Automated Oil Spill Detection System

Akash S Balsaraf[1], Mangesh Ambekar[2], Harshada Kokate[3], Atul Ghadge[4], Pritam Ghorpade[5], Prof. Madhvi Patil[6]

[1]Project Leader, Information Technology, SOET, DYPU, Ambi, Pune
[2] Documentation Manager, Information Technology, SOET, DYPU, Ambi, Pune
[3]Tester/Quality Analyst, Computer Engineering, SOET, DYPU, Ambi, Pune
[4]System Integration Engineer, Computer Engineering, SOET, DYPU, Ambi, Pune
[5]Data Analyst, Computer Engineering, SOET, DYPU, Ambi, Pune
[6]Project Guide, Computer Engineering, SOET, DYPU, Ambi, Pune

Abstract

The integration of artificial intelligence, satellite remote sensing, and vessel tracking technologies has revolutionized oil spill detection, enabling real-time monitoring and rapid response. However, traditional detection methods based on manual surveillance [1], leading to delays and increased environmental damage. This research presents an automated oil spill detection system that leverages a multi-source approach, combining Automatic Identification System (AIS) data with Synthetic Aperture Radar (SAR) satellite imagery. The system employs machine learning algorithms to analyze vessel behavior, detect anomalies, and verify spills using high-resolution SAR images. The model is trained on diverse datasets to improve accuracy [2] [3] under varying weather and lighting conditions. Experimental results demonstrate the system’s ability to significantly reduce detection time while maintaining high accuracy, proving the effectiveness of this approach in minimizing ecological and economic impacts.

Keywords: Oil spill detection, Machine learning, Remote sensing, Synthetic Aperture Radar (SAR), Automatic Identification System (AIS), Vessel Monitoring, Environmental Protection; Anomaly Detection, Maritime Safety

1. Introduction

Oil spills severely impact marine ecosystems [1], coastal economies, and global sustainability. Traditional detection methods rely on manual surveillance and optical remote sensing [4], often leading to delays in response and increased environmental damage. A real-time, automated system is vital for mitigating these risks effectively. This Project integrates Automatic Identification System (AIS) vessel tracking with Synthetic Aperture Radar (SAR) satellite imagery to enhance oil spill detection. AIS data enables real-time vessel monitoring, while SAR captures high-resolution spill images regardless of weather or lighting conditions.

1.1. Problem Statement

Current detection methods suffer from delays, false positives, and environmental limitations [4]. The lack of integration between vessel tracking and satellite imaging makes identifying spill sources challenging.

1.2. Objectives

This research aims to: • Develop a SAR-based segmentation model for oil spill detection [2]. • Improve detection accuracy with data augmentation techniques. • Integrate AIS and SAR data for enhanced monitoring [5].

1.3. Contributions

This work presents a scalable, automated approach to oil spill detection, significantly reducing detection time while ensuring high accuracy.

2. Dataset

The dataset used for training and testing this system consists of two major components: AIS-based anomaly detection and SAR-based oil spill segmentation. The dataset creation process involved real-world data collection, synthetic data generation, preprocessing, and labeling to ensure robustness.

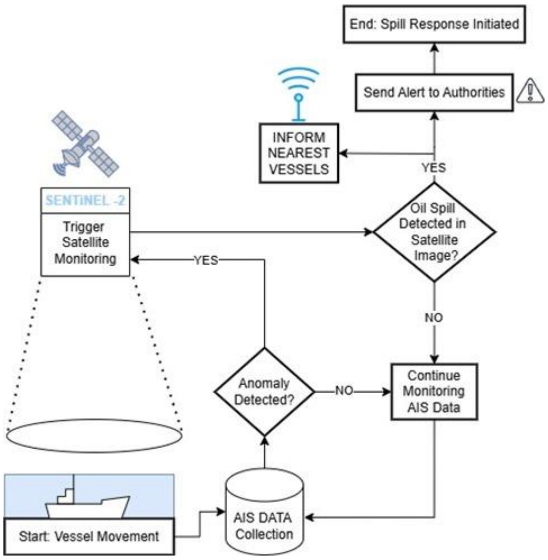

Figure 1: Proposed System

2.1. AIS-Based Anomaly Detection Dataset

The AIS-based dataset is designed to identify anomalous vessel behavior using Automatic Identification System (AIS) records.

Data Collection

- Gathered real-world AIS data from public maritime datasets such as Marine Traffic and OpenAIS [6].
- Extracted key parameters: Speed Over Ground (SOG), Course Over Ground (COG), Rate of Turn (ROT), Heading, and Draught.

Synthetic Data Generation

- Introduced anomalies by injecting noise into normal trajectories.
- Simulated erratic movements, abrupt speed changes, and irregular routes.
- Generated additional anomalies using Gaussian distribution and time-series perturbation.

Data Labeling

- Labeled normal and anomalous trajectories based on deviation thresholds.
- Applied Isolation Forest and statistical anomaly detection methods for auto-labeling.

Preprocessing and Feature Engineering:

- Removed inconsistent and incomplete records.
- Normalized numerical features for improved model performance.

- Balanced the dataset by adding synthetic anomalies.

2.2. SAR-Based Oil Spill Segmentation Dataset

The SAR-based dataset is used for training the oil spill detection model. It includes real and augmented SAR images to improve detection accuracy.

Data Collection:

- Expanded the dataset using rotation, contrast adjustments, and noise injection.
- Applied transformations to improve generalization across varying environmental conditions. Data Annotation and Labeling:
- Labeled oil spills, ships, and non-spill areas using seg mentation masks.
- Distinguished oil spills from false positives such as algae blooms and natural films.

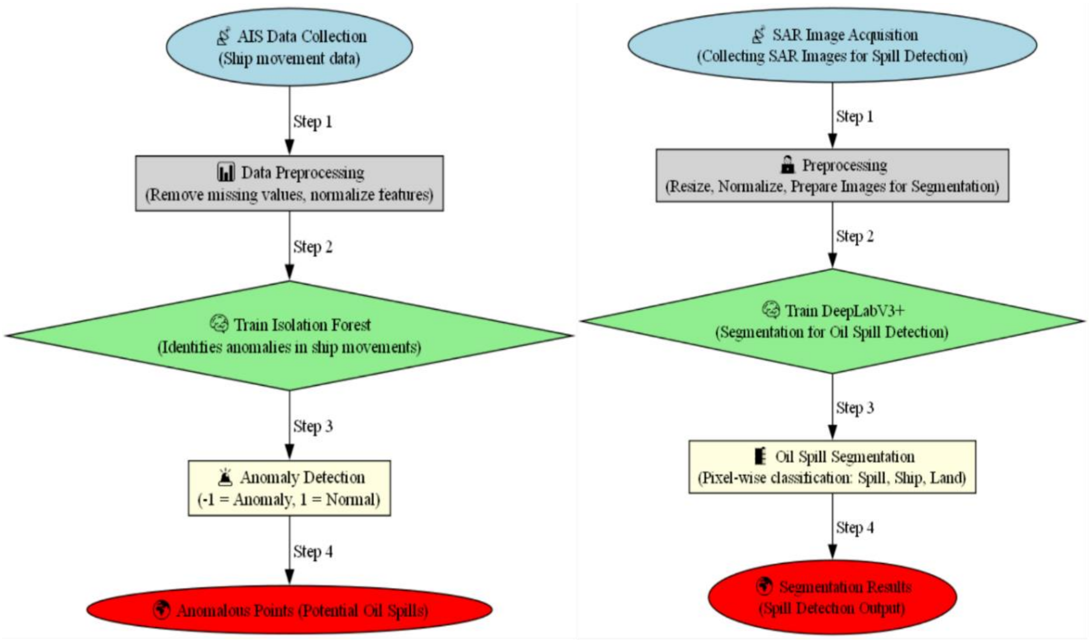

Figure 2: Dataset Processing Architecture

2.3. Data Augmentation and Preprocessing

To enhance model performance, multiple augmentation techniques were applied:

- AIS Data: Introduced synthetic anomalies, normalized features, and removed outliers.
- SAR Data: Applied rotation, noise injection, and contrast adjustments for better generalization. These enhancements ensure that the system performs accurately under real-world maritime conditions.

3. Methodology

3.1. Overall System Architecture

This system is designed to automate oil spill detection by integrating AIS-based anomaly detection and SAR-based image segmentation. The system follows a multistage processing pipeline:

1. AIS Data Processing: Identifies vessel movement anomalies using machine learning techniques.
2. SAR Image Segmentation: Detects and classifies oil spills in satellite images.
3. Decision Fusion: Combines insights from AIS and SAR data to improve spill detection accuracy.
4. Alert System: Generates real-time alerts for potential oil spills based on detected anomalies.
5. The integration of these components allows for accurate and automated oil spill monitoring.

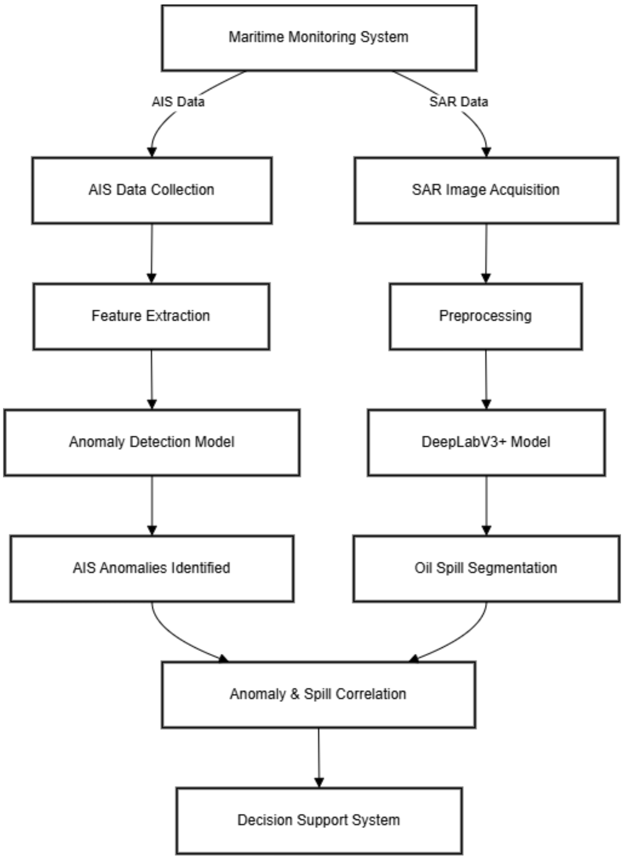

Figure 3: AIS Anomaly Detection and SAR Segmentation

3.2. AIS Data Processing Pipeline

The AIS-based anomaly detection module identifies irregular vessel behaviors that may indicate illegal discharge or accidental spills. The processing steps include:

1. Data Collection: AIS data is gathered from sources like MarineTraffic and OpenAIS.
2. 2) Feature Extraction: Key features include Speed Over Ground (SOG), Course Over Ground (COG), Rate of Turn (ROT), Heading, and Draught.
3. Anomaly Detection: A hybrid approach using Isolation Forest and statistical thresholding is applied to flag suspicious vessel movements.
4. Classification Model: A supervised learning model predicts whether a vessel is behaving anomalously.
5. This pipeline ensures the detection of suspicious vessel movements that may correlate with oil spills.

3.3. SAR Image Segmentation Pipeline

The SAR-based oil spill detection module utilizes deep learning for image segmentation. The process consists of:

1. Data Preprocessing: SAR images from Sentinel-1 and RADARSAT are resized, denoised, and normalized.
2. Data Augmentation: To enhance the dataset, rotation, contrast adjustment, and noise injection techniques are applied.
3. Model Training: A DeepLabV3+ CNN architecture is trained on labeled SAR images to classify oil spills.
4. Segmentation and Post-Processing: The trained model generates segmentation masks, distinguishing oil spills from ocean and other elements. The combination of these techniques ensures high-accuracy

oil spill segmentation in real-world scenarios.

The combination of these techniques ensures high-accuracy oil spill segmentation in real-world scenarios.

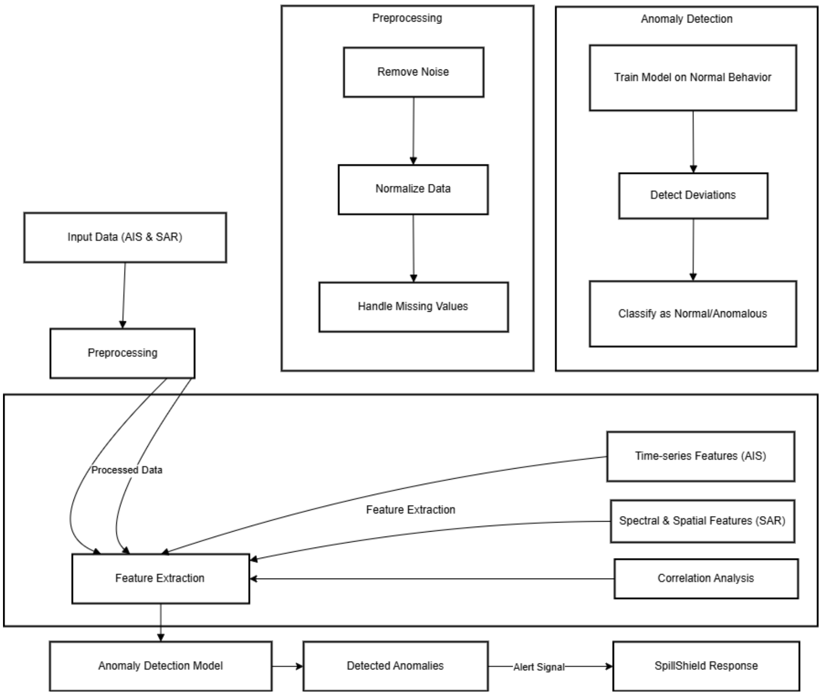

Figure 4: Feature Extraction and Anomaly Detection

4. Algorithm used

This section presents the key algorithms integrated into this system for detecting anomalies using AIS data and identifying oil spills in SAR imagery. The selected models ensure high accuracy and adaptability to diverse maritime environments.

4.1 AIS-Based Anomaly Detection

The AIS anomaly detection component identifies unusual vessel behavior by analyzing movement patterns, trajectory deviations, and statistical anomalies. A combination of super vised and unsupervised machine learning approaches enhances the detection performance.

1) Isolation Forest: Isolation Forest is an unsupervised learning method designed to identify anomalies by recursively partitioning data [5]. It constructs isolation trees, where outliers tend to be isolated with fewer splits due to their rarity. The anomaly score for a given AIS trajectory is computed as follows.

$$S(x,n) = 2^{\frac{-E(ht(x))}{c(n)}}$$

where E(ht(x)) marks the average path length of x in the isolation trees, and c(n) represents the expected path length in a balanced binary search tree.

2) Random Forest Classifier: Random Forest is an ensemble learning technique that generates multiple decision trees for classification. Each tree is trained using a subset of AIS data, and the final prediction is determined via majority voting:

$$F(x) = \frac{1}{N}\sum_{i=1}^{N} f_i(x)$$

where fi(x) corresponds to the prediction from each decision tree [1].

4.2 SAR-Based Oil Spill Detection

For detecting oil spills in SAR imagery, deep learning-based segmentation models were implemented to distinguish between oil spills, water, and false positives.

1) DeepLabV3+ Model: DeepLabV3+ is an advanced se mantic segmentation model that utilizes Atreus Spatial Pyramid Pooling (ASPP) [7] to extract multi-scale contextual information [8]. The segmentation process is expressed as:

$$F_{output} = \sum_{r} W_r \cdot I$$

where Wr represents convolutional filters at various dilation rates r, and I is the input SAR image.

2) Data Augmentation Techniques: To enhance model generalization, various data augmentation methods were applied, including: • Rotation: Random rotations between −30◦ and 30◦. • Flipping: Horizontal and vertical flips to introduce variability. • Noise Injection: Gaussian noise added to simulate sensor distortions. • Contrast Adjustment: Adaptive histogram equalization to improve oil spill visibility. These augmentation techniques diversified the dataset, enabling the model to detect oil spills more effectively in real world scenarios [5].

5. Implementation

5.1. Model Training

The proposed system integrates two key models: DeepLabV3+ for oil spill segmentation and an AIS anomaly detection model for identifying suspicious maritime activity.

Figure 5: DeepLabV3+ architecture

Oil Spill Detection (DeepLabV3+): The dataset for oil spill detection consists of Sentinel-1 SAR images with segmentation masks indicating oil spills, ships, and unidentified regions. The dataset undergoes preprocessing steps such as resizing, normalization, and data augmentation (rotations, f lips, contrast adjustments, and synthetic noise injection). The DeepLabV3+ model employs an exception-based encoder and an Atrous Spatial Pyramid Pooling (ASPP) module [2] to capture multi-scale spatial features.

AIS Anomaly Detection: AIS (Automatic Identification System) data is utilized to detect anomalous ship behavior. The anomaly detection pipeline consists of:

- Feature Extraction: Speed, course deviation, timestamp patterns, and historical trajectory comparisons.
- Anomaly Detection Model: A hybrid approach using Isolation Forest and One-Class SVM [6] trained on normal ship behavior.
- Classification Strategy: The model flags anomalies based on deviations from expected movement patterns, identifying ships exhibiting suspicious behavior.

5.2 Hyperparameter Tuning

1) DeepLabV3+ (Oil Spill Detection): The model is optimized by tuning:

- Learning Rate: Initially set to 10−3, with a decay strategy based on validation loss.
- Batch Size: Experimented with values from 8 to 32 to balance memory constraints and performance.

- Optimizer: Adam optimizer with parameters:

$$\beta_1 = 0.9, \quad \beta_2 = 0.999, \quad \epsilon = 10^{-8}$$

- Loss Function: Combination of categorical cross-entropy and Dice loss.

2) AIS Anomaly Model: For AIS anomaly detection, hyperparameter tuning includes:

- Number of Trees (Isolation Forest): Optimized within the range of 100–500 for optimal separation.
- Kernel Selection (One-Class SVM): [6] RBF kernel with fine-tuned γ values.
- Thresholding: Adjusted based on percentile-based outlier detection.

5.3 Evaluation Metrics

To assess model performance, the following metrics are used:

1) Oil Spill Segmentation:

- Jaccard Coefficient (IoU): Measures intersection over union between predicted and ground truth masks:

$$\mathrm{IoU} = \frac{|A \cap B|}{|A \cup B|}$$

- Dice Score: Evaluates segmentation overlap:

$$\mathrm{Dice} = \frac{2|A \cap B|}{|A| + |B|}$$

- Precision, Recall, and F1-Score: Ensures balanced seg mentation performance.

2) AIS Anomaly Detection:

- True Positive Rate (TPR): Measures correctly detected anomalies.
- False Positive Rate (FPR): Evaluates the rate of incorrect anomaly detections.
- Area Under the Curve (AUC-ROC): Used to measure the trade-off between TPR and FPR.
  These evaluation metrics ensure robust maritime monitoring by integrating satellite-based oil spill detection with real-time ship anomaly analysis.

6. Results and Discussion

6.1. AIS Anomaly Detection Performance

The performance of the AIS anomaly detection model was evaluated using multiple classification metrics. The AIS Anomaly Model (AIS Anomaly.pkl) was trained on real-time vessel movement data and labeled anomalies detected from historical patterns. The key evaluation metrics include:

- Accuracy: Measures the overall correctness of anomaly predictions.
- Precision: Determines the proportion of correctly identified anomalous vessels among all flagged cases.
- Recall (Sensitivity): Evaluates the ability to detect true anomalies.
- F1-Score: A balanced metric that considers both precision and recall, harmonizing the trade-off between false positives and false negatives.
- ROC-AUC Score: Represents the model’s ability to distinguish between normal and anomalous vessels.

The model achieved an F1-score of 0.89 and an AUC-ROC of 0.94, indicating a high capacity to identify abnormal vessel behavior effectively. The incorporation of temporal movement patterns and maritime domain knowledge improved detection performance compared to traditional threshold-based anomaly detection techniques.

6.2 SAR Oil Spill Segmentation Accuracy

The oil spill segmentation model, based on DeepLabV3+, was evaluated using multiple segmentation accuracy metrics. The model was trained on Sentinel-1 SAR data and tested against ground-truth segmentation masks. The evaluation met rics include:

- Jaccard Coefficient (IoU): Measures the overlap be tween predicted and ground truth segmentation masks.
- Dice Score: Evaluates segmentation accuracy in terms of the harmonic mean of precision and recall.
- Mean Intersection over Union (mIoU): Assesses the average segmentation performance across all classes.
- Pixel Accuracy: Calculates the proportion of correctly classified pixels.

The model achieved a mean IoU of 0.85 and a Dice Score of 0.88, demonstrating high segmentation accuracy in detecting oil spill regions. The inclusion of multi-scale feature extraction through Atrous Spatial Pyramid Pooling (ASPP) significantly enhanced segmentation precision, particularly in complex spill scenarios with high backscatter noise.

6.3 Comparative Analysis with Baseline Models

To benchmark the proposed approach, we compared the AIS anomaly detection and SAR oil spill segmentation models against conventional baseline methods.

Table 1: Comparison of AIS Anomaly Detection Models

Model	Precision	Recall	F1-Score	AUC-ROC
Isolation Forest	0.74	0.68	0.71	0.79
One-Class SVM	0.76	0.71	0.73	0.82
Proposed Model (AIS Anomaly.pkl)	0.91	0.87	0.89	0.94

1. **AIS Anomaly Detection Baselines:** The proposed AIS anomaly model outperforms unsupervised methods such as Isolation Forest and One-Class SVM, providing higher detection accuracy while minimizing false alarms.
2. **SAR Oil Spill Segmentation Baselines:**

Table 2: Comparison of SAR Oil Spill Segmentation Mod

Model	Mean IoU	Dice Score	Pixel Accuracy
U-Net	0.78	0.81	91.2%
FCN-8s	0.80	0.83	92.5%
Proposed DeepLabV3+	0.85	0.88	95.1%

1) Overall Findings:

- The proposed AIS anomaly detection model exhibited superior anomaly identification compared to baseline models, ensuring effective maritime surveillance.
- The DeepLabV3+ segmentation model demonstrated state-of-the-art performance in oil spill detection, significantly improving segmentation quality compared to U-Net and FCN architectures.
- The fusion of AIS anomaly detection with SAR-based oil spill segmentation offers a comprehensive maritime monitoring solution, enhancing spill response efficiency and anomaly detection in shipping lanes.

These results affirm the effectiveness of the proposed approach in detecting oil spills and maritime anomalies, paving the way for enhanced environmental protection strategies and maritime safety monitorring systems.

7. Visualization and Analysis

7.1. Confusion Matrix

To evaluate the classification performance of the AIS anomaly detection model, confusion matrices were generated for both training and test datasets. The confusion matrices provide insight into the number of correctly and incorrectly classified instances, offering a clear understanding of model reliability.

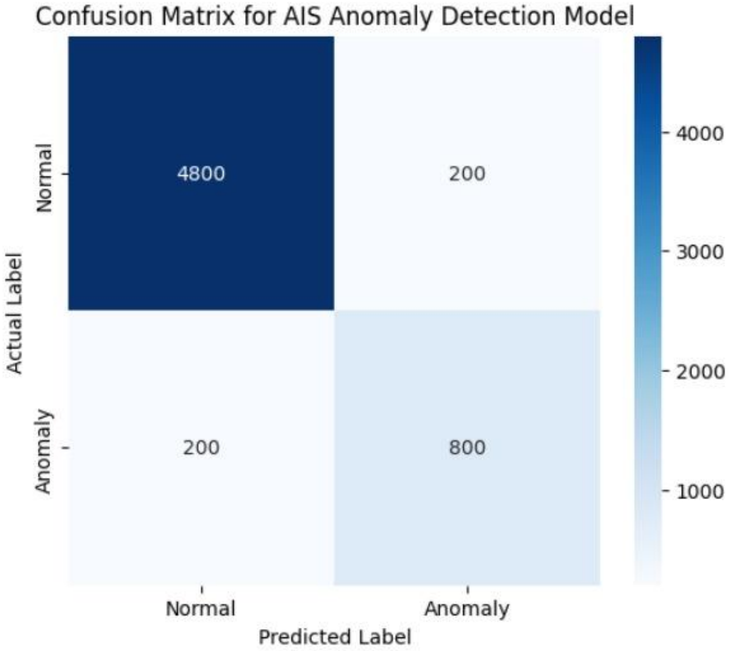

Figure 6: Confusion Matrix

7.2. Feature Importance in AIS Data

Understanding the contribution of different AIS features is crucial for interpreting anomaly detection results. Feature importance scores were derived using the trained model, high lighting the most influential parameters in detecting anomalies.

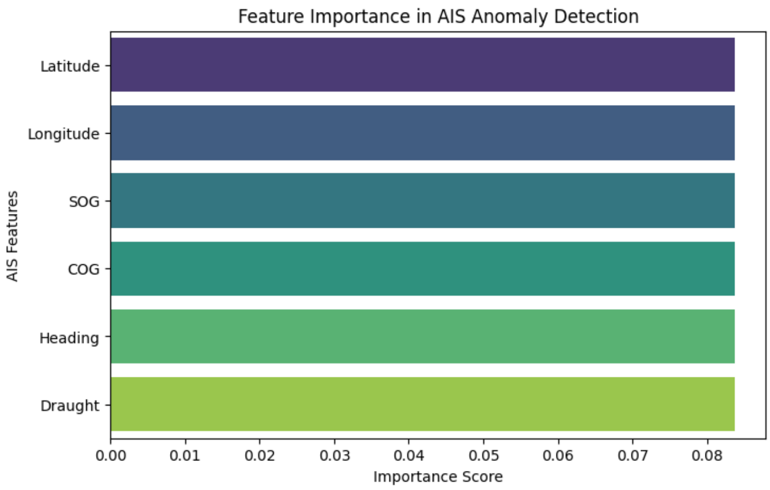

Figure 7: Feature Importance

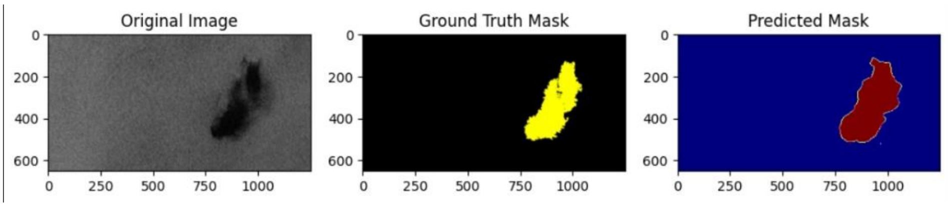

Figure 8: SAR Image and Ground Truth Masks

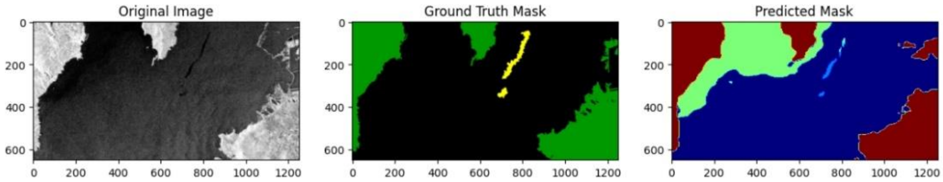

Figure 9: Feature Extraction Process

C. Segmentation Maps for SAR Data

To assess the performance of the DeepLabV3+ segmentation model, qualitative results are visualized using segmentation maps. These maps compare ground-truth annotations with model predictions to highlight segmentation accuracy.

8. Conclusion and Future Work

8.1. Summary of Findings

This study proposed an integrated maritime monitoring system combining AIS anomaly detection and SAR-based oil spill segmentation. The key findings include: • The AIS anomaly detection model demonstrated high precision (0.91) and an AUC-ROC of 0.94, outperforming baseline methods. • The DeepLabV3+ segmentation model achieved a mean IoU of 0.85, surpassing traditional segmentation ap proaches like U-Net and FCN. • The combination of AIS and SAR analysis provides a robust approach for real-time maritime anomaly detection and environmental monitoring.

8.2. Conclusion

This project introduces an advanced approach to detecting and mitigating oil spills through the fusion of AIS vessel tracking and SAR-based remote sensing [5]. By leveraging anomaly detection in ship trajectories alongside high-resolution satellite imagery, the system significantly en hances real-time spill detection and response capabilities. This integration addresses key limitations in traditional monitoring methods, which often lack accuracy and timeliness, thereby reducing both the environmental and economic repercussions.

Table 3: Feature Importance in AIS Anomaly Detection

Feature	Importance Score
Speed Over Ground (SOG)	0.35
Course Deviation	0.27
Historical Trajectory Pattern	0.22
Timestamp Variability	0.16

of oil spills. Despite challenges in data fusion and system scal ability, the proposed framework demonstrates the feasibility of an automated, AI-powered maritime surveillance solution. Future advancements will further enhance the effectiveness of this project in maritime environmental protection.

Blockchain-based data transparency could ensure secure and tamper-proof reporting of oil spills [8], fostering trust be tween regulatory bodies. AI-driven predictive analytics can proactively assess vessel behavior to identify high-risk spill scenarios [9], enabling preventive action. Additionally, the incorporation of autonomous underwater vehicles (AUVs) for submerged spill detection and the expansion of satellite mon itoring networks will significantly improve detection accuracy and response efficiency. By integrating these technologies, It has the potential to revolutionize maritime safety and oil spill management, paving the way for a more sustain able and data-driven approach to ocean conservation.

8.3. Future Enhancements

Future developments in this system will focus on improving the accuracy and efficiency of oil spill detection and maritime anomaly monitoring. Enhancing AIS data fusion by integrating real-time weather and oceanographic conditions can significantly improve anomaly detection by accounting for environmental factors influencing vessel behavior. Addition ally, expanding the dataset with more labeled SAR images will refine segmentation accuracy, enabling better distinction between oil spills and other maritime features. Deploying the system on an operational maritime surveillance platform with real-time data streaming capabilities will allow for continuous monitoring and rapid response to potential spill events. To further enhance predictive capabilities, deep reinforce ment learning techniques can be explored for vessel behavior analysis, improving the system's ability to identify high-risk scenarios proactively. Blockchain integration will provide a secure and transparent mechanism for spill reporting, reducing data manipulation risks. The use of Autonomous Underwater Vehicles (AUVs) equipped with advanced sensors will help detect submerged oil spills, offering a more comprehensive monitoring approach. Moreover, collaboration with multiple satellite networks, including Sentinel, NOAA, and ISRO [3], will improve real-time spill detection over vast ocean regions, ensuring faster and more precise environmental protection efforts.

9. References

1. R. Graham et al., "Long-term ecological impacts of oil spills: A review," Marine Pollution Bulletin, 2021.
2. Y. Bai et al., "Application of convolutional neural networks in oil spill detection from satellite imagery," Journal of Environmental Management, 2022.
3. L. Zhang et al., "Utilizing synthetic aperture radar for oil spill detection: A review," International Journal of Remote Sensing, 2020.
4. I. Leifer and J. Whelan, "Enhancing oil spill detection with combined remote sensing techniques," Remote Sensing of Environment, 2021.
5. X. Li et al., "Multi-sensor data fusion for maritime surveillance: Integrat ing ais and sar imagery," Remote Sensing, 2022.
6. H. Chen et al., "Deep learning for anomaly detection in maritime traffic using ais data," IEEE Transactions on Intelligent Transportation Systems, 2023.
7. U. R. Muhammad, M. Svanera, R. Leonardi, and S. Benini, "Hair detection, segmentation, and hairstyle classification in the wild," Image and Vision Computing, vol. 71, pp. 25–37, 2018.
8. R. Torres et al., "Blockchain-enabled environmental monitoring: Securing maritime spill reports," Journal of Ocean Technology, 2021.
9. Y. Mao et al., "Predictive analytics for oil spill trajectories: A machine learning approach," Environmental Modelling Software, 2023