2023 IEEE 9th Information Technology International Seminar (ITIS)
Batu Malang, Indonesia, October 18 – 20, 2023

Marine Fish Species Classification Using Transfer Learning And Residual Network

Budi Dwi Satoto
Department of Information System
University of Trunojoyo Madura
Bangkalan, Indonesia
budids@trunojoyo.ac.id

Bain Khusnul Khotimah
Department of Informatics Engineering
University of Trunojoyo Madura
Bangkalan, Indonesia
bainkk@trunojoyo.ac.id

Muhammad Syarief
Department of Information System
University of Trunojoyo Madura
Bangkalan, Indonesia
mohammad.syarief@trunojoyo.ac.id

Mohammad Yusuf
Department of Information System
University of Trunojoyo Madura
Bangkalan, Indonesia
mohammadyusuf@trunojoyo.ac.id

Mochammad Kautsar Sophan
Department of Informatics Engineering
University of Trunojoyo Madura
Bangkalan, Indonesia
kautsar@trunojoyo.ac.id

Devie Rosa Anamisa
Department of Informatics Engineering
University of Trunojoyo Madura
Bangkalan, Indonesia
devros_gress@trunojoyo.ac.id

***Abstract*— Sea fish is one source of food that can be found in coastal waters. Fishermen meet market demand and deliver products to local markets. Fish species are selected using visual observation. Artificial intelligence is expected to help this process by using deep learning. The algorithm that can be used is the Convolutional neural network. Classification of fish species: Research the shape and color of fish in depth so that the machine can help distinguish the type of fish caught. This research used nine classes of marine fish for consumption, including Tongkol, tuna, Kembung, Tenggiri, Kakap, Cakalang, Salmon, Sarden, and Baronang. Each of the nine classes of fish contains 60 training images, including 45 validation images and 10 test images, so the total images used are 534. The architecture used is a custom residual layer with data augmentation. Data limitations handled by augmentation techniques include scaling, rotation, sliding, and reflection. From the results of the research trial scenarios, it can be seen that this research can classify species with an average accuracy of 99.26%, precision of 0.9926, recall of 0.9927, and F-1 score of 0.9927 with an average computational learning time using layers custom about 3 minutes.**

Keywords—Fish Classification, deep learning, convolution neural network, data augmentation

I. Introduction

Marine fish are the main component of the marine ecosystem and have a significant role in the food cycle and maintaining the ecosystem's balance. Understanding marine fish and their classification helps scientists, fishermen, natural resource managers, and other stakeholders maintain fisheries' sustainability and a healthy marine ecosystem. The reason for choosing the topic of marine fish classification, one related to Fisheries Management, includes knowledge about the category of saltwater fish that helps in sustainable fisheries management. By identifying the species caught and understanding the relationships between different species, it is possible to develop effective management strategies to prevent overfishing, limit the catch of endangered species, and ensure the sustainability of fishery resources. Ecosystem Understanding: The classification of marine fish helps us understand their role in the marine ecosystem. As predators and prey, marine fish influence the structure and function of ecosystems. By understanding the classification of marine fish, it is possible to trace the flow of energy and nutrients in the food chain, as well as predict and overcome the impact of ecosystem changes that may occur. Resource Utilization can identify species that have commercial value and are essential in the fisheries sector. This information helps develop appropriate policies for managing fisheries resources, maintaining the sustainability of fisheries, and providing economic benefits for coastal communities. By understanding the classification of marine fish, we can expand our knowledge of marine biodiversity and maintain a healthy ecosystem [1] [2].

Classification of coastal marine fish is a widely studied problem in image segmentation, pattern recognition, and information retrieval. The government must balance the supply of fish for consumption with ecosystem conditions, trade, agriculture, marine science, and the fish industry, including nutrition and canning factories. This balancing effort has been applied in many domains, including target marketing. In image pattern recognition, the preprocessing and feature extraction stages are carried out to determine the fish family/species type. The survey used the Global Information System (GIS) on Fishes, the Fish4-Knowledge Database, and other FC databases. At the same time, the technique for understanding the characteristics of fish patterns is done by feature mapping. The number of extracted features can be identified using the appropriate feature extraction method, thereby increasing the classification accuracy value [3].

Subsequent research tries to label fish objects automatically using a camera without human intervention. The step is to apply the deep neural fish classification system. Various picture poses, environmental conditions, uncontrolled lighting, and limitations of the training dataset affect the data acquisition process. Identifying fish in nature is still a problem that has not yet received a proper solution. In the process, the use of independent branches concerning the variations of fish poses and scales and the results of extracting discriminatory features to distinguish fish [4].

Furthermore, the second branch uses context information in scenarios to infer fish species, assuming the fish caught correlate with the surrounding environmental conditions. The conditions are influenced by boats, fishing gear, daylight, and weather. The prediction results of the initial classification on both branches were then re-weighted, and the average value was sought as the final prediction. In the world competition "The Nature -Conservancy Fisheries Monitoring" by Kaggle, this model reached the top 0.7% on the last board among solutions submitted by 2,293 other competitive teams [5].

The potential for fish production in Indonesia is very high because the territory of Indonesia consists of waters (seas, lakes, rivers, and ponds). However, fish consumption in the community is still shallow. The community needs a tool that can be used quickly and accurately in choosing fish worthy of consumption. Community awareness of the freshness of the fish consumed must also be considered. The minimal use of technology makes fishing production slow. It is not uncommon for fish rotting processes to escape the attention of fish traders. The method of sorting fish by making a manual selection of fish ensures the freshness of the fish that reaches the hands of the consumer cannot be guaranteed. This research further wants to overcome that by developing a method to classify the freshness of fish based on the image of the fish. One of the algorithms that can be used is K-Nearest Neighbor (kNN), based on the summary of colors from fish images. The community needs a tool that can be used quickly and accurately in choosing fish worthy of consumption. Community awareness of the freshness of the fish consumed must also be considered. The minimal use of technology makes fishing production slow. It is not uncommon for fish rotting processes to escape the attention of fish traders [6].

Due to the country's extensive water resources (lakes, rivers, ponds, and seas), Indonesia has a very high potential for fish production. Fish intake in society is still deficient. Another issue is that consumers are unaware of how fresh the fish they eat is. Because it is typical for fish vendors to be unaware of the process of fish rotting, the public requires tools that can be utilized quickly and precisely to choose fish that is fit for consumption. In addition, the delayed production of fisheries is caused by the little technology usage. Since fish is manually sorted, it is impossible to guarantee that it will be fresh when it gets to consumers. Based on the findings of this research, a method for categorizing fish freshness was created. The classification technique k-Nearest Neighbor (kNN) is utilized based on the color summary of fish photos. For the kNN, the classification accuracy results were 91.36%. It demonstrates that the final method is workable. On the other hand, the fish's eyes' dark hue is the color that most accurately identifies how fresh the fish is. The reason for this is that among all the fish varieties used, black has the highest Information Gain [7].

The following paper's recommendation of an automatic fish species classification method is its primary goal. For ichthyologists and marine biologists to fully understand fish behavior, highly accurate fish classification is essential. The responsible agencies must record the quantity of fish per species and mark endangered species in large and small water bodies. Due to obstacles such as background noise, image distortion, other bodies of water in the image, image quality, and occlusion, most systems now concentrate on classifying fish found outside of water. This method uses cutting-edge methods based on Deep Learning, Convolutional Neural Networks, and Image Processing to attain an accuracy of 96.29%. Compared to earlier proposed methods, our technique guarantees a notable increase in discrimination accuracy [8].

From the paper review process above, this research is interested in taking the theme of fish classification for coastal areas because there are benefits that can be used. Namely, the process of selecting fish can be done with the help of artificial intelligence. The method used is deep learning. The technique is a convolutional neural network with a training process and deep understanding. In this research, the gap is that previous research still uses machine learning, where the advantage of machine learning is the short processing time because there is no learning process. Even so, the accuracy produced by machine learning is not as good as deep learning. In addition, the process of creating ground truth requires a long process. From these considerations, the authors developed a deep learning system with a residual transfer learning architecture with the aim that the learning system with previously trained datasets is added to the primary dataset to be prepared and introduced to increase accuracy. In addition, augmentation techniques are used to overcome the limitations of preliminary data. Residual architecture has a skip connection, saving the training process computational time. It is hoped that this research will obtain the results of the optimal kernel convolution multiplication so that it can improve accuracy.

II. RESEARCH BACKGROUND

A. Classification of Types of Sea Fish for Consumption

The differences in fish consumption in Southeast Asia and Europe mainly lie in the types of fish commonly consumed, taste preferences, and local culinary culture. Traditional Processing and Cooking Methods: In Southeast Asia, fish is often fried, grilled, or cooked in rich and aromatic sauces. In Europe, fish is boiled, grilled with spices or sauce, or served raw as sushi or sashimi. Cakalang fish has an abundance of omega-3 and B complex vitamins, which can be nutritious for the body. Mackerel fish has nutritional content of vitamin B1, vitamin B2, vitamin C, vitamin B6, vitamin B12, vitamin A, and vitamin E. The ASEAN region fish is shown in Fig. 1.

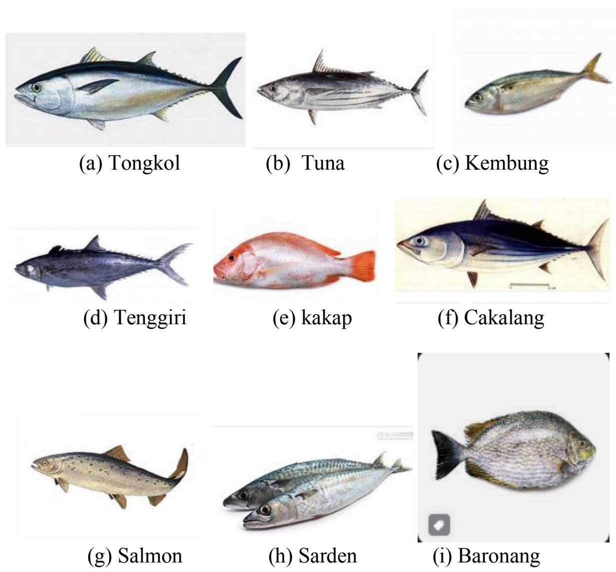


Fig. 1. Types of fish

Omega-3 Fatty Acid Fish is a Source of Vitamin B and Rich in Antioxidants. Sardines have nutritional content such as minerals and vitamins. Snapper fish has the benefit of helping maintain heart health because it contains omega-3 fatty acids that are nutritious for the body. Salmon has DHA fatty acid, which plays an essential role in the development and function of the brain [9].

B. Deep learning

The basic concept in deep learning is to utilize artificial neural networks, which in more complex cases consist of many layers (deep layers) to learn the patterns in the data. One of the methods of deep learning is Transfer Learning. This technique uses the knowledge that previous deep learning models have learned to overcome new problems that are similar or related. Transferring knowledge from previously trained models can reduce the time and resources required to introduce new models. Components Layer components in transfer learning include:

1. Pre-trained Base Model: This deep learning model has been pre-trained on a large dataset and a task similar to the job you want to solve. This model generally has convolution layers that extract general features from the data.

$$R[i,j] = \sum_{u=-K}^{K} \sum_{v=-K}^{K} h[u,v]W[i-u, i-v] \quad (1)$$

$R[i,j]$ = Convolution Result
u, v = index
K = maximum value of input matrix
W = Kernel Filter [10]

2. Frozen layers: These are the layers in the pre-trained base model that will not undergo weight renewal during the transfer learning training process. By locking these layers, it can retain the previously learned feature representation.

$$\hat{x}^{(k)} = \frac{x^{(k)} - E[x^{(k)}]}{\sqrt{Var[x^{(k)}]}} \quad (2)$$

$\hat{x}^{(k)}$ = Normalisation value from k-th hidden unit
$E[x^{(k)}]$ = Expectation value from k-th unit / mean value
$Var[x^{(k)}]$ = Variance value from the hidden unit [9].

After normalization, each remote unit will have a mean of zero and a variance of one, but having a standard of 0 and 1 is undesirable. Next, Max polling is shown in Fig. 2 [11].

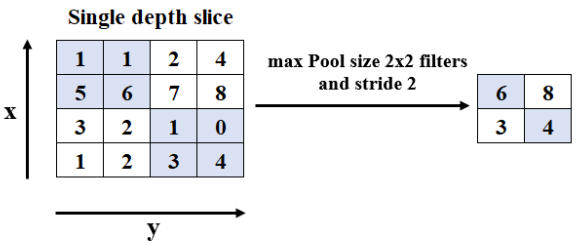


Fig. 2. Max Pooling layer [12]

The max pooling process involves shifting small windows (e.g., 2x2 or 3x3) on the input and taking the maximum value in each window. This process is repeated for each window that shifts across the information [13].

3. Fine-tuning layer: These layers will undergo weight renewal during transfer learning training at the end of the pre-trained base model. It allows the model to adapt feature representations for specific tasks [9].

4. Additional Layers: Besides the pre-trained base model, you can add extra layers after the base model to adapt the model to the task you want to complete. These additional layers can include convolution, pooling, and fully connected layers [14].

In transfer learning, the general steps are to load a pre-trained base model, lock the necessary layers, add additional layers, and train the model using specific data.

C. *Augmentation of data*

Data augmentation produces new variations from existing data by transforming or manipulating the data. Four methods can be used: scale, rotation, shears, and reflection [15].

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \quad (3)$$

x, y = Initial Position
x', y' =Next position
θ = angle of rotation
This technique increases the available data samples, which is helpful if the initial dataset is limited [16].

D. *Confusion matrix*

The confusion matrix compares the model's predicted and observed data labels [17]. It is shown in Fig. 3.

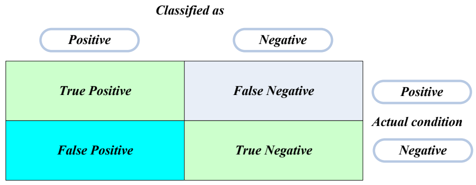


Fig. 3. Confusion Matrix [18]

E. *Measurement*

The ratio between the total number of correct predictions. With the total number of predictions.

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad (4)$$

Precision: The ratio of correctly predicted positive cases to all positive objects [19].

$$Precision = \frac{TP}{(TP + FP)} \quad (5)$$

Recall or Sensitivity (Recall/Sensitivity): The ratio of favorable cases predicted to be true to all positive cases.

$$Recall = \frac{TP}{(TP + FN)} \quad (6)$$

F1 value (F1 Score) is the Harmonic average value of precision and recall, which measures the balance between the two [20].

$$F-1\ Score = \frac{2 * (Presisi * Recall)}{(Presisi + Recall)} \quad (7)$$

F. *Error Value of MSE, RMSE, and MAE*

This measurement technique is usually used in the case of regression by measuring how well the model can predict continuous numeric values. The lower the MSE value, the closer the model prediction is to the actual value [21].

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \tilde{y}_i)^2 \quad (8)$$

$$MAE = \frac{1}{n} \sum_{j=1}^{n} |y_j - \hat{y}_j| \quad (9)$$

MSE = Minimum Square Error
MAE = Mean Absolute Error
n = number of sampling data
y_i = actual value
$\tilde{y}_i$ = prediction value [22]

III. RESEARCH METHODS

Transfer learning with a Residual Network is a deep and famous convolutional neural network (CNN) architecture. It is known as using residual blocks to enable more efficient learning through the "shortcut" or "skip connection" mechanism shown in Fig. 4. ResNet has generally been trained on large datasets such as ImageNet, which includes millions of images with various categories.

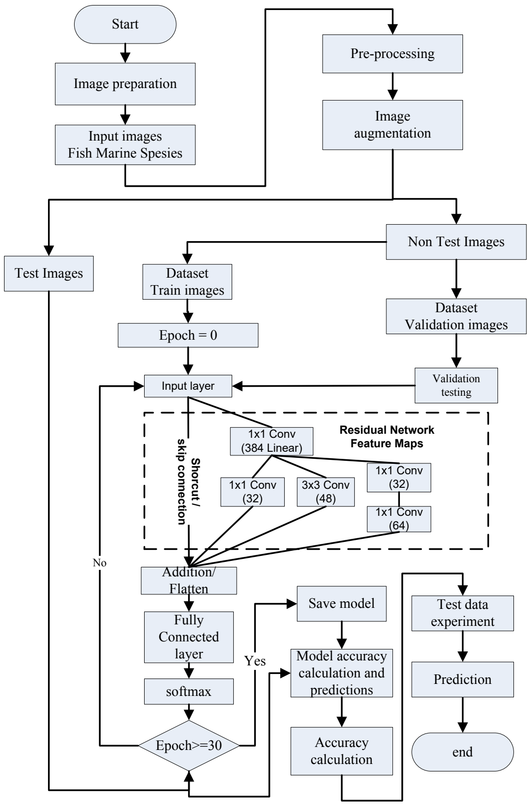


Fig. 4. Research Methods

using local training to help overcome overfitting problems. Overfitting can occur when the model is too complex compared to the limited training data. Residual transfer learning enables direct information flow from input to output via skip connections. These connections help retain the original information and facilitate gradient propagation during the model training. Inception helps lighten the convolution process by multiplying kernels with small sizes. The locally trained dataset was obtained from kaggle with the keyword "large-scale-fish-dataset" containing nine different types of seafood collected from supermarkets in Izmir, Turkey, published in ASYU 2020. The data is from Google and local wisdom because fish types in Europe and the ASEAN region have different names and types of characteristics

IV. RESULT AND DISCUSSION

A. Preprocessing

The step taken at the initial stage is to check all the files used in this research because the data have different sizes. The result of the image resize process is shown in Figure 3. For the clarity of the process, an image with 224 horizontal pixels, 224 vertical pixels, and bit depth three is used, which means RGB or 24 colors. It is shown in Fig. 5.

Fig. 5. Resize of image

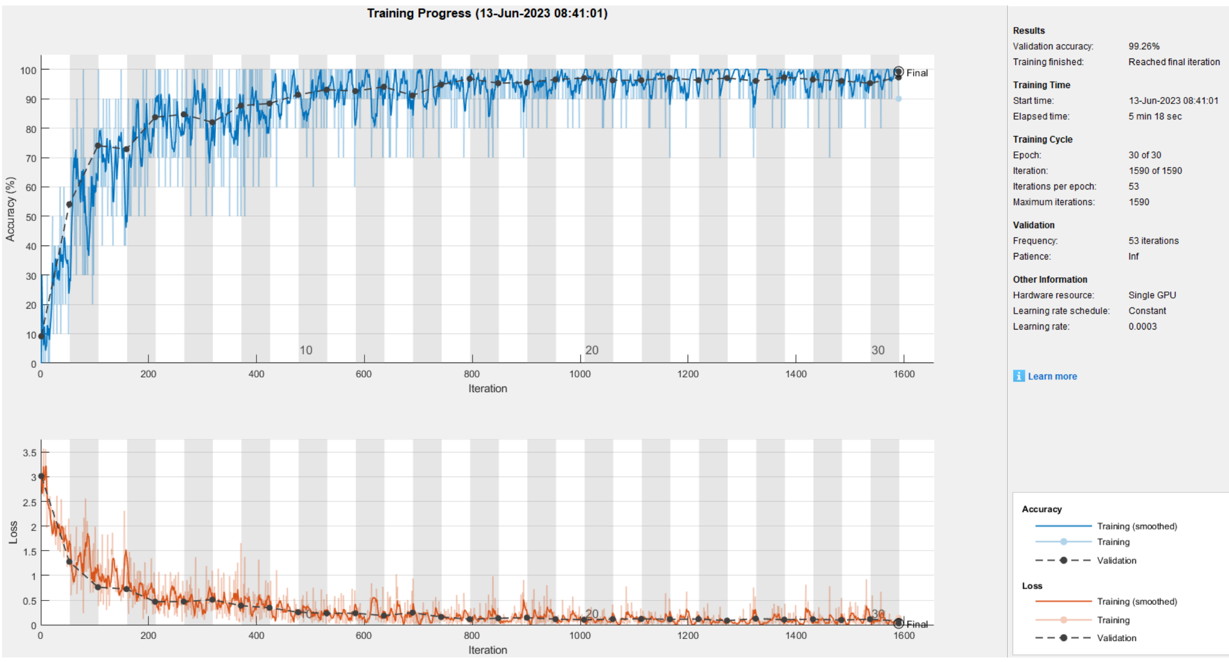


Fig. 6. Training Progress

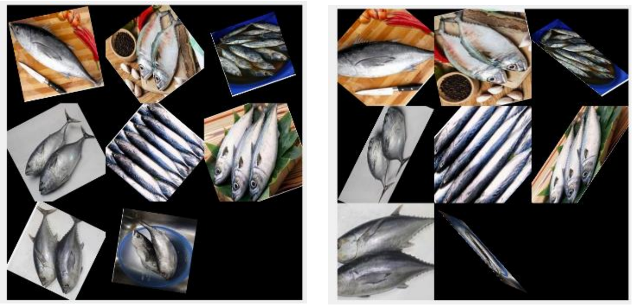
Residual transfer learning allows leveraging knowledge in a pre-trained model to benefit from previously learned feature representations. The results are combined with education

B. Training Option

In this research, the parameters used for training include: training options with Adam's optimization options, Initial Learning Rate 3e-4, MaxEpochs 100, Performed Shuffle do every-epoch, Using Validation Data, imdsValidation, Validation Frequency used 10, Displaying Verbose value = true, Plots, training-progress. It is shown in Fig. 6.

C. Augmentation of Data

This technique increases the amount of data due to data limitations. It is shown in Fig. 7. Several processes are done to achieve the desired result: resize, rotation, shears, and reflection.



(a) Rotate and Scaling (b) Shears and Reflection

Fig. 7. Data Augmentation

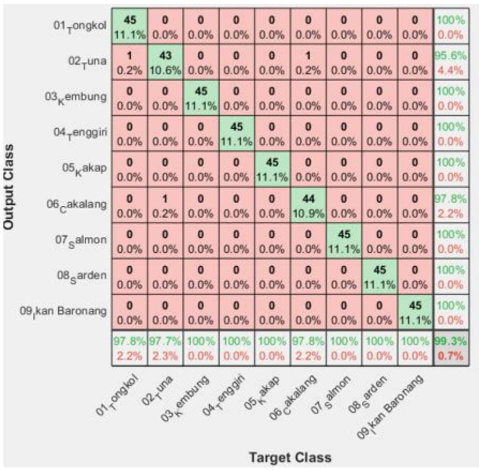
$$x' = x\cos\theta - y\sin\theta \quad (10)$$
$$y' = x\sin\theta - y\cos\theta \quad (11)$$

x, y = Initial position of a pixel
x', y' = Next position of a pixel
θ = angle of rotation

The result of the multiplication process of the input image matrix with the augmentation kernel is shown in Formula 10. It is done before the convolution process until it is obtained.

D. Confusion Matrix

The confusion Matrix makes it possible to calculate various performance metrics such as accuracy, precision, recall, and F1-score. The Matrix provides insight into the model's performance obtained from the difference between the input matrix and the model formed during training. It is shown in Fig. 8.



Output Class \ Target Class	01 Tongkol	02 Tuna	03 Kembung	04 Tenggiri	05 Kakap	06 Cakalang	07 Salmon	08 Sarden	09 Ikan Baronang	
01 Tongkol	45 11.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
02 Tuna	1 0.2%	43 10.6%	0 0.0%	0 0.0%	0 0.0%	1 0.2%	0 0.0%	0 0.0%	0 0.0%	95.6% 4.4%
03 Kembung	0 0.0%	0 0.0%	45 11.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
04 Tenggiri	0 0.0%	0 0.0%	0 0.0%	45 11.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
05 Kakap	0 0.0%	0 0.0%	0 0.0%	0 0.0%	45 11.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
06 Cakalang	0 0.0%	1 0.2%	0 0.0%	0 0.0%	0 0.0%	44 10.9%	0 0.0%	0 0.0%	0 0.0%	97.8% 2.2%
07 Salmon	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	45 11.1%	0 0.0%	0 0.0%	100% 0.0%
08 Sarden	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	45 11.1%	0 0.0%	100% 0.0%
09 Ikan Baronang	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	45 11.1%	100% 0.0%
	97.8% 2.2%	97.7% 2.3%	100% 0.0%	100% 0.0%	100% 0.0%	97.8% 2.2%	100% 0.0%	100% 0.0%	100% 0.0%	99.3% 0.7%

Fig. 8. Confusion Matrix

Three photographs were in the wrong category in the data training process, using 45 images per class. One photo of tuna entering the tongkol, one photo of tuna into cakalang, and One photo of cakalang into tuna. From the results obtained, the accuracy of the model is around 99.3%

E. Prediction

After getting the model's accuracy, it continued by making predictions. The step is the test data matrix, which is the experimental data compared to the model until the accuracy believability is obtained. It is shown in Fig. 9.

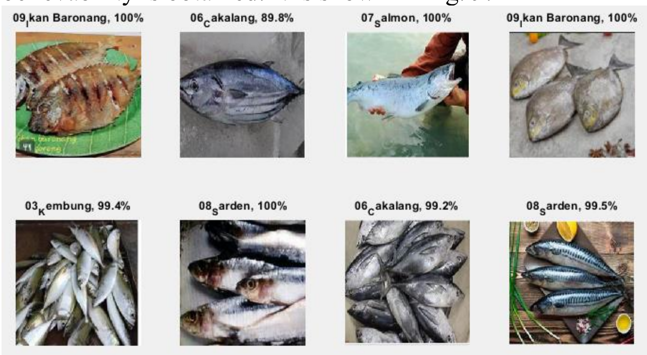


Fig. 9. Prediction Result of CNN 01

CNN's preliminary prediction results show that the average system is already working well. Accuracy confidence produced 99.4% bream, 89.8% bream, and Sardine 99.5%.

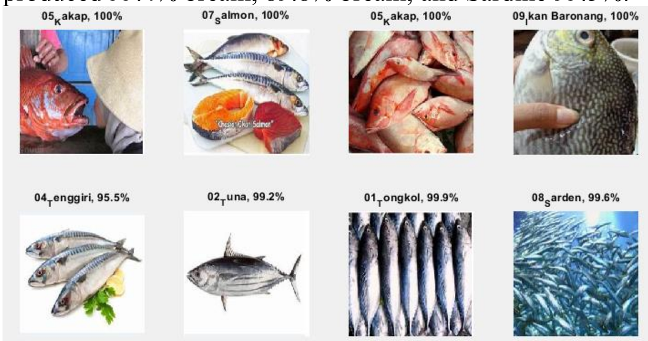


Fig. 10. Prediction Result of CNN 02

In the second experiment, Fig. 10, tenggiri 95.5%, tuna 99.2%, tongkol 99.9%, and sardines 99.6% can be recognized well. From the results of the prediction process, the system has a level of trust because the model used is good enough. If the model produced has high accuracy, then the accuracy of belief is also high.

F. Architecture Comparison and time-consuming

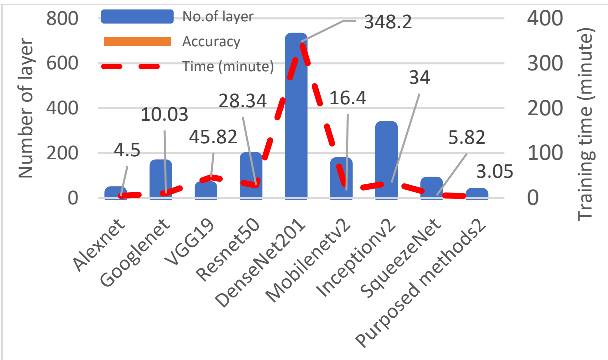


Fig. 11. Comparison between Architecture

The comparison between the architecture of deep learning shows that the proposed methods have an optimal time computation of around 3 minutes to train and get the model compared to each other. It is shown in Fig. 11.

G. Comparison of Architecture

A test scenario validated the system by comparing Resnet, alexnet, googlenet, mobilenetv2, VGG, densenet, squeezenet, Resnet50, mobiletnetv2, Inception, and purposed methods.

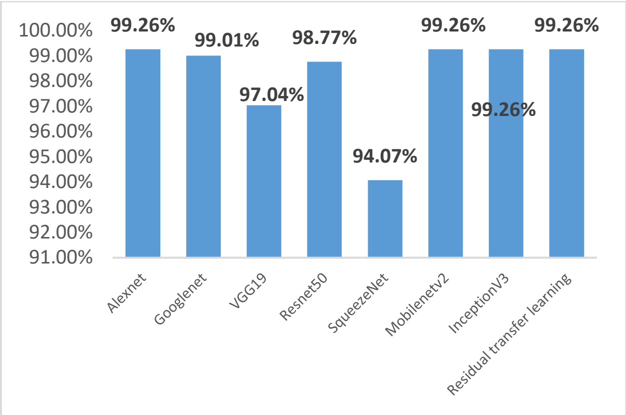


Fig. 12. Comparison of Accuracy

V. CONCLUSION

The Tests have been carried out, and the system can distinguish well the types of consumption fish that can be obtained in the coastal waters. The nine classes observed in this research included types of fish: Tongkol, tuna, Kembung, Tenggiri, Kakap, Cakalang, Salmon, Sarden, and Baronang. Each of the nine classes of fish contains 60 training images, including 45 validation images and 10 test images. The limitation of this research is the number of data for each category. It is overcome using data augmentation techniques such as scale, rotation, shears, and reflection. The research contribution uses fish data in the Southeast Asia region and the CNN architecture by adopting transfer learning and adding skip connections to the feature maps layer. The test scenario results show an average accuracy of 99.26% with an average computation time of 3 minutes to build a model—precision value 0.9926, Recall 0.9927, and F-1 score 0.9927.

ACKNOWLEDGMENT

The author thanks the engineering faculty of Trunojoyo University for helping to publish this article.

REFERENCES

[1] P. Xu, M. Xie, W. Zhou, and A. Suo, "Research on Fishery Resource Assessment and Sustainable Utilization (FRASU) during 1990–2020: A bibliometric review," *Glob. Ecol. Conserv.*, vol. 29, p. e01720, 2021, doi: https://doi.org/10.1016/j.gecco.2021.e01720.

[2] R. A. Davis and Q. Hanich, "Transparency in fisheries conservation and management measures," *Mar. Policy*, vol. 136, p. 104088, 2022, doi: https://doi.org/10.1016/j.marpol.2020.104088.

[3] S. Chen, W. Lin, C. Zeng, B. Liu, A. Serres, and S. Li, "Mapping the fishing intensity in the coastal waters of Guangdong province, China through AIS data," *Water Biol. Secure.*, vol. 2, no. 1, p. 100090, 2023, doi: https://doi.org/10.1016/j.watbs.2022.100090.

[4] K. Cai, X. Miao, W. Wang, H. Pang, Y. Liu, and J. Song, "A modified YOLOv3 model for fish detection based on MobileNetv1 as a backbone," *Aquac. Eng.*, vol. 91, p. 102117, 2020, doi: https://doi.org/10.1016/j.aquaeng.2020.102117.

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

[6] S. Winiari, F. Indikawati, A. Oktaviana, and H. Yuliansyah, "Consumable Fish Classification Using k-Nearest Neighbor," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 821, p. 12039, May 2020, doi: 10.1088/1757-899X/821/1/012039.

[7] N. M. S. Iswari, Wella, and Ranny, "Fish freshness classification method based on fish image using k-Nearest Neighbor," in *2017 4th International Conference on New Media Studies (CONMEDIA)*, 2017, pp. 87–91, doi: 10.1109/CONMEDIA.2017.8266036.

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

[9] W. Yu, K. Yang, H. Yao, X. Sun, and P. Xu, "Exploiting the complementary strengths of multi-layer CNN features for image retrieval," *Neurocomputing*, vol. 237, pp. 235–241, 2017, doi: https://doi.org/10.1016/j.neucom.2016.12.002.

[10] T. Alshalali and D. Josyula, "Fine-Tuning of Pre-Trained Deep Learning Models with Extreme Learning Machine," in *2018 International Conference on Computational Science and Computational Intelligence (CSCI)*, 2018, pp. 469–473, doi: 10.1109/CSCI46756.2018.00096.

[11] P. K. Koo and S. R. Eddy, "Representation learning of genomic sequence motifs with convolutional neural networks," *PLoS Comput. Biol.*, vol. 15, no. 12, p. e1007560, 2019, doi: 10.1371/journal.pcbi.1007560.

[12] F. Chou, Y. Tsai, Y. Chen, J. Tsai, and C. Kuo, "Optimizing Parameters of Multi-Layer Convolutional Neural Network by Modeling and Optimization Method," *IEEE Access*, vol. 7, pp. 68316–68330, 2019, doi: 10.1109/ACCESS.2019.2918563.

[13] L. C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 4, 2018, doi: 10.1109/TPAMI.2017.2699184.

[14] I. Dominguez-Catena, D. Paternain, and M. Galar, "Additional Feature Layers from Ordered Aggregations for Deep Neural Networks," in *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2020, pp. 1–8, doi: 10.1109/FUZZ48607.2020.9177555.

[15] S. S. Yadav and S. M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," *J. Big Data*, vol. 6, no. 1, p. 113, 2019, doi 10.1186/s40537-019-0276-2.

[16] S. C. Wong, A. Gatt, V. Stamatescu, and M. D. McDonnell, "Understanding Data Augmentation for Classification: When to Warp?," in *2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, 2016, pp. 1–6, doi: 10.1109/DICTA.2016.7797091.

[17] L. Zhang and I. A. Kakadiaris, "Local classifier chains for deep face recognition," in *2017 IEEE International Joint Conference on Biometrics (IJCB)*, 2017, pp. 158–167, doi: 10.1109/BTAS.2017.8272694.

[18] L. M. Rodrigues and M. Mestria, "Classification methods based on Bayes and neural networks for human activity recognition," in *2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, 2016, pp. 1141–1146, doi: 10.1109/FSKD.2016.7603339.

[19] S. Mishra, R. Sachan, and D. Rajpal, "Deep Convolutional Neural Network based Detection System for Real-time Corn Plant Disease Recognition," *Procedia Comput. Sci.*, vol. 167, pp. 2003–2010, 2020, doi: https://doi.org/10.1016/j.procs.2020.03.236.

[20] A. S. Paymode and V. B. Malode, "Transfer Learning for Multi-Crop Leaf Disease Image Classification using Convolutional Neural Network VGG," *Artif. Intell. Agric.*, vol. 6, pp. 23–33, 2022, doi: https://doi.org/10.1016/j.aiia.2021.12.002.

[21] F. Bougourzi, F. Dornaika, and A. Taleb-Ahmed, "Deep learning based face beauty prediction via dynamic, robust losses and ensemble regression," *Knowledge-Based Syst.*, vol. 242, p. 108246, 2022, doi: https://doi.org/10.1016/j.knosys.2022.108246.

[22] J. Guo, W. Wang, Y. Tang, Y. Zhang, and H. Zhuge, "A CNN-Bi_LSTM parallel network approach for train travel time prediction," *Knowledge-Based Syst.*, p. 109796, 2022, doi: https://doi.org/10.1016/j.knosys.2022.109796.