Development of a Predictive Model for Estimating Fish Age from Otolith Images in the Southern Gulf of St. Lawrence

WatSpeed Machine Learning Course - Final Report (Group 11)

Chen, Xin[1], Fishman, David[2], Shan, Xiaojin[3], Thaker, Rudra[4], Thankappan, Shinoj[5], and Wu, Xuan[6]

Emails: [1]x75chen@uwaterloo.ca, [2]dfishman@uwaterloo.ca, [3]x23shan@uwaterloo.ca, [4]r2thaker@uwaterloo.ca, [5]sthankap@uwaterloo.ca, [5]x34wu@uwaterloo.ca

December 2024

Contents

1 Objectives

Monitoring fish stocks is a critical component of sustainable fisheries management in the Southern Gulf of St. Lawrence. One key aspect of this work is understanding the population dynamics of various fish stocks, which involves accurate age determination. Age data are essential for modeling growth patterns, understanding reproduction, and assessing the health and sustainability of fish populations.

1.1 Goal of the analysis

The primary objective of this project is to develop and implement a machine learning-based predictive model capable of estimating the age of fish from otolith images. This model will automate the process of age determination, reduce the potential for human error, and provide quicker assessments for large datasets.

Specifically, the objectives of the project include:

- Data Collection and Preparation: Compile an archive of otolith images along with corresponding fish age, length, and weight data.
- Model Development: Create a predictive model using machine learning techniques to automatically identify annuli in otolith images and predict the age of the fish.
- Model Validation: Validate the model's accuracy using a separate set of otolith images and corresponding age data.

The training dataset will consist of fish otolith images originating from two Atlantic Canadian species of economic and ecological importance: American plaice (*Hippoglossoides platessoides*) and Atlantic herring (*Clupea harengus*).

1.2 Rationale behind the analysis

Fish age is commonly determined by examining biological materials such as otoliths (inner ear bones) and scales. These materials exhibit growth rings, or "annuli" (Figure 1 and Figure **??**), which can be counted similarly to tree rings. Each ring represents a period of growth, typically corresponding to one year in the life of the fish. However, manually counting these rings can be time-consuming, subjective, and prone to human error. Organizations

that age fish are typically limited by the amount of time it takes the team of human experts to read the otolith samples.

The development of an automated system for predicting fish age based on otolith images would significantly improve the accuracy and speed of age estimation. At a minimum, having a machine learning (ML) model to complement the work of human agers will provide an aspect of quality control to the aging process. At best, these models might actually help speed up the processing time required for aging fish.

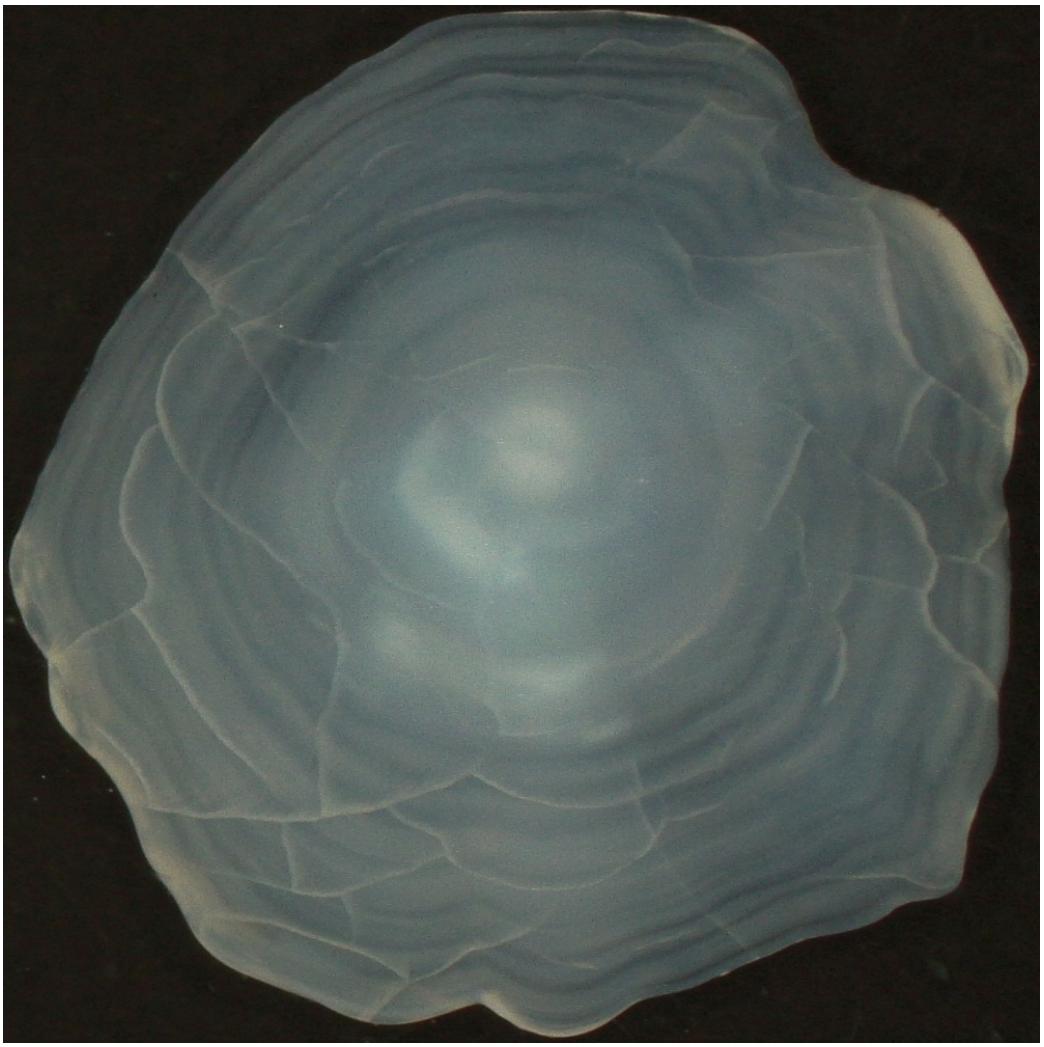

Figure 1: An example of an otolith image taken from an American Plaice.

Figure 2: An example of an otolith image taken from an Atlantic Herring.

2 Data Preparation

2.1 What was your data source?

Lorem ipsum dolor sit amet, consectetur adipisicing elit. A, accusantium cum dolor eveniet facere impedit laborum non recusandae sit velit. Doloribus ducimus est exercitationem libero maxime nostrum omnis quos temporibus. Ecosystem Survey [1]

2.2 How good was the data quality?

Lorem ipsum dolor sit amet, consectetur adipisicing elit. A, accusantium cum dolor eveniet facere impedit laborum non recusandae sit velit. Doloribus ducimus est exercitationem libero maxime nostrum omnis quos temporibus.

2.3 What did you need to do to procure it?

Lorem ipsum dolor sit amet, consectetur adipisicing elit. A, accusantium cum dolor eveniet facere impedit laborum non recusandae sit velit. Doloribus ducimus est exercitationem libero maxime nostrum omnis quos temporibus.

2.4 What tools or code did you need to use to prepare it for analysis?

2.4.1 Fish Specimen Data

Lorem ipsum dolor sit amet, consectetur adipisicing elit. A, accusantium cum dolor eveniet facere impedit laborum non recusandae sit velit. Doloribus ducimus est exercitationem libero maxime nostrum omnis quos temporibus.

2.4.2 Otolith Images

Lorem ipsum dolor sit amet, consectetur adipisicing elit. A, accusantium cum dolor eveniet facere impedit laborum non recusandae sit velit. Doloribus ducimus est exercitationem libero maxime nostrum omnis quos temporibus.

2.5 What challenges did you face?

Lorem ipsum dolor sit amet, consectetur adipisicing elit. A, accusantium cum dolor eveniet facere impedit laborum non recusandae sit velit. Doloribus ducimus est exercitationem libero maxime nostrum omnis quos temporibus.

array

Name	Type	D
uuid	String (or unique identifier)	A unique identifier for each data record. This is useful fo
fish_id	String	A unique identifi

age Integer The known age of the fish, which will be used as the target variable for the model. The primary target variable to predict based on the otolith image features. The model will learn to predict this based on otolith image analysis and possibly other features. length Float The length of the fish, which is often correlated with age and could be useful as a feature for the model. Features like length and weight can help the model better estimate the age since they provide additional biological context for each fish. weight Float The weight of the fish, which, like length, may correlate with age. Another potentially important feature for age estimation, as larger or heavier fish may be older. month Integer the specimen was collected (from the wild) normalized to 1 A potential feature for the model to understand seasonal variations in growth or age estimation. $\text{is}_{m}aleBoolean(0or1)Abinaryvalueindicatingwhetherthefishismale(1)ornot(0).Gendercouldinf$
$specificcharacteristicscouldinfluencethemodel, asgrowthpatternsmaydifferbetweenspecies.T$

3 Analysis

3.1 Our Analysis...

Lorem ipsum dolor sit amet, consectetur adipisicing elit. A, accusantium cum dolor eveniet facere impedit laborum non recusandae sit velit. Doloribus ducimus est exercitationem libero maxime nostrum omnis quos temporibus.

3.1.1 Methodology

Lorem ipsum dolor sit amet, consectetur adipisicing elit. A, accusantium cum dolor eveniet facere impedit laborum non recusandae sit velit. Doloribus ducimus est exercitationem libero maxime nostrum omnis quos temporibus.

3.1.2 Results

Lorem ipsum dolor sit amet, consectetur adipisicing elit. A, accusantium cum dolor eveniet facere impedit laborum non recusandae sit velit. Doloribus ducimus est exercitationem libero maxime nostrum omnis quos temporibus.

4 Conclusions

4.1 Was the model useful?

Lorem ipsum dolor sit amet, consectetur adipisicing elit. A, accusantium cum dolor eveniet facere impedit laborum non recusandae sit velit. Doloribus ducimus est exercitationem libero maxime nostrum omnis quos temporibus.

4.2 What did you learn about your data set?

Lorem ipsum dolor sit amet, consectetur adipisicing elit. A, accusantium cum dolor eveniet facere impedit laborum non recusandae sit velit. Doloribus ducimus est exercitationem libero maxime nostrum omnis quos temporibus.

References

[1] Government of Canada. Nafo division 4t groundfish research vessel trawl survey (september survey) dataset. https://open.canada.ca/data/en/dataset/1989de32-bc5d-c696-879c-54d422438e64, 2024. Accessed on 2024-04-13.