Turkish Journal of FISHERIES and AQUATIC SCIENCES

RESEARCH PAPER

A Novel Method Utilizing Otolith Outline Analysis for Identifying Fish Species

Vu Quyet Thanh[1],* , Nguyen Van Quan[2]

[1]Vietnam-Russia Tropical Science and Technology Research Center, Ha Noi, Vietnam.
[2]Vietnam Academy of Science and Technology, Ha Noi, Vietnam.

Abstract

Fish species can be identified based on the analysis of otolith morphological indices. Common methods used to determine populations or identify species based on otolith morphology analysis include Basic Dimension Parameters (BDP), Shape Indices (ShI), and Elliptic Fourier Descriptors (EFD) indices through various statistical techniques. In this paper, we introduce a novel method to distinguish between otolith morphological groups to identify differences among five commercially important fish species in Vietnam—*Sillago sihama, Siganus canaliculatus, Selaroides leptolepis, Johnius carouna, and Otolithes ruber*—based on 188 left otolith samples. This method involves the automated calculation of distances from the otolith center to 100 points along its contour, implemented through a Python-based image processing pipeline. This is the first comparison between three otolith morphology analysis methods, including two conventional methods and a new method. The results of this study demonstrated that the center-to-contour distance index is effective for classifying otolith morphology with the aim of identifying fish species. Additionally, the study results reaffirm that otolith morphological analysis is an effective tool for identifying fish species and distinguishing otolith morphological differences among the five species.

Introduction

Otoliths, also known as ear stones, are calcium carbonate structures found in the inner ears of bony fishes, excluding lampreys. These structures play a crucial role in balance and hearing. Each fish has three pairs of otoliths: 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 sagittal otoliths is associated with swimming capabilities (Volpedo & Echeverra, 2003) and stock distribution (Lombarte & Cruz, 2007; Tuset et al*.,* 2016). The morphometric properties of otoliths are extensively used for species differentiation and stock identification (Stransky et al*., 2008*; Bani et al*.,* 2013). Otoliths also preserve records of individual fish growth and development (Campana et al*.,* 1985; Hosseini-Shekarabi et al*.,* 2014; Yedier, 2021). Numerous studies have utilized sagittal otolith shapes to identify species, populations, and stocks (Osman et al*.,* 2020; Ghanbarifardi & Zarei, 2021). Variations in sagittal shapes often correlate with the biological characteristics of species and fish stocks (Campana & Casselman, 1993; Vu & Kartavsev, 2020). In recent years, otolith morphology analysis has been employed to distinguish between species or populations (He et al*.,* 2017; Vu & Kartavsev, 2020). The analysis of sagitta size and shape is particularly valuable for differentiating fish species and identifying fish populations. Besides conventional species identification methods like morphological

analysis and DNA sequencing, otolith morphology-based identification is especially useful for reconstructing historical fish species composition in archaeological studies (Lin et al., 2019) and for investigating the feeding habits of piscivorous animals, thereby enhancing our understanding of marine food webs, particularly at higher trophic levels (Dürr & González, 2002; Garcia-Rodriguez et al., 2011). Additionally, otoliths have the advantage of long-term sample preservation, requiring minimal space and simple storage conditions.

Many authors have identified species using common indices such as basic dimension parameter - BDP (length, width, perimeter, area), shape indices – ShI (circularity, roundness, rectangularity, form factor, aspect ratio, and ellipticity), and EFD, which are standard methods for otolith shape analysis (Burke et al., 2009; Mapp et al., 2017; He et al., 2017; Vu & Kartavsev, 2020). A recent study employed BDP, ShI, and EFD indices of otoliths as tools to distinguish between two species, *Hypomesus japonicus* and *H. nipponensis* (Vu & Kartavsev, 2020). The analysis of otolith shape has also been used to determine the stock structure of European anchovy (*Engraulis encrasicolus*) along the Tunisian coast, identifying three distinct stock units of *E. encrasicolus*, with significant implications for fisheries management. By analyzing BDP and ShI indices using various statistical methods, numerous studies have assessed the morphological differences in otoliths among fish species. These analyses are used to delineate fish stocks (Agüera & Brophy, 2011; Paul et al., 2013; Ferhani et al., 2021).

However, these methods have their own advantages and disadvantages in identifying and observing differences in otolith morphology, with a certain rate of misidentification still present (Lin & Al-Abdulkader, 2019). Therefore, the development of new methods, especially shape indices, is necessary to improve accuracy. In this paper, we introduce a novel method using the distance from the center to points on the otolith contour (DftC) index, implemented in the Python programming language, for species identification. This method will be compared with currently prevalent methods. The samples used in this study include five common species from the coastal waters around Soncha Island (Hue City). These species frequently appear in natural habitats and are commercially significant fish used for human consumption and as prey for piscivorous animals in the nearby forest. The biology and ecology of the five fish species under investigation highlight their varied growth patterns and ecological roles within coastal and estuarine environments. *Sillago sihama*, a species commonly found in coastal waters, estuaries, and lagoons, reaches sexual maturity at a total length of approximately 12 cm, with moderate growth and a lifespan of up to 4 years. This species inhabits sandy areas in sea inlets, along beaches, sandbars, mangrove creeks. Typically forming schools, *S. sihama* exhibits a unique behavior where adults bury themselves in the

sand when disturbed, enhancing their ability to evade predators. As a benthic feeder, it primarily consumes polychaete worms, small prawns (genus Penaeus), shrimps, and amphipods. This benthic shoaling species contributes significantly to nutrient cycling within these habitats and serves as an essential prey item for larger marine species. *Selaroides leptolepis* inhabits shallow coastal waters, including coral reefs and the inshore waters of the continental shelf. The species is associated with seagrass beds and feeds mainly on plankton, such as ostracods, gastropods, euphausiids, and small juvenile fishes. Being a diurnal feeder, it subsists on a variety of small marine organisms, forming schools that serve as key prey for larger marine predators. Reaching sexual maturity within one year at a length of about 12 cm, this species often resides in soft substrates, making it a schooling species that plays a significant ecological role in the marine food web. *Siganus canaliculatus* inhabits shallow coastal waters, often associated with coral reefs and seagrass beds. This herbivorous species primarily consumes algae and seagrass, occasionally feeding on incidental invertebrates while browsing. It is a shoaling species, playing a crucial role in maintaining coral reef balance by controlling algal growth. With rapid juvenile growth, it reaches maturity at around 12 cm, ensuring early reproductive readiness (Froese & Pauly, 2022). *Otolithes ruber* and *Johnius carouna* are important carnivorous species in estuarine and coastal ecosystems, with moderate growth rates. *O. ruber* is found on the continental shelf, inhabiting subtidal edges of mud flats, sheltered bays, and estuaries. This generalist carnivore preys on shrimp, small fish, and squids. It reaches maturity at approximately 16 cm and is a mid-level predator in these environments. On the other hand, *J. carouna*, a widely distributed species in shallow coastal waters and estuaries, typically reaches maturity at 15.3 cm (Sawusdee & Rattanara, 2021; Froese & Pauly, 2022). This species primarily feeds on crustaceans, small fish, worms, and insects, playing a significant role in nutrient cycling within the benthic community. Together, these species represent a range of trophic roles and growth patterns essential to the biodiversity and function of tropical marine ecosystems. While various methods have been developed to quantify otolith shape, most rely on general outline descriptors or frequency-domain transformations. In contrast, approaches that compute fixed-point radial distances from the otolith centroid are rare in the literature. This gap underscores the novelty of our study, which introduces and systematically evaluates a center-to-contour distance index (DftC) as a shape descriptor. In addition to its conceptual simplicity and reproducibility, the DftC index offers high potential for integration with machine learning and deep learning models. Because the distance values are structured, fixed-length, and continuous, they can be directly used as input features in supervised classification models, neural networks, or feature selection pipelines. Therefore, this approach not only enables conventional statistical evaluation but also

paves the way for scalable, automated fish species identification frameworks based on otolith morphology.

Materials and Methods

Sample Collection

Individuals were captured by trawling and purchased from fishing areas in Soncha (Figure 1), with samples collected in May and June 2024. All individuals used were of adult size according to FishBase (2022). A total of 188 left otoliths were collected (Figure 1), representing five fish species. The number of individuals per species and their corresponding standard lengths are detailed in Table 1.

Image Acquisition and Preprocessing

Each otolith was placed against a black background (Table 1). The left sagitta (hereafter referred to as otolith) was photographed using an Olympus SZ61 zoom stereo microscope. Digital images of the otoliths were captured under the stereomicroscope using Olympus CellSens (version 2.2) with an SC180 camera and saved in JPG format.

Basic Dimension Parameters (BDP) and Shape Indices (ShI)

The characteristics of the otoliths were measured and analyzed based on ShIs. To determine the morphometric characteristics of the otoliths, four basic dimensional parameters, namely area (A), perimeter (P), otolith width (OW), and otolith length (OL), were measured using Python (version 3.9). Six common ShIs were calculated using the ratios of OW, OL, A, and P (Agüera & Brophy 2011) as follows: aspect ratio = OL/OW; ellipticity = (OL – OW) / (OL + OW); circularity = P/A^2; rectangularity = A/(OL × OW); roundness = $4A/\pi OL^2$; and form factor = $4\pi A/P^2$ (Agüera & Brophy, 2011; He et al., 2017).

Elliptic Fourier Descriptors (EFD)

Fourier analysis was conducted using the pyefd package (Henrik, 2020), following the formula proposed by Kuhl & Giardina (1982). Specifically, the elliptic_fourier_descriptors function was employed with a default order of 16, generating 64 coefficients (16 orders × 4 coefficients per order). These coefficients were then used in Linear Discriminant Analysis (LDA) to
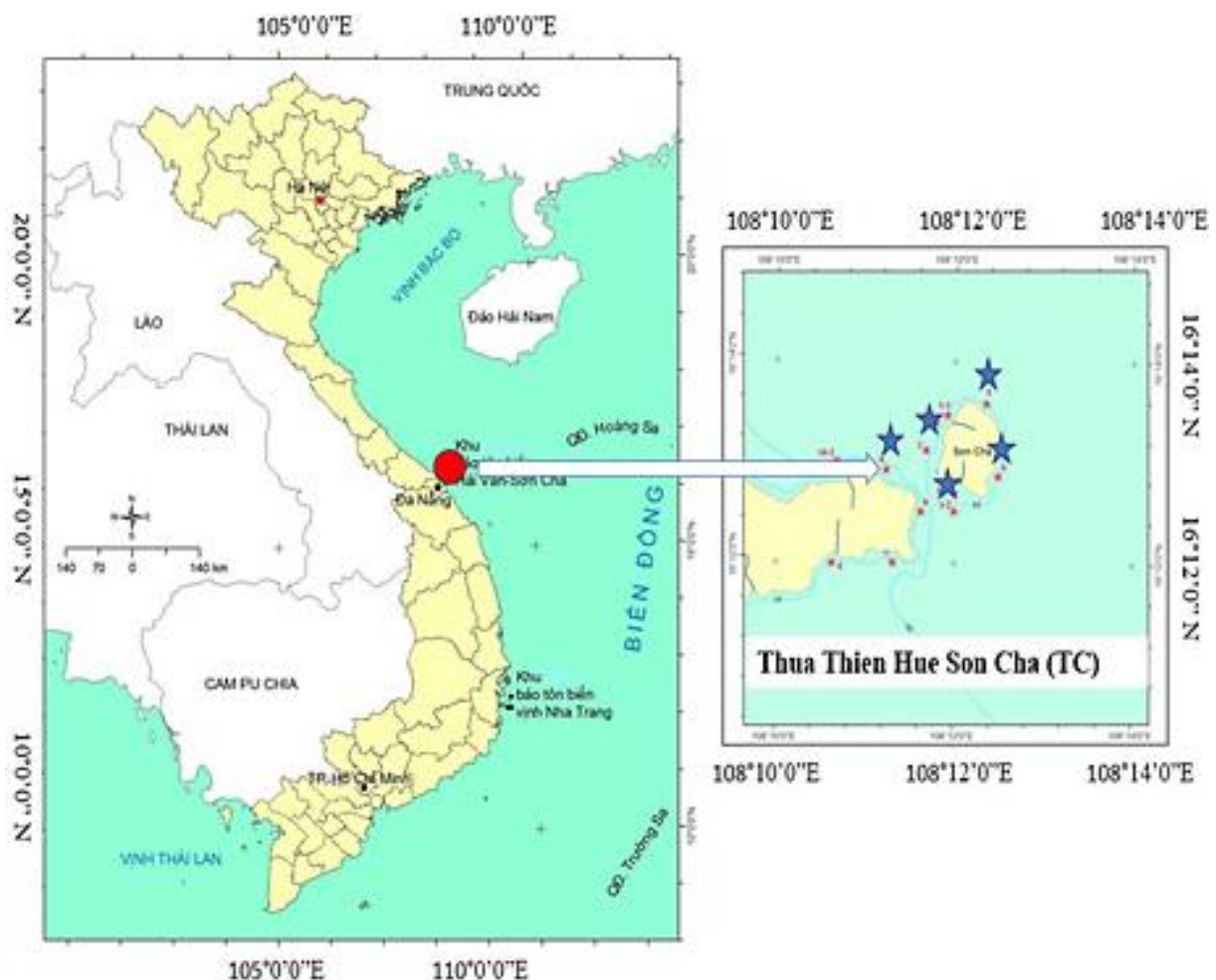

Figure 1. Sampling locations of otolith. The blue stars indicate specific fishing points.

assess data clustering. The command coeffs = elliptic_fourier_descriptors (contour, o rder=16, normalize=False) was used to compute these descriptors.

Proposed Method: Distance from Center to Contour (DftC)

Identify the center of the otoliths and calculate the distance from the center to points on the:

 # Compute spatial moments of the contour to determine its centroid

M = cv2.moments(max_contour)

if M['m00'] != 0:

cx = int(M['m10'] / M['m00'])
cy = int(M['m01'] / M['m00'])

 # Compute the Euclidean distance from the centroid to each point on the contour

distances = np.sqrt((max_contour[:, 0, 0] - cx) ** 2 + (max_contour[:, 0, 1] - cy) ** 2)

The center of the otoliths was identified, and the distance from the center to 100 points on the contour was calculated, resulting in a set of 100 distances. This set of distances was subsequently used to compare otolith differences among species using MANOVA and LDA analyses.

Data Analysis

Data Processing in Python

The otolith image was converted to grayscale, and the edge was detected using the Canny function. The largest contour was selected from the detected edges to reduce noise and ensure consistency. Subsequently, basic dimensional parameters (BDPs) were extracted in pixel units, including OL, OW, P, and A:

 # Load image and extract contour using OpenCV
image = cv2.imread(image_path, cv2.IMREAD_GRAYSCALE)

edges = cv2.Canny(image, 100, 300)

contours, _ = cv2.findContours(edges, cv2.RETR_TREE, cv2.CHAIN_APPROX_SIMPLE)
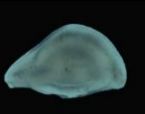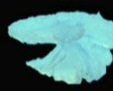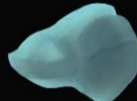
 # Select the largest contour to reduce noise

max_contour = max (contours, key=cv2.contourArea, default=None)

 # Compute basic dimensional parameters (BDPs)

if max_contour is not None:

x, y, width, height = cv2.boundingRect(max_contour)
OL = max (width, height)
OW = min (width, height)
P = cv2.arcLength(max_contour, True)
A = cv2.contourArea(max_contour)

Table 1. Information on the otolith samples of the 5 species collected for the study

Species Name	Species Code	Image	Sample Size	Size of the fish
Order: Eupercaria Family: Sillaginidae *Sillago sihama* (Fabricius, 1775)	SSI		29	16 - 19 cm
Order: Carangiformes Family: Carangidae *Selaroides leptolepis* (Cuvier, 1833)	CCV		50	12 – 16 cm
Order: Acanthriformes Family: Siganidae *Siganus canaliculatus* (Park, 1797)	DIA		30	18 – 25 cm
Order: Eupercaria Family: Sciaenidae *Otolithes ruber* (Bloch & Schneider, 1801)	XOR		29	20 -25 cm
Order: Eupercaria Family: Sciaenidae *Johnius carouna (Cuvier, 1830)*	CDD		50	16 – 19 cm

The Libraries of Python Used In the Analysis Include

The values used for statistical analysis were tested for homogeneity of variance and normal distribution using the statsmodels package (Seabold and Perktold, 2010). The packages include os for system operations (Python Software Foundation, 2024), numpy (abbreviated as np) for numerical computations (Harris et al., 2020), and cv2 for image processing (Bradski, 2000). For data manipulation and analysis, pandas (abbreviated as pd) (McKinney, 2010) and seaborn (abbreviated as sns) (Waskom, 2021) are used to handle and visualize data. The library matplotlib.pyplot (abbreviated as plt) (Hunter, 2007) is also used for plotting. Additionally, the code employs LinearDiscriminantAnalysis (LDA) from sklearn.discriminant_analysis and LabelEncoder from sklearn.preprocessing for classification and label encoding, along with confusion_matrix from sklearn.metrics to evaluate results (Pedregosa et al., 2011). MANOVA from statsmodels.multivariate.manova support multivariate statistical analysis (Seabold & Perktold, 2010), while pingouin (pg) provides additional statistical tools (Vallat, 2018).

Results

Determination and Extraction of Contour Results

Figure 2 shows the contours of all 188 otolith samples from five species that have been well defined, the contours are unique and closely fit the objects, with no noisy contours. The determined contours are unique to each object. Based on these identified contours, data on BDP-ShI, EFD, and DftC will be extracted for each species group for comparison.

Manova Analysis of Morphological Differences in Otoliths among Five Fish Species Using Bdp, Shl, Efd, and DftC

The MANOVA (Multivariate Analysis of Variance) test is a statistical method used to examine whether there are significant differences across multiple dependent variables among groups, here, different fish species. By analyzing various morphological characteristics of otoliths together, MANOVA can identify differences that may be linked to species. This
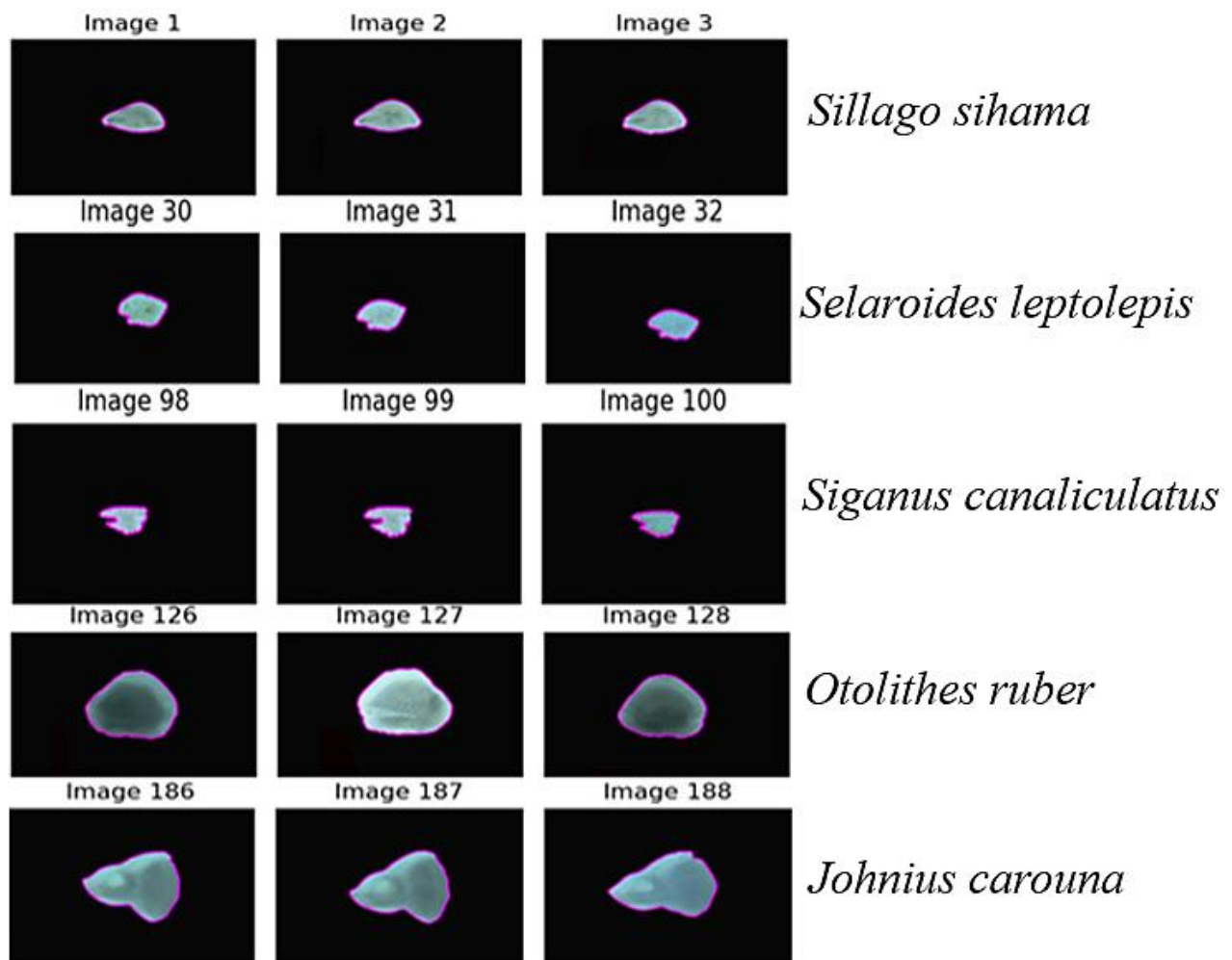

Figure 2. Results for the determination of otolith contours (contours outlined in purple).

method allows us to explore potential variations in otolith morphology across species based on indices such as BDP, ShI, EFD, and DftC. The results in Table 2 outline these morphological differences among five fish species, indicating statistical significance across these indices.

The MANOVA results for morphological otolith characteristics based on BDP, ShI, EFD, and DftC indices across five fish species consistently indicate statistically significant effects of the species variable on otolith metrics. Specifically, the analysis of Length, Width, Perimeter, and Area (BDP) yields low Wilks' lambda (0.027), high Pillai's trace (0.972), Hotelling-Lawley trace (35.872), and Roy's greatest root (35.872), all with F values of 1614.28 and P<0.001, confirming significant interspecies differences. This trend is reinforced in the between-group analysis, where Wilks' lambda (0.01) and other high F values (all P<0.001) further underscore substantial shape differences across species. Similarly, the ShI shape indices (including aspect ratio, ellipticity, circularity, rectangularity, roundness, and form factor) exhibit strong variance explained by species, with the Intercept statistics showing Wilks' lambda at 0.003, Pillai's trace at 0.996, and both Hotelling-Lawley and Roy's greatest root at 315.886, all with extremely high F values (9371.29) and P<0.001. Between-group analysis

here also confirms pronounced species differences, with Wilks' lambda at 0.0115 and significant p-values across all metrics. The EFDs follow this pattern, with Wilks' lambda at 0.01, Pillai's trace at 0.99, and both Hotelling-Lawley and Roy's greatest root at 93.68, all indicating a significant model with F values of 188.89 and P<0.001, and with between-group results (Wilks' lambda at 0.001) reinforcing substantial morphological variation between species. Finally, the DftC results demonstrate a similarly strong species impact on otolith shape, as indicated by Wilks' lambda (0.026), Pillai's trace (0.974), Hotelling-Lawley trace (37.655), and Roy's greatest root (37.655), each with an F value of 31.63 and P<0.001. Between-group statistics, with a near-zero Wilks' lambda (0.0001), confirm significant species distinctions (Table 2). Together, these MANOVA findings affirm statistically significant morphological differences in otolith shape among all five species. However, while MANOVA substantiates these overall differences, it does not reveal the extent of variation between populations or the potential for misclassification across groups. Thus, further analysis using LDA plots and confusion matrices is necessary to assess clustering precision and evaluate inter-population misclassification.

Table 2. MANOVA results for morphological otoliths based on BDP, ShI, EFD, DftC of five fish species.

		Statistic	Value	Num DF	Den DF	F Value	*Pr*>F
BDP	Intercept	Wilks' lambda	0.027	4	180	1614.28	0.001
		Pillai's trace	0.972	4	180	1614.28	0.001
		Hotelling-Lawley trace	35.872	4	180	1614.28	0.001
		Roy's greatest root	35.872	4	180	1614.28	0.001
	Between-group	Wilks' lambda	0.01	16	550.55	118.85	0.001
		Pillai's trace	1.6	16	732	30.53	0.001
		Hotelling-Lawley trace	43.64	16	354.04	488.31	0.001
		Roy's greatest root	42.61	4	183	1949.63	0.001
ShI	Intercept	Wilks' lambda	0.003	6	178	9371.29	0.001
		Pillai's trace	0.996	6	178	9371.29	0.001
		Hotelling-Lawley trace	315.886	6	178	9371.29	0.001
		Roy's greatest root	315.886	6	178	9371.29	0.001
	Between-group	Wilks' lambda	0.0115	24	622.17	67.35	0.001
		Pillai's trace	2.138	24	724	34.654	0.001
		Hotelling-Lawley trace	14.466	24	412.13	106.6	0.001
		Roy's greatest root	10.537	6	181	317.869	0.001
EFD	Intercept	Wilks' lambda	0.01	61	123	188.89	0.001
		Pillai's trace	0.99	61	123	188.89	0.001
		Hotelling-Lawley trace	93.68	61	123	188.89	0.001
		Roy's greatest root	93.68	61	123	188.89	0.001
	Between-group	Wilks' lambda	0.001	244	494.07	33.85	0.001
		Pillai's trace	3.73	244	504	28.22	0.001
		Hotelling-Lawley trace	78.61	244	437.85	39.16	0.001
		Roy's greatest root	33.30	61	126	68.79	0.001
DftC	Intercept	Wilks' lambda	0.026	100	84	31.631	0.001
		Pillai's trace	0.974	100	84	31.631	0.001
		Hotelling-Lawley trace	37.655	100	84	31.631	0.001
		Roy's greatest root	37.655	100	84	31.631	0.001
	Between-group	Wilks' lambda	0.0001	400	338.698	11.179	0.001
		Pillai's trace	3.599	400	348	7.814	0.001
		Hotelling-Lawley trace	100.797	400	300.192	20.802	0.001
		Roy's greatest root	79.173	100	87	68.881	0.001

Num DF: Numerator Degrees of Freedom; Den DF: Denominator Degrees of Freedom; F Value: The F value, a measure used to test the differences between groups; *Pr*>F: The p value, indicating the level of statistical significance

Morphological Otolith Analysis Based on BDP-ShI

The Linear Discriminant Analysis results for the five fish species, based on the BDP and ShI shape indices, reveal two discriminant functions with eigenvalues of 0.858 and 0.101 (Figure 3). The total eigenvalue sum is 0.96, indicating the overall variance explained by the model. The first discriminant function accounts for 89.43% of the variance, while the second function explains 10.57% of the variance. This indicates that the first discriminant function is the most significant, capturing the majority of the discriminative information between the species. The second function, although less significant, still contributes to the model's ability to differentiate the fish species. These results suggest that the shape indices BDP-ShI are effective in distinguishing between the five fish species, with the first discriminant function playing a dominant role in this differentiation (Figure 3). The LDA group scatter plot, interpreted along the LD1 axis, demonstrates a distinct separation, dividing the chart into two segments. On the positive side of the axis, *species O. ruber* and *J. carouna* are distributed, while the remaining species, *S. leptolepis* and *Siganus canaliculatus*, are positioned on the negative side. As analyzed previously, LD1 contributes significantly to explaining the variance, highlighting its importance in the LDA grouping. Regarding the vertical axis, the majority of the species are positioned above the value of -2, with only *S. canaliculatus* appearing on the opposite side. Overall, the LDA plot distinctly categorizes the morphological groups. There are only minor overlaps among a few points of *O. ruber* and *J. carouna*; minimal overlap between *S. sihama* and *S. leptolepis*; and between S. *leptolepis* and *S. canaliculatus* (Figure 3).

The analysis of shape features contributing to the Linear Discriminant Analysis plot for distinguishing among five fish species reveals that ellipticity and aspect ratio play the most significant roles along the LD1 axis. Ellipticity, with the highest contribution score of 79.53, is the primary feature differentiating the species on this axis, indicating substantial variation in otolith ellipticity among the groups. Aspect ratio follows with a contribution of 14.68, further enhancing the distinction along LD1. For the LD2 axis, form factor stands out with the largest contribution (57.57), showing it is the key feature separating species along this secondary axis, complemented by ellipticity and rectangularity with
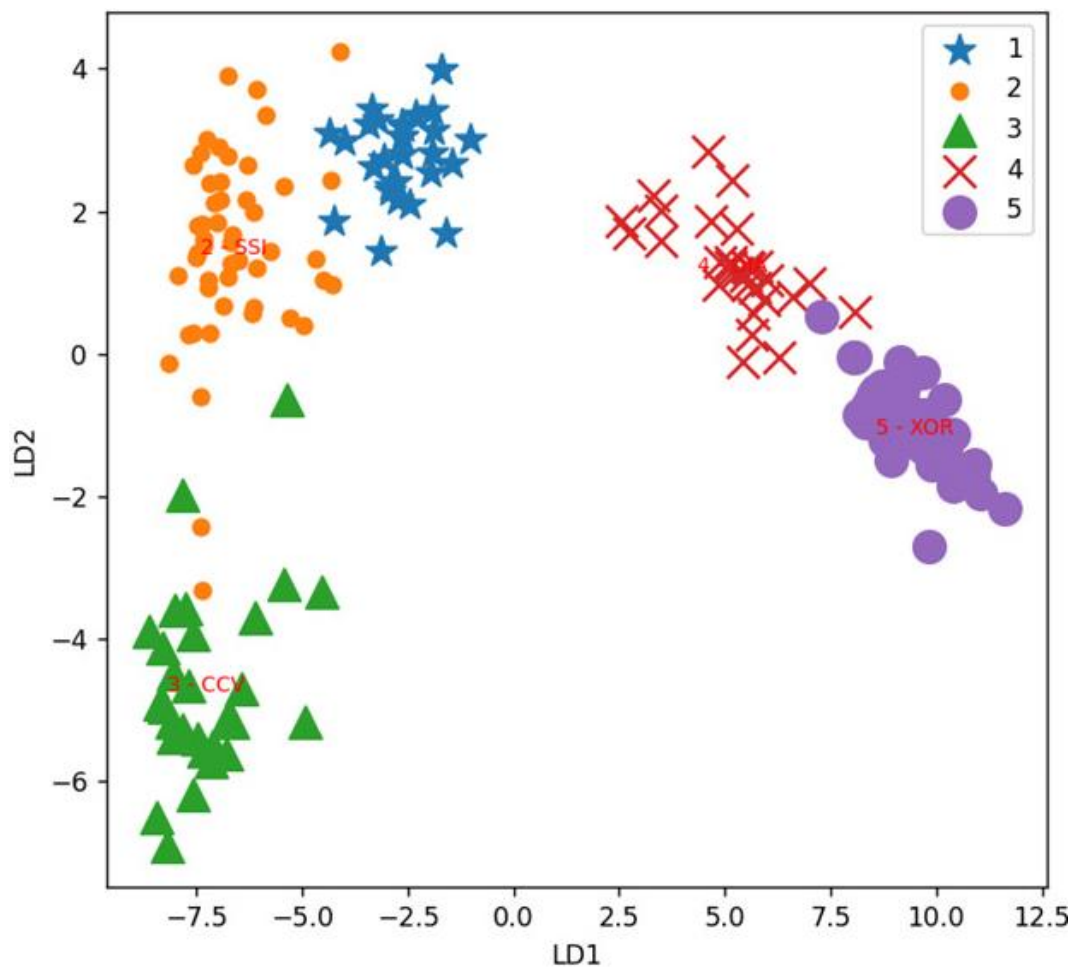

Figure 3. Scatter plot of linear discriminant analysis of otolith morphometric differences among five fish species based on BDP-ShI. 1 blue stars - *S. sihama*; 2. orange circles - *S. leptolepis*; 3. green triangles - *S. canaliculatus*; 4 - red crosses *O. ruber*; 5. purple circles - *J. carouna*.

notable contributions of 20.72 and 18.13, respectively (Table 3). Features such as area, circularity, and perimeter exhibit minimal contributions to both LD1 and LD2, suggesting limited influence on group separation. This analysis indicates that ellipticity, aspect ratio, and form factor are the most informative descriptors for otolith morphology in distinguishing fish species in LDA space, providing critical insights into morphological differences relevant to species identification.

To assess the level of misclassification among fish groups based on BDP-ShI, we will examine the confusion matrix for pairwise species comparisons in Figure 4. The confusion matrix from the Linear Discriminant Analysis clearly demonstrates the classification capabilities for five fish species based on BDP-ShI traits. *S. sihama* and *J. carouna* both achieve perfect classification accuracies

of 100%, indicating no misclassifications among these samples. *S. leptolepis*, however, shows a classification accuracy of 94% with 4% of its samples misclassified as *S. canaliculatus* and 2% as *S. sihama*, suggesting potential similarities that may impede clear differentiation. *S. canaliculatus* is accurately classified at a rate of 96.67%, with a small proportion, 3.33%, being incorrectly classified as *S. sihama*, highlighting a slight confusion that could be attributed to overlapping traits. *O. ruber* shows an accuracy of 96.55%, with a minor misclassification rate of 3.45% as *J. carouna* (Figure 4). These results illustrate the effectiveness of LDA in distinguishing between these species, although the slight misclassifications between *S. leptolepis*, *S. canaliculatus*.

Table 3. Contribution of shape features to linear discriminant functions (LD1 and LD2) for otolith morphological differentiation among five fish species

Factor	LD1	LD2	Contribution LD1	Contribution LD2
Length	0,104	0,064	0,076	0,023
Width	0,007	0,056	0,005	0,020
Perimeter	-0,012	-0,038	0,009	0,014
Area	0,000	0,000	0,000	0,000
Aspect ratio	20,037	7,410	14,678	2,649
Ellipticity	-108,564	-57,957	79,530	20,716
Circularity	0,000	0,002	0,000	0,001
Rectangularity	0,805	50,724	0,590	18,130
Roundness	-1,098	-2,469	0,805	0,883
Form factor	-5,879	-161,056	4,307	57,566

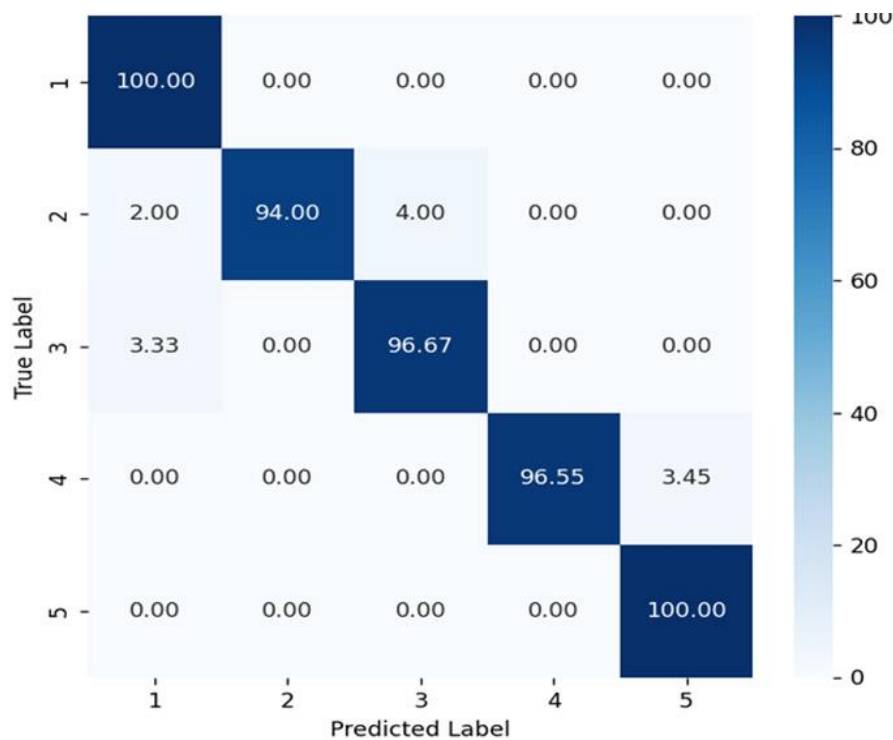Note: LD1 - Linear Discriminant 1; LD2 - Linear Discriminant 2.

Figure 4. Confusion matrix for species classification using linear discriminant analysis based on BDP-ShI. 1. *S. sihama,* 2. *S. leptolepis,* 3. *S. canaliculatus,* 4. *O. ruber,* 5. *J. carouna*.

Morphological Otolith Analysis Based on Efd

Eliptic Fourier Descriptors of five fish species reveal two discriminant functions with eigenvalues of 0.424 and 0.319 (Figure 5). The total eigenvalue sum is 0.743, indicating the overall variance explained by the model. The first discriminant function accounts for 57.02% of the variance, while the second discriminant function explains 42.98% of the variance. These results suggest that both discriminant functions play substantial roles in differentiating the otolith morphology among the fish species. Unlike previous analyses where one function might dominate, in this case, both functions contribute significantly to the model's ability to distinguish between species. This indicates that the EFD method captures a more balanced set of discriminative information, with no single function overwhelmingly dominating the variance explanation. Focusing on the LD1 axis, it is evident that *J. carouna* is clearly isolated on the far left. Nearby, *O. ruber* also forms a separate cluster, albeit slightly overlapping with *S. sihama* and *S. leptolepis*, which are closer to the center of the LD1 axis (Figure 5)*. S. canaliculatus*, distinctly separated and located towards the top of the LD2 axis, showcases a clear divergence from the other species, emphasizing its unique morphological traits based on EFD (Figure 5). The LD1 axis primarily aids in the segregation of *J. carouna*,

underlining major differences in their otolith shapes compared to the others. The slight overlap observed between *S. sihama*, *O. ruber*, and *S. leptolepis* near the center suggests some similarity in their otolith morphometrics, which may lead to closer clustering (Figure 5).

Each species has achieved a classification accuracy of 100%, as evidenced by a 100% value in the corresponding column for each species label, with zero values across other columns in the respective rows. The absence of misclassifications among the species indicates that the EFD indices utilized in the LDA are highly effective at differentiating these species (Figure 6).

Figure 7 illustrates the contributions of various variables (coefficients) to the first and second linear discriminants (LD1 and LD2) derived from a linear discriminant analysis based on Elliptic Fourier Descriptors. The bar plot indicates that certain coefficients, such as coef 62, coef 50, and coef 61, exhibit high contributions to LD1, as evidenced by their larger weights. This suggests these variables play a significant role in defining the separation along LD1. Similarly, for LD2, several coefficients, including coef 55 and coef 50, show prominent contributions, implying their relevance in distinguishing along this dimension. The distribution of weighting factors across coefficients
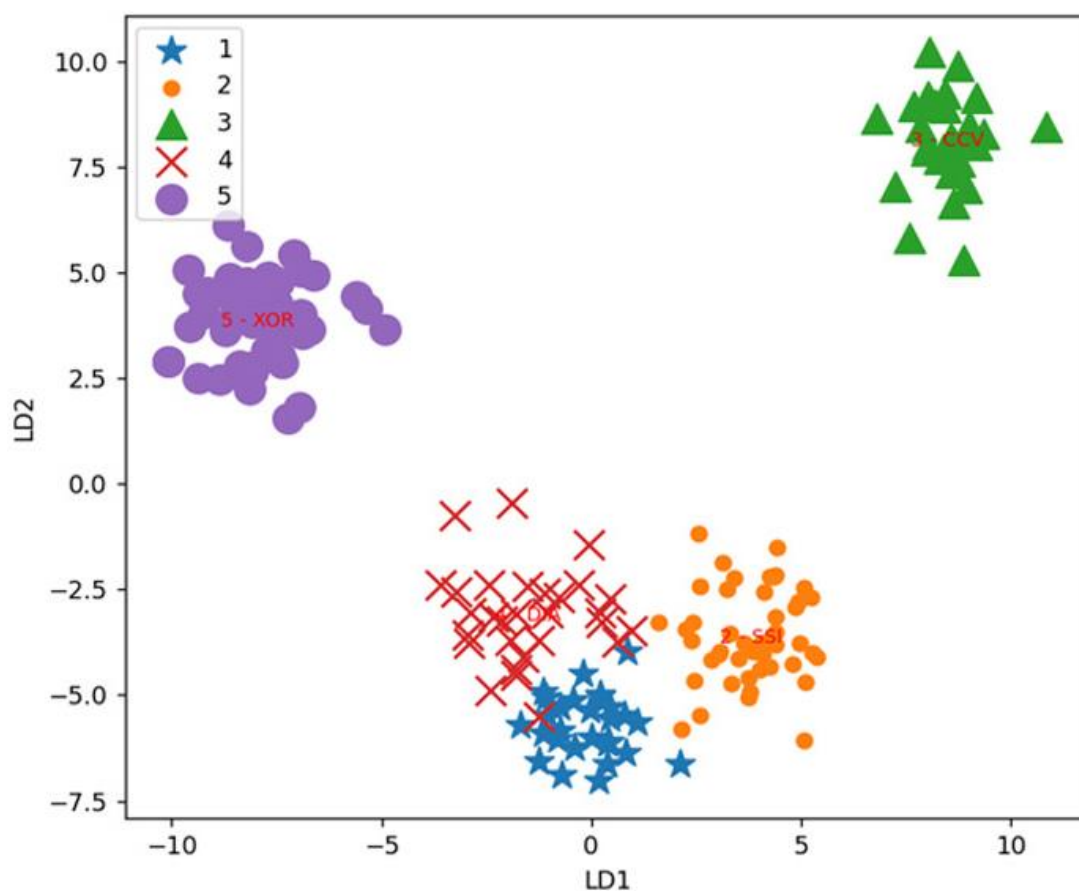

Figure 5. Scatter plot of linear discriminant analysis of otolith morphometric differences among five fish species based on coefficients of EFD. 1 blue stars - *S. sihama*; 2. orange circles - *S. leptolepis*; 3. green triangles - *S. canaliculatus*; 4 - red crosses *O. ruber*; 5. purple circles - *J. carouna*.

highlights a pattern where a subset of variables strongly influences each discriminant, while many others have minimal impact. This suggests that LDA may benefit from the most influential variables, allowing for potential dimensionality reduction by focusing on these high-contribution coefficients. The differential contribution between LD1 and LD2 may also indicate distinct feature sets that discriminate between the classes in different ways along each axis.

Morphological Otolith Analysis Based on DftC

The LDA results indicate the presence of two discriminant functions with eigenvalues of 0.785 and 0.098 (Figure 8). The total eigenvalue sum is 0.89, representing the overall variance explained by the model. The first discriminant function accounts for 88.89% of the variance, while the second discriminant function explains 11.11% of the variance (Figure 8). This suggests that the first discriminant function captures the majority of the discriminative information between the classes, making it the most significant component in distinguishing the different groups in the dataset. The second function, while less significant, still contributes to the model's ability to separate the classes. Figure 8 provides a clear separation of species along the LD1 and LD2 axes. On the LD1 axis, *J. carouna* is clearly isolated

on the far left, showcasing its unique morphometric features distinct from those of other species. *O. ruber* is also distinctly grouped, with no overlap with *J. carouna*, which is positioned closer to the center of the LD1 axis. This distinct positioning of *O. ruber* indicates its unique otolith characteristics that set it apart from *J. carouna*, despite these two species being in the same family and relatively similar in terms of body morphology. This result demonstrates a significant improvement over classifications based on BDP-ShI.

Each species has achieved a classification accuracy of 100%, as evidenced by a 100% value in the corresponding column for each species label, with zero values across other columns in the respective rows. The absence of misclassifications among the species indicates that the DftC indices utilized in the LDA are highly effective at differentiating these species (Figure 9).

Figure 10 illustrates the contributions of various distance variables from DftC to the first and second linear discriminants (LD1 and LD2) in a linear discriminant analysis (LDA). Notably, certain distances, such as distance 11, distance 10, and distance 8, show high contributions to LD1, indicating their importance in defining this discriminant. In contrast, a few other distances, including distance 75 and distance 99, contribute more significantly to LD2. The distribution of
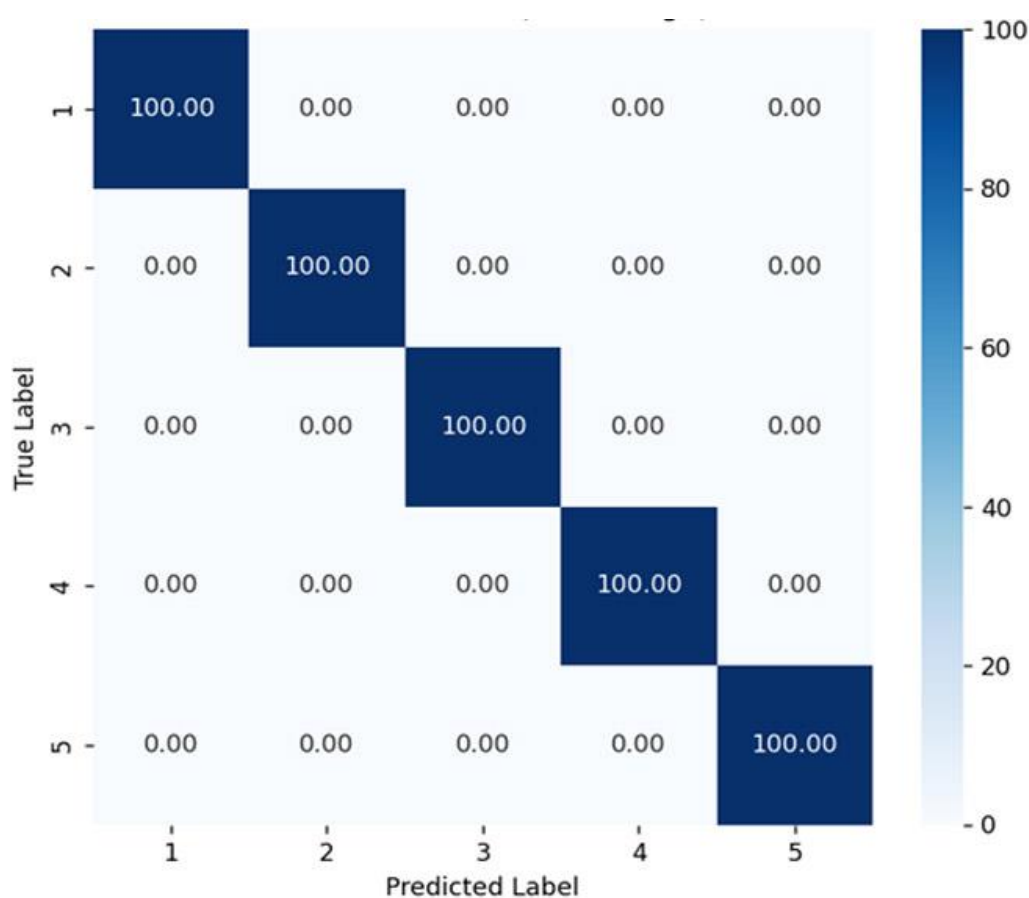

Figure 6. Confusion matrix for species classification using linear discriminant analysis based on coefficients of EFD. 1. *S. sihama,* 2. *S. leptolepis,* 3. *S. canaliculatus,* 4. *O. ruber,* 5. *J. carouna*.
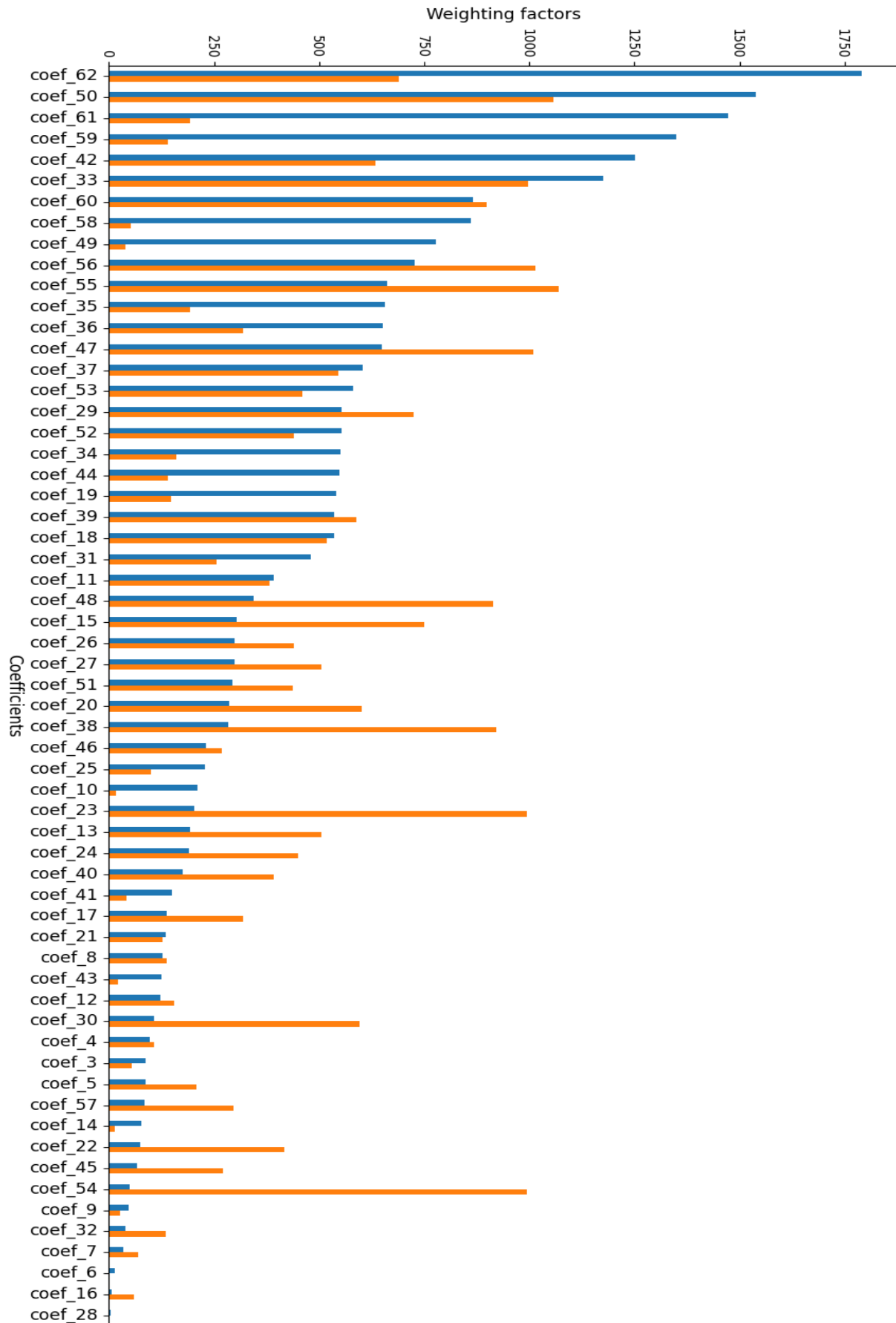
Figure 7. Histogram of variable contributions to LD1 and LD2 in linear discriminant analysis based on EFD coefficients. The blue bars represent the contributions of variables to LD1, while the orange bars indicate their contributions to LD2. The chart highlights the weighting factors of each variable in defining the separation along each discriminant, with certain coefficients showing higher importance for LD1 or LD2.
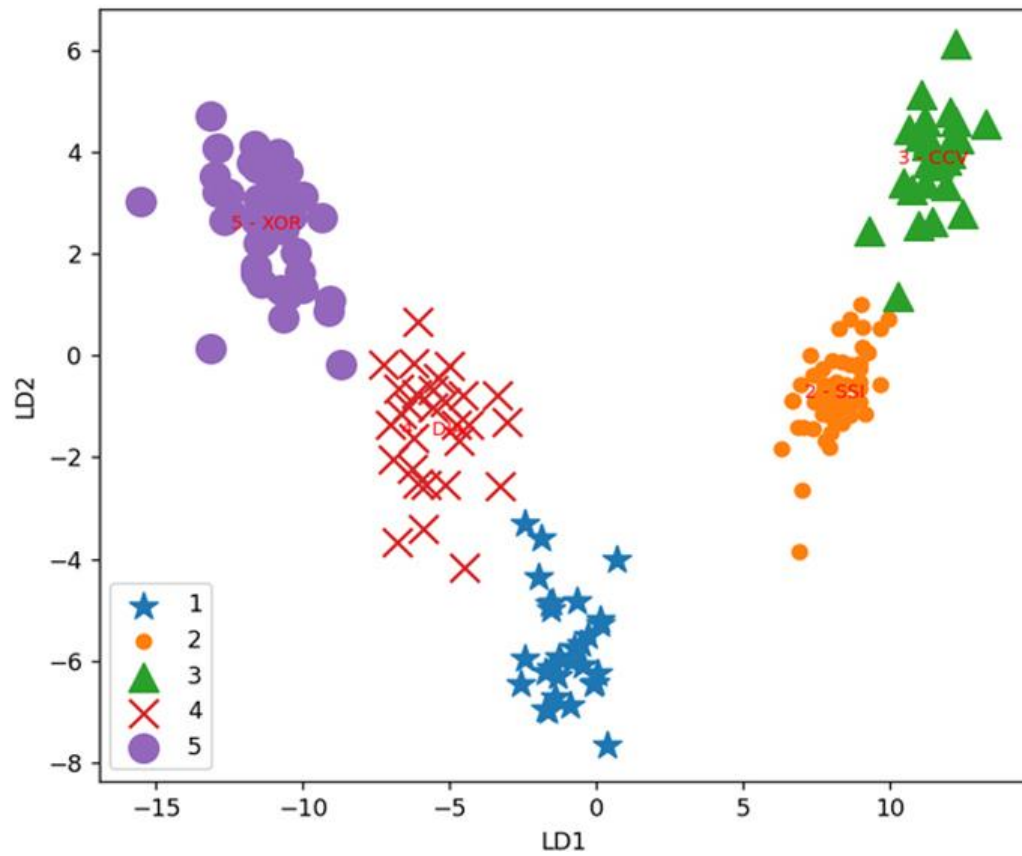
Figure 8. Scatter plot of linear discriminant analysis of otolith morphometric differences among five fish species based on DftC. 1 blue stars - *S. sihama*; 2. orange circles - *S. leptolepis*; 3. green triangles - *S. canaliculatus*; 4 - red crosses *O. ruber*; 5. purple circles - *J. carouna*.
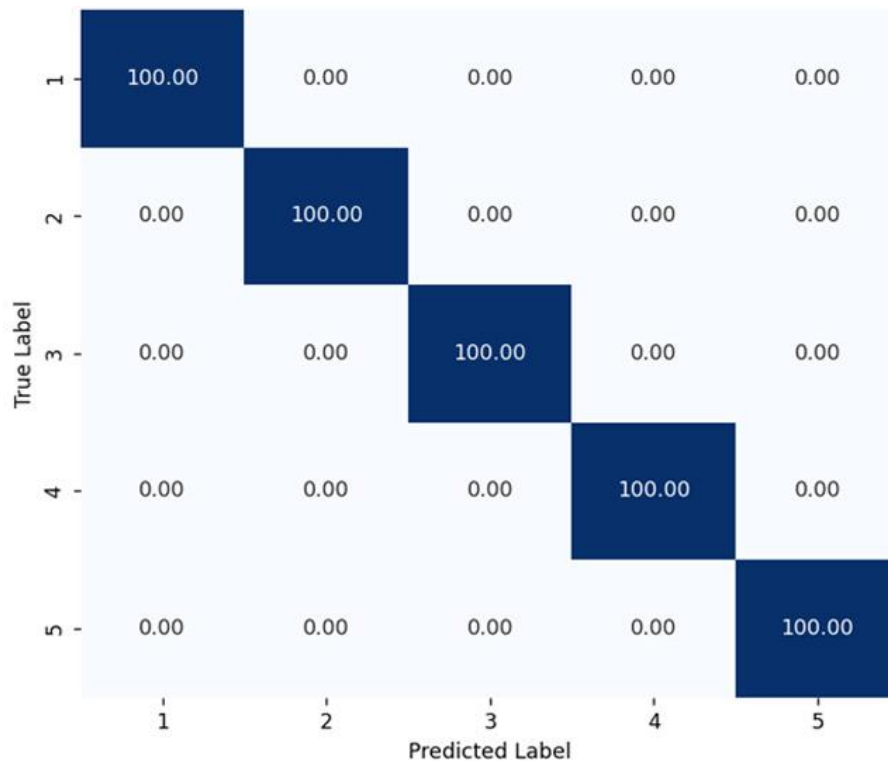

Figure 9. Confusion matrix for species classification using linear discriminant analysis based on DftC. 1. *S. sihama,* 2. *S. leptolepis,* 3. *S. canaliculatus,* 4. *O. ruber,* 5. *J. carouna.*
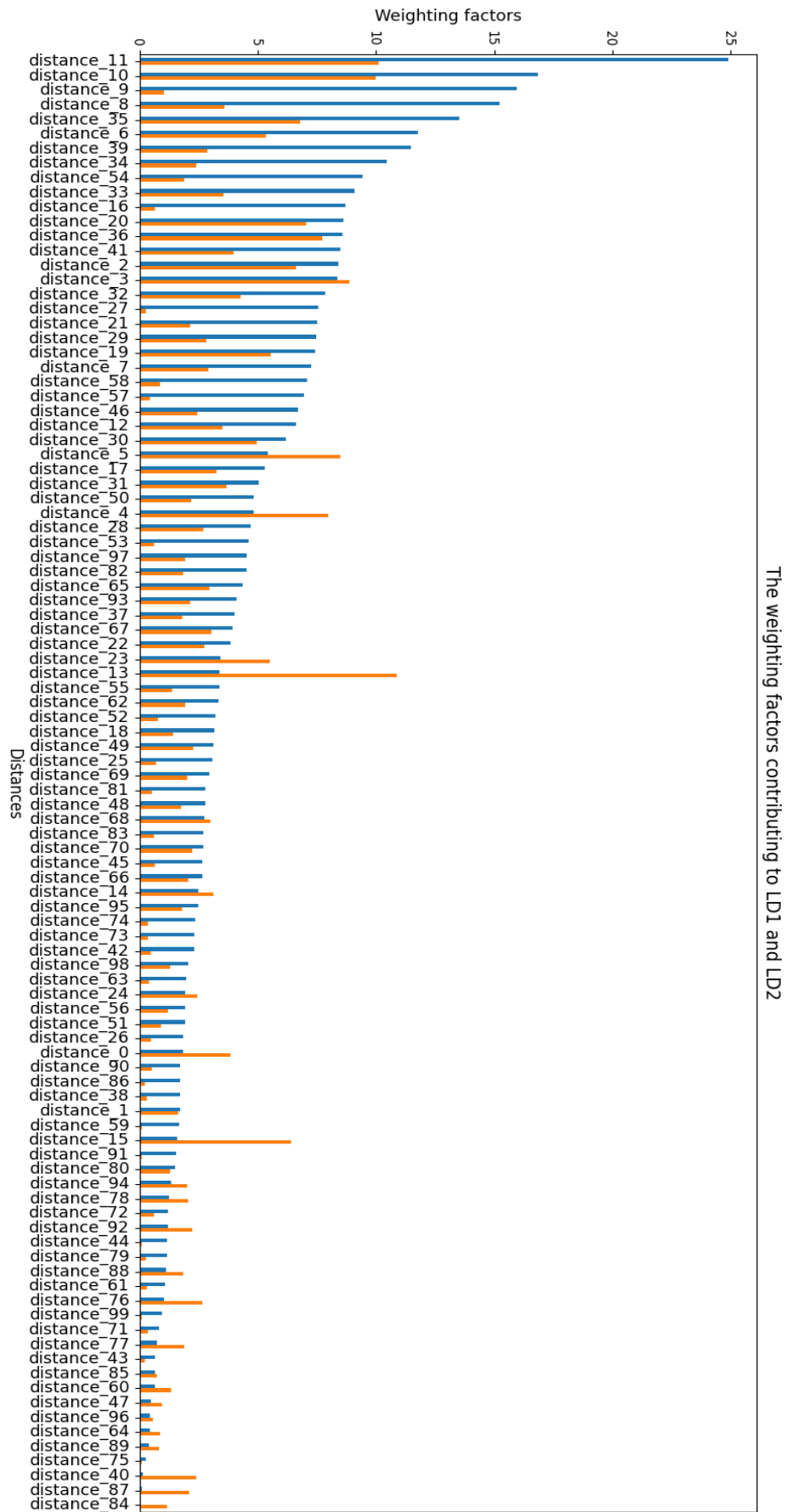
Figure 10. Histogram of variable contributions of DftC distances to LD1 and LD2 in linear discriminant analysis. The blue bars represent the contributions of variables to LD1, while the orange bars indicate their contributions to LD2. The chart highlights the weighting factors of each variable in defining the separation along each discriminant, with certain coefficients showing higher importance for LD1 or LD2.

weighting factors across the distances suggests that LD1 and LD2 each emphasize different sets of distance variables, which play crucial roles in distinguishing between classes along each discriminant. This differential weighting implies that the DftC distances are utilized variably across the two discriminants, potentially providing unique separation patterns based on specific distance features.

Discussion

Synthesizing the results from Figures 3, 4, 5, 6, 8, 9, and Table 2 we observe that among the three methods used to assess otolith morphology, DftC exhibits the clearest grouping and highest accuracy. To further evaluate and gain deeper insights into the effectiveness of these methods, we will examine Figure 11, which depicts the reconstructed mean outlines using 100 mean points on the outline.

It is evident that, in terms of size, utilizing BDP-ShI has the drawback of difficulty in distinguishing between certain species pairs: *S. sihama - S. leptolepis*; *S. sihama - S. canaliculatus*; *O. ruber - J. carouna*, as they share similar sizes and shapes compared to other pairs. Consequently, misidentifications occurred with frequencies of 2.22% for *S. sihama - S. canaliculatus*; 3.33% for *S. leptolepis - S. canaliculatus*; and 3.45% for *O. ruber - J. carouna* (Figure 4). Regarding the coefficients of EFD, although they are independent of otolith size, this also presents a limitation as otolith size

is an important characteristic for differentiation. DftC exhibits a balanced characteristic by preserving both size traits and outline variations. Thus, the grouping in the LDA plot based on DftC (Figure 8) is clearer compared to LDA based on EFD coefficients (Figure 5).

According to several references, the study by Salimi et al. (2016) used the STFT shape index to identify 14 fish species and successfully identified them with an accuracy of 70% to 100%. The authors consider this a relatively acceptable identification result; however, they emphasize the need to continue evaluating the identification capability using different recognition models. The lack of comprehensive characteristics to describe otolith shapes can make classification challenging (Simoneau et al., 2000). The research by Yuwen Chen and Guoping Zhu in 2023 utilized various machine learning models to identify otolith species based on ShI and wavelet transform for otoliths of four fish species (*E. carlsbergi*, *C. rastrospinosus*, *P. antarcticum*, *K. anderssoni*), achieving overall accuracy from 66% to 96%. Thus, with good clustering results and species identification accuracy reaching 100%, DftC has the potential to be a good classification index for otolith morphology research.

The study by Avigliano et al. (2016) on several fish species used canonical discriminant analysis (CDA) based on BDP, ShI, and additional indices to illustrate species differentiation among *H. cunninghami*, *K. stewarti*, *L. haplodactylus*, *N. compressus*, and *U. xenogrammus*, with correct classification rates ranging
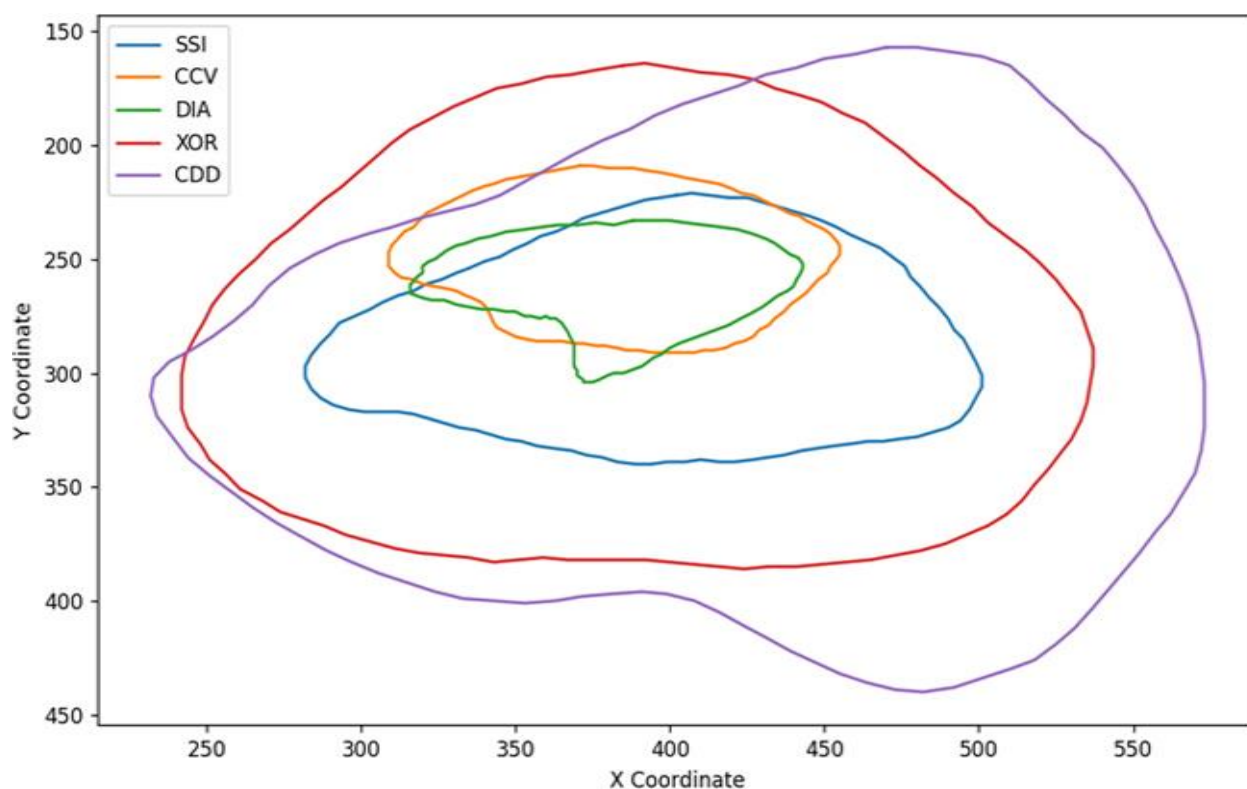

Figure 11. Average contour of five species. SSI- *S. sihama*; CCV- *S. leptolepis*; DIA - *S. canaliculatus*; 4 – XOR *O. ruber*; CDD- *J. carouna*.

from 66% to 100%. However, the accuracy rate was lower (50%) for *A. rufus*, *B. dorsalis*, *C. jojettae*, *F. flavonigrum*, *F. varium*, *F. gymnotum*, *M. bathytaton*, and *N. caerulpunctus*. In our research, species classification accuracy was relatively high, especially when using EFD and DftC indices, which showed differentiation up to 100%. For BDP and SHI indices, the minimum accuracy was 94%, likely due to the distinct nature of the study subjects. Charmpila et al. (2024) conducted a study on five fish species collected from southern Iran and the Persian Gulf region. When using otolith variables and shape indices, the authors reported an overall classification accuracy of 80.2%, with values ranging from 65% to 89% across species. The application of wavelet-based shape analysis to the same dataset resulted in significantly lower classification performance, with accuracy ranging from 13% to 78%, and an overall average of only 39.6%. Although the five species in their study belong to the same genus, the low classification success with wavelet analysis suggests limitations in capturing subtle intra-genus differences. Our newly proposed DftC method, with its fixed angular resolution and detailed radial structure, may offer improved discriminatory power in such cases (accuracy of 100%). More recently, Vu et al. (2025) evaluated the classification ability of six machine learning models and three deep learning architectures based on BDP, ShI, and their combinations. With BDP-based inputs, the Random Forest Classifie RFC and the Bagging Classifier BaC models achieved classification accuracies ranging from 84.7% to 86.8%. Classification based on ShI alone yielded much lower accuracies, with RFC reaching only 53.47% and other models falling below 50%. Deep learning models (Dense64, Dense128, Dense256) trained on BDP-ShI data yielded accuracies between 75.7% and 77.1%. Our newly proposed method (DftC) demonstrated higher classification accuracy compared to BDP-ShI, suggesting its strong potential to further enhance performance when integrated with machine learning and deep learning models. The DftC index provides a compact, structured, and biologically interpretable shape descriptor. Unlike Elliptic Fourier Descriptors (EFD), which often generate large coefficient sets requiring dimensionality reduction through PCA—potentially discarding important shape information—the DftC retains full spatial resolution with no transformation loss. Many studies have indicated that environmental factors, such as salinity, water temperature, and depth, contribute to inter- and intra-species differences in otolith groove area and length. However, variables like otolith size, rostrum shape, and groove morphology are primarily genetically controlled within a specific fish group (Reichenbacher & Reichard, 2014). Biologically, the five species in this study mostly belong to different families, thus reflecting substantial genetic differences. Regarding their habitats, these species represent the tropical coastal zone with distinct diets and habitats, leading to relatively high differentiation levels. Additionally, DftC demonstrated superior discriminatory power compared to commonly used indices. Continual improvements in research methodology, particularly the development of new indices, are essential in scientific research on otoliths. The DftC index set, comprising 100 parameters, effectively numerically characterizes otolith shape. This numerical dataset can be associated with each species and holds the potential to establish a standard otolith database for each species. Integrating this with machine learning or deep learning techniques could transform it into a highly useful index.

Conclusion

This study reaffirms that otolith morphological analysis is an effective tool for species classification, with the five fish species investigated exhibiting distinct otolith morphological characteristics.

All three methods (BDP-ShI, EFD, DftC) utilized for otolith morphological analysis in this study showed good classification capabilities for the five commercial fish species. Specifically, scatter plot of LDA and Confusion matrix based on DftC exhibited the clearest grouping (species clearly separated along LD1 and LD2 axes, with no overlap) and high species identification accuracy (100% correct prediction for all species). Based on EFD, relatively clear grouping was observed (with some minor overlap) and high species identification accuracy (100% correct prediction for all species). With BDP-ShI, however, there was still significant overlap in morphological features between species on the LDA plot, resulting in misclassification rates ranging from 3.33% to 6%.

Ethical Statement

The care and use of experimental animals complied with Vietnam's animal welfare laws, guidelines, and policies as approved by the Vietnamese authority under Decree No. 26/2019/NĐ-CP.

Funding Information

Author Contribution

First Author: Conceptualization, Formal Analysis, Investigation, Methodology, Writing -review and editing; Second Author: Data Curation, Visualization and Writing -original draft.

Conflict of Interest

The author(s) declare that we have no known financial or non-financial, professional, or personal conflicts of interest that could have appeared to influence the work reported in this paper.

References

Agüera, A., & 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. http://dx.doi.org/10.1016/j.fishres.2011.06.003

Avigliano, E., Jawad, LA., & Volpedo, AV. (2016). Assessment of the morphometry of saccular otoliths as a tool to identify triplefin species (Tripterygiidae). *Journal of the Marine Biological Association of the United Kingdom*, *96*(5), 1167-1180.

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

Bradski, G. (2000). The OpenCV Library. Dr. Dobb's *Journal of Software Tools*.

Burke, N., Brophy, D., Schön, P. J., & 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. https://doi.org/10.1093/icesjms/fsp140

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

Campana, S., & Casselman, J. (1993). Stock discrimination using otolith shape analysis. *Canadian Journal of Fisheries and Aquatic Sciences, 50*(5), 1062-1083. https://dx.doi.org/10.1139/f93-123

Charmpila, E. A., Teimori, A., & Reichenbacher, B. (2024). Otolith-based species identification in the killifish Aphaniops (Teleostei; Cyprinodontiformes; Aphaniidae) using both morphometry and wavelet analysis. *Acta Zoologica*, 00(0), 1–18.

Chen, Y., & 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. https://doi.org/10.1093/icesjms/fsad052

Dürr, J., & 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., & 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.

Froese, R., & Pauly, D. (2022). *FishBase - World Wide Web electronic publication*. Access www.fishbase.se date18/01/2022.

Garcia-Rodriguez, F. J., De La, & 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.

Ghanbarifardi, M., & Zarei, R. (2021). Otolith shape analysis of three mudskipper species of Persian Gulf. *Iranian Journal of Fisheries Sciences, 20*(2), 333-342.

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., ... & Oliphant, T. E. (2020). Array programming with NumPy. Nature, 585(7825), 357-362.

He, T., Cheng, J., Qin, J. G., Li, Y., & Gao, T. X. (2017). Comparative analysis of otolith morphology in three species of *Scomber*. *Ichthyological Research, 65*, 192-201. https://doi.org/10.1007/s10228-017-0605-4

Hosseini-Shekarabi, S. P., Valinassab, T., Bystydzieńska, Z., & Linkowski, T. (2014). Age and growth of *Benthosema pterotum* (Alcock, 1890) (Myctophidae) in the Oman Sea. *Journal of Applied Ichthyology, 31*, 51-56. https://doi.org/10.1111/jai.12620

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, *9*(3), 90-95.

Kuhl, F. P., & Giardina, C. R. (1982). Elliptic Fourier features of a closed contour. *Computer Graphics and Image Processing, 18*(3), 236-258. https://doi.org/10.1016/0146-664X(82)90034-X

Lin, C. H., & Chien, C. W. (2021). Late Miocene otoliths from northern Taiwan: insights into the rarely known Neogene coastal fish community of the subtropical northwest Pacific. *Historical Biology*, 34(2), 361–382. https://doi.org/10.1080/08912963.2021.1916012

Lin, Y. J., & 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.

Lombarte, A., & Cruz, A. (2007). Otolith size trends in marine fish communities from different depth strata. *Journal of Fish Biology, 71*, 53-76. https://doi.org/10.1111/j.1095-8649.2007.01465.x

Mapp, J., Hunter, E., Kooij, V. D., Songer, S., & Fisher, M. (2017). Otolith shape and size: The importance of age when determining indices for fish-stock separation. *Fisheries Research, 190*, 43-52. https://doi.org/10.1016/j.fishres.2017.01.017

McKinney, W. (2010). Data Structures for Statistical Computing in Python. Proceedings of the 9th Python in Science Conference, 51-56.

Osman, A., Farrag, M., Mehanna, S., & 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., & 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. https://doi.org/10.1111/jai.12145

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.

Reichenbacher, B., & Reichard, M. (2014) Otoliths of five extant species of the annual killifish Nothobranchius from the East African savannah. *PLoS ONE*, 9(11):e112459. doi: 10.1371/journal.pone.0112459.

Salimi, N., Loh, K. H., Kaur Dhillon, S., & 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.

https://doi.org/10.7717/peerj.1664

Santos, R. D. S., Costa-de-Azevedoa, M. C., Albuquerque, C. D., & 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. https://doi.org/10.1016/j.fishres.2017.07.027

Sawusdee, A. ., & Rattanarat, J. (2021). Population Dynamics of the Caroun Croaker Johnius carouna (Cuvier, 1830) in Coastal Fishing Ground in the Middle Gulf of Thailand . *Walailak Journal of Science and Technology (WJST)*, 18(17), Article 13709 (8 pages). https://doi.org/10.48048/wjst.2021.13709

Schulz-Mirbach, T., Ladich, F., Plath, M., Metscher, B. D., & 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. https://doi.org/10.1186/1742-9994-11-25

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. *Proceedings of the 9th Python in Science Conference*, 57-61.

Simoneau, M., Casselman, J. M., & 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.

Stransky, C., Baumann, H., Fevolden, S. E., Harbitz, A., Høie, H., Nedreaas, K. H., Salberg, A. B., & Skarstein, T. (2008). Separation of Norwegian coastal cod and Northeast Arctic cod by outer otolith shape analysis. *Fisheries Research, 90*, 26-35. https://doi.org/10.1016/J.FISHRES.2007.09.009

Tuset, V. M., Farre, M., Otero-Ferrer, J. L., Vilar, A., Morales-Nin, B., & Lombarte, A. (2016). Testing otolith morphology for measuring marine fish biodiversity. *Marine and Freshwater Research, 67*, 1037-1048. https://dx.doi.org/10.1071/MF15052

Vallat, R. (2018). Pingouin: statistics in Python. *Journal of Open Source Software*, *3*(31), 1026

Volpedo, A. V., & Echeverra, D. D. (2003). Ecomorphological patterns of the sagitta in fish on the continental shelf off Argentina. *Fisheries Research, 60*, 551-560. https://doi.org/10.1016/S0165-7836(02)00170-4

Vu, Q. T., & 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. https://doi.org/10.1134/S1063074020060115

Vu, Q. T., Pham, T D., & Nguyen, V. Q. (2025). Artificial intelligence models for identifying several fish species based on otolith morphology index analysis from nearshore areas of Vietnam. *Iranian Journal of Fisheries Sciences*, 24(2), 277–304. , https://doi.org/ 10.22092/ijfs.2025.132877

Waskom, M. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60), 3021.

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.

Henrik, B. (2020) PyEFD: Elliptical Fourier Descriptors in Python (Version 1.1.0). Available at. https://pyefd.readthedocs.io/en/latest/(8/8/2024)