International Journal Software Engineering and Computer Science (IJSECS) 3 (3), **2023, 270-280**
Published Online December 2023 in IJSECS (http://www.journal.lembagakita.org/index.php/ijsecs)
P-ISSN : 2776-4869, E-ISSN : 2776-3242. DOI: https://doi.org/10.35870/ijsecs.v3i3.1763.

Comparative Analysis VGG16 Vs MobileNet Performance for Fish Identification

Djarot Hindarto *
Informatics Study Program, Faculty of Communication and Informatics Technology, Universitas Nasional, City of South Jakarta, Special Capital Region of Jakarta, Indonesia.
E-mail: djarot.hindarto@civitas.unas.ac.id

Received: 30 October 2023; Accepted: 20 November 2023; Published: 10 December 2023.

Abstract: This research aims to conduct a comparative evaluation of the efficacy of two neural network architectures in the field of fish identification through the utilization of supervised learning techniques. The evaluation of VGG16 and MobileNet, which are prominent deep learning architectures, has been conducted about their speed, accuracy, and efficiency in resource utilization. To assess the classification performance of both architectures, we employed a dataset encompassing diverse fish categories. The findings indicated that the VGG16 model demonstrated superior accuracy in fish classification, albeit due to increased computational time and resource utilization. On the contrary, MobileNet exhibits enhanced speed and efficiency, albeit at a marginal cost to its accuracy. The findings of this study have the potential to inform the selection of deep learning models for fish recognition scenarios, considering the specific requirements of the task, such as prioritizing accuracy or efficiency. The findings mentioned above offer significant insights that can be utilized in the advancement of Artificial Intelligence (AI)-based applications within the domains of fisheries resource management and environmental monitoring. These applications specifically necessitate precise and effective fish recognition capabilities. The comparison findings indicate that the accuracy achieved by VGG16 was 0.99, whereas MobileNet also attained an accuracy of 0.99.

Keywords: Accuracy; Fish Classification; Supervised Learning; MobileNet; VGG16.

1. Introduction

Fish play a significant role in marine ecosystems and are a crucial protein source for human populations worldwide. Nevertheless, the pressing need for monitoring and sustainable management of fisheries resources has arisen due to overfishing, climate change, and the detrimental impact on the marine environment. The identification and monitoring of fish species captured constitute a fundamental component within the realm of fisheries resource management. The utilization of fish identification and classification technology has emerged as a crucial foundation for these endeavors, facilitating enhanced efficacy and precision in monitoring and management practices. The swift advancement of technology has given rise to deep learning methodologies and intense neural networks [1], which have demonstrated remarkable efficacy in image analysis-based fish species identification. The utilization of supervised learning methodologies in conjunction with deep learning techniques has facilitated the capacity of computer systems to comprehend and categorize various fish species with progressively enhanced levels of precision. This enables researchers, fisheries resource managers, and environmental monitors to gather more comprehensive and precise data.

Within this particular context, two deep learning architectures [2] that have garnered considerable interest and recognition are VGG16 [3] and MobileNet. The VGG16 [4] model, renowned for its exceptional precision, is a convolutional neural network famous for its capacity to discern intricate characteristics within visual data. On the other hand, MobileNet is recognized for its rapidity and effectiveness in utilizing computational resources. It is a specialized architecture tailored for mobile devices and systems with restricted resources. Both architectural approaches possess distinct advantages and disadvantages, and their application within the domain of fish identification gives rise to significant inquiries regarding their efficacy and pertinence. In recent years, there has been considerable progress in deep learning-based image recognition technology, which has facilitated the ability of systems to recognize and classify various fish species accurately. Nevertheless, there exists a scholarly discourse regarding the optimal timing and efficacy of employing distinct neural network architectures, such as VGG16 [5] and MobileNet, in specific scenarios. To effectively examine this inquiry, it is imperative to conduct a performance evaluation that compares the two architectures in the context of fish identification.

This study aims to elucidate the growing significance of deep learning technology in the identification and classification of fish species across multiple domains. Additionally, it explores the distinct roles that VGG16 and MobileNet, two diverse deep-learning architectures, can assume in this context. This research aims to enhance the comprehension of the performance and significance of various architectures, thereby offering improved guidance for the application of image recognition technology in fisheries resource management, environmental monitoring, and marine biology research.

The primary concern in fish identification through citrate lies in attaining a harmonious equilibrium between classification precision and computational efficacy. The demand for precise and dependable identification remains significant, particularly within fisheries resource management and marine environmental monitoring, which exhibit remarkable intricacy. The accuracy of fish identification is crucial for determining the conciseness and lack of embellishment in a statement. Incorrect identification can result in ineffective catch management. In distributed network systems and mobile devices, utilizing power-efficient computing resources becomes crucial due to the inherent limitations in available computing resources. The over utilization of resources, coupled with inefficient cost management and excessive power consumption, can lead to a deceleration in processing speed. Therefore, the integration of high precision in fish identification through the utilization of highly efficient computing is a crucial element in the advancement of image-based fish recognition and classification systems. By addressing these requirements, technological solutions like VGG16 [6] and MobileNet effectively achieve a harmonious equilibrium between precision and resource utilization. Each approach presents a distinct methodology for addressing this intricate issue.

The principal objective of this research is to perform an exhaustive comparative evaluation of the efficacy of two deep learning architectures, namely MobileNet and VGG16, in the context of fish identification. The utilization of image recognition for the identification of fish species holds significant importance in a range of applications, such as fisheries resource management, environmental monitoring, and marine biology research. The primary objective of this study is to examine the efficacy of VGG16 and MobileNet [7] in addressing two prominent and frequently opposing challenges in fish recognition, namely accuracy and computational efficiency.

Initially, we shall assess the efficacy of these two architectures in terms of accuracy by scrutinizing their ability to identify and categorize various fish species with minimal errors. In this context, precision holds utmost importance, particularly within fisheries resource management, as misidentification can significantly influence the formulation and implementation of management policies and strategies. Furthermore, an assessment will be performed to examine the computational efficiency of MobileNet and VGG16, including aspects like optimized resource utilization and processing speed. This is particularly pertinent in situations where computing resources are scarce, such as in distributed network systems or mobile devices, and where expeditious and energy-efficient processing is of utmost importance. This research aims to enhance our understanding of the accuracy and computational efficiency of VGG16 and MobileNet architectures. By doing so, it will offer more precise recommendations regarding the appropriate utilization of these architectures in the domain of fish identification. Additionally, the research will explore the potential contributions of VGG16 and MobileNet [8] towards monitoring objectives and the sustainable management of fisheries resources.

The following are the research questions that are raised by this investigation:

In supervised learning, what is the comparative performance of MobileNet and VGG16 in categorizing and identifying fish species? (Research Question 1). How do MobileNet and VGG16 compare regarding classification accuracy when it comes to the recognition of fish species? (Research Question 2). Through resolving these inquiries, this study will establish a solid groundwork for identifying deep learning models that are well-suited to requirements in the domain of fish species recognition while also considering factors related to precision.

2. Research Method

2.1. Dataset

In the realm of fish identification, the incorporation of publicly available datasets is an essential undertaking in the advancement of deep learning models for the categorization of fish species. Consisting of approximately 18,000 images, this dataset is divided into nine discrete categories, each corresponding to a distinct fish species. The dataset comprises a diverse collection of images representing a wide range of fish species, sizes, and lighting conditions. Utilizing this extensive and diverse dataset for training and evaluating fish classification models with a notable level of accuracy is a highly advantageous undertaking. The importance of public datasets in this study cannot be emphasized enough. This extensive dataset facilitates the training of deep learning models by incorporating a wide range of data, encompassing diverse fish species and variations in lighting conditions, orientation, and size. By comprehensively aggregating visual data, the model can acquire a more holistic comprehension of various attributes and effectively discern unique fish species across multiple scenarios. The careful consideration of introducing fish into dynamic natural environments, such as open water or submerged conditions, is of utmost importance.

The presence of nine distinct classes in the dataset poses a significant classification challenge for this research endeavor. The deep learning model should demonstrate the capability to distinguish between different types of fish based on their respective classes. Moreover, this extensive dataset allows researchers to evaluate the model's ability to overcome the challenges of differentiating between similar or dissimilar fish species. As a result, researchers will be able to assess the effectiveness and precision of the model in real-world scenarios involving fish species that share common characteristics. Furthermore, the availability of these extensive public datasets enables developers and the scientific community to assess and compare different strategies for fish recognition. This facilitates collaboration and progress in deep learning and image recognition. This study aims to enhance the understanding of the specific application of deep learning models in fish species identification by integrating pre-existing datasets with state-of-the-art methodologies and models. In a broader context, utilizing a publicly accessible dataset of 18,000 images and encompassing nine distinct fish categories provides a solid foundation for the present inquiry. This dataset enables the training, evaluation, and authentication of deep learning models, as well as the assessment of their ability to overcome the challenges associated with fish species classification. The model's utility is heightened in various fish introduction scenarios, such as fisheries resource management, environmental monitoring, and marine biology research, owing to diverse data encompassing a broad spectrum of fish species.

2.2. VGG16

The VGG16 model, created by a team of researchers from the University of Oxford, represents a significant advancement in the convolutional neural networks (CNNs) [14] and the domain of image recognition. The model is renowned for its considerable depth and intricacy, rendering it highly productive in tasks on object recognition and classification. The VGG16 architecture is composed of a total of 16 layers, comprising 13 convolutional layers and three fully connected layers. The convolutional layers in this network employ various filters to identify distinct visual attributes within images, such as edges, corners, and texture. Furthermore, as the network progresses deeper, the detected features become increasingly intricate. Incorporating Rectified Linear Unit (ReLU) activation after each convolution layer facilitates the introduction of non-linearities within the data. This enables the neural network to comprehend intricate patterns present in the image. One notable benefit of the VGG16 model is its understanding of complex visual characteristics. This attribute renders it highly proficient in object classification and image recognition, specifically in the domain of fish type identification. Its ability to discern intricate and abstract features in fish images contributes to its effectiveness. Despite the demonstrated efficacy of VGG16 in these tasks, it has its limitations, notably its substantial parameter counts and resource-intensive computational demands. In specific application contexts characterized by limited resources, such as mobile devices or systems with constrained computing power, this may result in reduced efficiency. In fish recognition, the VGG16 model engages in the classification process. This involves taking fish images as input and conducting hierarchical feature extraction through convolution layers and ReLU activations. Subsequently, the extracted features are connected to fully connected layers to ascertain the most appropriate fish type. This illustrates the utilization of VGG16 in object recognition tasks, showcasing its significant potential in diverse domains such as fisheries resource management and environmental monitoring.

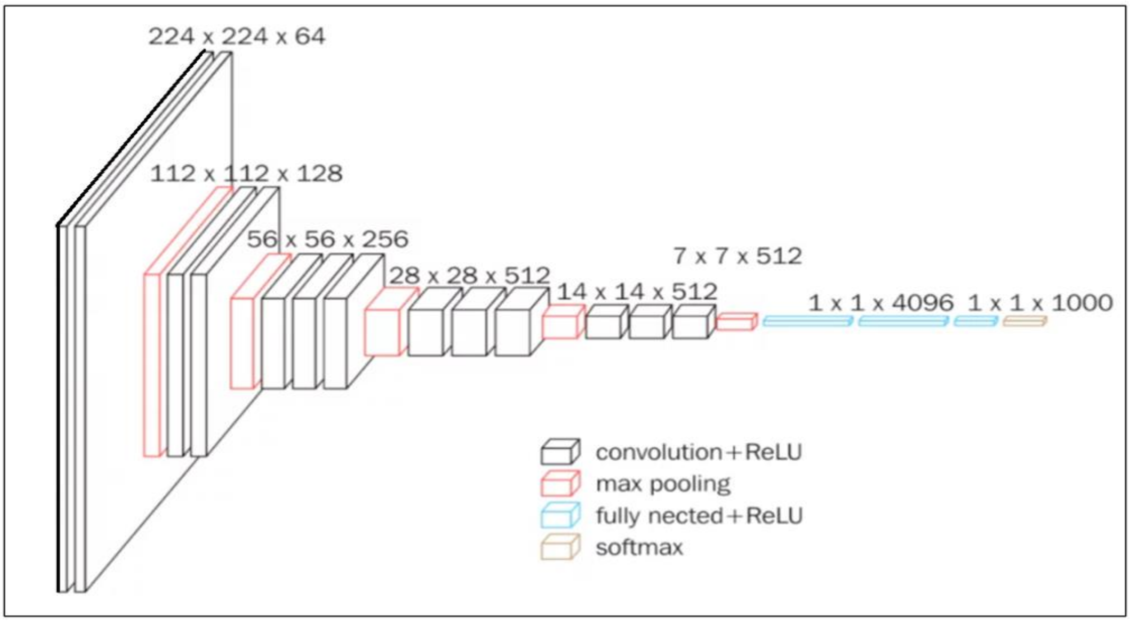

Figure 1. VGG16 Architecture

Figure 1, the cov1 layer receives a 224 x 224 RGB image as input. The image is processed by a series of convolutional (conv.) layers employing filters with a tiny receptive field of 3×3. This size is considered optimal for capturing the concepts of left/right, up/down, and center. One of the configurations additionally incorporates 1×1 convolution filters, which may be conceptualized as a linear process involving a non-linear transformation of the input channels. The convolution stride is maintained at 1 pixel, and the spatial padding of the conv. Layer input is designed to maintain the spatial resolution even after convolution, for instance, for 3×3 conv. Layers, the padding is set to 1 pixel. Spatial pooling is executed by utilizing five max-pooling layers, which are positioned after a subset of the conv. Layers (max-pooling does not follow all conv. layers). Maximal pooling is executed across a 2×2-pixel grid, utilizing a stride of 2. A stack of convolutional layers, the depth of which varies across architectures, is followed by three Fully-Connected (FC) layers: the initial two FC layers each consist of 4096 channels, while the third FC layer executes 1000-way ILSVRC classification and thus comprises 1000 channels (one for each class). The soft-max layer constitutes the last layer. In every network, the fully connected layers are configured identically. Each hidden layer is endowed with non-linear rectification (ReLU). It is worth mentioning that except for one network, none incorporate Local Response Normalization (LRN). While LRN does not enhance the performance of the networks on the ILSVRC dataset, it does result in increased computation time and memory consumption.

The classification of objects, specifically fish species, is executed by VGG16 through the subsequent procedures:

1. Feature Extraction: Initially, feature extraction is executed on fish images using VGG16. The initial convolution layer detects essential image characteristics, including edges and the base color. As the network deepens, the convolution layers progressively increase in complexity to enable the detection of ever more intricate features, such as the texture and shape of fish.
2. ReLU Activation: ReLU activation is employed after the convolution layer to eliminate non-linearities present in the data. This enables the network to tackle classification problems that are more intricate and abstract.
3. Pooling: The image dimensions are reduced using a pooling layer after the ReLU activation layer. By decreasing the number of parameters, the network must assimilate, it can concentrate on critical features present in the image.
4. Fully Connected Layers: Subsequently, the fully connected layers receive the outcomes of the feature extraction process. The network then classifies or sorts fish according to the extracted characteristics. By employing a supervised learning algorithm to train these fully connected layers, the model can identify patterns associated with the specific fish species.
5. SoftMax Activation Function: The SoftMax activation function generates predicted probabilities for each fish category in the output layer. The category exhibiting the most significant probability represents the model's final prediction.

2.3. MobileNet

MobileNet is a significant milestone in convolutional neural networks (CNN) [15], having been purpose-built to cater to the requirements of applications running on mobile devices and distributed systems characterized by restricted computational resources. The MobileNet architecture was constructed by a group of Google Research researchers, who primarily prioritized two critical factors: classification accuracy and computational efficiency. MobileNet provides technologically significant solutions, particularly when mobile devices are becoming more prevalent daily. An aspect that sets MobileNet apart is its methodology for optimizing computational efficiency. The MobileNet architecture acknowledges that numerous image-centric tasks, such as object recognition, frequently necessitate execution on mobile devices or distributed systems characterized by constrained computational capabilities. To surmount these challenges, MobileNet implements several optimization strategies that enhance the speed and power efficiency of image recognition. Utilizing separable depth wise convolution, which decreases the number of parameters requiring learning while maintaining the accuracy of the recognition outcomes, is a crucial technique. Additionally, the implementation of bottleneck layers aids in the reduction of data dimensions, thereby preserving efficiency.

In addition to its high efficiency, MobileNet demonstrates commendable accuracy in classification. Despite its emphasis on efficiency, MobileNet maintains a praiseworthy ability to identify and categorize objects within images accurately. MobileNet exhibits reliability in object recognition, encompassing intricate tasks like identifying fish species. MobileNet is a highly pertinent option for many applications, including environmental monitoring and fisheries resource management, due to its effective and precise operation. MobileNet executes classification in the domain of fish species identification via a succession of procedures. MobileNet obtains significant features from images of fish by employing convolutional layers. These attributes comprise data on the form, sheen, and additional visual qualities that distinguish the fish species. By introducing non-linearity into the feature extraction process via ReLU activation, the model can comprehend more intricate patterns within the image.

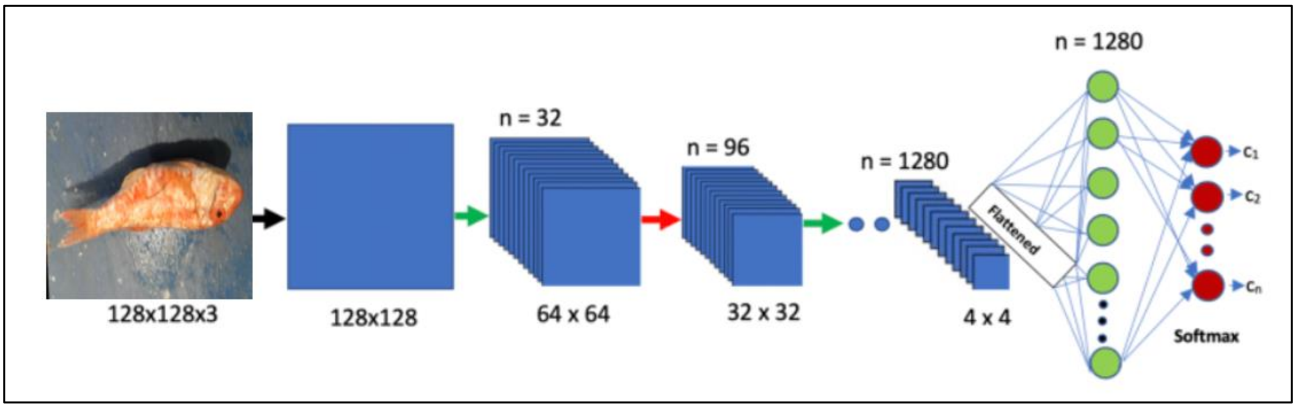

Figure 2. Architecture MobileNet

The MobileNet framework employed in the experiments incorporates Convolution and Classifier layers, as illustrated in Figure 2. The MobileNet framework is precisely engineered to support applications running on mobile devices and systems with restricted computational capabilities. The main job of the Convolution Layer is to pull out basic features from fish images. These characteristics comprise data on the form, sheen, and additional visual attributes employed to classify the fish species. A convolution layer generates a more compact and abstract representation of an image's features. Following this, the extracted feature representation is passed through the Classifier layer, which classifies the most appropriate fish species according to these features. Through supervised learning, this layer is taught to recognize patterns associated with fish species. MobileNet can generate accurate final predictions regarding the species of fish depicted in an image by utilizing the Classifier layer. Combining Convolution and Classifier layers, the MobileNet architecture generates an efficient and effective model for classifying fish types, which is crucial for environmental monitoring and fisheries resource management, among other applications.
In addition, the function of the bottleneck layer is to reduce the size of the preceding data and preserve computational efficiency. Following this, the obtained outcomes are transmitted to the fully connected layers, where the model uses the extracted features to determine the most appropriate fish species. By training these layers with a supervised learning algorithm, patterns associated with fish species are identified. A softmax activation function is subsequently implemented in the output layer to produce predicted probabilities for every fish category. The category exhibiting the most significant probability represents the model's final prediction. MobileNet's effective methodology for image recognition serves as a robust resolution to the computational efficiency obstacles encountered in many mobile device applications and distributed systems. Its exceptional classification accuracy renders it highly applicable to precise tasks, such as identifying fish species. MobileNet's capacity for rapid and efficient object recognition confers significant promise across diverse domains, encompassing sustainable fisheries resource management and environmental monitoring. MobileNet is a prominent accomplishment within deep learning, offering inventive and pertinent resolutions in the contemporary digital age.

3. Result and Discussion

3.1 Results

The outcome of this experimental study yields a classification model that demonstrates a commendable level of accuracy in identifying various fish types. Through the utilization of the pre-trained VGG16 model, this model demonstrates the capability to discern visual distinctions among closely related fish species. The findings of this study possess extensive implications across diverse domains, encompassing fisheries resource management, environmental monitoring, and marine biology research. In the conducted experiment, the achieved training accuracy was 0.99. In general, the conducted classification experiments involving the utilization of Pretrained VGG16 and fish datasets serve as a manifestation of the efficacy of transfer learning within the domain of fish species identification. This exemplifies the ability to modify pre-existing models to tailor them to datasets, resulting in elevated levels of accuracy in the classification of fish species. This approach enables optimizing computing resource utilization while preserving robust classification capabilities, rendering it a pertinent solution for diverse applications involving the recognition of fish species. The experiment can be seen in Figure 3.

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
vgg16 (Functional)           (None, 6, 6, 512)         14714688
_________________________________________________________________
global_average_pooling2d_1 ( (None, 512)               0
_________________________________________________________________
flatten_1 (Flatten)          (None, 512)               0
_________________________________________________________________
dense_1 (Dense)              (None, 9)                 4617
=================================================================
Total params: 14,719,305
Trainable params: 4,617
Non-trainable params: 14,714,688
_________________________________________________________________
```

Figure 3. Architecture VGG16

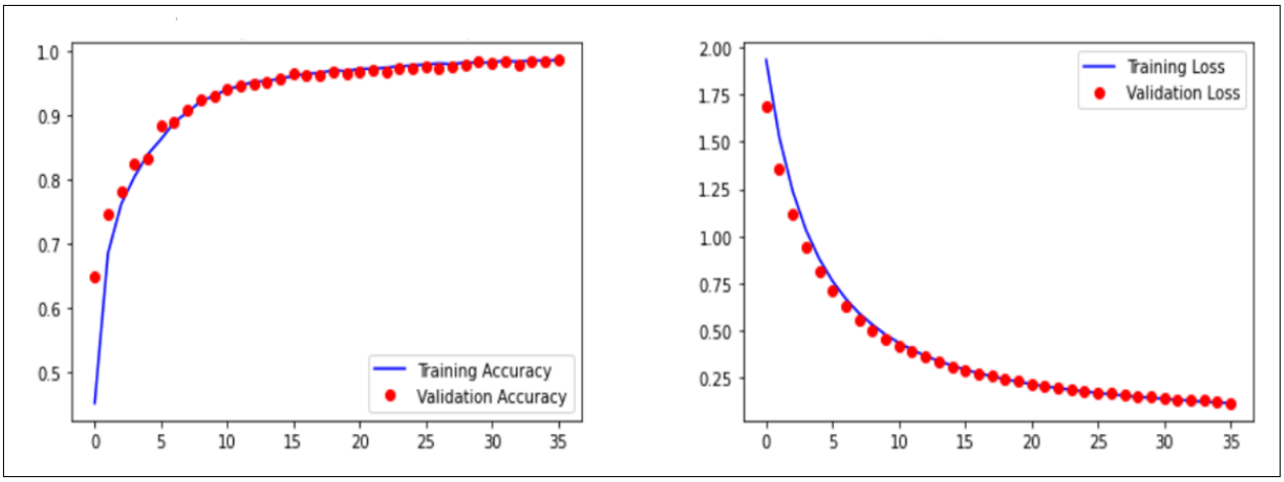

Figure 4. Accuracy and Loss Performance VGG16

Figure 4 depicts the performance of the VGG16 model in classification experiments, exhibiting an accuracy level of 0.99 and a loss value of 0.1153. The findings of this study demonstrate that the VGG16 model shows a high level of efficacy in accurately identifying and categorizing various fish species within the dataset employed. Proximity to 1, or 99%, accuracy rate denotes a minimal number of errors the model makes in classifying fish species. This exemplifies the considerable degree of precision in the process of classification. Furthermore, the observed low loss value of 0.1153 serves as an indicator of the model's proficiency in accurately assessing the proximity between its predictions and the proper labels. Decreased loss value indicates a higher alignment between the model's predictions and the actual labels. The VGG16 model exhibits exceptional proficiency in assessing classification accuracy due to its low loss value.

The findings suggest that transfer learning with VGG16 can be a potent approach for recognizing fish species. The model exhibits a notable degree of precision and minimal margin of error, suggesting its proficiency in identifying different fish species across diverse scenarios. Therefore, the modified VGG16 model has the potential to be a significant and applicable solution in a range of contexts, particularly in the fields of fisheries resource management and marine biology research.

```
inputs = pretrained_model.input

x = tf.keras.layers.Dense(128, activation='relu')(pretrained_model.output)
x = tf.keras.layers.Dense(128, activation='relu')(x) #relu is the activation fun
ction for neurla network task

outputs = tf.keras.layers.Dense(9, activation='softmax')(x) #softmax is the acti
vation function for classiciation task

model = tf.keras.Model(inputs=inputs, outputs=outputs)

model.compile(
    optimizer='adam',
    loss='categorical_crossentropy',
    metrics=['accuracy']
)
```

Figure 5. Scripts for MobileNet

The utilization of scripts or scripts associated with the MobileNet implementation in the experiment is illustrated in Figure 5. The MobileNet model is initialized, trained, and evaluated using this script to classify fish species. Critical operations are performed by this script, including loading a pre-configured MobileNet architecture, configuring training parameters, and accessing a pre-processed fish dataset. During the MobileNet model training procedure, the model is optimized for precisely identifying fish species. Furthermore, this script evaluates the MobileNet model's performance, generating confusion matrices, accuracy, and loss, quantifying the error rate and proportion of correct classifications in fish type identification. The scripts in question are critical in automating the MobileNet model training and testing procedure. The outcomes derived from the execution of MobileNet via this script offer valuable insights into the model's capability to classify fish species accurately. This MobileNet implementation is particularly applicable to the recognition of fish species and is capable of delivering pertinent results for a variety of applications. This script enables the investigation, advancement, and implementation of MobileNet models across diverse practical domains, potentially enhancing comprehension and governance of fishery resources and optimizing environmental monitoring.

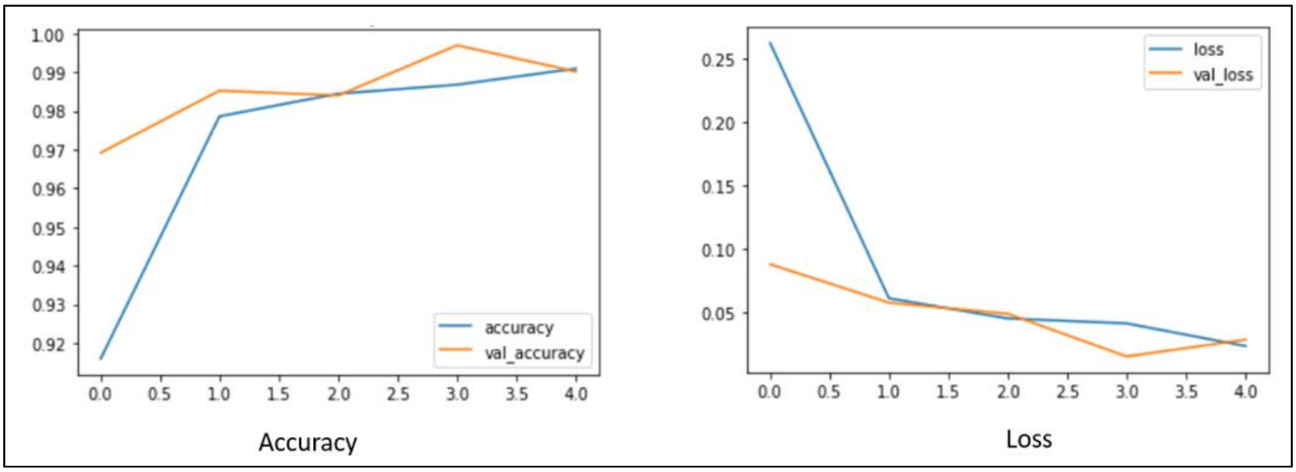

Figure 6. Performance for MobileNet

The outcomes of the MobileNet model's performance in the fish species classification experiment are depicted in Figure 6. The outcomes encompass several metrics that offer a comprehensive comprehension of the MobileNet model's performance in this classification endeavor. To begin with, MobileNet model accuracy is a critical metric to assess the model's proficiency in accurately classifying different fish species. With an accuracy of 0.99, the model demonstrated a capability to classify fish species with an approximate 99% degree of accuracy. MobileNet is an exceptionally effective model for identifying fish species, as shown by these results. Besides accuracy, loss is an additional significant metric. Loss is an indicator of the model's performance in estimating the proximity of the predictions to the actual labels. Within this framework, the MobileNet model achieves a low loss value of 0.1253, indicating its capability to generate predictions that closely approximate the existing labels. A significant indicator of classification quality is an ordinary loss.
Additional metrics, such as a confusion matrix, which offers a more comprehensive understanding of the MobileNet model's capability to classify various fish species, might be incorporated into Figure 5. A confusion matrix can identify the errors the model commits, such as whether specific fish species are misclassified with greater frequency. This data may facilitate a more comprehensive comprehension of the model's performance within particular contexts. Figure 5 illustrates that MobileNet is an exceptionally effective model for identifying fish species. The model's robust classification

capabilities are inferred from its high accuracy and low loss value. The implications of the study's findings have practical significance in various fields, including marine biology research, environmental monitoring, and fisheries resource management. The utilization of MobileNet plays a crucial role in the conservation of fisheries ecosystems and the marine environment by effectively and precisely identifying various fish species.

3.2 Discussion

In supervised learning, what is the comparative performance of MobileNet and VGG16 in categorizing and identifying fish species? (Research Question 1).

Dalam konteks pembelajaran terawasi, pertanyaan penelitian utama yang diajukan adalah sejauh mana perbandingan kinerja antara model MobileNet dan model VGG16 dalam tugas kategorisasi dan identifikasi spesies ikan. The comparison of performance is of utmost importance in evaluating and comprehending the capabilities of both deep learning architectures in the task of classifying fish species. VGG16 and MobileNet are two distinct deep-learning architectures that exhibit different characteristics. VGG16 is renowned for its high accuracy in image recognition and its ability to comprehend complex image features. However, VGG16 also shows weaknesses in terms of high computational resource usage and many parameters. On the other hand, MobileNet is designed explicitly for computational efficiency, employing optimization techniques that enable fast and power-efficient image recognition. The advantage of MobileNet lies in its efficiency, making it highly relevant in applications for mobile devices or systems with limited resources. The emergence of this research question is prompted by the divergent implications of employing both deep learning architectures in the context of fish species recognition. Dalam konteks pengelolaan sumber daya perikanan, pemantauan lingkungan, dan penelitian biologi laut, penentuan yang akurat terhadap identitas spesies ikan memiliki signifikansi yang tinggi. Kesalahan dalam klasifikasi jenis ikan dapat berdampak signifikan terhadap kebijakan dan tindakan pengelolaan. Hence, the primary consideration lies in the ability of both architectures to identify fish species with minimal error rates accurately.

However, utilizing computational resources is also a crucial consideration, particularly in situations where computational resources are limited, such as in mobile devices or distributed network systems. MobileNet, with its computationally efficient approach, can be considered a more relevant option as it can maintain acceptable performance within limited resources. Pertanyaan penelitian ini akan membantu dalam mengevaluasi konteks di mana penggunaan model VGG16 atau MobileNet akan lebih relevan dan efektif dalam melakukan identifikasi spesies ikan. Sebagai hasil dari penelitian ini, diharapkan akan diperoleh pemahaman yang lebih mendalam mengenai perbandingan kinerja kedua arsitektur ini dalam konteks identifikasi jenis ikan. The findings of this research will provide more explicit guidance to decision-makers, researchers, and practitioners in the monitoring of fisheries resources and environmental management. This will enable them to select the most suitable deep learning architecture according to their needs to maintain the sustainability of fisheries ecosystems and marine environments. Dalam konteks penting identifikasi jenis ikan, penelitian ini diharapkan dapat memberikan kontribusi yang signifikan dalam pemahaman dan pengembangan aplikasi deep learning.

How do MobileNet and VGG16 compare regarding classification accuracy when it comes to the recognition of fish species? (Research Question 2).

The comparison of classification accuracy between MobileNet and VGG16 holds significant implications within the domain of fish species recognition. Research Question 2 pertains to evaluating the classification accuracy of two deep learning architectures, which holds significance in assessing the efficacy and suitability of both models for fish-type identification. The VGG16 model has gained recognition for its exceptional accuracy in image recognition, particularly in tasks involving object classification. The capacity to comprehend intricate visual characteristics renders it highly appropriate for identifying fish species. In numerous instances, the VGG16 model demonstrates a remarkable capacity to distinguish between various fish species accurately. This can be attributed to its exceptional depth and capability to extract abstract image features. The VGG16 model has a notable track record of attaining excellent levels of accuracy across a diverse range of image recognition tasks, rendering it a prominent selection within the realm of deep learning. In addition, MobileNet demonstrates a commendable capability in accurately identifying various fish species. Despite being primarily designed for computational efficiency, MobileNet manages to maintain a considerable level of accuracy. MobileNet demonstrates effective classification capabilities by leveraging optimization techniques such as separable depthwise convolution and bottleneck layers. MobileNet is a highly pertinent choice in scenarios with constraints on computing resources, such as mobile devices or systems with limited power availability.

Nevertheless, assessing the accuracy comparison between these two architectures is a complex task. Several factors contribute to the variability of classification accuracy results:

1. The number of distinct fish classes
2. The size of the dataset
3. The existence of variations within the images under examination

The VGG16 model, due to its increased depth and complexity, exhibits superior performance in identifying fish species characterized by subtle visual distinctions. However, the utilization of this model necessitates a higher allocation of computational resources. Conversely, MobileNet can be a more efficient alternative in scenarios where achieving exceptionally high accuracy is only sometimes the foremost objective. The significance of Research Question 2 lies in its ability to elucidate the circumstances under which each deep learning architecture can yield enhanced advantages in fish species identification. In specific contexts, such as in the field of aquatic or environmental monitoring, achieving a high degree of accuracy is of paramount importance. In scenarios necessitating prompt identification of mobile devices or distributed systems with constrained computational capabilities, the determining factor lies in MobileNet's efficiency and capacity to deliver a satisfactory level of accuracy.

Research Question 2 aims to ascertain the comparative utility of different architectural approaches in identifying fish species within the framework of scientific inquiry and practical implementation. This publication will guide researchers, policymakers, and practitioners across diverse disciplines, encompassing fisheries resource management, marine biology research, and environmental monitoring. By acquiring a more comprehensive comprehension of the respective merits and drawbacks of the classification accuracy of VGG16 and MobileNet, individuals can make more knowledgeable and productive decisions when employing deep learning technology. This, in turn, aids in preserving the sustainability of fisheries ecosystems and the marine environment. Therefore, the significance of Research Question 2 lies in its contribution to the development of recommendations and decision-making processes on the implementation of deep learning technology for fish species recognition.

4. Related Work

Within the scope of this study, a literature review, commonly referred to as "Related Work," assumes a pivotal role in comprehending the theoretical underpinnings and prior investigations on the application of deep learning architectures, specifically MobileNet and VGG16, in the domain of fish species identification. A comprehensive overview of the methodologies and findings of prior studies will establish the essential context for assessing and appraising the advancements made by this research. This study employs deep learning with a Convolutional Neural Network (CNN) model to detect microalgae in light and scanning electron microscope images. The results showed that light and electron microscope images could be classified with 99% accuracy. The VGG16 and EfficientNetV2 models were the most accurate. Interestingly, the cheaper light microscope method worked better than the more expensive electron microscope at identifying algae [9]. The objective of this study is to identify plant diseases automatically using deep learning and MobileNet-V2, which has been improved with location-based soft attention. The experimental findings indicate that open-source datasets achieved an average accuracy rate of 99.71%, with an additional average accuracy of 99.13% under conditions of cluttered backgrounds. The efficacy and competitiveness of this method in the identification of plant diseases are evident [10]. Transfer learning reuses trained models. It is used in image classification, prediction, and natural language processing, with deep learning models like MobileNet, MobileNetV2, VGG16, VGG19, and ResNet50 outperforming machine learning on big data. ResNet50 is most accurate with ImageNet, but MobileNetV2 performs best with fewer parameters. Transfer learning could improve natural language processing accuracy [11]. This research identifies ore deposits as a critical mining task. Data augmentation and transfer learning with CNN models like MobileNet improve our dataset classification accuracy. Transfer learning achieves 94% accuracy with MobileNet and 96% with SENet. Automatic ore classification with high accuracy using less training data is possible with this method [12]. This study uses deep learning to detect late and early blight in potato plants. Four deep-learning models were trained, with VGG16 having the highest accuracy (92.69%). After fine-tuning the VGG16 model, it classified late and early blight diseases with 97.89% accuracy. This study details the customized VGG16 model architecture and compares it to other methods [13].

5. Conclusion

An examination of the performance of VGG16 and MobileNet in the domain of fish species identification revealed that both models achieved an exceptionally high degree of accuracy, as measured by a value of 0.99. Regarding fisheries resource management, environmental monitoring, and marine biology research, this signifies that both VGG16 and MobileNet are capable of accurately classifying fish species with exceedingly low error rates. Such an accomplishment is both noteworthy and crucial. Although the two methods achieve nearly identical levels of accuracy, they differ marginally in terms of loss; MobileNet has a value of 0.1253, and VGG16 has 0.1153. The low loss indicates that both metrics possess robust capabilities in assessing the degree of similarity between the predicted and actual labels. While the disparities in loss are significant, they are inconsequential and do not undermine that both entities accomplished outstanding outcomes in classifying fish species. When considering the utilization of computing resources, MobileNet exhibits efficiency advantages that render it a more pertinent option, particularly in scenarios involving constrained resources like mobile devices. However, both can produce accurate results when it comes to identifying fish species; therefore, the decision between the two should be predicated on the requirements of the application and the resources at hand. In summary, both VGG16 and MobileNet exhibit remarkable efficacy in accurately classifying fish species, thereby contributing to the preservation of fisheries ecosystems and the marine environment while also facilitating more comprehensive investigations in the field of marine biology.

References

[1] Iswahyudi, I., Hindarto, D. and Santoso, H., 2023. PyTorch Deep Learning for Food Image Classification with Food Dataset. *Sinkron: jurnal dan penelitian teknik informatika*, *8*(4), pp.2651-2661. DOI: https://doi.org/10.33395/sinkron.v8i4.12987.

[2] Mansour, M., Cumak, E.N., Kutlu, M. and Mahmud, S., 2023. Deep learning based suture training system. *Surgery Open Science*, *15*, pp.1-11. DOI: https://doi.org/10.1016/j.sopen.2023.07.023.

[3] Chen, Y., Chen, Y., Fu, S., Yin, W., Liu, K. and Qian, S., 2023. VGG16-based intelligent image analysis in the pathological diagnosis of IgA nephropathy. *Journal of Radiation Research and Applied Sciences*, *16*(3), p.100626. DOI: https://doi.org/10.1016/j.jrras.2023.100626.

[4] Nijaguna, G.S., Babu, J.A., Parameshachari, B.D., de Prado, R.P. and Frnda, J., 2023. Quantum Fruit Fly algorithm and ResNet50-VGG16 for medical diagnosis. *Applied Soft Computing*, *136*, p.110055. DOI: https://doi.org/10.1016/j.asoc.2023.110055.

[5] Zhu, F., Li, J., Zhu, B., Li, H. and Liu, G., 2023. Uav remote sensing image stitching via improved vgg16 siamese feature extraction network. *Expert Systems with Applications*, *229*, p.120525. DOI: https://doi.org/10.1016/j.eswa.2023.120525.

[6] Sun, J., Yang, F., Cheng, J., Wang, S. and Fu, L., 2024. Nondestructive identification of soybean protein in minced chicken meat based on hyperspectral imaging and VGG16-SVM. *Journal of Food Composition and Analysis*, *125*, p.105713. DOI: https://doi.org/10.1016/j.jfca.2023.105713.

[7] Wang, H., Lu, F., Tong, X., Gao, X., Wang, L. and Liao, Z., 2021. A model for detecting safety hazards in key electrical sites based on hybrid attention mechanisms and lightweight Mobilenet. *Energy Reports*, *7*, pp.716-724. DOI: https://doi.org/10.1016/j.egyr.2021.09.200.

[8] Liu, Y., Wang, Z., Wang, R., Chen, J. and Gao, H., 2023. Flooding-based MobileNet to identify cucumber diseases from leaf images in natural scenes. *Computers and Electronics in Agriculture*, *213*, p.108166. DOI: https://doi.org/10.1016/j.compag.2023.108166.

[9] Sonmez, M.E., Altinsoy, B., Ozturk, B.Y., Gumus, N.E. and Eczacioglu, N., 2023. Deep learning-based classification of microalgae using light and scanning electron microscopy images. *Micron*, *172*, p.103506. DOI: https://doi.org/10.1016/j.micron.2023.103506.

[10] Chen, J., Zhang, D., Suzauddola, M. and Zeb, A., 2021. Identifying crop diseases using attention embedded MobileNet-V2 model. *Applied Soft Computing*, *113*, p.107901. DOI: https://doi.org/10.1016/j.asoc.2021.107901.

[11] Gujjar, J.P., Kumar, H.P. and Chiplunkar, N.N., 2021. Image classification and prediction using transfer learning in colab notebook. *Global Transitions Proceedings*, *2*(2), pp.382-385. DOI: https://doi.org/10.1016/j.gltp.2021.08.068.

[12] Zhou, W., Wang, H. and Wan, Z., 2022. Ore image classification based on improved CNN. *Computers and Electrical Engineering*, *99*, p.107819. DOI: https://doi.org/10.1016/j.compeleceng.2022.107819.

[13] Chakraborty, K.K., Mukherjee, R., Chakroborty, C. and Bora, K., 2022. Automated recognition of optical image based potato leaf blight diseases using deep learning. *Physiological and Molecular Plant Pathology*, *117*, p.101781. DOI: https://doi.org/10.1016/j.pmpp.2021.101781.

[14] Hindarto, D., 2023. Performance Comparison ConvDeconvNet Algorithm Vs. UNET for Fish Object Detection. *Sinkron: jurnal dan penelitian teknik informatika*, *8*(4), pp.2827-2835. DOI: https://doi.org/10.33395/sinkron.v8i4.13135.

[15] Hindarto, D., 2023. Enhancing Road Safety with Convolutional Neural Network Traffic Sign Classification. *Sinkron: jurnal dan penelitian teknik informatika*, *8*(4), pp.2810-2818. DOI: https://doi.org/10.33395/sinkron.v8i4.13124.