

## Research Article

# Artificial intelligence models for identifying several fish species based on otolith morphology index analysis from nearshore areas of Vietnam

Vu Q.T.<sup>1\*</sup>, Pham T.D.<sup>1</sup>, Nguyen V.Q.<sup>2</sup><sup>1</sup> Joint Vietnam-Russia Tropical Science and Technology Research Center, Ha Noi, Vietnam<sup>2</sup> Vietnam Academy of Science and Technology, Hanoi, Vietnam\*Correspondence: [vuquyetthanh@gmail.com](mailto:vuquyetthanh@gmail.com)**Keywords**

BDP,  
ShI,  
East Vietnam Sea,  
Deep Learning,  
Machine learning,  
Otolith

**Abstract**

Fish species can be identified based on the analysis of morphological indices including basic dimension parameters and shape index. Several pattern recognition methods have been proposed to classify fish species through the morphological characteristics of otolith outlines. Machine learning methods have been applied in various fields, particularly in the differentiation of object shapes. Applying machine learning models to identify species based on basic dimension parameters and shape index of otoliths is highly promising. The purpose of this study is to apply machine learning models to classify marine fish species, aiming to determine which machine learning model and indices are suitable for otolith shape classification. A total of 720 samples of left otoliths (sagittae) from 12 fish species, with 60 individuals per species, were used to develop and evaluate the identification model using Python language. For the first time, a comparative evaluation of six machine learning models and three deep learning models was conducted to distinguish 12 fish species in the nearshore areas of northern and central Vietnam. The results of this study have identified machine learning and deep learning models based on high-performing basic dimension parameter (BDP) and/or shape index ShI indices for species identification. This lays the groundwork for developing software for automatic species or population identification based on otolith morphological analysis.

**Article info**

Received: June 2024

Accepted: August 2024

Published: March 2025



Copyright: © 2023 by the authors.  
Licensee MDPI, Basel, Switzerland.  
This article is an open access article  
distributed under the terms and  
conditions of the Creative Commons  
Attribution (CC BY) license  
(<https://creativecommons.org/licenses/by/4.0/>).

## Introduction

Otoliths, calcium carbonate structures found in bony fishes (excluding lampreys), serve the functions of balance and hearing. Each fish possesses three pairs of otoliths, comprising two smaller pairs (lapilli and asteriscii) and one larger pair (the sagitta) (Schulz-Mirbach *et al.*, 2014; Santos *et al.*, 2017). Research indicates that the shape of the sagittal otolith is correlated with swimming behavior (Volpedo and Echeverra, 2003), age determination (Hosseini-Shekarabi, 2014), and stock distribution (Lombarte and Cruz, 2007; Sadighzadeh *et al.*, 2014; Tuset *et al.*, 2016). Morphometric features of otoliths have been widely utilized for species differentiation and stock identification (Stransky *et al.*, 2008; Tuset *et al.*, 2012; Bani *et al.*, 2013; Sadighzadeh *et al.*, 2014; Vu *et al.*, 2022; Corrêa *et al.*, 2022). Additionally, otoliths retain records of individual fish growth and development (Hosseini-Shekarabi *et al.*, 2014; Yedier, 2021). The sagittal otolith shape has been extensively employed in species, population, and stock identification (Osman *et al.*, 2020; Ghanbarifardi and Zarei, 2021), often reflecting biological uniqueness (Stransky, 2005; Vu and Kartavsev, 2020).

In recent years, otolith morphological analysis has been employed for species or population differentiation (Portnoy and Gold, 2013; He *et al.*, 2017; Vu and Kartavsev, 2020). Apart from conventional species identification methods like morphological analysis and DNA sequencing, otolith-based species identification proves particularly valuable in reconstructing the historical composition

of fish assemblages in archaeology (Lin *et al.*, 2019) or determining fish prey items (otoliths remain intact in digestive waste). Analysis of sagitta size and shape serves as a useful tool for species discrimination and population determination (Levia *et al.*, 1994; Tuset *et al.*, 2003). A recent study utilized BDP and ShI indices of otoliths to differentiate between *Hypomesus japonicus* and *H. nipponensis* species (Vu and Kartavsev, 2020). Furthermore, otolith shape analysis was employed to assess the stock structure of European anchovy (*Engraulis encrasicolus*) along the Tunisian coast, revealing the presence of three distinct stock units with significant implications for fisheries management (Khemiri *et al.*, 2018).

Shape indices, including basic dimension parameter (BDP) and shape index (ShI), are widely utilized to assess differences among fish species (Bani *et al.*, 2013; He *et al.*, 2017; Vu and Kartavsev, 2020). Various statistical analysis methods have been employed to analyze BDP and ShI indices, with numerous studies evaluating the extent of otolith shape differences among fish species. Several studies have assessed the classification ability based on different shape indices. A recent publication by Lin and Al-Abdulkader (2019) used otolith shape analysis to identify species from certain families in the western Persian Gulf. This study revealed that classification based on shape indices still entails error rates, necessitating continued evaluation of accuracy when applying this method for species identification. Additionally, this study affirmed a significant increase in identification capability when combining

shape indices in linear discriminant analysis (LDA). Salimi *et al.* (2016) successfully identified 14 fish species using the STFT shape index set, achieving recognition rates ranging from 70% to 100%. The authors considered these identification results relatively acceptable; however, further evaluation of identification capability through various recognition models is still required. Inadequate features to describe otolith shape can pose challenges for classifiers (Simoneau *et al.*, 2000). Many authors have identified species based on common indices such as BDP, ShI, and a combination of BDP and ShI. BDP comprises length, width, perimeter, and area indices extensively used in otolith research. These indices are frequently employed in species classification as well as population differentiation. ShI, including circularity, roundness, rectangularity, form factor, aspect ratio, and ellipticity, are standard methods for otolith shape analysis (Burke *et al.*, 2009; Keating, 2014; Mapp *et al.*, 2017). The outcomes of these analyses contribute to delineating fish stocks (Agüera and Brophy, 2011; Paul *et al.*, 2013; Ferhani *et al.*, 2021). BDP-ShI complexity: Although BDP usage may provide effective classification in some cases, relying solely on BDP for classification may confuse, particularly when fish species have similar basic dimensions but differing shapes. Thus, employing a combination of BDP and ShI (BDP-ShI) enhances accuracy. The coastal areas surrounding Son Cha Island (Thua Thien Hue province) and Cat Ba Island (Hai Phong province) constitute regions of Vietnam's coastline renowned for their high

species diversity. Commonly encountered species inhabit natural habitats while also serving as commercially important food sources for humans and prey for carnivorous animals in adjacent forest ecosystems. Utilizing otoliths for species identification offers several advantages, including long-term sample storage, minimal space requirements, and simple preservation methods. Additionally, otoliths play a crucial role in studying the feeding habits of piscivorous animals, significantly contributing to our understanding of marine food webs, particularly at higher trophic levels (Dürr and González, 2002; Garcia-Rodríguez *et al.*, 2011). Otoliths are also applied in archaeology to identify past fish species compositions in ecological histories.

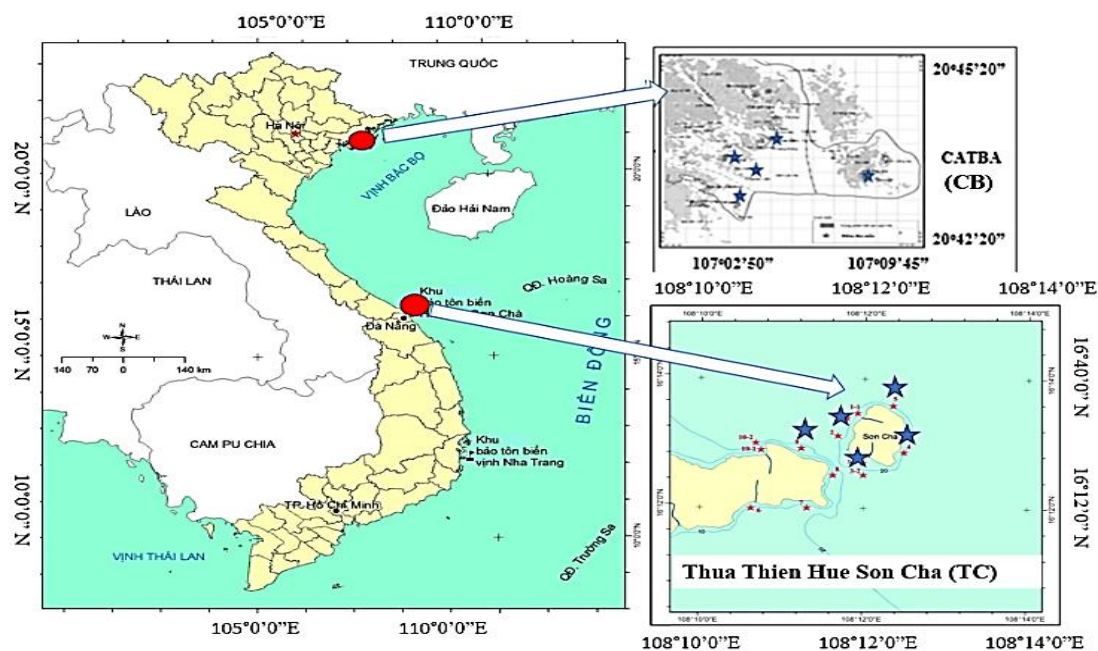
In recent years, advancements in information technology, particularly in artificial intelligence (AI), have opened new avenues for scientific research. Leveraging AI technology for otolith shape analysis holds significant promise for automated species identification. This technology enables the identification of species across large sample sizes in a short period. In this paper, we aim to utilize AI technology to identify the species composition of several commercially important fish species frequently caught around Cat Ba (Hai Phong) and Son Cha (Hue) Islands. Our goal is to assess the applicability of AI models in developing species identification software.

## Materials and methods

*Sample collection.* Individuals were captured by trawling and purchased from fishing grounds in two regions around Cat

Ba and Son Cha, with samples collected from April to June 2024. Fish specimens were identified as adults based on body size and Fishbase documentation (2024). A total

of 720 left otoliths were collected for this study (Fig. 1).



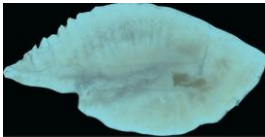
**Figure 1:** Sampling locations of otolith in Son Cha and Cat Ba. The blue stars indicate specific fishing points.

*Image acquisition, digitization, and shape index*

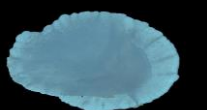
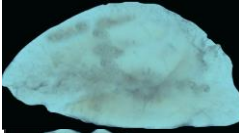
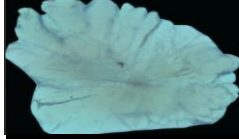
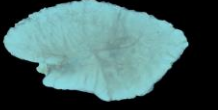
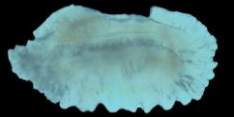
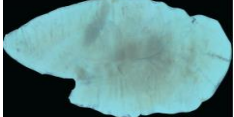


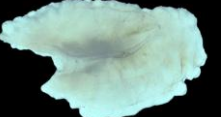
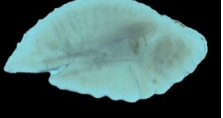
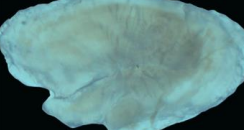
Each otolith was placed against a black background and was standardized by positioning them with the rostrum oriented to the left (Table. 1). The left sagitta was photographed using an Olympus SZ61 zoom stereo microscope. Digital images of otoliths were captured under a stereomicroscope using Olympus CellSens (version 2.2) with an SC180 camera and

saved in JPG and BMP formats. The characteristics of otoliths were measured (in mm) and analyzed based on SHIs. To determine the morphometric characteristics of otoliths, four basic dimensional parameters, namely area (A), perimeter (P), otolith width (OW), and otolith length (OL), were measured using the CellSens (version 2.2).

**Table 1: Information on otolith samples of 12 species collected from Son Cha and Cat Ba (Fig. 1).**

STT	Serial number Python code	Species name	Species code	Image	Sample quantity
1	0	Order: Eupercaria incertae sedis Family: Nemipteridae <i>Pentapodus setosus</i> (Valenciennes, 1830)	1CL		60

**Table 1 (continued):**

STT	Serial number Python code	Species name	Image	Sample quantity
2	1	Order: Siluriformes Family: Plotosidae <i>Plotosus lineatus</i> (Thunberg, 1787)		60
3	2	Order: Holocentriformes Family: Holocentridae <i>Sargocentron rubrum</i> (Forsskål, 1775)		60
4	3	Order: Gobiiformes Family: Gobiidae <i>Acentrogobius caninus</i> (Valenciennes, 1837)		60
5	4	Order: Carangiformes Family: Carangidae <i>Selaroides leptolepis</i> (Cuvier, 1833)		60
6	5	Order: Mugiliformes Family: Mugilidae <i>Crenimugil pedaraki</i> (Valenciennes, 1836)		60
7	6	Order: Carangiformes Family: Carangidae <i>Alepes djedaba</i> (Forsskål, 1775)		60
8	7	Order: Beloniformes Family: Hemiramphidae <i>Hemiramphus</i> sp.		60
9	8	Order: Eupercaria incertae sedis Family: Serranidae <i>Cephalopholis boenak</i> (Bloch, 1790)		60
10	9	Order: Eupercaria incertae sedis Family: Serranidae <i>Diploprion bifasciatum</i> Cuvier, 1828		60
11	10	Order: Carangiformes Family: Carangidae <i>Decapterus macrosoma</i> (Bleeker, 1851)		60
12	11	Order: Perciformes Family: Pomacentridae <i>Abudefduf bengalensis</i> (Bloch, 1787)		60

**Notes: Python code numbers: These are the codes used for training in Python. In the following sections of the article, the sequence number of the Python codes will be used to refer to the species name (using Latin names is difficult to follow).**

Six common ShIs were calculated using the ratios of OW, OL, A, and P (Aguera and Brophy 2011) as follows: Aspect ratio=OL/OW; Ellipticity=(OL-OW)/(OL+OW); Circularity=P/A<sup>2</sup>; Rectangularity=A/(OL×OW); Roundness=4A/πOL<sup>2</sup>; and Form factor=4πA/P<sup>2</sup> (Aguera and Brophy, 2011; He *et al.*, 2017). The BDP-ShI composite: is created by combining the BDP dataset with ShI.

#### Data analysis

Use LDA in Python to evaluate the input data.

Use 6 models to train the data. Random Forest Classifier (RFC) (Breiman, 2001), Extra Trees Classifier (ETC) (Geurts *et al.*, 2006), Gradient Boosting Classifier (GBC) (Friedman, 2001), Bagging Classifier (BaC), AdaBoost Classifier (ABC) (Solomatine and Shrestha, 2004), Hist Gradient Boosting (HGBC) (Pedregosa *et al.*, 2011), with the following key parameters (Fig. 2):

```

RFC: (n_estimators=100,
      random_state=42),
ETC: (n_estimators=100,
      random_state=42),
GBC: (n_estimators=100,
      random_state=42),
BaC: (n_estimators=100,
      random_state=42),
ABC: (n_estimators=100,
      random_state=42),
HGBC: (max_iter=100, random_state=42).
Design a deep learning model using the
Softmax function with classification layers
models = [Sequential([Dense(64,
activation='relu',
input_shape=(X_train.shape[1],)),
Dense(64, activation='relu',
Dense(len(df['Sp'].unique()),
activation='softmax')]),
Sequential([Dense(128, activation='relu',
input_shape=(X_train.shape[1],)),
Dense(128, activation='relu',
Dense(len(df['Sp'].unique()),
activation='softmax')]),
Sequential([Dense(256, activation='relu',
input_shape=(X_train.shape[1],)),
Dense(256, activation='relu',
Dense(len(df['Sp'].unique()),
activation='softmax')

```

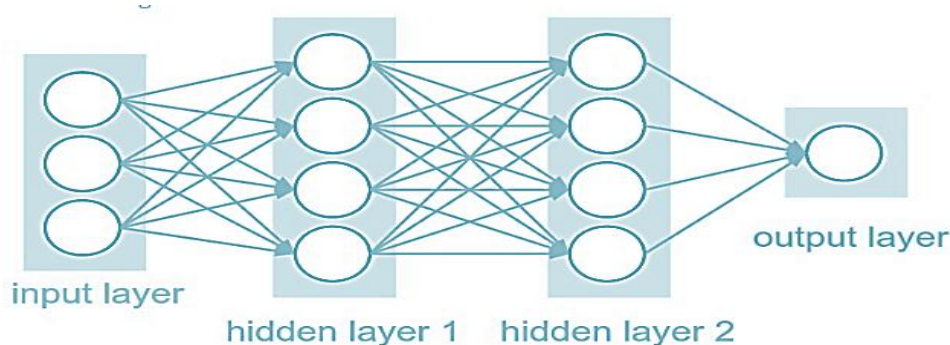


Figure 2: The neural network diagram.

### *Evaluation indicators Formulaic description*

$$\text{Accuracy} = (\text{TP} + \text{TN}) / (\text{TP} + \text{FN} + \text{FP} + \text{TN})$$

$$\text{Precision} = \text{TP} / (\text{TP} + \text{FP})$$

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

$$\text{F1-score} = (2 \times \text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$$

TP, True positive. TN, True negative. FN, False negative. FP, False positive.

Additionally, the accuracy of the models is also evaluated using a confusion matrix.

## Results

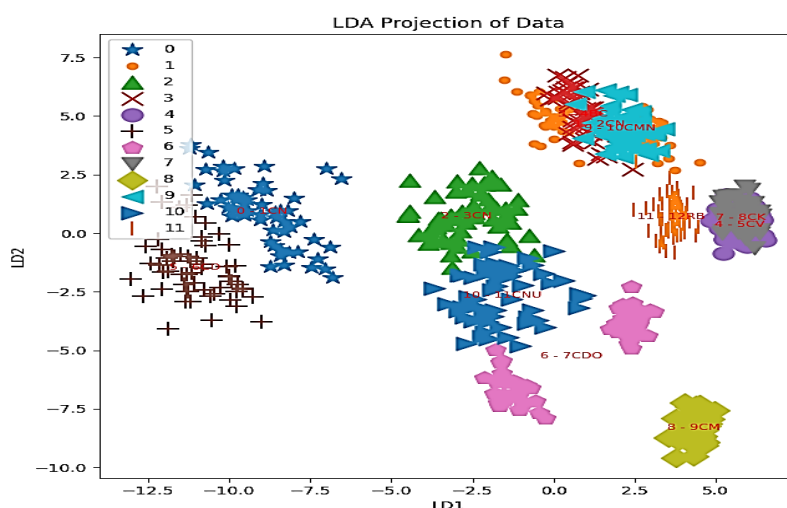
### *Evaluation of BDP input data*

The LDA scatter plot in Figure 3 illustrates the input data for AI based on the BDP index. This input data was generated by measuring 720 samples of 12 species. The results show that the data is clustered into different groups, with several clusters being very clearly classified such as species 0, 5, 8, and 11, some relatively clear like species 2, 6, and 10, and some overlapping and potentially difficult to classify such as species 1, 3, 4, 7, and 9. Thus, the input data shows that 33% is clearly grouped, 25%

relatively clear, and 41.7% overlapping and difficult to group. From this data, we will conduct experiments with 6 machine learning models and 3 deep learning models to train the data and evaluate which model is suitable for classifying the marine fish species collected in this area.

### *Classification performance using machine learning based on BDP*

The classification performance results using machine learning based on BDP are shown in Figure 4, specifically for each model as follows: In analyzing the performance of various classifiers on the dataset based on BDP, as depicted in Table 2, the following results were observed. The RFC achieved the highest accuracy of 86.81%, demonstrating robust precision (86.91%) and recall (86.81%), resulting in a balanced F1-score of 86.70%. Similarly, the BaC exhibited competitive performance with an accuracy of 85.42%, precision of 85.46%, recall of 85.42%, and an F1-score of 85.17%. The ETC and GBC closely followed with accuracies of 84.72%.



**Figure 3:** The LDA scatter plot shows the clustering of input data for the AI models based on BDP, with a cumulative explained variance percentage of 83.9%.

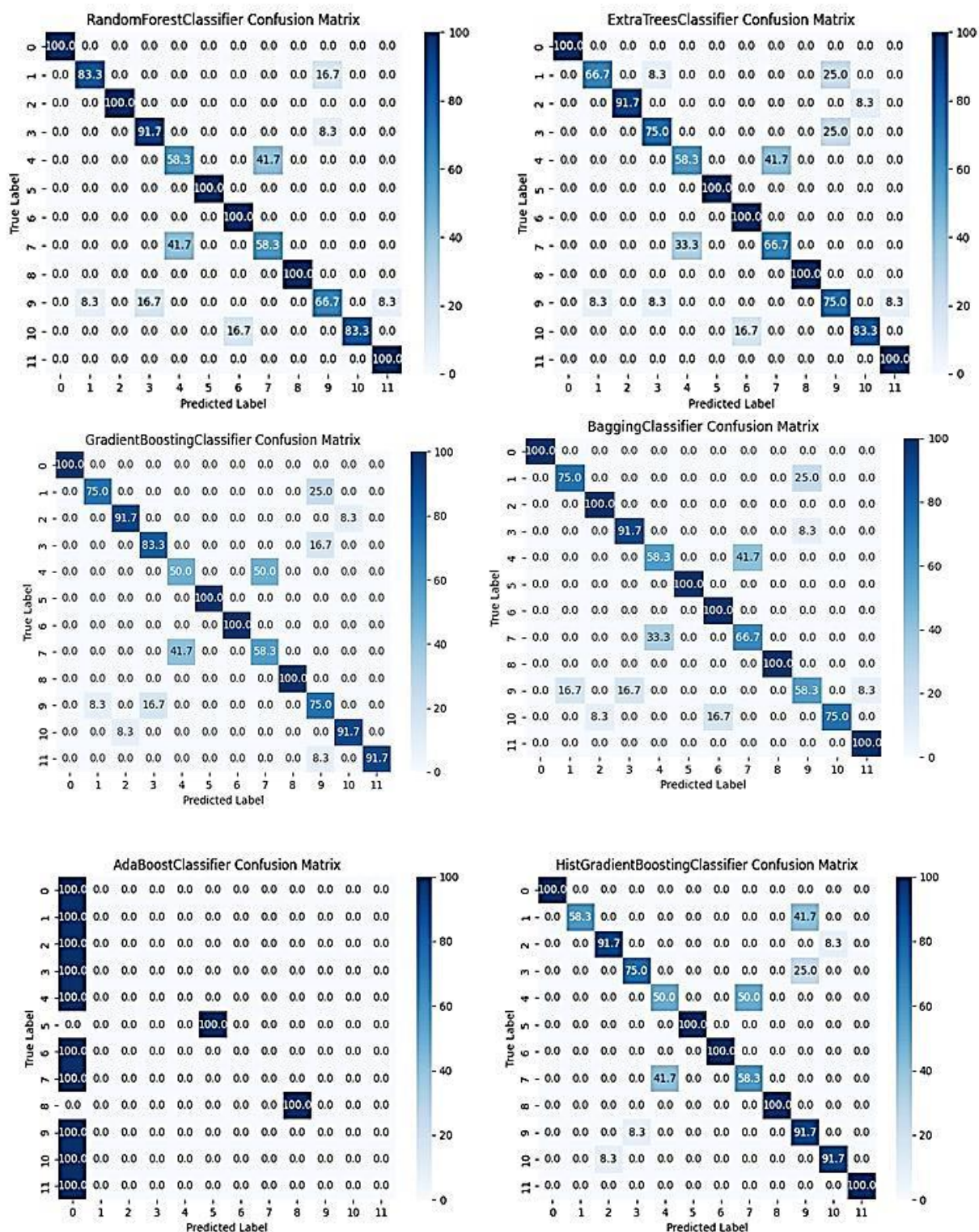


Figure 4: Heatmap of the confusion matrix for machine learning models based on the BDP. The y-axis represents the true species names, while the x-axis represents the predicted results.



**Table 2: Classification performance using machine learning models based on the BDP according to four evaluation metrics (Accuracy, Precision, Recall, F1-score).**

Classifier	Index assesses accuracy level			
	Accuracy	Precision	Recall	F1-score
RFC	86.81%	86.91%	86.81%	86.70%
ETC	84.72%	85.40%	84.72%	84.74%
GBC	84.72%	85.42%	84.72%	84.91%
BaC	85.42%	85.46%	85.42%	85.17%
ABC	25.00%	17.50%	25.00%	18.18%
HGBC	84.72%	86.63%	84.72%	84.83%

**Notes:** Random Forest Classifier (RFC), Extra Trees Classifier (ETC), Gradient Boosting Classifier (GBC), Bagging Classifier (BaC), AdaBoost Classifier (ABC), Hist Gradient Boosting (HGBC).

While the ETC demonstrated slightly lower precision (85.40%) and F1-score (84.74%) compared to the GBC (85.42% precision and 84.91% F1-score), both classifiers maintained consistent recall rates of 84.72%. The HGBC also achieved an accuracy of 84.72%, with higher precision (86.63%) and a well-rounded F1-score of 84.83%. In contrast, the ABC significantly underperformed relative to the others, attaining only 25% accuracy. Its precision was notably low (17.50%), coupled with a similarly low recall of 25.00% and an F1-score of 18.18%. Overall, the RFC and BaC demonstrated the most reliable performance, while the ABC struggled significantly.

#### *The predictive accuracy using machine learning based on the BDP*

The classification results using the RFC, as shown in Figure 4, indicate that 9 species were identified with over 80% accuracy, with 6 species achieving 100% accuracy (species 0, 2, 5, 6, 8, 11). However, species 4 and 7 were often confused with each other, showing a low identification accuracy of 53.3% and a confusion level of 47.7%. The ETC revealed that 7 species were identified with more than 80%

accuracy, and 5 species achieved perfect identification (species 0, 5, 6, 8, 11), while species 4 and 7 exhibited high confusion levels at 41.7%, with a corresponding identification accuracy of 53.3%. The GBC results showed that 8 species had an identification accuracy above 80%, with 4 species perfectly classified (species 0, 5, 6, 8). Two species had a moderate identification accuracy of 75% (species 9, 11), while species 4 and 7 had a high confusion rate of 50%, resulting in a low identification accuracy of 50%. The BaC accurately identified 7 species with more than 90% accuracy, and 6 species were perfectly classified (species 0, 2, 5, 6, 8, 11). Two species were moderately classified with 75% accuracy (species 1, 10), while species 4, 7, and 9 had lower accuracy, ranging from 50% to 67%. The ABC results show that only 2 species were classified with 100% accuracy, with the remaining species mostly misclassified. Lastly, the HGBC identified 8 species with over 90% accuracy, with 5 species perfectly classified (species 0, 5, 6, 8, 11). One species had a moderate classification accuracy of 75% (species 3), whereas species 1, 4, and 7 had lower accuracy, ranging from 50% to 60%. These results

illustrate the performance of various classifiers as depicted in Figure 4, with varying degrees of accuracy and confusion among the species.

The findings regarding the accuracy levels of the deep learning models based on BDP are depicted in Figure 5. Model 1 attained the highest accuracy at 77.08%, complemented by robust precision (83.93%) and recall (77.08%), resulting in a harmonious F1-score of 74.11%. Similarly, Model 3 demonstrated competitive performance, achieving an accuracy of 75.69%, a precision of 76.69%, and a comprehensive F1-score of 75.32%.

However, Model 2 displayed inferior performance, with an accuracy of only 70.14% and the lowest F1-score recorded at 68.75%. Overall, Models 1 and 3 exhibited the most robust performance, while Model 2 lagged significantly behind (Fig. 5).

*Predictive accuracy using deep learning based on the BDP*

According to the confusion matrix for deep learning models based on BDP in Figure 6, several key observations can be made. Using the Dense 64 model, 8 species were identified with 100% accuracy (species: 2, 4, 5, 6, 8, 9, 10, 11), while 4 species had low classification accuracy ranging from 16.7% to 66.7%.

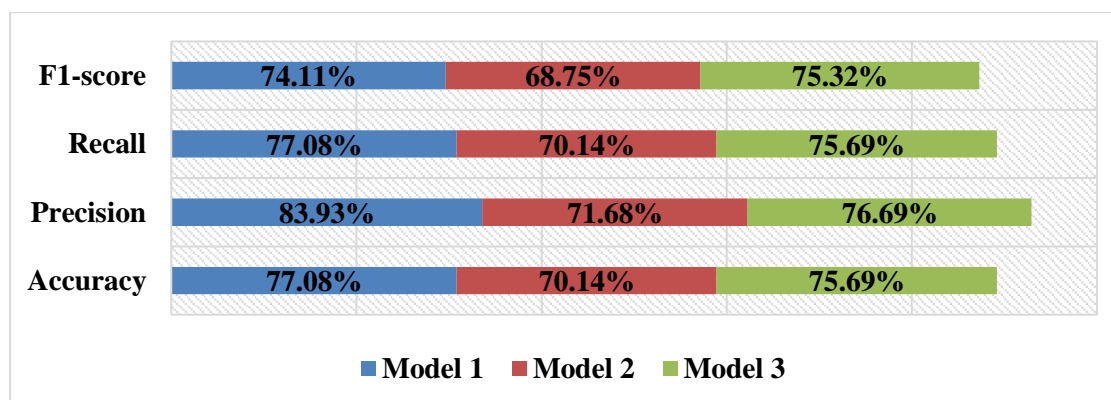
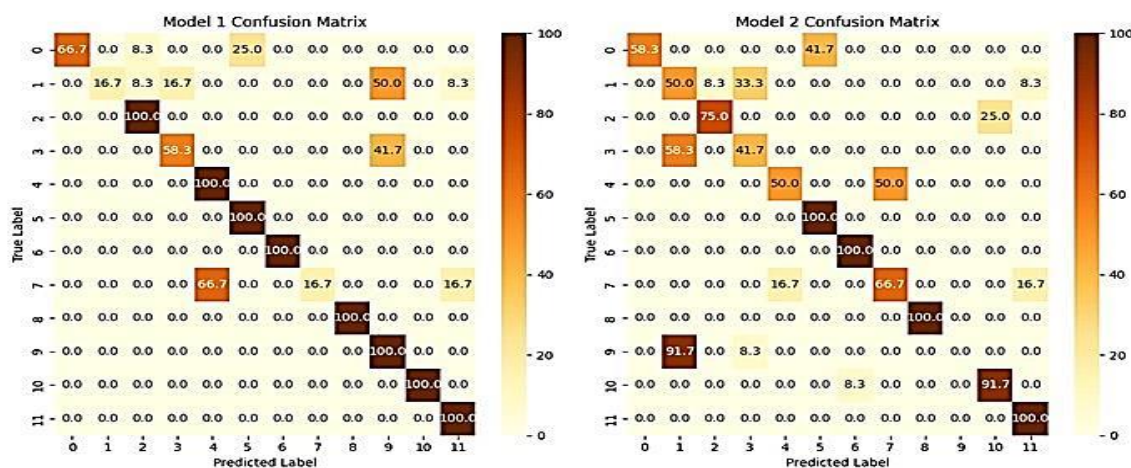


Figure 5: Classification performance using deep learning models based on the BDP according to four evaluation metrics (accuracy, precision, recall, F1-score).



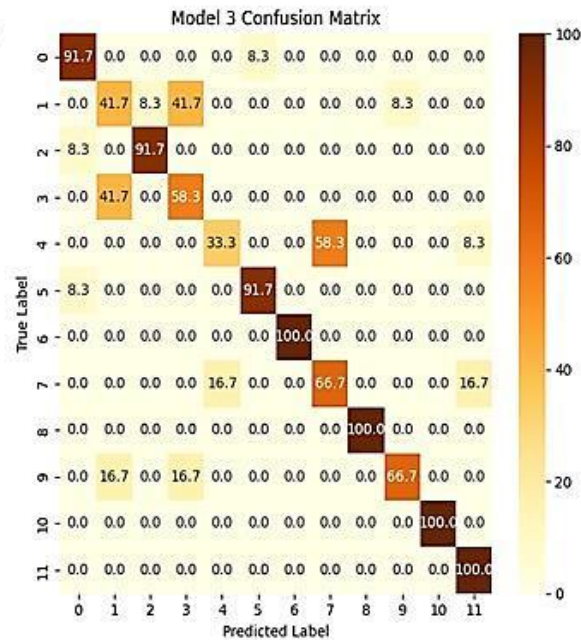
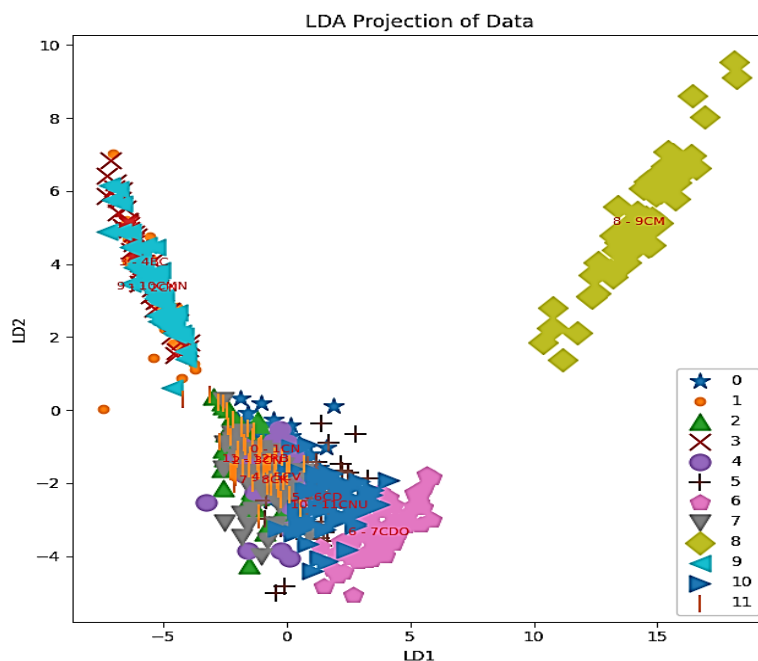


Figure 6: Heatmap of the confusion matrix for deep learning models based on BDP. The y-axis represents the true species names, while the x-axis represents the predicted results.

The Dense 128 model showed that 5 species were identified with over 90% accuracy, with 4 species achieving perfect accuracy (species: 5, 6, 8, 11) and 1 species achieving over 90% accuracy (species 10). Additionally, species 2 had a moderate classification accuracy of 75%, whereas the remaining 6 species had a low accuracy below 70%, with species 9 being completely misclassified as species 1 (0% classification rate). Finally, the Dense 256 model demonstrated that 7 species were identified with over 90% accuracy, with 4 species achieving perfect accuracy (species: 6, 8, 10, 11) and 3 species achieving over 90% accuracy (species: 0, 2, 5). All other species (species: 1, 3, 4, 7, 9) had low accuracy below 70%.

#### *Evaluation of input data for the ShI*

The LDA scatter plot illustrates the input data for machine learning based on the ShI. This input data was generated by calculating the ShI for 720 samples from 12 species (Fig. 7). Unlike the LDA analysis results based on BDP, only one cluster (species number 8) is clearly distinguished, while the other species overlap. Specifically, species 1, 9, and 3 form one group, and species 0, 2, 4, 5, 6, 7, 10, and 11 form another group. The machine learning results using 6 machine learning models and 3 deep learning models to train the data are presented below.



**Figure 7:** The LDA scatter plot shows the clustering of input data for the AI models based on ShI, with a cumulative explained variance of 91.5%.

#### *Classification performance using machine learning based on the ShI*

The findings regarding the accuracy levels of the ML models based on ShI are presented in Table 3. The RFC achieved an accuracy of 53.47%, accompanied by a precision of 55.60%, a recall of 53.47%, and an F1-score of 53.70%. The HGBC attained an accuracy of 52.08%, a precision of 56.61%, a recall of 52.08%, and an F1-score of 52.31%. The BaC and ETC

exhibited accuracies of 50.69% and 50.00%, respectively, along with similar precision, recall, and F1-scores. The GBC demonstrated a comparable accuracy of 49.31%, a precision of 51.46%, a recall of 49.31%, and an F1-score of 49.62%. The ABC notably underperformed with an accuracy of only 23.61%, precision of 12.05%, recall of 23.61%, and an F1-score of 14.04%.

**Table 3:** Classification performance using machine learning models based on the ShI according to four evaluation metrics (accuracy, precision, recall, F1-score).

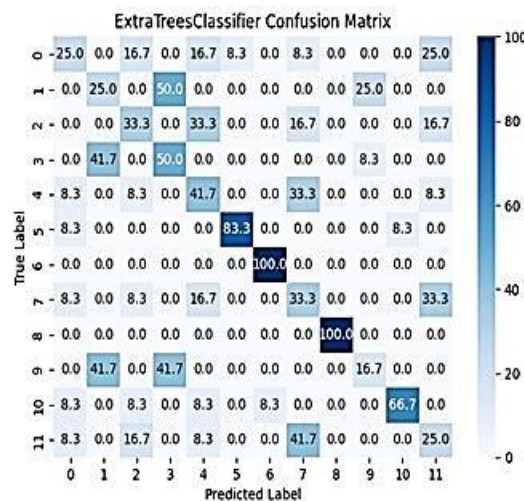
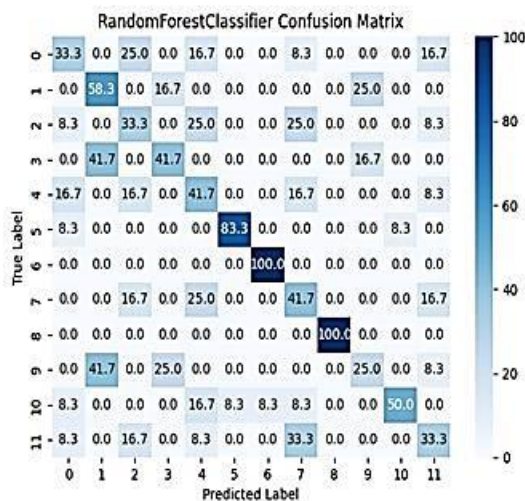
Classifier	Index assesses accuracy level			
	Accuracy	Precision	Recall	F1-score
RFC	53.47%	55.60%	53.47%	53.70%
ETC	50.00%	51.59%	50.00%	50.09%
GBC	49.31%	51.46%	49.31%	49.62%
BaC	50.69%	52.09%	50.69%	50.62%
ABC	23.61%	12.05%	23.61%	14.04%
HGBC	52.08%	56.61%	52.08%	52.31%

**Notes:** Random Forest Classifier (RFC), Extra Trees Classifier (ETC), Gradient Boosting Classifier (GBC), Bagging Classifier (BaC), AdaBoost Classifier (ABC), Hist Gradient Boosting (HGBC).

*The predictive accuracy using machine learning based on ShI*

The classification results for the RFC reveal that three species were accurately identified at a rate exceeding 80%, with two species achieving perfect accuracy of 100% (species 6 and 8). Conversely, the remaining species exhibited notably lower identification rates, falling below 60% (species 0, 1, 2, 3, 4, 7, 9, 10, 11). Similarly, the ETC revealed that three species were correctly classified at an accuracy surpassing 80%, with two species achieving flawless 100% accuracy (species 6 and 8), while the remaining species displayed inferior identification rates below 60% (species 0, 1, 2, 3, 4, 7, 9, 10, 11). The GBC, as shown in Figure 8, showed that two species were accurately identified at a perfect 100% rate (species 6 and 8), with one species attaining a moderate accuracy level of 75% (species 5), while the

remaining species exhibited suboptimal identification rates below 60% (species 0, 1, 2, 3, 4, 7, 9, 10, 11). The BaC indicated that three species were accurately classified at a rate exceeding 80%, with two species achieving a pristine accuracy of 100% (species 6 and 8), and the remaining species displaying notably lower identification rates falling below 60% (species 0, 1, 2, 3, 4, 7, 9, 10, 11). The ABC revealed the accurate identification of only two species at a perfect rate of 100% (species 0 and 8), with one species achieving a classification rate of 66.7% (species 9), while the remaining species were entirely misclassified (species 1, 2, 3, 4, 5, 6, 7, 10, 11). Lastly, the HGBC (Fig. 8) indicated the accurate identification of two species at a perfect rate of 100% (species 6 and 8), with one species achieving a moderate accuracy level of 75% (species 5), while the remaining species, totaling nine, displayed subpar identification rates below 60% (species 0, 1, 2, 3, 4, 7, 9, 10, 11).



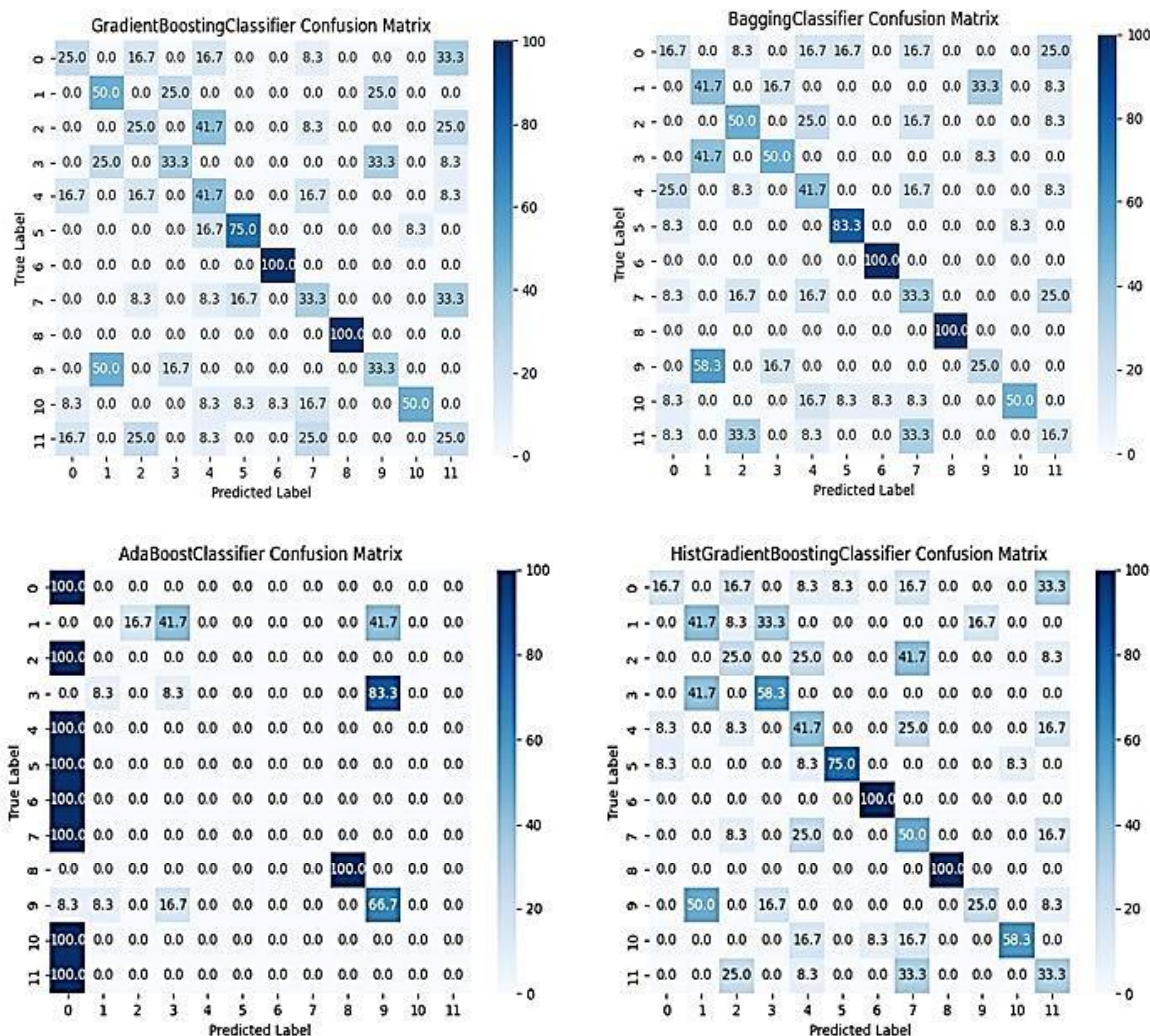
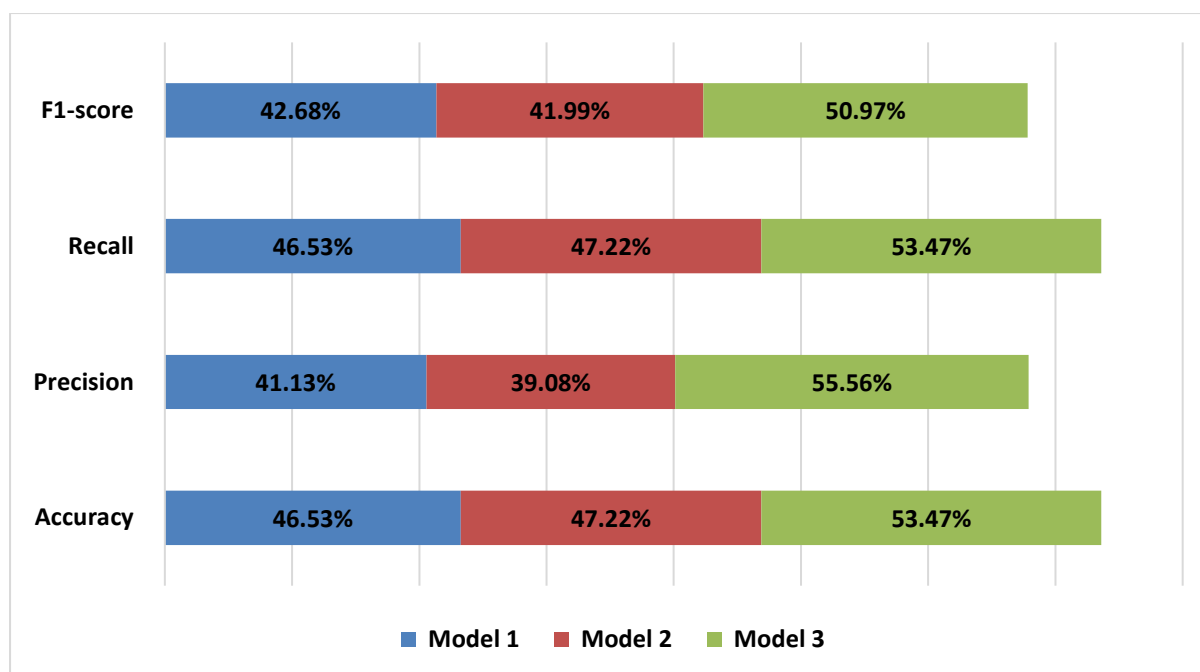


Figure 8: Heatmap of the confusion matrix for machine learning based on the ShI. The y-axis represents the true species names, while the x-axis represents the predicted results.

*Classification performance using deep learning based on the ShI*

The findings regarding the accuracy levels of the deep learning models based on ShI are depicted in Figure 9. Model 3 indicated the highest accuracy at 53.47%, accompanied by a precision of 55.56%, recall of 53.47%, and an F1-score of 50.97%. Despite its relatively higher accuracy compared to other models, Model 3 still struggles to deliver consistently accurate classifications, as indicated by its F1-score. Model 2 showed the next best

performance, achieving an accuracy of 47.22% with a precision of 39.08%, a recall of 47.22%, and an F1-score of 41.99%. Although Model 2 slightly surpasses Model 1 in terms of accuracy, its precision is the lowest among the three, indicating a higher propensity for false positives. Model 1 exhibited the lowest accuracy at 46.53%, with a precision of 41.13%, recall of 46.53%, and an F1-score of 42.68%. This model indicates a marginally better precision than Model 2, but overall its predictive performance remains weak (Fig. 9).



**Figure 9:** Classification performance using deep learning models based on the ShI according to four evaluation metrics (accuracy, precision, recall, F1-score).

#### *The predictive accuracy using deep learning models based on the ShI*

The classification results using deep learning models, as shown in Figure 10, exhibit varying levels of accuracy across different species. For Model 1 (Dense 64), it demonstrates the ability to classify 2 species (species 6, 8) with 100% accuracy, while achieving a moderate level of 75% accuracy for 2 species (species 9, 10). The remaining 8 species (species 0, 1, 2, 3, 4, 5, 7, 11) exhibit low classification accuracy, with 2 species (species 1, 2) being unclassifiable (0%). In the case of Model 2 (Dense 128), it also classifies 2 species

(species 6, 8) with 100% accuracy and achieves 75% accuracy for 1 species (species 5). However, 9 species (species 0, 1, 2, 3, 4, 7, 9, 10, 11) show low classification accuracy below 70%, with 3 species (species 2, 4, 9) being unclassifiable (0%). Finally, Model 3 (Dense 256) indicates the ability to classify 1 species (species 8) with 100% accuracy, over 90% accuracy for 1 species (species 6), and a moderate 75% accuracy for 1 species (species 10). The remaining 9 species (species 0, 1, 2, 3, 4, 5, 7, 9, 11) exhibit low classification accuracy below 70%, with 2 species (species 2, 4) being unclassifiable (0%).

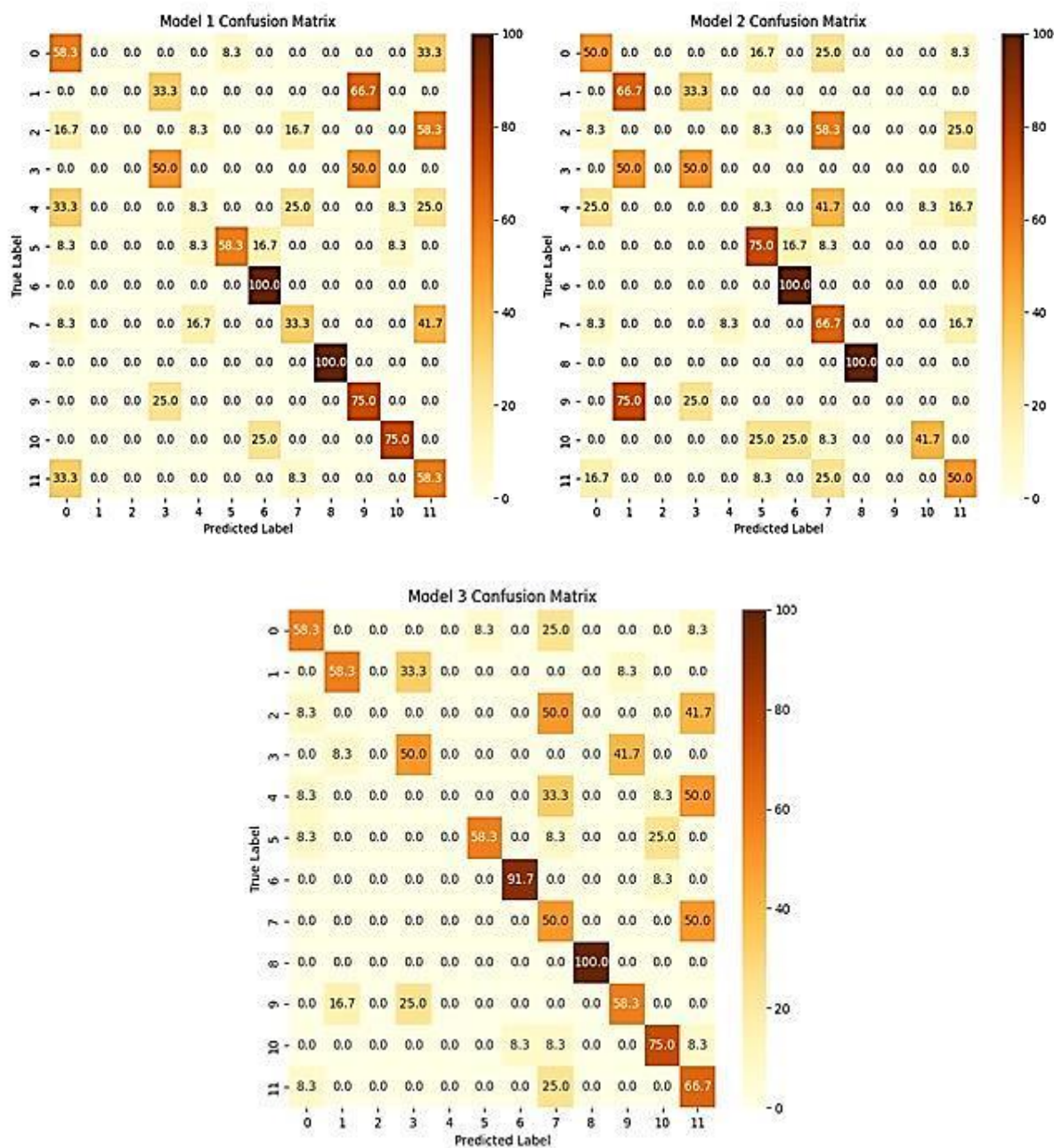


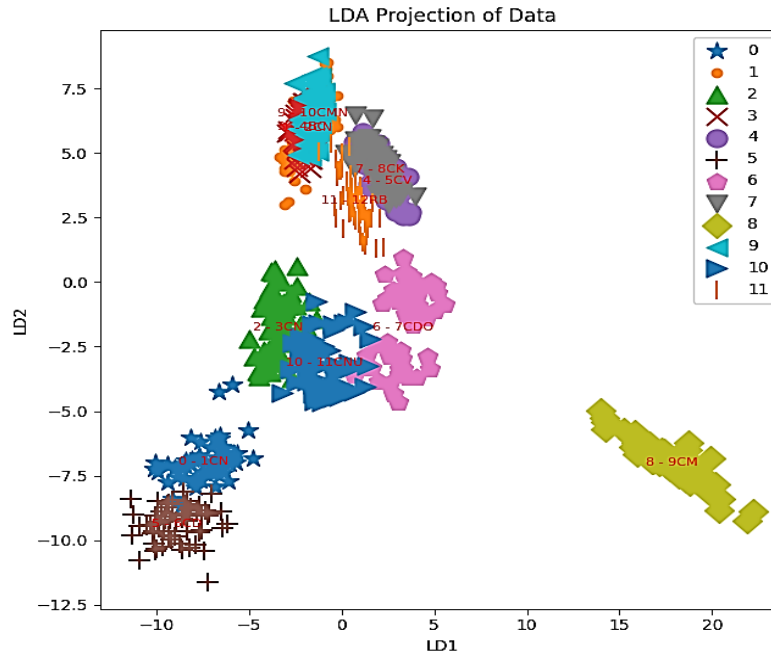
Figure 10: Heatmap of the confusion matrix for deep learning models based on ShI. The y-axis represents the true species names, while the x-axis represents the predicted results.

*Evaluation of input data for the BDP-ShI dataset*

The LDA scatter plot illustrates that the input data is a complex combination of ShI-BDP for AI models (Fig. 11). This input data was generated by merging the BDP and ShI metrics of 720 samples belonging to 12 species. The results indicate that the

data is grouped into distinct clusters, with several species being clearly categorized, such as species 0, 2, 5, 6, and 8. Some species, like 10 and 11, are relatively well-defined, while others, such as 7, 4, 1, 9, and 3, exhibit overlapping and may be challenging to classify.





**Figure 11:** The LDA scatter plot shows the clustering of input data for the AI models based on BDP-ShI.

Thus, we can observe that the input data consists of 41% clearly defined clusters, 18% relatively well-defined clusters, and 41% overlapping and challenging-to-classify clusters.

*Classification performance using machine learning based on ShI-BDP*

The classifiers' performances using machine learning models based on BDP-ShI show several key observations as presented in Table 4. The BaC achieved the

highest accuracy of 86.11%, accompanied by a precision of 86.81%, recall of 86.11%, and a matching F1-score of 86.11%. The RFC and GBC followed closely, with the RFC achieving an accuracy of 84.03% (precision: 84.77%, recall: 84.03%, F1-score: 84.13%) and the GBC reaching an accuracy of 84.72% with balanced precision, recall, and F1-scores around 84.59%.

**Table 4:** Classification performance using machine learning models based on the BDP-ShI according to four evaluation metrics (accuracy, precision, recall, F1-score).

Classifier	Index assesses accuracy level			
	Accuracy	Precision	Recall	F1-score
RFC	84.03%	84.77%	84.03%	84.13%
ETC	81.25%	81.83%	81.25%	81.18%
GBC	84.72%	84.72%	84.72%	84.59%
BaC	86.11%	86.81%	86.11%	86.11%
ABC	25.00%	17.50%	25%	18.18%
HGBC	82.64%	83.89%	82.64%	82.64%

*Notes:* Random Forest Classifier (RFC), Extra Trees Classifier (ETC), Gradient Boosting Classifier (GBC), Bagging Classifier (BaC), AdaBoost Classifier (ABC), Hist Gradient Boosting (HGBC).

The HGBC and ETC delivered moderate results, with the HGBC showing an

accuracy of 82.64% (precision: 83.89%, recall: 82.64%, F1-score: 82.64%) and the

ETC achieving an accuracy of 81.25% (precision: 81.83%, recall: 81.25%, F1-score: 81.18%). In contrast, the ABC significantly underperformed with an accuracy of only 25%, precision of 17.5%, recall of 25%, and an F1-score of 18.18%.

#### *The predictive accuracy using machine learning based on BDP- ShI*

The classification results using various machine learning models based on BDP-ShI indicate differing levels of accuracy (Fig. 12). The RFC accurately identified 7 species with over 80% accuracy, with 5 species achieving perfect accuracy of 100% (species 0, 5, 6, 8, 11), 2 species with moderate accuracy at 75% (species 1, 3), and 3 species with low accuracy below 70% (species 4, 7, 9). The ETC showed that 7 species were identified with over 80% accuracy, with 6 species at 100% accuracy (species 0, 2, 5, 6, 8, 11), while the remaining species had accuracy below 70% (species 1, 3, 4, 7, 9). The GBC identified 9 species with over 80% accuracy, including 6 species at 100% (species 0, 2, 5, 6, 8, 11), 2 species at 75% (species 1, 3), and 3 species below 70% (species 4, 7, 9). The BaC also identified 9 species with over 80% accuracy, with 5 species at 100% (species 0, 5, 6, 8, 11), 1 species at 75% (species 7), and 2 species with accuracy between 50% and 68% (species 1, 4). The ABC identified only 2 species with 100% accuracy (species 5, 8), while the remaining species were misclassified. The HGBC showed that 8 species had over 80% accuracy, with 5 species at 100% (species 0, 5, 6, 8, 11), and 4 species below 70% (species 1, 3, 4, 7).

#### *Classification performance using deep learning based on BDP- ShI*

Figure 13 illustrates the results concerning the accuracy levels of the deep learning models derived from BDP-ShI. Model 1 demonstrated the highest accuracy at 77.08%, coupled with a precision of 83.93%, recall of 77.08%, and an F1-score of 74.11%. Model 3 followed closely, achieving an accuracy of 75.69% with a precision of 76.69%, a recall of 75.69%, and an F1-score of 75.32%. Model 2 exhibited the lowest accuracy at 70.14%.

#### *Accuracy levels using deep learning based on BDP-ShI*

The classification results, as shown in Figure 14, demonstrate the performance of three deep learning models. Model 1 (Dense 64) shows the ability to classify 8 species with an accuracy of over 90%, including 7 species with perfect classification at 100% accuracy (species 2, 4, 5, 6, 8, 10, 11) and 1 species at a moderate accuracy of 75% (species 9). The remaining 3 species exhibit classification accuracies below 70% (species 0, 1, 7). Model 2 (Dense 128) is able to classify 6 species with 100% accuracy (species 5, 6, 7, 8, 10, 11), with 1 species at a moderate accuracy of 75% (species 2). The other species show accuracies below 70% (species 0, 1, 3, 4, 9), with species 4 and 9 being highly confused with each other. Model 3 (Dense 256) successfully classifies 6 species with 100% accuracy (species 0, 2, 6, 8, 10, 11) and 1 species at an accuracy of 83.3% (species 5).

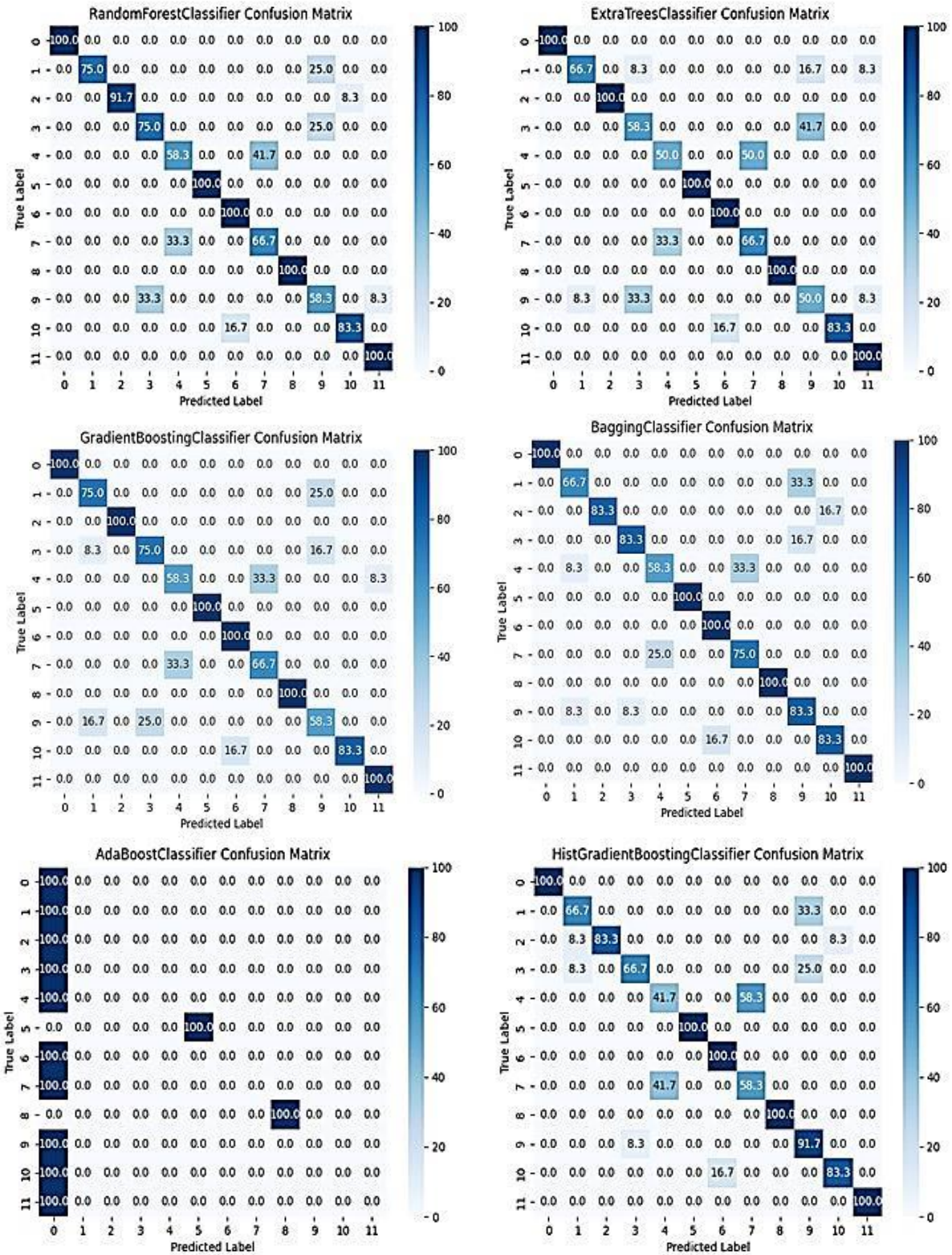


Figure 12: Heatmap of the confusion matrix for machine learning based on the BDP- ShI. The y-axis represents the true species names, while the x-axis represents the predicted results.

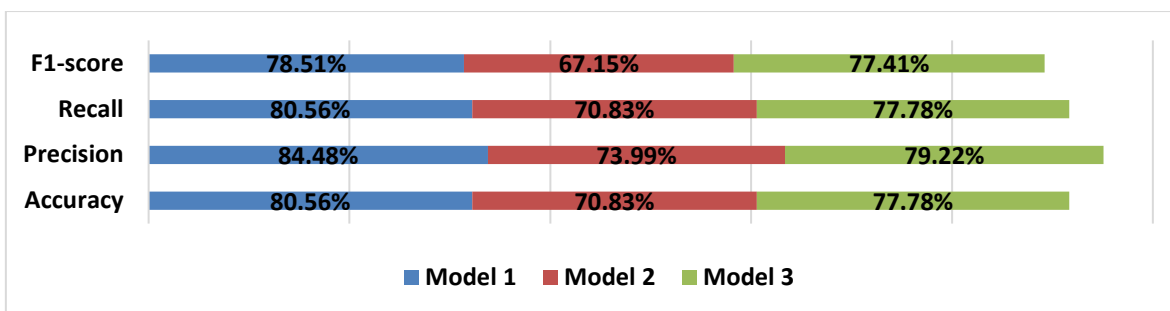


Figure 13: Classification performance using deep learning models based on the BDP-ShI according to four evaluation metrics (accuracy, precision, recall, F1-score).

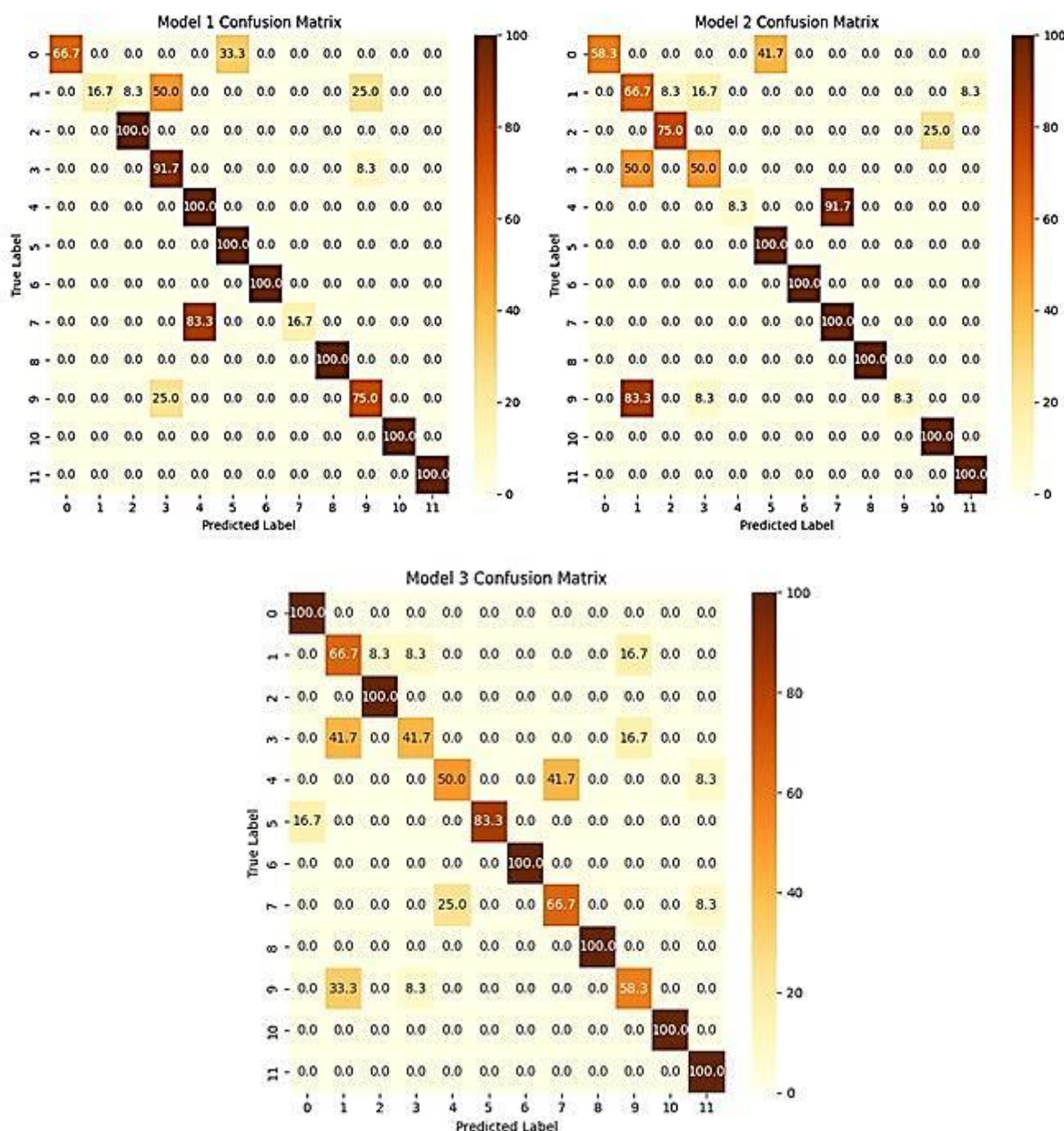


Figure 14: Heatmap of the confusion matrix for deep learning models based on BDP-ShI. The y-axis represents the true species names, while the x-axis represents the predicted results.

The remaining 5 species display low classification accuracies below 70% (species 1, 3, 4, 7, 9).

### Discussion

In recent studies, AI methods, have demonstrated significant improvements in accuracy and reliability over traditional methods in the field of otolith shape analysis. AI-based approaches, particularly convolutional neural networks (CNNs), have been shown to reduce error rates substantially. According to a study by Liu *et al.* (2021), deep learning models achieved an accuracy of 90.5% in identifying fish species based on otolith images, compared to a maximum of 75.3% accuracy reported with traditional geometric methods (Liu *et al.*, 2021). Furthermore, traditional methods have been reported to have error rates as high as 25% due to variability in manual measurements, whereas AI methods can significantly minimize these errors by automating and standardizing the process (Chen *et al.*, 2022). This evidence highlights the enhanced reliability and reduced error rates of AI techniques compared to traditional methods in otolith shape analysis.

This study evaluates various machine learning and deep learning models for species identification using otolith images, focusing on accuracy, precision, recall, and F1-scores. Models with over 80% accuracy are classified as high-performing, while those exceeding 75% are seen as having practical potential. The analysis highlights significant differences in model performance, with some excelling in accuracy and others showing limitations.

Evaluation based on the BDP dataset demonstrates that species identification using five models—RFC, ETC, GBC, BaC, and HGBC—achieved overall recognition accuracy above 80%. These results surpass the classification accuracy obtained with LDA on the input data. Among these, the RFC model yielded the highest accuracy of 86.81%, identifying 9 out of 12 species with over 80% accuracy. The RFC's precision of 86.91% suggests a reliable performance in minimizing false positives. Its recall rate matches the accuracy, indicating the model's strong ability to identify a substantial proportion of positive cases. With an F1-score of 86.70%, the RFC demonstrates a balanced trade-off between precision and recall. In contrast, the ABC model exhibits the lowest accuracy at 25.00%, indicating significant difficulties in accurate classification. Its precision of 17.50% may lead to a high number of false positives, and despite the recall rate matching its accuracy, it falls far short compared to the other models. The ABC's F1-score of 18.18% underscores its ineffectiveness in balancing recall and precision. The BaC model, with an accuracy of 85.42%, performs well, though not as strongly as the RFC. Its precision, at 85.46%, indicates a moderate rate of correctly predicting positive instances. The recall rate mirrors the accuracy, confirming its efficiency in identifying positive cases. With an F1-score of 85.17%, the BaC demonstrates a good balance between precision and recall. These results are similar to the findings of Yuwen Chen and Guoping Zhu (2023), who used various machine learning models to identify four species—*Chionodraco rastrospinosus*,

*Krefflichthys anderssoni*, *Electrona carlsbergi*, and *Pleuragramma antarcticum*. In their study, the RFC model also yielded the best performance, while the ABC model had the lowest performance. All models showed high confusion between species 4 and species 7. Although the RFC model provided the best overall results, it performed worse than other models for species 9, suggesting that different models might be more suitable for specific cases. The RFC emerged as the most robust model, showing superior performance across all metrics, with the BaC following closely behind, performing consistently.

The five models—RFC, ETC, GBC, BaC, and HGBC—used for species identification based on ShI all resulted in overall recognition accuracy below 55%, with low accuracy for identifying individual species. The accuracy for identifying individual species was also low. These results are not better than those obtained with LDA on the input data. The RFC model achieved the highest overall accuracy at 53.47%, while the remaining models had accuracies ranging from 49.32% to 52%. All five models exhibited high confusion, with only 3 species (5, 6, 8) achieving classification levels between 75% and 100%. These results indicate that using ShI alone for species identification is not suitable.

Evaluation based on the combined BDP-ShI metric demonstrates that species identification using the five models—RFC, ETC, GBC, BaC, and HGBC—resulted in overall recognition accuracy above 80%. These recognition results are better than those obtained with LDA on the input data. Among these models, the BaC achieved the

highest accuracy of 86.11%, identifying 9 out of 12 species with over 80% accuracy. All models exhibited high confusion with species 4. The BaC's highest accuracy of 86.11% indicates exceptional overall performance in predicting correct classes. Its precision, at 86.81%, highlights strong reliability in minimizing false positives. The recall rate matches the accuracy, showcasing the model's ability to identify a significant proportion of positive cases. With an F1-score of 86.11%, the BaC strikes an effective balance between precision and recall. On the other hand, the ABC exhibits the lowest accuracy at 25%, implying substantial difficulties in classifying instances accurately. Its precision, at 17.5%, could result in a high number of false positives. While the recall rate is consistent with its accuracy, it falls considerably short compared to the other models. With an F1-score of 18.18%, the ABC struggles to balance recall and precision. The RFC, with an accuracy of 84.03%, performs well but not as strongly as the BaC. Its precision of 84.77% indicates a satisfactory rate of correctly predicting positive instances. The recall rate matches the accuracy, affirming the model's efficiency in identifying positive instances. With an F1-score of 84.13%, the RFC demonstrates a good balance between precision and recall. Similarly, the GBC also delivers solid performance, with an accuracy of 84.72%. Its precision and recall rates are both 84.72%, resulting in an F1-score of 84.59%. The ETC and HGBC exhibit competitive, yet slightly lower, accuracies at 81.25% and 82.64%, respectively. The ETC maintains consistent precision at 81.83% and an F1-score of

81.18%. The HGBC achieves a higher precision of 83.89%, leading to a balanced F1-score of 82.64%.

The use of deep learning shows that species identification using three deep learning models—Model 1 (Dense 64), Model 2 (Dense 128), and Model 3 (Dense 256)—all resulted in overall recognition accuracy above 70%. These results are better than those obtained with LDA on the input data. Among these, Model 1 achieved the highest accuracy of 77.08% for all cases, accurately identifying 8 out of 12 species with 100% accuracy. Model 1 achieved the highest accuracy of 77.08%, indicating good overall performance in predicting correct classes. Its precision, at 83.93%, suggests reliability in avoiding false positives, which is crucial for certain applications. The recall rate matches the accuracy, indicating the model's capability to identify a substantial proportion of positive cases. With an F1-score of 74.11%, Model 1 demonstrates a balanced trade-off between precision and recall. In contrast, Model 2 exhibits the lowest accuracy at 70.14%, implying potential struggles in classifying instances accurately. Its precision, at 71.68%, may lead to a higher number of false positives. While the recall rate matches its accuracy, it falls short compared to Model 1. With an F1-score of 68.75%, Model 2 appears less effective in balancing recall and precision. Model 3, with an accuracy of 75.69%, performs reasonably well, albeit not as strongly as Model 1. Its precision, at 76.69%, indicates a moderate rate of correctly predicting positive instances. The recall rate mirrors the accuracy, confirming its efficiency in identifying positive instances. With an F1-

score of 75.32%, Model 3 demonstrates a good balance between precision and recall. In conclusion, Model 1 is the most effective, showing the best performance across all metrics.

Evaluation based on the ShI metric reveals that species identification using three models—Model 1 (Dense 64), Model 2 (Dense 128), and Model 3 (Dense 256)—all resulted in overall recognition accuracy below 50%. Among these, Model 3 achieved the highest accuracy, but only reached 53.47% for all cases. These results indicate that using deep learning based solely on the ShI metric for species identification is not suitable.

Evaluation based on the combined BDP-ShI metric demonstrates that species identification using three models—Model 1 (Dense 64), Model 2 (Dense 128), and Model 3 (Dense 256)—all resulted in overall recognition accuracy above 70%. These results are better than those obtained with LDA on the input data. Among these, Model 1 achieved the highest accuracy of 80.56% for all cases, accurately identifying 8 out of 12 species with over 90% accuracy, and 7 out of 12 species with 100% accuracy. Model 1 Performance: Model 1 achieved an accuracy of 80.56%, indicating a reliable performance in correctly classifying most instances. With a precision of 84.48%, Model 1 demonstrates a strong ability to make positive predictions accurately, minimizing false positives. The recall is consistent with the accuracy at 80.56%, highlighting the model's strong ability to identify the most positive instances. The F1-score is 78.51%, indicating a good balance between precision and recall. Model 2 Performance:

Model 2 achieved a lower accuracy of 70.83%, suggesting that it struggles to classify instances as effectively as Model 1. At 73.99%, Model 2 maintains a relatively good level of precision, although lower than Model 1, indicating a higher number of false positives. The recall is consistent with the accuracy at 70.83%, showing that the model can identify a reasonable proportion of positive cases but not as effectively as Model 1. The F1-score is 67.15%, suggesting a less favorable trade-off between precision and recall compared to Model 1. Model 3 Performance: Model 3 achieved a higher accuracy of 77.78%, which is better than Model 2 but still lower than Model 1. Precision stands at 79.22%, indicating that Model 3 can identify positive instances with relatively high accuracy while maintaining a lower false-positive rate. The recall matches the accuracy at 77.78%, indicating that Model 3 identifies a substantial proportion of positive cases. The F1-score is 77.41%, showing a well-balanced trade-off between precision and recall. Model 1 clearly outperforms the other models, achieving the highest accuracy and precision while maintaining strong recall and F1-scores. Model 2 exhibits the lowest performance across all metrics, indicating that it requires

significant improvements to be practical. Model 3 delivers a balanced performance, although not as strong as Model 1, with consistent precision, recall, and F1-scores. Overall, the performance results of the three deep learning models are quite similar to the findings of Salimi (2016), who evaluated the performance of various machine learning models in identifying 14 fish species based on otolith contours, with performance results ranging from 62% to 94%.

The training times for different machine learning models ranged from 0.1 to 4.3 seconds, with GBC taking the longest at 4.3 seconds for the BDP-ShI complex. The other models (HGBC, ABC, BaC, ETC, RFC) had training times of less than 0.3 seconds. The training times for the three deep learning models ranged from 4.2 to 5.6 seconds (Fig. 15). Thus, the training times for the deep learning models did not differ significantly. In this study, analyzing 720 otolith samples, the training times for all models were under 6 seconds. This short training duration does not significantly impact the choice of the optimal model for practical application (this may change with larger datasets).

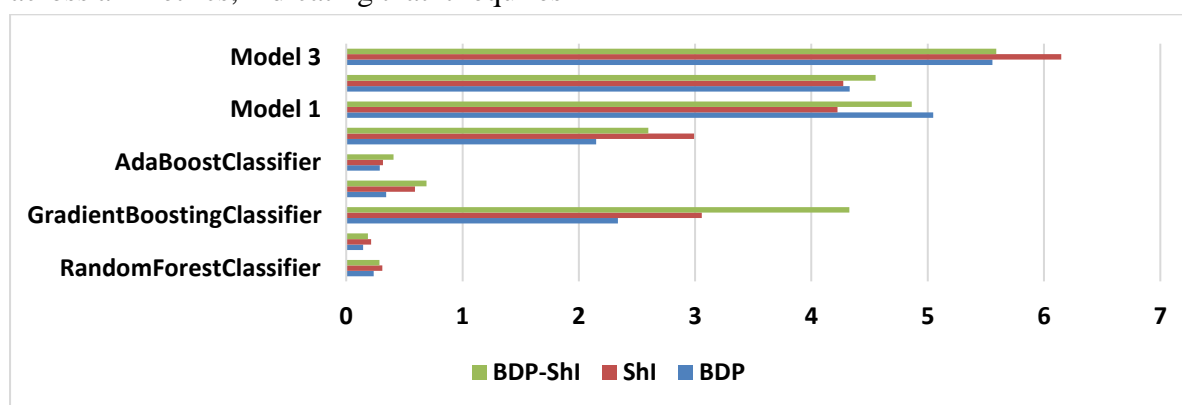


Figure 15: Training Time (Y-axis: Model Names, X-axis: Training Time in Seconds).



Thus, we can conclude the following for this study: Using the RFC model based on the BDP metric gives the best results for species identification. However, it can be challenging in some cases when species have similar otolith sizes. Using the BaC model based on the BDP-ShI complex provides the best and most comprehensive results for species identification using otolith analysis. This is because the BDP-ShI complex can minimize the dependence on similar sizes among species. The RFC model based on the BDP-ShI also gives stable results and its classification is not much different from the BaC model. Both models have potential for practical application. Using Model 1 (Dense 64) based on either the BDP metric or the BDP-ShI complex gives good results for species identification and can be applied in practice. Model 3 (Dense 256) also shows potential for classification due to its relatively high accuracy in species identification. Using the ShI metric alone for species identification with machine learning and deep learning models results in low accuracy and is not suitable for practical application.

### Acknowledgments

This work was funded by the grant in aid from Joint Vietnam-Russia Tropical Science and Technology Research Center, Ha Noi, Vietnam “ST.Đ1.09/23” and partially funded by the Vietnam Academy of Science and Technology’s grant in aid (ĐL0000.01/23-24).

### References

**Agüera, A. and Deirdre, B., 2011.** Use of sagittal otolith shape analysis to

discriminate Northeast Atlantic and Western Mediterranean stocks of *Atlantic saury*, *Scomberesox saurus saurus* (Walbaum). *Fisheries Research*, 110, 465-471. DOI:10.1016/j.fishres.2011.06.003

**Bani, A., Poursaeid, S. and Tuset, V.M., 2013.** Comparative morphology of the sagittal otolith in three species of south Caspian gobies. *Journal of Fish Biology*, 82, 1321-1332. DOI:10.1111/jfb.12073

**Breiman, L., 2001.** Random forests. *Machine Learning*, 45, 5–32.

**Burke, N., Brophy, D., Schön, P.J. and King, P.A., 2009.** Temporal trends in stock origin and abundance of juvenile herring (*Clupea harengus*) in the Irish Sea. *ICES Journal of Marine Science*, 66(8), 1749-1753. DOI:10.1093/icesjms/fsp140

**Chen, L., Zhang, Y. and Liu, Q., 2022.** Comparative analysis of traditional and AI-based approaches in fish otolith shape classification. *Marine Technology Society Journal*, 56(2), 45-58.

**Chen, Y. and Guoping Zhu, G., 2023.** 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*, 80(5), 1277–1290. DOI:10.1093/icesjms/fsad052

**Corrêa, G.M. de S., Coletto, J.L., Castello, J.P., Miller, N.R., Tubino, R. de A., Monteiro-Neto, C. and da Costa, M.R., 2022.** Identification of fish stock based on otolith as a natural marker: The case of *Katsuwonus pelamis* (Linnaeus, 1758) in the Southwest Atlantic Ocean. *Fisheries*

- Research, 255(1), 106436. DOI: 10.1016/j.fishres.2022.106436
- Dürr, J. and González, J. A., 2002.** Feeding habits of *Beryx splendens* and *Beryx decadactylus* (Berycidae) off the Canary Islands. *Fisheries Research*, 54, 363–374.
- Ferhani, K., Bekrattou, D. and Mouffok, S., 2021.** Inter-population morphological variability of the round sardinella (*Sardinella aurita* Valenciennes, 1847) on the Algerian Coast based on body morphometric, meristic, and otolith shape. *Iranian Journal of Fisheries Sciences*, 20(6), 1757-1774.
- Friedman, J. H., 2001.** Greedy function approximation: A gradient boosting machine. *Annals of Statistics*, 29, 1189–1232.
- Froese, R. and Pauly, D., 2024.** FishBase - World Wide Web electronic publication. Available at: <http://fishbase.org> (Accessed on 8/2/2024).
- Garcia-Rodriguez, F.J., De La, and Cruz-Agüero, J., 2011.** A comparison of indexes for prey importance inferred from otoliths and cephalopod beaks recovered from pinniped scats. *Journal of Fisheries and Aquatic Science*, 6, 186.
- Geurts, P., Ernst, D. and Wehenkel, L., 2006.** Extremely randomized trees. *Machine Learning*, 63(1), 3-42. DOI: 10.1007/s10994-006-6226-1
- Ghanbarifardi, M. and Zarei, R., 2021.** Otolith shape analysis of three mudskipper species of Persian Gulf. *Iranian Journal of Fisheries Sciences*, 20(2), 333-342.
- He, T., Cheng, J., Qin, J.G., Li, Y. and Gao, T.X., 2017.** Comparative analysis of otolith morphology in three species of Scomber. *Ichthyological Research*, 65, 192-201. DOI: 10.1007/s10228-017-0605-4
- Hosseini-Shekarabi, S.P., Valinassab, T., Bystydzińska, Z. and Linkowski, T., 2014.** Age and growth of *Benthoosema pterotum* (Alcock, 1890) (Myctophidae) in the Oman Sea. *Journal of Applied Ichthyology*, 31, 51-56. DOI:10.1111/jai.12620
- Keating, J. P., Brophy, D., Officer, R. A., 2014.** Otolith shape analysis of blue whiting suggests a complex stock structure at their spawning grounds in the Northeast Atlantic. *Fisheries Research*, 157, 1-6. DOI:10.1016/j.fishres.2014.03.009
- Lin, Y.J. and Al-Abdulkader, K., 2019.** Identification of fish families and species from the western Arabian Gulf by otolith shape analysis and factors affecting the identification process. *Marine and Freshwater Research*, 70, 1818–1827.
- Liu, J., Wang, H. and Zhao, X., 2021.** Deep learning for otolith-based fish species classification. *Journal of Fish Biology*, 98(5), 1337-1349.
- Lombarte, A. and Cruz, A., 2007.** Otolith size trends in marine fish communities from different depth strata. *Journal of Fish Biology*, 71, 53-76. DOI:10.1111/j.1095-8649.2007.01465.x
- Mapp, J., Hunter, E., Kooij, V.D., Songer, S. and Fisher, M., 2017.** Otolith shape and size: The importance of age when determining indices for

- fish-stock separation. *Fisheries Research*, 190, 43-52. DOI:10.1016/j.fishres.2017.01.017
- Osman, A., Farrag, M., Mehanna, S. and Osman, Y., 2020.** Use of otolithic morphometrics and ultrastructure for comparing between three goatfish species (family: Mullidae) from the northern Red Sea, Hurghada, Egypt. *Iranian Journal of Fisheries Sciences*, 19(2), 814-832.
- Paul, K., Oeberst, R. and Hammer, C., 2013.** Evaluation of otolith shape analysis as a tool for discriminating adults of Baltic cod stocks. *Journal of Applied Ichthyology*, 29(4), 743-750. DOI:10.1111/jai.12145
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M. and Duchesnay, É., 2011.** Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- Portnoy, D.S. and Gold, J.R., 2013.** Finding geographic population structure in marine fish species with high gene flow. *Proceedings of the 65th Gulf and Caribbean Fisheries Institute*, 65, 384-389.
- Sadighzadeh, Z., Valinassab, T., Vosugi, G., Motallebi, A.A., Fatemi, M.R., Lombarte, A. and Tuset, V.M., 2014.** Use of otolith shape for stock identification of John's snapper, *Lutjanus johnii* (Pisces: Lutjanidae), from the Persian Gulf and the Oman Sea. *Fisheries Research*, 155, 59-63.
- Salimi, N., Loh, K.H., Kaur Dhillon, S. and Chong, V.C., 2016.** Fully-automated identification of fish species based on otolith contour: Using short-time Fourier transform and discriminant analysis (STFT-DA). *PeerJ*, 4, e1664. DOI:10.7717/peerj.1664
- Santos, R.D.S., Costa-de-Azevedo, M.C., Albuquerque, C.D. and Araujo, F.G., 2017.** Different sagitta otolith morphotypes for the white mouth croaker *Micropogonias furnieri* in the Southwestern Atlantic coast. *Fisheries Research*, 195, 222-229. DOI:10.1016/j.fishres.2017.07.027
- Schulz-Mirbach, T., Ladich, F., Plath, M., Metscher, B.D. and Heb, M., 2014.** Are accessory hearing structures linked to inner ear morphology? Insights from 3D orientation patterns of ciliary bundles in three cichlid species. *Frontiers in Zoology*, 11, 1-25. DOI:10.1186/1742-9994-11-25
- Simoneau, M., Casselman, J.M. and Fortin, R., 2000.** Determining the effect of negative allometry (length/height relationship) on variation in otolith shape in lake trout (*Salvelinus namaycush*), using Fourier-series analysis. *Canadian Journal of Zoology*, 78, 1597-1603.
- Solomatine, D.P. and Shrestha, D.L., 2004.** AdaBoost. RT: a boosting algorithm for regression problems. *IEEE International Joint Conference on Neural Networks*, 2, 1163-1168.
- Stransky, C., 2005.** Geographic variation of golden redfish (*Sebastes marinus*) and deep-sea redfish (*S. mentella*) in the North Atlantic based on otolith shape analysis. *ICES Journal of Marine*

- Science*, 62(8), 1691–1698.  
DOI:10.1016/j.icesjms.2005.06.009
- Stransky, C., Baumann, H., Fevolden, S.E., Harbitz, A., Høie, H., Nedreaas, K.H., Salberg, A.B. and Skarstein, T., 2008.** Separation of Norwegian coastal cod and Northeast Arctic cod by outer otolith shape analysis. *Fisheries Research*, 90, 26-35.  
DOI:10.1016/J.FISHRES.2007.09.009
- Tuset, V.M., Azzurro, E. and Lombarte, A., 2012.** Identification of Lessepsian fish species using the sagittae otolith. *Scientia Marina*, 76, 289-299.  
<http://doi:10.3989/SCIMAR.03420.18E>
- Tuset, V. M., Farre, M., Otero-Ferrer, J. L., Vilar, A., Morales-Nin, B. and Lombarte, A., 2016.** Testing otolith morphology for measuring marine fish biodiversity. *Marine and Freshwater Research*, 67, 1037-1048.  
DOI:10.1071/MF15052
- Volpedo, A.V. and Echeverra, D.D., 2003.** Ecomorphological patterns of the sagitta in fish on the continental shelf off Argentina. *Fisheries Research*, 60, 551-560.  
DOI:10.1016/S0165-7836(02)00170-4
- Vu, Q.T. and Kartavsev, Yu. Ph., 2020.** Otolith shape analysis and its utility for identification of two smelt species, *Hypomesus japonicus* and *H. nipponensis* (Osteichthyes, Osmeridae) from the northwestern Sea of Japan with inferences in stock discrimination of *H. japonicus*. *Russian Journal of Marine Biology*, 46, 431-440.  
DOI:10.1134/S1063074020060115
- Vu, Q.T., Tran, V.D. and Nguyen, V.Q., 2022.** Otolith shape utilization for the stock identification of Caroun croaker, *Johnius carouna* (Cuvier, 1830) (Perciformes: Sciaenidae), from the east Vietnam sea. *Iranian Journal of Fisheries Sciences*, 21(4), 864-879.
- Yedier, S., 2021.** Otolith shape analysis and relationships between total length and otolith dimensions of European barracuda, *Sphyraena sphyraena*, in the Mediterranean Sea. *Iranian Journal of Fisheries Sciences*, 20, 4, 1080-1096.