Using artificial intelligence (AI) for supplementing Pacific halibut age determination from collected otoliths

PREPARED BY: IPHC SECRETARIAT (B. HUTNICZAK, J. FORSBERG, K. SAWYER VAN VLECK, & K. MAGRANE; 21 MAY 2024)

PURPOSE

In the continuously evolving realm of marine science, the IPHC recognizes the transformative potential of artificial intelligence (AI). This document summarizes the information available on the use of AI for determining the age of fish from images of collected otoliths and provides an update on the exploratory work of implementing an AI-based age determination model for Pacific halibut.

The purpose of this document is twofold. First, to provide a background in support of developing a protocol for creating a database of pictures with expert-provided labels for ageing use. Second, to propose a process for developing the necessary AI services for supplementing current Pacific halibut ageing protocol.

BACKGROUND

Otoliths are crystalline calcium carbonate structures, mostly in the form of aragonite, found in the inner ear of fish. They contain growth rings, that are often compared to tree growth rings. By analyzing the growth patterns in otoliths, scientists estimate the age of fish (Campana, 1999; Campana & Neilson, 1985), supporting the estimation of fish population demographics and population dynamics (Campana & Thorrold, 2001). In turn, fish age is a key input to stock assessment models that inform management decisions related to fish exploitation (Methot & Wetzel, 2013). It is estimated that the number of otoliths from captured fish that are read annually worldwide is on the order of one million (Campana & Thorrold, 2001).

The current method for determining ages of most fish species relies on manually extracting, preparing (embedding, sectioning), and reading otoliths. The simplest approach to reading the otolith is to immerse it in a clear liquid, such as water or alcohol solution, illuminate it from above, and view it against a dark background, using a stereo microscope. This method is suitable only for otoliths that are relatively thin with all annual bands visible from the surface. For species such as Pacific halibut, as the growth rate of the fish slows down, the outer growth bands become increasingly compressed and difficult to read from the surface of the whole otolith. To correctly determine the number of annual bands in such cases, otoliths are typically viewed in cross section which allows viewing the bands that are not visible from the surface view. In addition, the contrast between the growth rings can be enhanced through the baking process. Pacific halibut otoliths are aged using the 'break and bake' technique.

This manual ageing process is expensive, time-consuming,[1] and can be subject to bias[2] as well as imprecision due to variations in age estimations between readers and within readers over time. Recent advances in imaging technologies and machine learning suggest that AI can assist in this process by automating the analysis of otolith images[3] and identifying and measuring the growth rings to determine age. AI algorithms can be trained on a large dataset of otolith images with known ages to learn the patterns and variations in growth rings. Once trained, the AI model can analyze new otolith images and predict the age of the fish based on the identified patterns in the image.

Using AI for age determination of Pacific halibut could improve consistency and replicability of age estimates, as well as provide time and cost savings to the organization, providing age data for reliable management advice. However, it's important to note that the AI model's accuracy depends on the quality and diversity of the training data, as well as the expertise of the scientists involved in training and validating the model. Regular validation and calibration with manual age determinations is necessary to ensure the accuracy and reliability of the AI predictions. Thus, the proposed approach integrates AI-based age determination and traditional ageing methods for maximum accuracy of the estimates.

Model

The model framework (Figure 1) includes a continuous process of training the model using available labelled data (aged otoliths), querying the model to select the next sample, labeling or relabeling the selected sample, and enriching the model with newly labelled samples.

This model relies on automatized ageing that is supplementing the expert-derived age estimates continuously improving the model in the *Label* phase and the *Enrich* phase.

[1] While the actual reading may account only for a fraction of the total cost and time required to process the otolith from collection to age determination, skilled readers require years of training, which should be considered when conducting a cost-benefit analysis.

[2] While the count of annual rings on Pacific halibut otoliths was found to provide unbiased age estimate using validation against bomb radiocarbon isotopes (Piner & Wischniowski, 2004), an earlier oxytetracycline (OTC) mark-recapture study indicated biases among age readers (Blood, 2003). In the 1980s, the IPHC applied injections with the antibiotic oxytetracycline (OTC) during routine tagging operations to evaluate validity of ageing method (IPHC, 1985). Upon injection, the OTC is absorbed by the fish's bony structure, including the otoliths, and leaves a mark that is easily seen when viewed under an ultraviolet light. When an OTC-injected tagged fish is recovered, the otoliths are removed and examined under the ultraviolet light. By comparing the number of annuli laid since the OTC mark to the fish recovery, the accuracy of the age readings can be determined.

[3] Although the idea of taking pictures of Pacific halibut otoliths is not new. See 1960 report by G. Morris Southward, *Photographing Halibut Otoliths for Measuring Growth Zones* (Southward, 1962).

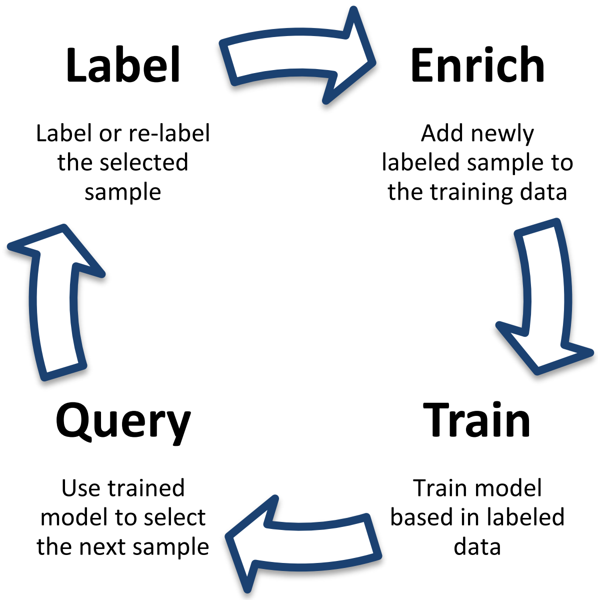

Figure 1: Model framework.

Modeling approach

Previous literature (see perspective piece by Malde et al., 2020) suggests adapting a pre-trained convolutional neural network (CNN) designed for image classification to estimate age using otolith images obtained via microscope camera. This type of model is trained on a large collection of images of otoliths previously aged by human readers. Moen et al. (2018) presents the first case of the use of deep learning and CNN to estimate age from images of whole otoliths of Greenland halibut (*Reinhardtius hippoglossoides*).[4]

Artificial neural networks (ANNs) are computational structures inspired by biological neural networks. They consist of simple computational units referred to as neurons, organized in layers. The neuron parameters (or weights) are estimated by training the model using supervised learning. This process consists of two steps: forward propagation, where the network makes a prediction based on the input; and back propagation, where the network learns from its mistake by calculating the gradient of a loss function, and then uses the gradient to update the neuron weights. The ANNs approach has been used for fish ageing by Robertson & Morison (1999) and Fablet & Le Josse (2005) with a limited success.

The neural networks approach significantly improved in recent years with the increase in the number of layers, applying an approach often referred to as deep learning. Deep learning neural networks are known for their generality. With sufficient training data, they can be used to classify raw data (e.g., an array of pixels) directly, without explicit design of low-level features. The deep learning algorithm lower layers learn to distinguish between primitive features automatically, typically identifying sharp edges or color transitions. Subsequent layers then learn to recognize more abstract features as combinations of lower layer features, and finally merge this information to provide a high-level classification.

[4] CNN was also applied for other tasks related to fisheries management, e.g. fish species identification (Allken et al., 2019).

In CNNs (LeCun et al., 1998; Simonyan & Zisserman, 2015), the layers are structured as stacks of filters, each recognizing increasingly abstract features in the data. Convolutional layers may be understood as an efficient way to transform an input image into another image, highlighting meaningful patterns, learned from data during training. The training is sequential, meaning the output of each layer is the input of the next layer, and the useful features are learned in the various layers during training. This approach is very effective for many image analysis problems, where objects are often recognized independent of their location. During network training, the performance is monitored over sequential epochs. Epochs represent the number of times that the training dataset is passed forward and backward through the network to refine model weights. Whenever the validation loss decreases, the trained model is saved, ending up with the network that corresponds to the minimum loss and highest accuracy on the validation set. The trained network is then evaluated on the testing set.

In the CNN model, prediction of age can be defined as a classification task (age as a class category) or image regression, that is a task of predicting a continuous variable from an image, in this case prediction of age as a numeric value from an otolith image. Both approaches can be tested for devising a method better suited for Pacific halibut. Considering fish age as a discrete parameter is a common approach used to identify the individual year class, i.e. grouping fish originating from the spawning activity in a given year (Moen et al., 2018), although this may be less appropriate for long-living species with a larger number of age categories in the sample. The oldest Pacific halibut on record were aged at 55 years (Keith et al., 2014).

Software options

The proposed approach follows that of (Moen et al., 2018; Moore et al., 2019) who chose TensorFlow and Keras libraries to implement and train the model. TensorFlow is currently the largest and most popular library available for deep learning. Keras is a high-level API which runs on top of TensorFlow and simplifies implementation of TensorFlow models.

The approach uses a transfer-learning technique to develop a CNN for otolith age estimation. Transfer learning is the process of repurposing a machine learning model that has been pre-trained for another, related, task. Specifically, it starts with the Inception v3 model from Google, pre-trained on the ImageNet database. ImageNet database contains over 14 million (14,197,122) annotated images classified intro 1000 categories. The CNN layers are loaded with pre-trained (with ImageNet data) and publicly available weights, as opposed to using random initialization. Various training meta-parameters contribute substantially to final accuracy by using a stochastic gradient descent (SGD) optimizer and by leaving all network layers as trainable.

For the application to otolith ageing for Pacific halibut, the input layer was scaled to match the images' resolution.[5] The output layer was changed from a multi-dimensional output vector representing class probabilities to a single numeric output, effectively transforming it to a new regression layer.[6] This design follows the following pattern: Input → InceptionV3 (feature

[5] Resolution is the total number of pixels along an image's width and height, expressed as pixels per inch (PPI). The Inception v3 model processes images that are 299 x 299 pixels in size. The original images, which were 2548 x 2548 pixels, were resized to 400 x 400 pixels.

[6] Alternatively, Politikos et al. (2021) replaced the last layer with a feed-forward network with two hidden layers replacing the default 1000-categories output layer with a fully-connected layer with six hidden nodes, corresponding to a limited number of age categories [Age-0 – Age-5+], with the last one representing fish of age 5 and older, In this case, the network outputs probabilities using the softmax function, a function that performs multi-class classification and transforms the outputs to represent the probability distributions over a list of potential outcomes. The IPHC uses in its stock assessment bins Age-2 – Age 25+ for the current age data and Age-2 - Age-20+ for the

extractor) → Classifier/Regressor → Output. At this point, the neural network is trained to minimize the mean squared error (MSE) between predicted ages and human expert age estimates,[7] using the otolith images as inputs.

A similar approach, although adopting classification approach, was applied for ageing Greek Red Mullet (*Mullus barbatus*) (Politikos et al., 2022) and the associated code is available on GitHub (github.com/dimpolitik/DeepOtolith). The available open-source code was adapted for testing the approach for Pacific halibut.

Use of auxiliary data

Precision of age predictions of otoliths using neural networks from geometric features could be potentially improved by using auxiliary data, for example, fish size or date and location of capture (Moen et al., 2018). Past IPHC work suggests a good deal of spatial variation in Pacific halibut growth ring patterns. This points to the importance of good spatial coverage in the training sample. Additionally, the project plans to explore the use of additional spatial covariates for better age prediction. Other available auxiliary data include year collected, which could be applied to account for variation between cohorts and prevalent environmental conditions throughout the aged fish life histories, and the collection dates, which provides insights into seasonal variation to the interpretation of the otolith edge.

Performance metrics

Performance of the CNN to correctly assign ages (rounded output of the regression layer) to otolith images in the test set is assessed via the root mean squared error (RMSE). Moen et al., (2018) also suggest calculating coefficient of variation (CV).[8]

For the production set, accuracy could be further refined using a mixed-method approach. A minimum number of otoliths (e.g., 10%) could be reexamined by human readers after the selection based on the model-derived confidence intervals, targeting samples where the confidence is low. The final bias relevant to products such as stock assessment could integrate the predicted age estimates derived following the re-label phase. In practice, mixed-method approach would eliminate the need for human experts to read ‘easy’ otoliths, while maintaining human-based decision control over more ‘difficult’ otoliths.[9]

Achieved accuracy

Moen et al., (2018), for Greenland halibut, achieved MSE for the left and right otoliths and pair of 3.27, 2.71 and 2.99, respectively. Age was correctly estimated for 48 out of the 164 tested otolith-pairs (29%). In addition, 63 cases (38%) were estimated to be one year off the read age. There was also a clear tendency for the system to predict a lower age for older individuals, when

historical surface read ages. The adoption of a larger number of age categories prompted the decision to incorporate a regression layer in place of class probabilities.

[7] In practice, the neural network minimizes the MSE of normalized age values, i.e., age values divided by the maximum age provided as input.

[8] The CV of the predicted age at true age is the primary input to the IPHC stock assessment. It is generally modelled as a parametric function of age accounting for the complex joint probability that both estimates can be incorrect (Punt et al., 2008).

[9] If there is a strong junction in the relative precision between old and younger fish due to the change in methods this may require a nonparametric approach to ageing imprecision. If an AI method is biased as a function of age (standard for surface reading methods) and the break and bake method is unbiased, integrating the methods may prove challenging.

compared to human readers. The variance of the predictions also increased with the age of the otolith.

The model developed by Moore et al. (2019), for prediction of age of snapper using CT scans,[10] gave the same age as the human reader for 47% of otoliths in a test dataset, with a further 35% of ages estimated within 1 year of the human reader estimate of age (n=687). For hoki, the model gave the same age as the human reader for 41% of individuals (n=882).

The age model for Greenland halibut by Politikos et al., (2022) gave RMSE of 1.69 years between age prediction and age reading by experts (n=8218, 26 age categories). For Greek red mullet, correct age was predicted for 69.2% individuals, with an additional 28.2% being within 1 year of error (n=5027).

Benson et al., (2023), using near-infrared spectroscopy of otoliths, supplemented by geospatial and biological data routinely collected on the survey, estimated age of walleye pollock. For the optimal multimodal CNN model, an RMSE of 0.83 for the training set and an RMSE of 0.91 for the test set indicated that at least 67% of estimated ages were predicted within ±1 year of age compared to traditional microscope-based ages.

However, it should be noted that neither the traditional ageing methods for Pacific halibut are perfectly accurate. Within- and between-reader agreement in age assignment is generally 60%-70% complete agreement, 80% to 90% within one year, and 100% within 3 years. The IPHC Secretariat's publications report on % agreement (see *Technical Report No. 46* and *No. 47*).

Database

The IPHC annually ages a considerable number of otoliths (see *Appendix B* for details). Since 1925, over 1.5 million otoliths have been aged and stored for potential future use. Otoliths collected by the IPHC for ageing purposes undergo additional processing. Otoliths are sectioned (broken in half) and baked to enhance the contrast between the growth rings. These stored and previously aged otoliths serve as a valuable resource for creating a database of images for training purposes. To optimize model training, the selection of otoliths included in the model covers a broad spectrum of fish sizes, ages, sexes, and collection locations.

Before photographing, processed otoliths were placed in a monochrome tray featuring an elongated groove designed to keep the otolith upright and immersed in water. The pictures were taken with AmScope 8.5MP eyepiece cameras,[11] under consistent lighting conditions and magnification. The input database includes images of standardized size, 2548 by 2548 pixels, which are later resized to the desired resolution based on the model's specification.[12]

---

[10] CT scanning uses X-ray technology to produce image slices through objects, which can be reconstructed into virtual, three-dimensional (3D) images that can be rotated and viewed in any orientation (Moore et al., 2019). Such images may provide more accurate estimates, but the cost of this approach is prohibitive at (based on trial conducted in New Zealand) $1,500 per day, with scan timed for an individual otolith between 40 min to one hour. However, as the technology progresses, this approach may provide an option for fully automating the entire ageing process by scanning a whole fish (e.g., along a conveyor belt). Deep learning methods (i.e., CNN) developed for age determination from surface images could serve as a base for age determination from CT scans.

[11] The camera fits in one of the microscope eyepieces, eliminating the need to purchase a separate camera mount for the microscope.

[12] Moen et al. (2018) used images 400 by 400 pixels, which required the input layer to be scaled to match the images size as Inception v3 classifies by default images with a size of 299 by 299 pixels. Ordoñez et al. (2020), using the same set of images, built a CNN with images resized to 224 by 224 pixels, the default input of the VGG-

It is important to note that <u>it may not be necessary to image the otoliths at resolutions sufficient for human viewers to resolve, because the CNN may be able to arrive at an age estimate without directly counting bands</u> (Moore et al., 2019).

Figure 2 shows an example of a range of images used in the CNN training dataset.

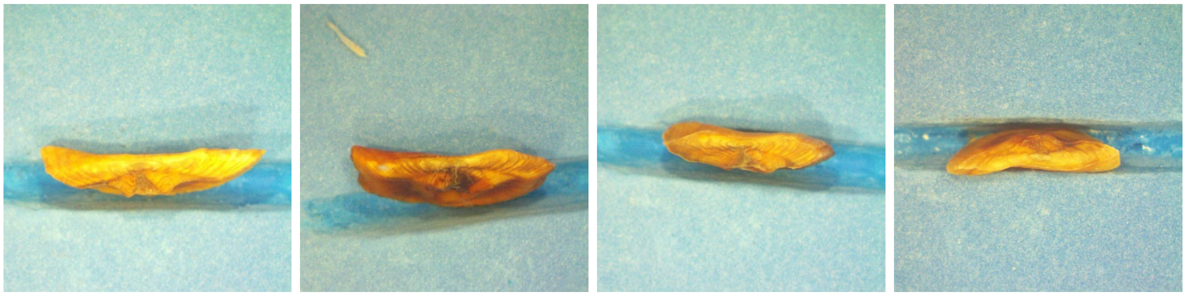

Figure 2: Examples of Pacific halibut otolith images taken for inclusion in the training set.

Note: In due course, the IPHC will create a database comprising labelled images of otoliths both pre- and post-processing and conduct a cost-benefit analysis of processing the otoliths for ageing using AI. The analysis will look at the accuracy improvement when using an image database containing images of processed (broken and baked) otoliths with enhanced contrast vs. those captured prior to processing (i.e. whole otoliths). In their research, Politikos et al. (2022) utilized digital images of otoliths that were not subject to any additional processing in the laboratory, immersed in water and placed under a stereomicroscope on a white background with transmitted light. However, it is important to note that even if results indicate that breaking and baking is not necessary for age determination using AI, a subsample chosen for the Label and Enrich phases would have to be fully processed for age determination with traditional methods by an expert reader.

Presorting otoliths

The adopted procedure excludes broken otoliths, applying manual presorting at the image-taking stage. Presorting has also occurred at the collection stage when crystalized otoliths[13] are omitted when collecting samples.

Ongoing research [Dimitris Politikos, personal communication] is investigating the initial stage of the aging process, specifically assessing whether an otolith is of sufficient quality for age determination. This research is pertinent for cases involving crystallized or broken otoliths and aims to potentially eliminate the need for subjective decisions by samplers regarding the usability of otoliths for age determination. This approach implements a two-stage classification system. In the first stage, the model assesses the otolith's suitability for ageing; in the second, it determines the age. Th algorithm-driven presorting could also incorporate expert knowledge for handling problematic otoliths.

19 model. Higher resolution images offer the flexibility to adapt the model in the future to more detailed and complex image analysis tasks, potentially improving the accuracy and effectiveness of image recognition capabilities.

[13] Crystalized otoliths have an altered composition – specifically, where the aragonite in the otolith is partially or mostly replaced by vaterite, a phenomenon known as otolith crystallization. Crystallized otoliths are not suitable for ageing.

In developing the model, the training dataset can be strategically supplemented with images of samples that represent a group of otoliths with which the original model struggles the most (Query phase).[14]

Image collection

The image collection is associated with labels storing:

1. Otolith reference number – using referencing system already in place;
2. Image name and location – exact path for image access;
3. Resolved age – human reader derived age (**rsvage**);
4. Year collected – to account for variation between cohorts and prevalent environmental conditions;
5. Date collected – to account for the 'edge effect' reflecting seasonal changes;
6. Geospatial characteristics (latitude and longitude) – to capture regional variation;
7. Resolved sex – to determine whether otolith characteristics (possibly not directly visible to human eye) could be used for sex determination.[15]

PRELIMINARY RESULTS

The initial model run utilized 2,286 images of otoliths collected during the 2019 IPHC fishery-independent setline survey (FISS). The 2019 FISS offers a valuable starting point for image database creation, being the most recent extensive survey expected to have captured the regional differences in otoliths, providing a robust dataset for initial modeling efforts.

The images were divided into training, validation, and test datasets. The training set (1,360) was used for training purposes. The validation set (240) was used to evaluate the model during the training process, allowing for adjustments without using the test set, which was reserved for the final evaluation. The test dataset (30%, 686) was used to assess the performance of the model after training, providing an unbiased evaluation of its generalization capability to new, unseen data. All images were resized to 400x400 pixels. Images of broken otoliths were excluded. The number of epochs was set to 1000, with EarlyStopping applied and patience set to 100. Learning rate was set to 0.0002 and batch size to 16.

Normalized age MSE in training set was 0.00032 and 0.00166 in validation set. The model was trained for 324 epochs (i.e., 224 effective epochs with patience=100). The model achieved RMSE in the test set of 1.82, and 1.84 when applied to rounded results. Correct age was predicted for 30.0% individuals, with an additional 38.5% being within 1 year of error. Figure 3 shows accuracy adjustment over the training process, while Figure 4 compares manually-derived age with AI predicted age. Figure 5 compares age composition derived manually with model predictions.

[14] About 1% of otoliths are partly crystallized and are assigned ages. The same is true for broken otoliths that are aged (1%)
[15] IPHC is currently using genotyping for Pacific halibut sex determination.

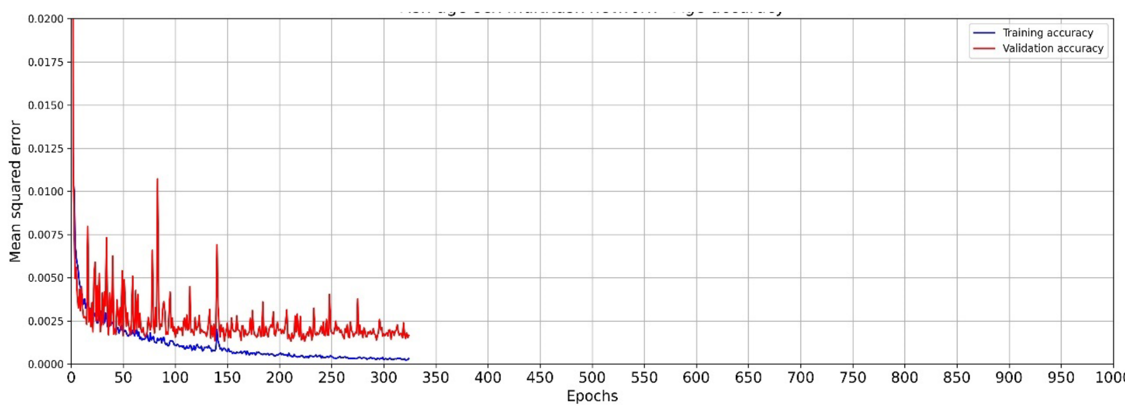

Figure 3: Age accuracy (measured as normalized age MSE) throughout the training process.

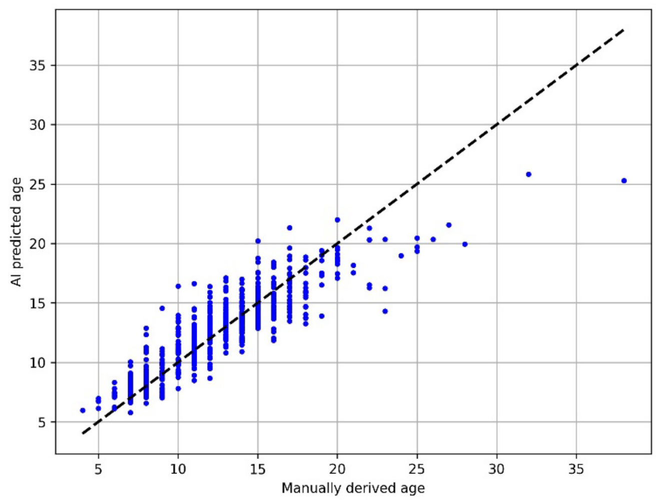

Figure 4: Comparison between manually derived age with AI predicted age.

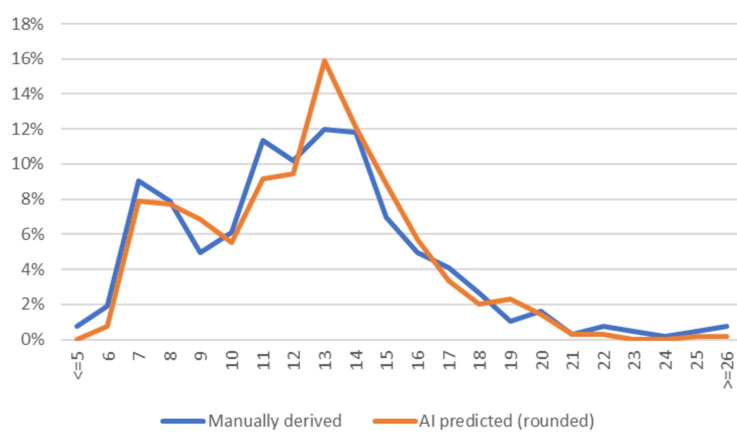

Figure 5: Comparison between manually derived age with AI predicted age – age composition.

CONCLUSIONS

In conclusion, the ongoing advancement of AI technologies in the field of marine science offers considerable potential to enhance the efficiency of age determination of Pacific halibut using otolith images. Preliminary results presented here suggest that AI could serve as a promising alternative to the current ageing protocol, which relies entirely on manual age reading. AI is also evolving rapidly, and adapting to new developments may further improve results over time However, it is important to continue verifying whether achieved accuracy of CNN-based predictions do not learn biased prediction rules based on changes in the relationship between age and covariates used by the model, noise or other irrelevant imaging artefacts present in the data (Ordoñez et al., 2020). Therefore, it is key to continuously diagnose performance problems and find ways to fix them (Belcher et al., 2023; Norouzzadeh et al., 2018). Moreover, the automated ageing process will still depend on trained readers for training the model with inputs that capture temporal changes, which is increasingly important in the face of changing environmental conditions and climate change.

LITERATURE

Allken, V., Handegard, N. O., Rosen, S., Schreyeck, T., Mahiout, T., & Malde, K. (2019). Fish species identification using a convolutional neural network trained on synthetic data. *ICES Journal of Marine Science*, *76*(1), 342–349. https://doi.org/10.1093/icesjms/fsy147

Belcher, B. T., Bower, E. H., Burford, B., Celis, M. R., Fahimipour, A. K., Guevara, I. L., Katija, K., Khokhar, Z., Manjunath, A., Nelson, S., Olivetti, S., Orenstein, E., Saleh, M. H., Vaca, B., Valladares, S., Hein, S. A., & Hein, A. M. (2023). Demystifying image-based machine learning: a practical guide to automated analysis of field imagery using modern machine learning tools. *Frontiers in Marine Science*, *10*(June), 1–24. https://doi.org/10.3389/fmars.2023.1157370

Benson, I. M., Helser, T. E., Marchetti, G., & Barnett, B. K. (2023). The future of fish age estimation : deep machine learning coupled with Fourier transform near-infrared spectroscopy of otoliths. *Canadian Journal of Fisheries and Aquatic Sciences*, *00*, 1–13. https://doi.org/dx.doi.org/10.1139/cjfas-2023-0045

Blood, C. L. (2003). I . Age validation of Pacific halibut II . Comparison of surface and break-and-burn otolith methods of ageing Pacific halibut. *IPHC Technical Report*, *47*.

Campana, S. E. (1999). Chemistry and composition of fish otoliths: Pathways, mechanisms and applications. *Marine Ecology Progress Series*, *188*, 263–297. https://doi.org/10.3354/meps188263

Campana, S. E., & Neilson, J. D. (1985). Microstructure of Fish Otoliths. *Canadian Journal of Fisheries and Aquatic Sciences*, *42*(5), 1014–1032. https://doi.org/10.1139/f85-127

Campana, S. E., & Thorrold, S. R. (2001). Otoliths, increments, and elements: keys to a comprehensive understanding of fish populations? *Canadian Journal of Fisheries and Aquatic Sciences*, *58*(1), 30–38. https://doi.org/10.1139/f00-177

Fablet, R., & Le Josse, N. (2005). Automated fish age estimation from otolith images using statistical learning. *Fisheries Research*, *72*(2–3), 279–290. https://doi.org/10.1016/j.fishres.2004.10.008

IPHC. (1985). Annual Report 1984. In *IPHC Annual Report*.

Keith, S., Kong, T., Sadorus, L. L., Stewart, I. J., & Williams, G. (2014). The Pacific Halibut:

Biology, Fishery, and Management. *IPHC Technical Report*, *59*. https://doi.org/10.1042/bj0490062

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient Based Learning Applied to Document Recognition. *Proc. of the IEEE*.

Malde, K., Handegard, N. O., Eikvil, L., & Salberg, A. B. (2020). Machine intelligence and the data-driven future of marine science. *ICES Journal of Marine Science*, *77*(4), 1274–1285. https://doi.org/10.1093/icesjms/fsz057

Methot, R. D., & Wetzel, C. R. (2013). Stock synthesis: A biological and statistical framework for fish stock assessment and fishery management. *Fisheries Research*, *142*, 86–99. https://doi.org/https://doi.org/10.1016/j.fishres.2012.10.012

Moen, E., Handegard, N. O., Allken, V., Albert, O. T., Harbitz, A., & Malde, K. (2018). Automatic interpretation of otoliths using deep learning. *PLoS ONE*, *13*(12), e0204713.

Moore, B. R., Maclaren, J., Peat, C., Anjomrouz, M., Horn, P. L., & Hoyle, S. (2019). Feasibility of automating otolith ageing using CT scanning and machine learning. *New Zealand Fisheries Assessment Report*, *58*.

Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., & Clune, J. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(25), E5716–E5725. https://doi.org/10.1073/pnas.1719367115

Ordoñez, A., Eikvil, L., Salberg, A. B., Harbitz, A., Murray, S. M., & Kampffmeyer, M. C. (2020). Explaining decisions of deep neural networks used for fish age prediction. *PLoS ONE*, *15*(6), 1–19. https://doi.org/10.1371/journal.pone.0235013

Piner, K. R., & Wischniowski, S. G. (2004). Pacific halibut chronology of bomb radiocarbon in otoliths from 1944 to 1981 and a validation of ageing methods. *Journal of Fish Biology*, *64*(4), 1060–1071. https://doi.org/10.1111/j.1095-8649.2004.0371.x

Politikos, D. V, Petasis, G., Chatzispyrou, A., Mytilineou, C., & Anastasopoulou, A. (2021). Automating fish age estimation combining otolith images and deep learning: The role of multitask learning. *Fisheries Research*, *242*, 106033. https://doi.org/https://doi.org/10.1016/j.fishres.2021.106033

Politikos, D. V, Sykiniotis, N., Petasis, G., Dedousis, P., Ordoñez, A., Vabø, R., Anastasopoulou, A., Moen, E., Mytilineou, C., Salberg, A. B., 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), 1–11. https://doi.org/10.3390/fishes7030121

Punt, A. E., Smith, D. C., KrusicGolub, K., & Robertson, S. (2008). Quantifying age-reading error for use in fisheries stock assessments, with application to species in Australia's southern and eastern scalefish and shark fishery. *Canadian Journal of Fisheries and Aquatic Sciences*, *65*(9), 1991–2005. https://doi.org/10.1139/F08-111

Robertson, S. G., & Morison, A. K. (1999). A trial of artificial neural networks for automatically estimating the age of fish. *Marine and Freshwater Research*, *50*(1), 73–82.

https://doi.org/10.1071/MF98039

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *ICLR 2015 - Conference Track Proceedings*.

Southward, G. M. (1962). Photographing Halibut Otoliths for Measuring Growth Zones. *Journal of the Fisheries Research Board of Canada*, *19*(2), 335–338. https://doi.org/10.1139/f62-018

APPENDIX A

Counts of otoliths aged by the IPHC.

Collection year	Ageing method	IPHC FISS*	Commercial (Market Sample)*	NOAA Trawl survey*	Tag recovery*	ADF&G recreational*	Clean collection
pre-1960	surface	70,984			10,068		
1960	surface	6,606			681		
1961	surface	4,727		4,576	842		
1962	surface	2,605		1,692	594		
1963	surface	8,257		2,209	440		
1964	surface	10,295	27,828	1,001	353		
1965	surface	5,169	27,252	1,186	493		
1966	surface	3,750	24,638	1,777	796		
1967	surface	6,325	29,797	2,271	1,151		
1968	surface	2,314	29,772	1,887	1,813		
1969	surface	1,510	23,361	1,019	1,869		
1970	surface	1,138	24,686	1,184	867		
1971	surface	2,702	16,374	2,294	732		
1972	surface	2,597	23,381	1,180	490		
1973	surface	1,747	16,683	893	244		
1974	surface	1,021	11,569	1,189	128		
1975	surface	1,212	14,128	1,136	131		
1976	surface	1,843	14,103	969	72		
1977	surface	1,853	13,514	1,102	83		
1978	surface	1,933	11,434	1,309	61		
1979	surface	2,021	7,219	730	93		
1980	surface	5,022	10,317	717	168		
1981	surface	7,942	8,267	460	129		
1982	surface	5,720	9,644	443	208		
1983	surface	5,822	9,262	1,355	286		
1984	surface	6,508	10,233	1,089	455		
1985	surface	5,872	12,986	1,192	778		
1986	surface	5,139	12,426	1,120	1,020		
1987	surface	42	16,137		859		
1988	surface	1,179	17,154	98	761		
1989	surface	6,130	14,122		710		
1990	surface	2,201	14,800	4,802	397		
1991	surface	1,315	13,461	2,598	280		
1992	surface/BB	7,530	14,564	222	182		
1993	surface/BB	3,384	13,747		147		
1994	surface/BB	2,618	13,311		99		
1995	surface/BB	4,512	12,297	433			
1996	surface/BB	10,893	13,452	2,211			

1997	surface/BB	14,784	15,501	834	148		
1998	surface/BB	8,587	14,395	1,145	98		
1999	surface/BB	11,971	12,858	3,029	70	3,672	
2000	surface/BB	14,122	13,982	1,209	46	2,706	
2001	surface/BB	14,731	13,181	2,952	27	2,609	
2002	BB	13,635	17,932	761	24	2,349	
2003	BB	12,626	13,915	3,876	79	2,754	
2004	BB	14,474	11,798	897	450	3,288	
2005	BB	12,651	14,650	2,028	643	3,183	
2006	BB	14,976	13,399	2,621	679	3,179	
2007	BB	16,285	13,964	3,930	455	3,026	
2008	BB	15,545	13,460	1,527	304	1,500	
2009	BB	15,706	13,583	4,922	276	1,500	
2010	BB	14,080	16,106	1,915	21	1,500	625
2011	BB	14,451	11,391	4,592	26	1,500	676
2012	BB	17,896	12,902	1,639	9	1,500	1164
2013	BB	12,717	11,039	2,044	19	1,503	1020
2014	BB	16,194	12,606	1,476	22	1,500	1096
2015	BB	15,815	12,312	2,133	24	1,500	1072
2016	BB	15,113	11,618	742	21	1,502	902
2017	BB	12,565	10,821	1,384	15	1,500	756
2018	BB	12,935	11,013	576	39	1,499	798
2019	BB	17,716	10,711	1,640	34	1,497	925
2020	BB	10,323	10,568		34	1,413	577
2021	BB	12,253	11,051	1,444	38	1,500	547
2022	BB	9,702	10,942	1,902	39	2,334	519
2023	BB	8,506	10,932	(3,147)		(1,958)	

Notes:

- Star (*) indicates blind side otolith.
- BB stands for ‘break and bake’ approach.
- All otoliths reported in this table were aged with the exception of the clean collection.
- All aged otoliths are stored in glycerol/thymol solution.
- Some small fish from trawl survey collection are still aged by surface method; otoliths with surface age>4 are broken and baked.
- Sample data not entered prior to 1960 for FISS, 1964 for commercial, 1961 for NOAA trawl survey.
- Clean collection is not aged, stored dry, and include paired otoliths.
- Tribal otoliths are included in the Market Sample series.
- Additionally, there are 144 not aged 2A recreational otoliths, all from Hein Bank collected between 2004 and 2009.
- Sex information available since 2017 (typically ca. 1 year of lag).
- Trawl and recreational otoliths lag one year in ageing.
- In brackets, otoliths available for ageing but ageing not completed.