

Fish Disease Detection System: A Case Study of Freshwater Fishes of Bangladesh

Juel Sikder¹, Kamrul Islam Sarek², Utpol Kanti Das³
Dept. of Computer Science and Engineering
Rangamati Science and Technology University
Rangamati, Bangladesh

Abstract—The proposed system is designed for automatic detection and classification of fish diseases in freshwater especially Rangamati Kaptai Lake and Sunamganj Hoar area of Bangladesh. Our experimental result is indicating that the proposed approach is significantly an accurate and automatic detection and recognition of fish diseases. This study presents fish disease detection based on the K-means and C-means fuzzy logic clustering method to segment the filtering image. Gabor's Filters and Gray Level Co-occurrence Matrix (GLCM) are used to extract the features from the segmented regions. Finally Multi-Support Vector Machine (M-SVMs) is used for classification of the test image. The proposed system demonstrated a comparison between K-means clustering and C-means fuzzy logic. The proposed methodology gave 96.48% accuracy using K-means and 97.90% using C-means fuzzy logic which is the highest accuracy rate to compare other existing methods. The proposed system has been experimented in the MATLAB environment on infected fish images of Rangamati Kaptai Lake and Sumangan Hoar area. It is a challenging task of fisheries farming in Hoar areas and Lake areas to detect fish diseases initially. The proposed methodology can detect and classify different fish diseases in early stages and also contributes to improved results for fish disease detection.

Keywords—K-means; c-means fuzzy logic; multi-SVM; detection

I. INTRODUCTION

Fisheries play a vital role in social and economic development and deserve the potential for future development in Bangladesh's agricultural economy. This sector represents one of the most prolific and energetic sectors in Bangladesh. It subsidizes 3.65% of the national GDP and approximately one-fourth (23.81%) of the GDP. Bangladesh is considered as one of the most suitable countries in the world for freshwater rural fisheries because of its resources and agro-climatic condition [1]. Fisheries and aquaculture paid 25% of agricultural GDP and 4.4% of national GDP in 2012 [1]. Fish biodiversity of the Rangamati Kaptai Lake and the Sunamganj Hoar area is decreasing day by day, due to territory deprivation and different man-made causes. It also worsens the normal health condition of fish and causes death of the fishes which results in great economic loss. Early detection of infected fish is a necessary step in preventing the spread of the disease. This paper recognizes and identifies the different fish diseases of Rangamati Kaptai Lake and Sunamganj Hoar which impacts the control of disease and maintenance of a strong relationship between living creatures, and the environment of Rangamati and Sunamganj. The prosperity of a fish culture depends on its precise regulation, fish life time, and cultural environment. However, unexpected use of feed and fertilizer are the main

reasons of failures in fish production. Traditional and semi-intensive culture systems are also responsible for the fish disease. Therefore, the search for a more efficient, less expensive, and more accurate method of diagnosing fish disease cases is very important [2]. In this research, we conducted the identification of local freshwater fish using a software based approach. This research also approaches a method for different fish disease detection. Firstly, the pre-processing stage applied on the test image. In the pre-processing stage, the test image is resized [3] then RGB images converted into HIS. The noise filtering technique is also used to remove noise. In the second stage, K-means and C-means fuzzy logic clustering method applied to segment the filtering image. In the third stage, the system extracts the features from the segmented and infected regions using Gabor's Filters and GLCM. In the fourth stage, classification of the test image is done by Multi-SVMs algorithm. The proposed system is tested using different types of fish diseases, especially Tail and Fin Rot, EUS, Red spot, Bacterial gill rot, Parasitic Disease (Argulus), Broken Antennae, and Rostrum. Finally, fish diseases are detected and recognized that can be prescribed in its early stages [4].

This research is designed in six sections. Section II represents some related works and Section III describes the data collection process. Section IV gives a detailed description of the proposed methodology. Section V briefing the experimental results and discussing about the mentioned method and outcomes of the system. Section VI presents conclusions and the opportunity of probable coming work in this area.

II. RELATED WORKS

Some related works associated with the proposed method described here. Waleed, Ahmed, et al. [5] completed a survey on automatic recognition of different fish diseases. They proposed three several categories of fish diseases named Columnaries, Ichthyophthirius (Ich) and Epizootic ulcerative syndrome (EUS) to recognize and identify automatically. They show the effect of several color spaces on CNN ultimate execution using Alexnet. Chacon, Mario I., Luis Aguilar, and Abdi Delgado [6] proposed a method related to an optional defuzzification function, fuzzy operators and an image fuzzification function. The Contextuality of the system is described in three phases, geometric measurements, edge detection and image binarization. On binarization the fuzzy method describes the way to solve indeterminacy on the classification of pixels. The fuzzy edge detection showed how a particular perception could be demonstrated by a fuzzy system. Lastly, fuzzy geometry is an illustration of how to obtain a variable state

below of lighting to gather more consistent measurements. Mohanaiah, P., P. Sathyanarayana, and L. GuruKumar [7] proposed a method that applies GLCM to extract second order statistical texture features. The Four features namely, Inverse Difference Moment, Correlation, Angular Second Moment and Entropy are computed using Xilinx FPGA. The image compression time can be diminished significantly by extracting the features of an image. These features are important in the motion approximation of videos and in real time pattern identification applications. Li, Wei, and Qian Du [8] proposed a system that showed the importance of extracted features using a simple Gabor filter for the NRS classifier. The accuracy of the experimented classifier results compared to conventional pixel-wise classifiers. Specifically, to extract spatial features in the PCA projected domain or a subset of original bands with band selection, a simple two-dimensional Gabor filter was implemented. Traditional classifiers, such as SVM, RLDA and SRC compared with applied classification techniques, i.e., PC-Gabor-NRS and BS-Gabor-NRS. Authors claimed that Experimental results can outperform other classification techniques. Lyubchenko, Valentin, et al. [9] proposed a color based image segmentation by determining the infected areas so that it can detect and diagnose fish diseases. To detect infected and healthy areas of the fish body, they used digital image processing and colored markers. Malik, Shaveta, Tapas Kumar, and A. K. Sahoo [4] proposed a method to segment and classify EUS and Non-EUS infected fish diseases. To enhance the image, they applied segmentation. Again, they collect useful information using different edge detection methods and also applied morphological operation. They used various feature descriptors to extract features to classify the fish disease using PCA. Sankar, M. Muni, et al. [10] proposed an approach to investigate diseased fish using image processing techniques. They applied capturing image sensing techniques for image acquisition and k-means clustering method. They also used edge detection techniques to detect edges and K-means clustering algorithm to cluster the input prawn image and detect diseases.

III. DATA COLLECTION

During the study of the historical 20 endangered fish species, IUCN has been initiating in Kaptai Lake. Among 20 species 16 were found to be available, 4 rarely available, and 34 were not available during the study period. Accessibility of three categories of disapprovingly endangered, threatened, and helpless fishes from Kaptai Lake are presented in considering exposed species, 30% were found accessible, 7% rarely available and 63% were not available during the study period. Available fish landing hubs and marketplaces would be helpful to provide primary knowledge of fish fauna, mostly to know about the fish fauna in the biggest artificial Kaptai Lake, Rangamati, Bangladesh.

In this study, Kaptai reservoir and fishery Ghat in Rangamati and Sunamganj Haor were selected to collect the different types of freshwater fishes. The data were also collected through the interview with fishermen, associated persons and cross-check interviews with informers. After that, the collected data were cross-checked with the landing data from Bangladesh Fisheries Development Corporation fish landing center especially Rangamati and Sunamganj, Bangladesh. Several large diverse fisheries are found in Rangamati town and Sunamganj

Community Based Resource Management Project (CBRMP). Different types of diseases and symptoms are described in Table I. After collecting different types of fishes from Rangamati Kaptai Lake and Sunamganj Haor, we created a database which was uploaded to kaggle.com [11]. We used different types of fish diseases and symptoms of the diseases in our database. In our study, we also used 126 fish species from 39 families initiated in the Sunamganj Haor area [9].

TABLE I. AVAILABILITY OF FRESHWATER FISH SPECIES KAPTAL LAKE AND SUNAMGANJ HAOR

Family	Lacal name	Common name	Scientific name
Schilbeidae	Bacha	Bacha	Eutropiichthys Vacha
Siluridae	Kani Pabda	pabo catfish	Ompok Bimaculatus
Cyprinidae	Mrigel, Mirka	Caurvery white carp	Cirrhinus Cirrhosis
Nandidae	Bheda	Mud perch	Nandus Nandus
Gobiidae	Bele	Tank goby	Glossogobius Giuris
Belonidae	Kaikka	Needle fish	Xenentodon Cancila
Sciaenidae	Poa	Pama	Otolithoides Pama
Clariidae	Magur	Air breathing catfish	Clarias Batrachus
Pangasidae	Pangas	Yellowtail catfish	Pangasius Pangasius
Cyprinidae	Rui	Rohu	Labeo Rohita
Notopterid	Chital	Humped featherback	Notopterus Chitala

IV. PROPOSED METHODOLOGY

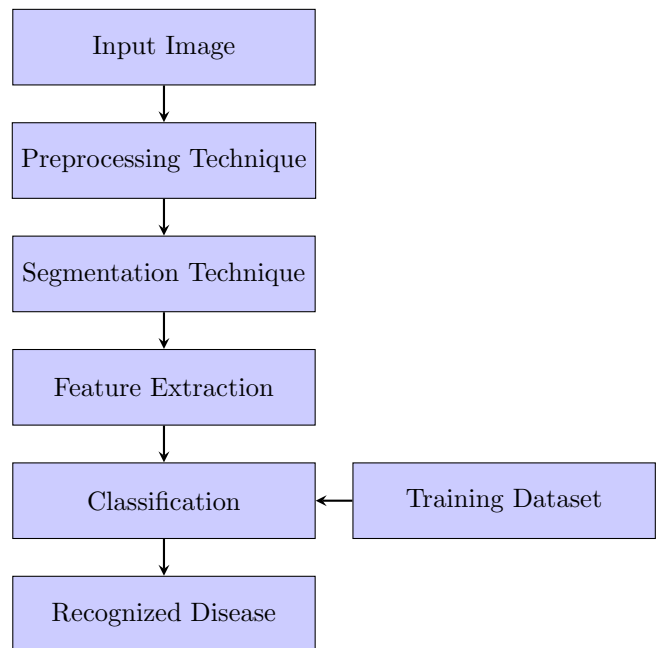


Fig. 1. Block Diagram of the Developed System.

“Fig. 1” describes the proposed block diagram.

A. Input Image

The system required a color image to start the procedure. An affected fish image passed as an input to the system.

B. Pre-processing

Noise Filtering is applied to remove several forms of noises and to filter the redundant information from the images [12]. Contrast Stretching is compulsory when the system needs to intensify or diminish the total number of pixels. Pre-processed results of the input images are shown in “Fig. 2”.

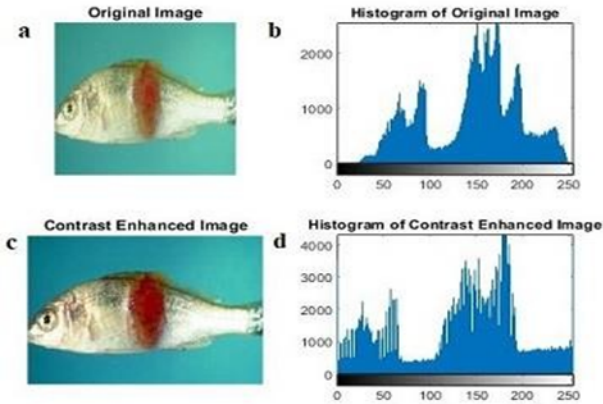


Fig. 2. (a) Input of the EUS Affected Original Fish Image (b) Histogram Result (c) Enhanced Result and (d) Histogram of Enhanced Image.

C. Image Segmentation

Fuzzy C-means (FCM) is a clustering method used to separate data into two or more clusters. We used FCM in our system. Firstly, it picks two initial cluster values from the input image, $I (r \times c)$. The system accesses every pixel as $I (i, j)$, where

$$0 \leq j < c \quad (1)$$

and

$$0 \leq i < r \quad (2)$$

Secondly, the average of median values of each row is calculated. Thirdly, the distance between every pixel of the image and cluster values calculated. Again, the distance between the pixels of the image and averaged medians are measured also. Fourthly, the final fuzzy field smooths the data obtained from cluster values and from averaged medians. Fifthly, the fuzzy rule is applied to segment the image between two clusters. Sixthly, the segmented region is reconstructed by updating cluster centers. Seventhly, steps third to sixth repeated for t times. The whole process is iteratively performed until the difference between two consecutive clusters reaches to a constant position and becomes minimum [13]. Finally, the segmented regions of both clusters for a test image are found.

K-means clustering is an unsupervised learning algorithm that is used enormously. The method is used to define k samples, one for each cluster [14]. The next step is to assign each point of input data into the nearest cluster. After selecting each data point, the first step is completed by re-calculating the new cluster centroid for the k -clusters. An iterative process has been generated until the new centroid reaches the constant state. As a result, the data points separate into k clusters step by step.

D. Feature Extraction

After segmentation, when the area of the interest region is detected then the significant features are extracted. GLCM and Gabor Filter are used to extract feature vectors from input fish disease image like texture. Texture element is extracted from the RGB colored image. The GLCM functions represent texture features of an image. Textures are the features whose focus on the distinct pixels that create an image. Texture features such as statistical feature is used which contains GLCM, grey level histogram, run length matrices and autocorrelation features for texture extraction. GLCM procedures extract 2nd-order statistical texture features [7]. Textural features include entropy, correlation, contrast, homogeneity and energy. The texture is particularly appropriate for this kind of study because of its possessions. The system also used Gabor filter features. The Gabor filters consist of parameters such as standard deviation, orientation and the radial center frequency [8]. The system combined the GLCM feature and the Gabor Filters into one feature set. This combination of the Gabor filters and the GLCM feature generate a better outcome on the fish texture dataset.

E. Classification Using Multi-SVMs

The proposed system used multi-SVMs to classify fish diseases. The extracted features of each fish disease image are stocked in the database. The SVM classifier will calculate the feature number of database images and the feature value of the input image, based on these values the classifier will separate the input image from that category. SVM will compare the test sample feature set to all training samples and select the shortest distance. Support vector machines were initially calculated for binary classification. Several approaches have been proposed where typically the system constructs a multi-class classifier by combining several binary classifiers. The SVM is learned by features given as an input to its training procedure. During training, the SVM identifies the appropriate boundaries in the 2-categories. Features are called according to class associative to a specific class. Using received features, the multi-SVM starts classification steps. The system uses multi-SVM as a two class SVM repeatedly. When the test image is passed to the classifier, it classifies it from the first two classes. After that classified class and a new class used to classify the test image. Similarly, it continued its classification steps until all categories were tested. From this last classification, the test image is classified as the true categorized class [14].

V. RESULT AND DISCUSSION

The proposed method is examined in MATLAB R2018a. The experimented database was uploaded to kaggle.com [11]. The database contains fish disease images of freshwater of Rangamati Kaptai Lake and Sunamganj Hoar area of Bangladesh. The system experimented using different fish diseases such as EUS, Red Spot, Argulus, Tail and Fin Rot, Broken antennae rostrum and Bacterial Gill rot. The system divided the input image into three clusters in both C-means fuzzy logic and K-means. All clusters of both C-means fuzzy logic and K-means clustering are given in “Fig. 3”.

From all of the clusters, one cluster chosen as the segmented cluster based the segmented region. All cluster regions

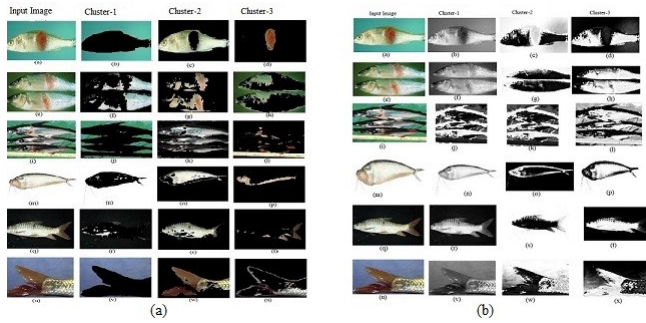


Fig. 3. a. C-means Clusters and b. K-means Clusters.

are calculated and the best cluster region chosen by user input. The segmented output is given in Fig. 4. In “Fig. 4”, the left side portion i.e., (a) represents the segmented image using C-means and the right side portion i.e., (b) represents the segmented images using K-means method.

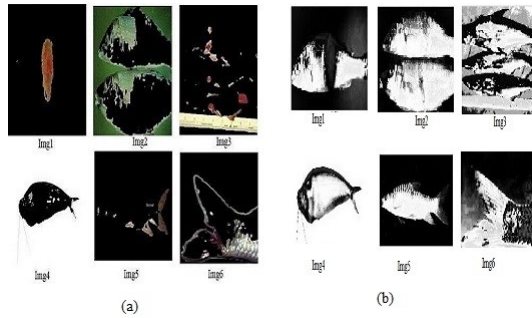


Fig. 4. Segmented Output (a. C-means and b. K-means).

TABLE II. GLCM FEATURE EXTRACTED FROM THE FISH IMAGES

Feature Name	Img1	Img2	Img3	Img4	Img5	Img6
Contrast	0.0473	0.0002	0.0001	0.0000	0.0001	0.0000
Correlation	0.0569	0.0002	0.0002	0.0001	0.0001	0.0001
Energy	0.0280	0.0001	0.0002	0.0000	0.0000	0.0000
Homogeneity	0.0577	0.0002	0.0002	0.0001	0.0001	0.0001
Mean	0.1127	0.0004	0.0005	0.0001	0.0002	0.0001
Standard Deviation	0.9833	0.0010	0.0010	0.0010	0.0010	0.0010
Entropy	0.9800	0.0010	0.0010	0.0010	0.0010	0.0010
RMS	0.9901	0.0010	0.0010	0.0010	0.0010	0.0010
Variance	0.9797	0.0010	0.0010	0.0009	0.0010	0.0010
smoothness	0.9604	0.0009	0.0009	0.0007	0.0010	0.0010
Kurtosis	0.2382	0.0001	0.0001	0.0007	0.0005	0.0002
Skewness	0.2339	0.0001	0.0001	0.0007	0.0005	0.0002
IDM	3.0432	0.0017	0.0023	0.0132	0.0042	0.0045

Table II illustrates the extracted feature of segmented images. The feature extracted using GLCM feature extractor. Combining those GLCM features and Gabor filtered features, passed to the classifier to create the desired classifier.

The results of the Multi-SVM classifier are shown in Table III. The performance of the classifier is calculated by the different procedures and The SVM classifier calculates the feature value of the database and of the test image. The SVM classifier uses that feature values to identify the test image class. Input feature is (1.0e+03) and maximum accuracy of EUS and Broken antennae and rostrum is (98.4127), the

TABLE III. MULTI-SVM CLASSIFIER RESULTS FOR THE FISH DISEASE IMAGES

S/N	Fish Image	Multi-SVM Classifier	Input feature	Maximum Accuracy	Row
1	Img1	EUS	1.0e+03	98.4127	500 iterations
2	Img2	Red spot	1.0e+03	95.2381	500 iterations
3	Img3	Broken antennae and rostrum	1.0e+03	98.4127	500 iterations
4	Img4	Argulus	1.0e+03	95.2381	500 iterations
5	Img5	Bacterial Gill rot	1.0e+03	96.8254	500 iterations
6	Img6	Tail and fin rot	1.0e+03	95.2381	500 iterations

accuracy of Linear Kernel is the highest 500 iteration.

TABLE IV. A COMPARISON BETWEEN K-MEANS CLUSTERING AND C-MEANS FUZZY LOGIC

S/N	Image	Fish disease	Accuracy of K-means	Number of Iteration	Accuracy of C-means
1	Img1	EUS	98.4127	500 iterations	99.25
2	Img2	Red Spot	95.2381	500 iterations	96.50
3	Img3	Broken antennae	98.4127	500 iterations	99.45
4	Img4	Argulus	95.2575	500 iterations	97.23
5	Img5	Bacterial Gill rot	96.8254	500 iterations	98.70
6	Img6	Tail and fin rot	94.7450	500 iterations	96.25

Table IV represents the proportional outcomes of the performance of between K-means clustering and C-means fuzzy logic segmentation. The system gives 96.48% accuracy using C-means fuzzy logic where K-means gives 97.90% accuracy.

VI. CONCLUSION

The paper proposed an image processing based approach and automatically detects different types of fish diseases. The testing of the proposed system has been experienced on the actual infected fish images collected from Rangamati Kaptai Lake and Sunamganj Hoar, Bangladesh. The experimental results indicate that the proposed approach is a valuable approach and can significantly support accurate and automatic detection of fish disease and the proposed combination gave better accuracy after applying the machine learning algorithm. The infected fish images were then classified with an average accuracy rate of K-means clustering and C-means fuzzy logic gives 96.48% and 97.90% respectively. In the future, it will present some guidelines such as establishing a better segmentation system, choosing improved classification algorithms and deep learning algorithms that can apply on different feature descriptors.

REFERENCES

- [1] FRSS. "Yearbook of Fisheries Statistics of Bangladesh." (2017).
- [2] Austin, Brian. "Methods for the diagnosis of bacterial fish diseases." Marine Life Science & Technology 1, no. 1 (2019): 41-49.
- [3] Juel Sikder, Utpol Kanti Das and Rana Jyoti Chakma, "Supervised Learning-based Cancer Detection" International Journal of Advanced Computer Science and Applications (IJACSA), 12(5), 2021. <http://dx.doi.org/10.14569/IJACSA.2021.01205101>
- [4] Malik, Shaveta, Tapas Kumar, and A. K. Sahoo. "Image processing techniques for identification of fish disease." In 2017 IEEE 2nd International Conference on Signal and Image Processing (ICSIP), pp. 55-59. IEEE, 2017.
- [5] Waleed, Ahmed, Hadeer Medhat, Mariam Esmail, Kareem Osama, Radwa Samy, and Taraggy M. Ghanim. "Automatic Recognition of Fish Diseases in Fish Farms." In 2019 14th International Conference on Computer Engineering and Systems (ICCES), pp. 201-206. IEEE, 2019.
- [6] Chacon, Mario I., Luis Aguilar, and Abdi Delgado. "Definition and applications of a fuzzy image processing scheme." In Proceedings of 2002 IEEE 10th Digital Signal Processing Workshop, 2002 and the 2nd Signal Processing Education Workshop., pp. 102-107. IEEE, 2002.

- [7] Mohanaiah, P., P. Sathyanarayana, and L. GuruKumar. "Image texture feature extraction using GLCM approach." *International journal of scientific and research publications* 3, no. 5 (2013): 1-5.
- [8] Li, Wei, and Qian Du. "Gabor filtering based nearest regularized subspace for hyperspectral image classification." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7, no. 4 (2014): 1012-1022.
- [9] Lyubchenko, Valentin, Rami Matarneh, Oleg Kobylin, and Vyacheslav Lyashenko. "Digital image processing techniques for detection and diagnosis of fish diseases." (2016).
- [10] Sankar, M. Muni, CH Nageswar Rao, G. Sailaja, N. Bhuvaneshwary, and P. Gunasekhar. "White spot syndrome virus detection in shrimp images using image segmentation techniques." *International Journal of Advanced Research in Computer Science and Software Engineering* 3, no. 9 (2013).
[11] <https://www.kaggle.com/Futpoldas/freshwaterfishdisease>.
- [12] Mahmud, Tanjim, Juel Sikder, Rana Jyoti Chakma, and Jannat Fardoush. "Fabric Defect Detection System." In *International Conference on Intelligent Computing & Optimization*, pp. 788-800. Springer, Cham, 2020.
- [13] Christ, MC Jobin, and R. M. S. Parvathi. "Fuzzy c-means algorithm for medical image segmentation." In *2011 3rd International Conference on Electronics Computer Technology*, vol. 4, pp. 33-36. IEEE, 2011.
- [14] Sikder, Juel, Utpol Kanti Das, and AM Shahed Anwar. "Cancer Cell Segmentation Based on Unsupervised Clustering and Deep Learning." In *International Conference on Intelligent Computing & Optimization*, pp. 607-620. Springer, Cham, 2020.