# Automated quality evaluation of fishes with image enhancement and machine learning techniques by uniquely constructed Datasets

**Bhaskar Nandi[1*], Trisha Bej[1], Soumit Chowdhury[2] , Yashraj Sharma[1] , Suvajit Majhi[1] , Kalyan Mahata[3] ,Anirban Goswam[4],Nabin Ghoshal[5]**

[1]Institute of Engineering and Management(IEM), University of Engineering and Management,Department of Basic Science & Humanities , Kolkata. 700160
[2]Department of Computer Science Engeneering, Jalpaiguri Government Engineering College , Maulana Abul Kalam Azad University of Technology :
[3]Dertment of IT,Government College of Engineering and Leather Technology, Maulana Abul Kalam Azad University of Technology, Kolkata,
[4]Department of IT,Techno India Main Saltlake,Maulana Abul Kalam Azad University of Technolog,Kolkata
[5]Department of Engg. & Technological Studies,University of kalyani,Kalyani,West Bengal 741235
E-mail :[1*] bhaskar.nandi@uem.edu.in

**ABSTRACT**

**Proper identification of freshwater fish species is the basis of aquaculture conservation, ecological studies, and management. This article reports some sophisticated image-processing procedures for generating a fine Bata fish (Labeo bata) dataset. Attempting to cover diverse states of the fish, the dataset was generated through systematic photography during morning, afternoon, and evening hours. Following background removal techniques, data augmentation techniques like rotation, flipping, and brightness adjustment were used to increase dataset diversity. Preprocessing techniques like noise removal, scaling, normalization, and format conversion were carried out to make all samples homogeneous. A CNN model was further incorporated using this updated dataset to achieve a decent accuracy of the existing work for classification of the different fish images. Hence this proposed approach will ensure consistency of the dataset for automatic fish classification and quality evaluation and plans to build a larger dataset of different fish images in the context of machine learning models for automatic fish quality categorization.**

**KEYWORDS:** ***Neural Networks,machine Learning, Image Feature Extraction ,Image Preprocessing,Food Quality Evaluation,Automated Image Analysis.***

## 1. INTRODUCTION :

Water taxonomy in fish is important for fishery management, conserving biodiversity, and ensuring food security.Traditional fish quality assessment relies on manual inspection, which is often subjective, time-consuming, and prone to human error.[2][5][8] To arrest these limitations, advanced computer vision techniques, combined with image enhancement and machine learning, are increasingly being utilized for accurate and automated quality assessment of fish. Labor-intensive inspection utilized in conventional techniques for identifying fish was slow as well as biased[5][3]. These advancements have improved the automated processes to perform better with higher power for classification and precision[9][10]. But the diversity and quality of the training dataset play an important role in determining how accurate such models will perform.

So in view of the practice approaches ,this proposed study captures systematic images and subsequently applying preprocessing techniques for data consistency and reliability, the present study aims to construct a appropriately framed dataset of Bata fish (Labeo bata).

## 2. PRESESENT EXISTING APPROACHES :

AI and machine learning are transforming food quality detection, ensuring safety through innovative and automated solutions. Studies have shown that deep learning models like CNN and MobileNet can accurately classify fish freshness using image analysis, with some achieving over 97% accuracy [1,2]. IoT



Paper ID: NCRTST25-JGEC-102-0010 & ISBN: 978-81-986641-7-4

sensors combined with AI have also been used to detect harmful chemicals like formaldehyde in seafood, making food safety monitoring more reliable [3]. Machine learning techniques have been successfully applied to detect food spoilage, with deep learning models like VGG16 and clustering algorithms proving highly effective [4]. A tool called Freshy Pal demonstrated how AI could identify bread spoilage with up to 98% accuracy [5]. In shrimp quality assessment, researchers improved detection using GANs for data augmentation, achieving over 98% accuracy in real-time applications [6]. Meat quality testing has also benefited from AI, with Res-net-based models and SVM classifiers showing impressive results, making quality control in poultry and meat industries more efficient [7,8,10]. Overall, AI-driven approaches are proving to be faster, more accurate, and less invasive than traditional methods, revolutionizing food safety and quality control on a global scale.

So in view of existing work, We have seen different researchers are working in different Fish databases for their research. Basically all researchers are using edible foreign fish species. But ,our dataset are based on local species of fishes which will take everyday life.

## 3. METHODOLOGY

The data was subjected to a series of preprocessing steps to enhance its quality and utility for future machine learning processes as shown in figure 1. The preprocessing pipeline, depicted in Figure below includes the following steps:


*Figure 1.Work flow diagram in Image Preprocessing Step*

•*Image Acquisition:* The conditions remained normal as images were taken at three time periods (morning, afternoon, and evening) and at each given time period, 5 catches of fish were caught and preserved accordingly along with their respective hedonic ratings for various time periods. Hedonic rating included colour, texture, and shape[2].

•*Background Removal:* Rembg library was utilized to remove the background and separate the fish in every image to have a uniform structure in the dataset[3][5].

•*Data Augmentation:* Rotation, flipping and brightness changes were utilized to change and make the dataset diverse. This operation is mathematically expressed as: I' = T ( I ) where I is the original image and T is the transformation function (rotation, flipping, brightness adjustment, etc.)[7,8].

•*Noise Removal:* For the sake of sanity of the dataset, noisy images with poor noise or defects were removed manually. A Gaussian filtering method [Web-1] was used for image denoising by the following:

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad \text{------------------------------(1)}$$

where σ is the standard deviation of the Gaussian function.

•*Resizing & Normalization:* The final images were normalized to one consistent pixel intensity value across the dataset and resized to a common dimension. The following equation was utilized to perform normalization[Web-2]:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad \text{--------------------------------------------(2)}$$



Paper ID: NCRTST25-JGEC-102-0010 & ISBN: 978-81-986641-7-4

where X represents original pixel value, Xmin and Xmax represent = minimum and maximum pixel value in the dataset and X' is the normalized pixel value.
•*Class labelling:* They were classified into three classes with the assistance of a hedonic rating scale with scores ranging from 1-10: Fresh, Edible, and Bad.

## 4. EXPERIMENTAL RESULTS AND DISCUSSIONS:

An evenly organized dataset with uniform image quality was achieved after preprocessing. The diversity of the dataset was enhanced by the augmentation process[3,5], and it became even more suitable for future use. Noise removal preserved the purity of each sample and prevented any potential biases in the classification models. The derived dataset now forms a sufficient basis for fish classification deep learning model training[7,9].

### 4.1.DATASET DESCRIPTION

This dataset contains 95 images of Bata fish captured at three occasions (morning, afternoon, and evening) over several days. The fish were captured under the same conditions which guaranteed minimal possible variation depending on extrinsic factors. We labeled the dataset into three quality classes: Fresh, Edible, and Bad, according to hedonic ratings. The image was labeled with its corresponding hedonic score by considering shape, texture and color.
The obtained graphs and values are portrayed below:
Average Skewness: 4.0252
Average Kurtosis: 37.9922
Number of outliners detected: 334
Variance Distribution Across Features: Shows how pixel values vary across images, helping identify stable vs. Dynamic features in Figure 2.

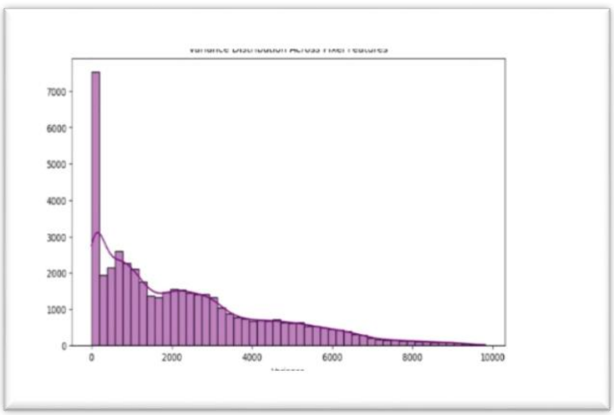


*Figure 2:Variance Distribution Across Features*

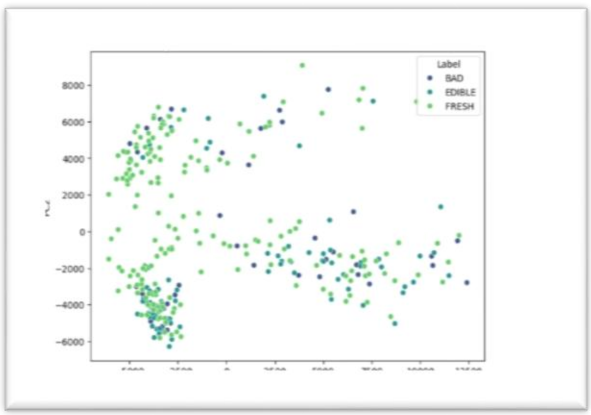


*Figure 3:PCA Projection of Images*



Paper ID: NCRTST25-JGEC-102-0010 & ISBN: 978-81-986641-7-4

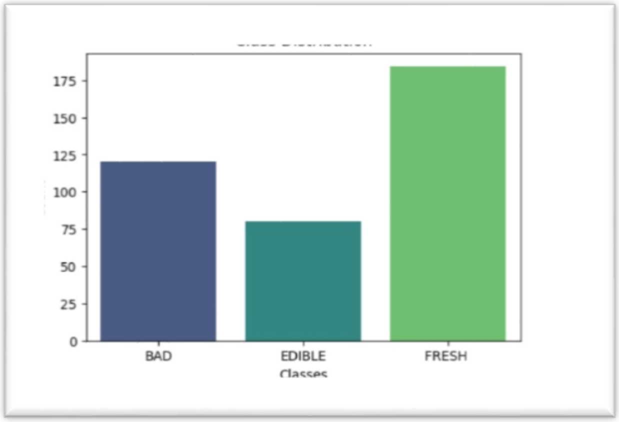


*Figure 4:Class Distribution*

- PCA Projection of Images: Reduces dimensionality [5]while preserving essential patterns, making visualization and classification easier in Figure 3.
- Class Distribution: Displays the count of images per class, ensuring dataset balance[7] for unbiased model training in Figure 4.

The provided table gives a detailed description about the Pixel and MI score of the dataset samples.

*Table 1:Pixel and MI score*

| Pixel | MI Score |
|---|---|
| 22903 | 0.35533 |
| 21041 | 0.34807 |
| 22148 | 0.347947 |
| 22153 | 0.347458 |
| 21401 | 0.344135 |
| 20636 | 0.340529 |
| 21013 | 0.340496 |
| 23267 | 0.338714 |
| 25863 | 0.337083 |
| 23997 | 0.336405 |

## 5. CONCLUSION

This presented study facilitate consistency and reliability for applications in any future computer-aided fish classification,The major contribution of this study are as follows.

1. Implementation of sophisticated noise reduction methods, data augmentation, background subtraction, and normalization for obtaining an optimal dataset of bata fish dataset.

2. Image dimensionality optimization with PCA method for fomatting the proposed dataset.

3. Analyze pixel distribution for different fish images to evaluate the fish quality by variance distribuition method related to pixel.

The present work is an important stepping block towards the application of computer vision for fish classification and quality inspection. Model training will be the area of research in the future. CNN-based classification models will automatically detect Quality Bata fish. Detection of freshness can be made more



Paper ID: NCRTST25-JGEC-102-0010 & ISBN: 978-81-986641-7-4

precise by using biochemical and environmental parameters. Incorporating more datasets with other fish species should make generalization of the model to large fisheries even more error-free.

## 6. ACKNOWLEDGEMENTS

We would like to express our sincere gratitude to IEM-UEM Group for providing the necessary resources and funding that made this research possible. We are also thankful to our Director and few feculty number for their constructive discussions and technical support.

Finally, we extend our heartfelt appreciation to our families and friends for their encouragement and unwavering support during this research journey.

## REFERENCES

[1] Roikhanah I., Harsono T., and Yuniarti H. (2021). "Formalin Fish Detection System Based on Digital Image Processing," in *Proc. 2021 Int. Electron. Symp. (IES)*, Surabaya, Indonesia, 2021, pp. 362–367, doi: 10.1109/IES53407.2021.9593972.

[2] . Karmaker R., Tinky M. A., and Akter S. (2024). "Deep Learning and Image Processing Approach for Fish Freshness Detection," in *Proc. 2024 IEEE Int. Conf. Power, Electr., Electron. Ind. Appl. (PEEIACON)*, Rajshahi, Bangladesh, 2024, pp. 1–6, doi: 10.1109/PEEIACON63629.2024.10800710.

[3] H. S., K. S., Suntharavadivelan V. M., and Prabhavathy P. (2024). "FinSecure: Utilizing IoT Sensors for Formaldehyde Detection and Fish Freshness Detection for Enhancing Safety in Fish Consumption Using Machine Learning and Deep Learning," in *Proc. 2024 Int. Conf. Signal Process., Comput., Electron., Power Telecommun. (IConSCEPT)*, Karaikal, India, 2024, pp. 1–8, doi: 10.1109/IConSCEPT61884.2024.10627800..

[4] Megalingam R. K., Sree G. S., Reddy G. M., Sri Krishna I. R., and Suriya L. U. (2019). "Food Spoilage Detection Using Convolutional Neural Networks and K-Means Clustering," in *Proc. 2019 3rd Int. Conf. Recent Dev. Control, Autom. Power Eng. (RDCAPE)*, Noida, India, 2019, pp. 488–493, doi: 10.1109/RDCAPE47089.2019.8979114.

[5] Pitroda H. and Kolhe A. (2023). "Freshy Pal: A Proposal of an Interactive Web-based Tool to Automatically Detect Microbial Growth on White Bread," in *Proc. 2023 IEEE 8th Int. Conf. Converg. Technol. (I2CT)*, Lonavla, India, 2023, pp. 1–8, doi: 10.1109/I2CT57861.2023.10126433.

[6] Prema K. and Visumathi J. (2022). "An Improved Non-Destructive Shrimp Freshness Detection Method Based on Hybrid CNN and SVM with GAN Augmentation," in *Proc. 2022 Int. Conf. Adv. Comput., Commun. Appl. Inform. (ACCAI)*, Chennai, India, 2022, pp. 1–7, doi: 10.1109/ACCAI53970.2022.9752599.

[7] Antony M. A. and Kumar R. S. (2021). "A Comparative Study on Predicting Food Quality using Machine Learning Techniques," in *Proc. 2021 7th Int. Conf. Adv. Comput. Commun. Syst. (ICACCS)*, Coimbatore, India, 2021, pp. 1771–1776, doi: 10.1109/ICACCS51430.2021.9441743.

[8] Liu C., Wen H., Liu X., Sun S., and Yang L. (2021). "Development of Portable Device for Non-Destructive Testing of Livestock Meat Quality based on Machine Vision," in *Proc. 2021 Int. Conf. Netw., Commun. Inf. Technol. (NetCIT)*, Manchester, United Kingdom, 2021, pp. 336–342, doi: 10.1109/NetCIT54147.2021.00074.

[9] B. N. B. J., S. P. S., and K. N. R. (2024). "Beef Quality Assessment using InceptionNet," in *Proc. 2024 Int. Conf. Invent. Comput. Technol. (ICICT)*, Lalitpur, Nepal, 2024, pp. 995–1000, doi: 10.1109/ICICT60155.2024.10544487.

[10] Gunasekara J., Rathnayake S., and Panduwawala P. (2023). "Quality Level Identification of Stored Chicken Using Machine Learning Algorithms," in *Proc. 2023 5th Int. Conf. Adv. Comput. (ICAC)*, Colombo, Sri Lanka, 2023, pp. 567–572, doi: 10.1109/ICAC60630.2023.10417269.

[11] https://images.app.goo.gl/V4zLNeNeMq1Ap6yJ9

[12] https://images.app.goo.gl/DbUXZ4FH7V87zpQ29



Paper ID: NCRTST25-JGEC-102-0010 & ISBN: 978-81-986641-7-4