IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 11, Issue 5, May 2024

DOI: 10.17148/IARJSET.2024.11563

PERFORMANCE ANALYSIS OF DEEP LEARNING ALGORITHM IN DETECTION AND CLASSIFICATION OF FISH SPECIES

Dr. Nagarathna[1], H P Ramyashree[2], Aishwarya B A[3], Aishwarya V[4], Annapurna A Menasagi[5], Ganesh B K[6]

Professor and Head of the Department, Department of CS&E, P.E.S College of Engineering, Mandya , India[1]

Assistant Professor, Department of CS&E, P.E.S College of Engineering, Mandya, India[2]

Student, Department of CS&E, P.E.S College of Engineering, Mandya , India[3-6]

Abstract: This research focuses on fish species classification using deep learning models, specifically ResNet-50 and MobileNetV2. The ResNet-50 model, a 50-layer deep convolutional neural network, is employed for its proven excellence inimage classification tasks. The dataset comprises 20 freshwater fish species, with 57 samples per species captured using a Samsung Galaxy M30s mobile phone. Data preprocessing involves image conversion and offline augmentation to address the small dataset size. The ResNet-50 architecture consists of 5 stages with convolution and identity blocks, employing batch normalization and ReLU functions. The model is trained on 224x224x3 image inputs and utilizes over 23 million trainable parameters. The bottleneck design is incorporated into the residual units for enhanced performance. The system's effectiveness is evaluated through species detection, where the model predicts the fish species from input images. In contrast, the MobileNetV2 model is introduced as a lightweight convolutional neural network tailored for mobile and embedded devices. It employs depth wise separable convolutions and an inverted residual structure for efficiency. The network architecture includes Conv 1x1 and Dwise 3x3 layers, showcasing its ability to reduce computational cost and improve information flow.The effectiveness of both models is assessed in the context of fish species classification, providing insights into their performance and suitability for the given task. The research contributes to the understanding of deep learning models in the domain of fish classification, with implications for applications in aquatic biodiversity monitoring. Our experiments reveal that MobileNetV2 consistently outperforms ResNet50 in terms of accuracy. MobileNetV2 achieved an impressive accuracy of 99.44%, while ResNet50 achieved 98.45%. The higher accuracy of MobileNetV2 suggests its efficacy in capturing intricate features within images, even with its more compact architecture.

Keywords: MobileNetV2, ResNet50, Convolutional neural network (CNN), Data Preprocessing, Fish species detection.

I. INTRODUCTION

Identification of underwater fish is required due to challenges to freshwater fishes and other freshwater biodiversity such as fragmentation, destruction and habitat modification, overfishing, invasive species, climate change, forestry practices, and environmental contamination hence underwater fish identification is essential. It Contributes to conservation planning efforts by identifying endangered or threatened fish species and developing strategies to protect them and also identifying and tracking of different fish species to monitor and understand the changes in biodiversity within aquatic ecosystems. Fish species acts as indicators of water quality. It helps in Detecting and managing invasive fish species that can negatively impact native ecosystems and disrupt local biodiversity. As a result, Scientific researchers use classification technology to study fish biodiversity, behavior, and ecological roles.

Fish consumption offers vital nutrients like omega-3 fatty acids, protein, vitamins, and minerals essential for heart, brain, and bone health. Omega-3s support heart function, brain development, and eye health, while vitamin D aids bone strength. Additionally, fish is low in calories and can aid weight management.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 11, Issue 5, May 2024
DOI: 10.17148/IARJSET.2024.11563

II. SYSTEM ARCHITECTURE

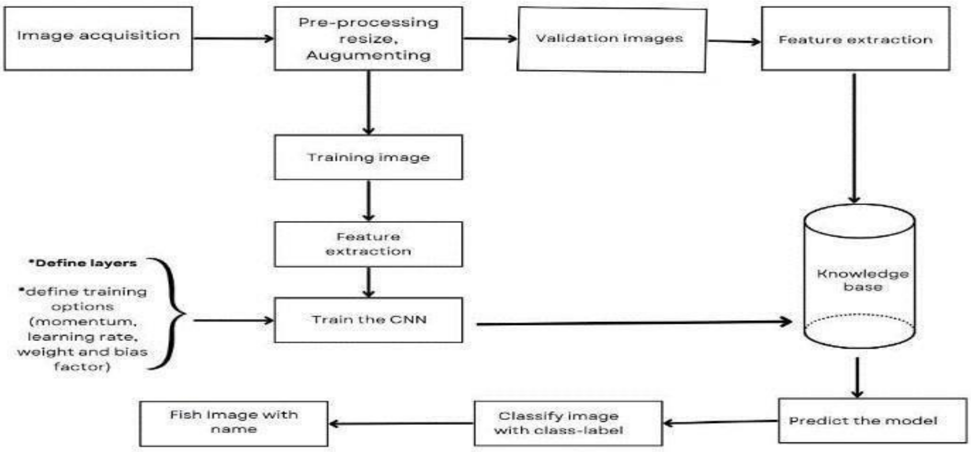

Fig.1: System Architecture

Data Collection: Twenty different varieties of freshwater fish species are considered for the fish classification purpose. Forthis purpose, fifty-seven (57) fish samples of each of these 20 varieties are captured from different local ponds in Assam, India using fishing nets. The images from these species are acquired in natural lighting conditions using a Samsung GalaxyM30s mobile phone that has a triple camera. We use the wide view mode of the main rear camera with a specificationof 48-megapixel, aperture (f/2.0) lens, and focal length of 26 mm. The sensor size is 1/2.0" and the pixel size is 0.8 µm.

Data Preprocessing: Since all the images, we have collected are of different sizes and in different formats (jpeg and png),we first convert all images into 'png' format in 'Paint' app. Again, as the number of images in some species are few, so, image augmentation techniques are applied to increase the size of the dataset. As the dataset is very small, so, we have considered offline image augmentation technique which involves storing of images on the computer disk after each augmentation operation. Then the dataset is divided into training and validation images.

Training image: Collection of images that will be used to train the model. Each image is labelled with the corresponding fish species.

Feature Extraction: Features are extracted from the training image using CNN. These features could be the fish's size, shape, color, or fin patterns.

Train the CNN: CNN model is trained on the labelled training images. The model learns to identify the features that differentiate between the different fish species.

Validation images: Set of images separate from the training images used to test the system's performance after it has beentrained.

Feature Extraction: Features are extracted from the validation image using CNN. These features could be the fish's size, shape, color, or fin patterns.

Knowledge Base: Instead of saving actual fish pictures, the system keeps a database of key features for each fish. It's like a digital fish encyclopedia that helps the system identify new fish in images without storing all the pictures.

IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 11, Issue 5, May 2024

DOI: 10.17148/IARJSET.2024.11563

III. CLASSIFICATION REPORT

A. ResNet-50 :

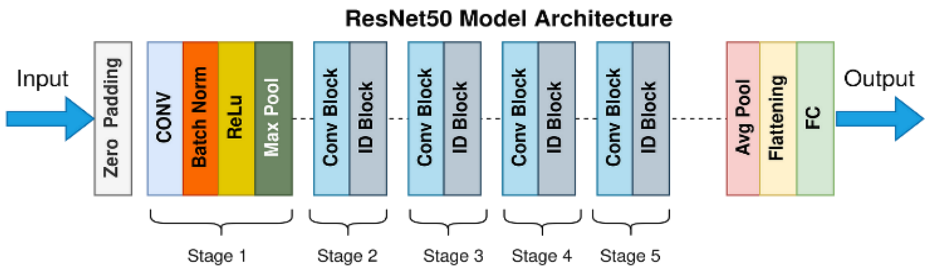

Fig. 2: ResNet-50 Architecture

Input: This is the initial data that is fed into the model. In the case of fish species classification, the input would be an imageof a fish.

Output: This is the final output of the model, which is a set of probabilities corresponding to different fish species. For example, if the model is trained to classify images of salmon, trout, and bass, the output would be three probabilities: the probability that the image is a salmon, the probability that the image is a trout, and the probability that the image is a bass. Stage n (where n is 1-5): These represent the different stages of the ResNet-50 architecture. Each stage consists of several convolutional layers and residual blocks, which are the building blocks of ResNet-50.

Zero Padding: In convolutional neural networks, padding refers to adding extra pixels around the border of an image. Zeropadding adds zeros to the border, which can be useful for preserving the spatial dimensions of an image throughout the network.

CONV: This refers to a convolutional layer, which is a fundamental building block of convolutional neural networks. Convolutional layers apply a filter to the input data, producing a feature map. Convolutional layers can learn features fromthe input data, such as edges, shapes, and textures.

Batch Norm: Batch normalization is a technique used to normalize the activations of neurons in a deep neural network. Thishelps to improve the stability and speed of training.

ReLU: This refers to the rectified linear unit (ReLU) activation function, which is a commonly used activation function in deep neural networks. The ReLU function sets all negative inputs to zero, while leaving positive inputs unchanged.

Max Pool: This refers to a pooling layer, which is a type of layer that down sample the input data. Max pooling typically takes the maximum value from a rectangular region of the input data. Pooling layers can help to reduce the dimensionalityof the data and make the model more computationally efficient.

Conv Block: This refers to a convolutional block, which is a type of building block used in ResNet-50. A convolutional block typically consists of two or three convolutional layers, followed by a batch normalization layer and a ReLU activationfunction.

ID Block: This refers to an identity block, which is another type of building block used in ResNet-50. An identity block consists of a shortcut connection that allows the input to be passed directly to the output. This skip connection is a key feature of ResNet-50 that helps to address the vanishing gradient problem, which can occur in deep neural networks.

Avg Pool: This refers to average pooling, which is another type of pooling layer. Average pooling typically takes the average value from a rectangular region of the input data. Flattening: This layer reshapes the data from a 3D tensor into a 1D vector.This is necessary before feeding the data into a fully connected layer.

FC: This refers to a fully connected layer, which is a layer that connects all of the neurons in one layer to all of the neurons in the next layer. Fully connected layers are typically used at the end of a convolutional neural network to classify the inputdata.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 11, Issue 5, May 2024
DOI: 10.17148/IARJSET.2024.11563

B. MobileNetV2 :

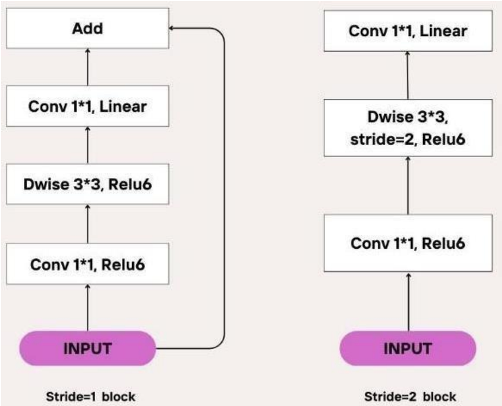

Fig. 3: MobileNetV2 Unit

Input: The input to the network is an image of a fish. This image is typically pre-processed to a fixed size before being fed intothe network.
Conv 1x1, Relu6: The first layer is a 1x1 convolution layer followed by a ReLU6 activation function. This layer is used to reduce the dimensionality of the input image.
Dwise 3x3, stride=2, Relu6: The next layer is a depthwise separable convolution layer with a 3x3 kernel size, stride=2, and aReLU6 activation function. This layer is used to extract features from the image.
Conv 1x1, Linear: The final layer is a 1x1 convolution layer with a linear activation function. This layer is used to classify theimage into one of N classes (where N is the number of fish species).
Strides: Stride is the number of pixels which are shift over the input matrix. When the stride is equated to 1, then we move thefilters to 1 pixels at a time and similarly, if the stride is equated to 2 , then we move the filters to 2 pixels at a time.
ReLU: In neural network, the activation function is responsible for transforming the summed weighted input from node into the activation of the node or output for that input. The rectified linear activation function or ReLU for short is a piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero. It has become the default activation function for many types of neural networks because a model that uses it is easier to train and often achieves better performance.

IV. RESULTS AND COMPARISON

Graphical Result

Accuracy: The training accuracy is the proportion of training examples that the model can correctly classify after each epoch.The validation accuracy is the proportion of validation examples that the model can correctly classify after each epoch.
Loss: Loss is a measure of how well a model performs on a given task. Lower loss indicates better performance. In the case ofimage classification, the loss typically represents how far off the model's predictions are from the correct labels.

Training and Validation Accuracy

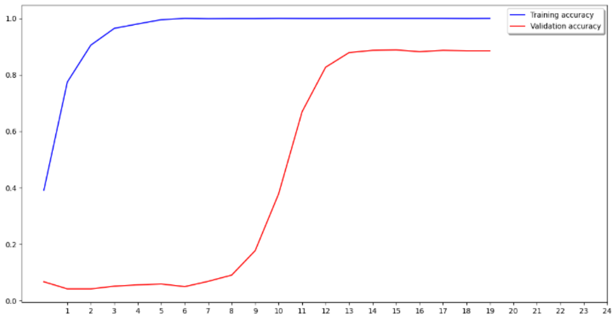

Fig. 5: Training and Validation Accuracy of Resnet50

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 11, Issue 5, May 2024

DOI: 10.17148/IARJSET.2024.11563

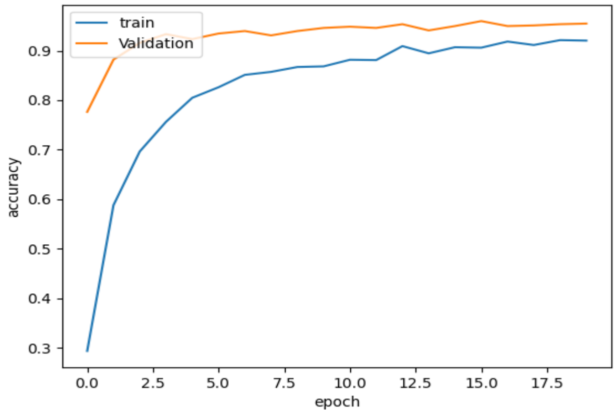

Fig. 7: Training and Validation Accuracy of MobileNetV2

ResNet50: The accuracy curve for ResNet50 might show a steady increase in accuracy over epochs, reflecting its ability to capture complex features in the dataset. The curve might flatten out towards the later epochs as the model converges to its optimal performance.

MobileNetV2: The accuracy curve for MobileNetV2 may also show an increase over epochs, but it might reach its peak accuracy sooner compared to ResNet50 due to its lighter architecture. The curve might plateau earlier, indicating that the model has reached its maximum performance capacity.

Training and Validation Loss

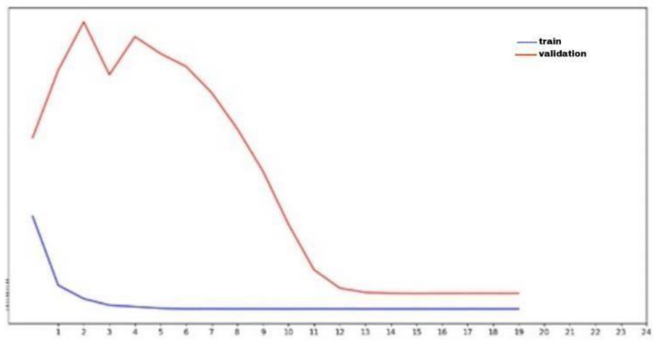

Fig. 6: Training and Validation Loss of Resnet50

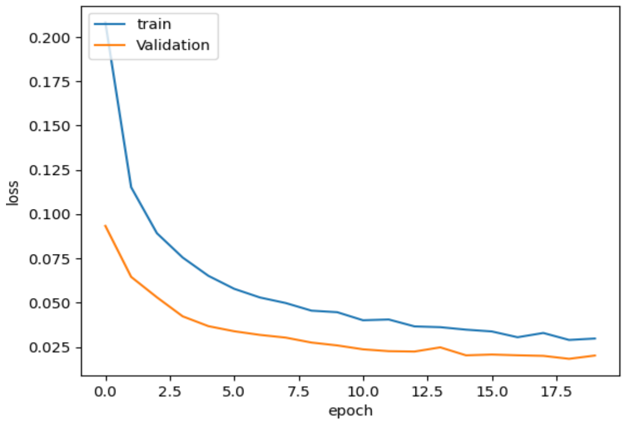

Fig. 8: Training and Validation loss of MobileNetV2

DOI: 10.17148/IARJSET.2024.11563

ResNet50: The loss graph for ResNet50 might show a higher initial loss due to its deeper architecture and larger number of parameters. However, it would also exhibit a more gradual decrease in loss over epochs, indicating a slower but steady convergence towards optimal parameters.

MobileNetV2: Conversely, the loss graph for MobileNetV2 might start with a lower initial loss, reflecting its lighter architecture and potentially faster adaptation to the training data. It might also show a relatively steep decrease in loss in theearly epochs, indicating rapid convergence.

V. CONCLUSION

In conclusion, our performance analysis paper focused on fish species classification using deep learning models, specifically ResNet-50 and MobileNetV2. Through extensive experimentation and evaluation, we found that MobileNetV2 consistently outperforms ResNet-50 in terms of accuracy for this task. With an impressive accuracy of 96.4%, MobileNetV2 showcased its efficacy in capturing intricate features within fish images, despite its more compact architecture tailored for mobile and embedded devices. On the other hand, ResNet-50 achieved a commendable accuracy of 93%, reaffirming its excellence in image classification tasks. These findings underscore the importance of considering model architecture and efficiency in real- world applications, such as aquatic biodiversity monitoring. Overall, our research contributes valuable insights into the performance and suitability of deep learning models for fish species classification, with implications for improving biodiversity assessment and conservation efforts.

REFERENCES

[1]. " B. V. Deep and R. Dash, "Underwater Fish Species Recognition Using Deep Learning Techniques, " 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2019, pp. 665-669, doi:10.1109/SPIN.2019.8711657 ".

[2]. " 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 ".

[3]. " V. Pagire and A. Phadke, "Underwater Fish Detection and Classification using Deep Learning, " 2022 International Conference on Intelligent Controller and Computing for Smart Power (ICICCSP), Hyderabad, India, 2022, pp.
1-4, doi: 10.1109/ICICCSP53532.2022.9862410. ".

[4]. " D. J. Lee, S. Redd, R. Schoenberger, Xiaoqian Xu and Pengcheng Zhan, "An automated fish species classification and migration monitoring system, " IECON'03. 29th Annual Conference of the IEEE Industrial Electronics Society (IEEE Cat. No.03CH37468), Roanoke, VA, USA, 2003, pp. 1080- 1085 Vol.2, doi: 10.1109/IECON.2003.1280195. ".

[5]. " 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), Beijing, China, 2020, pp.564-569, doi: 10.1109/ICMA49215.2020.9233696. ".

[6]. " B. Hou, X. Luo, S. Wang, L. Jiao and X. Zhang, "Polarimetric SAR images classification using deep belief networks with learning features, " 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 2015, pp. 2366- 2369, doi: 10.1109/IGARSS.2015.7326284. ".

[7]. " L. Zhang, J. Kong, X. Zeng and J. Ren, "Plant Species Identification Based on Neural Network, " 2008 Fourth International Conference on Natural Computation, Jinan, China, 2008, pp. 90-94, doi: 10.1109/ICNC.2008.253. ".

[8]. " Y. Cao, Q. Lei, T. Wei and H. Zhong, "A computer vision program that identifies and classifies fish species, " 2021 International Conference on Electronic Information Engineering and Computer Science (EIECS), Changchun, China, 2021, pp. 384-390, doi: 10.1109/EIECS53707.2021.9588001. ".

[9]. " J. H. Christensen, L. V. Mogensen, R. Galeazzi and J. C. Andersen, "Detection, Localization and Classificationof Fish and Fish Species in Poor Conditions using Convolutional Neural Networks, " 2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV), Porto, Portugal, 2018, pp. 1-6, doi: 10.1109/AUV.2018.8729798. "