Implementation Of Support Vector Machine And Harmony Search For Cataract Severity Classification In Fundus Images

Firdausa Yasmin Hermadiputri[1] , Eka Prakarsa Mandyartha[2] , Agung Mustika Rizki[3]
[1]Informatics, Faculty of Computer Science, UPN "Veteran" East Java
[2,3]Master of Informatics, Faculty of Computer Science, UPN "Veteran" East Java
firdausayasmin@gmail.com

ABSTRACT

Cataract is a condition that causes clouding of the lens of the eye and is a leading cause of blindness, including in Indonesia. Cataract diagnosis is often inconsistent between ophthalmologists due to personal experience. This research proposes a Support Vector Machine (SVM) based classification system and Harmony Search metaheuristic algorithm to optimize the weight vector 'w' on the SVM hyperplane as a supporting tool for cataract diagnosis. The research data comes from Kaggle which includes normal eye fundus images and cataracts with mild-moderate and severe levels. The research stages include image conversion from RGB to Grayscale, image enhancement with Histogram Equalization and GLCE, and feature extraction using GLCM and Haar Wavelet Transform, and unbalanced data is balanced by the SMOTEENN method. The results showed that Harmony Search successfully improved SVM accuracy compared to Conventional SVM using Gradient Descent. Accuracy increased by 18% from 0.53 to 0.71 on unbalanced data, and by 13% from 0.67 to 0.80 on balanced data. In addition, Harmony Search can improve computational time efficiency due to its ability to explore space globally.

Keywords: Cataract, Eye Fundus, Support Vector Machine, Harmony Search.

INTRODUCTION

Cataract is a condition in which the lens of the eye becomes cloudy resulting in blurred vision[1], [2]. This is caused by biochemical reactions and proteolytic breakdown of crystallins (lens proteins) called protein coagulation.[3]. Cataracts can be classified by severity or maturation into mild-moderate (incipient and immature) and severe (mature and hypermature) cataracts.[2]. Although commonly associated with the aging process, cataracts can also occur in children from birth or after eye injury, inflammation, trauma, and some other eye diseases.[2].

According to the World Health Organization (WHO), cataracts are the leading cause of visual impairment in the world and one of the leading causes of blindness. In Indonesia, the blindness rate due to cataracts is the highest in Southeast Asia, reaching 1.5% of the world's total 17 million population.[2], [4]. Indonesia's subtropical geography contributes to the high rate of blindness. The Rapid Assessment of Avoidable Blindness (RAAB) survey conducted in 15 regions in Indonesia (2014-2016) shows that East Java province has the highest blindness rate at 4.4%, with cataract as the main cause at 73.8%.[4].

Diagnosis of cataracts can be made through observation of the fundus image of the eye as it includes the retina, macula, optic disc, fovea and blood vessels.[5]. The fundus image is taken by reflecting light into the eye, so that the part can be seen, with the reflected light representing the intensity of RGB waves and the quantity of reflected light. Fundus images can be used to detect various eye diseases as they provide details of the affected area. However, observation of an eye fundus image for cataract diagnosis requires an ophthalmologist to classify or determine the severity of the cataract. These observations by ophthalmologists can be influenced by personal experience, potentially resulting in inconsistent accuracy between experts.

Therefore, this study proposes a cataract classification or grading system as a computer-aided system that supports ophthalmologists in cataract diagnosis through fundus images. The proposed system combines digital image processing, pattern recognition, and machine learning. The methods used are Support Vector Machine (SVM) and Harmony Search algorithm. SVM is chosen because it is able to work effectively with clear separation boundaries between classes, suitable for complex and limited datasets.[6]. Harmony Search algorithm, as a metaheuristic algorithm, plays an important role in achieving the optimal weight vector w from the SVM hyperplane to optimize feature weighting in eye fundus images.[7]. The optimal weight vector is the vector that minimizes the SVM objective function [8].[8].

Through the development of a cataract severity classification system as a computer-aided system using the Support Vector Machine (SVM) method with the One-Versus-Rest (OvR) approach and the Harmony Search algorithm, it is expected to create an effective and efficient system in classifying or grading cataract disease and become a tool for ophthalmologists in diagnosing cataract disease by utilizing digital image processing technology, pattern recognition, and machine learning.

METHODS

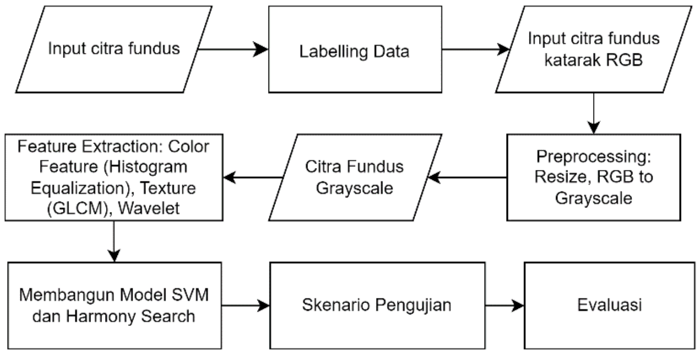


Figure 1. Research Methodology

Data Collection and Labelling

The first step in this research starts with accessing Kaggle. The data used contains 4217 eye fundus images with various conditions such as normal eyes, diabetic retinopathy, glaucoma, and cataracts. This dataset has a fairly high usability rate of 8.75. Data collection was done through direct download from Kaggle, and appropriate labelling of each fundus image associated with a specific eye condition.

Once the data collection process is complete, the next step is labelling, where the entire dataset is labeled based on cataract severity. The purpose of this labelling process is to prepare the dataset so that it can be used in the development of a classification model that can distinguish between different levels of cataract severity. Labelling was done in three main categories: (1) non-cataract or normal, which includes fundus images of eyes that show no signs of cataract; (2) mild-moderate cataract, which includes early or mild stages of cataract (incipient) and more advanced stages of development that have not yet reached full maturity (immature); and (3) severe cataract which includes mature and hypermature stages, representing advanced cataract conditions. The labelling process was done in consultation with an eye specialist from RSUD dr Soegiri Lamongan to ensure consistency and accuracy in labelling.

Image Preprocessing

In the preprocessing stage for converting the dataset from RGB to Grayscale format, several approaches were taken to compare the optimal results. The first approach is to take the intensity value of one particular channel, such as red, green, or blue, from each image in the dataset. This process involves opening the image and extracting the values of each RGB (Red, Green, Blue) channel. Next, the RGB intensity averaging approach is performed, where the R (Red), G (Green), and B (Blue) values of each pixel are summed and then divided by 3. The final approach is to use the luminosity formula, which assigns different weights to each channel with coefficients of 0.21 for R, 0.72 for G, and 0.07 for B.[9]. After applying these three methods, the quality of the resulting grayscale image is evaluated. The purpose of this step is to improve the clarity of image details and save storage space for large and complex datasets such as eye fundus images.

Image Enhancement

After the preprocessing stage, the next step is to perform image enhancement using the histogram equalization algorithm. The process begins by loading the image dataset and converting it into a numpy array to facilitate data processing. After that, the number of pixels at each gray level in the image is calculated, which covers a range of values from 0 to L-1 (with L being the total number of gray levels which is 256). Next, the histogram of the original image is calculated to determine the probability distribution of the gray levels. This probability is calculated by the formula

$$P_i(r_i) = ni/n$$

Where pi is the probability distribution of the color intensity of each pixel, ni is the frequency of color occurrence at a pixel, and n is the total number of pixels in the image.
Then, the Cumulative Distribution Function (CDF) is calculated to create a cumulative distribution function from the image histogram using the equation

$$s_k r_k = \Sigma_{(i=0)}^{k} P_i(r_i), k = 0, 1, 2, \ldots, L - 1$$

The last step is to apply histogram equalization or flatten the intensity distribution of the image color histogram with the formula

$$h(v) = INT((L - 1) \text{ x } CDF$$

This approach aims to flatten the histogram distribution of the original image[10]This approach aims to flatten the histogram distribution of the original image [10], improve image contrast and clarity efficiently, and clarify visual details before further analysis. The histogram equalization algorithm was chosen because this algorithm is better at improving the contrast of image data by looking at statistical values, namely PSNR and MSE, where the higher the PSNR, the better the quality and the lower the MSE value, the lower the error value.[11].

Feature Extraction

After completing the image enhancement process, the next step is to perform feature extraction from the eye fundus image. This process involves two main techniques, namely Grey-Level Co-occurrence Matrix (GLCM) and wavelet transform.

First, GLCM is used to extract texture features from the image. The steps start with loading the image dataset, initializing the GLCM feature result store, and adjusting the directions such as 0°, 45°, 90°, and 135°. Each image in the dataset is then processed to generate GLCM with steps such as initialization for each direction, creation of a co-occurrence matrix with a size determined by the image bits, calculation of the color distribution at each pixel, frequency of occurrence of neighboring pixel intensity pairs in the direction, normalization of the co-occurrence matrix to obtain a normalized value, and calculation of the 4 main GLCM features of contrast, correlation, energy, and homogeneity from the normalized matrix.[12].

Secondly, wavelet transform is used to obtain more in-depth details from the image. This process involves the use of a Haar wavelet filter with steps such as initialization of variables to store the Haar wavelet transform results, selection of the desired or maximum decomposition level, non-overlapping filtering process with a 2x2 matrix to obtain sub-bands such as LL (Low-Low), LH (Low-High), HL (High-Low), and HH (High-High), and storage of the Haar wavelet transform coefficients for further analysis.[10].

The combination of these two feature extraction techniques aims to produce a more robust and detailed feature representation of the eye fundus image, which will be used in the classification process or as input.

Data Resampling

After obtaining feature information using GLCM and Haar wavelet transform, the next step is to perform dataset balancing or data resampling using the SMOTE + ENN method. SMOTE is used to add synthetic data in the minority class, while ENN is used to remove irrelevant samples in the majority class.

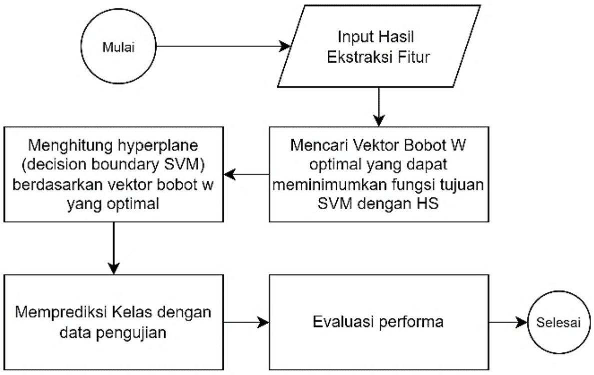
The process starts with a data pre-processing step, where the dataset is divided into features (X) and labels (y) to determine the proportion between minority and majority classes. Next, SMOTE is applied using the KNN algorithm to identify the nearest neighbor of each data in the minority class. The stages of SMOTE include minority class identification, random data selection, calculation of the distance between the random data and the nearest neighbor, and generation of synthetic samples by taking the difference between the random data and the nearest neighbor multiplied by a random value between 0 and 1. This process is repeated until reaching the number of synthetic samples corresponding to the desired proportion of the minority class.

After the SMOTE process is completed, the next step is to apply ENN to remove the samples that do not correspond to the nearest majority class. ENN uses the KNN approach to find the nearest neighbor of each data and evaluate the nearest majority class. Samples that do not correspond to the nearest majority class are removed from the dataset, and this process is repeated until it reaches the desired balance of proportions between minority and majority classes.

SMOTEEN was chosen as the data resampling method because it has better robustness than other resampling methods by looking at the symmetrical distribution of the data and the median value in the boxplots diagram.[13].

Implementation of SVM Model and Harmony Search

The implementation of the hybrid SVM and Harmony Search method is used to identify cataracts based on severity. Support Vector Machine algorithm is used for data classification method in 3 different classes, namely non-cataract, cataract/mild, and severe cataract. Then, Harmony Search which is a metaheuristic algorithm is used to optimize the weight vector w on the Support Vector Machine hyperplane and get the optimal fundus image feature weighting value. The meaning of the optimal weight vector w is a vector that can minimize the objective function of the SVM algorithm. Harmony Search is implemented in the training or 'fit' function in the Support Vector Machine model. Figure 2 is the flowchart of the proposed system (SVM model) and Figure 3 is the harmony search algorithm applied in the training stage of the SVM model.



SVM Model Flowchart

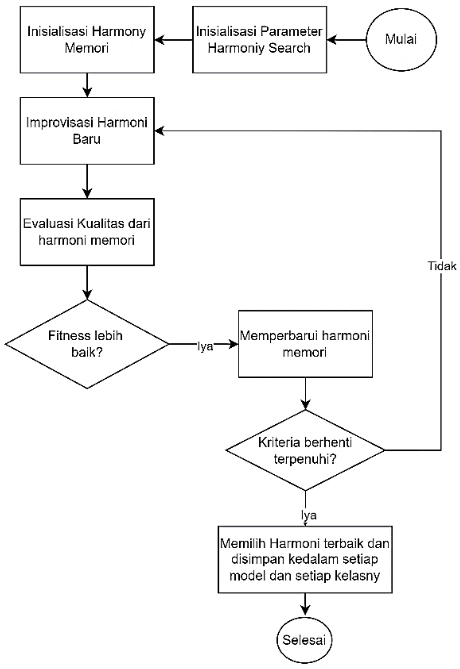


Figure 3. Harmony Search Algorithm for SVM Training

Training and Testing

The system was trained and tested using resampled datasets. To obtain a high level of accuracy and efficient computing time, training and testing were conducted with five different data sharing proportions of 50%, 40%, 30%, 20%, and 10%. The hyperparameters used in the model were explored using several values based on the range of values in previous studies[14]. The lambda hyperparameter with values of 0.01, 0.1, and 1, the number of iterations in the Harmony Search optimization algorithm was varied between 50, 100, and 500 to understand the impact of training duration on the final result. HMCR (Harmony Memory Consideration Rate) with values of 0.3, 0.4, and 0.5, pitch adjust rates (PAR) in the range of 0.7, 0.8, and 0.9. Finally, the bandwidth for tuning the optimization process with values of 0.001 and 0.01. Each of these combinations is thoroughly evaluated to optimize the performance of the classification model, with a focus on the accuracy and stability of the model across different test data conditions.

Model Evaluation

The last step is to analyze the evaluation metrics on the classification report to determine the performance of the system model, namely on accuracy and computation time.

RESULTS

Data Collection and Data Labelling

This study classifies eye fundus images into three classes: non-cataract, mild-moderate cataract, and severe cataract. The dataset used amounted to 1023 non-cataract/normal eye fundus images and 1060 cataract fundus images that had not been grouped by severity. Therefore, in this study, re-labelling was carried out with an eye specialist at one of the hospitals in Lamongan to make it more valid. It was found that there were 208 fundus images labeled as cataract/light-moderate and 855 fundus images labeled as cataract/severe. Figures 4a, 4b, 4c show the difference in fundus images for each class. It can be seen that normal/non-cataract eye fundus images show parts of the eye fundus very clearly, cataract/moderate eye fundus images show parts of the eye fundus already have opacities but still thin, and cataract/severe eye fundus images show parts of the eye fundus are not visible at all.

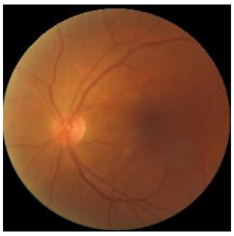


Figure 4(a) non-cataract **Image (b)** cataract/moderate **Picture (c)** cataract/severe

Data Preprocessing

After that, the processing stage is carried out by converting the RGB image format into gray scale with an approach to getting the most suitable results in this study. Figure 5 is the result of the RGB to graylevel conversion stage. It can be analyzed that the best or clearest graylevel image results are by taking the green channel value component because the human eye is more sensitive to green so that the details in the fundus image can be clearer.

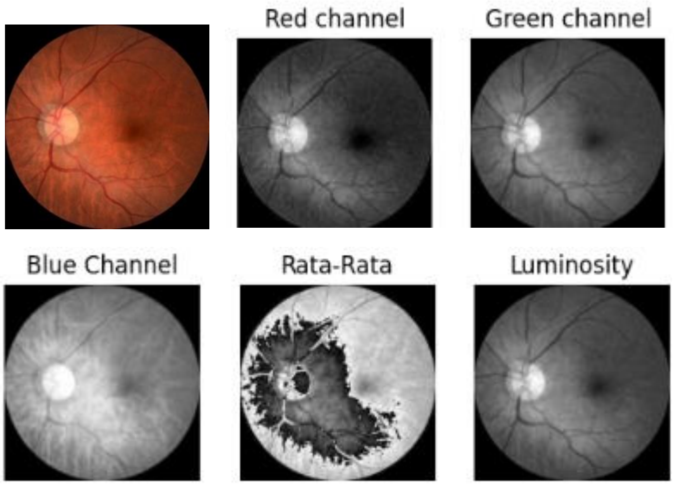
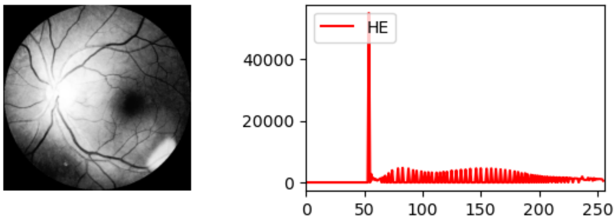


Figure 5. Preprocessing Results

Image Enhancement

Then the quality or contrast enhancement is carried out on the image using the Histogram Equalization algorithm by flattening the pixel color distribution. The result of image enhancement is shown in Figure 6



Histogram Equalization Result

Feature Extraction

The next step is to perform feature extraction using GLCM to find out the structure/texture information in the image and Haar wavelet transformation to find out the details in the image with the frequency or signal in the image. The results of this stage are used as model input, including 32 features generated from both feature extraction algorithms. Figure 7 shows an example of the results of the feature extraction stage.

```
        LL_energy    HL_energy    LH_energy   HH_energy    All_energy
0     5.859588e+09   4312883.50   4415159.50  1396135.50  5.869712e+09
1     5.862841e+09   3645120.50   4650225.50  1236925.50  5.872373e+09
2     5.829596e+09  14568708.75  13144706.75  8935782.75  5.866245e+09
3     5.830596e+09  11928353.50  10551402.50  7411449.50  5.860487e+09
4     5.813047e+09   3632604.75   3765074.75  1364797.75  5.821810e+09
...            ...          ...          ...         ...           ...
2100  1.451041e+09   1545064.50   1783364.50   581109.50  1.454950e+09
2101  1.460511e+09   1214954.25   1566087.25   551571.25  1.463844e+09
2102  1.460333e+09   1148764.75   1458217.75   706832.75  1.463647e+09
2103  1.502133e+09   1857855.25   1761683.25  1389515.25  1.507142e+09
2104  1.454873e+09    721332.75   1034426.75   253317.75  1.456882e+09

[2105 rows x 32 columns]
```

Figure 7. Feature Extraction Results

Data Resampling

Due to the unbalanced amount of data in each class, data resampling is performed using the SMOTE algorithm as oversampling and ENN as undersampling. With this algorithm, the advantage is that irrelevant data in the majority class will be removed. Figure 8 is the result of resampling data using SMOTEENN at each data division size in the training data.

```
Test size: 0.5
Sebelum SMOTE-ENN: Counter({-1: 504, 1: 435, 0: 113})
Setelah SMOTE-ENN: Counter({0: 294, 1: 217, -1: 190})

Test size: 0.4
Sebelum SMOTE-ENN: Counter({-1: 615, 1: 518, 0: 130})
Setelah SMOTE-ENN: Counter({0: 385, 1: 264, -1: 231})

Test size: 0.3
Sebelum SMOTE-ENN: Counter({-1: 711, 1: 610, 0: 152})
Setelah SMOTE-ENN: Counter({0: 456, 1: 306, -1: 270})

Test size: 0.2
Sebelum SMOTE-ENN: Counter({-1: 824, 1: 692, 0: 168})
Setelah SMOTE-ENN: Counter({0: 511, 1: 375, -1: 347})

Test size: 0.1
Sebelum SMOTE-ENN: Counter({-1: 921, 1: 783, 0: 190})
Setelah SMOTE-ENN: Counter({0: 565, 1: 421, -1: 403})
```

Figure 8. Results of SMOTEENN Stages

Testing Results

After building the system model, training and testing are carried out on data that has been resampled using SMOTEENN. Because this research uses a hybrid method, to find out the performance of Harmony Search, a comparison is made with Conventional SVM using Gradient Descent, the results of which can be seen in table 1.

Table 1. Testing Results

No.	SVM Model	Highest Accuracy	Average Accuracy	Computation Time
1.	SVM Gradient Descent	0.53	-	1.743
2.	SVM Gradient Descent and SMOTEENN data resampling	0.67	-	1.151
3.	**Hybrid SVM and HS**	0.73	0.71	0.24
4.	**Hybrid SVM and HS and SMOTEENN data resampling**	0.82	0.80	0.276

DISCUSSION

In the multiclass classification tests conducted, it is seen that the use of SVM optimized with Harmony Search (HS) provides the most significant results. In the first model using SVM with Gradient Descent (GD) without resampling data, the highest accuracy achieved was only 0.53 with a computation time of 1.743 seconds. This model is less effective because it does not consider data imbalance, so the accuracy is low and the computation time is long. When data resampling using SMOTEENN is applied to SVM with GD, the highest accuracy increases to 0.67 and the computation time is slightly faster to 1.151 seconds. This shows that data balancing plays an important role in improving model performance in both accuracy and computation time.

However, a more significant improvement is seen in the proposed system. The system achieved the highest accuracy of 0.73 with an average accuracy of 0.71, and a computation time of 0.24 seconds. The use of Harmony Search in optimizing SVM parameters allows for a more efficient search and faster convergence compared to the pure Gradient Descent method. This shows that Harmony Search is able to find more optimal parameters for SVM, which has an impact on improving accuracy and computational time efficiency. Finally, the Hybrid SVM model with Harmony Search and SMOTEENN data resampling gave the best results, with the highest accuracy of 0.82 and an average of 0.80, and the fastest computation time of 0.276 seconds. These results prove that optimizing SVM parameters using Harmony Search and SMOTEENN data balancing techniques is very effective in improving model performance for multiclass classification. Harmony Search not only helps in better exploration and exploitation of the parameter space, but ensures that the model can reach the optimal solution with more efficient computation time. Therefore, in this study, the Hybrid SVM method with Harmony Search proved to be very superior in optimizing the weight vector w in SVM to achieve the best performance.

CONCLUSION

Based on the severity of the fundus image, the proposed system successfully classifies and identifies cataract disease more accurately and efficiently with the highest accuracy reaching 0.83 and computation time 0.24 compared to the SVM method using Gradient Descent which only reaches the highest accuracy of 0.67 in 1.151 time. Broadly speaking, the results show that the system can be used as a computer-aided system that can assist ophthalmologists in making an initial diagnosis of cataract disease. This is because harmony search can explore and exploit more widely than gradient descent which depends on the objective function. In the future, it is hoped that this system can be integrated as a cataract severity classification system that can assist eye specialists in making an initial diagnosis of cataracts.

LIMITATIONS

This research has limitations, which include the lack of datasets or dataset imbalance, especially in mild/moderate cataracts because the number of sufferers of this level is less than the severe level and the characteristics of these two levels are almost the same so that data resampling is carried out. Then, the hyperparameter values used have not been explored in the entire range (only using the values recommended in previous studies. Therefore, in future research, it is expected that the number of datasets at each severity level can be more balanced and a wider selection of hyperparameter values is made without relying on recommended values to conduct further analysis.

REFERENCE

M. K. Behera, S. Chakravarty, A. Gourav, and S. Dash, "Detection of nuclear cataract in retinal fundus image using radial function-based SVM," in *PDGC 2020 - 2020 6th International Conference on Parallel, Distributed and Grid Computing*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020, pp. 278-281. doi: 10.1109/PDGC50313.2020.9315834.

K. L. Moreau and J. A. King, "Protein misfolding and aggregation in cataract disease and prospects for prevention," *Trends in Molecular Medicine*, vol. 18, no. 5. pp. 273-282, May 2012. doi: 10.1016/j.molmed.2012.03.005.

J. Education and D. Counseling, "The Relationship between Cataract Incision Type and Dry Eye Syndrome in Post-Cataract Surgery Patients at Majalengka Eye Clinic, Majalengka Regency in 2021."

L. Rif'Ati, A. Halim, Y. D. Lestari, N. F. Moeloek, and H. Limburg, "Blindness and Visual Impairment Situation in Indonesia Based on Rapid Assessment of Avoidable Blindness Surveys in 15 Provinces," *Ophthalmic Epidemiology*, vol. 28, no. 5, pp. 408-419, 2021, doi: 10.1080/09286586.2020.1853178.

J. Research, M. Dan, P. Mathematics, A. E. Suwanda, D. Juniati, and U. N. Surabaya, "Classification of Eye Diseases Based on Retinal Fundus Image Using Fractal Box Counting Dimension and Fuzzy K-Means," 2022.

T. B. Sasongko, "Comparison and Analysis of SVM and PSO-SVM Algorithm Model Performance (Case Study of High School Interest Path Classification)," 2016.

L. Frank, "Support Vector Machines (SVMs)". Available: https://www.researchgate.net/publication/380881348

Institute of Electrical and Electronics Engineers and IEEE Instrumentation and Measurement Society, *IST 2017: IEEE International Conference on Imaging Systems and Techniques: Beihang University, Beijing, China, October 18-20, 2017, Beijing China: 2017 conference proceedings*.

R. Andrian, A. Junaidi, and D. Indah Lestari, "APPLICATION OF MEASURING PLANT LEAF AREA USING ANDROID-BASED DIGITAL IMAGE PROCESSING." 2022.

J. Xiong *et al.*, "Application of Histogram Equalization for Image Enhancement in Corrosion Areas," *Shock and Vibration*, vol. 2021, 2021, doi: 10.1155/2021/8883571.

E. P. Mandyartha, H. E. Wahanani, and P. W. Atmaja, *IMPLEMENTATION OF GLOBAL-LOCAL CONTRAST ENHANCEMENT FOR CONTRAST IMPROVEMENT OF RETINA FUNDUS IMAGES*. [Online]. Available: https://www.mathworks.com/help/images/ref/hist

W. I. Praseptiyana, A. W. Widodo, and M. A. Rahman, "Utilization of Gray Level Co-occurrence Matrix (GLCM) Feature for Melasma Detection in Facial Image," 2019. [Online]. Available: http://j-ptiik.ub.ac.id

Y. A. Sir and A. H. H. Soepranoto, "Data Resampling Approach to Handle Class Imbalance Problem," *Journal of Computers and Informatics*, vol. 10, no. 1, pp. 31-38, Mar. 2022, doi: 10.35508/jicon.v10i1.6554.

J. Wang, H. Ouyang, S. Li, W. Ding, and L. Gao, "Equilibrium optimizer-based harmony search algorithm with nonlinear dynamic domains and its application to real-world optimization problems," *Artificial Intelligence Review*, vol. 57, no. 7, Jul. 2024, doi: 10.1007/s10462-024-10793-4.