# Bangladeshi Indigenous Fish Classification using Convolutional Neural Networks

**Conference Paper** · July 2021

DOI: 10.1109/ICIT52682.2021.9491681



**4 authors**, including:

Krishno Dey
Daffodil International University



Most. Hasna Hena
Daffodil International University

# Bangladeshi Indigenous Fish Classification using Convolutional Neural Networks

Krishno Dey, Md. Mustahid Hassan, Md. Masud Rana, Mst. Hasna Hena
*Computer Science and Engineering*
*Daffodil International University, Dhaka, Bangladesh*
{krishno.taf, mustahid15-9388, masud15-9413, hena.cse}@diu.edu.bd

*Abstract*—**Fish is an important part of Bangladeshi culture and cuisine. Fish is the main and sometimes the only source of protein in the rural household of Bangladesh. Moreover, thousands of peoples of Bangladesh are directly and indirectly dependent on the fish industry. With time many traditional indigenous Bangladeshi fish has lost their existence and many of them are in danger of losing their existence. Furthermore, the young generation of Bangladesh is unable to recognize these traditional indigenous fishes besides they are also missing out on the protein and nutrition provided by these indigenous fishes. Hence an automatic fish classification system can help us not only to recognize traditional fishes but also in the production and preservation of these indigenous fishes. So, in this paper, we propose a convolutional neural network (CNN) based automatic fish classification system. In this paper, we mainly focus on the classification of traditional indigenous fishes of Bangladesh. We used a dataset of eight classes of indigenous fish which contains 8000 images after performing 8 types of augmentation methods. We fed our data to VGG16, Inception V3, MobileNet pre-train models with slight modification in the output layer. We also applied this dataset to a 5 layers CNN model which we name FishNet, where convolutional layers of the CNN model uses "Adam optimizer", "ReLU" and the "SoftMax" activation function. FishNet surpasses VGG16 in all the performance measures and goes toe to toe with InceptionV3 and MobileNet models. Finally, we see all the models provide excellent performance measures.**

*Index Terms*—**Computer Vision, Fish Classification, Fish Identification, Deep Learning, Convolutional Neural Networks.**

## I. Introduction

Fishes are one of the most common food in the world and an integral part of the different cuisine around the world. Fish industry around the world worth billions of thousands of dollars. As time goes by many fishes are losing their existence due to global warming and pollution. We can use modern research techniques to prevent these fishes from losing their existence. With not much effective research approach is not applied to the sector leaves us with tremendous scope to apply effective research operations.

These indigenous fishes are an important part of the diet of Bangladeshi people and one of the main sources of protein for people of Bangladesh. Moreover, indigenous fishes are irreplaceable sources of protein and nutrition. These types of fishes contain omega-3, omega-6 fatty acids, and different vitamins, which are essential for human health [1], [2]. They also lower the risk of heart attacks and strokes, prevents cancer, boosts brain health and skin health [3].

There are 260 species of freshwater indigenous fishes in Bangladesh
citeb4. But many of them have already lost their existence and many of them fighting to keep their existence. Moreover most of us cant even recognize these indigenous fishes. Therefore, an automatic indigenous fish recognition system can really help us to recognize and preserve these indigenous fishes.

In recent times machine learning has blessed us with CNN. CNN can be applied to computer vision based problems and image classifications. CNN can be trained to detect many complex features. It helps us to classify images easily with excellent accuracy measures.

Here we classified 8 different indigenous fishes using our proposed classification model. Dataset we used here contains images of eight classes of indigenous fishes, total of 8000 images. We split our dataset into train, validation and test set to pull off a better result. Then we fed the dataset into a 5 layers convolutional neural networks(CNN) model.

The rest of the paper is organized as follows. Section II represents the literature review of our study. Research Methodology is discussed in section III. Section IV discussed the performance evaluation and analyzed the result of the proposed models and conclude our paper stating the possible future work in section V.

## II. LITERATURE REVIEW

A lot of work has been done in fish identification from image. Many of the researchers have proposed deep CNN model for classification of fish image. But most of the work only focuses on under water fish classification. Moreover a very little work has been done on Bangladeshi fish classification. But only few of them uses deep CNN.

Zheng [5] et al. used a combination of CNN and hand designed image feature for new feature creation. They specially focuses on low resolution images and applied the proposed method on two distinct marine animal datasets. Their proposed method achieved greater accuracy measures than just using

978-1-6654-2870-5/21/$31.00 ©2021 IEEE

CNN alone. Ahmad [6] et al. introduced a system that uses CNN for feature extraction and SVM for classification of underwater fish images. They achieve an average accuracy of 90%. Muhammad [7] et al. used 2 layers CNN on low resolution underwater fish image dataset. They mainly emphasized on augmentation to increase the size of the dataset and dropout to overcome overfitting. They achieved and accuracy of 99.7% on test data. Dhruv [8] et al. introduced a combination of CNN, deep learning and image processing to encounter the noises of the underwater images and achieved an accuracy of 96.29%. Guang [9] et al. introduced a two branch based automated fish detection system. Where one branch detect, align fishes from input no images and pass the fish image to the classifiers. The other branch makes use of fish instances and context information of the input image to infer the type of fish. Their main goal was to automatically label fish using camera without any human interaction. Minsung [10] et al. introduced convolutional neural network based on you only look once (YOLO) algorithm. Their introduced method successfully identified 93% of underwater fish video images. Xiu [11] et al. proposed a fast R-CNN (Regions with Convolutional Neural and Networks) for fast and robust classification of huge amount of underwater fish images (imageCLEF dataset). They showed how their approach can be very helpful for marine biologist to identify various classes of fishes quickly. Md. [12] Aminul et al. introduced a new feature descriptor, Hybrid Local Binary Pattern (HCBL) for Bangladeshi indigenous fish. They used SVM for classification and achieved an average accuracy of 90%.

## III. RESEARCH METHODOLOGY

Here, we mainly focuses on classification of eight categories indigenous fishes of Bangladesh. Our proposed methodology is shown in figure 1.

### *A. Dataset*

Here we used dataset namely BDIndigenousFish2019 publicly available at Githoub, created by Md. Aminul et al. [12]. Dataset contains eight classes of indigenous fish images. Dataset consists of total 2610 images all are in .jpg format. Description dataset is given below with their frequency, common engilsh name and scientific name [13]. in Table I. Figure 2 shows sample view of images of the dataset.

### *B. Data Pre-processing*

Here we re-sized all the images to 224 x 244 pixels for fitting into VGG16, MobileNet and FishNet. We also had to re-size the image to 299x299 pixel because the input shape of Inception V3 is 299x299. We re-scaled the value of each pixel into a range of 0 to 1 from the range of 0-255. Proper augmentation of image data is very important to increase the size of the data set which eventually helps us to achieve good accuracy and combat overfitting [14]–[16] . Seven different types of augmentation were applied.

- Rotation of 40 degree.

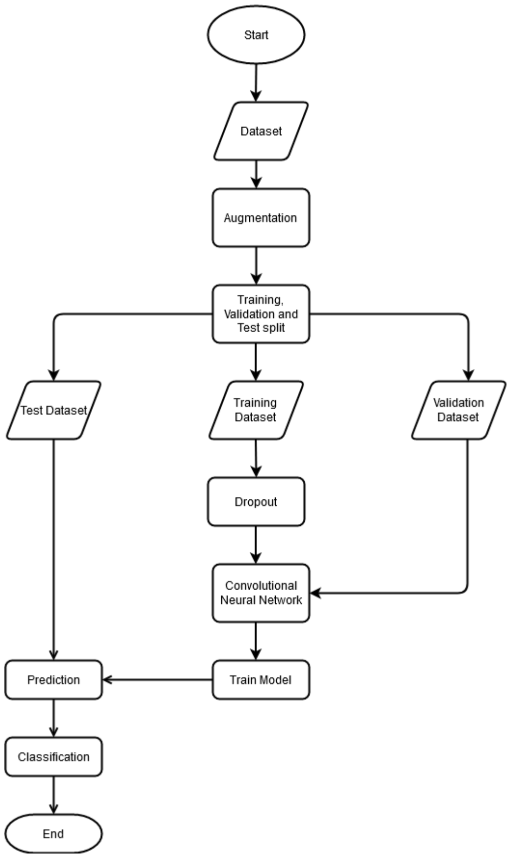


Fig. 1: Research Methodology

TABLE I: Frequency Distribution in of Dataset

| SL | Common English Name | Scientific Name | Local Name | No. of Images |
|---|---|---|---|---|
| 1 | Lesser spiny eel | Macrognathus aculeatus | Tara baim | 500 |
| 2 | Bronze featherback | Notopterus notopterus | Pholi | 300 |
| 3 | Climbing perch | Anabas testudineus | Koi | 380 |
| 4 | Stinging catfish | Heteropneustes fossilis | Shingi | 400 |
| 5 | Snakehead murrel | Channa striata | Shol | 120 |
| 6 | Olive barb | Puntius sarana | Sarpunti | 200 |
| 7 | Spotted snakehead | Channa punctata | Taki | 390 |
| 8 | Tyangra | Mystus tengara | Tengra | 390 |

- Width shift range of 0.2.
- Height shift range of 0.2.
- Shear range of 0.2.
- Zoom range of 0.2.
- Horizontal flip is equal to true.
- Fill mode is equal to nearest.

So, after augmentation total of 8000 images were created from original of 2610 images where each of eight class contain 1000 images each. We divided our dataset into training, validation and test set for obtaining a better result. From each of eight classes 700 images (70%) used for training, 150 images (15%) used for validation, 150 images (15%) used for

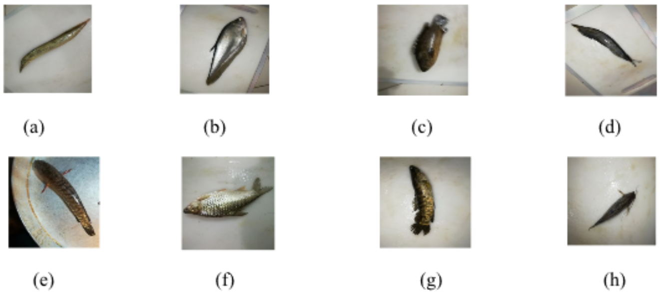


Fig. 2: A small slice of dataset, a. Lesser spiny eel, b. Bronze featherback, c. Climbing perch, d. Stinging catfish, e. Snakehead murrel, f. Olive barb, g. Spotted snakehead, h. Tyangra.

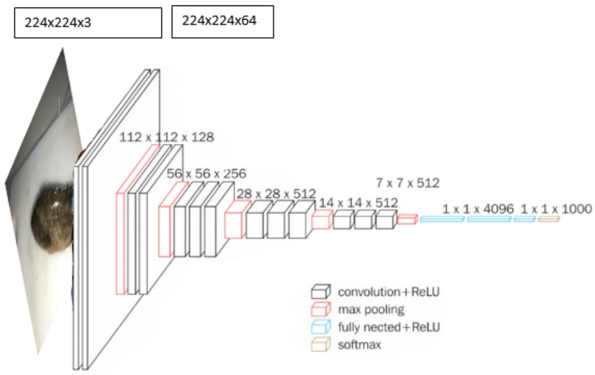


Fig. 3: Model configuration of VGG16

testing. Table II shows frequency distribution after splitting augmented dataset into training set, validation set and test set.

TABLE II: Dataset Division after Augmentation

| **Total Augmented Images** | **Training (70%)** | **Validation (15%)** | **Testing (15%)** |
|---|---|---|---|
| 8000 | 5600 | 1200 | 1200 |

### C. VGG16

A very deep convolutional neural network model VGG16 was proposed by K. Simonyan and A, Zisserman back in 2014. Vgg16 achieved top-5 test accuracy with accuracy of 92.7%. Vgg16 has the depth of 23. The main concept was that it use several 3×3 convolution filters instead of large filters to increase the depth of the model, while maintaining considerable number of parameters. Vgg16 was proved to be an improvement over other prior works like AlexNet, this was achieved by pushing the depth to 16-19 weight layers. Vgg16 basically replaced kernel size filters 11 and 5 of the first and second layers of AlexNet with multiple 3×3 kernel-sized filters one after another. Furthermore, vgg16 is a very deep convolutional neural network with input shape of 244×244. Figure 3 show the architecture of a VGG16 model.

### D. Inception V3

Inception v3 is convolutional neural networks architecture introduced by google for classification of ImageNet dataset where it yields 78.1% accuracy. Inception V3 is widely used for image analysis and object detection is the third edition to the Google's convolutional neural networks. Use of label smoothing, factorized 7×7 convolutions and use of an auxiliary classifier to propagate label information lower down the network make inception v3 significant improvement over the previous versions. We can consider the inception v3 as combination of symmetric and asymmetric building block which also includes convolution average pooling, max pooling, concats, dropouts, batchnorm, softmax and fully connected layers. Moreover we can consider the inception v3 is consist of two parts, where fist part extract feature from input and the second part classify image based on their features. Figure 4 shows the simple architecture of Inception V3 model.

### E. MoblieNet

MobileNet is a very light weight deep convolutional neural network which is smaller in size and faster than any other popular classification model available nowadays. It is widely used for image detection, face attributes, image analysis and any other task that CNN model performs. MobileNet inherits from a class of very small, low latency and low power model available. It uses a streamlined architecture which uses depth wise separable convolutions instead of 3x3 filters. Input shape of the model 224x224 .While MobileNet is smaller in size and faster but there is a trade-off like every other thing, it is slightly other deep learning models. Figure 5 shows the architecture of MobileNet.

### F. FishNet

We name our proposed model as FishNet and details configuration is given in the followings.

For efficient classification of indigenous fish images we are proposing a 5 layered CNN. The model consist of 5 convolutional layers and 2 hidden layers in fully connected layer. At, all the 5 convolutional layers and hidden layers, ReLU activation function is used except the output layer where Softmax activation function is used. ReLU is known for its simplicity, it convert an input into either zero or one. ReLU is mostly used in hidden layers. On the other hand softmax mostly used in output layer of multiclass

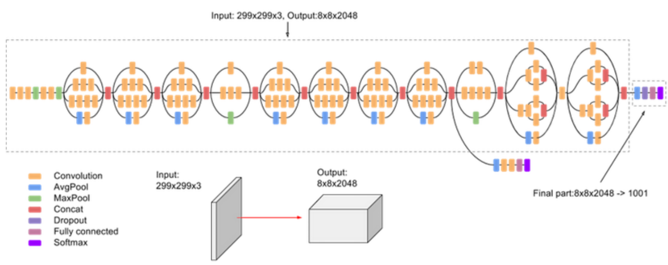


Fig. 4: Model configuration of Inception V3

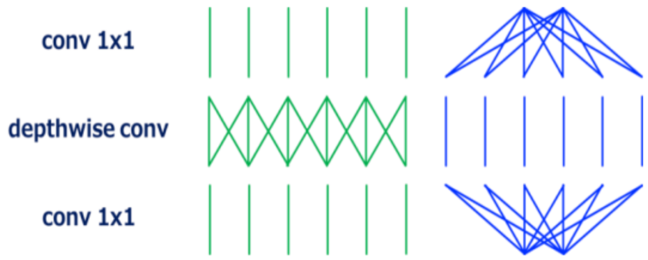


Fig. 5: Model configuration of MobileNet

classification problem. Input shape of the model is 224x224. Figure 6 shows the model configuration.

Dropout layers actually randomly drop few of inputs form the previous layer of neural network. Dropout is an efficient way to reduce overfitting [17]. Dropout used in 3 hidden layers in fully connected layers are respectively 40%, 40% and 30%.

### *G. Training*

Here we train all 4 mention CNN model for 20 epoch with the learning rate of 0.0001. Just 20 epoch was enough for training accuracy and loss to converge to a stable value. We used 5600 images for training and 1200 images for training.

Figure 7 shows the comparison of accuracies of all the four models on training data over the 20 epoch and figure 8 shows the comparison of the losses of all four models over the 20 epoch.

From the figure 7 we can observe that the performances of both Inception V3 and MobileNet is quite similar, we hardly them from the graph provided in the figure 7. On the other other hand the performance of the FishNet model is slightly below Inception V3 and MobileNet but well over the VGG16 model.

From figure 8 we can feel assure that there is no overfitting or underfitting.

Then we observe the performance measures of the models. Performance results are thoroughly discussed in the next section.

## IV. PERFORMANCE EVALUATION

All the Model was run for just 20 epoch with batch size of 10 and 20 epoch was quite enough for our model to converge and give a more stable accuracy.

For compilation of the models we use Adam optimizer.. Adam optimizer [18] is based on stochastic gradient descent (SGD) method and provide efficient result than other SGD based methods. Unlike classical SGD method learning rate is adapted as learning upholds. Adam optimizer is found to robust dealing with large dataset and noisy data. As we have not used any noise removal techniques, Adam was a straightforward choice. We used categorical-crossentropy loss function as we dealing with a multiclass problem. Loss function basically represent the difference between the predicted output and actual output. An optimizer always tempt to minimize the loss function by continuously updating the weights of the model during training. Along with a good optimizer a proper learning rate is also very important. Learning rate or step size defines the interval of jumps the optimizer make after each epoch for adjusting the weights of the model.

After a lot of experimentation the learning rate of 0.0001 provided the best result for our problem. All four model uses the same learning rate and provide excellent performance measures.

After training all the models we have compared the training and validation parameters to observe the performance of the models. We have compared training accuracy vs validation accuracy. In case of all four models we found training accuracy and validation accuracy is pretty much the same. The differences between training loss and validation loss was also very little which made sure there is no overfitting or underfitting any of the four models. Table III shows the accuracy and losses on train, validation and test set.

TABLE III: Accuracy and Losses on Train, Validation and Test set

| **Models** | **Training** | | **Validation** | | **Test** |
|---|---|---|---|---|---|
| | ***Accuracy*** | ***Loss*** | ***Accuracy*** | ***Loss*** | ***Accuracy*** |
| VGG16 | 98.29% | 07.34% | 97.42% | 08.24% | 97.25% |
| Inception V3 | 100% | 0.010% | 100% | 0.013% | 100% |
| MobileNet | 99.84% | 01.06% | 99.92% | 0.26% | 99.83% |
| FishNet | 99.61% | 01.67% | 99.00% | 03.25% | 99.00% |

This subsection is dedicated to discussion of the result obtained from our proposed models. We used performance matrices like precision, recall and f1-score. For testing we used 1200 images, 150 images each form total of eight classes of fishes. We got an accuracy of 97.25%, 100%, 99.83% and 99.00% on test data respectively for VGG16, Inception V3, MobileNet and FishNet model.

We cannot evaluate a model with just accuracy we also need precision, recall, f1 score, receiver operating characteristics curve etc. and we can determine all this through confusion matrix. Confusion matrix is one very popular in the field of machine learning for visualization and performance evaluation of algorithms. It helps us the check the correctness of a model or algorithms. For further evaluation of the models, we calculated performance matrices from confusion matrices for all four models. Table IV shows average precision, recall and f1-score of all four models.

We also did ROC curve analysis for all four models and observed 99.96%, 100%, 100% and 99.99% area of respectively VGG16, Inception V3, MobileNet and FishNet

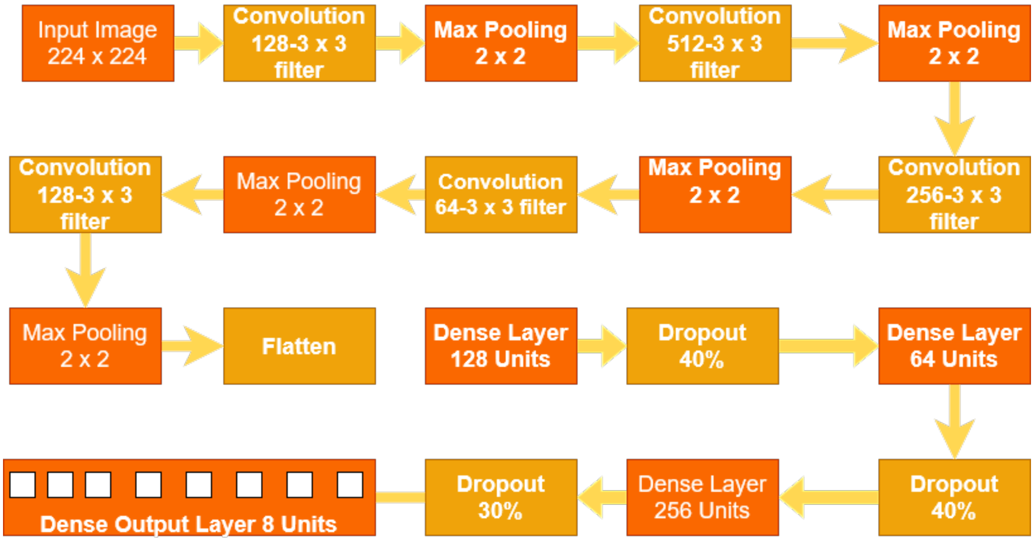


Fig. 6: Model configuration of FishNet

TABLE IV: Average precision, recall and f1-score

| Models | Trainable parameters | Precision | Recall | f1-score |
|---|---|---|---|---|
| VGG16 | 138.35 millions | 97.369% | 97.25% | 97.23% |
| Inception V3 | 23.85 millions | 100% | 100% | 100% |
| MobileNet | 4.25 millons | 99.83% | 99.83% | 99.83% |
| FishNet | 2.82 millons | 99.01% | 99.00% | 99.00% |

models are under the ROC curve.

From the above performance analysis it pretty clear that FishNet provides better performance than VGG16 model where Inception V3 and MobileNet provide slightly better performance than FishNet. If you consider the trainable parameters FishNet has only 2.82 millions where Inception V3 has 23.85 and MobileNet has 4.25 millions. FishNet is faster and easier to train than those two models. So, from this perspective our proposed model FishNet is arguably superior than both Inception V3 and MobileNet.

Finally, we can say all four models produces excellent performance measures and can be deployed in real life automated system for robust classification of fish images.

In table V, we have shown a comparison of our work with the some of the previous work in fish classification using CNN.

TABLE V: Accuracy and Losses on Train, Validation and Test set

| Works | Data Size | Classifier | Accuracy(%) |
|---|---|---|---|
| This Work | 8000 | VGG16 | 97.25 |
| This Work | 8000 | Inception V3 | 100 |
| This Work | 8000 | MobileNet | 99.83 |
| This Work | 8000 | FishNet | 99.00 |
| Salman et al. | - | CNN | 90.00 |
| Rachmatullah et al. | 22443 | CNN | 99.70 |
| Rathi et al. | 27142 | CNN | 96.29 |
| Sung et al. | 929 | CNN with YOLO | 93.00 |
| Islam et al. | 2610 | SVM | 90.00 |

## V. CONCLUTION & FUTURE WORK

In this paper we have presented a deep CNN model to classify eight categories of Bangladeshi indigenous fishes. From the performance evaluation it is proved that FishNet model provide a satisfactory accuracy in classification. FishNet also provides rich precision, recall and F1 scores. In future we want to apply our FishNet model on more categories of indigenous fishes. With large and quality dataset

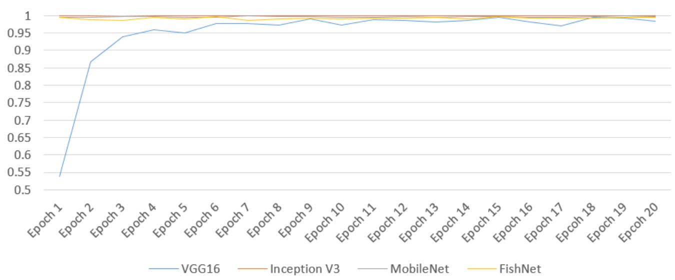


Fig. 7: Accuracy comparison of all four models

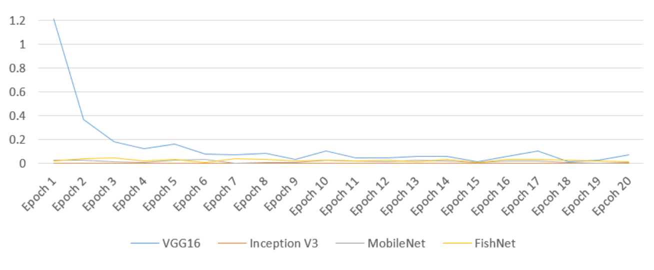


Fig. 8: Loss comparison of all four models

we can enhance the performance our model even more.

We feel if we can manage to collect a lot of images of these fishes with the collaboration of BFRI (Bangladesh Fisheries Research Institute) we can build a classification system for a large-scale indigenous fish classification. We want to deploy our models to web and mobile application so they become accessible to BFRI and as well as general people. Then we can continuously update our dataset and retrain our model with time. That would greatly benefit BFRI, they would no longer be required to detect each fish manually. General people will also be able to learn about Bangldeshi indigenous fishes. The classification system would detect fishes automatically, easily, and accurately. So in the future, we want to build a classifier for large scale classification of Bangladeshi indigenous fishes.

## REFERENCES

[1] N. Roos; M.M. Islamyand; S.H. Thilsted "Small Indigenous Fish Species in Bangladesh: Contribution to Vitamin A,Calcium and Iron Intakes", Animal Source Foods and Nutrition in Developing Countries, vol. 6, pp 1-6, June 24–26, 2002.

[2] Healthline.com, '11 Evidence-Based Health Benefits of Eating Fish', 2019. [Online]. Available: https://www.healthline.com/nutrition/11-health-benefits-of-fish. [Accessed: 21- Jan- 2021].

[3] Organicfacts.net, '14 Amazing Benefits Of Fish', 2020. [Online]. Available: https://www.organicfacts.net/fish.html/. [Accessed: 01- Jan- 2021].

[4] Bdfish.org, 'Learn about indigenous fishes of Bangladesh', 2012. [Online]. Available: http://en.bdfish.org/2009/11/small-indigenous-species-sis-of-fishes/. [Accessed: 01- Dec- 2020].

[5] Z. Cao, J. C. Principe, B. Ouyang, F. Dalgleish and A. Vuorenkoski, "Marine animal classification using combined CNN and hand-designed image features," OCEANS 2015 - MTS/IEEE Washington, Washington, DC, USA, 2015, pp. 1-6, doi: 10.23919/OCEANS.2015.7404375.

[6] A. Salman, A. Jalal, F. Shafait, A. Mian, M. Shortis, J. Seager, E. Harvey 'Fish species classification in unconstrained underwater environments based on deep learning". Limnology and Oceanography: Methods, vol.16, 2016.

[7] M. N. Rachmatullah and I. Supriana, "Low Resolution Image Fish Classification Using Convolutional Neural Network," 2018 5th International Conference on Advanced Informatics: Concept Theory and Applications (ICAICTA), Krabi, Thailand, 2018, pp. 78-83, doi: 10.1109/ICAICTA.2018.8541313.

[8] D. Rathi, S. Jain and S. Indu, "Underwater Fish Species Classification using Convolutional Neural Network and Deep Learning," 2017 Ninth International Conference on Advances in Pattern Recognition (ICAPR), Bangalore, India, 2017, pp. 1-6, doi: 10.1109/ICAPR.2017.8593044.

[9] G. Chen, P. Sun and Y. Shang, "Automatic Fish Classification System Using Deep Learning," 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI), Boston, MA, USA, 2017, pp. 24-29, doi: 10.1109/ICTAI.2017.00016.

[10] M. Sung, S. Yu and Y. Girdhar, "Vision based real-time fish detection using convolutional neural network," OCEANS 2017 - Aberdeen, Aberdeen, 2017, pp. 1-6, doi: 10.1109/OCEANSE.2017.8084889.

[11] Xiu Li, Min Shang, H. Qin and Liansheng Chen, "Fast accurate fish detection and recognition of underwater images with Fast R-CNN," OCEANS 2015 - MTS/IEEE Washington, Washington, DC, USA, 2015, pp. 1-5, doi: 10.23919/OCEANS.2015.7404464.

[12] M. A. Islam, M. R. Howlader, U. Habiba, R. H. Faisal and M. M. Rahman, "Indigenous Fish Classification of Bangladesh using Hybrid Features with SVM Classifier," 2019 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), Rajshahi, Bangladesh, 2019, pp. 1-4, doi: 10.1109/IC4ME247184.2019.9036679.

[13] Wikipedia.org,'Learn about indigenous fish of Bangladesh and their scientific names', [Online]. Available: $https : //en.wikipedia.org/wiki/list_of_fishes_in_bangladesh$. [Accessed: 12-oct-2020].

[14] A. Mikołajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," 2018 International Interdisciplinary PhD Workshop (IIPhDW), Świnouście, Poland, 2018, pp. 117-122, doi: 10.1109/IIPHDW.2018.8388338.

[15] A. Fawzi, H. Samulowitz, D. Turaga and P. Frossard, "Adaptive data augmentation for image classification," 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 2016, pp. 3688-3692, doi: 10.1109/ICIP.2016.7533048.

[16] C. Shorten; T.M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning", Journal Big Data 6, 60 (2019). https://doi.org/10.1186/s40537-019-0197-0

[17] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting,"Journal of Machine Learning Research (JMLR), vol. 15, no. 1, pp. 1929–1958, Jun. 2014.

[18] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," inProc. Int. Conf. Learn. Representations, 2015, pp. 1–15, [Online]. Avail-able: https://arxiv.org/pdf/1412.6980.pdf