A hybrid Model for The Detection of Retinal Disorders Using Artificial Intelligence Techniques

Ahmed M. Salaheldin
Cairo University
Manal Abdel Wahed
manal.wahed@eng1.cu.edu.eg
Cairo University
Neven Saleh
Higher Institute of Engineering, EL Shorouk Academy

Article

Keywords:

Posted Date: August 21st, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3262058/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Biomedical Physics & Engineering Express on July 2nd, 2024. See the published version at https://doi.org/10.1088/2057-1976/ad5db2.

Abstract

The prevalence of vision impairment is rising at an alarming rate. The goal of the study is to create an automated method that uses Optical Coherence Tomography (OCT) to classify retinal disorders into four categories, namely, Choroidal Neovascularization, Diabetic Macular Edema, Drusen, and normal cases. The study proposed a new framework that combines machine learning and deep learning-based techniques. The utilized classifiers were Support Vector Machine (SVM), K-Nearest Neighbor (K-NN), Decision Tree (DT), and Ensemble Model (EM). A feature extractor was also employed, which was the InceptionV3 convolutional neural network. The performance of the models has been measured over nine criteria using a dataset of 18000 OCT images. For the SVM, K-NN, DT, and EM, the analysis exhibited state-of-the-art performance with classification accuracies of 99.43%, 99.54%, 97.98%, and 99.31%, respectively. A promising methodology has been introduced for the automatic identification and classification of retinal disorders leading to reducing human error and saving time alike.

Introduction

The incidence of vision problems is rapidly increasing on a global scale. In 2020, data from the World Health Organization revealed that approximately 2.2 billion individuals worldwide are affected by various retinal abnormalities, spanning from mild to severe blindness. Within this statistic, an alarming one billion cases, roughly half of the total, remain either unaddressed or could have been prevented. The demographic most impacted by vision impairment and blindness consists predominantly of individuals over the age of 50. However, vision loss can affect people of all ages. The substantial economic burden of global productivity losses attributed to vision impairment is a pressing concern [1]. Given the profound impact of vision impairment on daily living, extensive research into the early detection of various ophthalmic conditions is imperative. For instance, retinal disorders such as choroidal neovascularization (CNV), diabetic macular edema (DME), and Drusen underscore the urgency for enhanced investigative approaches [1,2].

Optical Coherence Tomography (OCT) imaging modality has emerged as a valuable tool in numerous studies aimed at categorizing retinal diseases. An array of artificial intelligence-driven methodologies, encompassing traditional machine learning techniques and deep learning algorithms, have been harnessed for this purpose. A central concern of medical experts in the classification process revolves around the thickness of the retina and each of its distinct layers, a pivotal indicator of underlying disease.

Numerous studies have proposed diverse approaches to address this underlying challenge through the implementation of various artificial intelligence techniques [3]. Hussain et al. [2], for instance, introduced a method classifying OCT images into three categories: normal, age-related macular degradation (AMD), and diabetic macular edema (DME). Employing the Random Forest (RF) algorithm and validated through a 15-fold cross-validation method, their model achieved an accuracy of 95% and an impressive area under the curve (AUC) of 0.99. Alsaih et al. [4] presented a study focusing on the classification of spectral-domain OCT images into DME and normal classes. Employing three distinct classifiers—linear support vector

machine (SVM), RF, and kernel SVM—alongside principal components analysis (PCA) for dimensional reduction, their most effective model, a linear SVM with PCA, demonstrated a sensitivity and specificity of 87.5%.

In yet another contribution, Abdulrahman and Khatib [5] designed an algorithm to classify OCT images into four categories: CNV, DME, Drusen, and normal. Utilizing a genetic algorithm for feature extraction and a SVM-based classifier, their model exhibited an accuracy of 90.65% through local and global feature extraction. The pursuit of accurate classification techniques continued with Srinivasan et al. [6] who proposed a method for categorizing AMD, DME, and normal classes. Employing a one-vs-one SVM classifier and cross-validation, their model achieved the identification of 100% of AMD cases, 100% of DME cases, and 86.67% of normal cases.

Liu et al. [7] introduced a novel computerized approach for diagnosing macular pathologies. Their model categorized spectral domain OCT images into three groups: normal macula, macular diseases, and glaucoma. Leveraging a nonlinear SVM, their model demonstrated robust performance, evidenced by high cross-validation AUC values on both development and testing datasets. In a similar vein, Dash and Sigappi [8] developed a model for DME classification utilizing OCT images. This study utilized two techniques: Local Binary Pattern (LBP) with RF and Scale-Invariant Feature Transform (SIFT) with RF, achieving accuracy rates of 100% and 88%, respectively.

In the work [9], the authors devised a hybrid strategy to categorize retinal disorders. The classification into CNV, DME, drusen, and normal was carried out by amalgamating the SqueezeNet CNN for feature extraction and employing diverse classifiers. This approach achieved 97.47% classification accuracy with the K-NN classifier. Khan et al. [10] introduced a deep learning model tailored for the automatic identification of distinct retinal disorders, encompassing age-related macular degeneration, branch retinal vein occlusion, central retinal vein occlusion, central serious chorioretinopathy, and diabetic macular edema. Their model entails three key stages: initial feature extraction via training pretrained models—DenseNet-201, InceptionV3, and ResNet-50; subsequent enhancement of features through ant colony optimization; and eventual classification training using the K-NN classifier. Notably, the proposed model achieved a remarkable accuracy of 99.1%.

The question that arises here is whether combining machine learning and deep learning improves the detection and classification of retinal disorders. Additionally, whether introducing a deep learning algorithm enhances the results of detection and/or classification. According to the related works mentioned, Inception V3 is rarely considered. Therefore, the study was proposed to seek the answers to the two questions, considering Inception V3 for the second question. Therefore, the aim is to propose a hybrid model for the classification of retinal disorders, specifically CNV, DME, Drusen, and normal categories. This approach introduces an innovative combination of machine learning and deep learning methods. Four potential classifiers are considered for addressing the highlighted challenge. The study proposes an automated diagnostic tool that holds promise in facilitating early recognition and diagnosis of retinal abnormalities by ophthalmologists. In doing so, this solution has the potential to significantly

mitigate human error, saving both time and effort. The study's contributions encompass various aspects, including (i) the design of an automated detection system using both machine learning and deep learning algorithms, (ii) the introduction of deep learning algorithm (InceptionV3) to address the given problem, (iii) the introduction of new criteria for performance evaluation, (iv) the successful implementation of four machine learning methods, ultimately identifying the optimal one, (v) the study's outcomes demonstrate superior performance in comparison to existing efforts in terms of accuracy and sensitivity, and finally (vi) the study proposes a platform that guides ophthalmologists, ultimately reducing human error and optimizing time management.

The following is an outline of the article's structure. Section 2 discusses the materials and methods used. Section 3 depicts the study's findings, defining the classification of retinal abnormalities based on the technique used and a set of criteria. Section 4 highlights and discusses the research outcomes, as well as creating a benchmark based on the pertinent studies. Section 5 ends the research and discusses future attempts.

Materials and Methods

The presented methodology aims to categorize retinal disorders across four distinct classes: CNV, DME, Drusen, and the normal state. This endeavor employs both a machine learning-based and a deep learning-based approach. To ascertain the optimal approach, an evaluation of four machine learning-based methods is undertaken, including SVM, K-NN, DT, and EM. The selection of optimal hyperparameters is facilitated through Bayesian optimization, thus streamlining the optimization process. For feature extraction rooted in deep learning, Inception V3 is harnessed. The schematic representation of the proposed methodology is illustrated in Fig. 1, depicting the sequential stages of significance detailed in the subsequent subsections.

Data Preparation

The dataset employed in this study originates from Kermany et al. [11], spanning the timeframe from July 1st, 2013, to March 1st, 2017. The authors implement the preprocessing methodology advocated by the study [12]. This procedure entails a sequence of stages, commencing with the enhancement of image contrast, followed by the application of anisotropic diffusion filtration [13], prioritizing high contrast edges over their low-contrast counterparts. This technique offers the advantage of noise reduction without significant loss of image content. Notably, edges, lines, and other distinctive features are crucial for comprehending OCT images in their entirety, ultimately facilitating the creation of multi-dimensional images tailored to match the input layer dimensions of the CNN. An illustrative instance of a processed image is illustrated in Fig. 2.

Features Extraction

A method for automated feature extraction is employed, leveraging the transfer learning capabilities of the InceptionV3 CNN, as recommended by the authors of the work [12]. Specifically, the network operates as a feature extractor, concentrating on the global average pool layer. With each image, this network yields a substantial count of 2048 features, tailored to the context. However, due to the high volume of features generated, a thoughtful feature optimization strategy becomes imperative to streamline the subsequent classification process effectively.

Training Process

The dataset encompasses a total of 18,000 OCT images, meticulously categorized into four distinct classes. For the classifier training process, many trials of data portioning were conducted. For the best combination, the data is distributed as follows: 68% (12240 images) for training, 17% (3060 images) for validation, and 15% (2700 images) for testing. To discern the most effective classifier, four distinct models are introduced—SVM, K-NN, DT, and EM—undergoing rigorous evaluation to unveil superior performance. In pursuit of optimal model configurations, the Bayesian optimization technique is skillfully employed, intricately fine-tuning hyperparameters for each classifier. The optimization process consists of 30 iterative cycles, each supported by classifier-specific hyperparameters shown in Table 1 and guided by the basic ideas of the Bayes Principle [14]. Additionally, the training duration for each classifier is meticulously computed, bearing relevance to the overall evaluation procedure.

Table 1
Hyperparameters of each classifier for classifying retinal disorders.

SVM		**K-NN**	
Kernel function	Linear	Number of neighbors	12
Box constraint level	0.0010132	Distance metric	Spearman
Multiclass method	One-vs-One	Distance weight	Inverse
Training time	9328.9 sec	Training time	4067.9 sec
Prediction speed	8100 obs/sec	Prediction speed	90 obs/sec
DT		**EM**	
Maximum number of splits	61	Ensemble method	AdaBoost
Split criterion	Maximum deviance reduction	Maximum number of splits	16
Training time	349.5 sec	Number of learners	39
Prediction speed	13000 obs/sec	Learning rate	0.90426
		Training time	6787.5 sec
		Prediction speed	9700 obs/sec

Results

This work presents a system that efficiently classifies retinal diseases, including CNV, DME, Drusen, and normal instances, by combining the strengths of four different classifiers. Within a dataset of 18,000 OCT images, an advanced hybrid artificial intelligence framework is harnessed. The development of these algorithms is facilitated through the utilization of MATLAB R2020b's deep learning and machine learning toolboxes. Each classifier's efficacy is measured through a comprehensive evaluation, manifested in the form of a confusion matrix, clearly depicted in Fig. 3. Moreover, Table 2 outlines the nuanced class-specific performance metrics for each classifier. It encompasses true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Moreover, accuracy, sensitivity, specificity, and precision are rigorously computed via the designated formulas (1)–(4) [15,16].

Accuracy = (TP + TN) / (TP + FN + FP + TN) (1)

Sensitivity = TP / (TP + FN) (2)

Specificity = TN / (TN + FP) (3)

Precision = TP / (TP + FP) (4)

Table 2
Class-by-class system performance evaluation matrix per classifier

Criteria Classifiers	Class	TP	TN	FP	FN	Accuracy	Sensitivity	Specificity	Precision
SVM	CNV	671	2020	5	4	**99.67**	**0.9941**	**0.9975**	**0.9926**
	DME	668	2022	3	7	99.63	0.9896	0.9896	0.9896
	Drusen	664	2014	11	11	99.19	0.9837	0.9946	0.9837
	Normal	666	2013	12	9	99.22	0.9867	0.9941	0.9823
K-NN	CNV	673	2021	4	2	**99.78**	**0.9970**	0.9980	**0.9941**
	DME	667	2022	3	8	99.59	0.9881	**0.9985**	0.9881
	Drusen	668	2017	8	7	99.44	0.9896	0.996	0.9882
	Normal	668	2016	9	7	99.41	0.9896	0.9956	0.9867
DT	CNV	663	1999	26	12	**98.59**	**0.9822**	0.9872	**0.9623**
	DME	639	2012	13	36	98.19	0.9467	**0.9936**	0.9467
	Drusen	646	1993	32	29	97.74	0.957	0.9842	0.9528
	Normal	643	1987	38	32	97.41	0.9526	0.9812	0.9442
EM	CNV	669	2018	7	6	**99.52**	**0.9911**	0.9965	**0.9896**
	DME	665	2020	5	10	99.44	0.9852	**0.9975**	0.9852
	Drusen	665	2013	12	10	99.19	0.9852	0.9941	0.9823
	Normal	664	2012	13	11	99.11	0.9837	0.9936	0.9808

The system demonstrated remarkable classification prowess, achieving an overarching accuracy of 99.43% for SVM, 99.56% for K-NN, 97.58% for DT, and 99.31% for EM classifiers. SVM demonstrates consistent performance for all classes, with the CNV class having an exceptional accuracy of 99.67%, along with noteworthy specificity and precision values of 99.75% and 99.26%, respectively. K-NN shines with exceptional accuracy, prominently seen in the CNV class, with an impressive sensitivity of 99.78%, specificity of 99.80%, and precision of 99.41%. DT maintains consistent performance, experiencing slight dips in accuracy and precision in the Drusen and normal classes.

EM demonstrates consistent accuracy levels, with the CNV class showcasing noteworthy specificity (99.65%) and precision (98.96%). Collectively, the classifiers demonstrate commendable performance, with the CNV class consistently attaining the highest accuracy, underscoring its pivotal role in the classification process. The varied performance metrics across classifiers underscore the need for a comprehensive evaluation approach, encompassing diverse criteria to holistically judge their efficacy. The total performance evaluation of each classifier is determined by adding the other five criteria, which are

error rate, false-positive rate, false-negative rate, negative predictive value, and F1-score, using Equations (5)–(9) [15,16].

Error Rate = (FP + FN) / (TP + FN + FP + TN) (5)

False Positive Rate = FP / (FP + TN) (6)

False Negative Rate = FN / (FN + TP) (7)

Negative Predictive Value = TN / (TN + FN) (8)

F1-Score = (2 × (Sensitivity × Precision)) / (Sensitivity + Precision) (9)

The total performance of the model for each classifier is determined considering the nine criteria as shown in Table 3. The classification accuracy, sensitivity, specificity, and precision are calculated for all classes for each classifier.

Table 3
Overall performance evaluation matrix per classifier

Classifier Criteria	SVM	K-NN	DT	EM
Classification accuracy	0.9943	**0.9956**	0.9798	0.9931
Sensitivity	0.9885	**0.9911**	0.9596	0.9863
Specificity	0.9962	**0.9970**	0.9865	0.9954
Precision	0.9871	**0.9893**	0.9515	0.9845
Error Rate	0.0057	**0.0044**	0.0202	0.0069
False Positive Rate	0.0038	**0.0030**	0.0135	0.0046
False Negative Rate	0.0115	**0.0089**	0.0404	0.0137
Negative Predictive Value	0.9962	**0.9970**	0.9866	0.9954
F1-Score	0.9878	**0.9902**	0.9555	0.9854

Discussion

The research introduces an innovative model for diagnosing and categorizing prevalent ophthalmological disorders: CNV, DME, Drusen, and normal cases. Employing SVM, K-NN, DT, and EM as classifiers, each model's performance undergoes comprehensive evaluation across nine criteria. The criteria are accuracy, sensitivity, specificity, precision, error rate, false positive rate, false negative rate, and F1-score. Hyperparameter tuning employs the Bayesian optimization algorithm, with validation encompassing 3060 OCT images and testing involving 2700 OCT images, categorized into four classes.

Delving into class-specific performance, the CNV class emerges as a standout, boasting the highest accuracy across all classifiers. Remarkably, this class consistently commands the best sensitivity across classifiers, while the DME class tends to yield the highest specificity values, except for CNV. The CNV class also attains the highest precision, solidifying its superiority among other classes. This dominance is attributed to CNV's distinct hyper-reflective appearance in OCT images, setting it apart from its counterparts.

Analyzing overall performance, the K-NN classifier prevails across all nine criteria, as indicated in Table 3. Moreover, Table 4 conducts a comparative analysis with related studies, encompassing factors like classifier selection, sample size, accuracy, sensitivity, and specificity. Impressively, the proposed models exhibit elevated accuracy levels compared to cited works [2,4,6–8], with SVM notably surpassing its counterparts. Furthermore, when contrasted with Srinivasan et al. [6], our SVM model outshines by a 4% accuracy improvement. This is due to the usage of histogram of oriented gradients as it uses a sliding window technique to extract features from every pixel of an image. Hence, the accuracy is not highly reliable compared to the CNNs.

Distinguished by its comprehensive approach, this study expands the horizons of related research. While existing studies primarily deploy SVM and RF with limited sample sizes, our work introduces four classifiers, three novel to this problem, and leverages a substantial 2700 OCT images. Notably, with a four-class setting, we achieve remarkable classification accuracies of 99.43%, 99.54%, 97.98%, and 99.31%, respectively. Compared to Abdulrahman & Khatib [5] and Liu et al. [7], achieving 90.65% and 90.5%, respectively, our models shine brighter. Obviously, the large number of OCT images conducted in the proposed study compared with the related studies impacts positively on the results.

Additionally, the study introduces five evaluation criteria that were not yet employed for relevant works. The criteria are error rate, false positive rate, false negative rate, negative predictive value, and 1-score. Almost all of the related works use only accuracy, sensitivity, specificity, and precision. In another context, while Dash and Sigappi attain 100% accuracy using LBP, their scope is confined to two classes and only 40 OCT images [8]. This study, however, demonstrates a pioneering leap in complexity and accuracy across a multi-class scenario, demonstrating significant progress in retinal disorder classification.

Through meticulous comparison with pertinent studies that adopted hybrid techniques for classification, our proposed models notably outperform the results attained in the study [9]. This achievement can be attributed to the enhanced complexity and the adept utilization of the InceptionV3 model as a feature extractor. This facet empowers the extraction of an extensive array of features, prioritizing high-level attributes that significantly contribute to the classification process.

According to Table 4, Model (1), in which the SVM was implemented as a classifier, attains superior evaluation metrics, such as accuracy and sensitivity, in comparison to Khan et al.'s [10] study. Also, it demonstrates a higher number of tested samples. It's important to note that the increased sample size

might potentially influence the outcome, warranting further investigation into the interplay between sample size and performance.

Table 4
Comparison of different retinal classification models

Study	Classes	Classifier	Tested samples	Accuracy	Sensitivity	Specificity
Hussain et al. (2018)	2	RF	251	95%	94.67%	100%
Alsaih et al. (2017)	2	SVM	238	NA	87.5%	87.5%
Abdulrahman and Khatib (2018)	4	SVM	160	90.65%	NA	NA
Srinivasan et al. (2014)	3	SVM	45	95.5%	NA	NA
Liu et al. (2011)	4	SVM	131	90.5%	NA	NA
Dash and Sigappi (2018)	2	RF	40	100% for LBP, 88% SITF	100% for LBP, 91% for SITF	NA
Salaheldin et al (2022)	4	SqueezeNet + KNN	1800	97.47%	94.94%	98.31%
Khan et al. (2023)	5	DenseNet 201 + Cubic SVM	601	99.1%	98.2%	NA
Model (1)	4	SVM	2700	99.43%	98.85%	99.62%
Model (2)	4	K-NN	2700	99.56%	99.11%	99.70%
Model (3)	4	DT	2700	97.98%	95.96%	98.65%
Model (4)	4	EM	2700	99.31%	98.63%	99.54%

Conclusion

The classification of retinal disorders is one of the most important issues that ophthalmologists are concerned about because of its impact on people's daily lives. Using 18000 OCT images, the authors proposed and designed an automated expert system to classify retinal disorders into CNV, DME, Drusen, and normal cases. SVM, K-NN, DT, and EM are the four classifiers presented. The Inception V3 CNN was used to extract the features. For the highlighted problem, the authors present a novel technique that combines machine learning and deep learning approaches. The proposed four models achieved significant results when compared to relevant studies. Due to the results, the K-NN model achieved the best classification accuracy of 99.56%. The study introduces nine performance computation evaluation

criteria, some of them were not used for this problem. Indeed, the study provides a novel platform to assist ophthalmologists in the diagnosis and classification of retinal disease, thereby reducing effort and time. This research can be applied to other areas of interest, including neuro ophthalmic diseases such as Papilledema and Pseudo Papilledema detection. Furthermore, additional classifiers can be used for the same dataset to benchmark the results.

Declarations

Author contributions

The authors' contributions to the paper are outlined as follows: A. M. Salaheldin, M. A. Wahed, and N. Saleh were responsible for the study's conception and design. A. M. Salaheldin and N. Saleh carried out the analysis and interpretation of the results. A. M. Salaheldin, M. A. Wahed, and N. Saleh participated in manuscript preparation. All authors critically reviewed the results and provided their approval for the final version of the manuscript.

Data availability

The data that support the findings of this study are openly available in https://www.kaggle.com/paultimothymooney/kermany2018

Competing interests

The authors declare no competing interests.

References

1. Bourne, R. *et al.* Trends in prevalence of blindness and distance and near vision impairment over 30 years: an analysis for the Global Burden of Disease Study. Lancet Glob. Heal. 9, e130–e143 (2021).
2. Hussain, M. A. *et al.* Classification of healthy and diseased retina using SD-OCT imaging and Random Forest algorithm. PLoS One 13, e0198281–e0198281 (2018).
3. Saleh, N. & Salaheldin, A. M. A benchmarking platform for selecting optimal retinal diseases diagnosis model based on a multi-criteria decision-making approach. J. Chinese Inst. Eng. Trans. Chinese Inst. Eng. A 45, 27–34 (2022).
4. Alsaih, K. *et al.* Machine learning techniques for diabetic macular edema (DME) classification on SD-OCT images. Biomed. Eng. Online 16, 68 (2017).
5. Abdulrahman, H. & Khatib, M. Classification of Retina Diseases from OCT using Genetic Programming. Int. J. Comput. Appl. 177, 41–46 (2020).
6. Srinivasan, P. P. *et al.* Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. Biomed. Opt. Express 5, 3568–3577 (2014).

7. Liu, Y.-Y. *et al.* Computerized macular pathology diagnosis in spectral domain optical coherence tomography scans based on multiscale texture and shape features. Invest. Ophthalmol. Vis. Sci. 52, 8316–8322 (2011).
8. Dash, P. & Sigappi, A. N. Detection and recognition of diabetic macular edema from oct images based on local feature descriptor. Int. J. Pure Appl. Math. 119, 1–7 (2018).
9. Saleh, N., Abdel Wahed, M. & Salaheldin, A. M. Computer-aided diagnosis system for retinal disorder classification using optical coherence tomography images. 67, 283–294 (2022).
10. Khan, A., Pin, K., Aziz, A., Han, J. W. & Nam, Y. Optical Coherence Tomography Image Classification Using Hybrid Deep Learning and Ant Colony Optimization. *Sensors* vol. 23 (2023).
11. Kermany, D. S. *et al.* Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell 172, 1122–1131.e9 (2018).
12. Saleh, N., Abdel Wahed, M. & Salaheldin, A. M. Transfer learning-based platform for detecting multi-classification retinal disorders using optical coherence tomography images. Int. J. Imaging Syst. Technol. 32, 740–752 (2022).
13. Perona, P. & Malik, J. Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. 12, 629–639 (1990).
14. Banerjee, A. Intuitive Hyperparameter Optimization: Grid Search, Random Search and Bayesian Search! *Medium* https://towardsdatascience.com/intuitive-hyperparameter-optimization-grid-search-random-search-and-bayesian-search-2102dbfaf5b (2019).
15. Taylor, J. R. *An introduction to error analysis*. (1982).
16. Salaheldin, A. M., Abdel Wahed, M. & Saleh, N. Machine Learning-Based Platform for Classification of Retinal Disorders Using Optical Coherence Tomography Images. in *Artificial Intelligence and Sustainable Computing* (eds. Pandit, M., Gaur, M. K., Rana, P. S. & Tiwari, A.) 269–283 (Springer Nature Singapore, 2022). doi:10.1007/978-981-19-1653-3_21.

Figures

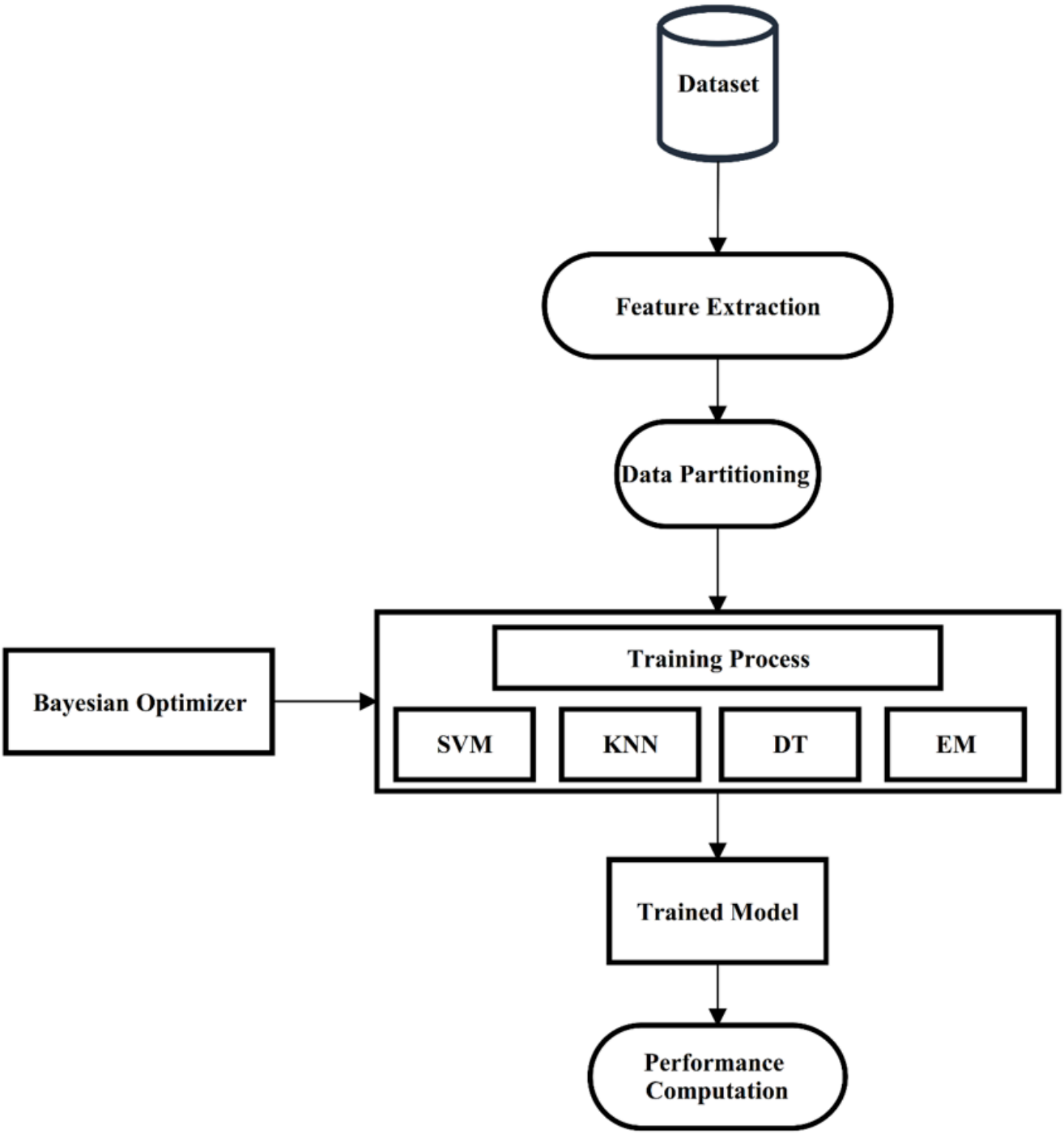

Figure 1

Block diagram of the proposed methodology

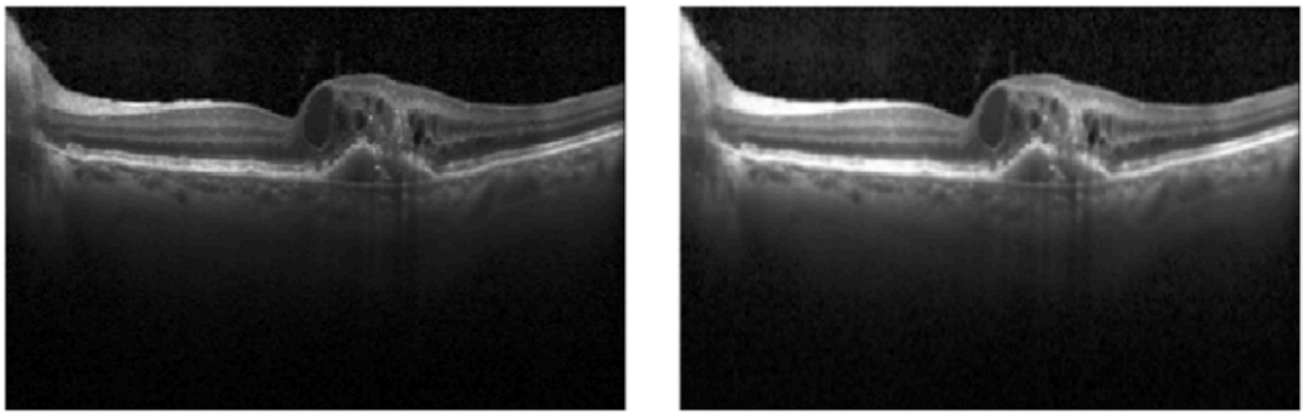

Figure 2

A sample of the CNV class (a) before processing and (b) after processing

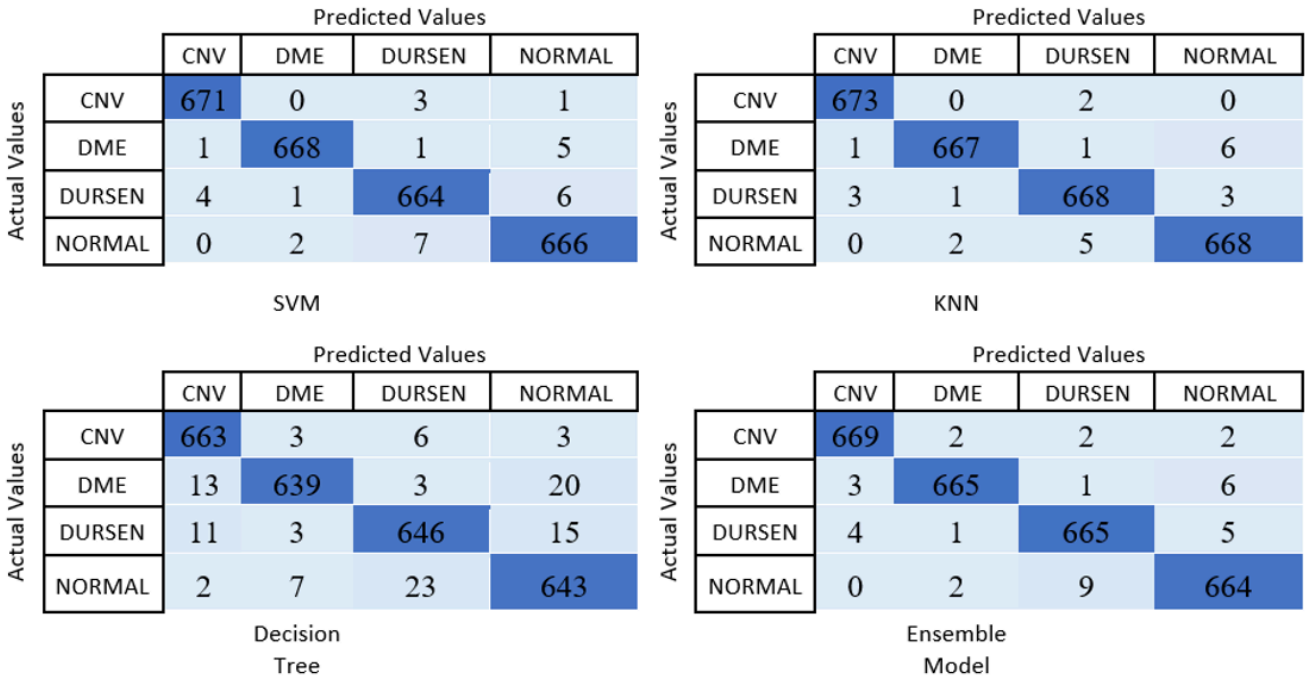

Figure 3

The proposed classifiers confusion matrices