

Article

Not peer-reviewed version

Otolith Image Based Age Classification of Japanese Jack Mackerel *Trachurus japonicus* Using Convolutional Neural Networks

[Min-Su You](#) and [Chul-Woong Oh](#)*

Posted Date: 28 May 2026

doi: 10.20944/preprints202605.1964.v1

Keywords: age classification; convolutional neural network; deep learning; fish ageing; Otolith image; *Trachurus japonicus*



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Otolith Image Based Age Classification of Japanese Jack Mackerel *Trachurus japonicus* Using Convolutional Neural Networks

Min-Su You ¹ and Chul-Woong Oh ^{1,2,*}

¹ Department of Industry 4.0 Convergence Bionics Engineering, Pukyong National University, Busan 48513, Republic of Korea

² Department of Marine Biology, Pukyong National University, Busan 48513, Republic of Korea

* Correspondence: ohcw@pknu.ac.kr

Abstract

Reliable age information is needed for fisheries assessment, but conventional otolith reading requires trained readers and considerable time. This study evaluated whether convolutional neural networks could classify reader assigned age classes of Japanese jack mackerel *Trachurus japonicus* directly from sagittal otolith images. Otolith images from fish aged 0 to 4 years were used to compare three image only backbones: Inception v3, Xception, and EfficientNet B4. The models were trained under the same data split, preprocessing, augmentation, and evaluation framework. In Stage 1, Inception v3 showed the highest validation macro F1 score (0.933) and was selected as the image only baseline. After additional optimization, the selected model reached a validation macro F1 score of 0.944, validation exact accuracy of 0.935, and validation agreement within one age class of 1.000. On the independent test set, the optimized image only model achieved exact accuracy of 0.866, macro F1 score of 0.873, and agreement within one year of 1.000. These results indicate that otolith images contain useful age related visual information. Convolutional neural networks may support age class screening in *T. japonicus*, although they should complement rather than replace expert otolith reading.

Keywords: age classification; convolutional neural network; deep learning; fish ageing; Otolith image; *Trachurus japonicus*

1. Introduction

Age information is one of the basic requirements in fisheries science. It is used to estimate growth, mortality, recruitment, and age composition. These estimates are then used in stock assessment and fisheries management [1,2]. In many teleost fishes, age is still determined by counting annual increments in sagittal otoliths. This method remains the standard approach, but it is not simple. It requires trained readers, careful preparation of otoliths, and repeated checking when growth marks are unclear [3,4]. When many samples are collected for monitoring programs, this process can become slow and labor intensive.

The Japanese jack mackerel *Trachurus japonicus* is an important small pelagic fish in Korean waters and adjacent seas. It is distributed in the East China Sea, the Yellow Sea, the southern coast of Korea, and waters around Japan [5]. In Korea, this species is one of the major species caught by large purse seine fisheries and is included in fisheries monitoring and management programs [6]. Reliable age information is therefore needed to describe its age structure and to support stock assessment. Previous studies have estimated age and growth of *T. japonicus* using sagittal otolith annuli in Korean and Japanese waters [6–8]. Even so, routine otolith reading still requires considerable time and reader experience.

Image based approaches have been explored as a way to support fish age estimation. Early methods often relied on manually selected image features, image processing, or statistical learning

[9,10]. These approaches showed that otolith images contain age related information, but their performance often depended on how the image features were selected. Deep learning offers a different approach. Convolutional neural networks (CNNs) can learn image features directly from otolith images, without requiring all relevant features to be defined in advance [11,12]. Recent studies have reported promising results for otolith based age estimation in several commercially important fishes, including Greenland halibut, Atlantic cod, red mullet, hoki, and snapper [13–17].

The choice of CNN architecture may affect model performance. Inception based models use convolutional filters at different spatial scales and can capture image patterns with different local resolutions [18]. Xception uses depthwise separable convolutions and can provide efficient feature extraction from image data [19]. EfficientNet models use compound scaling of network depth, width, and input resolution. EfficientNet B4 has also shown useful performance in recent otolith age interpretation studies [14,20]. However, no single architecture can be assumed to be optimal for all otolith datasets. Performance may depend on species, image quality, age range, sample size, and the visual structure of the otolith increments.

For *T. japonicus*, it remains unclear whether reader assigned age classes can be classified directly from sagittal otolith images using CNNs. It is also not known which commonly used backbone architecture provides the most suitable image only baseline for this species. Establishing this baseline is useful before adding other biological or morphometric information. It also allows the image based signal to be evaluated separately from measured otolith size or shape variables.

The aim of this study was to compare the performance of three CNN backbones for age class classification of *T. japonicus* using sagittal otolith images only. Inception v3, Xception, and EfficientNet B4 were trained and evaluated under the same preprocessing, augmentation, and validation framework. The purpose was not to replace expert otolith reading. Instead, this study tested whether otolith image based deep learning can provide a practical support tool for screening reader assigned age classes in this commercially important pelagic fish.

2. Materials and Methods

2.1. Otolith Image Dataset and Age Labels

Digital images of sagittal otoliths from *T. japonicus* were used in this study. The age classes were not newly assigned for the deep learning analysis. Instead, they were linked to the conventional otolith readings obtained from sagittal otolith annuli. The model task was therefore to classify reader assigned age classes from otolith images.

A total of 1,444 otolith images were retained for age class analysis. The dataset included age classes 0 to 4. The number of samples was 134 at age 0, 270 at age 1, 567 at age 2, 265 at age 3, and 208 at age 4 (Table 1). Each retained sample consisted of one otolith image and one age label. Samples with incomplete image information or unclear image quality were not used in the analysis.

Table 1. Sample composition of *Trachurus japonicus* otolith images used for age class classification.

Age class	Number of images	Percentage (%)
0	134	9.3
1	270	18.7
2	567	39.3
3	265	18.4
4	208	14.4
Total	1444	100.0

2.2. Dataset Partitioning

The full dataset was divided into training, validation, and test sets. The training set contained 70% of the data, and the validation and test sets each contained 15%. Stratified random sampling was used so that the relative proportions of age classes were preserved across all subsets.

The data split was generated once and then fixed for all experiments. This was done so that the benchmark models were evaluated on the same training, validation, and test subsets. Under this design, differences in performance could be interpreted mainly as differences among model architectures or training settings rather than differences in dataset composition.

2.3. Image Preprocessing and Data Augmentation

All otolith images were preprocessed in a standardized way before they were passed to the CNNs. Each image was padded to a square format while preserving the original aspect ratio. Padding was used because simple resizing could distort the otolith outline. The padded images were then resized according to the expected input size of each backbone.

The image size was 299×299 pixels for Inception v3 and Xception, and 380×380 pixels for EfficientNet B4. Images were converted to tensors and normalized using the ImageNet mean and standard deviation because all networks were initialized with ImageNet pretrained weights [13,14,20].

During training only, mild image augmentation was applied. This included random rotation within $\pm 10^\circ$ and small random changes in brightness and contrast. Flipping was not applied because otolith orientation was treated as meaningful. The augmentation strategy was deliberately conservative and was intended to reflect plausible variation in image acquisition while preserving age related structures (Table 2).

Table 2. Image augmentation strategies used in previous otolith deep learning studies and in the present study.

Study	Species	Augmentation	Note
Moen et al. [13]	Greenland halibut	Rotation 0–360°; horizontal/vertical reflection; vertical shift	Relatively strong geometric augmentation
Martinsen et al. [14]	Greenland halibut	Horizontal translation 10%; rotation up to 36°	Moderate augmentation
Present study	Japanese jack mackerel	Rotation $\pm 10^\circ$; small brightness and contrast changes	Mild augmentation chosen to reflect realistic image acquisition variation

2.4. Convolutional Neural Network Backbones

Three image only CNN backbones were compared: Inception v3, Xception, and EfficientNet B4. The term backbone refers to the main feature extraction part of the network. In this study, each backbone received an otolith image and transformed it into a feature vector for age classification.

Inception v3 was included because it is a well established image classification architecture that uses factorized convolutions and multi scale feature extraction [18]. Xception was included because it replaces standard convolutional operations with depthwise separable convolutions, which can improve parameter efficiency [19]. EfficientNet B4 was included because EfficientNet models use compound scaling of network depth, width, and input resolution, and this family has shown strong performance in image analysis tasks [20].

All three backbones were initialized with ImageNet pretrained weights and then fine tuned on the otolith image dataset. Transfer learning was used because pretrained visual features can provide

a more stable starting point than random initialization when the target dataset is much smaller than large computer vision datasets [21,22].

2.5. Image Classification Model

The original classification layer of each backbone was removed and replaced with the same task specific classifier head. This allowed the three backbones to be compared under a common output structure. Let h denote the image feature vector extracted by the backbone. The feature vector was passed to a fully connected classifier with one hidden layer, a rectified linear unit activation function, dropout, and a final output layer.

The output logits were converted to class probabilities using a softmax function. The number of classes was five, corresponding to age classes 0, 1, 2, 3, and 4. The model used only the otolith image as input. No morphometric or tabular variables were included in this manuscript.

2.6. Model Training and Optimization

All models were implemented in Python using the PyTorch framework, and backbone architectures were loaded through the timm library. Training was conducted on a Windows based personal computer with GPU acceleration using an NVIDIA GeForce RTX 3060 graphics card. Random seeds for Python, NumPy, and PyTorch were fixed before model fitting to improve reproducibility.

Models were trained using weighted cross entropy loss to reduce the effect of class imbalance. The network parameters were optimized using AdamW, which applies decoupled weight decay to adaptive gradient based optimization [23,24]. The maximum number of training epochs was 50 for each experiment. Model performance was monitored using validation macro F1 score.

Hyperparameter tuning was conducted in two stages. In Stage 1, the three backbones were compared under three learning rates: 3×10^{-5} , 1×10^{-4} , and 3×10^{-4} . The backbone with the highest validation macro F1 score was selected as the image only baseline. In Stage 2, the selected backbone was further optimized by searching combinations of learning rate, weight decay, and dropout. The learning rate was reduced when validation macro F1 did not improve for three consecutive epochs. Early stopping was applied when no improvement was observed for seven consecutive epochs. The model state with the highest validation macro F1 was retained.

2.7. Model Evaluation

Model selection was based primarily on validation macro F1 score. Macro F1 was used because the age class distribution was not uniform and because this metric gives equal weight to each class [25,26]. Validation exact accuracy, validation loss, and validation agreement within one age class were also examined.

After model selection, the optimized image only model was evaluated on the independent test set. Test performance was summarized using exact accuracy, macro F1 score, and agreement within one year. Agreement within one year was included because age classes are ordered and misclassification to an adjacent age class has a different practical meaning from misclassification to a distant age class.

3. Results

3.1. Dataset Composition

The retained dataset consisted of 1,444 otolith images from fish aged 0 to 4 years (Table 1). Age 2 was the most frequent class, followed by age 1, age 3, age 4, and age 0. The stratified split preserved this age structure across the training, validation, and test subsets.

3.2. Stage 1 Backbone Comparison

In Stage 1, Inception v3, Xception, and EfficientNet B4 were compared under three learning rate settings. Performance differences among the best performing settings were relatively small, but Inception v3 achieved the highest validation macro F1 score among all Stage 1 trials (Table 3; Figure 1). The best result was obtained by Inception v3 at a learning rate of 3×10^{-5} , which reached a validation macro F1 score of 0.933 at epoch 15.

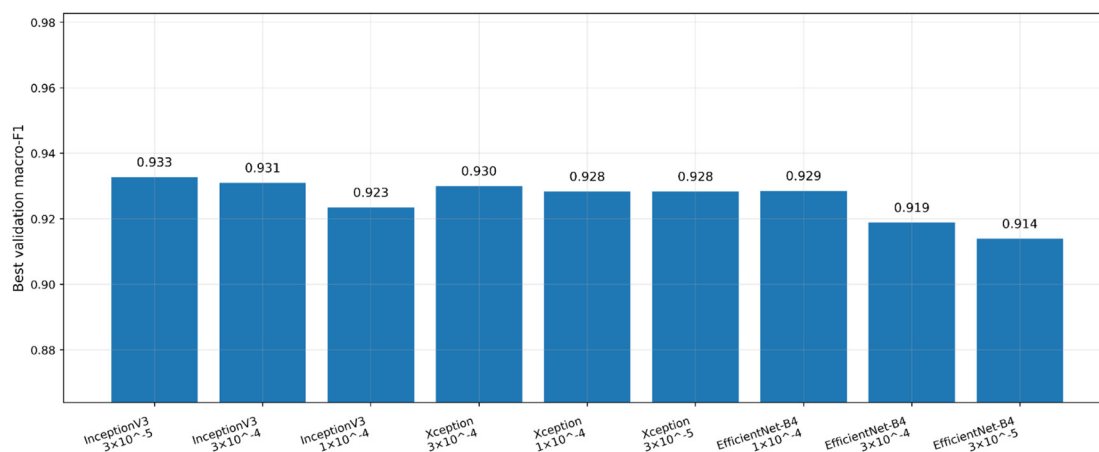


Figure 1. Comparison of the best validation macro F1 scores obtained in Stage 1 across three image only backbones and three learning rate settings.

Table 3. Stage 1 comparison of image only CNN backbones and learning rate settings.

Backbone	Learning rate	Best epoch	Validation macro F1	Validation loss	Validation exact accuracy
Inception v3	3×10^{-5}	15	0.932716	0.219860	0.926407
Inception v3	1×10^{-4}	13	0.923461	0.230347	0.917749
Inception v3	3×10^{-4}	14	0.930988	0.233401	0.926407
Xception	3×10^{-5}	27	0.928360	0.224112	0.922078
Xception	1×10^{-4}	6	0.928362	0.206149	0.926407
Xception	3×10^{-4}	18	0.930014	0.212586	0.922078
EfficientNet B4	3×10^{-5}	26	0.913933	0.284962	0.900433
EfficientNet B4	1×10^{-4}	21	0.928537	0.211526	0.926407
EfficientNet B4	3×10^{-4}	13	0.918861	0.254291	0.909091

Inception v3 also showed stable performance at 3×10^{-4} , with a validation macro F1 score of 0.931 at epoch 14. Xception showed competitive performance across all learning rates. Its best result was obtained at 3×10^{-4} , where the validation macro F1 score reached 0.930 at epoch 18. EfficientNet B4 showed its best performance at 1×10^{-4} , reaching a validation macro F1 score of 0.929 at epoch 21. The other two learning rate settings for EfficientNet B4 showed lower scores, suggesting greater sensitivity to the tested optimization scale.

Validation loss and macro F1 rankings were not identical. The lowest validation loss was observed for Xception at 1×10^{-4} , but this setting did not produce the highest macro F1 score. Because the main objective was balanced classification across age classes, macro F1 was retained as the primary model selection criterion. Inception v3 with a learning rate of 3×10^{-5} was therefore selected as the Stage 1 image only baseline.

Validation exact accuracy reached 0.926 for several settings, including Inception v3 at 3×10^{-5} , Xception at 1×10^{-4} , and EfficientNet B4 at 1×10^{-4} . Validation agreement within one age class was extremely high across nearly all models. This indicates that most errors, when they occurred, were between neighboring age classes rather than distant age classes (Figures 2–4).

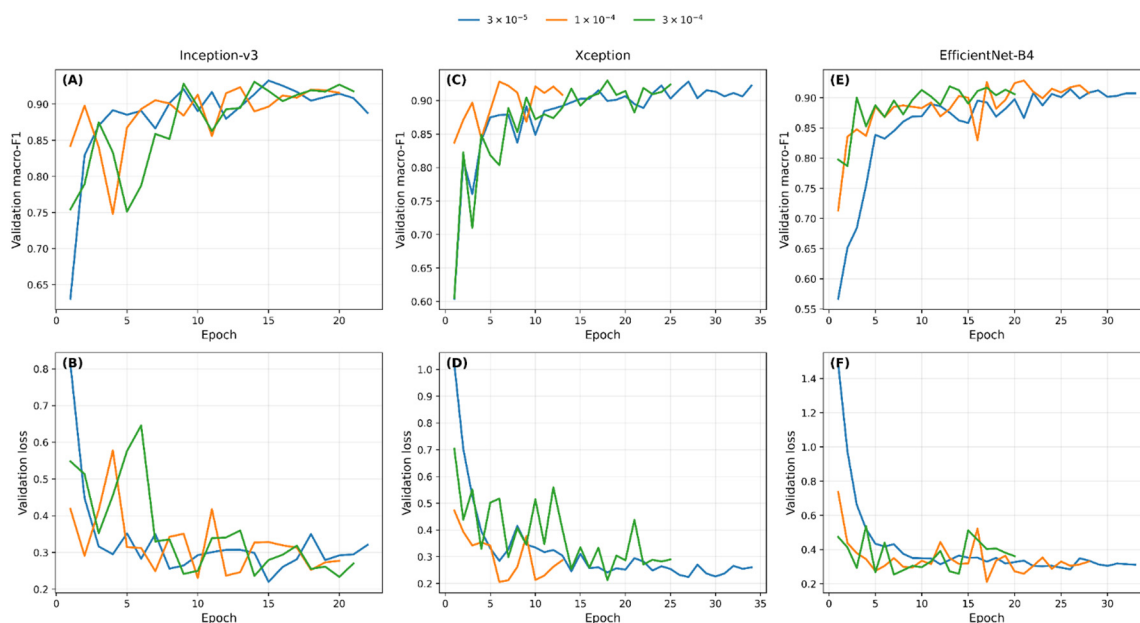


Figure 2. Validation macro F1 and validation loss during Stage 1 training for Inception v3, Xception, and EfficientNet B4 under three learning rate settings.

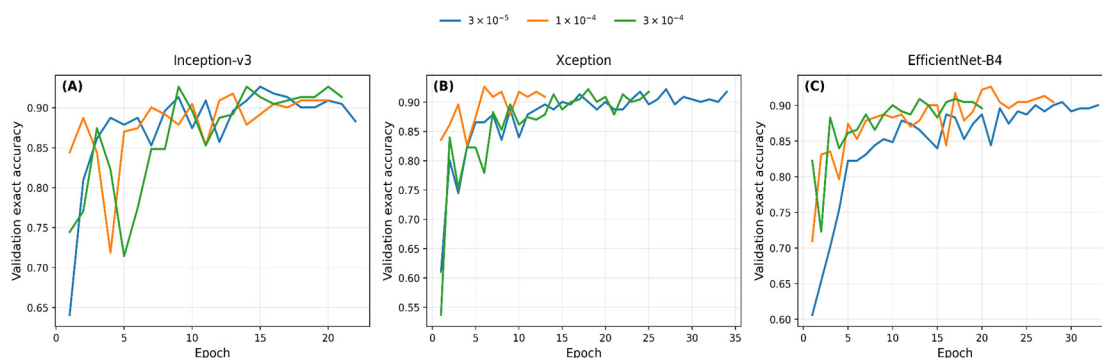


Figure 3. Validation exact accuracy trajectories of the three image only backbones under different learning rate settings in Stage 1.

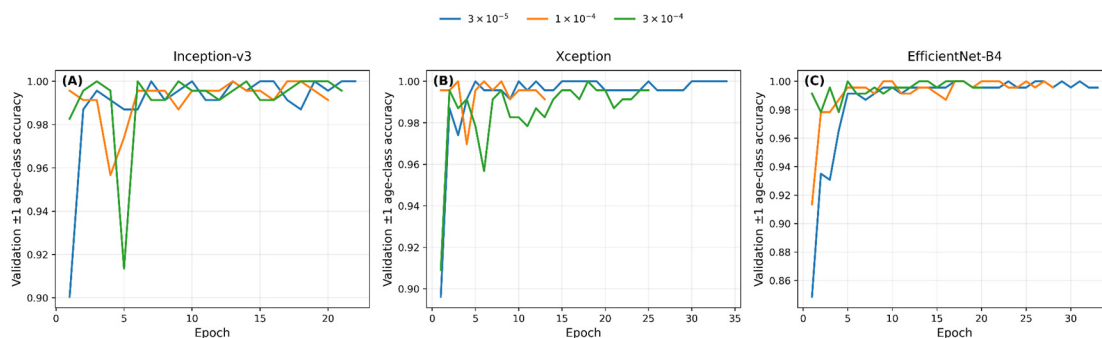


Figure 4. Validation agreement within one age class for the three image only backbones under different learning rate settings in Stage 1.

3.3. Stage 2 Optimization of the Selected Image Only Backbone

After the Stage 1 comparison, Inception v3 was further optimized by testing combinations of learning rate, weight decay, and dropout. The best validation macro F1 score was obtained with a learning rate of 1×10^{-4} , weight decay of 1×10^{-4} , and dropout of 0.3 (Table 4; Figure 5). This setting reached a validation macro F1 score of 0.944 at epoch 11, with validation exact accuracy of 0.935 and validation agreement within one age class of 1.000.

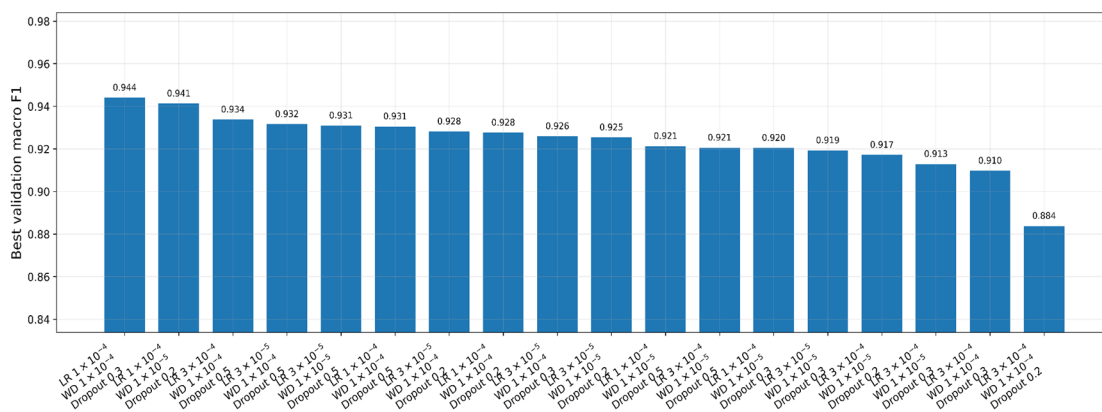


Figure 5. Comparison of the best validation macro F1 scores obtained from Stage 2 hyperparameter combinations for the selected Inception v3 image only model.

Table 4. Top five Stage 2 hyperparameter combinations for the selected Inception v3 image only model.

Ran k	Learnin g rate	Weigh t decay	Dropou t	Best epoc h	Validatio n macro F1	Validatio n loss	Validatio n exact accuracy	Validatio n ± 1 accuracy
1	1×10^{-4}	1×10^{-4}	0.3	11	0.944	0.181	0.935	1.000
2	1×10^{-4}	1×10^{-5}	0.2	13	0.941	0.158	0.935	1.000
3	3×10^{-4}	1×10^{-4}	0.5	15	0.934	0.279	0.931	0.996
4	3×10^{-5}	1×10^{-4}	0.5	26	0.932	0.274	0.926	0.996
5	3×10^{-5}	1×10^{-5}	0.5	15	0.931	0.273	0.926	1.000

The top five hyperparameter combinations all achieved validation macro F1 scores higher than 0.931. This shows that the optimized Inception v3 model was relatively robust across several regularization settings, although the best combination provided the most balanced validation performance.

3.4. Independent Test Performance of the Image Only Model

The optimized image only model was evaluated on the independent test set after model selection had been completed. The model achieved test exact accuracy of 0.866 and test macro F1 score of 0.873 (Table 5). Agreement within one year was 1.000, indicating that all test set errors were within one age class of the reader assigned age.

Table 5. Independent test performance of the optimized image only model.

Model	Test exact accuracy	Test macro F1	Test ± 1 year accuracy
Image only Inception v3	0.866	0.873	1.000

These results indicate that sagittal otolith images contained age related visual information that could be learned by CNNs. The high within one year agreement also suggests that the model captured the ordered nature of age classes rather than producing random class assignments.

4. Discussion

This study showed that reader assigned age classes of *T. japonicus* could be classified from sagittal otolith images using CNNs. The result supports the idea that otolith images contain visual signals related to age. These signals may include broad otolith shape, image texture, and internal growth patterns. However, the model output should be interpreted as classification of reader assigned age classes, not as independent validation of true age. This distinction is important because the supervised labels were derived from conventional otolith readings.

The Stage 1 comparison showed that all three backbones produced high validation performance. This means that the age related visual signal was not limited to a single CNN architecture. Among the tested models, Inception v3 showed the highest validation macro F1 score and was selected as the image only baseline. Inception based architectures use multiple convolutional filters and can capture visual patterns at different spatial scales [18]. This may be useful for otolith images because age related information can occur as local increment features as well as broader shape patterns. At the same time, the differences among the best backbones were moderate. The result should therefore be interpreted as the best architecture for the present dataset and training framework, rather than as general evidence that Inception v3 is always superior for otolith age classification.

EfficientNet B4 did not clearly outperform Inception v3 or Xception in this dataset, even though EfficientNet based models have performed well in some otolith age interpretation studies [14]. One possible reason is that EfficientNet B4 may require careful adjustment of image size, learning rate, and regularization to reach stable performance. In the present analysis, its best setting was competitive, but the other learning rate settings were lower. This suggests that architecture choice and optimization settings should be evaluated together.

The Stage 2 results showed that hyperparameter optimization improved the selected Inception v3 model. The best setting combined a learning rate of 1×10^{-4} , weight decay of 1×10^{-4} , and dropout of 0.3. This indicates that the final performance depended not only on the backbone but also on regularization and optimization. Weight decay helps limit excessive parameter growth, and AdamW applies weight decay separately from the adaptive gradient update [24]. Dropout also reduces co adaptation among features and can help limit overfitting [27]. These regularization methods were useful because the number of otolith images was small compared with general computer vision datasets.

The use of validation macro F1 as the primary selection criterion was appropriate because the age class distribution was not uniform. Overall accuracy can be dominated by the most common age classes. Macro F1 gives equal weight to each class and therefore provides a clearer view of class balanced prediction [25,26]. In the present dataset, age 2 was the most abundant class, whereas age 0

was less common. A metric that gives more equal weight to all classes was therefore important for model selection.

The independent test result showed lower exact performance than the best validation result, but the model still achieved a test macro F1 score of 0.873 and within one year agreement of 1.000. This pattern suggests that exact age class prediction remained difficult for some neighboring ages, while large age errors were not observed. Such a pattern is biologically reasonable because adjacent age classes can have similar otolith structures. In applied age reading, a prediction that differs by one year has a different meaning from an error of several years. For this reason, both exact accuracy and within one year agreement should be considered when evaluating otolith age classification models.

The present model may be useful as a support tool for age classification of *T. japonicus*. It could help screen large image datasets, provide a second opinion for readers, or flag samples that require careful rechecking. Similar support oriented use has been suggested for automated fish age estimation systems [16,17,28]. The goal should not be to remove expert readers from the process. Instead, CNN based tools can help improve efficiency and consistency when used together with conventional otolith reading.

Several limitations should be noted. First, the model was trained using reader assigned age labels. Any uncertainty in the original readings can be transferred to the model. Second, the model was developed using a single image dataset. Otolith image models can be sensitive to differences in lighting, magnification, camera settings, background, preparation method, and local reading protocol. Previous work has shown that models trained on one otolith image source may not transfer directly to another without adaptation [29]. Third, the sampled age range was limited to ages 0 to 4. Performance outside this age range remains unknown.

Future work should include external validation using otolith images from different sampling periods, imaging systems, or laboratories. Multiple reader datasets would also be useful for estimating label uncertainty. Model interpretation methods, such as class activation maps or saliency based approaches, could help determine whether the network focused on biologically meaningful otolith regions rather than background artifacts. Such interpretability will be important if CNN models are to be used in routine fisheries monitoring.

5. Conclusions

This study evaluated three CNN backbones for age class classification of *T. japonicus* using sagittal otolith images only. Inception v3, Xception, and EfficientNet B4 all showed high validation performance, indicating that otolith images contained useful age related visual information. Inception v3 gave the highest validation macro F1 score in the benchmark comparison and was further improved through hyperparameter optimization. The optimized image only model achieved moderate to high exact test performance and complete agreement within one year. These results suggest that otolith image based CNNs can support age class screening in *T. japonicus*. The approach should be treated as a complementary decision support tool rather than a replacement for expert otolith reading.

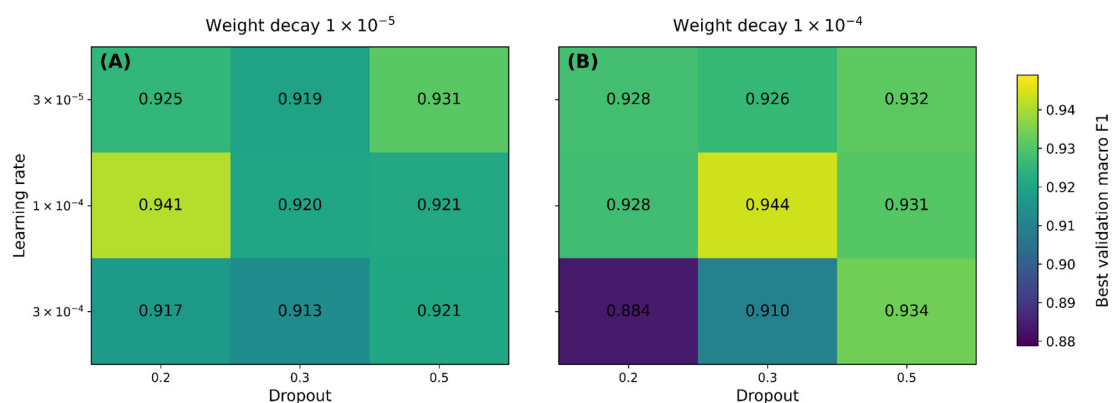


Figure 6. Heatmap of best validation macro F1 scores according to learning rate and dropout under two weight decay settings in Stage 2.

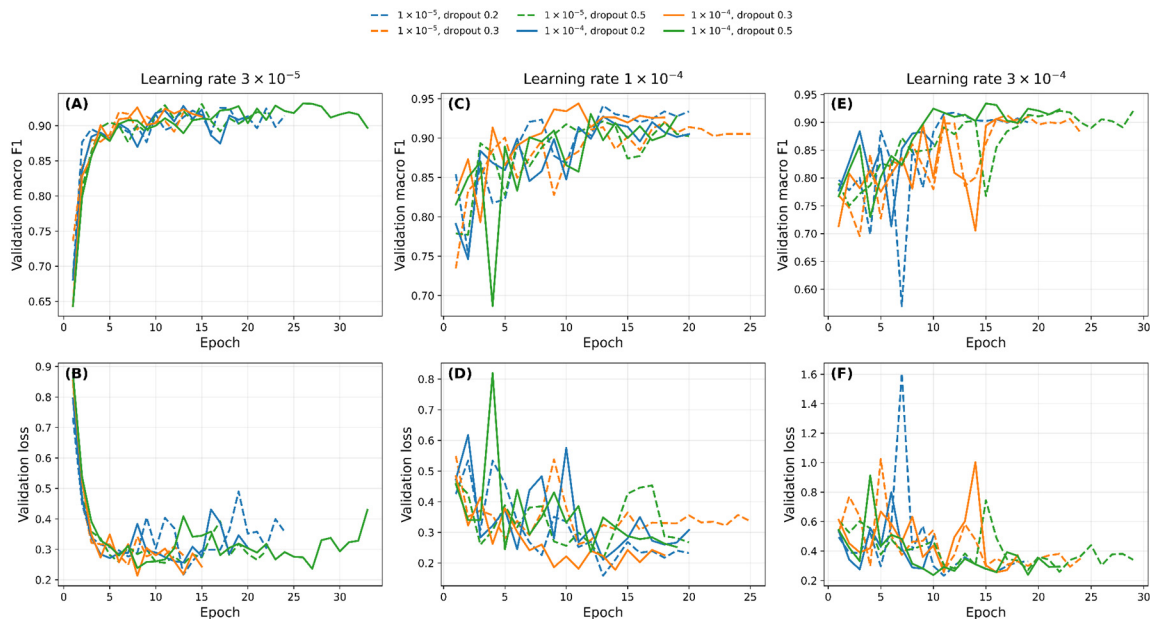


Figure 7. Validation macro F1 and validation loss curves during Stage 2 hyperparameter optimization of the selected Inception v3 image only model.

Author Contributions: Conceptualization, M.S.Y. and C.W.O.; methodology, M.S.Y.; software, M.S.Y.; validation, M.S.Y.; formal analysis, M.S.Y.; investigation, M.S.Y.; resources, C.W.O.; data curation, M.S.Y.; writing—original draft preparation, M.S.Y.; writing—review and editing, C.W.O.; visualization, M.S.Y.; supervision, C.W.O.; project administration, C.W.O. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by a Research Grant from Pukyong National University (2025).

Institutional Review Board Statement: Not applicable. The specimens used in this study were obtained from commercial fishery landings.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments: The authors thank the members of the laboratory who assisted with sample handling and otolith image preparation.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Beamish, R.J.; McFarlane, G.A. The forgotten requirement for age validation in fisheries biology. *Trans. Am. Fish. Soc.* 1983, 112, 735–743.
2. Campana, S.E. Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. *J. Fish Biol.* 2001, 59, 197–242. <https://doi.org/10.1111/j.1095-8649.2001.tb00127.x>.
3. Campana, S.E.; Moksness, E. Accuracy and precision of age and hatch date estimates from otolith microstructure examination. *ICES J. Mar. Sci.* 1991, 48, 303–316. <https://doi.org/10.1093/icesjms/48.3.303>.

4. Hüssy, K.; Radtke, K.; Plikshs, M.; Oeberst, R.; Baranova, T.; Krumme, U.; Sjöberg, R.; Walther, Y.; Mosegaard, H. Challenging ICES age estimation protocols: Lessons learned from the eastern Baltic cod stock. *ICES J. Mar. Sci.* 2016, *73*, 2138–2149. <https://doi.org/10.1093/icesjms/fsw107>.
5. Yamada, U.; Tokimura, M.; Horikawa, H.; Nakabo, T. *Fishes and Fisheries of the East China and Yellow Seas*; Tokai University Press: Kanagawa, Japan, 2007.
6. Lee, D.J.; Kang, S.; Jung, K.-M.; Cha, H.K. Age and growth of jack mackerel *Trachurus japonicus* off Jeju Island, Korea. *Korean J. Fish. Aquat. Sci.* 2016, *49*, 648–656.
7. Yoda, M.; Shiraiishi, T.; Yukami, R.; Ohshimo, S. Age and maturation of jack mackerel *Trachurus japonicus* in the East China Sea. *Fish. Sci.* 2014, *80*, 61–68. <https://doi.org/10.1007/s12562-013-0687-5>.
8. Yoda, M.; Tanaka, S.; Takahashi, M. Age, growth and reproductive cycle of the jack mackerel *Trachurus japonicus* in the Southwestern Sea of Japan. *Jpn. Agric. Res. Q.* 2023, *57*, 175–182.
9. Fablet, R.; Le Josse, N. Automated fish age estimation from otolith images using statistical learning. *Fish. Res.* 2005, *72*, 279–290. <https://doi.org/10.1016/j.fishres.2004.10.008>.
10. Morison, A.K.; Robertson, S.G.; Smith, D.C. An integrated system for production fish aging: Image analysis and quality assurance. *N. Am. J. Fish. Manag.* 1998, *18*, 587–598. [https://doi.org/10.1577/1548-8675\(1998\)018<0587:AISFPA>2.0.CO;2](https://doi.org/10.1577/1548-8675(1998)018<0587:AISFPA>2.0.CO;2).
11. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* 2015, *521*, 436–444.
12. LeCun, Y.; Bottou, L.; Bengio, Y.; Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* 1998, *86*, 2278–2324.
13. Moen, E.; Handegard, N.O.; Allken, V.; Albert, O.T.; Harbitz, A.; Malde, K. Automatic interpretation of otoliths using deep learning. *PLoS ONE* 2018, *13*, e0204713. <https://doi.org/10.1371/journal.pone.0204713>.
14. Martinsen, I.; Harbitz, A.; Bianchi, F.M. Age prediction by deep learning applied to Greenland halibut (*Reinhardtius hippoglossoides*) otolith images. *PLoS ONE* 2022, *17*, e0277244. <https://doi.org/10.1371/journal.pone.0277244>.
15. Moen, E.; Vabø, R.; Smoliński, S.; Denechaud, C.; Handegard, N.O.; Malde, K. Age interpretation of cod otoliths using deep learning. *Ecol. Inform.* 2023, *78*, 102325. <https://doi.org/10.1016/j.ecoinf.2023.102325>.
16. Politikos, D.V.; Petasis, G.; Chatzisprou, A.; Mytilineou, C.; Anastasopoulou, A. Automating fish age estimation combining otolith images and deep learning: The role of multitask learning. *Fish. Res.* 2021, *242*, 106033.
17. Politikos, D.V.; Sykiniotis, N.; Petasis, G.; Dedousis, P.; Ordoñez, A.; Vabø, R.; Anastasopoulou, A.; Moen, E.; Mytilineou, C.; Salberg, A.B.; Chatzisprou, A.; Malde, K. DeepOtolith v1.0: An open-source AI platform for automating fish age reading from otolith or scale images. *Fishes* 2022, *7*, 121. <https://doi.org/10.3390/fishes7030121>.
18. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016*; pp. 2818–2826.
19. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017*; pp. 1251–1258.
20. Tan, M.; Le, Q. EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019*; pp. 6105–6114.
21. Yosinski, J.; Clune, J.; Bengio, Y.; Lipson, H. How transferable are features in deep neural networks? In *Advances in Neural Information Processing Systems; 2014; Volume 27*.
22. Tajbakhsh, N.; Shin, J.Y.; Gurudu, S.R.; Hurst, R.T.; Kendall, C.B.; Gotway, M.B.; Liang, J. Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Trans. Med. Imaging* 2016, *35*, 1299–1312.
23. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* 2014, arXiv:1412.6980.
24. Loshchilov, I.; Hutter, F. Decoupled weight decay regularization. *arXiv* 2017, arXiv:1711.05101.
25. Sokolova, M.; Lapalme, G. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* 2009, *45*, 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>.

26. Tharwat, A. Classification assessment methods. *Appl. Comput. Inform.* 2021, 17, 168–192.
27. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 2014, 15, 1929–1958.
28. Bojesen, T.A.; Denechaud, C.; Malde, K. Annotating otoliths with a deep generative model. *ICES J. Mar. Sci.* 2024, 81, 55–65. <https://doi.org/10.1093/icesjms/fsad170>.
29. Ordoñez, A.; Eikvil, L.; Salberg, A.B.; Harbitz, A.; Elvarsson, B.P. Automatic fish age determination across different otolith image labs using domain adaptation. *Fishes* 2022, 7, 71. <https://doi.org/10.3390/fishes7020071>.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.