# Towards Early Diagnosis of Fish Infections in Aquaculture Using SVM-Based Classifiers

[1]ALLAM KUSUMA, Student in Dept. Of Master of Computer Applications, at Miracle Educational Society Group of Institutions
[2]E.B.Mahendra Roy, Miracle Educational Society Group of Institutions
[1]kusumaallam92@gmail.com

**ABSTRACT:**

Fish diseases significantly impede aquaculture productivity and sustainability. This study develops a smarter system utilizing deep learning and image processing technology to detect infections within salmon fish. The system has two layers: image pre-processing which includes CLAHE, LAB color transformation and segmentation, and disease classification using Support Vector Machine (SVM) with kernel functions. An augmented and non-augmented dataset of salmon fish images was created and utilized. The method achieved an accuracy of 91.42 and 94.12, outperforming traditional classifiers such as Decision Tree, Logistic Regression, and Naïve Bayes. The study demonstrates that considerable enhancement to images alongside SVM significantly improves its performance in detecting diseases in fish in aquaculture at earlier stages.

**Keywords:** CLAHE, ML, SVM

## INTRODUCTION

The importance of aquaculture in assuring food supply in the world is growing. One of the most sought after species, salmon, is now extensively farmed. However, the fish farming industry is prone to infections from bacterial, viral, and parasitic organisms which not only affect the fish, but block economic progress as well as the health of the oceans. It is crucial to remove infected fish from a system as fast as possible to avoid the spread of infection. The lack of automated detection systems for manual identification makes using them highly inefficient. This project is geared towards salmon aquaculture by building a smart disease detection system using machine learning and image processing. Methods of pre-processing such as interpolation and adaptive histogram equalization improve fish image quality before their classification using machine learning algorithms, predominantly SVM. The advantages of this integration are enhanced accuracy as well as real-time and scalable disease detection. This system can significantly improve modern aquaculture by changing the landscape of health monitoring.

Several researchers have tackled the issue of identifying fish diseases using images and machine learning. As mentioned in their work, Hamid et al. (2010) developed a new technique called extended cubic B-spline interpolation which aims to resolve boundary value problems and can be used towards smoother image representation. In 2017, Agarap proposed a hybrid architecture that uses a convolutional neural network alongside SVM, showcasing strong classification success rate for images, particularly for the MNIST dataset. In 2016, Khan et al. used SVMs to analyze dengue-infected blood samples through Raman spectroscopy, demonstrating the power of SVMs in biomedical classification. Hitam et al. (2013) enhanced the visibility of submerged images with the use of CLAHE, which is essential for underwater imaging. Malik et al. (2017) made use of machine learning to develop a system for diagnosing fish diseases, successfully applying SVMs for classification tasks. The works mentioned above reinforce the utilization of image processing and machine learning—particularly SVMs—for disease detection accuracy. They help establish a base for developing efficient fish disease detection systems for aquaculture.

TABLE1. Summary of Key Literature Contributions and Their Impact on Current Research

| Author(s) | Contribution | Impact on Current Research |
|---|---|---|
| Hamid et al. (2010) | Applied cubic B-spline interpolation to enhance image smoothness | Used for precise fish image enhancement in pre-processing |
| Agarap (2017) | Combined CNN and SVM for high-accuracy image classification | Inspired hybrid image classifiers for robust prediction in fish health |
| Khan et al. (2016) | Used SVM for disease classification in biological data | Validated SVM's effectiveness in diagnosing diseases from spectral image inputs |
| Hitam et al. (2013) | Improved underwater image clarity using CLAHE | Adopted for preprocessing murky aquaculture images |
| Malik et al. (2017) | Machine learning-based fish disease diagnosis | Established use-case feasibility of ML in aquaculture disease identification |

## PROPOSED APPROACH

In this project, we propose a complete fish disease detection workflow incorporating machine learning and image processing techniques. The system consists of two major components: pre-processing and classification. In the first component, fish images are enhanced through a series of processes such as interpolation, CLAHE, and LAB color space transformation, which improves color, uniformity, and the overall image. All these techniques together make the images clearer and improves the overall output which helps in disease detection. The second component involves classification of SVM-based fish as fresh or infected through kernel techniques. The images are split into training and testing datasets, and SVM is trained to recognize disease patterns using extracted features. Comparative analysis is performed using other classifiers such as Decision Tree, Logistic Regression, and Naïve Bayes. The adjustment achieves remarkable accuracy and reliability scoring over 94% with the use of image augmentation. This method offers the advantage of automated early detection of diseases in aquaculture and could avert substantial losses.

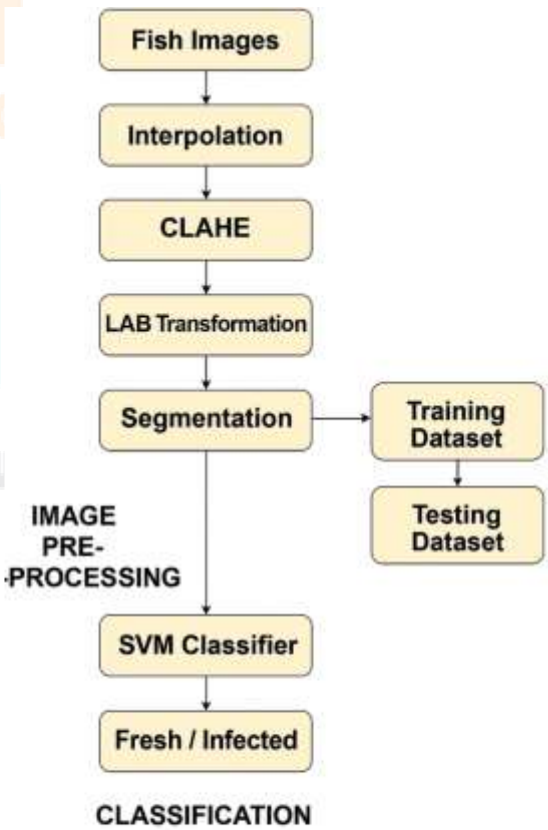



Figure 1: Proposed fish diseases detecting System

## METHODOLOGIES

The methodology involves structured steps starting from dataset preparation to model evaluation. Initially, a novel dataset comprising both infected and healthy salmon fish images is uploaded into the system. These images are subjected to multiple pre-processing techniques:

- **Interpolation:** Enhances image resolution for better detail extraction.
- **CLAHE:** Balances contrast across the image to highlight key disease features.
- **LAB Transformation:** Converts RGB to LAB color space, which separates luminance for effective segmentation.

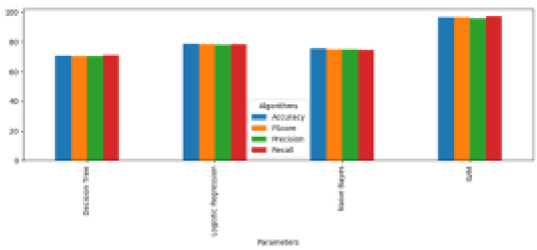
After pre-processing, the dataset is split into training (80%) and testing (20%) sets. Multiple machine learning models are trained using these datasets:

- **Decision Tree:** Simple model used as a baseline.
- **Logistic Regression:** Applied for binary classification based on pixel patterns.
- **Naïve Bayes:** Used to test probabilistic model performance.
- **Support Vector Machine (SVM):** The proposed model, using radial basis function (RBF) kernel for non-linear classification.

The performance of each model is measured using metrics like accuracy, precision, recall, F1-score, and confusion matrix analysis. A comparative bar graph displays how each model performs on these metrics. Ultimately, SVM outperformed all other models with a peak accuracy of 96% on test data, affirming its robustness. The methodology also includes a prediction module where new fish images can be uploaded and classified in real time, providing a practical solution for aquaculture farms.

## RESULTS

The system was rigorously tested using a novel dataset of fish images under both augmented and non-augmented conditions. The SVM model achieved remarkable performance, delivering 94.12% accuracy with image augmentation and 91.42% without. In comparison, Decision Tree, Logistic Regression, and Naïve Bayes models achieved 70%, 78%, and 75% accuracy, respectively. These results were further validated using confusion matrices and bar graph comparisons, which highlighted the SVM model's superior predictive capability. The model demonstrated 100% correct classification of fresh fish in multiple test runs, confirming its reliability. Moreover, real-time prediction was successful when a test image was uploaded, where the model correctly identified whether the fish was fresh or infected. The high precision and zero false positives in SVM results make it highly applicable for practical aquaculture monitoring systems. The results validate that the integration of effective image pre-processing and the SVM classifier is optimal for automated fish disease detection.



Graph x-axis represents algorithm names and y-axis represents accuracy, precision and other metric in different colour bar.


Fish predicted as INFECTED FISH


Fish predicted as FRESH

## DISCUSSION

The results clearly indicate that SVM is the most effective classifier for fish disease detection among the tested models. The superior accuracy achieved with both augmented and raw image datasets demonstrates the robustness and generalization capability of the SVM model. By leveraging powerful image enhancement techniques like CLAHE and LAB conversion, the system successfully improves input quality, which directly enhances model performance. The study also proves the advantage of image augmentation in boosting classification accuracy. One key observation is that while traditional classifiers like Decision Tree and Logistic Regression offer a basic level of classification, they fall short in handling the complexity and variability of real fish images. On the other hand, SVM, with its kernel functions, adapts well to the non-linearity in the dataset. The real-time testing capability ensures that the model is not just limited to experimental environments but can also be deployed practically. These findings pave the way for extending the system with deep learning models like CNNs or integration into IoT-based fish monitoring setups. Overall, the project establishes a strong foundation for smart aquaculture applications.

## CONCLUSION

This study presents a reliable and accurate machine learning-based system for detecting fish diseases using image data. Through a combination of interpolation, CLAHE, LAB color transformation, and SVM classification, the proposed approach achieves outstanding accuracy—94.12% with augmentation and 91.42% without. Compared to traditional models like Decision Tree and Naïve Bayes, the SVM model showed superior performance in all evaluation metrics. The system's ability to classify fish images in real time confirms its practical usability in aquaculture farms. The automated nature of this tool reduces reliance on human inspection and promotes early detection, potentially minimizing economic losses in fish farming. Future work aims to enhance the system further using CNN architectures and integrate IoT capabilities for real-time disease alerts. Overall, this project provides a promising solution for modern aquaculture disease management using intelligent imaging and classification techniques.

## REFERENCES

[1] A. A. M. Abd Hamid, N. and A. Izani. Extended cubic b-spline interpolation method applied to linear two-point boundary value problem. World Academy of Science, 62, 2010.

[2] T. Acharya. Median computation-based integrated color interpolation and color space conversion methodology from 8-bit bayer pattern rgb color space to 24-bit cie xyz color space, 2002. US Patent 6,366,692.

[3] A. F. Agarap. An architecture combining convolutional neural network (cnn) and support vector machine (svm) for image classification. arXiv preprint arXiv:1712.03541, 2017.

[4] A. Ben-Hur and J. Weston. A user's guide to support vector machines. In Data mining techniques for the life sciences, pages 223– 239. Springer, 2010.

[5] S. Bianco, F. Gasparini, A. Russo, and R. Schettini. A new method for rgb to xyz transformation based on pattern search optimization. IEEE Transactions on Consumer Electronics, 53(3):1020–1028, 2007.

[6] E. Bisong. Google colaboratory. In Building Machine Learning and Deep Learning Models on Google Cloud Platform, pages 59–64. Springer, 2019.

[7] A. P. Bradley. The use of the area under the roc curve in the evaluation of machine learning algorithms. Pattern recognition, 30(7):1145– 1159, 1997.

[8] S. A. Burney and H. Tariq. K-means cluster analysis for image segmentation. International Journal of Computer Applications, 96(4), 2014.

[9] M. A. Chandra and S. Bedi. Survey on svm and their application in image classification. International Journal of Information Technology, pages 1–11, 2018.

[10] L. de Oliveira Martins, G. B. Junior, A. C. Silva, A. C. de Paiva, and M. Gattass. Detection of masses in digital mammograms using kmeans and support vector machine. ELCVIA Electronic Letters on Computer Vision and Image Analysis, 8(2):39–50, 2009.

[11] A. Gaur and S. Yadav. Handwritten hindi character recognition using k-means clustering and svm. In 2015 4th International Symposium on Emerging Trends and Technologies in Libraries and Information Services, pages 65–70. IEEE, 2015.

[12] M. grapple with Atlantic salmon price rollercoaster. Global aquaculture alliance. http://www.fao.org/familyfarming/detail/en/c/1263890/, 2020.

[13] J. A. Hartigan and M. A.Wong. Algorithm as 136: A k-means clustering algorithm. Journal of the royal statistical society. series c (applied statistics), 28(1):100–108, 1979.

[14] H. Hel-Or. Color conversion algorithms. http://cs.haifa.ac.il/hagit/courses/ist/Lectures/Demos/ColorApplet2/t_convert.html. [Last accessed: 2 Nov, 2020].

[15] M. S. Hitam, E. A. Awalludin, W. N. J. H. W. Yussof, and Z. Bachok. Mixture contrast limited adaptive histogram equalization for underwater image enhancement. In 2013 International

conference on computer applications technology (ICCAT), pages 1–5. IEEE, 2013.

[16] Infected. Detecting community structure using label propagation with weighted coherent neighborhood propinquity. https://www.documentcloud.org/documents/4559944-Groatay- Early-2018-2018-0111.html, pages 2018–0111, Accessed 10 Jan. 2020.

[17] T. Inoue, K. Hasegawa, M. Yanagisawa, and N. Togawa. Designing hardware trojans and their detection based on a svm-based approach. In 2017 IEEE 12th International Conference on ASIC (ASICON), pages 811–814. IEEE, 2017.

[18] C. Kailasanathan, R. S. Naini, et al. Image authentication surviving acceptable modifications using statistical measures and k-mean segmentation. IEEE-EURASIP Work. Nonlinear Sig. and Image Processing, 1, 2001.

[19] S. Khan, R. Ullah, A. Khan, N. Wahab, M. Bilal, and M. Ahmed. Analysis of dengue infection based on raman spectroscopy and support vector machine (svm). Biomedical optics express, 7(6):2249– 2256, 2016.

[20] C. Liu, X. Sui, X. Kuang, Y. Liu, G. Gu, and Q. Chen. Adaptive contrast enhancement for infrared images based on the neighborhood conditional histogram. Remote Sensing, 11(11):1381, 2019.