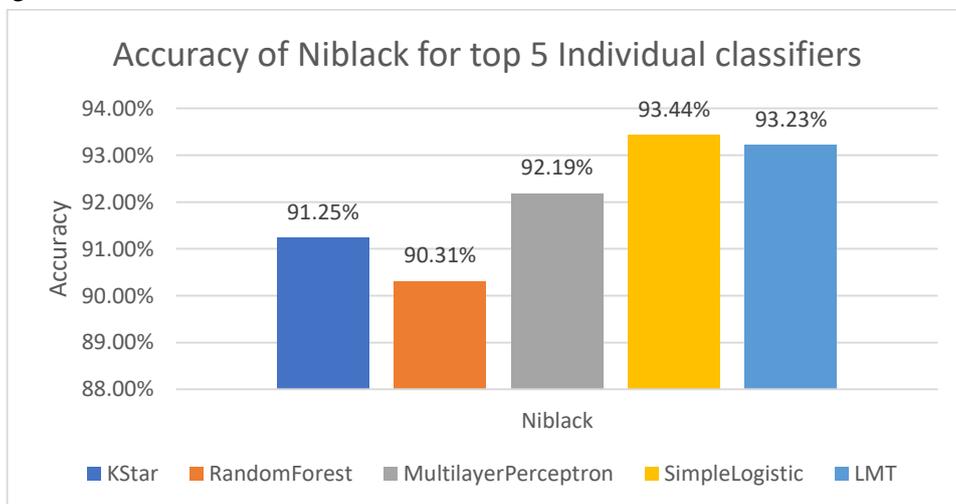


Figure. 7. Chart representing the accuracies for Niblack thresholding of the best five algorithms

Now the accuracies received by only the global and local features and the accuracies obtained by the feature fusions of these global and local features are compared. Figure 8 represents the side-by-side accuracies of the TSBTC N-ary and Feature Fusions. Here it can be noticed that each classifier that performed great for TSBTC N-ary gives better performance for feature fusion.

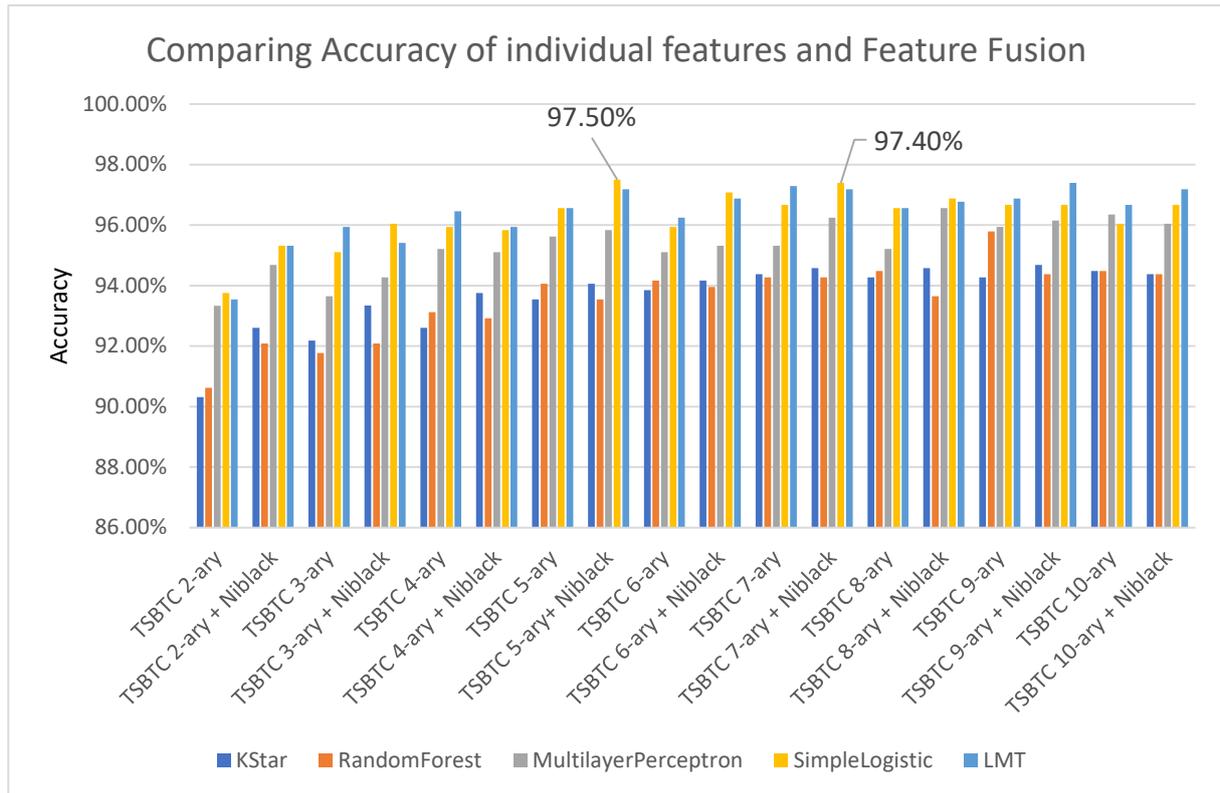


Figure. 8. Chart representing the comparison between TSBTC N-ary and TSBTC N-ary + Niblack of their accuracies for the five best classifying algorithms

The top 5 performing algorithms (LMT, SimpleLogistic, Multilayer Perceptron, KStar, RandomForest) are chosen for the next step, where they are used for the ensemble. Then ensemble classification is performed on the feature fused data. There are a total of 9 ensembles of classification algorithms created where the best six ensembles are ‘SimpleLogistic + Multilayer Perceptron + LMT + KStar + RandomForest’, ‘SimpleLogistic + Multilayer Perceptron + LMT + RandomForest’, ‘SimpleLogistic + LMT + Multilayer Perceptron’, ‘SimpleLogistic + LMT + RandomForest’, ‘SimpleLogistic + Multilayer Perceptron + RandomForest’ and ‘SimpleLogistic + LMT’. The acronyms SL (SimpleLogistic), MP (Multilayer Perceptron), RF (Random Forest) and Ks (KStar) represent the various classifiers present in that ensemble in the chart.

As revealed in figure 9, the ensemble that gives the highest accuracy is ‘SimpleLogistic + Multilayer Perceptron + LMT + RandomForest’ alias ‘SL + MP + LMT + RF’ for TSBTC 7-ary. This performance appraisal is due to feature fusion and the ensemble of the best-performing ML algorithms.

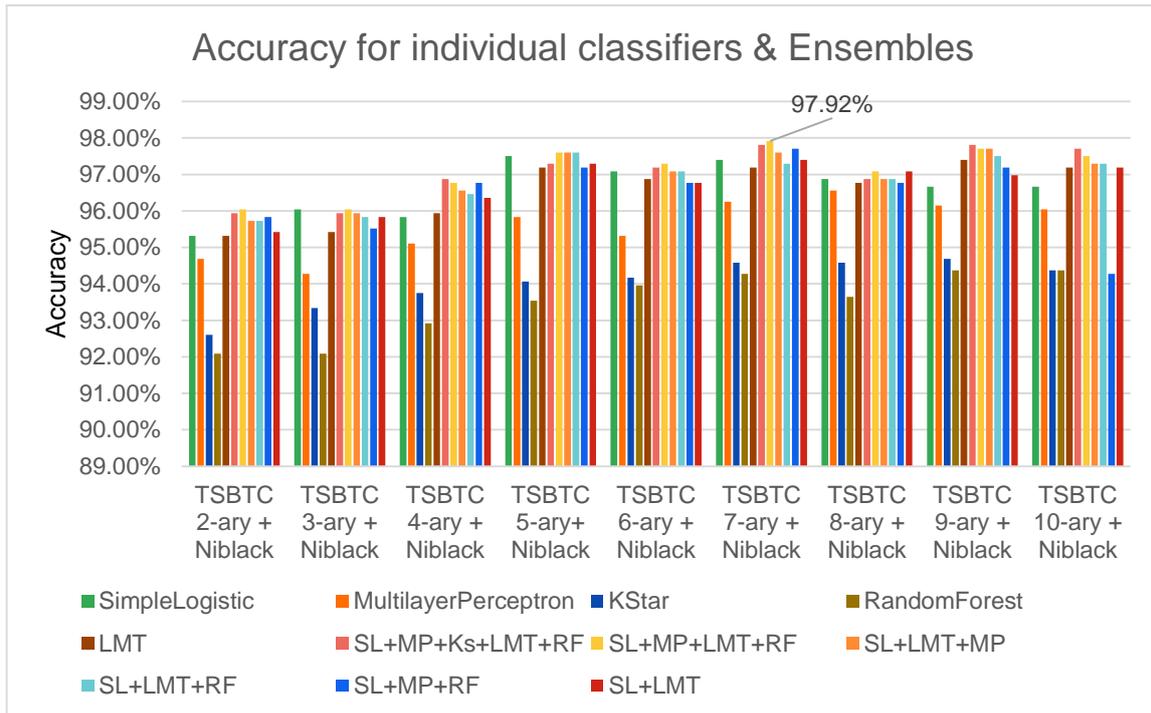


Figure. 9. Chart representing accuracies of the feature fusion of TSBTC N-ary + Niblack for the best individual algorithms and ensembles

Figure 10 depicts the direct comparison between the accuracy classification of the best performing TSBTC N-ary, i.e. TSBTC 7-ary and Niblack's thresholding individually and their feature-fusion for the presented ensembles.

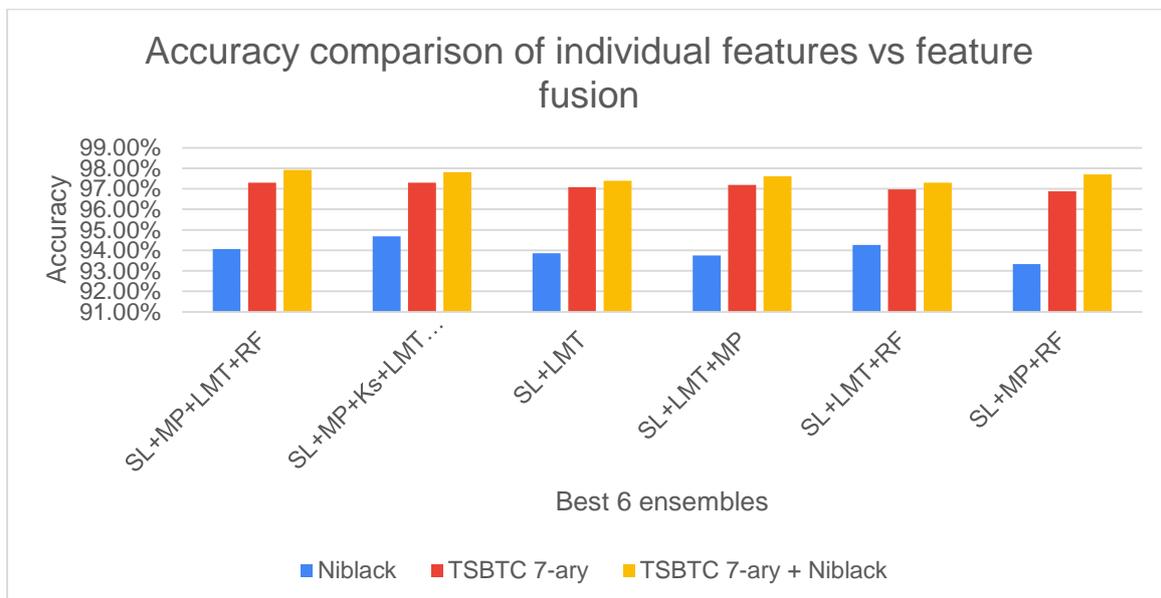


Figure. 10. Chart representation for the improvement in accuracies due to feature fusion compared to individual feature extracts

Other than accuracy, the classification also provides us with other necessary data like FP rate, TP rate, Precision, F-measure, Recall, ROC Area, MCC and PRC Area. This data is used to calculate the sensitivity and specificity for each classification.

For each classification, the sensitivity and specificity values remain consistent with accuracy. Sensitivity is nearly equal to accuracy, and specificity values are almost 100%. The relation of sensitivity, specificity and accuracy for feature fusion using ensemble classification can be seen in figure 11.

The results shown in figure 11 are calculated using majority voting, where individual classifier votes for a resultant class, and the majority class wins. To put it statistically, the identified target class of the ensemble is the mode of the distribution of distinct predicted classes. Also, the results were obtained for 10 folds cross-validation, which means the data is divided into ten different folds, and 90 percent of data is used to train the identification algorithm. The then-trained algorithm is tested on the remaining data portion; thus, the results such as accuracy, sensitivity and specificity are generated from this testing.

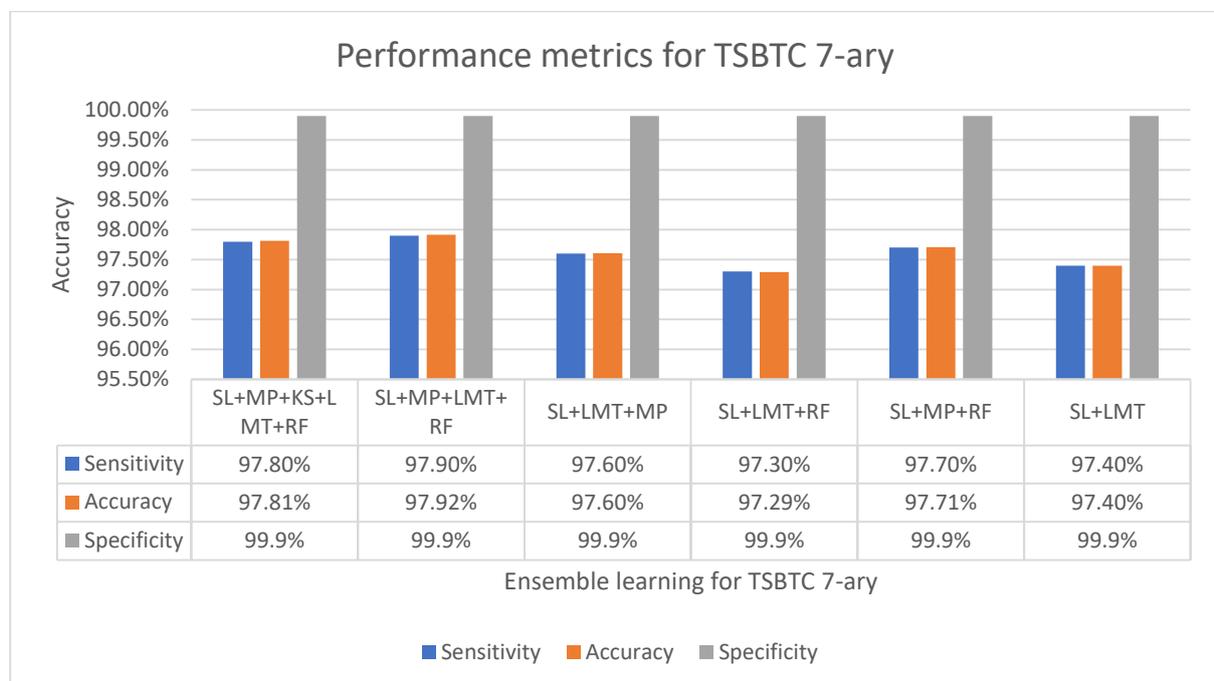


Figure. 11. Performance metrics (Sensitivity, Specificity, Accuracy) of the best-performing ary (TSBTC 7-ary)

For the individual feature extractions of TSBTC N-ary and Niblack thresholding, the maximum accuracies achieved are 97.29% and 94.69%, respectively. The accuracy achieved in the feature fusion of TSBTC 7-ary and Niblack (97.92%) is higher than the accuracies of both TSBTC 7-ary and Niblack individually. This increase in accuracy proves that Feature Fusion helps in this case with improvement in accuracy and also other performance metrics. Figure 12 represents the average accuracy achieved by individual feature extractors and feature fusion for all classification algorithms used.

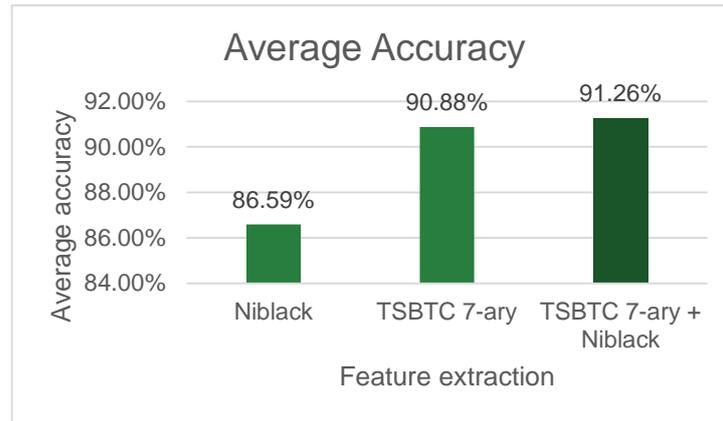


Figure. 12. Average accuracies for individual Niblack, TSBTC 7-ary and feature fusion of Niblack and TSBTC 7-ary

The maximum accuracy of 97.92% is achieved in TSBTC 7-ary + Niblack feature fusion for the ensemble of SimpleLogistic + Multiplayer Perceptron + LMT + RandomForest, as shown in figure 13. The ensemble classifiers give better accuracies than individual classifiers, so the ensemble proves fruitful. Figure 13 shows the accuracies of TSBTC 7-ary + Niblack with every classification algorithm in order of individual classifiers followed by ensembles, each sorted.

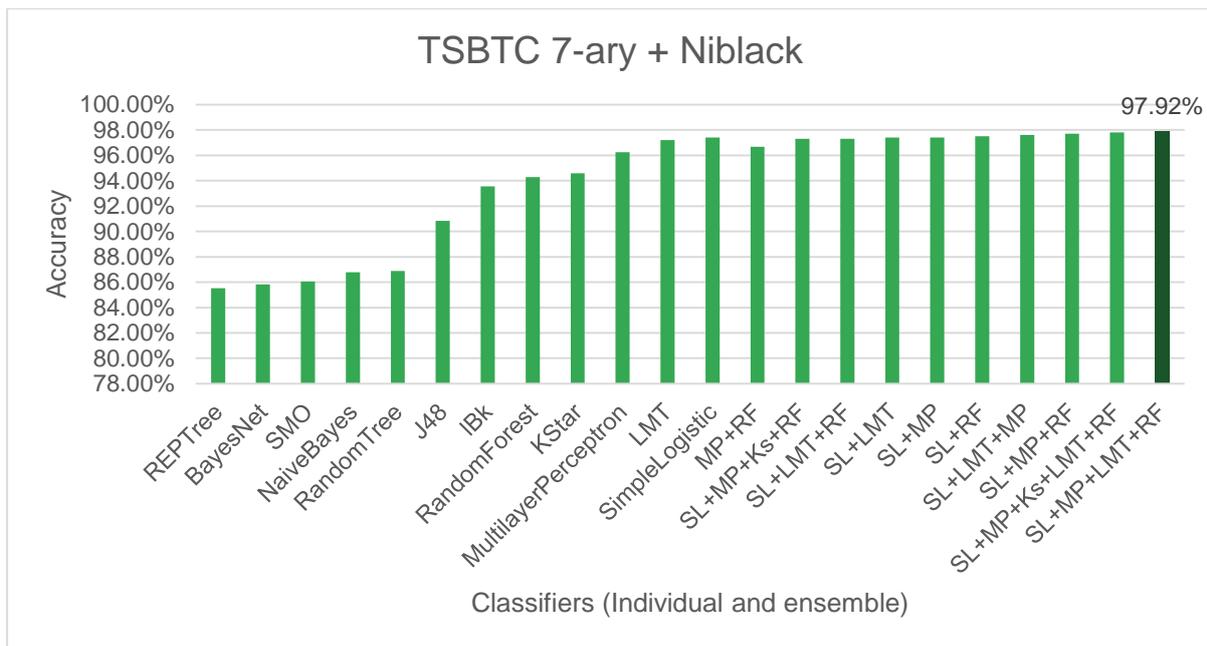


Figure. 13. Performance of TSBTC 7-ary + Niblack ensemble for all the classifiers used

The average accuracy growth is excellent initially but decreases after the peak. Table 2 shows the accuracies of all the ensembles and its average. This table shows that the peak of the average accuracy is achieved at TSBTC 7-ary+Niblack for the best ensembles, and then the accuracy values decrease. Figure 14 shows the graphical representation for the average accuracies of ensembles of TSBTC N-ary + Niblack.

Ensembles of machine learning Algorithms	TSBTC 10-ary + Niblack	TSBTC 9-ary + Niblack	TSBTC 8-ary + Niblack	TSBTC 7-ary + Niblack	TSBTC 6-ary + Niblack	TSBTC 5-ary+ Niblack	TSBTC 4-ary + Niblack	TSBTC 3-ary + Niblack	TSBTC 2-ary + Niblack
SL+MP+KS+LMT+RF	97.71%	97.81%	96.88%	97.81%	97.19%	97.29%	96.88%	95.94%	95.94%
SL+MP+LMT+RF	97.50%	97.71%	97.08%	97.92%	97.29%	97.60%	96.77%	96.04%	96.04%
SL+MP+KS+RF	94.17%	97.08%	96.88%	97.29%	96.77%	96.77%	96.04%	95.31%	95.83%
SL+LMT+MP	97.29%	97.71%	96.88%	97.60%	97.08%	97.60%	96.56%	95.94%	95.73%
SL+LMT+RF	97.29%	97.50%	96.88%	97.29%	97.08%	97.60%	96.46%	95.83%	95.73%
SL+MP+RF	94.27%	97.19%	96.77%	97.71%	96.77%	97.19%	96.77%	95.52%	95.83%
SL+LMT	97.19%	96.98%	97.08%	97.40%	96.77%	97.29%	96.35%	95.83%	95.42%
Average Accuracy	96.49%	97.43%	96.92%	97.57%	96.99%	97.34%	96.55%	95.77%	95.79%

Table 2: Performance of ensembles of ML algorithms in the proposed method of histopathological image classification using feature fusion

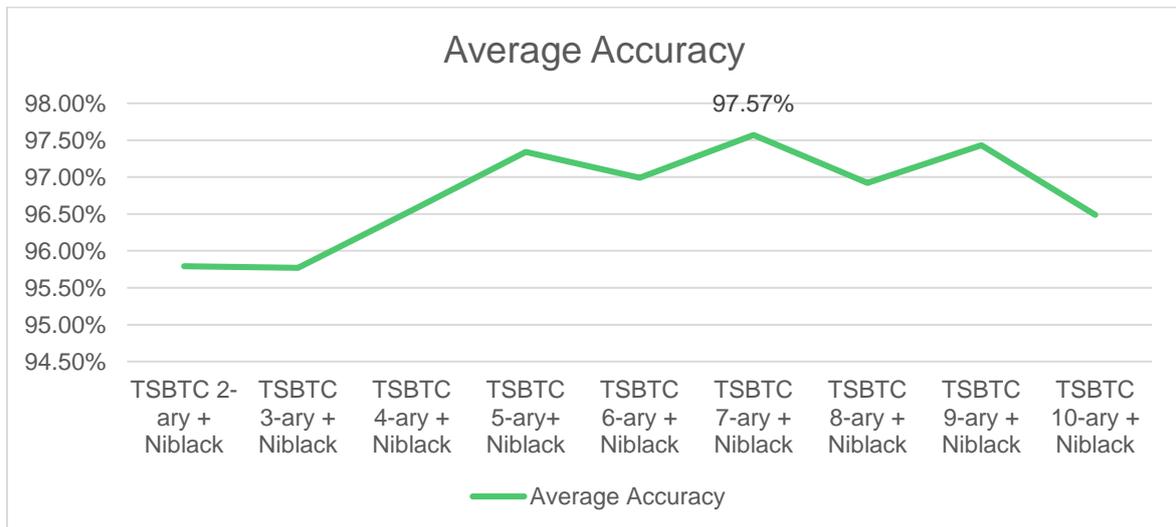


Figure. 14. Variation in average accuracy achieved for all feature fusions

Table 3 presents a performance comparison of the proposed method with a few of the relevant existing methods that have performed classification using the KIMIA Path960 dataset.

Authors	Logic Used	Dataset used	Technique used	Performance Metrics
Ambarish et al. [2], 2020	Deep Learning models and different optimization algorithms	KIMIA Path960	Custom ResNet50 + AdamW	Accuracy 99.90%
			Custom ResNet50 + Adam	Accuracy 99.77%
			Custom ResNet50 + AdaMax	Accuracy 99.79%
			Custom ResNet50 + Radam	Accuracy 99.27%

Meghna et al. [1], 2017	CNN, BoVW and LBP s	KIMIA Path960	Deep features	Accuracy	94.72%
			LBP	Accuracy	90.62%
			BoVW	Accuracy	96.50%
Taha et al. [4], 2018	LBP, HOG and Deep Features	KIMIA Path960	SVM + features from LBP	Accuracy	90.52%
			SVM + deep features	Accuracy	81.14%
			ANN + HOG	Accuracy	34.37%
Anish et al. [3], 2021	Feature Blending	KIMIA Path960	NN + GLCM + Mean of ordered grey values	AUC	0.999
				Precision	0.951
				Recall	0.951
				F1 score	0.951
			RF + GLCM + Mean of ordered grey values	AUC	0.997
				Precision	0.927
				Recall	0.926
				F1 score	0.926
			SVM + GLCM + Mean of ordered grey values	AUC	0.998
				Precision	0.919
				Recall	0.917
				F1 score	0.916
Proposed Method	Fusion of Niblack thresholding and Thepade sorted block truncation code (TSBTC) features.	KIMIA Path960	Ensemble of (SL + MP + RF + LMT) + TSBTC 7-ary features and Niblack features	Accuracy	97.92%
				Sensitivity	0.979
				Specificity	0.999

Table 3: Performance comparison of the Proposed Method with relevant existing Methods of histopathological image classification from the KIMIA Path 960 dataset.

5 Conclusion & Future scope

Feature extraction serves the purpose of extracting information that is beneficial for machine learning tasks. The analysis of histopathological images has traditionally used a variety of local features, together with GLCM and LBP. However, deep learning techniques like convolutional neural networks begin the analysis from feature extraction. This study of histopathological images using Machine learning has shown remarkable results. Diagnosis has become much simpler and less time-consuming due to this development. This classification can still be improved further using better algorithms which give better results.

The proposed paper tries to improve the classification accuracy for this histopathological data by experimenting with different data features and trying feature fusions. The suggested approach has shown improved results, including sensitivity, specificity, and accuracy. The deployment of a group of machine learning algorithms has been found to improve classification accuracy more than the stationing of a single machine learning algorithm. The best results are shown when the accuracy of TSBTC 7-ary is fused with Niblack thresholding for an ensemble of (SimpleLogistic + LMT + Multilayer Perceptron + RandomForest); The highest accuracy achieved is 97.92%; sensitivity achieved being 0.979 and specificity being 0.999.

The upcoming applications of ML aim to improve machine performance with minimal human participation. The prime goals of machine learning integration across various domains are to reduce time and expenses while improving real-time results. Artificial intelligence (AI) is a subcategory that helps software systems increase their prediction accuracy without being specifically created. Machine learning systems use old datasets to foretell outcomes accurately. A few standard machine learning technology implementations are observed in spam filtering, fraud detection, smart healthcare systems, speech recognition, computer vision, and smart transportation. Feature fusions and classification using ensembles will play a big part in optimal classification to get more accurate results.

References

- [1] M. Dinesh Kumar, M. Babaie, S. Zhu, S. Kalra and H. R. Tizhoosh, "A comparative study of CNN, BoVW and LBP for classification of histopathological images," 2017 IEEE Symposium Series on Computational Intelligence (SSCI), 2017, pp. 1-7, <https://doi.org/10.1109/SSCI.2017.8285162>
- [2] A. Ganguly, R. Das and S. K. Setua, "Histopathological Image and Lymphoma Image Classification using customized Deep Learning models and different optimization algorithms," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2020, pp. 1-7 <https://doi.org/10.1109/ICCCNT49239.2020.9225616>
- [3] A. Anurag, R. Das, G. K. Jha, S. D. Thepade, N. DSouza and C. Singh, "Feature Blending Approach for Efficient Categorization of Histopathological Images for Cancer Detection," 2021 IEEE Pune Section International Conference (PuneCon), 2021, pp. 1-6., <https://doi.org/10.1109/PuneCon52575.2021.9686523>
- [4] T. J. Alhindi, S. Kalra, K. H. Ng, A. Afrin and H. R. Tizhoosh, "Comparing LBP, HOG and Deep Features for Classification of Histopathology Images," 2018 International Joint Conference on Neural Networks (IJCNN), 2018, pp. 1-7. <https://doi.org/10.1109/IJCNN.2018.8489329>
- [5] Rakhlin, Alexander & Shvets, Alexey & Iglovikov, Vladimir & Kalinin, Alexandr. (2018). Deep Convolutional Neural Networks for Breast Cancer Histology Image Analysis. 10.1101/259911. https://doi.org/10.1007/978-3-319-93000-8_83
- [6] M. F. Wahid, M. F. Shahriar and M. S. I. Sobuj, "A Classical Approach to Handcrafted Feature Extraction Techniques for Bangla Handwritten Digit Recognition," 2021 International Conference on Electronics, Communications and Information Technology (ICECIT), 2021, pp. 1-4, <https://doi.org/10.1109/ICECIT54077.2021.9641406>
- [7] Thepade, Sudeep & Chaudhari, Piyush. (2020). Land Usage Identification with Fusion of Thepade SBTC and Sauvola Thresholding Features of Aerial Images Using Ensemble of Machine Learning Algorithms. *Applied Artificial Intelligence*. 35. 1-17. <https://doi.org/10.1080/08839514.2020.1842627>

[8] Badre, S. R., and S. D. Thepade. 2016. Novel video content summarisation using thepade's sorted n-ary block truncation coding. *Procedia Computer Science* 79:474-82. <https://doi.org/10.1016/j.procs.2016.03.061>

[9] S. Thepade and S. Badre, "Performance comparison of color spaces in novel video content summarisation using Thepade's sorted n-ary block truncation coding," *International Conference & Workshop on Electronics & Telecommunication Engineering (ICWET 2016)*, 2016, pp. 153-158, <https://doi.org/10.1049/cp.2016.1136>

[10] S. D. Thepade, S. Sange, R. Das and S. Luniya, "Enhanced Image Classification with Feature Level Fusion of Niblack Thresholding and Thepade's Sorted N-ary Block Truncation Coding using Ensemble of Machine Learning Algorithms," *2018 IEEE Punecon*, 2018, pp. 1-7, <https://doi.org/10.1109/PUNECON.2018.8745410>

[11] S. D. Thepade and Y. Bafna, "Improving the Performance of Machine Learning Classifiers for Image Category Identification Using Feature Level Fusion of Otsu Segmentation Augmented with Thepade's N-Ary Sorted Block Truncation Coding," *2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA)*, 2018, pp. 1-6, <https://doi.org/10.1109/ICCUBEA.2018.8697440>

[12] Rukhsar Firdousi, S. Parveen, "Local Thresholding Techniques in Image Binarization", *Int. Journal of Engineering And Computer Science*, Vol. 3 Issue 3, March 2014.

[13] Chowdhury, Shovan & Schoen, Marco. (2020). Research Paper Classification using Supervised Machine Learning Techniques. 1-6. 10.1109/IETC47856.2020.9249211. <https://doi.org/10.1109/IETC47856.2020.9249211>

[14] Q. B. Baker and A. Abu Qutaish, "Evaluation of Histopathological Images Segmentation Techniques for Breast Cancer Detection," *2021 12th International Conference on Information and Communication Systems (ICICS)*, 2021, pp. 134-139, <https://doi.org/10.1109/ICICS52457.2021.9464594>

[15] S. D. Thepade and M. M. Kalbhor, "Image cataloging using Bayes, Function, Lazy, Rule, Tree classifier families with row mean of Fourier transformed image content," *2015 International Conference on Information Processing (ICIP)*, 2015, pp. 680-684, <https://doi.org/10.1109/INFOP.2015.7489469>

[16] S. Thepade, R. Das and S. Ghosh, "Performance comparison of feature vector extraction techniques in RGB color space using block truncation coding for content based image classification with discrete classifiers," *2013 Annual IEEE India Conference (INDICON)*, 2013, pp. 1-6, <https://doi.org/10.1109/INDCON.2013.6726053>

[17] M. Mathur, V. Jindal and G. Wadhwa, "Detecting Malignancy of Ovarian Tumour using Convolutional Neural Network: A Review," 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC), Waknaghat, India, 2020, pp. 351-356, <https://doi.org/10.1109/PDGC50313.2020.9315791>

[18] S. D. Thepade and P. H. Patil, "Novel visual content summarization in videos using keyframe extraction with Thepade's Sorted Ternary Block truncation Coding and Assorted similarity measures," 2015 International Conference on Communication, Information & Computing Technology (ICCICT), Mumbai, India, 2015, pp. 1-5. <https://doi.org/10.1109/ICCICT.2015.7045726>

[19] Sudeep D. Thepade, Shalakra V. Bang, Piyush R. Chaudhari, Mayuresh R. Dindorkar, "Covid19 Identification from Chest X-ray Images using Machine Learning Classifiers with GLCM Features", *Electronic Letters on Computer Vision and Image Analysis*, 19(3): 85-97; 2020. <https://doi.org/10.5565/rev/elcvia.1277>

[20] Sudeep D. Thepade, Gaurav Ramnani and Shubham Mandhare, "Hybrid Wavelet Transform Based Melanoma Identification Using Ensemble of Machine Learning Algorithms", *Electronic Letters on Computer Vision and Image Analysis* 19(3):1-17; 2020 <https://doi.org/10.5565/rev/elcvia.1236>