JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

AN AUTOMATED FISH CLASSIFICATION SYSTEM USING CNN

[1]Mr. P. Darwin, [2]Pedagandham N Sai Siri Srujana, [3]Danduprolu. Devi SriHarika, [4]Mohammad Raheem Sharief

[1]Associate Professor, Dept of ECE, Godavari Institute of Engineering and Technology(A), Rajahmundry, AP
[2,3,4]Students, Dept of ECE, Godavari Institute of Engineering and Technology(A), Rajahmundry, AP

***Abstract:* —** In aquatic ecosystems and fishery management, precise species identification is all that counts. However, the existing Support Vector Machine (SVM) based models have several limitations such as slow processing speed and low accuracy in dealing with complex data. Additionally existing fish classifiers using Alex Net, tend to have an accuracy of 95.5%. In order to overcome these challenges and limitations, we propose a new fish classification system built on deep learning and the CNN (Convolutional Neural Network) model resulting in a significant boost in accuracy. As a result, this system self-explains complex themes and structures from fish imagery thus increasing classification precision. Also, our system measures accuracy aspects including precision, recall and F1 score which gives a high reliability of results. It allows to improve the fish species identification, avoid misclassifications and prove data accuracy. The Automated Fish Species Identification system employs advanced technologies such as deep learning and CNN to analyse data mainly presented through images. The species of fish can be detected quickly and accurately, as a result of this many activities including those of monitoring fish populations, evaluation of ecosystem health assessment, sustainable fisheries management and conservation of aquaculture biodiversity etc. can depend upon having these technologies embedded in their systems. By using the automatic identification process, our system is fast and economical in terms of saving time and resources while providing accurate data that can be used for sound decision-making concerning aquatic environment.

***IndexTerms* - CNN, Key metrics, SVM model, Marine Research.**

I. INTRODUCTION

Fishes are the most crucial part of aquatic ecosystem. Most importantly fishes serve as an important indicator of ecosystem health, especially aquatic ecosystem. The fish species population can serve as an important indicator of the environmental disturbance, such as climate change, pollution, i.e. land or water pollution etc. Moreover, observing the fish species is very much necessary as researchers can determine the catch limits for species that are endangered, by the over-exploitation of certain species can be controlled. By knowing which fish species lives where we can build strategies to protect these species and their habitat. In the recent times there has been an increase in the demand for fish species classification[3][6-7][10][13-15]systems for various purposes, including the ones mentioned above. The fish classification systems previously proposed, utilized several technologies including, Squeeze-and-Excitation Networks [5], Object Recognition Framework [12], Pretrained AlexNet [14][18] model and improved AlexNet [9][13] model etc. Specifically, the improved AlexNet [14][18] system had at training accuracy of 94%, but due to certain drawbacks that include overfitting risks if there is limited data also AlexNet [14][18] demands for high computational resources, limiting its implementation feasibility. Here, Convolutional Neural Network [7][10][13-15][18] (CNN) being a deep learning network, that is popularly used as an image classifier has been used to [11] improve the accuracy [6] of the proposed [3][10][12-13][15] system to 98.64%. With the help of CNN, the model extracts the necessary feature from the input image, mapping them to respective class and accurately classifying the image. The theoretical background related to the proposed model is discussed in detail in section 2. The species present in the dataset are mentioned in section 3 with the help of fig.1. Section 4 explains [14] the architecture of the proposed [10][18] system [16]. Section 5 talks about the results obtained and section 6 and 7 talk about the future scope of the project and conclusion respectively.

1.1 OBJECTIVES:

Our project aims to achieve two key objectives

1. effectively deploying a deep learning algorithm [10]
2. optimizing the overall performance of classification algorithms by improving its Accuracy.

II. THEORETICAL BACKGROUND:

2.1 CNN LAYERS:

1) **CONVOLUTIONAL 2D LAYER:** This layer performs the convolutional operation on the input image using a set of learnable filters to extract features [13][18] from the image. The convolutional layer generates feature maps from the images. This feature map is two dimensional. Moreover, the number of feature maps are always equal to the number of Convolutional filters.

2) **ZERO PADDING:** In the matrix form of the image, zeros are added at the edges of input image. This helps in covering the entire input image while extracting the features using convolutional filters. Moreover, the output generated, i.e. the feature map [3][14] will have the same size as input image.
3) **BATCH NORMALIZATION:** The network architecture of CNN has Batch Normalization [14] Layer. This layer is used for stabilizing the network, helps in speeding up the training process [7][17-18] and uses high learning rates. Batch Normalization [14] Layer is used between Convolutional and ReLU [8] layer.
4) **RELU LAYER:** This activation layer performs a threshold operation on the feature map [3][6][14] received. In ReLU [8] layer the feature map [3][6][14] values are set to either zero or input value. If the input value is less than zero, then the activation layer sets the value to zero, similarly. if the input value is more than zero, the layer returns the input value.

$$f(x) = \begin{cases} 0, & x \leq 0 \\ x, & x > 0 \end{cases} \quad (1)$$

5) **MAX POOLING:** Pooling in general helps in reducing the size of input image. Here the focus is specifically on max pooling, it is important to note that stationary convolutional filter is used in this process and convolutional area don't overlap. In max pooling [18] the largest value of each convolutional layer is taken into consideration. Max pooling [18] helps in tacking the overfitting [3][7][18] problem.
6) **FULLY CONNECTED LAYER:** This is a layer that is placed just before the output layer in the CNN architecture. This layer is also known as dense or linear layer. In simple terms the fully connected [3][6][10][13-14][18] layer is known to combine all the features extracted in the previous layers [2].
7) **SOFTMAX LAYER & CLASSIFICATION LAYER:** SoftMax layer is used for holding the probability of a class in decimal values. Here it is important to note that the probabilities of each class when summed up gives the value one. In other words, the SoftMax Layer converts real values to probability. Moreover, the Classification layer compares the predicted probabilities with actual labels.

2.2. PERFORMANCE METRICS: True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN) play a crucial role with regards to Accuracy:

- True Positives (TP): This signifies the number of samples correctly identified as positive by the classifier [9].
- False Positives (FP): This represents the count of negative samples that were erroneously classified as positive by the classifier [9].
- True Negatives (TN): This denotes the number of samples accurately classified as negative by the classifier [9].
- False Negatives (FN): This indicates the instances where positive samples were inaccurately classified as negative by the classifier [9].

1) **ACCURACY*:*** Accuracy is an evaluation metric, usually used to [11] show how the model behaves over all classes. It is useful if every class has an equal stake. It is computed as the number of correct predictions divided by the total number of predictions [10].

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (2)$$

2) **PRECISION***:* Precision is computed as the quotient of Positive samples correctly classified to the total number of samples classified as Positive (regardless of correctness or incorrectness). The precision assesses the model's ability to correctly classify a sample as positive.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

3) **RECALL***:* Recall is the fraction of Positive samples correctly classified as Positive in the total number of Positive samples. The recall measures the model's ability to find Positives. The more is the recall, more true positives are detected.

$$\text{Recall} = \frac{TP}{TP+FP} \quad (4)$$

4) **F1 SCORE***:* F1 score is a term that involves both precision and Recall [3][10]. A good F1 score indicates lesser false positive values and low false negative values. The perfect F1 score is considered to be one, which is an ideal condition.

$$\text{F1 Score} = \frac{2*\text{Precision}*\text{Recall}}{\text{Precision}+\text{Recall}} \quad (5)$$

III. DATASET:

The modified dataset consists of eight different Fish Species, which have been mentioned below with the help of a table:

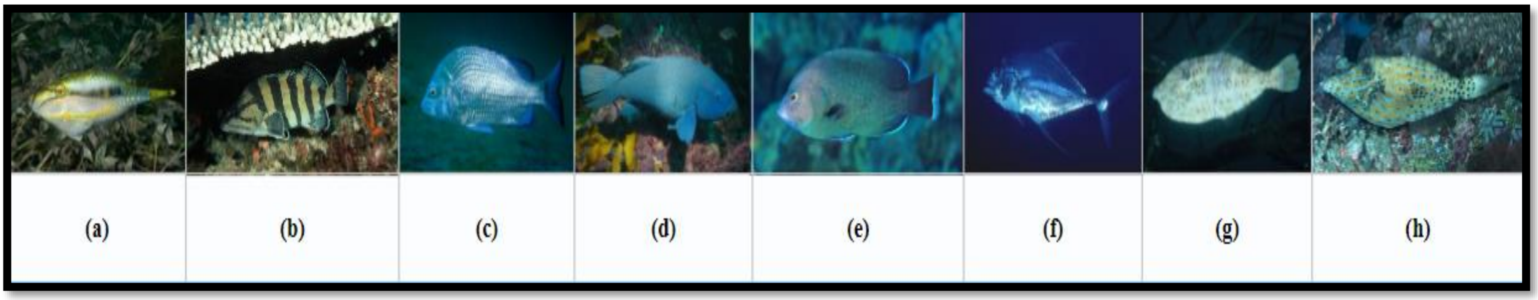


Fig.1. (a) acanthaluteres spilomelanurus (b) acanthistius cinctus (c) acanthopagrus berda (d) achoerodus gouldii (e) aethaloperca rogaa (f) alectis ciliaris (g) aluterus monoceros (h) aluterus scriptus

IV. PROPOSED SYSTEM:

The system gathers a collection of photos from a centralized location. Before feeding the image to the Convolutional Neural Network [7][10][13-15][18] (CNN), Preprocessing of data is implemented. This step includes resizing of the image following a standard ratio. Secondly, training options [1] for neural network using Stochastic Gradient Descent with Momentum [1] (SGDM) need to be specified. The next step is to put the deep learning algorithm [10] such as CNN to use. Using these above steps we happen to aim to increase the fish classification systems accuracy. In this study we suggest a CNN based approach to improve the accuracy [6] of [5][10] the system. The modules related to the proposed system include – Input image, preprocessing, training process, CNN layers, Training options [1], classification of fishes and finally performance estimation

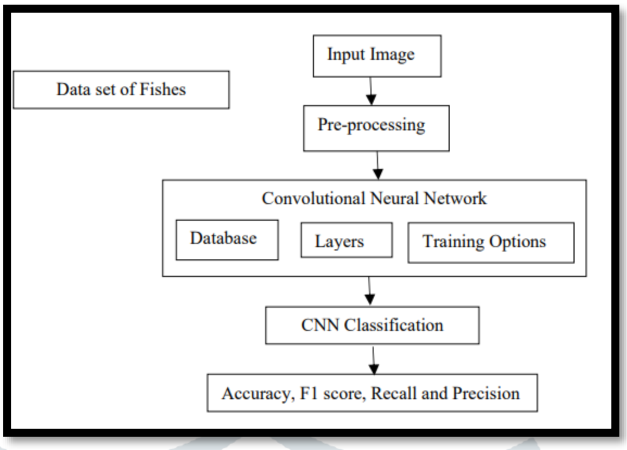


Fig.2. system architecture

4.1 MODULES:

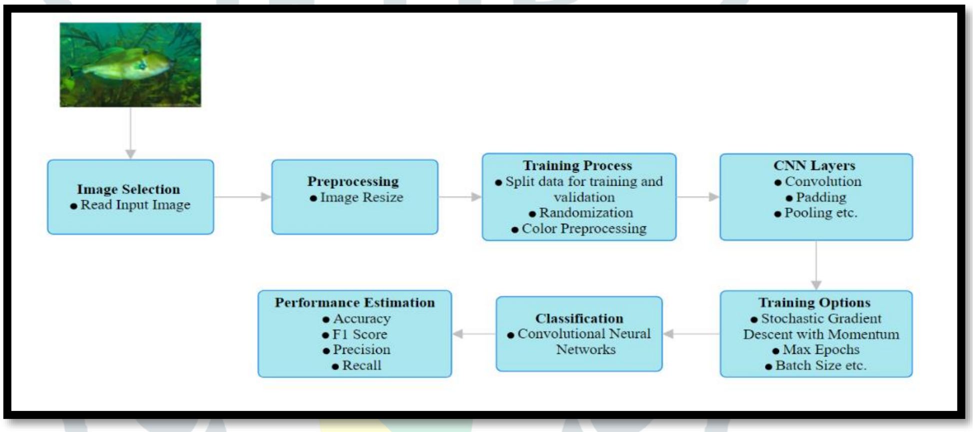


Fig.3. block diagram with brief description of the modules present

4.2 MODULE DESCRIPTION:

1) **INPUT IMAGE:** In this case, the input parameter is underwater images of fish species [3]. For this study, the Ground-Truth dataset [10] has been available via Kaggle data repository [10] (https://www.kaggle.com/sripaadsrinivasan/fish-species-image-data). This dataset was compiled with careful curation and derived from underwater images taken in the deep oceans. It includes about 4000 of meticulously checked images including the ones mentioned in Fig.1. , which reflects a wide variety of fish species. In addition, manual identification and labelling of fish species incorporated in the dataset are supported by marine biologists' recommendations [10]. The input dataset is an aggregation of images that are either in .png or.jpg format for reading and loading the input images, imread () function is used based on research protocol.
2) **PREPROCESSING:** As part of preprocessing stage [17] we resize the image to 227x227 pixels in our methodology. This step is very crucial as the fish species images may vary in resolution and size. This can lead to challenges during model training and inference. Resizing ensures uniformity and reduces computational complexity [6]. Resized images help in generalization and accuracy improvement [6] of the deep learning model.

Fig.4. (a) input image (b) resized image

3) **TRAINING PROCESS:** In the training process [7][17] we store large collection of images with their respective labels in a systematic manner with the help of imageDatastore (). These large set of images into two separate sets one for training and the other for validation, this task is done with the help of splitEachLabel (). The proportion of data we want in the training set [3][7][12][15] in this case is 80% and the rest is needed for validation set [7]. Before splitting the data, randomization of the order of data is done to avoid the overfitting [3][7][15][18] problem. The model is trained with training data, and once the training is done, we use the validation data to evaluate how well the model does based on the hyper-parameters of the model and train again. Moreover, data augmentation [7] and Colour preprocessing plays a major role in the training process [7][17].

4) **CNN LAYERS:** In the Convolutional Neural Network [7][10][13-15][18], we use Convolution2dLayer () to perform convolution operation, this particular function has been used three times, where each time a 3x3 filter is used to [11] extract features [13][18] from the images. In each of the three times the above function is used, the number of filters [4] has been increased from 16 to 32 and at last to 64. Zero padding is performed each time the Convolution2dLayer () is used. Apart from this, other operations and functions used include, batchNormalizationLayer, reluLayer and maxPooling2dLayer (). In the function used for maxpooling, the pooling window size is taken to be 2x2 with a stride of 2. The above-mentioned operations and maxpooling function are all used three times. Moreover fullyConnectedLayer (), softMaxLayer and ClassificationLayer are present in the Convolutional Neural Network [7][10][13-15][18].

5) **TRAINING OPTIONS:** Training Options [1] functions are used to [11] specify the options for training a CNN model using Stochastic Gradient Descent with Momentum [1] (SGDM). Some of the training options [1] used include Environment Execution option which is set to auto meaning the code will set the best available hardware (GPU or CPU) for training. Moreover InitalLearningRate, determining how well the model adjusts to the parameters quickly is set to 0.001 and maxepoches option are used, set to 100. The number of images used in minibatches are 25, the batches are shuffled after every epoch and the verbose is set to true as we need the printing progress info while running the code. The validation data and frequency are specified in order to show how often the input needs to be evaluated during the training process.

Training on single CPU.
Initializing input data normalization.

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Validation Accuracy	Mini-batch Loss	Validation Loss	Base Learning Rate
1	1	00:00:03	0.00%	13.46%	2.6907	6.4839	0.0010
7	50	00:01:30	100.00%	100.00%	0.0030	0.0036	0.0010
13	100	00:03:03	100.00%	100.00%	3.0901e-05	0.0002	0.0010
19	150	00:04:46	100.00%	100.00%	5.5764e-05	0.0002	0.0010
25	200	00:06:29	100.00%	100.00%	4.7305e-06	0.0001	0.0010
32	250	00:08:03	100.00%	100.00%	8.0586e-07	0.0001	0.0010
38	300	00:09:38	100.00%	100.00%	1.7166e-07	0.0002	0.0010
44	350	00:11:10	100.00%	100.00%	9.5367e-08	0.0003	0.0010
50	400	00:12:48	100.00%	100.00%	2.4224e-06	0.0139	0.0010
57	450	00:14:28	100.00%	100.00%	6.5903e-06	1.9152e-05	0.0010
63	500	00:16:01	100.00%	100.00%	0.0001	0.0056	0.0010
69	550	00:17:43	100.00%	100.00%	2.8133e-07	9.0698e-06	0.0010
75	600	00:19:21	100.00%	98.08%	7.1526e-08	0.1683	0.0010
82	650	00:20:57	100.00%	98.08%	4.4346e-07	0.0263	0.0010
88	700	00:22:41	100.00%	100.00%	1.6212e-07	1.4078e-05	0.0010
94	750	00:24:21	100.00%	100.00%	3.3379e-08	1.2017e-05	0.0010
100	800	00:26:08	100.00%	100.00%	3.9101e-07	7.7926e-06	0.0010

Fig.5. training progress

6) **CLASSIFICATION:** To classify the input image using a trained deep learning neural network, classify () is utilized. Now this function returns the predicted class label and classification score. msgbox () is used to [11] display a message box containing the predicted class label. Moreover char () is used to [11] convert the predicted class label to a character array before displaying it.

7) **PERFORMANCE ESTIMATION METRICS:** The performance of the Fish Classification with the use of Convolutional Neural Network [7][10][13-15][18] is taken based on the following performance metrics,

- Accuracy
- Precision
- Recall
- F1 - Score

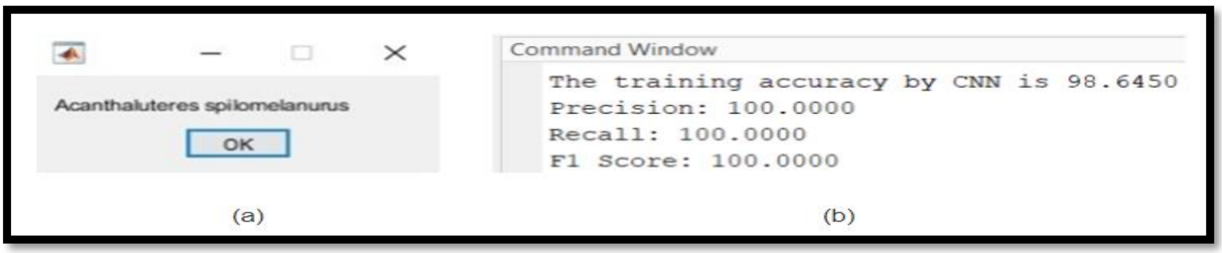
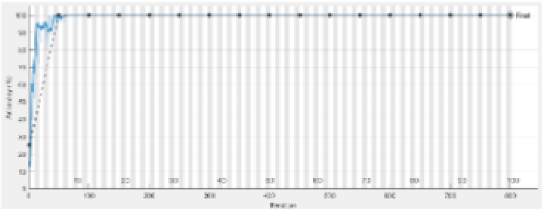


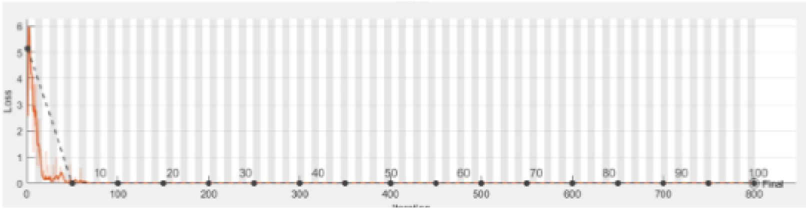
Fig.6. (a) display of fish species (b) display of accuracy in command window

V. RESULTS:

The following results using performance estimation have been obtained where accuracy obtained is 98.6450% as displayed in Fig.6. (b) From the below figures, i.e. Fig.7. (a) and (b), we can say that the fish species chosen as the input image, "acanthaluteres spilomelanurus", When trained, displayed the validation Accuracy [10][15] to be 100%. Moreover the Fig.7. (c) and (d) represent the validation accuracy of all the eight fish species present in the dataset, during the training process and accuracy of the existing systems mentioned in the reference section respectively. The average validation accuracy of the proposed system turns out to be 97.113%



(a)



(b)

Fish Species	Validation Accuracy
Acanthaluteres spilomelanurus	100%
Acanthistius cinctus	100%
Acanthopagrus berda	100%
Achoerodus gouldii	94.23%
aethaloperca rogaa	94.23%
Alectis ciliaris	96.15%
Aluterus monoceros	96.15%
aluterus scriptus	96.15%

(c)

Existing system	Accuracy
SVM	78.59%
Probabilistic Neural Network	89.65%
Video monitoring using CNN	93%
Squeeze and Excitation Ntwork	~ 93%
Improved AlexNet	94%
VGG16	96%
CNN & deep learning	96.29%
Image processing and statistical research analysis	96.67%
Pre-trained AlexNet SPP-SS: UC merced dataset SIRI-WHU google image dataset WHU RS dataset	 96.67±0.94% 95.07±1.09% 95.00±1.12%

(d)

Fig.7. (a) validation accuracy (b) loss (c) validation accuracy of fish species (d) accuracy of existing fish classification systems

VI. FUTURE SCOPE:

The proposed system can further be developed to identify fish species in real time video, and introducing new features to the system, such that the system can display the information related to the particular species identified and display the ideal condition for the fish and the present environmental condition in which the fish is living. Moreover, the dataset taken can be further expanded with the addition of new fish species, that come under endangered species category. So that when fishermen plan to catch fishes, they can be altered by authorities using the system. As a next step the interface can be made user-friendly, so that this system can be accessible to the fishermen as well.

VII. CONCLUSION:

The marine, underwater fish species images used initially as input are from the publicly available data. In order to classify these images precisely, we used the preprocessing methods for the images i.e. image resizing. In the end we have used the Convolutional Neural Network [7][10][13-15][18] layers for extracting the features and classifying the images. The performance was predicted based on the parameters including accuracy, precision, recall, and F1 score. In conclusion we have successfully developed the Automated fish species classification system with the help of Convolutional Neural Network (CNN) to classify the modified dataset containing fish species. This proposed system holds great potential for applications in fisheries management and environmental conservation by providing a useful tool to the stakeholders to gain information related to the fish species.

REFERENCES

[1] A. Agarwal, K. Patni and R. D, "Lung Cancer Detection and Classification Based on Alexnet CNN," 2021 6th International Conference on Communication and Electronics Systems (ICCES), Coimbatre, India, 2021, pp. 1390-1397, doi: 10.1109/ICCES51350.2021.9489033.

[2] Annapurani. K, Divya Ravilla, “CNN based image classification model,” Special Issue, vol. 8, pp. 1106–1114, Oct. 2019

[3] A. Salman, A. Jalal, F. Shafait, A. Mian, M. Shortis, J. Seager, and E. Harvey, “Fish species classification in unconstrained underwater environments based on deep learning,” Limnology and Oceanography: Methods, vol. 14, no. 9, pp. 570–585, 2016.

[4] B S, Rekha & G N, Dr. Srinivasan & Reddy, Sravan & Kakwani, Divyanshu & Bhattad, Niraj. (2020). Fish Detection and Classification Using Convolutional Neural Networks. 10.1007/978-3-030-37218-7_128.

[5] C. Qiu, S. Zhang, C. Wang, Z. Yu, H. Zheng, and B. Zheng, “Improving transfer learning and squeeze-and-excitation networks for small-scale fine-grained fish image classification,” IEEE Access, vol. 6, pp. 78503– 78512, 2018.

[6] D. Rathi, S. Jain, and S. Indu, “Underwater fish species classification using convolutional neural network and deep learning,” in 2017 Ninth international conference on advances in pattern recognition (ICAPR), pp. 1–6, Ieee, 2017.

[7] F. Kratzert and H. Mader, “Fish species classification in underwater video monitoring using convolutional neural networks,” 2018.

[8] Gharghory, Sawsan. (2020). Performance Analysis of Efficient Pre-trained Networks based on Transfer Learning for Tomato Leaf Diseases Classification. International Journal of Advanced Computer Science and Applications. 11. 10.14569/IJACSA.2020.0110830.

[9] J. Duan, B. Yang, X. Wei, N. Liu, and Y. Zhang, “Anomaly detection method for operation and maintenance data based on one-class learning,” IOP Conference Series: Earth and Environmental Science, vol. 440, p. 032146, feb 2020.

[10] J. J. Mol and S. Albin Jose, “An automated fish species classification system using improved alexnet model,” in 2022 6th International Conference on Electronics, Communication and Aerospace Technology, pp. 121–126, 2022.

[11] L. Li and J. Hong, “Identification of fish species based on image processing and statistical analysis research,” in 2014 IEEE International Conference on Mechatronics and Automation, pp. 1155–1160, IEEE, 2014.

[12] M.-C. Chuang, J.-N. Hwang, and K. Williams, “A feature learning and object recognition framework for underwater fish images,” IEEE Transactions on Image Processing, vol. 25, no. 4, pp. 1862–1872, 2016.

[13] M. K. Alsmadi and I. Almarashdeh, “A survey on fish classification techniques,” Journal of King Saud University-Computer and Information Sciences, vol. 34, no. 5, pp. 1625–1638, 2022.

[14] R. Larsen, H. Olafsdottir, and B. K. Ersbøll, “Shape and texture based classification of fish species,” in Image Analysis: 16th Scandinavian Conference, SCIA 2009, Oslo, Norway, June 15-18, 2009. Proceedings 16, pp. 745–749, Springer, 2009.

[15] S. N. M. Rum and F. A. Z. Nawawi, “Fishdetec: a fish identification application using image recognition approach,” International Journal of Advanced Computer Science and Applications, vol. 12, no. 3, 2021.

[16] S. Ogunlana, O. Olabode, S. Oluwadare, and G. Iwasokun, “Fish classification using support vector machine,” African Journal of Computing & ICT, vol. 8, no. 2, pp. 75–82, 2015.

[17] U. Andayani, A. Wijaya, R. Rahmat, B. Siregar, and M. Syahputra, “Fish species classification using probabilistic neural network,” in Journal of Physics: Conference Series, vol. 1235, p. 012094, IOP Publishing, 2019.

[18] X. Han, Y. Zhong, C. Liqin, and L. Zhang, “Pre-trained alexnet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification,” Remote Sensing, vol. 9, p. 848, 08 2017.