DOI: 10.31398/tpjf/31.2.2023-0042

SHORT COMMUNICATION

Skipjack Tuna (*Katsuwonus pelamis*) Otolith Developmental Stage Classification Using Deep Learning

Malona V. Alinsug[1,2*], Ian Val P. Delos Reyes[3], Amiel Christian C. Maquiling[1], Anthony John Bercades[1], Custer C. Deocaris[2]

[1] School of Graduate Studies, Mindanao State University – General Santos City 9500, South Cotabato, Philippines
[2] Atomic Research Division, Philippine Nuclear Research Institute, Department of Science & Technology, Commonwealth Ave, Diliman 1101, Quezon City, Philippines
[3] Institute of Computing, Davao del Norte State College, Panabo City, Davao del Norte, Philippines

ABSTRACT

The Philippines is the second biggest source of skipjack tuna (*Katsuwonus pelamis*), contributing to the country's economic development. However, its sustainability faces challenges due to overfishing and a lack of proper management practices. Otoliths are important tools for managing fish stocks, but their analysis is time-consuming and requires a high level of expertise. In this paper, we explored the use of convolutional neural networks (CNNs) to recognize patterns and classify them according to developmental stages. The results showed that the CNN model achieved an accuracy of 100% in classifying otoliths by developmental stage using the RMSprop optimizer, demonstrating the potential of deep learning to provide a standardized and reliable protocol for managing fish stocks in countries like the Philippines, where there is a shortage of trained fish experts. This study provides an innovative approach to guide future efforts in conserving fish populations and promoting sustainable fishing practices.

*Corresponding Author: *malonava@pnri.dost.gov.ph*
Received: *June 24, 2023*
Accepted: *July 22, 2024*

Keywords: *Artificial intelligence, convolutionalneural networks, fish development stageclassification, sustainable fisheries management*

The Philippines is one of the world's largest producers and exporters of skipjack tuna (*Katsuwonus pelamis*) (Williams and Terawasi 2010; Tahiluddin and Terzi 2021). However, skipjack tuna fisheries' sustainability in the Philippines faces numerous challenges due to overfishing and a lack of proper management practices (Zaragosa et al. 2004). According to the Bureau of Fisheries and Aquatic Resources (BFAR) and the National Fisheries Research and Development Institute (NFRDI), overfishing, continuous fishing of juveniles, and climate change are the major threats to skipjack tuna fisheries in the country. Furthermore, as this tuna species is identified as a highly migratory fish under the United Nations Convention on the Law of the Sea, its mobility in the Pacific region adds complexity to fisheries management (Sibert and Hampton 2003). Thus, addressing these challenges is crucial to ensure the long-term sustainability of skipjack tuna fisheries in the Philippines, which requires improved management practices, the use of sustainable fishing techniques, and international cooperation.

Otoliths or ear-stones are hard, calcium carbonate structures located within the inner ear of bony fishes. These structures are considered essential for managing fish stocks as they contain vital information about the fish's age, growth, and habitat, which can help determine fishing quotas and season lengths. Nazir and Khan (2021) note that otoliths are used for fish stock discrimination, providing an accurate way to identify different fish stocks in a region. They also reveal information about mortality rates, diet, and exposure to pollutants (Morison et al. 2005). However, using otoliths presents challenges, such as the complexity of interpreting the results, high cost, laborious, and time-consuming analysis. Moreover, as this task requires accuracy and precision, there is a shortage of trained fish experts in certain countries who can provide the analysis. Avigliano (2022) discusses the importance of optimizing the methodological design for fish stock delineation from otolith chemistry and reviewing spatiotemporal analysis scales to overcome these issues. However, given the challenges and large quantity of samples required, a standardized and reliable protocol is necessary to ensure accurate and reproducible data. Monitoring fish populations and managing the fisheries is challenging, given the dynamic nature of fish populations and the vast geographical expanse of the Philippines' marine ecosystems (Lehodey et al. 2013). Thus, innovations to assist the scaling up of this approach can help realize otolith analysis as a practical tool for managing fish stocks and guiding sustainable fishery management in human resource-constrained countries like the Philippines.

Convolutional Neural Networks (CNNs) are deep learning algorithms that can recognize image patterns and classify them accordingly. CNNs have been used for various applications, including image and speech recognition, natural language processing, and medical image analysis (Ordoñez et al. 2020). Recent studies have shown that CNNs can also classify fish otolith images accurately, providing a novel approach to fish age estimation (Moen et al. 2018). In this paper, we aim to develop a CNN-based model for the classification of skipjack tuna otoliths in the Philippines. To our knowledge, no previous study has utilized CNNs to classify skipjack tuna otoliths by developmental stage.

Our study used 50 otoliths per developmental stage, with 10–15 images taken per otolith, to yield a total of 1620 collected images from actual photos. The otoliths were from 170 skipjack tunas collected from various fish ports of Mindanao, including Sarangani Bay, Moro Gulf, Davao Gulf, and Celebes Sea. To reduce sampling bias, the researchers collected tuna across all seasons from 2021 to 2022 and analyzed body morphometrics and otolith shape features (manuscript in preparation). The images of skipjack tuna otoliths were at different developmental stages based on the criteria by Collete and Nauen (1983), which were accepted by the industry: juvenile (S1) with fork length (FL) of less than 40 cm, mature (S2) FL at 40.0–80.0 cm, and adult (S3) FL at 80.01 cm and above. Fork length is a standard measure used to determine the length of a fish, typically used for tunas, which is the measurement from the tip of the fish's nose to the fork in the tail, where the tail splits into two separate points. Each developmental stage has been validated by an independent otolith expert consulting for the project. The otoliths of skipjack tuna exhibit distinct differences across developmental stages (Figure 1A).

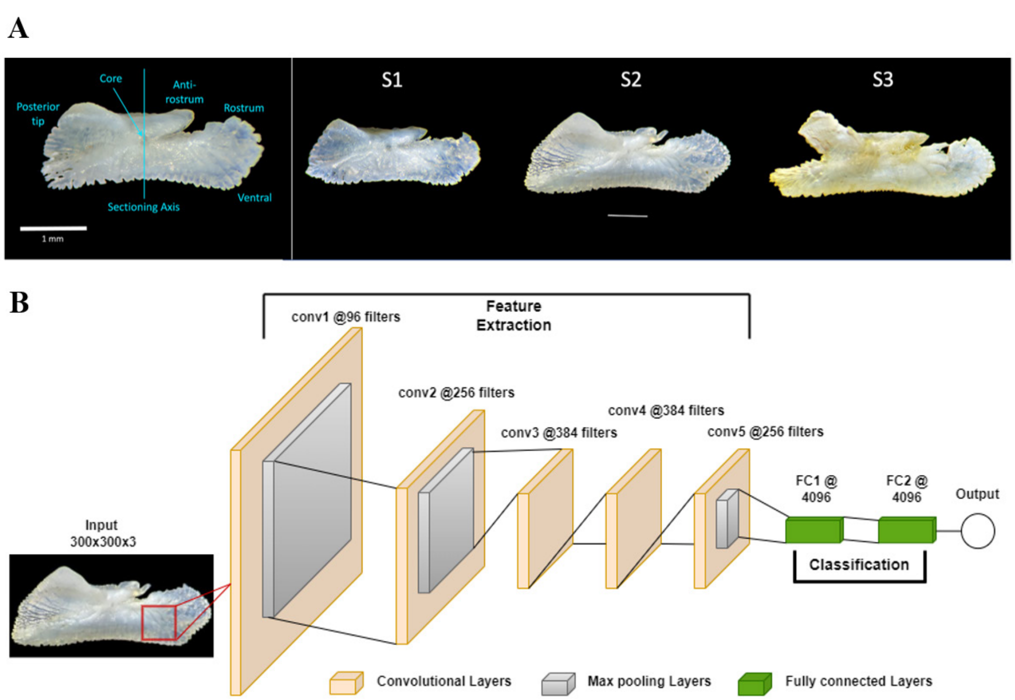

Figure 1. CNN architecture of skipjack tuna otoliths. (A) Otoliths of skipjack tuna exhibit distinct differences across developmental stages. (B) CNN architecture is used to classify the developmental stages based on skipjack tuna otolith images. An input image is passed through five convolutional layers and three max-pooling layers, where filters are applied to perform feature extraction. The classification operation is carried out by the two fully connected layers with a 4096-dimensional feature vector based on the features extracted from the previous layers and their various filters.

The overall size of the otoliths increases significantly from S1, the smallest, to S3, the largest, with the most notable growth occurring between the juvenile stages S1 and S2. The rostrum, a beak-like projection on the anterior part of the otolith, becomes longer, more pointed, and finely serrated as the fish ages from S1 to S3. The sulcus, a groove running along the ventral edge, deepens and straightens in curvature with each successive stage. Moreover, the otolith shape becomes more elongated, with more pronounced growth on the dorsal side compared to the ventral side from S1 to S3. The opacity of the otoliths also increases with maturity, reflecting the accumulation of additional layers over time. While these changes are apparent to an expert, discerning them requires careful analysis and specialized knowledge.

The CNN model architecture used in the study included five convolutional layers, three max-pooling layers, and two fully connected layers (Figure 1B). The input layer contained the raw pixel values of the image data, and each colored image was reshaped into a single column to be fed into the convolutional layer. The convolutional layers were used to extract various features from the input images, while the pooling layers summarized the feature present in a region of the feature map generated by the convolutional layer. The fully connected layers were applied after all the convolutional and max-pooling layers to classify the image into a particular category and associate features to a specific label. Each otolith developmental stage had an image size of 300 x 300 pixels. The datasets for training, validation, and testing were described in Table I. Data augmentation techniques like rotation, horizontal flip, vertical flip, random mirroring, and cropping the images to possibly reduce the amount of overfitting.

Python was used to write all programming codes, employing Keras neural network application programming interfaces and Tensorflow libraries. The computations were carried out on a GeForce RTX 3060 GPU in a 3.7 GHz Core i7 CPU with 32 GB RAM. As the choice of an optimization algorithm can significantly impact the efficiency of training deep learning models, we investigated four optimizers: Adam, RMSProp, SGD, and Adadelta. Adam uses little memory and is invariant to the diagonal rescale of the gradients (Bock and Weis 2019). RMSProp is a technique for accelerating learning by automatically changing the learning rate for each parameter and converging quickly (Xu et al. 2021). SGD updates the parameters more often based on randomly chosen training data, enabling it to converge quickly and be computationally fast (Nakkiran et al. 2019). Adadelta dynamically adapts over time using only first-order information, providing a low computing overhead and being robust to noisy gradient information, varied model architectural choices, diverse data modalities, and selection of hyperparameters while requiring no manual adjustment of the learning rate (Zeiler 2012).

The performance of the CNN models with optimizers was evaluated based on validation loss and accuracy across the training cycles. Validation loss is the difference between the CNN model's predicted value and the actual value using the cross-entropy function:

$$Loss = -\sum_{i=1}^{n} y_i \log(\acute{y}_i)$$

where n (is the number of classes), i (is class), y_i (is the true value of a class) and $\acute{y}_i$ (is the predicted value for that class). Validation accuracy, on the other hand, measures the model performance during training by calculating the number of rows where the neural network correctly predicted the target class, as shown in the equation:

$$Accuracy = \frac{c}{N}$$

where c (is the number of correct identifications) and N (is the size of the evaluated training set or the validation set).

The performance of four optimizers was comparatively evaluated based on computational effectiveness and speed during 30 epochs of training, as depicted in Figure 2. Lower validation loss indicated superior CNN model performance, while higher accuracy values signified the model's ability to classify data accurately. Our results revealed that RMSProp optimization was the optimal approach to mitigate the overfitting issue associated with CNN, as it achieved the lowest validation loss (0.01%) and highest accuracy (100%) compared to the other optimizers. The decrease in validation loss signaled increased confidence in the model's ability to classify samples accurately, and RMSProp outperformed the other optimization techniques as it exhibited a continuous decrease in validation loss from the 12 to

Table 1. Dataset distribution of otolith images for CNN analysis.

Skip Jack Tuna	Training	Validation	Testing	Total
S1 vs S2	504	144	72	720
S1 vs S3	504	144	72	720
S2 vs S3	504	144	72	720
S1 vs S2 vs S3	756	216	108	1080

the 17th epoch. Additionally, RMSProp demonstrated superiority over the other techniques, as it experienced a consistent increase in validation accuracy from the 17th epoch onwards.

After training, we then tested the robustness of making accurate predictions using our CNN model. Of note, training, and validation data included labels to monitor the performance metrics of the model; at

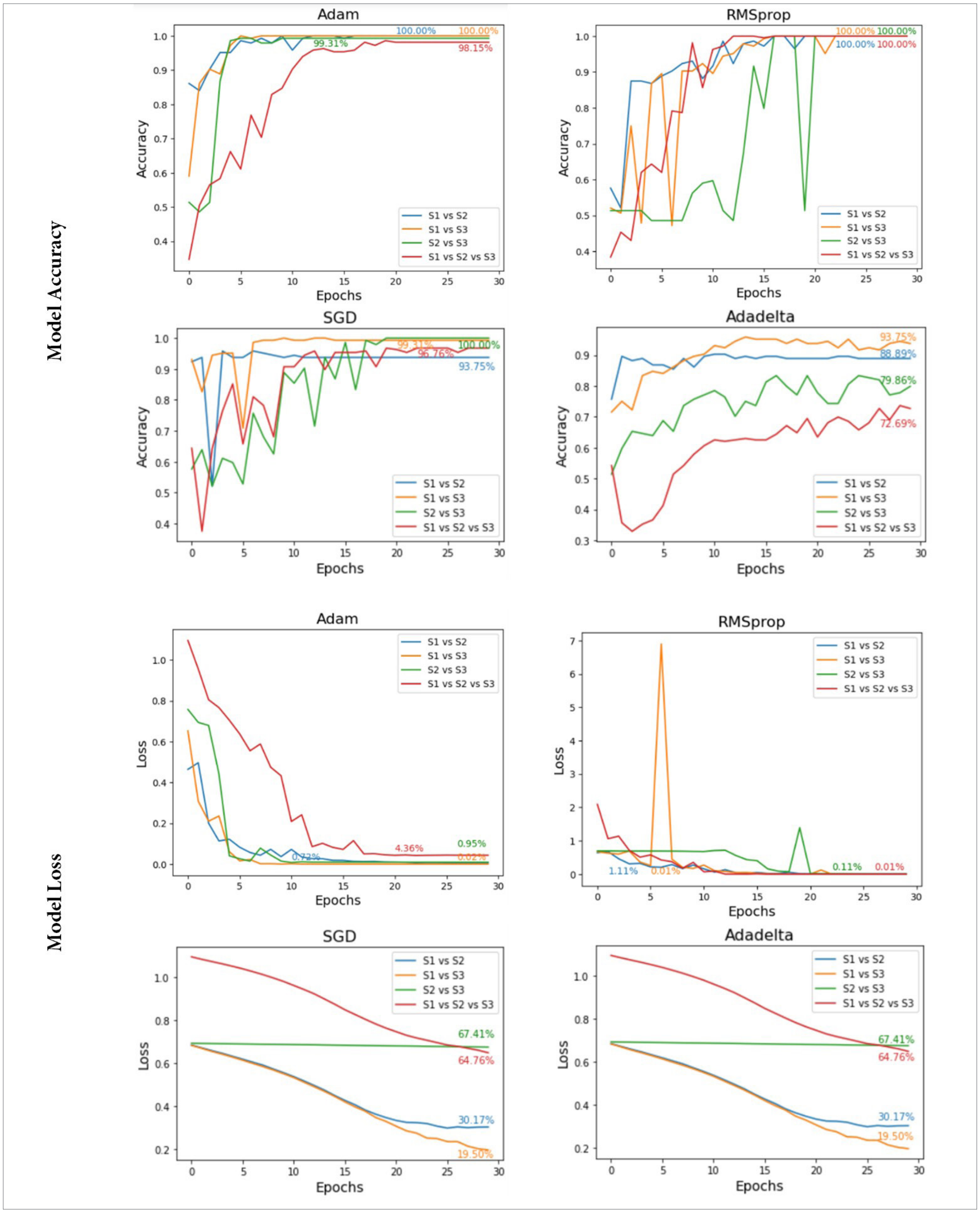

Figure 2. Comparative model performance during training on skipjack tuna otoliths at the different development stages, S1, S2, and S3. The model accuracy and model loss results from different CNN optimizers for the 0th to 30th epochs are displayed.

the testing phase, all the input data was unlabeled. Test (or unseen) datasets are used to confirm whether our CNN model was effectively trained. Model performance was verified based on precision, recall, and F1-score. Precision measures how accurate the model is at categorizing a positive sample. It is obtained using the equation:

$$Precision = \frac{TP}{TP + FP}$$

where TP is true positive and FP is false positive. Precision ranges from 0 (for no precision) to 1 (for full or perfect precision). Recall represents the number of correct positive predictions made from all the positive predictions expressed as:

$$Recall = \frac{TP}{TP + FN}$$

where FN refers to false negative. Recall values range from 1 (optimum) to zero.

The F1-Score is a metric of the balance between precision and recall. A high F1-Score indicates that the model has both high precision and recall. The score ranges from 0 to 1, with 1 being the best possible score. This metric can be calculated using the following equation:

$$\text{F1-Score} = 2 \times \left(\frac{Precision \times Recall}{Precision + Recall}\right)$$

The performance of a classification model can be effectively illustrated through confusion matrix or error matrix plots, as presented in Figure 3, for better analyses. These plots summarize the prediction results of confidently differentiating the development stages from S1 to S3 using different optimizers, providing an overview of the correct classification for the 108 test images. The numbers are shown, with rows corresponding to the predicted ranking and columns specifying the actual classification. Likewise, the results indicate that the RMSprop optimization technique achieves the highest average classification accuracy of 100%, outperforming Adam, SGD, and Adadelta optimizers used in the study, with accuracy rates of 98.77%, 98.44%, and 78.40% across S1, S2, and S3. The CNN model classification performance was evaluated using precision, recall, and F1-score, as shown in Table 2. The results consistently indicated that RMSprop outperformed the other optimizers, achieving 100% precision, recall, and F1-score. Pair-wise comparisons of developmental stages showed that S2 vs. S3 and S1 vs. S3 had the highest precision, recall, and F1-score, while the comparison of S1 vs. S2 had the lowest score for all the optimizers except RMSprop. This observation suggests that the changes in skipjack tuna during the transition from juvenile to pre-adult stages are more subtle than other stages.

	Adam		Actual			
Predicted	Frigate	Total	S1	S2	S3	Accuracy
	S1	35	35	0	0	98.15%
	S2	36	2	34	0	98.15%
	S3	37	0	0	37	100.00%
		108	Average			98.77%

	RMSprop		Actual			
Predicted	Frigate	Total	S1	S2	S3	Accuracy
	S1	37	37	0	0	100.00%
	S2	31	0	31	0	100.00%
	S3	40	0	0	40	100.00%
		108	Average			100.00%

	SGD		Actual			
Predicted	Frigate	Total	S1	S2	S3	Accuracy
	S1	35	35	0	0	96.30%
	S2	40	4	31	5	91.67%
	S3	33	0	0	33	95.37%
		108	Average			94.44%

	Adadelta		Actual			
Predicted	Frigate	Total	S1	S2	S3	Accuracy
	S1	40	40	0	0	88.89%
	S2	36	7	25	4	72.22%
	S3	32	5	19	8	74.07%
		108	Average			78.40%

Figure 3. Confusion matrix and comparative classification accuracy of different CNN optimizers using test datasets on skipjack tuna otoliths at the different development stages, S1, S2, and S3.

Table 2. Comparative classification performance using the different CNN optimizers on skipjack tuna otoliths at the different development stages, S1, S2, and S3. Performance parameters (precision, recall, and F1-score) of the trained CNN model on the test dataset are indicated.

Development Stage	Optimizer	Precision (%)	Recall (%)	F1-Score (%)
S1 vs S2	Adam	98.48	98.75	98.60
	RMSprop	100.00	100.00	100.00
	SGD	90.71	89.86	90.12
	Adadelta	89.81	81.03	82.63
S1 vs S3	Adam	100.00	100.00	100.00
	RMSprop	100.00	100.00	100.00
	SGD	99.37	99.24	99.30
	Adadelta	90.39	90.48	90.28
S2 vs S3	Adam	99.30	99.32	99.31
	RMSprop	100.00	100.00	100.00
	SGD	100.00	100.00	100.00
	Adadelta	69.06	69.06	69.06
S1 vs S2 vs S3	Adam	98.20	98.15	98.12
	RMSprop	100.00	100.00	100.00
	SGD	92.20	92.50	91.63
	Adadelta	66.80	64.81	61.94

Several other forms of artificial intelligence have been carried out on fish otoliths with different levels of success. One such study conducted by Moore et al. (2019) used computed tomography (CT) scanning and machine learning to automate the aging process of fish otoliths with impressive accuracy. Their model accurately estimated the age of 47% of snapper *(Pagrus auratus*) and 41% of hoki (*Macruronus novaezelandiae*) within one year of the human reader estimate of age. Similarly, Moen et al. (2018) employed deep learning to interpret otoliths of Greenland halibut (*Reinhardtius hippoglossoides*) with a precision comparable to human experts. Taking these findings further, Politikos et al. (2021) developed a multitask deep learning model that combines otolith images and age data to accurately estimate the age of several fish species, including European hake (*Merluccius merluccius*), European anchovy (*Engraulis encrasicolus*), and Atlantic mackerel (*Scomber scombrus*). The results of this study indicate the potential of deep learning tools to enhance the accuracy and efficiency of fish stock assessments. Moreover, to make these AI tools more accessible, Politikos et al. (2022) developed DeepOtolith v1.0, an open-source platform that employs deep learning algorithms to automate fish age reading from otolith or scale images. With a user-friendly interface and the ability to analyze large datasets efficiently, DeepOtolith v1.0 has the potential to revolutionize the field of fish stock assessments.

In conclusion, our study successfully utilized a convolutional neural network (CNN) model to classify the developmental stage of skipjack tuna otoliths. We employed a rigorous methodology to collect and process otolith images, and the CNN architecture was designed with careful consideration for optimal image analysis. The evaluation of the CNN model performance during training, using validation loss and accuracy metrics, provided a quantitative assessment of the model's success. It is worth noting that while the study did not compare the results with human expert validation, our results, nonetheless, could be a promising step towards further application of computer vision integrated with deep learning in the field of fish otolith analysis and tuna fisheries management. Finally, while CNNs remain popular as a robust image classifier, exploring newer architectures like state-of-the-art vision transformers (ViTs) and hybrid approaches can lead to improved image classification. However, in order to combine different predictors to improve the precision, future work also needs to verify that the predictors are independent and unbiased when being applied to the same samples.

ACKNOWLEDGEMENTS

This research was funded by the Department of Science & Technology – Philippine Council for Agriculture, Aquatic, and Natural Resources Research

Development (DOST-PCAARD) Tuna Program under the program leadership of Dr. MV Alinsug.

AUTHOR CONTRIBUTIONS

Alinsug MV: Conceptualization, methodology, data collection, formal analysis, writing – draft & editing, funding acquisition. **Delos Reyes IV**: Conceptualization, methodology, software, validation, formal analysis, writing – draft. **Maquiling AC**: Data collection, software, validation, investigation. **Bercades AJ**: Data collection, software, validation, investigation. **Deocaris CC**: Methodology, formal analysis, writing – original & editing.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

ETHICS STATEMENT

The authors obtained a permit from the Department of Agriculture-Bureau of Fisheries and Aquatic Resources to sample tuna at Sarangani Bay, Moro Gulf, Davao Gulf, and Celebes Sea throughout the duration of the research project. This permit ensures compliance with relevant regulations governing the collection and use of tuna for research purposes in the Philippines.

REFERENCES

Avigliano E. 2022. Optimizing the methodological design in fish stock delineation from otolith chemistry: review of spatio-temporal analysis scales. Rev Fish Sci Aquac. 30(3):330–345. https://doi.org/10.1080/23308249.2021.1961679

Bock S, Weis M. 2019. A proof of local convergence for the Adam optimizer. Proceedings of the International Joint Conference on Neural Networks. 1–8. https://doi.org/10.1109/IJCNN.2019.8852239

Collette B, Nauen C. 1983. Scombrids of the world: an annotated and illustrated catalogue of tunas, mackerels, bonitos and related species known to date. FAO Species Catalogue, Vol. 2.

Lehodey P, Senina I, Calmettes B, Hampton J, Nicol S. 2013. Modelling the impact of climate change on Pacific skipjack tuna population and fisheries. Climatic Change. 119:95–109. https://doi.org/10.1007/s10584-012-0595-1

Moen E, Handegard NO, Allken V, Albert OT, Harbitz A, Malde K. 2018. Automatic interpretation of otoliths using deep learning. PLoS One. 13(12):e0204713. https://doi.org/10.1371/journal.pone.0204713

Moore BR, Maclaren J, Peat C, Anjomrouz M, Horn PL, Hoyle S. 2019. Feasibility of automating otolith ageing using CT scanning and machine learning. New Zealand Fish Assess Rep 2019/58. p. 23.

Morison AK, Burnett J, McCurdy WJ, Moksness E. 2005. Quality issues in the use of otoliths for fish age estimation. Mar Freshw Res. 56(5):773–782. https://doi.org/10.1071/MF04217

Nakkiran P, Kaplun G, Kalimeris D, Yang T, Edelman B, Zhang H, Barak B. 2019. SGD on neural networks learns functions of increasing complexity. Proc 33rd Conf Neural Infor Processing Sys. 314:3496–3506.

Nazir A, Khan MA. 2021. Using otoliths for fish stock discrimination: Status and challenges. Acta Ichthyol Piscat. 51(2):199–218. https://doi.org/10.3897/aiep.51.64166

Ordoñez A, Eikvil L, Salberg A-B, Harbitz A, Murray SM, Kampffmeyer MC. 2020. Explaining decisions of deep neural networks used for fish age prediction. PLoS ONE. 15(6):e0235013. https://doi.org/10.1371/journal.pone.0235013

Politikos DV, Petasis G, Chatzispyrou A, Mytilineou C, Anastasopoulou A. 2021. Automating fish age estimation combining otolith images and deep learning: The role of multitask learning. Fish Res. 242:106033. https://doi.org/10.1016/j.fishres.2021.106033

Politikos DV, Sykiniotis N, Petasis G, Dedousis P, Ordoñez A, Vabø R, Anastasopoulou A, Moen E, Mytilineou C, Salberg AB, Chatzispyrou A, Malde K. 2022. DeepOtolith v1. 0: An Open-Source AI Platform for Automating Fish Age Reading from Otolith or Scale Images. Fishes. 7(3):121. https://doi.org/10.3390/fishes7030121

Sibert J, Hampton J. 2003. Mobility of tropical tunas and the implications for fisheries management. Mar Policy. 27(1): 87-95. https://doi.org/10.1016/S0308-597X(02)00057-X

Tahiluddin A, Terzi E. 2021. An overview of fisheries and aquaculture in the Philippines. J Anatol Environ Anim Sci. 6(4): 475-86. https://doi.org/10.35229/jaes.944292

Williams PG, Terawasi P. 2010. Overview of tuna fisheries in the western and central Pacific Ocean, including economic conditions–2009. Presented during the Scientific Committee 6th Regular Sessions of the Western and Central Pacific Fisheries Commission in Naku'alofa, Tonga on 10-19 August 2010. WCPFC-SC6-2010/GN WP-1. pp. 66 https://meetings.wcpfc.int/file/5814/download

Xu D, Zhang S, Zhang H, Mandic DP. 2021. Convergence of the RMSProp deep learning method with penalty for nonconvex optimization. Neural Networks. 139:17–23. https://doi.org/10.1016/j.neunet.2021.02.011

Zaragosa EC, Pagdilao CR,Moreno EP. 2004. Fisheries for tuna and other large pelagic fishes. turbulent seas: The status of Philippine marine fisheries. Cebu City: Coastal Resource Management Project. pp. 38–41.

Zeiler MD. 2012. ADADELTA: An Adaptive Learning Rate Method. http://arxiv.org/abs/1212.5701