A Model of Visual Intelligent System for Genus Identification of Fish in the Siluriformes Order

Taufik Budhi Pramono¹, R Ardharyan Islamy², Saprudin³, Joni Johanda Putra¹, Teddy Suparyanto⁴, Kartika Purwandari⁴, Tjeng Wawan Cenggoro⁴,⁵, Bens Pardamean⁴,⁶

¹Aquaculture Study Program, Faculty of Fisheries and Marine Science, Jenderal Soedirman University, Purwokerto, Central Java, Indonesia 53123
²Aquatic Resources Management, Brawijaya University, Malang, East Java, Indonesia
³Graduate Program Aquatic Resources, Faculty of Fisheries and Marine Science, Jenderal Soedirman University, Purwokerto, Central Java, Indonesia 53123
⁴Bioinformatics and Data Science Research Center, Bina Nusantara University, Jakarta, Indonesia 11480
⁵Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta, Indonesia 11480
⁶Computer Science Department, BINUS Graduate Program - Master of Computer Science Program, Bina Nusantara University, Jakarta, Indonesia 11480

Correspondent author: kartika.purwandari@binus.edu

Abstract. Ambiguity in the fish naming is present in several fish species database, especially for fish in Siluriformes order. To fix the ambiguity, a visual intelligent system is needed to automate the fish naming correction in the database. In this study, we developed a deep-learning-based model as the core of the intelligent system. The proposed model achieved 89% accuracy for the classification of three genera in Siluriformes order: Mystus, Hemibagrus, and Glyptothorax.

Keywords: fish image classification, intelligent system, deep learning, siluformes.

1. Introduction

Species identification is important to preserve biodiversity in a natural environment [1]. However, specifically for fish, the identification of fish species is generally carried out by taxonomic identification using references. The key references for identifying fish species in Indonesia are Saanin [2] and Kottelat et.al. [3]. Along with the development of technological science, the current identification process is starting to be carried out visually from images by comparing the characteristics from databases of fish species such as IUCN, Fishbase, and BOLD system. However, several images in the databases were not correctly named, especially for the Siluriformes order. For example, the naming of the Senggaringan fish species from the Siluriformes order from the Klawing River, Purbalingga Regency, Indonesia is still ambiguous. Different studies identified it as different species, namely Mystus nigriceps [4–7] and Mystus micrachantus [8]. It was eventually confirmed genetically that it actually is Mystus singaringan by Pramono et.al. [9–11]. To fix the misnaming issue, the most practical approach is to develop an intelligent system that can scan through the fish images in the databases and list all the potential misnaming.
Currently, a trend to integrate deep learning [12] for visual intelligent systems is emerging [13,14]. This trend was set by the astonishing winning of AlexNet [15] in the ImageNet Large Scale Visual Recognition Challenge [16,17] in 2012. Not only for image, deep learning also excels in other complex data types such as audio [18], text [19,20], and even biological data [21]. Deep learning is popular for image analysis because of its ability to automatically learn visual features, hence eliminating the needs to use handcrafted visual features which is essential in classical image analysis models [22,23].

Since 2016, deep learning is started to be integrated in fish classification models. The early deep learning-based fish classification models still incorporate Support Vector Machine [24] on top of a deep learning architecture [25–27]. Following the trend in other computer vision research, the fish classification models were started to employ an end-to-end deep learning architecture. The first full deep learning model for fish classification was developed by Chen et al. [28], in which combined deep learning for image-level classification and instance-level classification with adaptive prediction. Despite the superior performance of a full deep learning model, it normally requires a massive training data to reach its optimal performance [29,30]. Because of the limited training data for fish classification, strategies that can reduce the training data size requirement were heavily utilized to develop full deep learning models for fish classification. The strategies were transfer learning [31] and data augmentation [32]. Not only they were studied for modeling research, but deep-learning-based fish classification models were also physically implemented for underwater-drone [33] and aquarium [34].

This study was conducted to address the misnaming issue by developing the intelligent system. Specifically, the contribution of this study is to develop an Artificial Intelligence (AI) model for genus identification of fish in the Siluriformes order. The model is based on deep learning. Specifically, the model is Convolutional Neural Networks (CNN) [35,36], which is specialized for image processing tasks. The model developed in this study is expected to be the basis of a general morphological fish identification system.

2. Method
In this study, a deep learning model based on transfer learning is proposed for the automated detection of the fish images. Transfer learning is utilized in this study because of its capability to improve a deep learning model with limited data [36,37]. For the deep learning architectures, we compared the performance of ResNet-18, ResNet-34, and ResNet-50 pre-trained networks, which are popular for transfer learning. These architectures are the member of ResNet [38] architectures family, which has special links called residual connection. These links connect a layer to the consecutive layers. A layer can access the feature maps of other layers because of these residual connections. Since each layer learns the flow of information directly within the network, the overfitting problem which may arise during the training of models can be averted. The architectures were employed in a transfer learning scheme as illustrated in Fig. 1.
To train the networks, we used a dataset that contains 179 images of fish from three genera in Siluriformes order: Mystus, Hemibagrus, and Glyptotheorax. The number of images for Mystus, Hemibagrus, and Glyptotheorax is respectively 87, 55, and 37. Based on the Pareto ratio, we split 20% of the whole dataset for testing. The remaining dataset was split again for training and validation also with the ratio based on Pareto principle (80% training and 20% validation. This process resulted in 112 images for training, 30 images for validation, and 37 images for testing. To reduce the deep learning requirement for big training data further, we applied random rotation, color jitter, and random horizontal flip as the augmentation strategies during the training process. Fig. 2 shows a random sample of images after the augmentation strategies are applied.

All Resnet models were trained using the stochastic gradient descent with momentum (SGDM) optimizer, with an initial learning rate of 0.001 and a mini-batch size of 8. All models were trained for 50 epochs.

3. Result and Discussion

Fig. 3 shows the confusion matrix of ResNet-18 3(a), ResNet-34 3(b), and ResNet-50 3(c). On the one hand, ResNet-18 misclassified 2 images of Glyptotheorax as Hemibagrus and Mystus, 1 image of Hemibagrus as Mystus, and 1 image of Mystus as Hemibagrus. The testing accuracy is 89%. On the other hand, ResNet-34 misclassified 2 images of Glyptotheorax as Mystus, 3 images of Hemibagrus as Mystus, and 1 image of Mystus as Hemibagrus. Meanwhile, ResNet-50 misclassified 3 images of Glyptotheorax as Mystus, 1 image of Hemibagrus as Mystus, and 4 images of Mystus as Hemibagrus. Based on these confusion matrices, the accuracy of each model is calculated and presented in Table 1.

It is evident from Table 1 that ResNet-18 achieved the best test accuracy of 89%, followed by ResNet-34 and ResNet-50. It should be noted that ResNet-18, ResNet-34, and Resnet-50 respectively have 18, 34, and 50 layers. Generally, an architecture with more layers can outperform a model with fewer layers. However, if the training data size is not adequate, the larger model can overfit the dataset, which can greatly reduce its accuracy. Thus, we can argue that the trend shown in Table 1 is due to training data inadequacy.

To assess the performance of all models further, we present the per-class accuracy in Table 2. It is apparent that the accuracy for Mystus class is the highest across all models, which has the largest number of images among all classes. We can see that the number of images in the training dataset plays an important role in the accuracy of the corresponding class. This phenomenon is common for deep learning that is trained with an imbalanced dataset [39–41]. With the number of Mystus images that is about two times larger than the number of Glyptotheorax images, we can consider that the dataset in this study is imbalanced.
Fig. 3 Confusion Matrix for fish dataset with ResNet-18 (a), ResNet-34 (b), and ResNet-50 (c)

**Table 1. Testing accuracy from each model.**

| Network     | Testing Accuracy |
|-------------|------------------|
| ResNet-18   | 89%              |
| ResNet-34   | 83%              |
| ResNet-50   | 78%              |

**Table 2. Testing accuracy from each class using 3 network model.**

| Network     | Glyptothorax | Hemibagrus | Mystus |
|-------------|--------------|------------|--------|
| ResNet-18   | 75%          | 91%        | 94%    |
| ResNet-34   | 75%          | 73%        | 94%    |
| ResNet-50   | 63%          | 91%        | 78%    |
To assess if the models suffered overfitting, we plotted the training and validation performance of all models as shown in figure 4 until 9. Based on the plots, we can identify the existence of an overfitting issue in all models. This confirms the trend shown in Table 1. We conjecture that the cause of this overfitting issue is the small training data used to train the models.

**Figure 4.** Training and validation accuracy from ResNet-18 model

**Figure 5.** Training and validation loss from ResNet-18 model

**Figure 6.** Training and validation accuracy from ResNet-34 model

**Figure 7.** Training and validation loss from ResNet-34 model

**Figure 8.** Training and validation accuracy from ResNet-50 model

**Figure 9.** Training and validation loss from ResNet-50 model
4. Conclusion
The best deep learning model proposed in this study for fish classification gave an accuracy of 89%. The best model used ResNet-18 architecture, which is the smallest model tested in this study. The larger models were outperformed by ResNet-18 due to the overfitting issue, which might be caused by the small dataset in this study. For future works, more data is needed to be used to alleviate the overfitting issue. A more extensive data augmentation can also be employed to harness the information in a dataset with limited size.

References

[1] Kurniawan ID, Soesiloahadi RCH, Rahmadi C, Caraka RE, Pardamean B. The difference on Arthropod communities’ structure within show caves and wild caves in Gunungsewu Karst area, Indonesia. Ecol Environ Conserv 2018;24:72–81.
[2] Saanin H. Taksonomi dan Kunci Identifikasi Ikan. Bandung: Bina Cipta; 1986.
[3] Kottelat M, Whitten AJ, Kartikasari SN, Wirjoatmodjo S. Freshwater fishes of western Indonesia and Sulawesi 1993.
[4] Heltonika B. Food study and its relationship with reproduction of Senggaringan fish (Mystus nigriceps) in Klawing River, Purbalingga, Central Java. Master’s Thesis, Bogor Agric Univ Bogor, Indonea 2009.
[5] Hendri A. Manipulasi Fotothermal dalam memacu pematangan gonad ikan Senggaringan (Mystus nigriceps). Tesis, 2010.
[6] Pramono TB. Profil Reproduksi ikan Senggaringan (Mystus nigriceps): Dasar Pengembangan Domestikasi dan Budidaya. Tesis, 2010.
[7] Suryaningsih S, Sagi M, Nitimulyo KH, Hadisusanto S. Spawning aspects of javaean barb Puntius orphoides (Valenciennes, 1842) in Klawing River, Purbalingga, Central Java. J Iktiologi Indonesia 2012;12:35–48.
[8] Suryaningsih S, Sukmaningrum S, Simanjuntak SBI, Kusbiyanto. Diversity and longitudinal distribution of freshwater fish in Klawing River, Central Java, Indonesia. Biodiversitas J Biol Divers 2018;19:85–92.
[9] Pramono TN, Arfiati D, Widodo MS, Yanuhar U. Identification of Mystus fish with genetic approach. Sumberd Akuatik Indopasifik J 2017;1:1–9.
[10] Pramono TB, Arfiati D, Widodo MS, Yanuhar U. Iktiofauna di Hilir Sungai Klawing Kabupaten Purbalingga, Jawa Tengah. J Perikan Samakia 2018;9:30–4.
[11] Pramono TB, Arfiati D, Widodo MS, Yanuhar U. Sexual dimorphism in morphometric characters of Mystus singaringan from Klawing River in Central Java, Indonesia: Strategic instruction for conservation development. Biodiversitas J Biol Divers 2019;20:1133–9.
[12] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature 2015;521:436–44. doi:10.1038/nature14539.
[13] Prabowo H, Cenggoro TW, Budiarto A, Perbangsa AS, Muljo HH, Pardamean B. Utilizing Mobile-based Deep Learning Model for Managing Video in Knowledge Management System. Int J Interact Mob Technol 2018:62–73.
[14] Harsawardana, Rahutomo R, Mahesworo B, Cenggoro TW, Budiarto A, Suparyanto T, et al. AI-Based Ripeness Grading for Oil Palm Fresh Fruit Bunch in Smart Crane Grabber. IOP Conf Ser Earth Environ Sci 2020;426:012147. doi:10.1088/1755-1315/426/1/012147.
[15] Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. Neural Inf Process Syst 2012.
[16] Deng J, Dong W, Socher R, Li L, Kai Li, Li Fei-Fei. ImageNet: A large-scale hierarchical image database. 2009 IEEE Conf. Comput. Vis. Pattern Recognit., 2009, p. 248–55. doi:10.1109/CVPR.2009.5206848.
[17] Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, et al. ImageNet Large Scale Visual Recognition Challenge. Int J Comput Vis 2015;115:211–52. doi:10.1007/s11263-015-0816-y.
[18] Park DS, Chan W, Zhang Y, Chiu C-C, Zoph B, Cubuk ED, et al. SpecAugment: A Simple Data Augmentation Method for Automatic Speech Recognition. Interspeech 2019, ISCA; 2019, p. 2613–7. doi:10.21437/Interspeech.2019-2680.

[19] Vaswani A, Brain G, Shazeer N, Parmar N, Uszkoreit J, Jones L, et al. Attention Is All You Need. Adv Neural Inf Process Syst 2017:5998–6008.

[20] Devlin J, Chang M-W, Lee K, Toutanova K. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. Proc. NAACL-HLT 2019, 2019, p. 4171–86.

[21] Lumbanraja FR, Mahesworo B, Cenggoro TW, Budiarto A, Pardamean B. An evaluation of deep neural network performance on limited protein phosphorylation site prediction data. Procedia Comput Sci 2019;157:25–30. doi:10.1016/j.procs.2019.08.137.

[22] Farida, Caraka RE, Cenggoro TW, Pardamean B. Batik Parang Rusak Detection Using Geometric Invariant Moment. 2018 Indones. Assoc. Pattern Recognit. Int. Conf., IEEE; 2018, p. 71–4. doi:10.1109/INAPR.2018.8627000.

[23] Muchtar K, Rahman F, Cenggoro TW, Budiarto A, Pardamean B. An Improved Version of Texture-based Foreground Segmentation: Block-based Adaptive Segmenter. Procedia Comput Sci 2018;135:579–86. doi:10.1016/j.procs.2018.08.228.

[24] Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20:273–97.

[25] Qin H, Li X, Liang J, Peng Y, Zhang C. DeepFish: Accurate underwater live fish recognition with a deep architecture. Neurocomputing 2016;187:49–58. doi:10.1016/j.neucom.2015.10.122.

[26] Salmon A, Jalal A, Shafait F, Mian A, Shortis M, Seager J, et al. Fish species classification in unconstrained underwater environments based on deep learning. Limnol Oceanogr Methods 2016;14:570–85. doi:10.1002/lom3.10113.

[27] Sun X, Shi J, Dong J, Wang X. Fish recognition from low-resolution underwater images. Proc - 2016 9th Int Congr Image Signal Process Biomed Eng Informatics, CISP-BMEI 2016 2017:471–6. doi:10.1109/CISP-BMEI.2016.7852757.

[28] Chen G, Sun P, Shang Y. Automatic fish classification system using deep learning. Proc - Int Conf Tools with Artif Intell ICTAI 2018;2017-Novem:24–9. doi:10.1109/ICTAI.2017.00016.

[29] Pardamean B, Cenggoro TW, Chandra BJ, Rahutomo. Data Annotation System for Intelligent Energy Conservator in Smart Building. IOP Conf Ser Earth Environ Sci 2020.

[30] Cenggoro TW, Tanzil F, Aslamiah AH, Karuppiah EK, Pardamean B. Crowdsourcing annotation system of object counting dataset for deep learning algorithm. IOP Conf Ser Earth Environ Sci 2018;195. doi:10.1088/1755-1315/195/1/012063.

[31] Qiu C, Zhang S, Wang C, Yu Z, Zheng H, Zheng B. Improving transfer learning and squeeze-and-excitation networks for small-scale fine-grained fish image classification. IEEE Access 2018;6:78503–12. doi:10.1109/ACCESS.2018.2885055.

[32] Allken V, Handegard NO, Rosen S, Schreyeck T, Mahiout T, Malde K. Fish species identification using a convolutional neural network trained on synthetic data. ICES J Mar Sci 2019;76:342–9. doi:10.1093/icesjms/fsy147.

[33] Meng L, Hirayama T, Oyanagi S. Underwater-Drone with Panoramic Camera for Automatic Fish Recognition Based on Deep Learning. IEEE Access 2018;6:17880–6. doi:10.1109/ACCESS.2018.2820326.

[34] Khalifa NEM, Taha MHN. Aquarium Family Fish Species Identification System Using Deep Neural Networks. Proceedings of the International Conference on Advanced Intelligent Systems and Informatics 2018. vol. 2018. Springer International Publishing; 2019. doi:10.1007/978-3-319-99010-1.

[35] LeCun Y, Boser B, Denker JS, Henderson D, Howard RE, Hubbard W, et al. Backpropagation Applied to Handwritten Zip Code Recognition. Neural Comput 1989;1:541–51. doi:10.1162/neco.1989.1.4.541.

[36] Pardamean B, Cenggoro TW, Rahutomo R, Budiarto A, Karuppiah EK. Transfer Learning from Chest X-Ray Pre-trained Convolutional Neural Network for Learning Mammogram Data.
Procedia Comput Sci 2018;135:400–7. doi:10.1016/j.procs.2018.08.190.

[37] Pardamean B, Muljo HH, Cenggoro TW, Chandra BJ, Rahutomo R. Using transfer learning for smart building management system. J Big Data 2019;6:110. doi:10.1186/s40537-019-0272-6.

[38] He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition. 2016 IEEE Conf. Comput. Vis. Pattern Recognit., vol. 7, IEEE; 2016, p. 770–8.
doi:10.1109/CVPR.2016.90.

[39] Johnson JM, Khoshgoftaar TM. Survey on deep learning with class imbalance. J Big Data 2019;6:27.

[40] Fanny, Cenggoro TW. Deep Learning for Imbalance Data Classification using Class Expert Generative Adversarial Network. Procedia Comput Sci 2018;135:60–7.
doi:10.1016/j.procs.2018.08.150.

[41] Cenggoro TW, Isa SM, Kusuma GP, Pardamean B. Classification of imbalanced land-use/land-cover data using variational semi-supervised learning. 2017 Int. Conf. Innov. Creat. Inf. Technol., IEEE; 2017, p. 1–6. doi:10.1109/INNOCIT.2017.8319149.