Exploratory Data Analysis (EDA) Report:
Fish Counting – Domain Adaptive Object Detection

The Caltech Fish Counting – Domain Adaptive Object Detection (CFC-DAOD) dataset provides a benchmark for domain-adaptive object detection in sonar images. The goal of this project is to develop a robust fish detection model that can solve the domain adaptation challenges and improve detection accuracy under different environmental conditions. Specifically, I plan to train a model on the Kenai Left Bank (source domain) and adapt it to detect fish in the Kenai Channel (target domain) using pseudo-labeling. This report presents an exploratory data analysis (EDA) of the dataset which includes information about data distribution, bounding box characteristics, and domain shift challenges. The findings will guide preprocessing and model selection for efficient fish counting.

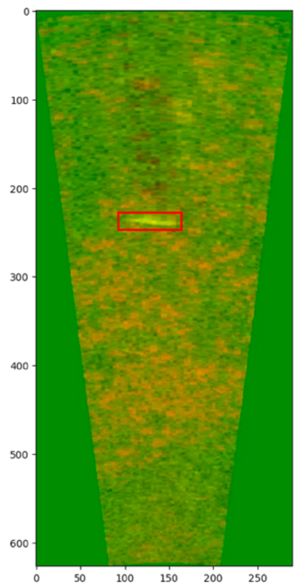
The dataset follows the COCO annotation format which includes images, annotations, and categories. The images have metadata such as file name, image ID, image height and image width. The annotations have information about bounding box coordinates (x, y, width, height) for detected fish. The categories for both train and test contain only one fish species which the category id is 1 and the name is 'fish'. Table 1 shows the training set consists of 76619 images with 132010 annotations, while the test set includes 13090 images with 41761 annotations.

Dataset	Images	Annotations	Categories
Source (Train)	76619	132010	1
Target (Test)	13090	41761	1

Table 1: Summary of the dataset composition, including the number of images, annotations, and categories for the training (Source) and test (Target) sets.

A distinguishing feature of this dataset is the 'Baseline++' format which has been introduced in the original CFC paper. It could enhance fish visibility by integrating three key image processing techniques into a single three-channel representation. The three key image processing techniques are original image data, background subtraction, and frame difference. This unique preprocessing method improves contrast and highlights fish features. It makes it easier for detection models to distinguish fish from their surroundings. The effectiveness of fish detection is evaluated using mean Average Precision (mAP) at an Intersection over Union (IoU) threshold of 0.5. It provides a standardized measure of model performance in accurately localizing fish in the dataset.

Understanding the distribution of fish counts and bounding box sizes in the dataset is crucial for designing effective detection models. Below, I have an in-depth analysis of different patterns. From Figure 1, it illustrates an example of a fish detection scenario. The red bounding box highlights a detected fish in the sonar image. The image shows a cone-like region of interest with a noisy background, which suggests the challenges in detecting fish due to occlusion and low contrast.



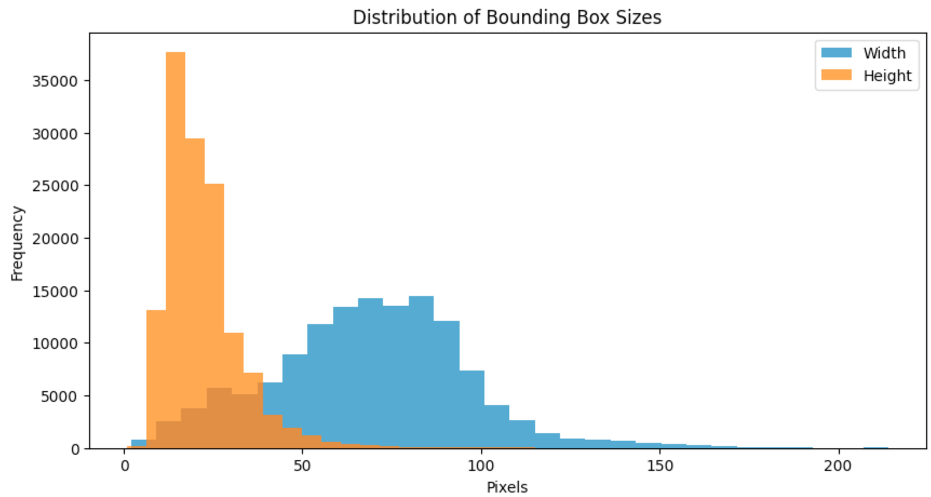
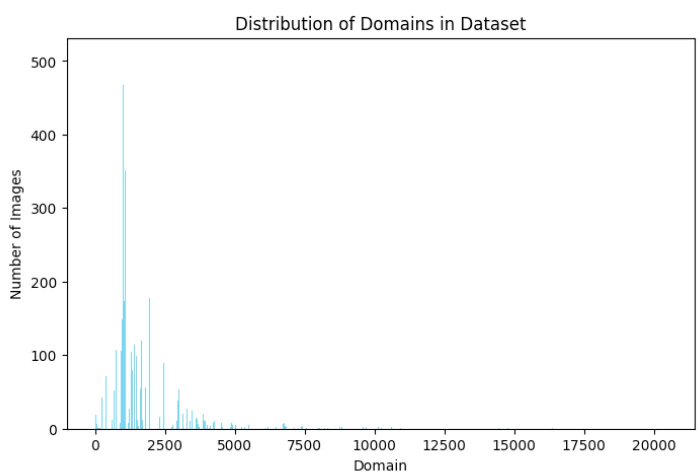


Figure 1: Fish Detection Example with Bounding Box

Figure 2: Distribution of Bounding Box Sizes

In Figure 2, the histogram visualizes the distribution of bounding box dimensions in the dataset. The blue histogram indicates that bounding box widths peak around 80 pixels, but there is significant variation. The orange histogram shows that bounding box heights peak around 20 pixels, with a strong skew toward smaller values. Thus, most of them are small bounding boxes which shows that models need to be optimized for detecting small objects in complex backgrounds.

Figure 3 shows the distribution of images across different domains which indicates a highly imbalanced dataset. The majority of domains are between 0 to 5000, while a small number of images have domains higher than 5000. This suggests potential domain shifts that could impact model generalization.



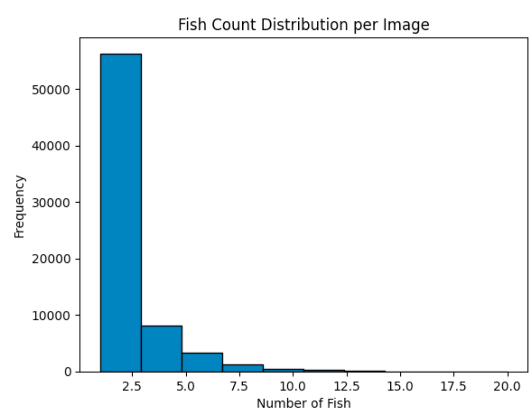


Figure 3: Distribution of Domains in the Dataset

Figure 4: Fish Count Distribution per Image

The histogram in Figure 4 represents the number of fish detected per image. The majority of images contain 1 to 3 fish, with a sharp drop-off for higher counts. A few images contain up to 10 fish, indicating potential outliers or unique scenarios. This means that most images feature sparse fish populations. It might require models to perform well on low-density detection while still considering occasional high-density cases.

From the above observations, the dataset primarily consists of small bounding boxes, with widths peaking at around 80 pixels and heights at around 20 pixels. The fish count per image is typically low, mostly between 1 to 3 fish per image. The dataset exhibits domain imbalances, which could impact model performance across different environments. Specialized detection strategies for small objects are necessary to handle the dataset's characteristics effectively.

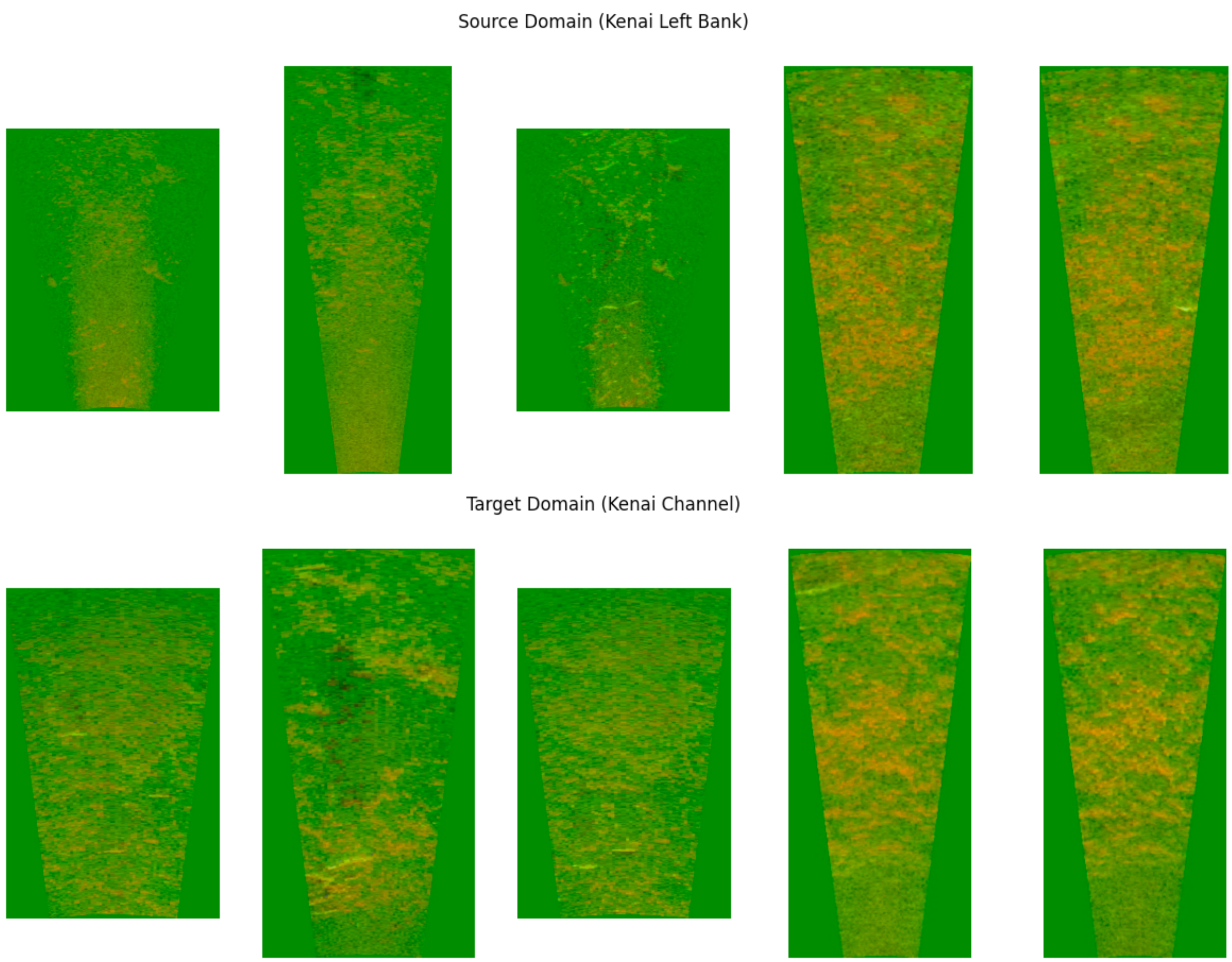


Figure 5: Visual Comparison of Source and Target Domains

To analyze domain differences, I visualized images from both domains. The images above provide a comparison of the source domain (Kenai Left Bank) and the target domain (Kenai Channel). The source domain images have a more uniform background with distinguishable fish regions. The visibility is clearer in some figures while other figures contain very insignificant noise. The green and orange color distributions indicate relatively stable sonar imaging conditions. The target domain shows more different background textures and contrast shifts.

Fish visibility is inconsistent, with some images showing lower contrast between fish and the background.

In conclusion, the target domain contains more background noise compared to the source domain. Differences in brightness and contrast mean that the sonar imaging conditions were different across domains. Fish distribution appears to be uneven so the model needs to correct for different densities. Also, fish could overlap, so it is harder for models to differentiate between them.

To address these challenges, YOLO (You Only Look Once) was selected as the object detection model due to its real-time detection effectiveness and strong performance in small-object localization. The model is first trained on the source domain labelled dataset with labelled data using transfer learning to fine-tune pre-trained weights for sonar images. To further minimize the domain shift, pseudo-labeling can be utilized. The learned model predicts the unlabeled target domain data and detections of high confidence are utilized sequentially as pseudo-labels to incrementally update the model. This process allows the model to learn incrementally to conform to the target domain's unique features. Training the YOLO-based detection model on the Source Train labelled dataset and validating it on the Target Test dataset to approximate domain shift impacts comes next. To ensure generalization, pseudo-labeling techniques will iteratively update the model with data from the target domain. Data augmentation strategies such as brightness changes and contrast adjustment will also be used to better reduce domain variance. This exploratory data analysis (EDA) provides a good foundation to generate a robust fish detection model.