Available online at https://journal.rescollacomm.com/index.php/ijqrm/index

International Journal of Quantitative Research and Modeling

e-ISSN 2721-477X
p-ISSN 2722-5046

Vol. 6, No. 2, pp. 218-226, 2025

Comparison of Random Forest and SVM Algorithms in Classification of Diabetic Retinopathy Based on Fundus Image Texture Features

Renda Sandi Saputra[1*], Moch Panji Agung Saputra[2]

[1]*University of informatics and business, Bandung, Indonesia*
[2]*Doctoral of Mathematics Study Program, Faculty of Mathematics and Natural Sciences, Universitas Padjadjaran, Sumedang 45363, Indonesia*

Corresponding author email: rendasandi@gmail.com

Abstract

Diabetic Retinopathy (DR) is a microangiopathic complication of diabetes mellitus that can cause visual impairment to permanent blindness. Early detection of DR is essential to prevent disease progression, but conventional methods require time, cost, and expertise that are not always available. This study aims to compare the performance of the Random Forest (RF) and Support Vector Machine (SVM) algorithms in DR classification based on texture features extracted from retinal fundus images. The dataset used consists of 3,000 retinal fundus images obtained from the Kaggle platform, divided into 2,400 training data and 600 test data. Image preprocessing includes conversion to grayscale, resizing to a resolution of 128×128 pixels, and normalization. Feature extraction is performed using a combination of Local Binary Pattern (LBP) and Gray Level Co-occurrence Matrix (GLCM) to produce a 14-dimensional feature vector. Performance evaluation uses accuracy, precision, recall, F1-score, ROC curve, and 5-fold cross-validation metrics. The results showed that Random Forest significantly outperformed SVM with an accuracy of 96% compared to 64%, an AUC value of 0.99 compared to 0.72, and an average cross-validation accuracy of 94.5% compared to 63.42%. Random Forest also showed balanced performance in both classes with precision, recall, and F1-score of 0.96, while SVM experienced classification imbalance especially in the disease class. This study proves that Random Forest is a more optimal algorithm for an automatic DR detection system based on fundus image texture features and can support increasing the accessibility of DR screening in areas with limited specialist medical personnel.

Keywords: Diabetic Retinopathy, Random Forest, Support Vector Machine, texture features, fundus images

1. Introduction

Diabetic retinopathy (DR) is a microangiopathic complication of diabetes mellitus that can cause visual impairment to permanent blindness. According to data from the Centers for Disease Control and Prevention (CDC), in 2021, an estimated 9.6 million people in the United States were living with RD, and 1.84 million of them had vision-threatening RD. Globally, the prevalence of RD among people with diabetes reaches 27%, causing approximately 0.4 million cases of blindness worldwide (Zegeye et al., 2023).

Early detection of RD is essential to prevent disease progression and more serious complications. However, conventional methods such as fundus examination by an ophthalmologist require time, cost, and expertise that are not always available, especially in remote areas. In addition, manual interpretation of fundus images is prone to subjectivity and inter-examiner variability.

Advances in information technology, especially digital image processing and artificial intelligence, open up opportunities to develop an automatic RD detection system (Bauskar, 2020). This system is expected to provide a more efficient, consistent, and accessible solution for the wider community, including in areas with limited specialist medical personnel.

Retinal fundus images contain important information that can be used to detect signs of RD, such as microaneurysms, hemorrhages, and exudates. Fundus image processing involves several stages, including pre-processing to improve image quality, segmentation to highlight important areas, and feature extraction to obtain relevant information (Seeram & Kanade, 2024).

Texture features in fundus images, such as pixel intensity distribution patterns, can provide indications of retinal abnormalities. Texture feature extraction methods, such as Gray Level Co-occurrence Matrix (GLCM), have been used to capture this information and are used in the classification process.

In the context of medical image classification, Random Forest (RF) and Support Vector Machine (SVM) algorithms have shown good performance. RF is an ensemble method that combines multiple decision trees to improve accuracy and reduce overfitting. Meanwhile, SVM works by finding the optimal hyperplane that separates the data classes with the largest margin (Boateng et al., 2020).

A study conducted by Bhattacharjee & Mahmud (2024) has applied the RF and SVM algorithms to classify the severity of diabetic retinopathy (RD) by utilizing features extracted from retinal fundus images. In the study, the RF algorithm showed quite promising classification performance with an accuracy of 76.5%, sensitivity of 77.2%, and specificity of 93.3%. These results indicate that RF has quite significant potential in identifying RD characteristics, especially through a texture analysis-based approach to fundus images.

Although various studies have explored the use of RF and SVM in RD classification systems, there are still gaps that have not been widely studied, especially regarding the direct comparison of the performance of the two algorithms in the context of fundus image texture features. The majority of previous studies have focused more on separate testing between algorithms or using general features that do not specifically describe the textural structure of the retina. In fact, texture information in fundus images has high diagnostic value in detecting microstructural changes due to diabetes complications. In addition, the evaluation of classification performance in most studies is still limited to accuracy metrics, without considering other important indicators such as sensitivity, specificity, and area under the curve (AUC) which provide a more comprehensive picture of model performance. Therefore, research is needed that systematically compares the performance of RF and SVM based on texture features, and comprehensively evaluates the performance of each model to support the development of a more accurate, efficient, and applicable early detection system for RD in a clinical context.

This study aims to compare the performance of the RF and SVM algorithms in RD classification based on texture features extracted from fundus images. The evaluation was carried out using various performance metrics to provide a comprehensive picture of the advantages and disadvantages of each algorithm.

This study is also expected to contribute to providing an in-depth comparative analysis between RF and SVM in the context of RD classification based on texture features of fundus images. The results of this study are expected to be a reference for the development of a more accurate and efficient automatic RD detection system, as well as support efforts to improve eye health services, especially in early detection of RD.

2. Literature Review

2.1. Diabetic Retinopathy and Fundus Images

Diabetic Retinopathy (DR) is one of the microvascular complications of diabetes mellitus that can cause visual impairment to permanent blindness. The main cause of DR is damage to small blood vessels in the retina due to chronic high blood glucose levels. Early detection of DR is very important to prevent disease progression and further complications (Mayya et al., 2021). Fundus examination by a retinal specialist is the conventional method used, but this method requires costs, time, and expertise that are not always available in remote areas. Therefore, the development of an automatic detection system based on fundus images is a promising alternative solution to increase the accessibility and efficiency of DR screening.

2.2. Texture Feature Extraction

2.2.1. Local Binary Pattern (LBP)

Local Binary Pattern (LBP) is a descriptive texture method used to characterize local patterns in images based on a comparison of the intensity of the central pixel to its neighbors. LBP produces a descriptor value in the form of a binary histogram that can be used to represent texture. This method is computationally lightweight and robust to lighting changes, making it suitable for application in medical image analysis.

2.2.2. Gray Level Co-occurrence Matrix (GLCM)

GLCM is a statistical method for extracting texture features from images based on the distribution of pixel intensity value pairs at a certain distance and orientation. Commonly used features of GLCM include contrast, homogeneity, energy, and entropy, which reflect the texture structure of the retina. Shafi et al. (2021) showed that the use of GLCM can improve the accuracy of RD classification. Aris et al. (2022) also developed an automatic RD detection system using a combination of GLCM and the Support Vector Machine (SVM) algorithm, which produced an accuracy of up to 90%.

2.3. Machine Learning for Medical Image Classification

2.3.1. Random Forest

Random Forest (RF) is an ensemble learning algorithm that forms a number of decision trees and combines their prediction results to improve accuracy and reduce overfitting. RF is known for its ability to handle high-dimensional data and is robust to noise. In RD classification, RF has been widely used with competitive results. Bhattacharjee & Mahmud (2024) reported that RF was able to achieve 76.5% accuracy, 77.2% sensitivity, and 93.3% specificity in classifying the severity of RD based on features extracted from fundus images.

2.3.2. Support Vector Machine

Support Vector Machine (SVM) is a machine learning algorithm that works by finding the optimal hyperplane that separates classes in a feature space. SVM is very effective for binary classification and has been widely used in medical image processing. In a study by Aris (2022), an automatic RD detection system using GLCM for feature extraction and SVM as a classifier showed an accuracy of 90%. Another study by Devi et al. (2025) stated that Multi-Class SVM even achieved an accuracy of up to 98.6% in RD level classification, outperforming other algorithms such as RF and K-Nearest Neighbors (KNN).

2.4. Related Studies and Research Gaps

Several previous studies have compared the performance of the RF and SVM algorithms in RD classification. Shafi et al. (2021) reported that RF had an accuracy of 95.19% in texture feature-based classification, while SVM showed slightly lower accuracy. However, a study by Devi et al. (2025) showed that Multi-Class SVM outperformed RF with an accuracy of 98.6% compared to 87.2% for RF. Bhattacharjee & Mahmud (2024) showed that RF was superior in terms of specificity by 93.3%. This difference in results indicates that the selection of the optimal classification algorithm is highly dependent on the prioritized evaluation metrics, such as accuracy, sensitivity, or specificity. However, there are still limitations in studies that directly compare the performance of the two algorithms based on texture features obtained from methods such as LBP and GLCM. Therefore, this study attempts to fill this gap by comparing the performance of RF and SVM based on texture features extracted from fundus images using the LBP and GLCM methods.

3. Methods

This study was designed to conduct a comparative analysis of the performance of the Random Forest (RF) and Support Vector Machine (SVM) algorithms in the classification of Diabetic Retinopathy (DR) based on texture features extracted from retinal fundus images. The methodological approach applied includes comprehensive stages starting from image data collection, pre-processing, texture-based feature extraction, classification model training, to performance evaluation using relevant metrics. Figure 1 presents a research flowchart that illustrates the systematic sequence of all stages.

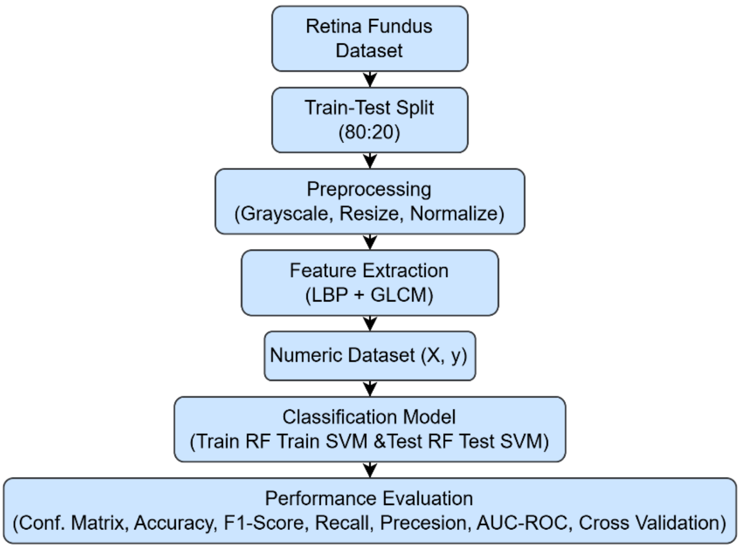

Figure 1: research stages

3.1. Dataset

The dataset used in this study consists of 3,000 retinal fundus images obtained from the Kaggle platform (Rath, 2020), then stored and managed via Google Drive. The dataset structure is organized into two classes (normal and Diabetic Retinopathy (DR)).

This dataset is divided into two subsets, namely:

1. Training data: 2,400 images, consisting of 1,200 normal images and 1,200 DR images.
2. Test data: 600 images, consisting of 300 normal images and 300 DR images.

All images are color images (RGB) with varying initial resolutions. For further processing purposes, the images are then resized to a fixed resolution and normalized during the pre-processing stage.

3.2. Image Preprocessing Stages

The initial stage of image processing is carried out through a preprocessing process that includes reading the image, converting to grayscale format, resizing to a resolution of 128×128 pixels, and normalizing pixel values to the range 0-1. The $cv2.imread()$ function is used to read the image, followed by $cv2.cvtColor()$ for color conversion, and $cv2.resize()$ to resize the image size. Normalization is done by dividing the pixel value by 255. This process aims to equalize the size and intensity of the image before the feature extraction stage is carried out.

3.3. LBP and GLCM Feature Extraction

Feature extraction is carried out by combining two texture descriptive methods, namely Local Binary Pattern (LBP) and Gray Level Co-Occurrence Matrix (GLCM). LBP is applied with parameters $P = 8$ and $R = 1$, and uses the 'uniform' method. The LBP histogram is calculated with 10 main bins and normalized. Next, GLCM is calculated with a pixel distance of 1 and an angle of 0°, then four statistical features are derived, namely contrast, correlation, energy, and homogeneity using the $graycoprops$ function. $Graycoprops$ functions include the following:

$$Contrast = \sum_{i,j} |i-j|^2 P(i,j). \tag{1}$$

$$Correlation = \sum_{i,j} \frac{(i-\mu_i)(j-\mu_j)P(i,j)}{\sigma_i \sigma_j}. \tag{2}$$

$$Energy = \sum_{i,j} P(i,j)^2. \tag{3}$$

$$Homogeneity = \sum_{i,j} \frac{P(i,j)}{1+|i-j|}. \tag{4}$$

The combination of LBP histogram and GLCM features forms the final feature vector that will be used in the classification model training process.

3.4. Dataset Loading and Vectorization

The training and testing datasets are loaded recursively using the $os.listdir()$ and $os.path.join()$ functions in each class directory. Each image that has been successfully processed and its features extracted will be stored in the feature array X, while its labels are stored in the array y. This process produces two data pairs (X_train, y_train) for training and (X_test, y_test) for testing that are ready to be used by the machine learning model.

3.5. Model Training (Random Forest and SVM)

Two machine learning models are used in this study, namely Random Forest and Support Vector Machine (SVM). The Random Forest model is trained using 1000 decision trees $(n\_estimators = 1000)$ with a random state value of 42 to maintain reproducibility. Meanwhile, the SVM model uses a radial basis function (RBF) kernel and the $probability = True$ parameter is activated to allow for ROC curve evaluation. Both are trained using X_train and y_train data, and then predictions are made on X_test.

3.6. Model Performance Evaluation

Model performance was evaluated using a set of standard metrics commonly applied in medical image classification. These metrics were derived from confusion matrix analysis and include Accuracy, Precision, Recall (Sensitivity), F1-Score, and the Area Under the Receiver Operating Characteristic Curve (AUC-ROC).

The formula for each metric is as follows:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%. \quad (5)$$

$$Precision = \frac{TP}{TP+FP} \times 100\%. \quad (6)$$

$$Recall = \frac{TP}{FP+FN} \times 100\%. \quad (7)$$

$$F1 - Score = 2 \times \frac{Precesion \times Recall}{Precesion + Recall} \times 100\%. \quad (8)$$

Where TP is True Positive, TN is True Negative, FP is False Positive, and FN is False Negative.

3.7. Cross-Validation

To test the stability and generalization of the model, a five-fold (5-fold) cross-validation (k-fold cross-validation) was performed. The accuracy of each fold was calculated using the $cross\_val\_score$ function from $sklearn.model\_selection$. The average accuracy of the five folds was calculated for each model and compared visually using a line graph. This graph displays the performance fluctuations between folds while also showing the average accuracy value as an indicator of model stability.

4. Results and Discussion

4.1. Preprocessing and Feature Extraction Results

The initial process carried out in this study is preprocessing of the retinal fundus image, which aims to improve visual quality and simplify the feature extraction stage. The preprocessing result image is shown in Figure 2.

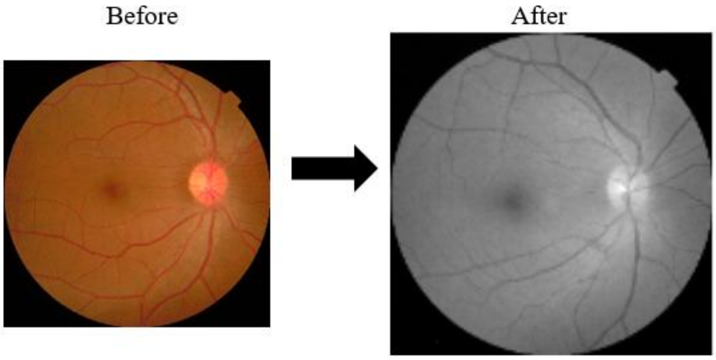

Figure 2: example of results before and after preprocessing

Figure 2 shows an example of the results of a retinal fundus image that has gone through the preprocessing stage. This process aims to improve the visual quality of the image and prepare it to be more optimal for the feature extraction and classification stages. The preprocessing stages applied include converting color images into grayscale format to simplify visual information, resizing the 128x128 image and normalizing pixel intensity values to maintain a uniform data distribution, and resizing the image to a fixed size to ensure consistency of input to the classification model. The results of this preprocessing clarify the anatomical structure of the retina, such as the optic disc and blood vessels, which are important in early detection of eye disease.

After the preprocessing stage, the feature extraction process is carried out to convert the image representation into a numeric vector that can be used as input for the classification algorithm. In this study, the resulting feature vector consists of 14 components that reflect the characteristics of texture, intensity, and spatial patterns in the fundus image. The feature values are: 0.0757, 0.0764, 0.0503, 0.0848, 0.0929, 0.0936, 0.0558, 0.0934, 0.2131, 0.1639, 83.8799, 0.9783, 0.0951, and 0.3902. These feature vectors reflect statistical and morphological information relevant to the detection of retinal abnormalities. All features obtained will be used in the training and evaluation process of the classification model in the next stage.

4.2. Model Results and Classification

The classification model training process was carried out using two popular machine learning algorithms, namely Random Forest (RF) and Support Vector Machine (SVM). The main objective of this process is to classify retinal fundus images into two classes, namely normal and diseased, based on previously extracted features. Both models were trained using the same dataset so that performance comparisons can be carried out fairly and objectively.

In the Random Forest model, the prediction results for the test samples showed high accuracy. A total of 10 test data that were tested randomly were all correctly predicted by the RF model, producing the following prediction result vector: [0, 0, 0, 0, 0, 0, 0, 0, 0, 0]. This indicates that RF is able to recognize patterns in data well, especially in the context of binary classification based on texture and intensity features of retinal images.

In contrast, the Support Vector Machine model showed relatively lower prediction performance on the same test samples. The SVM prediction results show inaccuracy in classifying some data, with the results: [1, 0, 0, 1, 0, 0, 0, 1, 0, 1]. These incorrect predictions reflect that the SVM model faces difficulties in distinguishing between normal and disease classes in some cases, possibly due to the sensitivity of the SVM to the data margin distribution and feature imbalance.

4.3. Classification Model Evaluation Results

4.3.1. Confusion Matrix

Visualization of the confusion matrix of both models is presented in Figure 3.

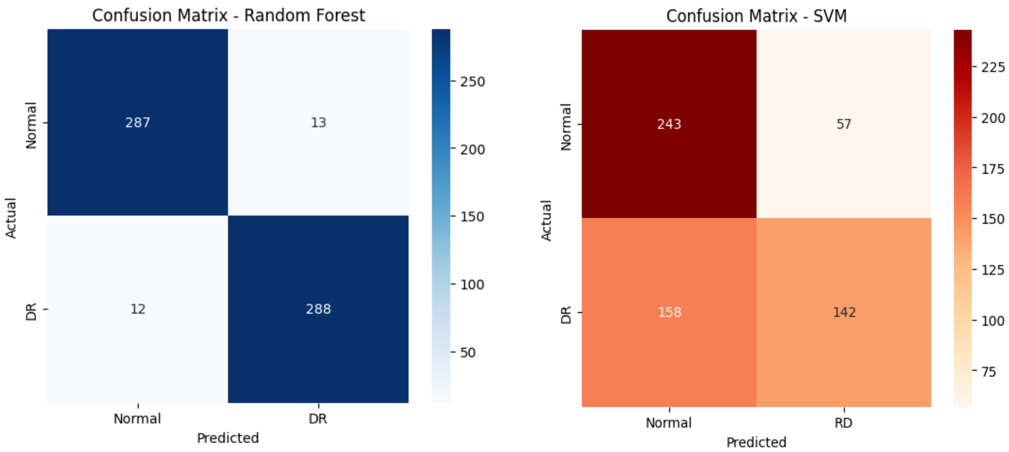

Figure 3: confusion matrix RF and SVM

Figure 3 shows that the RF model produces a nearly perfect classification distribution, with only 13 false positives and 12 false negatives out of a total of 600 data. In contrast, the SVM struggles to classify class 1 (disease), with 158 false negatives, which is very risky in a medical context.

4.3.2. Classification Report

The results of the performance evaluation of the two models based on the classification report are shown in Table 1.

Table 1: Classification report RF and SVM

Model	Class	Precision (%)	Recall (%)	F1-Score (%)	Accuracy (%)
Random Forest	0	96 %	96 %	96 %	
	1	96 %	96 %	96 %	96 %
SVM	0	61 %	81 %	69 %	
	1	71 %	47 %	57 %	64 %

Table 1 shows the results of the performance evaluation of the two classification models, namely Random Forest (RF) and Support Vector Machine (SVM), based on standard classification metrics such as precision, recall, F1-score, and accuracy. In the Random Forest model, both class 0 (normal) and class 1 (disease) showed consistently high precision, recall, and F1-score values, which were 96 % for all of these metrics, with an overall accuracy of 96%. This indicates that the RF model is able to classify data in a balanced and reliable manner for both classes.

In contrast, the SVM model showed much lower and unbalanced performance between the two classes. For class 0, the precision reached 61 % and the recall was 81 %, but the F1-score was only 69 %. Meanwhile, for class 1, the precision and recall were 71 % and 47 %, respectively, with an F1-score of only 57 %. The overall accuracy of the SVM was recorded at 64%, indicating the model's limitations in handling data with complex distributions or characteristics. This significant difference in evaluation metrics reinforces the finding that Random Forest is a more optimal model in the context of texture feature-based retinal fundus image classification.

4.3.3. ROC Curve and AUC Value

Comparison of model performance based on ROC curve is shown in Figure 4.

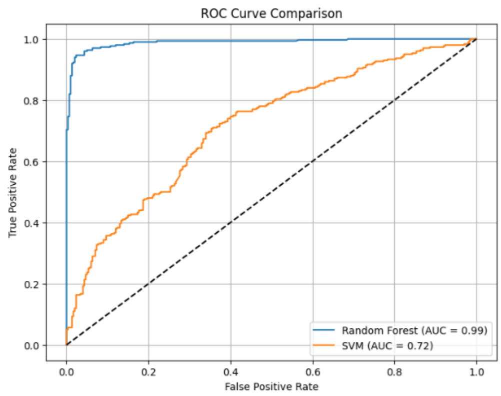

Figure 4: ROC curve and AUC for RF and SVM

Figure 4 shows a comparison of the ROC curves between Random Forest and SVM, where Random Forest has a much better classification performance with an AUC value of 0.99 compared to SVM which only achieved an AUC of 0.72.

4.3.4. Cross Validation

Evaluation of model stability was conducted through 5-fold cross validation. The results are shown in Table 2 and Figure 5.

Table 2: Cross validation results

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average
Random Forest	94.79%	93.33%	95.63%	95.83%	92.91%	94.50%
SVM	68.13%	65.63%	62.50%	59.79%	61.04%	63.42%

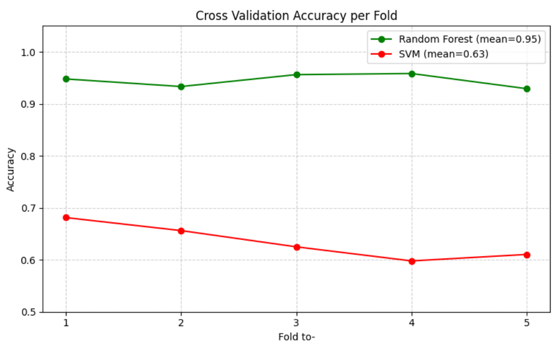

Figure 5: RF vs SVM cross validation graph

Table 2 and Figure 5 show that the Random Forest (RF) model consistently produces high accuracy on each fold, with values ranging from 92.91% to 95.83% and an average accuracy of 94.50%. This indicates that RF has very good stability and is able to maintain consistent performance across various data subsets. The RF curve in Figure 5 also appears relatively flat and is at a high level of accuracy, indicating low variance between folds. In contrast, the performance of the Support Vector Machine (SVM) shows greater fluctuations and a much lower average accuracy, which is only 63.42%. The lowest accuracy was recorded on the fourth fold (59.79%) and the highest on the first fold (68.13%). The SVM graph in Figure 5 shows a significant downward trend in accuracy between folds, reflecting the lack of stability of the model in handling data variations in each subset.

4.4. Classification Model Comparison

Based on the results of a comprehensive evaluation, the Random Forest (RF) model consistently performed better than the Support Vector Machine (SVM) model in the task of classifying retinal fundus images. This comparison covers several important aspects in evaluating machine learning models, such as overall accuracy, generalization ability through cross-validation, and accuracy in classifying minority classes (i.e., classes with disease conditions).

In terms of overall accuracy, Random Forest recorded an accuracy value of 96%, much higher than the accuracy of the SVM model which only reached 64%. This shows that RF is able to correctly classify the majority of test data and is more efficient in distinguishing between normal and disease classes.

The generalization ability of the model, as measured by 5-fold cross-validation, also shows significant superiority of the RF model. The average cross-validation accuracy of RF reached 94.5%, with relatively small variations between each fold, indicating the stability of the model to changes in training data. In contrast, the SVM model only obtained an average cross-validation accuracy of 63.4%, with larger performance fluctuations in each fold, reflecting less consistent performance.

In addition, in terms of classification accuracy on the minority class, which is the class with less data representation and higher complexity (disease class), the Random Forest model shows balanced and high precision and recall. The RF model is able to correctly classify a large number of disease images, with stable f1-scores across both classes. In contrast, the SVM model shows an imbalance in performance between classes, with a tendency to misclassify the disease class. This is potentially risky in a medical context, where misdetection of a disease can be fatal. The comparison results show that Random Forest is superior not only in numerical accuracy, but also in reliability and stability of performance, making it a more feasible model to be used in a retinal fundus image classification system for early detection of pathological conditions such as diabetic retinopathy.

4.5. Performance Analysis and Interpretation

Performance analysis of the two classification models used, namely Random Forest (RF) and Support Vector Machine (SVM), shows fundamental differences in their abilities to process and recognize patterns in medical image data. In general, Random Forest shows significant advantages in terms of both accuracy and generalization ability, while SVM experiences several limitations in the context of the data used in this study.

Random Forest has several key advantages that make it very effective in image classification tasks. First, this model is able to handle non-linear data well, thanks to its structure consisting of many decision trees built based on random subsets of data and features. This ensembling strategy makes RF resistant to overfitting, because the final prediction results are obtained from the aggregation of decisions from various trees, not from a single model. In addition, RF records high performance in various evaluation metrics, such as accuracy, precision, recall, and f1-score, both in the majority and minority classes. Another advantage that is no less important is its ability to provide interpretation of the importance of each feature (feature importance), which is useful in the medical context to find out which variables most influence the classification results.

In contrast, the Support Vector Machine model showed several weaknesses that significantly affected its performance in this experiment. One of the main weaknesses of SVM is its sensitivity to outliers and its inability to handle non-linearly separable data without a complex kernel transformation process. In this study, image data with high intensity variations and morphological structures caused the SVM model to have difficulty in forming an optimal decision boundary. In addition, when the number of features increases or the dataset size becomes large, the performance of SVM tends to decrease drastically due to the increasing computational complexity. The performance of SVM is also proven to be less than optimal in handling unbalanced or complex data distributions, as reflected in the low accuracy and f1-score in the disease class. Considering the experimental results and technical characteristics of both models, Random Forest can be concluded as a more robust, flexible, and reliable model for the task of retinal fundus image classification in this study. This model is not only superior quantitatively, but also easier to interpret and apply to machine learning-based detection systems in the medical field.

4.6. Comparison with Previous Research

The results of this study indicate that the Random Forest (RF) algorithm consistently outperforms the Support Vector Machine (SVM) in the classification of Diabetic Retinopathy (RD) based on fundus image texture features. This finding is in line with a study conducted by Shafi et al. (2021), which reported that RF was able to achieve an accuracy of 95.19% in RD classification using GLCM-based texture features, while SVM showed slightly lower performance. Bhattacharjee and Mahmud (2024) also showed that RF has an advantage in terms of specificity, which is 93.3%, although its accuracy is lower than other studies using different approaches.

However, there are also studies that report the opposite results. A study by Devi et al. (2025) showed that the Multi-Class SVM algorithm can achieve an accuracy of up to 98.6% in the classification of RD severity, surpassing the performance of RF which only reached 87.2% in the same scenario. This study shows that the selection of the optimal algorithm is highly dependent on the context of the problem, the characteristics of the dataset, and the evaluation metrics that are prioritized, such as accuracy, sensitivity, or specificity.

In the context of this study, the feature extraction method used is a combination of Local Binary Pattern (LBP) and Gray Level Co-occurrence Matrix (GLCM), which has been shown to provide a strong texture representation of retinal fundus images. Unlike several previous studies that only used one extraction method, this combination approach is able to capture local and spatial information more comprehensively. This is thought to be one of the factors supporting the high performance of the RF model in this experiment.

In addition, this study also fills a research gap that has not been widely explored, namely the direct comparison of RF and SVM performance on the same dataset with texture features resulting from a combination of LBP and GLCM. Several previous studies, such as by Aris et al. (2022), have indeed shown that the combination of GLCM and SVM can produce accuracy of up to 90%, but have not directly compared it with the RF model in a similar feature extraction context. Thus, this study provides a significant empirical contribution in strengthening the understanding of the effectiveness of machine learning-based classification models in early detection of Diabetic Retinopathy using fundus images.

5. Conclussion

The results showed that the Random Forest algorithm consistently outperformed the Support Vector Machine in all aspects of the evaluation. Random Forest achieved 96% accuracy with balanced precision, recall, and F1-score in both classes (normal and disease) of 96%. In contrast, SVM only achieved 64% accuracy with significant performance imbalance between classes, especially the difficulty in classifying the disease class with a recall of only 47%.

Evaluation using ROC curve showed the superiority of Random Forest with an AUC value of 0.99 compared to SVM which only reached 0.72. The results of 5-fold cross-validation also proved the stability of Random Forest with an average accuracy of 94.5% and small variations between folds, while SVM showed large fluctuations with an average accuracy of only 63.42%.

The superiority of Random Forest can be explained by its ability to handle non-linear data, resistance to overfitting through ensemble methods, and effectiveness in handling complex data distributions. In contrast, SVM has limitations in forming optimal decision boundaries on data with high intensity variations and complex morphological structures.

This study provides a significant empirical contribution in understanding the effectiveness of machine learning-based classification models for early detection of Diabetic Retinopathy. The results of the study indicate that Random Forest is a more robust, flexible, and reliable algorithm choice for an automatic detection system for Diabetic Retinopathy based on fundus images, so that it can support efforts to improve eye health services, especially in early detection in areas with limited specialist medical personnel.

References

Aris, A. A. B., Zulkifli, Z., & Hatim, S. M. (2022, September). Diabetic Retinopathy Detection Using Grey Level Co-Occurrence Matrix. In *2022 3rd International Conference on Artificial Intelligence and Data Sciences (AiDAS)* (pp. 316-321). IEEE.

Bauskar, S. (2020). View of Unveiling the Hidden Patterns AI-Driven Innovations in Image Processing and Acoustic Signal Detection. *Journal of Recent Trends in Computer Science and Engineering*, *8*(1), 10-70589.

Bhattacharjee, I., & Mahmud, T. (2024). Diabetic Retinopathy Classification from Retinal Images using Machine Learning Approaches. *arXiv preprint arXiv:2412*.02265.

Boateng, E. Y., Otoo, J., & Abaye, D. A. (2020). Basic tenets of classification algorithms K-nearest-neighbor, support vector machine, random forest and neural network: A review. *Journal of Data Analysis and Information Processing*, *8*(4), 341-357.

Centers for Disease Control and Prevention. (2024, May 15*). VEHSS modeled estimates: Prevalence of diabetic retinopathy (DR)*. Vision and Eye Health Surveillance System. Accessed from https://www.cdc.gov/visionhealth/vehss/estimates/diabetic-retinopathy.html

Devi, T. M., Karthikeyan, P., Muthu Kumar, B., & Manikandakumar, M. (2025). Diabetic retinopathy detection via deep learning based dual features integrated classification model. *Technology and Health Care*, *33*(2), 1066-1080.

Mayya, V., Kamath, S., & Kulkarni, U. (2021). Automated microaneurysms detection for early diagnosis of diabetic retinopathy: A Comprehensive review. *Computer Methods and Programs in Biomedicine Update*, *1*, 100013.

Rath, S. R. (2020). *Diabetic Retinopathy 224x224 Gaussian Filtered [Dataset]*. Kaggle. Accessed from https://www.kaggle.com/datasets/sovitrath/diabetic-retinopathy-224x224-gaussian-filtered

Seeram, E., & Kanade, V. (2024). Image Processing and Analysis. In *Artificial Intelligence in Medical Imaging Technology: An Introduction* (pp. 83-103). Cham: Springer Nature Switzerland.

Shafi, A. S. M., Rahat Khan, M., & Rahman, M. M. (2021). Statistical texture features based automatic detection and classification of diabetic retinopathy. In *Proceedings of International Joint Conference on Advances in Computational Intelligence: IJCACI 2020* (pp. 27-40). Springer Singapore.

Zegeye, A. F., Temachu, Y. Z., & Mekonnen, C. K. (2023). Prevalence and factors associated with Diabetic retinopathy among type 2 diabetic patients at Northwest Amhara Comprehensive Specialized Hospitals, Northwest Ethiopia 2021. *BMC ophthalmology*, *23*(1), 9.