

# FishDeTec: A Fish Identification Application using Image Recognition Approach

Siti Nurulain Mohd Rum<sup>1</sup>, Fariz Az Zuhri Nawawi<sup>2</sup>

Faculty of Computer Science and Information Technology  
Universiti Putra Malaysia, Serdang, Selangor, Malaysia

**Abstract**—The underwater imagery processing is always in high demand, especially the fish species identification. This activity is as important not only for the biologist, scientist, and fisherman, but it is also important for the education purpose. It has been reported that there are more than 200 species of freshwater fish in Malaysia. Many attempts have been made to develop the fish recognition and classification via image processing approach, however, most of the existing work are developed for the saltwater fish species identification and used for a specific group of users. This research work focuses on the development of a prototype system named FishDeTec to the detect the freshwater fish species found in Malaysia through the image processing approach. In this study, the proposed predictive model of the FishDeTec is developed using the VGG16, is a deep Convolutional Neural Network (CNN) model for a large-scale image classification processing. The experimental study indicates that our proposed model is a promising result.

**Keywords**—Component; Freshwater Fish; fish species recognition; FishDeTec; Convolutional Neural Network (CNN); VGG16

## I. INTRODUCTION

Rivers, ponds, and lakes are full of intrigue and mystery, and the interesting topic has always been the underwater discovery. Estimating the quantity of fish and its presence from image sources may help biologists to understand the underwater habitats and natural environment to aid preservation. There are more than 30,000 species of fish worldwide [1] and it is almost impossible to identify each one by just simply looks at their physical outlook as most of them have similar shapes. There are cases where people die due to the lack of information to differentiate between the non-poison and poison fish and it is happening every day [1-3]. In recent years, image recognition and classification techniques have attracted many scientists to improve the scientific field. It would be time consuming and tedious job for human to analyse and process the massive data generated by the underwater images. In fishery education, minimizing human error during the fish observation and analysis process is important and requires an automatic system detection. Image classification is the process of taking inputs such as images and producing output type categories or probabilities for a particular class. Studies in fish image recognition are as significant area, especially in the marine biology and aquaculture. Fish in general, having a skull and spine, usually breathe through the gills attached to the skin. They have a slender body shape suitable for swimming and fins to make them move faster through the water. The fish category can be

categorized into two types, namely saltwater fish, and freshwater fish. When comparing saltwater with freshwater fish, there are differences between these two types of fish in terms of its physiology, structural adaptation, and size. The freshwater fish is able to survive in a variety of habitats. There are some species can live in mild temperatures at 24 degrees Celsius, while others can survive at very low temperatures, between 5 to 15 degrees Celsius. Freshwater fish can be found in lakes, wetlands, and shallow rivers, where the water salinity is less than 0.05 percent. Saltwater fish can be found in a diversity of habitats, ranging from cold Antarctica and the Arctic Ocean to the warm tropical oceans. The most suitable habitats for saltwater fish include coral reefs, mangroves, salt ponds, deep sea and seagrass beds, and a number of fish thrive in each of these conditions. The size of freshwater fish can be from the small Filipino gobies which is less than an inch in size to the white sturgeon that weighing about 400 pounds, are one of the world's biggest freshwater fish. Freshwater fish include catfish, cisco, charr, gar, mooneye, shiner, trout (blueback, apache, brook, cutthroat and brown), sunfish, pike, whitefish and Salmon [4]. Fig. 1 is some freshwater fish species available in Malaysia. Saltwater fish include certain types of bass, albacore, common dolphin, butterfish, bluefish, eels, flounder, mackerel, cod, herring, marlin, shark, yellowtail, tuna, and snapper [5]. Many studies have been made to detect the fish images especially in saltwater habitat such as [6-8]. Several attempts have been made to recognise the visual images, but it is still an unsolved problem due to segmentation errors, distortion and occlusion and overlap of objects in coloured images [9, 10]. The problem of object classification lies in the main challenge of estimating the prevalence of each species of fish. Solutions to automatically detect the fish classification should be able to overcome problems related to fish size and orientation, feature variability, picture quality and segmentation.

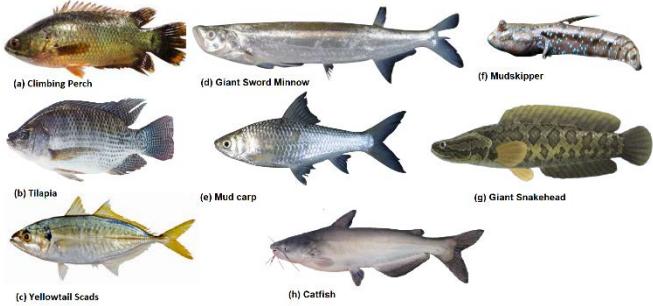


Fig. 1. Some of Freshwater Fish Species available in Malaysia.

Although progress has been made in the field of data generated in real time as well as the improvement of long-distance resolution, the existing works are still limited in their ability to detect or classify the freshwater fish, especially the species found in Malaysia. In other words, a work that specifically provide the identification of Malaysia freshwater fish species through the image processing could not be located. Motivated by this factor, the main objective of this study is to propose the development of a mobile application named FishDeTec to identify and classify the image of freshwater fish species found in Malaysia. In the next sections, we discussed the existing works done by others and followed by the entire processes required for the development of the proposed system.

## II. RELATED WORK

The Convolutional Neural Network (CNN) is a multi-layered/deep neural network designed to detect visual patterns using minimal pre-processing image's pixel. CNN is a unique neural network architecture and consist of two major components, namely convolutional and pooling layers. It can be used to capture the image vision in near-infinite ways. There are number of CNN architectures, which are the key to build algorithms to control and power AI as a whole in the near future. Some of them are LeNet [11], VGGNet [12], AlexNet [13], ZFNet [14], ResNet [15] and GoogLeNet [16]. The VGG16 also known as OxfordNet presented in Fig. 2, is the architecture of CNN is named after the Oxford Visual Geometry Group, which created it. It was used in 2014 to win the ImageNet Large Scale Visual Recognition Challenge (ILSVR) competition (ImageNet).

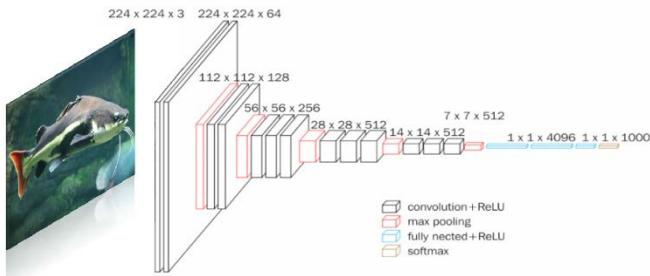


Fig. 2. The Architecture of VGG16.

The VGG16 is a CNN model developed by Karen Simonyan and Andrew Zisserman from the Oxford University [12]. This model has reached the accuracy at 92.7% and is ranked in the top 5 in ImageNet's, which is a data set that has more than 14 million images from 1000 classes. It was one of the popular models submitted to ILSVRC-2014. The improvisation was made by replacing the large size of kernel filters for the first and second convection layer, respectively with  $3 \times 3$ filters to improve the AlexNet [13] model. The VGG16 has been trained for weeks and used the GPU of NVIDIA Titan Black. The differences architectures of VGG16 and AlexNet is presented in Fig. 3. The main reason why VGG16 is a preferred CNN method over the AlexNet CNN in this study is because it supports the processing for a large-scale data set with a deeper network layers and smaller filter to produce a better performance. The VGG16 itself is the improvement of Alex Net. In addition, the

complexity for detecting the fish species from images required an effective modelling approach as most of them have similar shapes and features. The VGG16 has been widely used for Transfer Learning (TL) because of its performance. The main purpose of TL is to transfer the knowledge obtained from the source domain from the large dataset to the target domain, which is a smaller dataset. This is a good criterion for the underwater image's classification specifically the fish recognition. Today, there are several attempts have been made specifically for the development of the fish image recognition model using Machine Learning (ML) and Deep Learning (DL) approaches. The work by Puspa Eosina et al. [17] for example, presents the Soble's method for detecting and classifying freshwater fish in Indonesia. They used 200 numbers of freshwater images from 10 difference species to evaluate their model. However, to enhance the accuracy of the model, additional techniques are still needed such as texture or colour and retrieval of content-based technique. A study by [18] proposed a DL technique combined the Dense Neural Network (DNN) and Spatial Pyramid Pooling (SPP) by putting the SPP in front of the DenseNet layer.

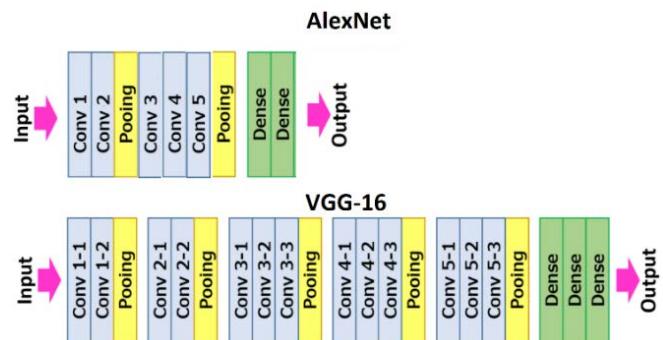


Fig. 3. The Difference Architecture between AlexNet and VGG16 Models.

They proposed a method to remove the noise in the dataset before the training step taken into action through the image processing implementation. This can eliminate the underwater obstacles, dirt, and non-fish bodies from the images. They use the DL approach by implementing the CNN model for the fish species classification. They also provide in their article the description of the performance of the different activation functions with the comparison of ReLU, SoftMax, and tanh and the activation function of the ReLU was found to be exceptionally precise. The authors in [6] propose a method for automated classification of fauna images using the CNN on Animal-10 dataset, with the accuracy at 91.84%. They discuss the implementation of the VGG16 architecture of CNN and the activation role of Leaky ReLU in image classification. The neural network learned to categorize the animal's image. The proposed classification of fauna images using a CNN can be widely used for the classification of fauna images, which would enable the ecologists and researchers to further analyze to preserve the environment and habitats. They used six different species of freshwater fish to test the model. The work by [19] demonstrates 96.29% accuracy for the automatic classification of aquatic fish species relative to traditional approaches. They emphasize on the image of the fish outside of the water, this is to get minimal background noises of the image such as distortion, occlusion, and image quality. The

FishApp [20] is a cloud based application for fish species recognition. It consists of a smartphone application designed for the Android and iOS mobile operating systems that allows the user to take and send photographs of a whole fish for remote inspection and a remote cloud-based computing system that incorporates a sophisticated image processing pipeline and a DL neural network to interpret images and identify them into predefined fish classes. The DeepFish [21] is a framework developed to classify fish from photographs collected in the marine observation network installed underwater cameras. In their work, they used the low rank matrices and sparse to extract the foreground. The deep neural network is used to extract the image of fish. The principal component analysis known as PCA is used in this architecture specifically in two layers of convolutional and block wise histograms in pooling layer and in a non-linear layer, it used binary hashing. They used the linear SVM for the classification and achieved the accuracy at 98.64%. Apart from the scholarly work, there are also a number of mobile applications available for fish identification such as FishVerify [22] and FishID+ [23]. The FishVerify for example, is a mobile application developed to help the local community in Florida, to identify fish. The users are provided with the instance identification of fish from live scan or photo as well the fishing rules and regulation in Florida. The FishID+ application has the same purpose as FishVerify, however the main target user is the fish collector. It uses the DL to verify the freshwater fish, aquarium fish and focuses only on the small fish. The database contains of more than 240 numbers of species of fish, including cichlids, clownfish, tetras, tangs and many other common aquarium fish. Based on our study, a specific mobile application developed for recognizing freshwater fish in Malaysia could not be located. Motivated by this factor, this study presents the development of FishDeTec, utilizing the Convolutional Neural Network (CNN) for the model development in identifying Malaysia freshwater fish.

### III. METHODOLOGY

In this section, we describe in detail the methodology used to develop the proposed system. The designed methodology is suitable for executing this research work as it supports the TL as introduced in VGG16, in which the knowledge gained from the pre-trained model can be used to improve the generalization about another task. The technologies used to realize this research work is, Python for the programming language, Google Colab as the editing and compiling tool for Python. The TensorFlow Mobile is an open-source library used to perform the image recognition task that can be easily embedded in Android Studio application through the library. The Android Studio is the platform used to develop the android mobile application. While the Firebase database is used to store all the information about the fish as well as the images. The VGG16 is the CNN used for the fish image recognition. This technique is preferred method as it is a network trained on more than a million images from the ImageNet database. Fig. 4 is the steps required for the execution of this study. It starts with the requirement analysis. During this stage, the limitation of existing work is analyzed from the literature review. Based on the analysis, it

can be concluded that most of the existing models were developed to recognize the saltwater fish, therefore this study is focusing on the development of fish specifically the Malaysia's freshwater fish. The next step is data set preparation and pre-processing. Usually, the process of acquiring data is messy where it comes from different sources. At this stage, images of several freshwater fish species of are collected. Most images used today is 24 bit or higher. The 8-bit gray scale image consists of black and white, and each pixel has a value ranging from 0 to 255. The RGB color picture indicates that the pixel color is a blend of red, green, and blue. The colors vary from 0 to 255 each. This generator of RGB colors demonstrates how RGB can produce any color. A pixel, then, comprises a range of the three RGB values (102, 255, 102) that correspond to color #66ff66. Images with a size of 0.48 megapixels consist of 800 pixels wide and 600 pixels high. However, the VGG16 was originally trained on 224 x 224 size of images.

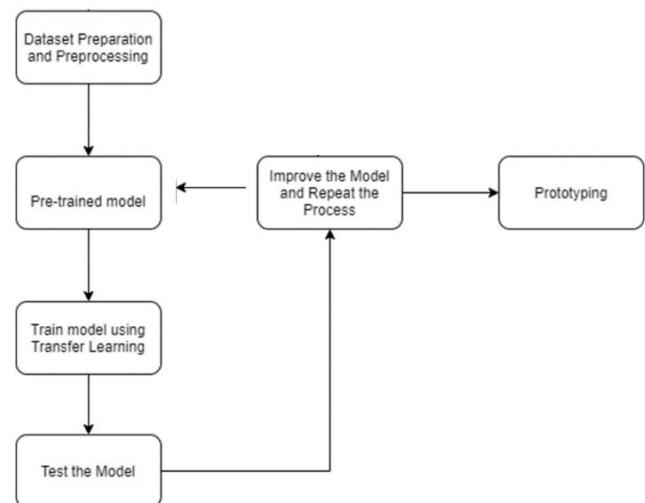


Fig. 4. The Methodology for the Development of FishDeTec.

To feed all images into a neural network model, it requires the cleaning process, standardization, and data augmentation. It is not practical to come up with a specific algorithm to suit different conditions, therefore all images are converted into the form that allows a common algorithm to fix it. A general pre-processing method (e.g., reflection, rotation, histogram, Gaussian blurring, equalization, and translation) requiring the enhancement of current datasets of perturbed copies of existing images is called data augmentation. This task is carried out to enlarge the data set and to expose the neural network to the wide range of image variations. This allows the model to recognize the object when it appears in any shape and form. The data is split into training data and testing data. The total number of images used in this study is 200 for eight type of fish species. The dataset distribution is illustrated in Table I.

The purpose of the ImageDataGenerator is to conveniently import label data into the model. It provides many functions, such as rescale, flip, zoom, rotate etc. The best part of this class is that the data contained on the disk are not affected. This class converts data on the go while feed it to the model.

The next step is to pre-trained the model using the VGG16 neural network. During this stage, the Tensorflow and Keras are used to develop the model. Since we deployed the script using the cloud infrastructure (Google Colab), the memory performance is not an issue. Most of the coding is based on the Keras only, which at the back end uses the TensorFlow. The next step is to train the developed model using TL. The TL is a well-known approach in computer vision as it enables the model accuracy development in a time-saving way[24]. The TL is not a new concept which is very specific in the DL. There is a significant difference between the traditional approach of learning model and ML and using methodologies that adapt the TL principles. In traditional ML, the knowledge is not retained, whereby in TL, learning the new tasks relied on the previous learned tasks. During this step, we do not train all layers of the model. We just freeze all the layers and train the lower layer of the model, which is using the weight of the trained model and this makes retraining very easy. The process is repeated to improve the model. Once, it has achieved the desire target of accuracy, the model is then embedded into the android mobile application development using Java. We have also added the information about the fish species in the Firebase database such as the scientific name, the name of the fish in Malay, etc. Table II is the technologies used to develop the whole system.

TABLE I. THE DISTRIBUTION OF DATASET

Fish	Number of Images in the training set
Climbing Perch	20
Catfish	25
Tilapia	20
Yellowtail Scads	25
Giant Snakehead	25
Giant Sword Minnow	18
Mud carp	25
Mudskipper	20

TABLE II. TECHNOLOGIES USED TO DEVELOP FISHDETEC

Technology used	Description
Python	The programming language used to develop the fish identification model.
Java	The programming language used to develop the android application
TensorFlow Mobile	The open source library used for the DL purpose.
Keras	The API for the DL used for Python
VGG16	The type of CNN used to detect the fish image species.
Google Colab	The development environment for Python that runs in the browser using Google Cloud.
Android Studio	The integrated development environment for Android operating system.
Firebase Database	The cloud hosted database used to store the information about the fish.

#### IV. RESULT AND DISCUSSION

In this section, we present the performance accuracy of the proposed model. Accuracy is a mechanism for calculating the efficiency of a classification model. Typically, it is presented as a percentage. Accuracy is the value of prediction of which it is equal to the real value and easier to interpret. It is often depicted in graphed to represent the accuracy of a developed model. Whereby a loss function is a prediction value for how much it is varying from true value. It is not represented in percentage; it is the errors sum made for each sample in validation set. The entropy loss and log loss are the most common function for loss. The loss function can be used in regression and classification problems. Fig. 5 is the graph that show the performance of four parameters, namely the accuracy, validation accuracy, loss, and validation loss based on the developed model. When training a model, the accuracy and loss for validation data in the model usually will vary in different situations. Usually, errors should be smaller for each epoch increasing, and accuracy should be greater. Three cases can possibly happen, for the first case, if the loss parameter is start increasing and the accuracy is decreasing, this indicate that the model is not learning. For the second case, if the loss parameters start increasing as well as the accuracy parameter, this may be the cause of diverse probability values or overfitting in cases where softmax is being used in output layer. For third case, if the loss parameter is starts decreasing, the accuracy parameter is starts increasing, this indicate that the built model is learning and working fine. It is clearly showing that in Fig. 5, the proposed model is categorized under the third case.

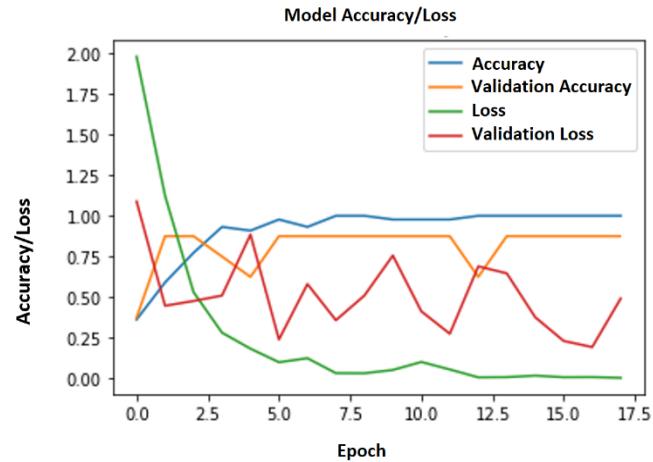


Fig. 5. The Accuracy and Loss Performance of the Proposed Model.

Based on the graph in Fig. 5, after 15 epochs, the model achieved almost 87% accuracy on validation dataset with the training loss 0.0125 and validation loss 0.25. To compare them with the existing works, the experimental result obtained in this study yielded a moderate performance as the accuracy is just between 60 to 80 percent throughout the 15 iterations. This may happen due to some factors such as small size of the data set and lack of image variation in every training sample. However, the accuracy rate obtained is still acceptable as the backend model for the FishDeTec application as shown in Fig. 6. The system is very simple and easy to use. With just one click, the user will be provided with the information about

the fish in English or Bahasa Melayu for every captured fish's image. All user needs to do is to just take picture of the fish and feed it into the system.

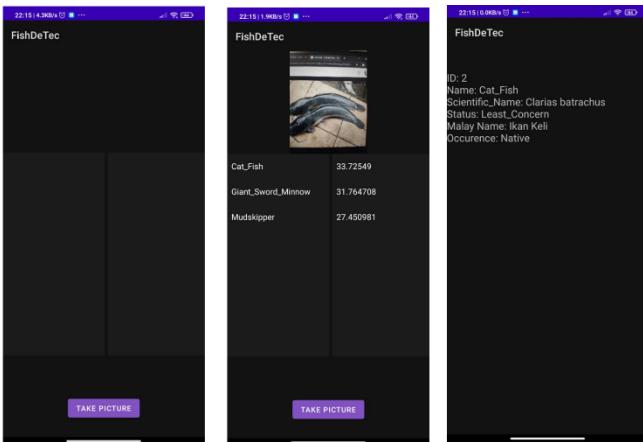


Fig. 6. The Final Look of FishDeTec.

## V. CONCLUSION

In this paper, we have presented the development of FishDeTec, a mobile application for identifying Malaysia freshwater fish. The model for detecting the fish species is developed using the VGG16, a Convolution Neural Network model introduced by K. Simonyan [7]. The TL in VGG16 is used to identify the freshwater fish species. We executed our proposed model using dataset consist of eight different types of freshwater species available in Malaysia, with 178 images in total. To minimize the risk of overfitting in various image variations, we have conducted an augmentation procedure. The foundation of the augmentation techniques used in this research was image transformations such as zoom, rotation, and flipping. The model achieved its accuracy at 60-80 percent, when tested in the eight different types of species freshwater fish. To compare with the existing works, more works need to be done, especially on the validation accuracy. The validation loss need to be reduced. The experimental results show that the pre-trained modeled yielded the moderate performance. For future work, to resolve and reduce the error rate of the result as well as the limited number of images in the data set, pre-training models on the combination of ImageNet and image enhancement could be used to solve the problem. To increase the model validation, more species with a greater number of images for each species are required.

## REFERENCES

- [1] Y. Yong, L. Quek, E. Lim, and A. Ngo, "A case report of puffer fish poisoning in Singapore," Case reports in medicine, vol. 2013, 2013.
- [2] J. K. Sims and D. C. Ostman, "Pufferfish poisoning: emergency diagnosis and management of mild human tetrodotoxinosis," Annals of emergency medicine, vol. 15, pp. 1094-1098, 1986.
- [3] L. Chen, Z. Li, and Z. Zhao, "Forensic medical identification of death due to poisoning of tetrodotoxin in puffer fish," Fa yi xue za zhi, vol. 15, pp. 131-2, 189, 1999.
- [4] W. J. Matthews, Patterns in freshwater fish ecology: Springer Science & Business Media, 2012.
- [5] W. A. Wurts, "Why can some fish live in freshwater, some in salt water, and some in both," World Aquaculture, vol. 29, p. 65, 1998.
- [6] D. A. Konovalov, A. Saleh, M. Bradley, M. Sankupellay, S. Marini, and M. Sheaves, "Underwater fish detection with weak multi-domain supervision," 2019 International Joint Conference on Neural Networks (IJCNN), 2019, pp. 1-8.
- [7] K. M. Knausgård, A. Wiklund, T. K. Sørdsalen, K. Halvorsen, A. R. Kleiven, L. Jiao, et al., "Temperate Fish Detection and Classification: a Deep Learning based Approach," arXiv preprint arXiv:2005.07518, 2020.
- [8] N. E. M. Khalifa, M. H. N. Taha, and A. E. Hassanien, "Aquarium family fish species identification system using deep neural networks," International Conference on Advanced Intelligent Systems and Informatics, 2018, pp. 347-356.
- [9] X. Bai, X. Yang, and L. J. Latecki, "Detection and recognition of contour parts based on shape similarity," Pattern Recognition, vol. 41, pp. 2189-2199, 2008.
- [10] J.-S. Kim and K.-S. Hong, "Color-texture segmentation using unsupervised graph cuts," Pattern Recognition, vol. 42, pp. 735-750, 2009.
- [11] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, et al., "Backpropagation applied to handwritten zip code recognition," Neural computation, vol. 1, pp. 541-551, 1989.
- [12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.
- [13] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," Advances in neural information processing systems, vol. 25, pp. 1097-1105, 2012.
- [14] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," European conference on computer vision, 2014, pp. 818-833.
- [15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770-778.
- [16] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, et al., "Going deeper with convolutions," Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1-9.
- [17] P. Eosina, G. F. Laxmi, and F. Fatimah, "The Sobel Edge Detection Techniques for Freshwater Fish Image Analysis," The 4th International Seminar on Sciences, 2017, p. 83.
- [18] H. Wang, Y. Shi, Y. Yue, and H. Zhao, "Study on Freshwater Fish Image Recognition Integrating SPP and DenseNet Network," 2020 IEEE International Conference on Mechatronics and Automation (ICMA), 2020, pp. 564-569.
- [19] 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), 2017, pp. 1-6.
- [20] F. Rossi, A. Benso, S. Di Carlo, G. Politano, A. Savino, and P. L. Acutis, "FishAPP: A mobile App to detect fish falsification through image processing and machine learning techniques," 2016 IEEE international conference on automation, quality and testing, robotics (AQTR), 2016, pp. 1-6.
- [21] H. Qin, X. Li, J. Liang, Y. Peng, and C. Zhang, "DeepFish: Accurate underwater live fish recognition with a deep architecture," Neurocomputing, vol. 187, pp. 49-58, 2016.
- [22] FishVerify. Species Identification and Regulation Guide. Available: [https://www.fishverify.com/#the\\_app](https://www.fishverify.com/#the_app).
- [23] F. ID+. Fish ID+ Fish Identifier for VIC - Apl di Google Play. Available: <https://play.google.com/store/apps/details?id=com.phonegap.fishidplus&hl=ms&gl=US>.
- [24] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: A comprehensive review," Neural computation, vol. 29, pp. 2352-2449, 2017.