FISH SPECIES CLASSIFICATION USING OTOLITH IMAGES

A Thesis submitted In Partial Fulfillment Of The Requirement For The Degree of

Master Of Technology

in

Computer Technology

Department of Computer Science
and Engineering, Jadavpur University

By

Soumya Bhattacharya
Registration Number:154170 of (2020-2021)
Examination Roll Number:M6TCT23024

Under The Guidance Of

Dr.Nibaran Das
Professor

Department of Computer Science and Engineering
Jadavpur University, Kolkata-700032
India

September, 2023

FACULTY OF ENGINEERING AND TECHNOLOGY
JADAVPUR UNIVERSITY

CERTIFICATE OF RECOMMENDATION

This is to certify that the thesis entitled "**FISH SPECIES CLASSIFICATION USING OTOLITH IMAGES**" has been satisfactorily completed by **Soumya Bhattacharya** (University Registration No.:154170, Exam Roll No.:M6TCT23024). It is a bonafide piece of work carried out under my guidance and supervision and be accepted in partial fulfilment of the requirement for the Degree of **Master of Technology in Department of Computer Science and Engineering, Faculty of Engineering and Technology, Jadavpur University, Kolkata 700032**.

Dr. Nibaran Das (Thesis Supervisor)
Professor
Department of Computer Science and Engineering,
Jadavpur University, Kolkata-700032.

Countersigned

Prof. Nandini Mukherjee
Head, Computer Science and Engineering,
Jadavpur University, Kolkata-700032.

Prof. Saswati Mazumdar
Dean, Faculty of Engineering and Technology,
Jadavpur University, Kolkata-700032.

FACULTY OF ENGINEERING AND TECHNOLOGY
JADAVPUR UNIVERSITY

CERTIFICATE OF APPROVAL

This is to certify that the thesis entitled "**FISH SPECIES CLASSIFICATION USING OTOLITH IMAGES**" is a bonafide record of work carried out by Soumya BHattacharya in partial fulfilment of the requirements for the award of the degree of Master of Technology in the Department of Computer Science and Engineering, Jadavpur University during the period of 2020 to 2023. It is understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn there in but approve the thesis only for the purpose for which it has been submitted.

Signature of Examiner
Date:

Signature of Supervisor
Date:

FACULTY OF ENGINEERING AND TECHNOLOGY
JADAVPUR UNIVERSITY

DECLARATION OF ORIGINALITY AND COMPLIANCE OF ACADEMIC ETHICS

I hereby declare that this thesis entitled "**FISH SPECIES CLASSIFICATION USING OTOLITH IMAGES**" contains literature survey and original research work by the undersigned candidate, as part of his Degree of Master of Technology in Computer Technology.
All information in this document has been obtained and presented in accordance with academic rules and ethical conduct.
I also declare that, as required by these rules and conduct, I have fully cited and referenced all materials and results that are not original to this work.

Name: Soumya Bhattacharya

University Registration No. : 154170 of 2020-2021

Examination Roll No. : M6TCT23024

Thesis Title: FISH SPECIES CLASSIFICATION USING OTOLITH IMAGES

Signature:

____

Date:

____

ACKNOWLEDGEMENT

First and foremost, I would like to start by thanking God Almighty for showering me with the strength, knowledge and potential to embark on this wonderful journey and to persevere and complete the embodied research work satisfactorily.

I am pleased to express my deepest gratitude to my thesis guide, **Dr. Nibaran Das**, Computer Science and Engineering, Jadavpur University, Kolkata for his invaluable guidance, constant encouragement and inspiration during the period of my dissertation.

I am highly indebted to **Jadavpur University** for providing me the opportunity and the required infrastructure to carry on my thesis.

I am also grateful to the **Center for Microprocessor Applications for Training Education and Research Lab, Computer Science and Engineering department, Jadavpur University** for giving me the proper laboratory facilities as and when required.

I am thankful to all the teaching and non-teaching staff whose helping hands have smoothed my journey through the period of my research.

Last but not the least; I would like to thank my family members, classmates, seniors and friends for giving me constant encouragement and mental support throughout my work.

Soumya Bhattacharya

University Registration No. : 154170 of 2020-2021

Examination Roll No. : M6TCT23024

Master of Technology

Department of Computer Science and Engineering

Jadavpur University

ABSTRACT

Otoliths are calcium carbonate structures found in the inner ear of fish. They are used for balance and hearing, and their shape and size can be used to identify fish species and age. The primary objective is to facilitate an automatic otolith identification using the power of deep learning method. The proposed method focused on the utilisation of **transfer learning** to achieve better accuracy over other traditional methods, which is essential in case of availability of limited labeled data.

The objective of using transfer learning in fish otolith identification and classification tasks is to improve the performance of machine learning models by transferring knowledge from a pre-trained model that has been trained on a large dataset of labeled images. This approach proved to be effective in this study with a small dataset of otolith images as it was able to achieve better accuracy beating other models and traditional machine learning methods which relied heavily on feature engineering and feature extractions.

The classification accuracy of the proposed method of fine tuning a pre trained Xception model on imagenet dataset on the otolith dataset is 93.75% which is better than previously used inception-V3 and VGG16 models

List of Keywords

FC Fish Classification
IMP Image Pre-Processing
AI Artificial Intelligence
ML Maxchine Learning
MP Morphological Properties
SVM Support Vector Machine
TP Texture Properties
CC Calcium Composition
FE Feature Extraction
ME Marine Ecology
NBC Naive Bayes Classifier
EIA Environmental Impact Assessment
MLP Multi Layer Perceptron
NNs Neural Networks
CNN Convolutional Neural Network
DCNN Deep Convolutional Neural Network

Contents

List of Figures

List of Tables

List of Algorithms

Chapter 1

INTRODUCTION

1.1 Otolith

Fisheries and aquatic ecology studies often require the identification and analysis of fish species in order to understand their population dynamics, distribution, and life history. One common method for species identification and age determination is the examination of otoliths, calcified structures found in the inner ear of fish. Otoliths vary in shape, size, and patterns among species and can provide valuable information for fisheries management and ecological research. However, the manual identification of otoliths is time-consuming and requires expert knowledge, which can be a limiting factor in large-scale studies. Scientific Rationale: Fish otoliths are used for species classification over other modalities like fish images or any other because they have several advantages:

- Durability: Otoliths are very durable structures and can be preserved for long periods of time, even in harsh environmental conditions. This makes them ideal for use in studies of ancient fish populations.

- Availability: Otoliths are present in all bony fishes, and can be collected from a variety of sources, including commercial fisheries, recreational fishing, and environmental surveys.

- Specificity: Otolith morphology can vary significantly between different fish species, even closely related species. This makes otoliths a very reliable tool for species identification.

- Non-destructive sampling: Otoliths can be removed from fish without harming the fish itself. This is important for studies of living fish populations.

Other methods of species classification, such as fish images, can be more challenging to use. Fish images can be difficult to identify accurately, especially for species that are similar in appearance. Additionally, fish images can be affected by factors such as lighting, camera angle, and water clarity.

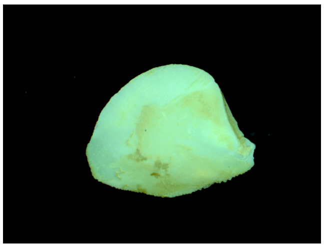
The proposed work aims to develop a novel approach of utilising the area of deep learning by leveraging the power of **transfer learning** to efficiently classify fish species by otolith identification. This approach is critical as otolith image data is not abundantly available, hence a limited collection of labeled data is something researchers have to rely upon. The proposed method successfully solves the classification tasks with much better accuracy than previous traditional models

To develop and optimize computer vision algorithms for otolith feature extraction, including shape, size, and patterns.

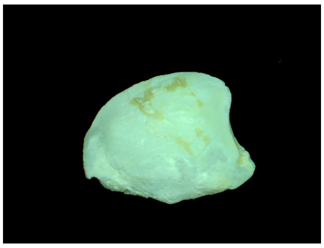
To implement machine learning algorithms, such as convolutional neural networks (CNNs), for otolith classification based on the extracted features.

The successful development of an automated otolith identification system will have significant implications for fisheries research, management, and conservation. By streamlining the identification process, researchers will be able to process larger sample sizes and obtain more accurate data on fish populations. This, in turn, will inform more effective management strategies and contribute to the long-term sustainability of fish stocks and ecosystems.

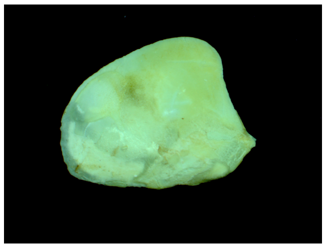
Furthermore, the proposed work has the potential to serve as a foundation for additional research and applications in related fields. For instance, the developed algorithms could be adapted for the identification of other biological



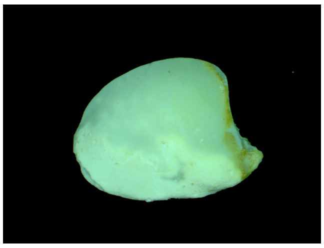
(a) ariMac family



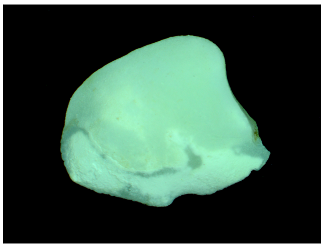
(b) ariMac family



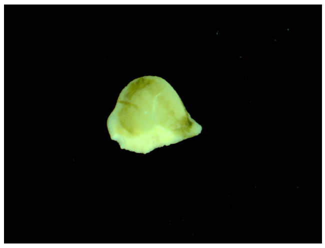
(c) Crytru family



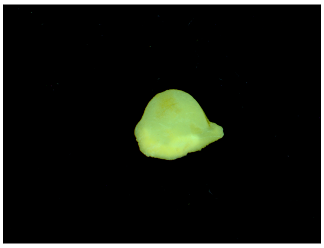
(d) Crytru family



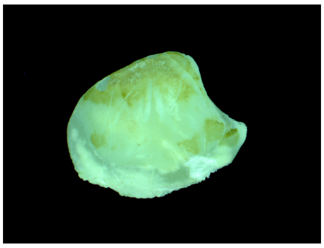
(e) HexSag family



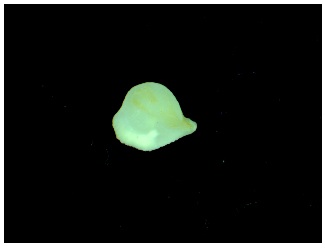
(f) Ostmil family



(g) PliArg family



(h) Nemcae family



(i) Ostmil family

Figure 1.1: Collection of fish otolith images from dataset

structures, such as scales or bones, in ecological and paleontological research. The techniques used in this work could also be applied to other computer vision tasks, contributing to the broader field of artificial intelligence and machine learning.

In conclusion, the proposed work seeks to harness the power of computer vision and machine learning to address the challenges associated with otolith identification in fisheries research. By automating this critical process, the project aims to improve the efficiency and accuracy of species identification and age determination, ultimately contributing to the long-term sustainability of fish stocks and aquatic ecosystems.

1.2 Objectives of Otolith images in fish species classification

The primary idea of this project is to aid the human-intensive fishery industry in order to classify the fish species for both biological research and consumer interests. The traditional industry is heavily based on human expertise.

As the required skill set for this job is in decline the best alternative is to train machines to do the task for us as this is highly efficient. The primary goal is to look for various properties of Otoliths and exploit the best possible ones in order to determine the fish species and other aspects such as fish age and diet. Another critical aspect of the project is to understand marine ecology and the balance that it sustains.

To preserve the various existing species and their properties in a scale-able database. When required researchers can retrieve the information in order to aid their research works. The propensity of various methods and applications could be measured to select the optimum methodology for location or habitat specific environment as well as the type of data that are available in real world.

1.3 Expected output and outcome of the proposal

The expected output of the proposed work includes:
A comprehensive data-set of annotated otolith images, representing a wide range of fish species and age classes. This data-set would be valuable for training and testing the machine learning algorithms and may also serve as a resource for other researchers in the field.
Optimized computer vision algorithms for otolith feature extraction, capable of accurately identifying shape, size, and patterns in otolith images. These algorithms will form the basis of the automated identification system.
A machine learning-based otolith classification system, leveraging convolutional neural networks (CNNs) or other suitable algorithms such as Support vector Machines (SVMs). This system will be capable of accurately identifying fish species and estimating age based on otolith features.
A robust validation of the developed system using a diverse set of otolith samples, demonstrating its accuracy, reliability, and generalizability.

The expected outcomes of the proposed work include:

- Increased efficiency and accuracy of otolith identification in fisheries research, enabling researchers to process larger sample sizes and obtain more reliable data on fish populations.
- Improved fisheries management and conservation strategies, informed by the enhanced understanding of fish population dynamics, distribution, and life history provided by the automated otolith identification system.
- Wider applications of the developed computer vision and machine learning techniques to other biological structures and object recognition tasks, contributing to the advancement of artificial intelligence research and its

application in ecology, paleontology, and beyond.

- Capacity-building in the fields of fisheries research and artificial intelligence, through the development of new tools, techniques, and expertise.

By automating the otolith identification process [1], the proposed work will significantly impact fisheries research, management, and conservation. This will lead to more effective strategies for ensuring the long-term sustainability of fish stocks and aquatic ecosystems, while also fostering advances in computer vision and machine learning.

1.3.1 Timeline of the proposed work

The project unfolded over a 12-month period and was delineated into several stages, each characterized by a set of distinctive objectives and tasks. Here is a retrospective account of each phase:

- Data Collection and Annotation - Otolith samples from various fish species and age classes were collected. High-quality otolith images were prepared and curated. Images were annotated with pertinent information including species identification and age data.

- Development of Feature Extraction Algorithms - Suitable computer vision techniques for otolith feature extraction were researched and selected. Algorithms designed to extract data on shape, size, and pattern from the otolith images were developed and optimized. A subset of the collected otolith images was used to test and refine the algorithms.

- Implementation of Machine Learning Algorithms - Appropriate machine learning algorithms, such as convolutional neural networks (CNNs) or support vector machines (SVMs) with various kernel functions, discriminant analysis, or multilayer perceptrons (MLPs), were chosen for otolith classification. The machine learning model was trained using the annotated otolith dataset. Efforts were undertaken to optimize the model, enhancing its accuracy and efficiency.

- Validation and Testing - The developed system was validated employing a diverse set of otolith samples, not previously included in the training dataset. The system's accuracy, reliability, and generalizability were assessed. Based on the validation outcomes, the algorithms and models underwent further refinements.

- Documentation and Dissemination - Comprehensive documentation encompassing the developed algorithms, system, and validation outcomes was prepared. Research findings were published in peer-reviewed journals and showcased at pertinent conferences. User-friendly software or tools, envisioned for application in fisheries research and management settings, were developed.

1.3.2 Origin of the Proposal

Fisheries and aquatic ecology studies often require the identification and analysis of fish species in order to understand their population dynamics, distribution, and life history. One common method for species identification and age determination is the examination of otoliths, calcified structures found in the inner ear of fish. Otoliths vary in shape, size, and patterns among species and can provide valuable information for fisheries management and ecological research. However, the manual identification of otoliths is time-consuming and requires expert knowledge, which can be a limiting factor in large-scale studies. Scientific Rationale: The rapid advancement of computer vision and machine learning techniques offers a promising solution to the challenges associated with manual otolith identification.

These technologies have been successfully applied to various object recognition and classification tasks, and their application to otolith identification could significantly improve the efficiency and accuracy of species identification and age determination in fisheries research. The proposed work aims to develop a novel computer vision and machine learning-based system for otolith identification, which will automate the process and minimize human error. The specific objectives of this work are:

To collect and curate a comprehensive dataset of otolith images, representing a wide range of fish species and age classes.

To develop and optimize computer vision algorithms for otolith feature extraction, including shape, size, and patterns.

To implement a transfer learning method to classify the species by otolith identification.

To validate the accuracy and robustness of the developed system using a diverse set of otolith samples.

The successful development of an automated otolith identification system will have significant implications for fisheries research, management, and conservation. By streamlining the identification process, researchers will be able to process larger sample sizes and obtain more accurate data on fish populations. This, in turn, will inform more effective management strategies and contribute to the long-term sustainability of fish stocks and ecosystems.

Furthermore, the proposed work has the potential to serve as a foundation for additional research and applications in related fields. For instance, the developed algorithms could be adapted for the identification of other biological structures, such as scales or bones, in ecological and pale-ontological research. The techniques used in this work could also be applied to other computer vision tasks, contributing to the broader field of artificial intelligence and machine learning. In conclusion, the proposed work seeks to harness the power of computer vision and machine learning to address the challenges associated with otolith identification in fisheries research. By automating this critical process, the project aims to improve the efficiency and accuracy of species identification and age determination, ultimately contributing to the long-term sustainability of fish stocks and aquatic ecosystems.

1.3.3 Survey on recent developments in otolith identification

The use of otoliths for species identification and age determination has long been a crucial aspect of fisheries research and management. In recent years, significant progress has been made in the development and application of computer vision and machine learning techniques to automate otolith analysis, improve accuracy, and increase efficiency. Recent developments in otolith identification and related techniques

Otolith Shape Analysis: Researchers have increasingly focused on the analysis of otolith shape [2], as it carries valuable information about species identification and age determination. Various techniques, such as geometric morphometrics and elliptic Fourier analysis, have been used to quantify otolith shape and identify species-specific characteristics.

Computer Vision for Feature Extraction: Several studies have explored the use of computer vision techniques [3] to automate the extraction of otolith features, such as shape, size, and patterns. Image processing software like ImageJ, along with custom algorithms and plugins, have been developed for this purpose. These techniques enable the identification of relevant otolith features without manual intervention, thus increasing efficiency and reducing human bias.

Machine Learning for Otolith Classification: Machine learning techniques and deep learning techniques [4], particularly convolutional neural networks (CNNs), have emerged as a powerful tool for otolith classification based on the extracted features, also Support Vector Machines (SVMs) can perform competively enough along with the CNNs with their different kernel functions tweaked according to the data that is available or what is needed. Researchers have successfully applied machine learning algorithms to classify otoliths by species [5], stock, or age, demonstrating the potential of these techniques for automating otolith analysis.

Integration of Expert Knowledge: While significant progress has been made in automating otolith analysis using computer vision and machine learning techniques, the integration of expert knowledge remains essential. Researchers continue to explore methods for incorporating human expertise into the automated systems, such as through the use of expert-annotated training data or ensemble learning techniques that combine multiple classification algorithms.

In conclusion, the research and development in the field of otolith identification using computer vision and machine learning [6] have advanced considerably in recent years. This progress has led to the development of more accurate and efficient methods for otolith analysis, with the potential to transform fisheries research and management. However, challenges remain in terms of data-set availability, algorithm optimization, and integration of expert knowledge. Further research and collaboration among researchers, fisheries managers, and artificial intelligence experts will be crucial for addressing these challenges and realizing the full potential of computer vision and machine learning techniques in otolith analysis.

(Orlando et al., 2017):

In a collaborative study between French and American researchers, Orlando and Carlson developed an otolith shape analysis method based on elliptical Fourier descriptors [7] to identify fish species. Their work demonstrated that this method could be used to discriminate among different fish species effectively.

(Horstmann et al., 2018):

Horstmann and his team used machine learning algorithms to analyze otolith shape data for fish stock identification. They applied various classification techniques [8], including Support Vector Machines (SVM), Random

Forest, and k-Nearest Neighbors (k-NN), to otolith data from the North Sea and Baltic Sea, demonstrating the potential of machine learning for otolith-based fish stock identification.

(Monteiro-Neto et al., 2018):

In a collaborative study between Brazilian and French researchers, Monteiro-Neto and Gerlotto [9] utilized artificial neural networks (ANNs) to identify fish species based on otolith shape analysis. Their work showed that ANNs could effectively classify fish species using otolith shape descriptors, with accuracy rates above 90%.

(Prista et al., 2021):

Prista and his team developed an open-source software called "otoliths" for otolith shape analysis using R programming language. This software allows researchers [10]to perform various otolith shape analyses, such as geometric morphometrics, elliptic Fourier analysis, and machine learning-based classification.

(Albouy and Rodriguez, 2020):

In a study by Canadian researchers, Albouy and Rodriguez applied convolutional neural networks (CNNs) [11] to otolith images to automate fish species identification. They trained a deep learning model on a dataset of otolith images, achieving an accuracy rate of over 95% in species identification .

These researchers and their contributions reflect the growing interest and progress in the field of otolith identification using computer vision and machine learning. Their work demonstrates the potential of these techniques to improve [12]the efficiency and accuracy of otolith analysis, with applications in fisheries research and management.

These studies [13] and the researchers behind them have significantly advanced the field of otolith identification using computer vision and machine learning [14]. Their work not only demonstrates the potential of these techniques to improve the efficiency and accuracy of otolith analysis, but also showcases the importance of international collaboration in driving progress in this area. As a result, fisheries research and management stand to benefit from the development of more accurate and efficient methods for species identification and age determination, ultimately contributing to the long-term sustainability of fish stocks and aquatic ecosystems.

1.4 Importance of the proposed idea in the context of current status

Transfer learning is a machine learning technique where a model developed for one task is reused as the starting point for a model on a second task. This can be done by using the trained model as a feature extractor, or by fine-tuning the entire model on the new task.

Transfer learning is important because it can save time and resources, and improve the performance of machine learning models.

Save time and resources Training a machine learning model from scratch can be time-consuming and computationally expensive, especially for complex tasks. Transfer learning allows you to leverage the knowledge and features learned from a pre-trained model, which can significantly reduce the time and resources required to train a new model.

Improve performance Transfer learning can also improve the performance of machine learning models, especially for tasks where there is limited data available. A pre-trained model has already learned a lot about the world, and this knowledge can be transferred to the new task, even if the new task is very different.

Other benefits Transfer learning can also offer other benefits, such as:
Making machine learning more accessible: Transfer learning can make machine

learning more accessible to people who do not have the resources or expertise to train models from scratch.

- Improving the interpretability of models: Transfer learning can help to improve the interpretability of machine learning models, by making it easier to understand how the model is making predictions.

- Reducing the risk of bias: Transfer learning can help to reduce the risk of bias in machine learning models, by using a pre-trained model that has been trained on a large and diverse dataset.

Transfer learning is used in a wide variety of applications, including: **Image classification:** Transfer learning is commonly used for image classification tasks, such as identifying objects in images or classifying medical images. **Natural language processing:** Transfer learning is also used for natural language processing tasks, such as sentiment analysis, machine translation, and text summarization. **Speech recognition:** Transfer learning can be used to improve the performance of speech recognition systems, especially for low-resource languages.

1.4.1 Location specificity of the project

The project of otolith identification using computer vision and machine learning is not inherently location-specific. The development of the automated system can be conducted anywhere with sufficient computational resources and access to relevant expertise. However, for the validation and testing phases [15], the project would benefit from being located near regions with a rich diversity of fish species and active fisheries, where a variety of otolith samples can be easily collected.
Given this, there are several potential locations for conducting this project:
Coastal Regions of India: India, with its vast coastline stretching along the Arabian Sea and the Bay of Bengal, is home to diverse marine fish species. Indian coastal cities like Kochi, Visakhapatnam, and Chennai have numerous

fisheries research institutes such as the Central Marine Fisheries Research Institute (CMFRI), Central Institute of Fisheries Technology (CIFT), and the ICAR-Central Inland Fisheries Research Institute (CIFRI) respectively. These locations could provide a rich source of otolith samples for developing and testing the automated system.

Andaman and Nicobar Islands: These islands are known for their high marine biodiversity, making them an ideal location for collecting a wide range of otolith samples.

Freshwater Ecosystems in the Himalayan Region: India's Himalayan region offers a variety of freshwater ecosystems that are home to numerous fish species. Institutes like the ICAR-Directorate of Coldwater Fisheries Research in Bhimtal, Uttarakhand, could serve as potential locations for the project.

The selected location would ideally have an active fisheries research community and facilities for otolith sample collection and processing. The location should also be near fisheries or aquatic ecosystems with a wide range of fish species, to ensure a diverse and representative otolith sample set for the development and testing of the automated system [16]. It is also important to note that collaboration with fisheries research institutes and organizations both locally and internationally would be crucial for the success of this project, providing access to additional otolith samples and expert knowledge.

Chapter 2

LITERATURE SURVEY

A literature survey, often referred to as a literature review, is a critical examination and interpretation of the existing body of knowledge related to a particular research topic or question. It serves as a foundational pillar for any research endeavor and is indispensable for various reasons.
Some works that had been done on this domain using various **machine learning** techniques are:

In the 2023 study "Integrating Machine Learning with Otolith Isoscapes: Reconstructing Connectivity of a Marine Fish over Four Decades" by Arai et al., [17] a meticulous examination of the Northwest Atlantic Mackerel's migration patterns from 1975 to 2019 was conducted. Utilizing both previously published [61,63] and unpublished datasets of otolith stable oxygen ($\delta^{18}O$) and carbon ($\delta^{13}C$) isotope analyses, the research delineated the fish's geographical origins and contingent mixing levels within US waters over the 44-year period. The team employed a machine-learning multi-model ensemble classifier grounded in Bayesian model averaging, a methodology that showcased a high classification accuracy of 84.9%. This approach adeptly revealed the geographic shifts over multi-decadal scales, providing significant insights into the migratory behaviors and historical geographical origins of the species. The high-precision results underscore the potential of integrating machine learning with otolith isoscapes in marine biology studies, paving the way for future research into the dynamic habitats of marine species over extended timescales.

In the 2023 study "Using Machine Learning to Alleviate the Allometric Effect in Otolith Shape-based Species Discrimination: The Role of a Triplet Loss Function" by Chen et al. [18], the researchers tackled the issue of discrimination confusion caused by the allometric growth of otoliths using a dataset of 159 fish samples collected from the Antarctic region.

By introducing a triplet loss function in the machine learning techniques applied, they were able to reduce intraspecific variation and increase interspecific variation, significantly improving classification performance. The classifiers used in this approach surpassed a classification accuracy of 85%, showcasing the potential of this method in enhancing otolith classification, especially in the context of limited sampling, which is crucial for studies in trophic ecology and fish life history.

In the 2021 research article "Otoliths as Taxonomic Tool to Identify Catfishes of the Genus Mystus (Teleostei: Bagridae) from India" by Nair et al. [19], the use of lapillus otoliths was explored as a tool for clarifying the taxonomy of bagrid catfishes from the genus Mystus. The study analyzed the otoliths of five distinct species, comparing various shape indices including circularity, ellipticity, rectangularity, form factor, and roundness.

The research found significant variations in the shape indices studied across the five species, indicating the potential utility of these indices in species classification. Utilizing this approach, the classification matrix achieved a correct classification rate of 68.6%, thus showcasing the promise of using otolith shape indices as a taxonomic tool in the identification of Mystus species. The findings represent a notable step toward the precise classification of Mystus genus catfishes leveraging the unique characteristics of otoliths.

In the 2022 study titled "Population Structure of Indian Mackerel (Rastrelliger kanagurta) in Java and Bali Island, Indonesia Inferred from Otolith Shape" by Wujdi et al. [20], the researchers scrutinized the viability of using otolith shape variability as a tool for stock discrimination. Drawing from a sample of 159 Indian mackerels collected from four fishing ports in Java and Bali, they employed a range of analytical tools including Wavelet coefficients, ANOVA-like permutation tests, and Canonical Analysis of Principal Coordinates (CAP) to model and assess otolith outlines.

Despite the comprehensive approach, the linear discriminant analysis (LDA) yielded a relatively low classification success rate of 44.26%, suggesting that otolith shape variability alone might not be sufficiently reliable for stock discrimination. Nonetheless, the study inferred the presence of at least two distinct stocks contributing to the fishery, a finding with significant repercussions for species management and conservation efforts. The research underscores the complex nature of stock discrimination and hints at the necessity for further studies to enhance understanding and management of the Indian mackerel population in the region.

In the 2021 research by Vasconcelos et al. [21], titled "Thinking of Fish Population Discrimination: Population Average Phenotype vs. Population Phenotypes," the team investigated the genetic polymorphism and phenotypic variation in blue jack mackerel populations. Utilizing a dataset of 670 otoliths from specimens collected between 2005 and 2018, the study focused on the impact of phenotypic variation throughout a latitudinal gradient on stock delimitation.

A critical part of the research was analyzing the otolith shape of Trachurus picturatus to understand the role of average phenotype and morphotypes in stock delimitation. The findings highlighted a clear separation in stocks with a remarkable classification success rate of about 90%, showcasing the potential of using phenotypic variations for precise fish population discrimination.

In the 2019 research conducted by Artetxe-Arrate et al. [22], titled "Otolith Microchemistry: A Useful Tool for Investigating Stock Structure of Yellowfin Tuna (Thunnus albacares) in the Indian Ocean," the team explored the yellowfin tuna stock structure in the Indian Ocean using otolith microchemistry. Utilizing a carefully curated dataset of 67 young-of-the-year (YOY) and age-1 yellowfin tuna, the study analyzed the otoliths for carbon and oxygen stable isotopes (δ^{13}C and δ^{18}O) and specific trace elements.

The findings revealed regional variations in the levels of trace elements Ba, Mg, and Mn in YOY otoliths, which enabled an 80% accurate classification of the tuna to their natal origins through linear discriminant analysis. The study affirms the potential of otolith microchemistry in examining fish stock structures, providing a vital tool in marine biology research.

In the 2019 research by Matta et al. [23], titled "Spatial and Temporal Variation in Otolith Elemental Signatures of Age-0 Pacific Cod (Gadus macrocephalus) in the Gulf of Alaska," the authors analyzed the elemental signatures in the otoliths of Pacific cod to understand regional population dynamics in the Gulf of Alaska.

Utilizing quadratic discriminant analysis, they identified nursery habitats with an overall classification success rate of 59%, which increased to 78% when distinguishing between larger spatial scales (eastern vs. western Gulf of Alaska). The study highlighted the potential of otolith micro-chemistry as a tool to discern seasonal and spatial variations influencing Pacific cod populations.

In the 2017 study by Jones et al. [24], titled "Classification of Otoliths of Fishes Common in the Santa Barbara Basin Based on Morphology and Chemical Composition," the authors investigated the morphological and chemical attributes of otoliths from 905 fishes in the Santa Barbara Basin to identify distinct taxonomic groups. The study analyzed a variety of morphometric features and chemical compositions, using advanced imaging techniques. The researchers created two classifier groups based on different combinations of these attributes. Utilizing a comprehensive approach, involving both morphometric and elemental traits, yielded the best classification model with a minimal misclassification error of 4.6%. The findings underscore the random forest techniques' superiority over discriminant function methods in classifying taxonomic groups with high precision.

In the 2016 study by Salimi et al. [25], titled "Fully-automated Identification of Fish Species Based on Otolith Contour: Using Short-time Fourier Transform and Discriminant Analysis (STFT-DA)," the researchers devised a fully automated model for distinguishing fish species using otolith contours. This innovative approach leveraged the Short-time Fourier transform (STFT) to extract crucial features from otolith contours, marking its first application in otolith shape research.

Utilizing images of right sagittal otoliths from 14 fish species across three families and applying Discriminant Analysis (DA) for classification, the team managed to achieve an accuracy rate exceeding 90%. This research not only demonstrated the high efficacy of the STFT-DA model but also paved the way for advanced, automated fish species identification through otolith contour analysis.

In the 2016 study "Evaluation of Otolith Shape as a Tool for Stock Discrimination in Marine Fishes Using Baltic Sea Cod as a Case Study" by Hüssy et al. [26], the researchers examined the feasibility of using otolith shape for stock discrimination in mixed-stock scenarios, focusing on two genetically distinct cod stocks in the Baltic Sea.

Through the analysis of 2940 otolith images obtained using a precise setup involving a Leica MZ12 microscope and a Leica DFC290 camera, the study sought to discern the influences of environmental, ontogenetic, and genetic factors on otolith shape. Implementing Linear Discriminant Analysis for this purpose, the researchers found a notable level of success, with an accuracy rate hovering around 80%, especially for the western genotype. The study thus highlights the potential viability of otolith shape analysis as a tool for stock discrimination in marine fishes, albeit acknowledging the complex interplay of various influences on otolith shape.

Along with the advancement of traditional machine learning, **Deep learning** approach was also adopted in the study of otolith recognition and studies of chemical properties:

In the 2023 study by Habouz et al. [27], titled "Efficient Semi-supervised Learning Model for Limited Otolith Data using Generative Adversarial Networks," the researchers analyzed 450 otolith images from 15 different species collected from the Moroccan Atlantic area. Utilizing an innovative semi-supervised classification model based on generative adversarial networks (GANs), they achieved a notable 80% accuracy in classification, outperforming traditional convolutional neural network systems even with a small training dataset. This research, conducted with the technical support of the National Institute of Fisheries Research (INRH) and leveraging advanced tools like the Leica S8 APO stereo microscope and Leica LAS EZ software, paves the way for more accurate otolith image classifications in marine biology.

In the 2023 study "Assessing the Utility of Computer Vision for Age Determination of Gulf Menhaden (Brevoortia patronus)" by Riedel et al. [28], the team explored the automation of fish age determination through computer vision. Utilizing scales and otoliths collected from 1,500 Gulf Menhaden specimens harvested by state agencies across Texas, Louisiana, Mississippi, and Alabama from 2016 to 2018, the researchers applied advanced computational approaches, including convolutional neural networks (CNN) and deep neural networks (DNN). A subset of the data, labeled "S2" and encompassing 500 individual samples from age-0 and 1-year-old fish, served as a significant focus in this investigation. Employing DNN on this dataset yielded the most promising results, achieving a peak validation accuracy of 81.00%. This demonstrated the substantial potential of computer vision in enhancing the precision and efficiency of age determination in marine biology studies, representing a pioneering step toward the automation of fish ageing processes in scientific research.

In the 2020 paper by Moen et al. [29], "Automatic Interpretation of Otoliths Using Deep Learning," the researchers utilized deep learning models to estimate the age of Greenland halibut from otolith images, using a substantial dataset from the Institute of Marine Research. Applying a pre-trained convolutional neural network (CNN), the study found that the model could predict fish age with precision comparable to human experts. The optimum CNN model, selected based on calculated CV values, achieved the lowest mean squared error (MSE) value of 2.65 when using the average predictions from paired otoliths, showcasing its potential utility in ichthyology.

Chapter 3

MATHEMATICAL BACKGROUND

3.1 Introduction

Computer vision, machine learning, and deep learning are interdisciplinary fields that leverage mathematical concepts to design algorithms capable of learning from and performing predictive data analytics. In this document, we delve into the mathematical foundations that underpin these fields.

3.2 Linear Algebra

Linear algebra forms the backbone of the computational methods used in computer vision and machine learning. Matrices and vectors are used extensively to represent data and perform transformations.

3.2.1 Matrices and Vectors

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

3.2.2 Eigenvalues and Eigenvectors

Eigenvalues and eigenvectors are central to many algorithms, helping in dimensionality reduction and understanding linear transformations.

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \tag{3.1}$$

3.3 Calculus

Calculus, specifically differential calculus, is pivotal in optimizing functions which is a key process in machine learning and deep learning.

3.3.1 Gradient Descent

Gradient descent is an optimization algorithm used to find the minimum of a function. The update rule for gradient descent is given by:

$$\mathbf{w}_{\text{new}} = \mathbf{w}_{\text{old}} - \alpha \nabla f(\mathbf{w}_{\text{old}}) \tag{3.2}$$

where α is the learning rate and $\nabla f(\mathbf{w}_{\text{old}})$ is the gradient of the function at the old weight values.

3.4 Probability and Statistics

Understanding probability and statistics is vital for designing and evaluating machine learning algorithms.

3.4.1 Bayesian Theorem

Bayesian theorem is fundamental in many machine learning algorithms, offering a way to update probability estimates:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{3.3}$$

3.4.2 Mean and Variance

The mean and variance are statistical measures used to describe data distributions:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i, \quad \sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \tag{3.4}$$

The fields of computer vision, machine learning, and deep learning are grounded in these and many other mathematical concepts. Understanding these foundations is crucial for delving deeper into the sophisticated algorithms and techniques used in these fields.

3.5 Optimization

Optimization techniques are pivotal in training machine learning models to find the best possible solutions.

3.5.1 Convex Optimization

Convex optimization deals with the problem of minimizing a convex function over a convex set. The basic problem in convex optimization is given by:

$$\text{minimize} \quad f(\mathbf{x})$$

$$\text{subject to} \quad \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \quad \mathbf{h}(\mathbf{x}) = \mathbf{0}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is a convex function, $\mathbf{g}(\mathbf{x})$ represents inequality constraints, and $\mathbf{h}(\mathbf{x})$ represents equality constraints.

3.6 Numerical Analysis

Numerical analysis involves algorithms for solving mathematical problems numerically.

3.6.1 Finite Difference Methods

Finite difference methods are used to approximate solutions to differential equations, and are defined as:

$$\frac{df(x)}{dx} \approx \frac{f(x+h) - f(x)}{h} \tag{3.5}$$

where h is a small increment in x.

3.7 Deep Learning

Deep learning is a subset of machine learning that utilizes neural networks with many layers (hence "deep") to analyze various levels of representation in data.

3.7.1

Forward Propagation

During forward propagation, the input data is passed through the network and the output is computed using the weights and activation functions. The output of a node is given by the equation:

$$y = f\left(\sum_{i=1}^{n} w_i x_i + b\right) \tag{3.6}$$

where:

- y is the output of the node
- f is the activation function
- w_i are the weights
- x_i are the inputs
- b is the bias
- n is the number of inputs

The activation function f introduces non-linear properties to the model, which can be one of several functions such as sigmoid, tanh, ReLU, etc. The choice of activation function can affect the performance of the neural network.

3.7.2 Convolutional Neural Networks (CNNs)

In CNNs, the convolution operation is central, defined as:

$$(S * K)(i, j) = \sum_{m} \sum_{n} S(i - m, j - n) K(m, n) \tag{3.7}$$

where S is the input image and K is the kernel applied to the image.

Activation Functions

Activation functions introduce non-linear properties to the network, allowing it to learn from the error and make necessary adjustments. Here we describe popular activation functions:

Sigmoid Activation Function

The sigmoid function maps any input into a range between 0 and 1. The equation is given by:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{3.8}$$

Derivative:

$$\sigma'(x) = \sigma(x)(1 - \sigma(x)) \tag{3.9}$$

Hyperbolic Tangent (tanh) Activation Function

The tanh function maps any input into a range between -1 and 1. The equation is given by:

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3.10}$$

Derivative:

$$\tanh'(x) = 1 - \tanh^2(x) \tag{3.11}$$

Rectified Linear Unit (ReLU) Activation Function

ReLU is defined as the positive part of its argument. The equation is given by:

$$f(x) = \max(0, x) \tag{3.12}$$

Derivative:

$$f'(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x \leq 0 \end{cases} \tag{3.13}$$

Leaky ReLU Activation Function

Leaky ReLU allows a small, positive gradient when the input is negative. The equation is given by:

$$f(x) = \max(0.01x, x) \tag{3.14}$$

Derivative:

$$f'(x) = \begin{cases} 1 & \text{if } x > 0 \\ 0.01 & \text{if } x \leq 0 \end{cases}$$

representation of activation functions:

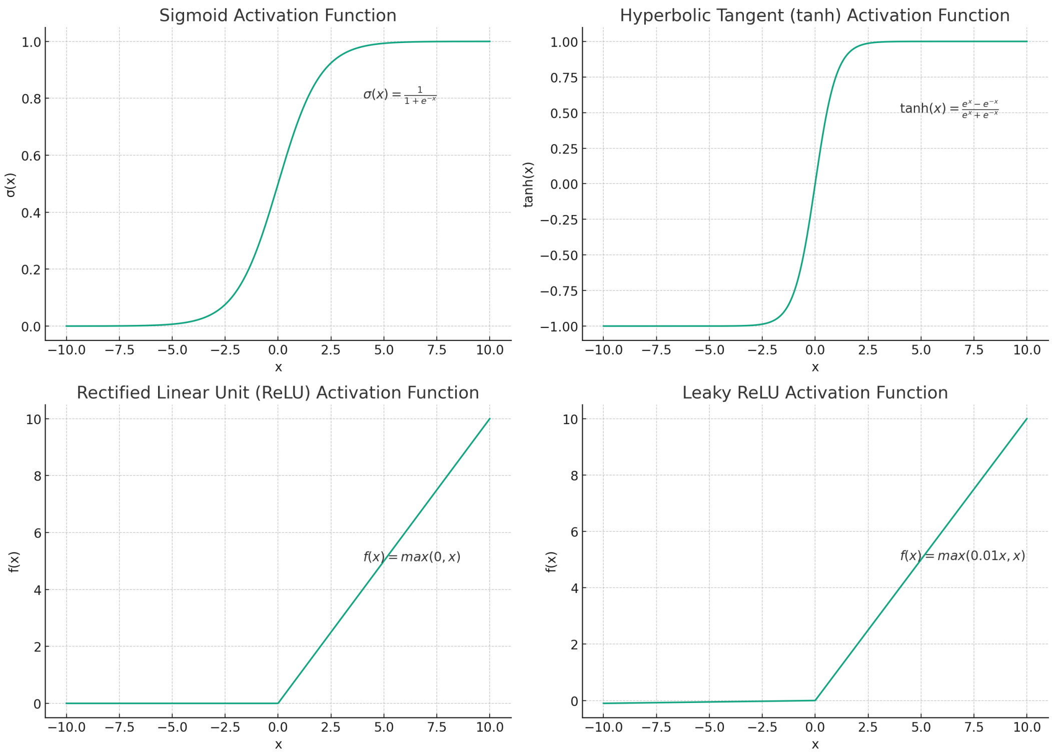


Figure 3.1: Activation functions

3.8 Conclusion

Through a deep understanding of these mathematical foundations, researchers and practitioners can develop more advanced and effective algorithms in computer vision, machine learning, and deep learning. Continuous exploration in these fields opens up unlimited possibilities for technological advancements.

Chapter 4

PROPOSED WORK and METHODOLOGY

4.1 Introduction

In fish otolith classification, the primary objective is to categorize different fish species based on the analysis of their otoliths — calcified structures in the fish's inner ear. These otoliths contain species-specific information that can be leveraged to classify them into different categories. Machine learning, and more recently, deep learning approaches have been employed to enhance the accuracy and efficiency of otolith classification.

The task of fish otolith classification using machine learning involves leveraging a variety of sophisticated algorithms to identify different species based on the characteristics of their otoliths. Recently transfer learning method had been utilized more by researchers compared to previous traditional machine learning methods.

4.2 Dataset description

The Research Institute for Nature and Forest (INBO) [30] as curated a unique dataset featuring photographs and specific data on fish otoliths, which were retrieved from the southern sector of the North Sea. These photographs, which were taken with a binocular microscope (Olympus DP25FW, 6.3X magnification) attached with a digital camera and showcase a black background along with a scale ruler, provide insightful reference material regarding the otoliths and the respective fish species they belong to. This dataset stands as a pivotal tool for researchers, aiding in the accurate identification of various fish species in the southern region of the North Sea through the study of their otoliths. Moreover, it can significantly contribute to the studies focusing on the feeding ecology of seabirds in that area.
The dataset contains a total of 168 saggital otolith images. The data is distributed into training and testing set. The training set is of 71% and testing set is of 29%. the distribution table is given below.

Dataset Break Down

	Training Samples	Testing Samples
Number of Samples	120	48

Table 4.1: Data distribution

The methods of traditional machine learning were first used in the experiments where the desirable features were extracted from the pre-processed images.

The work flow is presented in the following diagram.

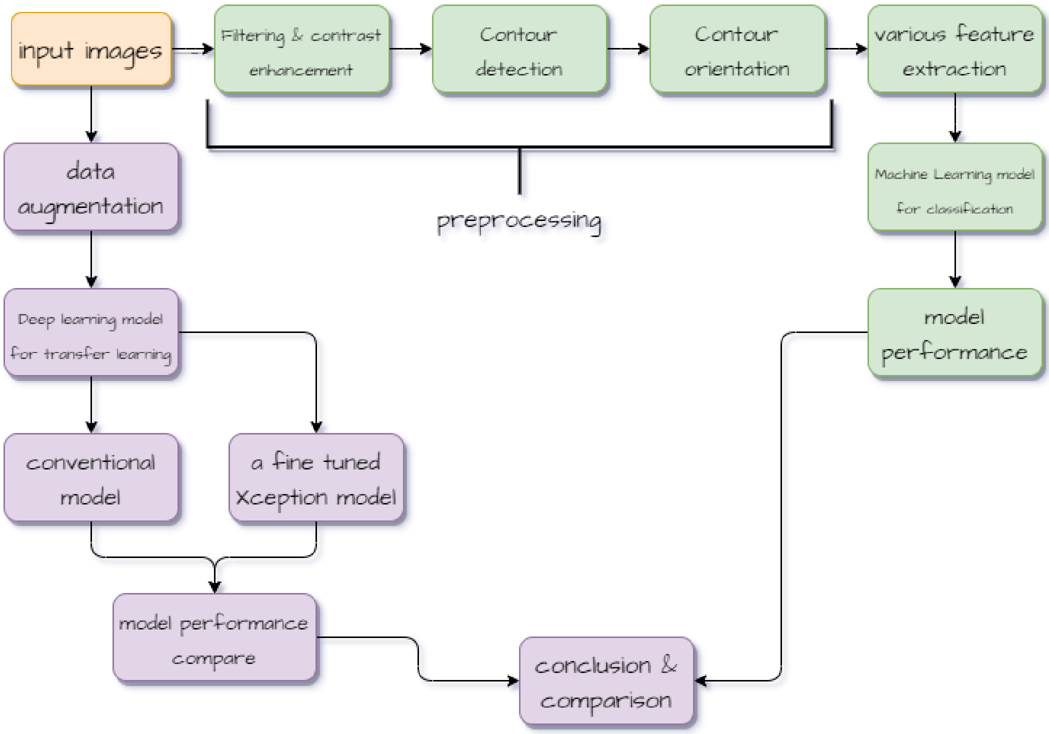


Figure 4.1: The complete workflow diagram

4.3 Morphological Feature extraction

At first, the Otolith images were examined and then a contrast enhancement was implemented on them for better edge detection and morphological analysis. Then, using the moment M, the centroid of the otolith was calculated.

The spatial moments of a contour are given by:

$$M_{ij} = \sum_x \sum_y x^i y^j f(x, y) \tag{4.1}$$

where M_{ij} is the (i,j) moment and $f(x, y)$ is the intensity of the pixel at location (x, y).

The centroid of the contour (C_x, C_y) was then given by:

$$C_x = \frac{M_{10}}{M_{00}}, \quad C_y = \frac{M_{01}}{M_{00}} \tag{4.2}$$

where C_x and C_y are the x and y coordinates of the centroid respectively, and M_{00}, M_{10}, and M_{01} are the zeroth and first order moments.

After that, the area was computed, i.e., the number of pixels in the region of interest. Once these steps were completed, the following algorithm was implemented to perform the thresholding and binarization of the image to obtain better results, plotting the contour points. Two types of filtering were used to reduce noise:

1. Gaussian filter 2. Mean filter

The kernel size was (5,5), and for the Gaussian filter, the standard deviation was 1, so the argument for the Gaussian filter was ((5,5),1). The main contribution of these techniques was to characterize the semantic information contained in the image independently of the background. These methods were often preceded by prefiltering. In the proposed system, the mean filter and mathematical morphology operators were used. Afterward, the contour of the otolith was traced using a threshold method in order to extract the pixels belonging to the contour. In an image of $N \times M$ pixels, the background was clearly different from the Otolith itself. It was supposed that the grey level background value was between a_1 and a_2, and the value of the otolith pixels was from a_3 to a_4 (where a_1, a_2, a_3, and a_4 are fixed values). A pixel belonged to the contour if the grey level of one of its neighbors was between a_1 and a_2.

This algorithm is a part of image processing algorithm designed to identify contour points (boundary pixels) in an image. It works by examining the gray level (intensity) of each pixel and its neighboring pixels to determine if the pixel is a part of a contour.

The transformed image is shown as follows:

Algorithm 1 Image Processing Algorithm

Require: A: the original image

Ensure: C: a set of pixel

$Af \leftarrow$ the image A filtered; $a1, a2, a3, a4$ to be determined
for all pixel $P = Af(x, y)$ in Af **do**
    if $a1 <$ gray level $P < a2$ **then**
        $C(x, y) \leftarrow 255$ (The pixel is not a contour point (white))
    else if $a3 <$ gray level $P < a4$ **then**
        if the gray level of one of its connected neighbor is between $a1$ and $a2$ **then**
            $C(x, y) \leftarrow 0$ (the pixel is a contour point (black))
        else
            $C(x, y) \leftarrow 255$ (the pixel is not a contour point (white))
        end if
    end if
end for

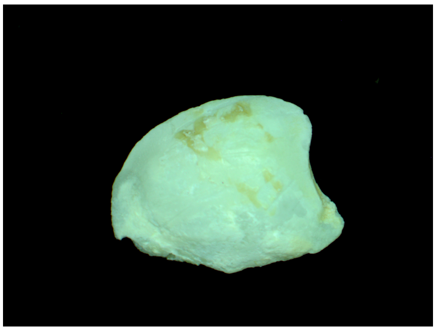


Figure 4.2: Original Otolith image of AriMac family.

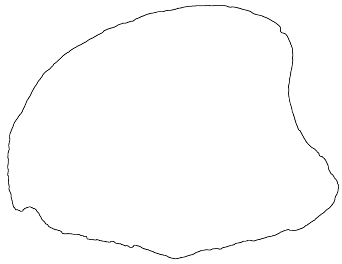


Figure 4.3: image after pixel transform

a step-by-step breakdown of the algorithm:

Input: The input to this algorithm is an image ‘A‘.

Output: The output is a new image ‘C‘, where the pixel intensities represent whether the corresponding pixels in the original image are part of a contour.

Initialization: The algorithm begins by applying some unspecified filter to the original image ‘A‘, resulting in a filtered image ‘Af‘. Four variables ‘a1‘, ‘a2‘, ‘a3‘, and ‘a4‘ are also initialized with values to be determined.

Main Loop: The algorithm then loops over each pixel ‘P=Af(x,y)‘ in the filtered image ‘Af‘.

- If the gray level of the current pixel ‘P‘ falls between ‘a1‘ and ‘a2‘ (exclusive), then ‘P‘ is determined to be not a contour point, and the corresponding pixel in the output image ‘C‘ is set to 255 (white).

- If the gray level of ‘P‘ falls between ‘a3‘ and ‘a4‘ (exclusive), then the algorithm checks the gray level of ‘P‘’s connected neighbors. If any neighbor’s gray level falls between ‘a1‘ and ‘a2‘, ‘P‘ is determined to be a contour point, and the corresponding pixel in ‘C‘ is set to 0 (black). If none of the neighbors’ gray levels fall between ‘a1‘ and ‘a2‘, ‘P‘ is not a contour point, and the corresponding pixel in ‘C‘ is set to 255 (white).

This algorithm assumes that the gray levels of the pixels on the contours of the object in the image are different from the gray levels of the pixels inside and outside the object. By comparing each pixel’s gray level (and the gray levels of its neighbors) to certain thresholds, it determines whether the pixel

is part of a contour.
Otolith Contour Orientation:
The aim of this step is to make a standard orientation for all otoliths images contour. After contour extraction, we detect the principal axis (AB) of the otolith contour:C represents the outline of an otolith.
A= (Xa, Ya) and B= (Xb, Yb) are two points belong to the contour.
The distance AB is the maximum distance between two points of C. Xa, Xb and Ya, Yb are the horizontal and vertical coordinates of A and B respectively. To calculate the distance AB the Euclidian distance was used.

The Euclidean distance between any two points (x1, y1) and (x2, y2) is given by the formula: $D = sqrt((x_2 - x_1)^2 + (y_2 - y_1)^2)$ To find the maximum Euclidean distance among the points of a contour,compute the distance between every pair of points and find the maximum of these. The Euclidean distance between any two points (x_1, y_1) and (x_2, y_2) is given by:

$$D = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{4.3}$$

To find the maximum Euclidean distance among the points of a contour, the distance between every pair of points was computed, and the maximum of these distances was identified.

$$D_{\max} = \max_{i,j} \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{4.4}$$

where $D_{\max}$ is the maximum distance, i and j are indices of the points in the contour, and (x_i, y_i) and (x_j, y_j) are the coordinates of these points.
Given a set of Euclidean distances $D = \{d1, d2, ..., dn\}$ calculated from all pairs of points in the contour, the median Euclidean distance is the middle value when the distances in D are arranged in ascending order. If n is odd, the median distance is the value at position (n+1)/2. If n is even, the median distance is the average of the values at positions n/2 and n/2 + 1.
Given a set of Euclidean distances $D = \{d_1, d_2, \ldots, d_n\}$ calculated from all pairs of points in the contour, the median Euclidean distance is the middle

value when the distances in D are arranged in ascending order.
If n is odd, the median distance D_{median} is the value at position $(n+1)/2$:

$$D_{\text{median}} = d_{\frac{n+1}{2}} \tag{4.5}$$

If n is even, the median distance D_{median} is the average of the values at positions $n/2$ and $n/2+1$:

$$D_{\text{median}} = \frac{d_{\frac{n}{2}} + d_{\frac{n}{2}+1}}{2} \tag{4.6}$$

where D_{median} is the median distance and d_i is the i-the distance when sorted in ascending order.

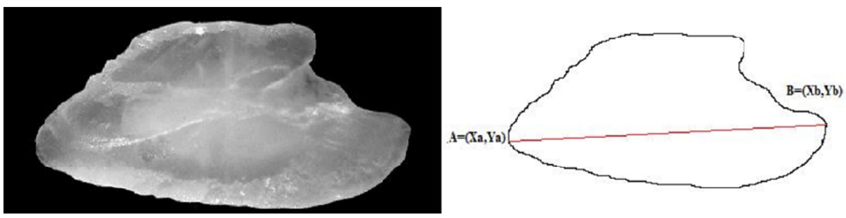


Figure 4.4: The principal axis of the otolith contour: Coris julis

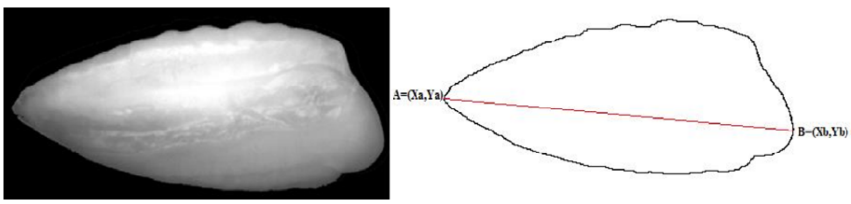


Figure 4.5: The principal axis of the otolith contour: Scomber colias

Otolith contour transformation by rotation was done, this step is included to normalise some of the out of orientation otolith images that were captured during dataset generation. This process would help in further image reading in future methods and algorithms for Otolith recognition and classification.

Algorithm 2 Otolith Contour Transformation Algorithm

Require: C: all the coordinate points (X, Y) forming the contour of the otolith

Ensure: C': the transformation results which we want to perform

AB $\leftarrow$ the principal axis of C
$A = (Xa, Ya) \leftarrow$ the first point of the major axis AB
$\theta \leftarrow$ the angle between AB and the horizontal axis of contour C
if $\theta \neq 0$ **then**
  for each point $P = (X, Y)$ from C **do**
    if $X = Xb$ and $Y = Yb$ **then**
      $X' \leftarrow Xb$, $Y' \leftarrow Yb$
    else
      $L \leftarrow Yb - Y$, $D \leftarrow \sqrt{(X - Xb)^2 + (Y - Yb)^2}$
      $T \leftarrow \arccos(D/L)$
      if $Xa < Xb$ **then**
        {The rotational direction}$\theta, \leftarrow T - \theta$
      else
        $\theta, \leftarrow T + \theta$
      end if
      $X' \leftarrow Xb - D * \sin(\theta_{\text{prime}})$
      $Y' \leftarrow Yb + D * \cos(\theta_{\text{prime}})$
    end if
  end for
else
  $C' \leftarrow C$
end if
Take $n \geq \min(X')$ and $m \geq \min(Y')$ (X', Y') is a point belongs to C'
for each $P' = (X', Y')$ from C' **do**
  if $X' < 0$ OR $Y' < 0$ **then**
    $X' \leftarrow X' + n$
    $Y' \leftarrow Y' + m$
  end if
end for

4.4 Wavelet Transform and Feature extraction

The wavelet transform is a mathematical tool that decomposes a signal into a series of wavelets. Unlike the Fourier transform, which decomposes signals into sines and cosines, the wavelet transform uses wavelet functions that are localized in both time and frequency. This makes it especially useful for analyzing signals that have non-stationary or time-varying characteristics.

There are two primary types of wavelet transforms:

Continuous Wavelet Transform (CWT): This provides a continuous representation of the signal using wavelets. It is defined as:

$$CWT_x(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} x(t)\psi^* \left(\frac{t-b}{a}\right) dt \tag{4.7}$$

Where:

- $CWT_x(a,b)$ is the wavelet coefficient at scale a and position b.
- $x(t)$ is the signal to be analyzed.
- $\psi^*(t)$ is the complex conjugate of the wavelet function $\psi(t)$.
- The integration is performed over all time, t.

Discrete Wavelet Transform (DWT): This provides a discrete representation of the signal using wavelets. It is particularly useful for digital signal processing applications. The DWT is defined in terms of two sets of functions: scaling functions $\phi(t)$ and wavelet functions $\psi(t)$. The DWT can be represented as:

$$DWT_x[j,k] = \int_{-\infty}^{\infty} x(t)\psi\left(2^j t - k\right) dt \tag{4.8}$$

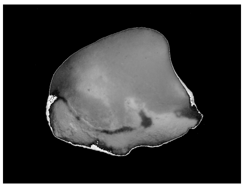
Where:

- $DWT_x[j,k]$ is the wavelet coefficient at scale 2^j and position k.
- The integration is again performed over all time, t.

The wavelets used in the wavelet transform can be of various types, with the most commonly used wavelets being the Daubechies, Haar, and Morlet wavelets, among others.

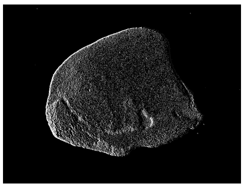
4.4.1 Feature Extraction for Classification using Wavelet Transform

Single-level two-dimensional Discrete Wavelet Transform (2D DWT) on the grayscale images using the Haar wavelet was performed. This process decomposed the image into different frequency components, yielding four sets of coefficients: the approximation coefficients (cA) and three sets of detail coefficients — horizontal (cH), vertical (cV), and diagonal (cD). The cA coefficients represented the low-frequency content of the image, encapsulating its broader structures in a smoothed and down-sampled representation. On the other hand, the detail coefficients (cH, cV, cD) highlighted finer details, capturing high-frequency contents along the horizontal, vertical, and diagonal directions, respectively. These coefficients were individually visualized as grayscale images, offering detailed perspectives into the different features and structures present in the original image. This method of decomposing images into various frequency components was a powerful tool in image processing, widely utilized in tasks such as image compression and feature extraction for analyzing images.

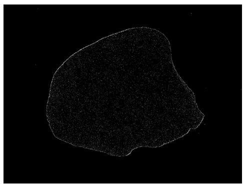


(a) Approximation detail

(b) Horizontal details



(c) Vertical details



(d) Diagonal details

Figure 4.6: DWT transformation on the original image.

4.5 Classification using SVM

Support vector machines (SVMs) are a type of supervised machine learning algorithm that can be used for both classification and regression tasks. However, they are most commonly used for classification. SVMs work by finding the optimal hyperplane that separates the data points into two classes with the maximum margin. This hyperplane is then used to classify new data points.

SVMs were first introduced in the paper "A Support Vector Network for Classification" by Corinna Cortes and Vladimir Vapnik in 1995. Since then, they have become one of the most popular machine learning algorithms due to their high performance and generalizability.

- **SVM for classification:**

  To classify data using SVM, the algorithm first needs to be trained on a set of labeled data points. This means that each data point must be assigned to one of two classes. Once the algorithm has been trained, it can be used to classify new data points by finding the side of the hyperplane on which they fall.

- **SVM for regression:**

SVMs can also be used for regression tasks, but this is less common. In regression, the goal is to predict a continuous value, such as the price of a house or the temperature on a given day. To do this, SVMs use a technique called support vector regression (SVR). SVR works by finding a hyperplane that best fits the training data. This hyperplane is then used to predict the value of new data points.

- **Advantages of SVMs:**

  SVMs have a number of advantages over other machine learning algorithms, including:

  a. High performance: SVMs have been shown to achieve high performance on a wide range of classification and regression tasks.
  b. Generalizability: SVMs are able to generalize well to new data, meaning that they are not overfitting to the training data.
  c. Robustness to outliers: SVMs are robust to outliers, which means that they are not overly influenced by a few data points that are very different from the rest of the data.

- **Disadvantages of SVMs:**

  SVMs also have a few disadvantages, including:

  a. Computational complexity: Training an SVM can be computationally expensive, especially for large datasets.
  b. Black box model: SVMs are a black box model, which means that it is difficult to understand how they make decisions.

The input was the extracted features in the form of a .CSV file which was fed to the model for the desired classification task. There are total of Six class in the dataset and the classification results are obtained accordingly.

The RBF kernel was selected as it yielded best results in all of the kernel functions.

The Radial Basis Function (RBF) kernel, also known as the Gaussian kernel, is one of the most commonly used kernel functions in SVM. It allows SVM to map the input data into a higher-dimensional feature space where linear separation is possible.

Mathematically, let's define a training data-set consisting of n data points:

$$D = \{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$$

where x_i represents the feature vector of the i-th data point, and y_i represents its corresponding class label (either -1 or $+1$).
The goal of SVM is to find a hyperplane in the feature space that separates the data points into two classes while maximizing the margin between the hyperplane and the nearest data points. The hyperplane is defined by the equation:

$$w^T x + b = 0 \tag{4.9}$$

where w is the weight vector normal to the hyperplane, x is the feature vector, and b is the bias term.

To find the optimal hyperplane, SVM solves the following optimization problem:

$$\min_{w,b} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n} \xi_i \tag{4.10}$$

subject to:

$$y_i(w^T x_i + b) \geq 1 - \xi_i$$

$$\xi_i \geq 0$$

where C is a regularization parameter that controls the trade-off between maximizing the margin and minimizing the classification errors, and ξ_i represents the slack variables that allow for some misclassifications.

Now, the RBF kernel is used to transform the input data into a higher-dimensional space using the following formula:

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \tag{4.11}$$

where x and x' are two feature vectors, $\|x - x'\|$ represents the Euclidean distance between the vectors, and σ is a parameter that determines the width of the Gaussian distribution. The RBF kernel measures the similarity between two samples in the feature space.

By applying the RBF kernel, the SVM algorithm constructs a non-linear decision boundary in the original input space, allowing it to capture complex relationships between features and improve the classification accuracy for non-linearly separable data. The RBF kernel uses a Gaussian distribution to measure the similarity between data points in the feature space, leading to flexible and accurate classification boundaries.

Support Vector Machines (SVMs) are originally designed for binary classification problems, where the goal is to separate two classes. However, SVMs can be extended to solve multi-class classification problems using different strategies. Two commonly used approaches are the One-vs-One (OvO) and One-vs-All (OvA) strategies.

- **One-vs-One (OvO)**
  In the OvO strategy, SVM constructs multiple binary classifiers, each trained to differentiate between two classes. For a multi-class problem with k classes, $k(k-1)/2$ binary classifiers are created. Each classifier is trained using a subset of the training data that contains only two classes, and the decision is made by a voting

scheme. During testing, each classifier predicts the class label, and the class with the highest number of votes is chosen as the final prediction. OvO requires training and testing $k(k-1)/2$ binary classifiers.

- **One-vs-All (OvA):**
  In the OvA strategy, SVM constructs k binary classifiers, where each classifier is trained to separate one class from the rest of the classes. During training, the samples from the target class are labeled as positive (+1) and samples from all other classes are labeled as negative (-1).

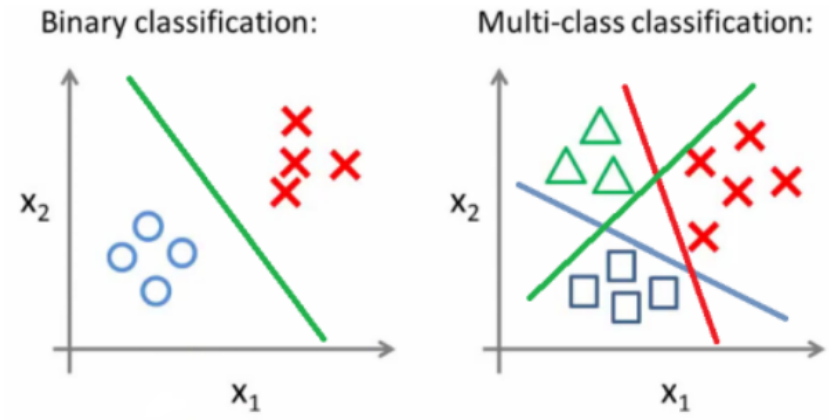


Figure 4.7: Binary classification vs. Multi-class classification

4.6 Transfer learning for classification

Deep learning is a type of machine learning that uses artificial neural networks to learn from data. Neural networks are inspired by the structure and function of the human brain, and they are able to learn complex patterns from large amounts of data.

Utilizing pre-existing models to foster learning in deep neural networks is a strategy known as transfer learning. This approach is favored primarily because it permits the use of smaller datasets in training, thereby conserving both time and computational resources.

In essence, transfer learning leverages knowledge acquired from one problem to solve a related issue. This is particularly useful when there is a scarcity of data available for the new task at hand. It is a cornerstone technique in deep learning, drastically lowering the threshold of data and resources normally needed.

However, the success of this method hinges heavily on the generality of the features learned during the initial task. For instance, a model educated through a diverse set of images, such as those in the ImageNet database, would be a prime candidate for transfer learning in similar endeavors, ensuring a level of versatility and adaptability in processing new tasks. It's a pathway to efficient and expedient model training, permitting new problems to be tackled with a foundational knowledge base.

4.6.1 Xception model

This experiment was done as a Transfer Learning with Xception model acting as the base.

The Xception model, short for "Extreme Inception," is a deep learning architecture developed by François Chollet [31], the creator of the Keras library. The Xception model is an extension of the Inception architecture but replaces the standard Inception modules with depthwise separable convolutions. The architecture aims to improve upon Inception by enhancing the efficiency and effectiveness of the model.

Mathematical Details

Depthwise Separable Convolution:

The main building block of the Xception architecture is the depthwise separable convolution, which is essentially a variant of the standard convolution operation. It separates the learning of spatial features and the learning of channel-wise features, making it more efficient in terms of computational resources.

A depthwise separable convolution can be broken down into two steps:

Depthwise Convolution: Apply a single convolutional filter for each input channel. If there are C channels, then C filters are used. Mathematically, for input X and depthwise filter W_d, the depthwise convolution is:

$$\text{DepthConv}(X, W_d) = X * W_d \tag{4.12}$$

Pointwise Convolution: Apply a 1x1 convolution across the channels. This is also known as channel mixing. Mathematically, for the output Y from depthwise convolution and pointwise filter W_p, the pointwise convolution is:

$$\text{PointConv}(Y, W_p) = Y * W_p \tag{4.13}$$

The whole operation can be represented as:

$$\text{Depthwise Separable Convolution}(X) = \text{PointConv}(\text{DepthConv}(X, W_d), W_p) \tag{4.14}$$

Architecture

The Xception architecture mainly consists of:

- Initial standard convolutions for input pre-processing.
- A series of depthwise separable convolution blocks.
- Global average pooling.
- A fully connected layer for classification.

The architecture can be visualized as follows:

$$
\begin{array}{ccc}
\text{Input} & \rightarrow & \text{Conv2D} & \rightarrow \\
 & & \downarrow & \\
 & & \text{Depthwise Separable Conv Blocks} & \rightarrow \\
 & & \downarrow & \\
 & & \text{Global Avg Pooling} & \rightarrow \\
 & & \downarrow & \\
 & & \text{Fully Connected} & \rightarrow \\
 & & \downarrow & \\
 & & \text{Output} &
\end{array}
$$

The graphical representation of the architecture is as follows:

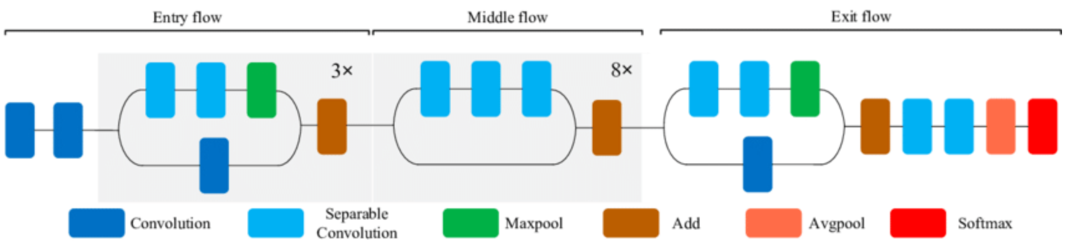


Figure 4.8: Model Architecture

The Xception architecture is composed of 36 convolutional layers, which are all depthwise separable convolutions. The layers are arranged in three flows: the entry flow, the middle flow, and the exit flow. The entry flow is responsible for extracting low-level features from the input image. It consists of two convolutional layers, followed by a max pooling layer.

The middle flow is responsible for extracting high-level features from the input image. It consists of 12 residual blocks, each of which contains a sequence of depthwise separable convolutions and max pooling layers. The exit flow is responsible for combining the features extracted by the middle flow and classifying the input image. It consists of three convolutional layers, followed by a global average pooling layer and a fully connected layer.

The Xception architecture has been shown to achieve state-of-the-art results on a variety of image classification tasks, including the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [32].

- **Entry flow:**
  2x7x7 depthwise separable convolutions with 64 filters each.
  Max pooling layer with stride 2.
- **Middle flow:**
  12 residual blocks, each of which contains the following layers:
  Depthwise separable convolution with 128 filters
  Depthwise separable convolution with 256 filters
  Max pooling layer with stride 2.
- **Exit flow:**
  3x3 depthwise separable convolutions with 256 filters each.
  Global average pooling layer.
  Fully connected layer with 1000 output neurons.

4.6.2 Why is Xception used over InceptionV3?

There are a few reasons why Xception is sometimes used over InceptionV3 in deep learning projects:

Accuracy: Xception has been shown to achieve slightly higher accuracy than InceptionV3 on some image classification tasks.
Efficiency: Xception is more efficient than InceptionV3, both in terms of training time and inference time. This makes it a better choice for mobile devices or other applications where computational resources are limited.
Simplicity: Xception is a simpler architecture than InceptionV3, which makes it easier to understand and debug.

4.6.3 Model fine tuning and Data augmentation

In the initial stage, the **ImageDataGenerator** class was sourced from the TensorFlow Keras library, serving as a pivotal tool in the loading and pre-processing of image data for the diligent training and validation of deep learning frameworks.
Subsequently, a pair of **ImageDataGenerator** entities were conceived; one was dedicated to training data, and the other was for test data. These entities underwent configuration to tailor the images to dimensions of 224x224 pixels, alongside a normalization process that dictated a zero mean and a standard deviation of unity.
In an endeavor to bolster the breadth of the training dataset, a series of data augmentation strategies were enacted through the **ImageDataGenerator** entities. This encompassed:
Rotation: Subjecting the images to random rotations within a maximum limit of 15 degrees.
Shear: Implementing a random shearing effect, confined to a peak of 0.1.
Zoom: Facilitating random zooming actions with an upper bound of 0.2.
Horizontal Flip: Invoking a random horizontal flip on the images.
Width and Height Shift: Altering the width and height randomly, with a restriction up to 0.1.

This meticulous process was orchestrated to foster a dataset robust enough to withstand variations, thereby nurturing a deep learning model with a heightened sensitivity to minor fluctuations in data, and enhancing its generalization capabilities on unseen data.
While the training data underwent these augmentations, the test data remained untouched to maintain its fidelity to the real data that the model would eventually encounter.

Following the augmentation process, the **flowfromdirectory** method was initiated to facilitate the streamlined loading of training and test data from the designated directories, while also setting the batch size and class mode. The batch size delineated the volume of images to be processed in a single training iteration, and the class mode, being categorical, meant that the labels were transformed into one-hot vectors, ensuring a more nuanced classification during the training phase.

The following classes were imported from the TensorFlow Keras library: **Xception**, which was utilized to load a pre-trained Xception model; **Dense**, which facilitated the creation of a dense layer; and the boldly highlighted **GlobalAveragePooling2D**, which enabled the establishment of a global average pooling layer. Additionally, the **Model** class was employed to forge a Keras model and was central to the operational flow. Lastly, the **Adam** class was pivotal in crafting an Adam optimizer.

Chapter 5

RESULTS and DISCUSSION

It was observed that the morphological features were dominated by the physical measurements of the Otoliths which were recorded at the time of preservation. Combining these features with the other morphological features such as contour area, median distance, contour orientation, and the wavelet transform coefficients the best possible classification was possible.

The SVC model was initialized with the kernel='rbf' parameter, which indicated the usage of the Radial Basis Function (RBF) kernel. The gamma='auto' parameter set the gamma value automatically based on the input data. The C=10 parameter set the regularization parameter C to a high value, indicating a strong penalty for misclassifications. The probability=True parameter enabled probability estimation for the SVC model.

The *make_pipeline* function was used to create a pipeline that included data preprocessing with $StandardScaler()$, and the SVC model.
During training, the pipeline was fitted to the training data using X_{train} and y_{train}. Then, the accuracy of the model was calculated using the score method with X_{test} and y_{test}. Predictions were made using X_{test} with the predict method, and a classification report was generated using the $classification\_report$ function.

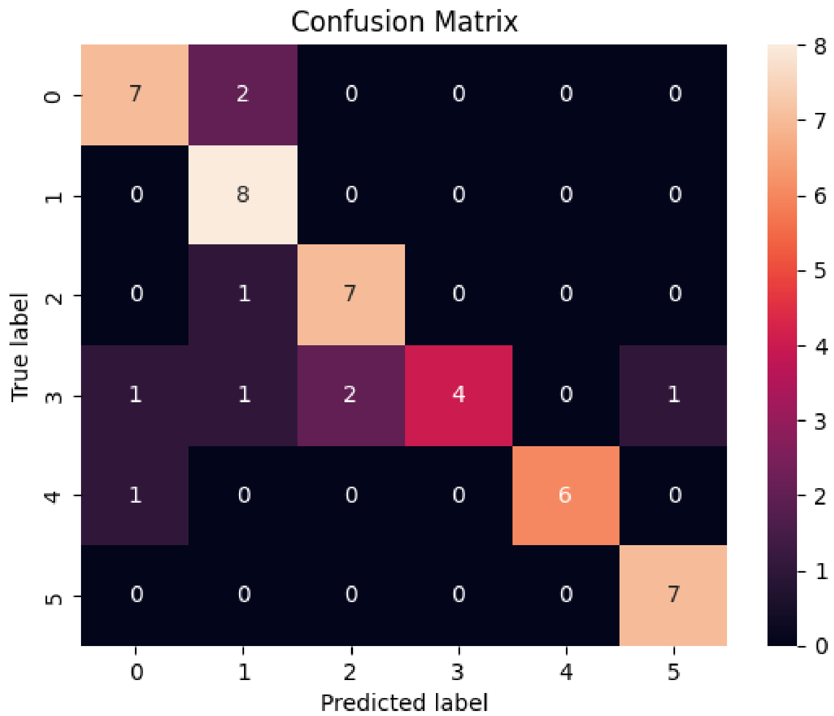


Figure 5.1: Confusion matrix of morphological analysis

Table 5.1: Classification Report of morphological analysis

Class	Precision	Recall	F1-Score	Support
Alosa fallax	0.78	0.78	0.78	9
Ammodytes marinus	0.67	1.00	0.80	8
Ammodytes tobianus	0.78	0.88	0.82	8
Clupea harengus	1.00	0.44	0.62	9
Hyperoplus lanceolatus	1.00	0.86	0.92	7
Sprattus sprattus	0.88	1.00	0.93	7
Accuracy			0.81	48

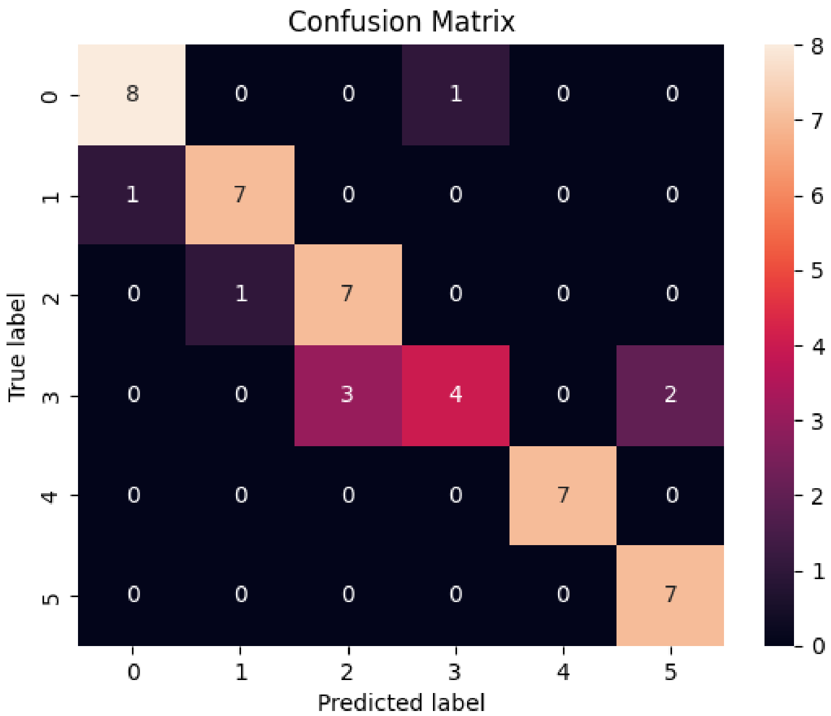
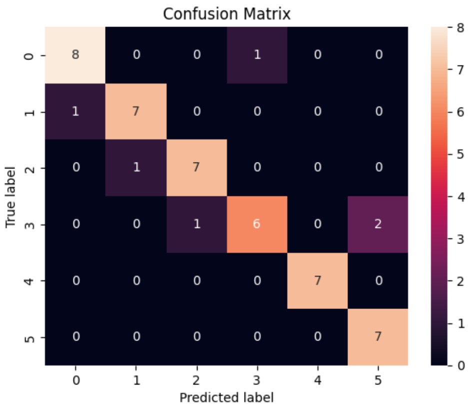


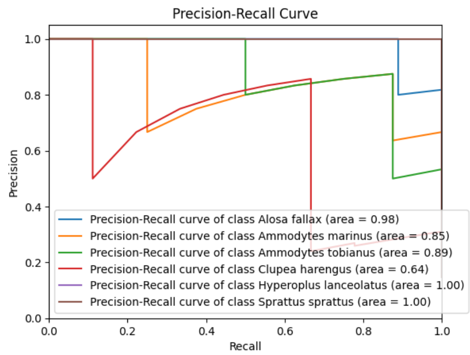
Figure 5.2: Confusion matrix of wavelet analysis

Table 5.2: Classification Report of wavelet analysis

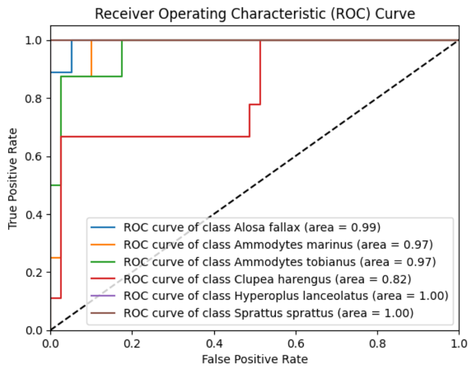
Class	Precision	Recall	F1-Score	Support
Alosa fallax	0.89	0.89	0.89	9
Ammodytes marinus	0.88	0.88	0.88	8
Ammodytes tobianus	0.70	0.88	0.78	8
Clupea harengus	0.80	0.44	0.57	9
Hyperoplus lanceolatus	1.00	1.00	1.00	7
Sprattus sprattus	0.78	1.00	0.88	7
Accuracy			0.83	48



(a) Confusion matrix



(b) Precision-recall curve



(c) ROC curve

Figure 5.3: Performance measure of morphological and wavelet transform

Table 5.3: Classification Report of morphological and wavelet transform

Class	Precision	Recall	F1-Score	Support
Alosa fallax	0.89	0.89	0.89	9
Ammodytes marinus	0.88	0.88	0.88	8
Ammodytes tobianus	0.88	0.88	0.88	8
Clupea harengus	0.86	0.67	0.75	9
Hyperoplus lanceolatus	1.00	1.00	1.00	7
Sprattus sprattus	0.78	1.00	0.88	7
Accuracy			0.88	48

The results of all these experiments performed was tabulated as follows:

Table 5.4: Machine learning model accuracy comparison

Experiments	Results and accuracy
Morphological features	81.25%
Wavelet transform	83.33%
Morphological analysis & Wavelet transform	87.5%

The different experiments that were done concluded that, the deep learning model and fine tuning it resulted much better accuracy in Otolith classification and this model proved to be reliable even with very less number of otolith image data. The dataset contained a total of 168 images of six classes and the proposed fine tuned model has successfully achieved the classification task beating the other Deep Learning models such as **InceptionV3** and deep learning models that were previously used in fish otolith classification.
All the source code for this entire work was compiled and executed on Google colab service Python 3 with support of T4 GPU units.

A pre-trained **Xception** model was loaded and modified by adding a **GlobalAveragePooling2D** layer, a Dropout layer, and a BatchNormalization layer. The **GlobalAveragePooling2D** layer reduces the dimensionality of the output from the Xception model, the Dropout layer prevents overfitting, and the **BatchNormalization** layer normalizes the output from the **Xception** model.

The model was fine-tuned by experimenting with max-pool and average pooling. Finally, a dropout value of 0.75 was found to be best for the classification. This is better than the dropout values of 0.5 and 0.75.

The pre-trained model was not used with the pre-trained weights, but rather the model was trained on the Otolith data-set. This method gave much better accuracy compared with the pre-trained weights of ImageNet dataset.

To optimize the model and prevent over-fitting, Reduce-LR-On-Plateau, early-stopping, and learning-rate-reduction callbacks were used. The categorical cross entropy loss function was used.

In this experiment, **Xception** model was utilized to analyze a dataset of 168 images of fish otoliths. Initially, a model—which underwent training and fine-tuning processes demonstrated a good accuracy of 93.75%. The hyperparameters were fine tuned and the most favourable values were identified.

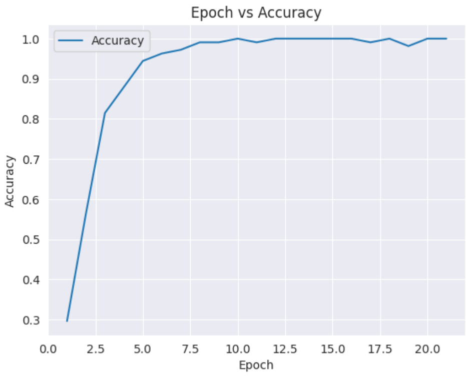
Table 5.5: Hyperparameters

Parameter	**Value**
Batch size	32
Learning rate	3×10^{-4}
No. of epochs	100

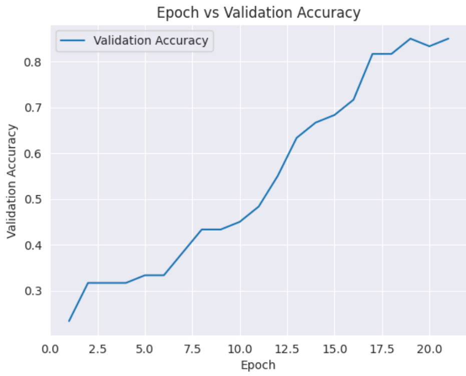
The test results are summarized as follows: The model performance metrics are calculated as follows where the graphs were plotted to have a visual understanding of the proposed model.

Table 5.6: Test Results

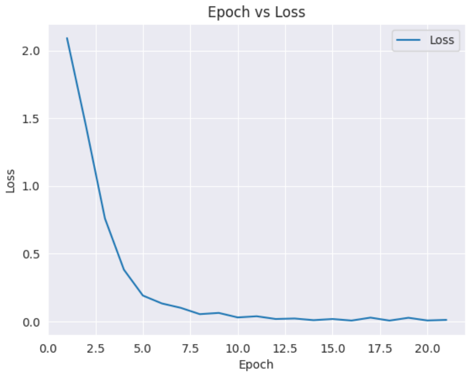
Metric	**Value**
Test loss	0.17406
Test accuracy	0.9375



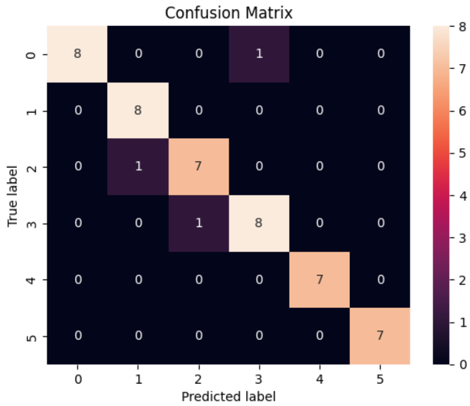
(a) epoch vs accuracy



(b) epoch vs validation accuracy



(c) epoch vs loss



(d) confusion matrix

Figure 5.4: Performance measure of the model.

The effect of utilizing transfer learning proved to be effective and feasible, especially in the case of limited availability of Otolith images as in this case. This method could be implemented to study the other properties such as variation in age growth and the food habits of the fishes as well as stock identifications.

Chapter 6

CONCLUSION and FUTURE WORK

6.1 Conclusion

The intent of this research was to employ transfer learning techniques for the classification of fish species using otolith images. The results obtained from the experiment provided insightful observations regarding the efficiency of different deep learning models in performing the said task.

The analysis of a dataset consisting of 168 fish otolith images using two distinct models revealed noteworthy differences in their performance. The initial model that underwent rigorous training and fine-tuning processes achieved an impressive accuracy of 93.75%. On the other hand, the machine learning approach by using SVM managed an accuracy of 87.5%, falling short by 6.25% compared to the former. This signifies the importance of model selection and fine-tuning in order to obtain the best results, especially when dealing with specific tasks like otolith image classification.

The parameters utilized during the experiment, such as a batch size of

32, a learning rate of 3×10^{-4}, and training over 100 epochs, seemed to be optimal for this dataset as it led to a test loss of 0.17406 and a commendable test accuracy of 93.75%. These metrics further emphasize the potential of the fine-tuned model in achieving high accuracy rates.

One of the primary takeaways from this research is the viability of transfer learning, especially in scenarios where the dataset is limited, as was the case with otolith images in this study. The success of this approach not only ensures efficient species classification but also opens doors to various other applications.

6.2 Future Work

The success of this research points towards several avenues for future exploration and advancement:

- Expansion of Dataset: To enhance the robustness of the model, future research could incorporate a larger and more diverse set of otolith images. This would aid in generalizing the model and improving its accuracy across varied samples.

- Exploring Other Deep Learning Models: While the Xception model and Inception v3 were explored in this study, there exists a plethora of other deep learning models that might offer even better performance for this specific task.

- Study of Other Properties: Leveraging the success of this method, future research could delve into studying other properties of fish, such as variations in age growth, food habits, and stock identifications using otolith images.

- Integration with Other Modalities: Integrating otolith images with other data modalities, like acoustic data or genetic data, might lead to a more comprehensive understanding and better classification results. Real-world Application Development: There lies potential in developing real-world applications, such as mobile apps or web platforms, where researchers or fisheries can instantly classify fish species using otolith images.

In conclusion, the initial success promises a future rife with potential advancements and broader applications that can significantly benefit marine biology, fisheries, and related fields.

Bibliography

[1] D. Ponton, "Is geometric morphometrics efficient for comparing otolith shape of different fish species?" *Journal of Morphology*, vol. 267, no. 6, pp. 750–757, 2006.

[2] M. Castonguay, P. Simard, and P. Gagnon, "Usefulness of fourier analysis of otolith shape for atlantic mackerel (scomber scombrus) stock discrimination," *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 48, no. 2, pp. 296–302, 1991.

[3] V. Parisi-Baradad, A. Lombarte, E. García-Ladona, J. Cabestany, J. Piera, and O. Chic, "Otolith shape contour analysis using affine transformation invariant wavelet transforms and curvature scale space representation," *Marine and freshwater research*, vol. 56, no. 5, pp. 795–804, 2005.

[4] K. Mahé, K. MacKenzie, D. Ider, A. Massaro, O. Hamed, A. Jurado-Ruzafa, P. Gonçalves, A. Anastasopoulou, A. Jadaud, C. Mytilineou *et al.*, "Directional bilateral asymmetry in fish otolith: a potential tool to evaluate stock boundaries?" *Symmetry*, vol. 13, no. 6, p. 987, 2021.

[5] S. E. Campana, "Chemistry and composition of fish otoliths: pathways, mechanisms and applications," *Marine ecology progress series*, vol. 188, pp. 263–297, 1999.

[6] S. E. Campana and J. M. Casselman, "Stock discrimination using otolith shape analysis," *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 50, no. 5, pp. 1062–1083, 1993.

[7] S. Bermejo, "Fish age classification based on length, weight, sex and otolith morphological features," *Fisheries Research*, vol. 84, pp. 270–274, 05 2007.

[8] M. K. Alsmadi and I. Almarashdeh, "A survey on fish classification techniques," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 5, pp. 1625–1638, 2022.

[9] M. K. Alsmadi, K. B. Omar, S. A. Noah, and I. Almarashdeh, "Fish recognition based on robust features extraction from color texture measurements using back-propagation classifier," *Journal of Theoritical and Applied Information Technology*, vol. 18, no. 1, pp. 11–18, 2010.

[10] N. Salimi, K. H. Loh, S. K. Dhillon, and V. C. Chong, "Fully-automated identification of fish species based on otolith contour: using short-time fourier transform and discriminant analysis (stft-da)," *PeerJ*, vol. 4, p. e1664, 2016.

[11] S. Yedier, S. Kontaş Yalçınkaya, D. Bostanci, and N. Polat, "Morphologic, morphometric and contour shape variations of sagittal otoliths of lepidorhombus spp. in the aegean sea," *Anatomia, Histologia, Embryologia*, vol. 52, no. 2, pp. 279–288, 2023.

[12] T. Schulz-Mirbach, F. Ladich, M. Plath, and M. Heß, "Enigmatic ear stones: what we know about the functional role and evolution of fish otoliths," *Biological Reviews*, vol. 94, no. 2, pp. 457–482, 2019.

[13] A. Lombarte and J. Lleonart, "Otolith size changes related with body growth, habitat depth and temperature," *Environmental biology of fishes*, vol. 37, pp. 297–306, 1993.

[14] F. Capoccioni, C. Costa, J. Aguzzi, P. Menesatti, A. Lombarte, and E. Ciccotti, "Ontogenetic and environmental effects on otolith shape variability in three mediterranean european eel (anguilla anguilla, l.) local stocks,"

Journal of experimental marine biology and ecology, vol. 397, no. 1, pp. 1–7, 2011.

[15] M. Vignon and F. Morat, "Environmental and genetic determinant of otolith shape revealed by a non-indigenous tropical fish," *Marine Ecology Progress Series*, vol. 411, pp. 231–241, 2010.

[16] J. F. Downhower, L. S. Blumer, P. Lejeune, P. Gaudin, A. Marconato, and A. Bisazza, "Otolith asymmetry in cottus bairdi and c. gobio," *Pol. Arch. Hydrobiol*, vol. 37, no. 1-2, pp. 209–220, 1990.

[17] K. Arai, M. Castonguay, V. Lyubchich, and D. H. Secor, "Integrating machine learning with otolith isoscapes: Reconstructing connectivity of a marine fish over four decades," *Plos one*, vol. 18, no. 5, p. e0285702, 2023.

[18] Y. Chen and G. Zhu, "Using machine learning to alleviate the allometric effect in otolith shape-based species discrimination: the role of a triplet loss function," *ICES Journal of Marine Science*, p. fsad052, 2023.

[19] S. M. Nair, S. S. Gangan, S. Raut, R. Raghavan, A. P. Kumar, L. Singh, A. Jaiswar, S. Sau, T. Nagesh, R. Trivedi *et al.*, "Otoliths as taxonomic tool to identify catfishes of the genus mystus (teleostei: Bagridae) from india," *Indian Journal of Fisheries*, vol. 68, no. 1, pp. 1–8, 2021.

[20] A. Wujdi, H. J. Kim, and C. W. Oh, "Population structure of indian mackerel (rastrelliger kanagurta) in java and bali island, indonesia inferred from otolith shape," *Sains Malays*, vol. 51, pp. 39–50, 2022.

[21] J. Vasconcelos, A. Jurado-Ruzafa, J. L. Otero-Ferrer, A. Lombarte, R. Riera, and V. M. Tuset, "Thinking of fish population discrimination: population average phenotype vs. population phenotypes," *Frontiers in Marine Science*, vol. 8, p. 740296, 2021.

[22] I. Artetxe-Arrate, I. Fraile, D. A. Crook, I. Zudaire, H. Arrizabalaga, A. Greig, and H. Murua, "Otolith microchemistry: a useful tool for investigating stock structure of yellowfin tuna (thunnus albacares) in the indian

ocean," *Marine and Freshwater Research*, vol. 70, no. 12, pp. 1708–1721, 2019.

[23] M. E. Matta, J. A. Miller, J. A. Short, T. E. Helser, T. P. Hurst, K. M. Rand, and O. A. Ormseth, "Spatial and temporal variation in otolith elemental signatures of age-0 pacific cod (gadus macrocephalus) in the gulf of alaska," *Deep Sea Research Part II: Topical Studies in Oceanography*, vol. 165, pp. 268–279, 2019.

[24] W. A. Jones and D. M. Checkley Jr, "Classification of otoliths of fishes common in the santa barbara basin based on morphology and chemical composition," *Canadian Journal of Fisheries and Aquatic Sciences*, vol. 74, no. 8, pp. 1195–1207, 2017.

[25] N. Salimi, K. H. Loh, S. K. Dhillon, and V. C. Chong, "Fully-automated identification of fish species based on otolith contour: using short-time fourier transform and discriminant analysis (stft-da)," *PeerJ*, vol. 4, p. e1664, 2016.

[26] K. Hüssy, H. Mosegaard, C. M. Albertsen, E. E. Nielsen, J. Hemmer-Hansen, and M. Eero, "Evaluation of otolith shape as a tool for stock discrimination in marine fishes using baltic sea cod as a case study," *Fisheries Research*, vol. 174, pp. 210–218, 2016.

[27] Y. El Habouz, Y. El Mourabit, M. Iggane, H. El Habouz, G. Lukumon, and F. Nouboud, "Efficient semi-supervised learning model for limited otolith data using generative adversarial networks," *Multimedia Tools and Applications*, pp. 1–14, 2023.

[28] R. Reidel and R. T. Leaf, "Assessing the utility of computer vision for age determination of gulf menhaden (brevoortia patronus)," *Available at SSRN 4417898.*

[29] E. Moen, N. O. Handegard, V. Allken, O. T. Albert, A. Harbitz, and K. Malde, "Automatic interpretation of otoliths using deep learning," *PLoS One*, vol. 13, no. 12, p. e0204713, 2018.

[30] H. Verstraete, W. Courtens, R. Daelemans, M. Van de walle, N. Vanermen, P. Desmet, and E. W. Stienen, "Photos and measurements of otoliths from fish caught in the southern part of the North Sea," Oct. 2020. [Online]. Available: https://doi.org/10.5281/zenodo.4066594

[31] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[32] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, pp. 211–252, 2015.