I used “Fish Species Image Data” and done some EDA on it.

# EDA

Image Statistics: Calculate and display statistics like the mean, standard deviation, and Labels.

Number of labels: 193

Image Mean: [107.53350298 123.19331434 122.46424965]

Image Std Dev: [74.10129476 72.92373349 78.332349 ]

Sample Image Display: Display a grid of sample images.

A collage of different fish

Description automatically generated

Class Distribution Visualization: Plot the class distribution for a better understanding of the dataset balance.

A collage of different fish

Description automatically generated

Label Count Distribution: Plot the count distribution of all labels.

A graph of blue lines

Description automatically generated

Image Aspect Ratios: Visualize the aspect ratios of images to understand the variety in image dimensions.

A blue graph with black text

Description automatically generated

Class-wise Mean Image: Create and visualize mean images for each class.

A group of fish with different colors

Description automatically generated

Image Size Distribution: Plot the distribution of image sizes (widths and heights).

A graph of different colored lines

Description automatically generated

Class-wise Standard Deviation Image: Visualize the standard deviation images for each class to understand intra-class variance.

A group of images of a fish

Description automatically generated

Principal Component Analysis (PCA): Reduces the dimensionality of the image data to 2 components and visualizes it, showing how images cluster based on their labels.

A group of colorful confetti

Description automatically generated

T-SNE Visualization: Another dimensionality reduction technique, often providing better separation for visualization of high-dimensional data.

A group of colorful squares

Description automatically generated

Correlation Matrix: Shows the correlation between pixel values, helping to understand the relationships between different pixels.

Class-wise Pixel Intensity Distribution: Plots the distribution of pixel intensities for each class, giving insight into the characteristics of different classes.

A group of graphs showing different sizes of data

Description automatically generated with medium confidence

Edge Detection Visualization: Uses edge detection to highlight the edges in images, which can be useful for analyzing the structure and details within images.

A collage of images of a fish

Description automatically generated

Image Brightness and Contrast Distribution: Analyzes the brightness and contrast of images, providing insights into the overall lighting and variation within the dataset.

A red graph with black text

Description automatically generated

A blue graph with black text

Description automatically generated