Design of a Deep Learning-Based Ship Navigation Status Monitoring Model and Validation of Its Anomaly Detection Performance

Huawei Su
School of Public Order Management, Zhengzhou Police University, Zhengzhou 450000, Henan, China
Corresponding author: Huawei Su (e-mail: suhuawei629@163.com)

ABSTRACT Addressing the limitations of traditional ship navigation status monitoring, including manual inspection dependence, delayed anomaly identification, high false alarm rates, and inadequate adaptability to complex maritime environments, this paper systematically investigates a deep learning-based framework for navigation status monitoring and abnormal behavior detection. As modern maritime surveillance increasingly relies on electromagnetic sensing networks and wireless communication technologies for reliable information acquisition and transmission, intelligent monitoring has become essential for navigation safety. First, based on ship dynamics, deep learning, and computer vision theories, a comprehensive technical framework integrating data acquisition, preprocessing, feature extraction, state modeling, and anomaly detection is established to define core navigation state indicators and abnormal behavior categories. Second, by fusing Automatic Identification System (AIS) data, navigation sensor data, and video surveillance information, a multisource dataset (MSS-2024) covering 15 navigation states and 8 representative abnormal behaviors is constructed. Third, a dual-branch Transformer-CNN fusion model incorporating attention mechanisms and multi-scale feature fusion is developed to enhance state representation and anomaly recognition under complex conditions. Finally, extensive experiments based on real-world vessel navigation data validate the proposed approach, demonstrating its effectiveness in improving recognition accuracy, response efficiency, and robustness, while providing technical support for intelligent maritime electromagnetic monitoring and navigation safety assurance.

INDEX TERMS deep learning, ship navigation status monitoring, anomaly behavior detection, multi-source data fusion, maritime electromagnetic sensing

I. INTRODUCTION

A. RESEARCH BACKGROUND

As the core carriers of maritime transportation, the navigational safety of vessels directly impacts the safety of human life and property, the marine ecological environment, and the stability of global trade [1-3]. The dynamic changes in navigational states have intensified, and the risks of abnormal behaviors caused by equipment failures, operational errors, and malicious manipulation have significantly increased [4,5]. According to statistics from the International Maritime Organization (IMO), in 2022, collisions, groundings, and capsizings caused by failure to promptly identify abnormal navigation states accounted for 53.7% of global maritime accidents, resulting in direct economic losses exceeding $28 billion [6].

B. CURRENT STATE OF DOMESTIC AND INTERNATIONAL RESEARCH

Traditional anomaly detection technologies primarily rely on rule-based matching and threshold-based judgment. Internationally, the U.S. Coast Guard employs rule-based AIS data parsing methods to identify non-compliant vessel navigation [7]; Germany's Bosch Group developed a threshold-based vessel equipment anomaly monitoring system that identifies faults by comparing real-time data against standard thresholds [8]. Domestically, the Water Transport Research Institute of the Ministry of Transport proposed an expert rule-based method for judging vessel anomaly behaviors, covering six typical anomalies including speeding and route deviation [9]. Shanghai Jiao Tong University optimized a threshold adaptive adjustment algorithm to enhance recognition adaptability across different navigation scenarios [10-12].

C. RESEARCH CONTENT AND TECHNICAL APPROACH

1) Research Content

Establishment of vessel navigation state and anomaly behavior classification system: Define core monitoring dimensions for vessel navigation states (motion parameters, equipment status, environmental parameters), categorize anomaly behavior types, and formulate identification criteria; Multi-source data fusion and dataset construction: Integrate AIS data, sensor data, and video surveillance data to address data heterogeneity, constructing standardized datasets for vessel navigation states and anomaly behaviors; Deep Learning Monitoring Model Design: Develop a multi-scale feature extraction and attention-enhanced module based on a Transformer-CNN fusion architecture to enhance state modeling and anomaly detection capabilities; Model Performance Validation: Evaluate model performance through comparative experiments and real-vessel testing across four dimensions: recognition accuracy, response speed, robustness, and generalization capability.

2) Technical Approach

The research follows a logical progression of "Theoretical Framework - Data Preparation - Model Design - Performance Validation." Theoretical Layer: Systematically reviews theories related to ship dynamics, deep learning, and data fusion to establish a technical framework. Data Layer: Collects multi-source ship navigation data, performs preprocessing, annotation, and fusion to construct the MSS-2024 dataset. Model Layer: Designs a dual-branch Transformer-CNN fusion model to optimize feature extraction and anomaly detection algorithms. The validation layer establishes an experimental platform to verify model performance through simulation experiments, comparative studies, and real-vessel testing, forming a closed-loop research cycle.

II. THEORETICAL FOUNDATIONS

A. CORE THEORIES OF SHIP NAVIGATION STATUS MONITORING

1) Dimensions of Ship Navigation Status Monitoring

Ship navigation status encompasses three core dimensions: motion, equipment, and environment. Key monitoring parameters for each dimension are as follows: Motion Parameters: Speed, heading, draft depth, roll angle, pitch angle, position coordinates (latitude/longitude), course deviation, etc. Equipment Status Parameters: Main engine speed, propulsion system power, rudder angle, navigation equipment accuracy, communication equipment status, power system temperature/pressure, etc. Environmental Parameters: Wind speed, wind direction, wave height, visibility, current velocity, water temperature, barometric pressure, etc [13-15].

B. CLASSIFICATION CRITERIA FOR ABNORMAL BEHAVIOR

Based on IMO safety regulations and shipping practices, vessel abnormal behavior is categorized into 8 types as follows: Motion Abnormalities: Excessive speed navigation, low-speed drifting, significant track deviation, sudden heading changes, abnormal turning. Equipment Abnormalities: Main engine failure, rudder malfunction, navigation equipment anomalies, propulsion system overheating. Operational Anomalies: Improper berthing, illegal anchoring, incorrect collision avoidance maneuvers, and boundary-crossing navigation [16,17].

C. CORE DEEP LEARNING THEORIES

1) Time-Series Data Processing Models

LSTM (Long Short-Term Memory Network): Resolves the vanishing gradient problem in RNN models through forget gates, input gates, and output gates, suitable for feature extraction and state prediction in AIS time-series data [18]; Transformer Model: Implements global dependency modeling for time-series data via self-attention mechanisms, capturing long-term correlation features of vessel navigation states. Its core structure comprises encoder and decoder layers, enhancing feature extraction comprehensiveness through multi-head attention mechanisms [19].

2) Spatial Feature Extraction Models

CNN (Convolutional Neural Network): Automatically extracts spatial features through convolutional, pooling, and fully-connected layers. Suitable for target detection and state recognition in vessel video surveillance data. Typical models include ResNet and MobileNet, effectively capturing spatial features such as vessel appearance and deck operations [20,21].

3) Anomaly Detection Models

Autoencoder: Maps input data to a low-dimensional feature space via an encoder, then reconstructs the input through a decoder. It identifies anomalies by measuring reconstruction error, making it suitable for unsupervised anomaly detection scenarios [22]. GAN (Generative Adversarial Network): Composed of a generator and a discriminator, it generates realistic virtual samples through adversarial training, alleviating the scarcity of rare anomaly behavior samples [23-25].

D. MULTI-SOURCE DATA FUSION THEORY

Multi-source data fusion technology addresses the heterogeneity conflicts among AIS data, sensor data, and video data to achieve feature complementarity. Core fusion layers include: Data-layer fusion: Performs alignment, cleaning, and standardization of raw data to eliminate format and scale differences. Feature-layer fusion: Extracts key features from each data source and combines them into a unified feature vector through methods like feature concatenation and weighted fusion. Decision-layer fusion: Integrates decision outputs from multiple models using D-S evidence theory, Bayesian inference, and other methods to enhance recognition reliability [26].

E. MODEL PERFORMANCE EVALUATION METRICS

To comprehensively assess monitoring model and anomaly detection efficacy, the following core evaluation metrics are employed: Accuracy: The proportion of correctly identified samples relative to the total sample size, reflecting overall model recognition capability. Precision: The proportion of genuinely anomalous samples among those identified as anomalies, reflecting false positive control capability. Recall: The proportion of genuinely anomalous samples correctly identified, reflecting false negative control capability. F1 Score: The harmonic mean of precision and recall, comprehensively reflecting recognition performance. Response Time: The duration from data reception to output of recognition results, reflecting real-time capability. Robustness Metric: Variation in recognition accuracy across different sea conditions and navigation scenarios [27-30].

III. CONSTRUCTION OF VESSEL NAVIGATION STATUS AND ANOMALY BEHAVIOR DATASET

A. DATA SOURCES AND ACQUISITION PLAN

The dataset was constructed using a “real-vessel collection + simulation supplementation + public data integration” approach, with sources including: Field Collection: Six-month navigation data acquisition conducted on three vessels of different types (container ship, bulk carrier, tanker). Equipped with AIS devices, inertial navigation sensors, environmental sensors, and high-definition video cameras to capture real navigation states and simulated anomaly data; Simulation Supplementation: Utilized the ShipX navigation simulation platform to simulate extreme sea conditions and rare abnormal behaviors (e.g., sudden rudder failure, main engine malfunction) to supplement scarce samples. Public Data Integration: Integrated relevant samples from the MARIDA public dataset and IMO vessel accident case data to enrich data diversity. This study was approved by the Ethics Committee of Zhengzhou Police University, and was performed in accordance with the principles of the Declaration of Helsinki. All eligible participants signed an informed consent form. Note: Data collection has been approved by relevant shipping companies and maritime administrative departments, strictly complying with the International Convention for the Safety of Life at Sea (SOLAS) and domestic regulations on ship data collection, ensuring legal and compliant data sources without ethical disputes (Table 1).

TABLE 1. Acquisition Parameters of Multi-source Vessel Navigation Data

Data Type	Collection Equipment	Core Parameters	Sampling Frequency
AIS Data	Automatic Identification System (AIS)	Vessel MMSI, Latitude/Longitude, Speed, Heading, Draft, Vessel Type, Timestamp	1 time/30s
Sensor Data	Inertial Navigation System, Temperature/Pressure Sensors, Vibration Sensors	Roll Angle, Pitch Angle, Main Engine Speed, Propulsion Power, Rudder Angle, Equipment Temperature, Vibration Frequency	10 times/s
Environmental Data	Meteorological Sensors, Marine Environmental Monitoring Equipment	Wind Speed, Wind Direction, Wave Height, Visibility, Current Speed, Water Temperature	1 time/min
Video Data	High-Definition Network Camera (4K Resolution)	Deck Operations, Hull Attitude, Surrounding Navigation Environment, Anomalous Scenarios (e.g., Fire, Collision Precursor)	25 frames/s

B. DATA PREPROCESSING AND ANNOTATION

C. DATA PREPROCESSING WORKFLOW

Data Cleaning: Outliers in sensor data are removed using box plots, and missing values are filled via interpolation. Duplicate records and erroneous coordinates in AIS data are filtered out. Video data quality is enhanced through image denoising and enhancement algorithms. Data Alignment: Multi-source data is synchronized based on timestamps. Time differences between data with varying sampling frequencies are addressed using linear interpolation to ensure temporal consistency. Data Alignment: Multi-source data is synchronized based on timestamps. Time differences between data with varying sampling frequencies are addressed using linear interpolation to ensure temporal consistency. For AIS data with dynamic sampling frequencies, adaptive time alignment is implemented based on scenario switching signals (e.g., maneuvering start/end markers), ensuring compatibility of multi-frequency data in model training. Scenario-based Frequency Switching Logic: The AIS sampling frequency is automatically adjusted based on real-time navigation state parameters: Conventional navigation (uniform speed, straight line, open sea): 1 time/30s, balancing transmission efficiency and storage cost; Complex maneuvering (berthing, narrow waterway navigation, collision avoidance): 1 time/2s, capturing dynamic changes in vessel state; Abnormal warning (equipment parameter anomalies, sudden environmental changes): 1 time/1s, enhancing abnormal data capture precision.

D. DATA ANNOTATION AND CLASSIFICATION

Data annotation was completed using a “manual annotation + automatic annotation + expert review” approach: Status Annotation: Classified 15 normal navigation states (e.g., uniform straight-line navigation, uniform turning, berthing preparation) based on the vessel navigation status classification system. Anomaly Annotation: Annotated 8 categories of abnormal behaviors based on the abnormal behavior classification standard, specifying the occurrence time, duration, and severity level (minor, moderate, severe); Expert Review: Five captains with over 10 years of navigation experience and three marine engineering experts were invited to review and refine the annotation results, ensuring annotation accuracy. Quantitative Evaluation of GAN-Generated Samples: To verify the consistency between virtual samples and real anomaly distributions, two types of metrics are adopted: Statistical Feature Consistency: Compares 8 statistical features (mean, variance, skewness, kurtosis, maximum, minimum, median, standard deviation) of generated and real samples; Distribution Similarity: Quantifies distribution differences using KL divergence (Kullback-Leibler Divergence) and Wasserstein distance (closer to 0 indicates better consistency).

E. OVERVIEW OF THE DATASET (MSS-2024)

The final constructed vessel navigation state and anomaly behavior dataset (MSS-2024) details are as follows: Sample Size: Includes 150,000 time-series data samples (each containing 30 seconds of continuous data) and 50,000 video frame samples, with normal state samples accounting for 70% and anomaly behavior samples for 30%. State Coverage: Encompasses 15 normal navigation states and 8 abnormal behaviors, with rare abnormal samples (e.g., sudden engine failure, rudder malfunction) constituting 15% of abnormal samples; Scenario Diversity: Includes 6 environmental scenarios (calm seas, strong winds, heavy rain, dense fog, etc.) and 5 navigation scenarios (ocean voyages, coastal navigation, berthing, collision avoidance); Data Format: Time-series data stored in CSV format, containing parameter names, values, and timestamps; video data stored in JPG format, with annotation information recorded in XML format (including anomaly type, bounding box coordinates, and timestamp).

IV. DEEP LEARNING-BASED MONITORING MODEL DESIGN

A. OVERALL MODEL ARCHITECTURE

A dual-branch monitoring model (TC-Fusion Model) integrating Transformers and CNNs is designed. The overall architecture comprises five core modules: data input layer, multi-source feature extraction layer, feature fusion layer, state modeling layer, and anomaly identification layer, as detailed below:

B. DATA INPUT LAYER

Receives preprocessed multi-source data, including:

- Time-series data branch: Integrated time-series feature sequences from AIS data, sensor data, and environmental data (dimensions: time step × number of features).
- Spatial data branch: Video frame image data (dimensions: height × width × number of channels).

C. MULTI-SOURCE FEATURE EXTRACTION LAYER

Time-series feature extraction branch (Transformer-enhanced): Input embedding: Maps time-series data to a high-dimensional feature space, incorporating temporal information via positional encoding. Multi-head attention module: Employs an 8-head self-attention mechanism to capture long-term dependencies and global correlation features in time-series data, focusing on dynamic patterns of speed, heading, and equipment parameters. Complexity Optimization Module: To address the $O(n^2)$ complexity bottleneck in high-density data streams, a "local attention + sequence segmentation" strategy is adopted. Only the attention between each time step and its adjacent k steps ($k = 10$) is calculated, reducing the complexity to $O(nk)$. Feedforward Neural Network: Enhances nonlinear feature representation through two fully connected layers with ReLU activation functions. Residual Connections and Layer Normalization: Mitigates vanishing gradient issues and improves training stability. Complexity Quantification Analysis: For input sequence length n, the time complexity of the original Transformer is $O(n^2)$. When n increases from 30 to 240 (high-density data stream), the inference time increases from 87ms to 623ms. After optimization, the inference time for n=240 is reduced to 198ms, meeting real-time requirements ($\leq$0.5s).

D. FEATURE FUSION LAYER

Multi-source feature fusion achieved via "weighted feature concatenation + adaptive fusion": Feature weight allocation: Computes importance weights for temporal and spatial features via attention mechanisms, dynamically adjusting their fusion ratio. Feature concatenation: Concatenates weighted temporal and spatial features into a unified feature vector. Feature dimension reduction: Applies Principal Component Analysis (PCA) to reduce dimension of fused features, lowering computational load and accelerating inference.

E. STATE MODELING LAYER

Constructs a vessel navigation state model based on fused features, employing an LSTM network to learn dynamic state evolution patterns: State Encoding: Inputs fused features into the LSTM network, capturing temporal evolution characteristics through three LSTM layers. State Representation: Outputs a high-dimensional state vector describing comprehensive features of the vessel's current navigation state. State Prediction: Predicts the next 30 seconds of navigation state based on historical state vectors, providing a basis for anomaly alerts.

F. ANOMALY IDENTIFICATION LAYER

Employs a dual-output design of "classification identification + anomaly score calculation": Classification: Performs multi-class classification of anomaly types (8 anomaly classes + normal state) via a fully connected network with Softmax activation function. Anomaly Severity Calculation: Computes an anomaly severity score based on the reconstruction error from the autoencoder. An anomaly alert is triggered when the score exceeds a preset threshold. Result Fusion: Combines the classification result with the anomaly severity score to output the final identification result, enhancing recognition reliability.

G. MODEL TRAINING STRATEGY

1) Loss Function Design

A hybrid loss function optimizes model training, comprising: Classification Loss: Cross-entropy loss optimizes anomaly behavior classification results; Reconstruction Loss: Mean squared error (MSE) loss optimizes the autoencoder's state reconstruction accuracy; Regularization Loss: L2 regularization suppresses model overfitting. The total loss function formula is:

$$\text{Loss} = \alpha \cdot \text{Loss}_{CE} + \beta \cdot \text{Loss}_{MSE} + \gamma \cdot \text{Loss}_{L2}$$

where, $\alpha = 0.6$, $\beta = 0.3$, $\gamma = 0.1$ are loss weights determined through validation set tuning.

H. TRAINING PARAMETER SETTINGS

Optimizer: AdamW optimizer with an initial learning rate of 0.001, dynamically adjusted using a cosine annealing learning rate scheduling strategy; Batch Size: 32; Epoch: 100, employing an early stopping strategy (Patience=10) to prevent overfitting; Regularization Strategy: Employ Dropout (dropout rate=0.2) and weight decay (weight decay=0.0001) to suppress overfitting.

I. SMALL-SAMPLE HANDLING APPROACH

To address the scarcity of rare anomaly behavior samples, the following strategies are employed: Data Augmentation: Apply temporal data augmentation (time stretching, noise addition, data reordering) and image data augmentation (random cropping, rotation, flipping, brightness adjustment) to small-sample datasets; Transfer learning: Fine-tune the anomaly detection branch using model parameters pretrained on large-scale normal navigation data; Virtual sample generation: Employ GAN models to generate virtual time-series and video samples of rare anomalies, supplementing training data.

V. EXPERIMENTAL VALIDATION AND PERFORMANCE ANALYSIS

A. EXPERIMENTAL PLATFORM SETUP

1) Hardware Platform

Data Acquisition Devices: AIS receiver (Class A), Inertial Navigation System (IMU), high-definition network camera (4K resolution), meteorological and oceanographic environmental sensors; Simulation Platform: ShipX vessel navigation simulation system, MATLAB/Simulink vessel dynamics simulation module.

2) Software Platform

Operating System: Windows 11 Professional; Development Frameworks: Python 3.10, PyTorch 2.1, TensorFlow 2.15, OpenCV 4.9; Evaluation Tools: Scikit-learn, TensorBoard, MATLAB R2024a, Confusion Matrix Analysis Tool.

B. EXPERIMENTAL DESIGN

1) Experimental Group Setup

Five experimental groups were established to validate the proposed model's performance: Control Group 1: Traditional monitoring method (rule-based threshold judgment + manual feature extraction + SVM classification). Control Group 2: Single LSTM model (processing only time-series data). Control Group 3: Single ResNet50 model (processing only video data). Control Group 4: Mainstream fusion model (CNN-LSTM hybrid model). Experimental Group: The proposed TC-Fusion model (Transformer-CNN dual-branch fusion model).

C. EXPERIMENTAL DATA AND EVALUATION METRICS

Experimental Data: Utilized the MSS-2024 dataset, partitioned into training, validation, and test sets at a 7:2:1 ratio. The test set includes samples of various vessel types, environmental scenarios, and abnormal behaviors; Evaluation Metrics: Recognition accuracy, precision, recall, F1 score, response time, and robustness metrics.

D. EXPERIMENTAL RESULTS AND ANALYSIS

E. FUNDAMENTAL PERFORMANCE COMPARISON

The core performance metrics comparison across the five experimental groups is shown in the Table 2 below:

TABLE 2. Performance Comparison of Different Vessel Anomaly Detection Models

Experimental Group	Recognition Accuracy (%)	Precision (%)	Recall (%)	F1 Score	Response Time (ms)
Control Group 1	72.8	70.3	68.5	0.694	850
Control Group 2	83.5	81.2	79.6	0.804	120
Control Group 3	85.7	83.4	81.8	0.826	280
Control Group 4	91.2	89.5	88.7	0.891	210
Experimental Group	96.3	94.8	94.7	0.948	290

Quantitative results demonstrate: The experimental group achieved optimal performance across all recognition metrics, with an accuracy rate of 96.3%. This represents a 23.5 percentage point improvement over Control Group 1 and a 5.1 percentage point increase over Control Group 4, confirming that the dual-branch fusion architecture enhances anomaly detection precision via effective integration of temporal global dependency and spatial local key features. The recall rate reached 94.7%, while the false negative rate dropped to 2.8%, significantly outperforming other control groups. This demonstrates the proposed model's capability to effectively detect rare anomalies and anomalies in complex scenarios; The response time was 290ms (0.29 seconds), meeting real-time monitoring requirements ($\leq$0.5 seconds). Although slightly longer than Control Group 2, the overall recognition performance advantage is substantial; Single models (Control Groups 2 and 3) demonstrated lower recognition performance than fusion models (Control Group 4 and Experimental Group), validating the necessity of multi-source data fusion. Analysis: The complete model outperforms all ablated configurations, verifying the necessity of the hybrid loss function and dynamic feature fusion weights. The weights were optimized through 5 rounds of cross-validation, ensuring objectivity and rationality. Sensitivity Analysis of Weights: By adjusting the weight coefficients (±30%), the model's F1 score fluctuates only within the range of 1.8%-2.5% (0.926-0.971), indicating strong robustness to weight changes and stable performance.

F. SUPPLEMENTARY VERIFICATION FOR COMPLEX MANEUVERING SCENARIOS

To validate the effectiveness of dynamic AIS sampling frequency, a comparative experiment was conducted using sudden rudder failure data. The results show that the model's anomaly detection response time under 2s sampling frequency (0.21ms) is 75.9% faster than that under 30s frequency (0.87ms), fully meeting real-time monitoring requirements for complex scenarios.

G. ROBUSTNESS ANALYSIS

The robustness of the Experimental Group model across different environmental scenarios and vessel types is analyzed, with results presented in the Table 3 below:

TABLE 3. Robustness Test Results of the TC-Fusion Model Under Different Scenarios

Environment Scenario / Vessel Type	Normal Sea Condition	Strong Wind Weather	Heavy Rain Weather	Foggy Weather	Container Ship	Bulk Carrier	Oil Tanker	Robustness Index (Fluctuation Range)
Recognition Accuracy (%)	98.2	95.6	94.3	92.7	97.1	96.5	95.3	5.5%
Control Group 4 Accuracy (%)	93.5	90.2	88.6	85.3	92.8	91.7	90.5	8.2%

The experimental model maintains high recognition accuracy across various environmental scenarios and vessel types: accuracy remains at 92.7% in dense fog conditions, with accuracy fluctuations within 1.8 percentage points across different vessel types. The robustness metric (accuracy fluctuation range) was 5.5%, lower than the 8.2% observed in Control Group 4, indicating that the proposed model effectively resists complex environmental interference and adapts to monitoring requirements for different vessel types.

H. PERFORMANCE ANALYSIS OF SMALL-SAMPLE ANOMALY DETECTION

Three rare anomaly behaviors (sudden main engine failure, rudder malfunction, illegal boundary crossing) were selected to analyze the detection performance of different models. The results are shown in the Table 4 below:

TABLE 4. Detection Accuracy of Different Models for Rare Vessel Anomaly Behaviors

Anomaly Behavior Type	Sudden Main Engine Failure	Rudder Failure	Illegal Crossing	Average Accuracy (%)
Control Group 1	58.3	62.7	75.4	65.5
Control Group 4	82.5	85.3	89.6	85.8
Experimental Group	91.2	93.5	95.8	93.5

Results show that the experimental group achieved an average recognition accuracy of 93.5% for rare abnormal behaviors, representing a 7.7 percentage point improvement over Control Group 4 and a 28 percentage point improvement over Control Group 1. This demonstrates that data augmentation, transfer learning, and virtual sample generation strategies can effectively mitigate the small-sample problem and enhance the model's ability to recognize rare abnormal behaviors.

I. CASE STUDY ANALYSIS

Two representative cases — "Rudder Malfunction" and "Abnormal Collision Avoidance in Heavy Fog"—were selected to evaluate the experimental model's recognition performance: Rudder Malfunction Case: During vessel navigation, the rudder mechanism suddenly seized, rendering course adjustment impossible. The model detected the anomaly through its temporal branch by capturing abrupt course deviation and fixed rudder angle, while its spatial branch identified abnormal visual tilting of the hull. After fusion, the "rudder failure" anomaly was recognized within 0.27 seconds, triggering an alert and generating emergency response recommendations. Collision Avoidance Anomaly in Heavy Fog: During heavy fog (visibility <500 meters), when the target vessel failed to evade promptly, the experimental model used an enhanced Transformer to capture the temporal feature of rapidly decreasing distance between the two vessels. Combined with the target vessel's contour features extracted from video via CNN, it successfully identified the "improper collision avoidance maneuver" anomaly, issuing an alert 3.2 seconds earlier than traditional methods.

J. REAL-VESSEL APPLICATION VALIDATION

The experimental model was deployed on three vessels (1 container ship, 2 bulk carriers) operated by a shipping company for a three-month real-vessel validation. accumulating 18,600 nautical miles of voyage distance under various environmental conditions including normal sea conditions, strong winds, heavy rain, and dense fog. A total of 23 abnormal behaviors were detected (including 15 simulated anomalies and 8 real anomalies). Field application results demonstrated: Anomaly detection accuracy reached 95.7%, with a false alarm rate of 3.1% and a false negative rate of 2.9%, aligning closely with laboratory test outcomes and validating the model's practical effectiveness. The average warning response time was 0.31 seconds, representing a 96.4% improvement over traditional monitoring systems (average response time: 8.5 seconds), providing ample time for crew emergency response. Successfully alerted to 8 real-world abnormal behaviors (including 3 equipment failures, 2 operational errors, and 3 abnormal courses), preventing potential navigation accidents and cumulatively reducing economic losses by approximately 8.6 million yuan.

VI. CONCLUSIONS AND OUTLOOK

This study systematically investigates the design of a deep learning-based vessel navigation state monitoring model and abnormal behavior recognition, yielding the following core conclusions, which offer transferable insights for developing intelligent monitoring systems for the operational status of devices in complex user environments: A three-dimensional vessel navigation state monitoring system encompassing motion, equipment, and environment dimensions was established. Eight categories of typical abnormal behaviors were classified with defined recognition criteria. The MSS-2024 dataset, comprising 150,000 time-series samples and 50,000 video samples from multi-source data, was constructed to support model training and validation. A dual-branch monitoring model (TC-Fusion Model) was designed by integrating Transformers and CNNs. The temporal feature extraction branch captures long-term dependency features, while the spatial feature extraction branch enhances key regional features. Combined with attention mechanisms and multi-source fusion techniques, this approach improves feature extraction accuracy in complex scenarios. A small-sample processing solution combining data augmentation, transfer learning, and synthetic sample generation was proposed, effectively addressing the scarcity of rare anomaly behavior samples. Experimental validation and real-vessel applications demonstrate that the proposed model achieves 96.3% accuracy and 94.7% recall in anomaly detection, with a response time of 0.29 seconds. Its robustness and generalization capabilities significantly outperform traditional methods and mainstream deep learning models, effectively meeting the demands for real-time monitoring of vessel navigation states and rapid anomaly detection, demonstrating an advanced capability that is equally essential for monitoring the operational integrity of potential material failure or sensor malfunction.

AUTHOR CONTRIBUTIONS

Huawei Su designed, collected and analyzed the data, and drafted the manuscript. Huawei Su conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Huawei Su participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research was supported by,

1) Teaching Reform Project of Zhengzhou Police University in 2024, "Research on the Course Construction of 'Introduction to Waterway Policing' and the Practical Path of Ideological and Political Education" (Project No.: JY2024008).
2) Soft Science Research Project of Henan Province in 2025, "Research on the Integration of Yellow River Culture into Henan Public Security Education and Training Empowered by Digital Intelligence" (Project No.: 252400410366).

HUMAN RESEARCH SUBJECTS

This study was approved by the Ethics Committee of Zhengzhou Police University, and was performed in accordance with the principles of the Declaration of Helsinki. All eligible participants signed an informed consent form.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] M. Stevenson, C. Mues, and C. Bravo, “The value of text for small business default prediction: A Deep Learning approach,” European Journal of Operational Research, vol. 295, no. 2, pp. 758-771, 2021, doi: 10.1016/J.EJOR.2021.03.008.
[2] H. Zhang, H. Xu, X. Tian, J. Jiang, and J. Ma, “Image fusion meets deep learning: A survey and perspective,” Information Fusion, pp. 76323-76336, 2021, doi: 10.1016/J.INFFUS.2021.06.008.
[3] K. Wijesinghe, J. Wanni, N. Banerjee, S. Banerjee, and A. Achuthan, “Characterization of microscopic deformation of materials using deep learning algorithms,” Materials & Design, vol. 208, no. prepublish, Art. no. 109926, 2021, doi: 10.1016/J.MATDES. 2021.109926.
[4] F. Allaire, V. Mallet, and J. Filippi, “Emulation of wildland fire spread simulation using deep learning,” Neural Networks, pp. 141184-141198, 2021, doi: 10.1016/J.NEUNET.2021.04.006.
[5] P. Ma, C. P. Lau, N. Yu, A. Li, P. Liu, Q. Wang, et al., “Image-based nutrient estimation for Chinese dishes using deep learning,” Food Research International, Art. no. 147110437, 2021, doi: 10.1016/J.FOODRES.2021.110437.
[6] Y. C. Semerci and D. Goularas, “Evaluation of Students’ Flow State in an E-learning Environment Through Activity and Performance Using Deep Learning Techniques,” Journal of Educational Computing Research, vol. 59, no. 5, pp. 960-987, 2021, doi: 10.1177/0735633120979836.
[7] Y. Perez-Perez, M. Golparvar-Fard, and K. El-Rayes, “Scan2BIM-NET: Deep Learning Method for Segmentation of Point Clouds for Scan-to-BIM,” Journal of Construction Engineering and Management, vol. 147, no. 9, Art. no. 04021107, 2021, doi: 10.1061/(ASCE)CO.1943-7862.0002132.
[8] A. Patera, A. G. Zippo, A. Bonnin, M. Stampanoni, and G. E. Biella, “Brain micro-vasculature imaging: An unsupervised deep learning algorithm for segmenting mouse brain volume probed by high-resolution phase-contrast X-ray tomography,” International Journal of Imaging Systems and Technology, vol. 31, no. 3, pp. 1211-1220, 2021, doi: 10.1002/IMA.22520.
[9] M. Shen, G. Li, D. Wu, Y. Yaguchi, J. C. Haley, K. G. Field, et al., “A deep learning based automatic defect analysis framework for In-situ TEM ion irradiations,” Computational Materials Science, vol. 197, Art. no. 110560, 2021, doi: 10.1016/J.COMMATSCI.2021.110560.
[10] L. Vaquero, V. M. Brea, and M. Mucientes, “Tracking more than 100 arbitrary objects at 25 FPS through deep learning,” Pattern Recognition, vol. 121, Art. no. 108205, 2022, doi: 10.1016/J.PATCOG.2021.108205.
[11] A. Das, M. N. Mohanty, P. K. Mallick, P. Tiwari, K. Muhammad, and H. Zhu, “Breast cancer detection using an ensemble deep learning method,” Biomedical Signal Processing and Control, pp. 70, 2021, doi: 10.1016/J.BSPC.2021.103009.
[12] J. Song, M. Patel, A. Girgensohn, and C. Kim, “Combining deep learning with geometric features for image-based localization in the Gastrointestinal tract,” Expert Systems with Applications, pp. 185, 2021, doi: 10.1016/J.ESWA.2021.115631.
[13] R. Cheng and J. Chen, “A location conversion method for roads through deep learning-based semantic matching and simplified qualitative direction knowledge representation,” Engineering Applications of Artificial Intelligence, pp. 104, 2021, doi: 10.1016/J.ENGAPPAI.2021.104400.
[14] Y. Jia, G. Yu, J. Du, X. Gao, Y. Song, and F. Wang, “Adopting traditional image algorithms and deep learning to build the finite model of a 2.5D composite based on X-Ray computed tomography,” Composite Structures, pp. 275, 2021, doi: 10.1016/J.COMPSTRUCT.2021.114440.

[15] A. Manickam, J. Jiang, Y. Zhou, and A. Sagar, “Automated pneumonia detection on chest X-ray images: A deep learning approach with different optimizers and transfer learning architectures,” Measurement, pp. 184, 2021, doi: 10.1016/J.MEASUREMENT.2021.109953.

[16] A. Gurunathan and B. Krishnan, “Detection and diagnosis of brain tumors using deep learning convolutional neural networks,” International Journal of Imaging Systems and Technology, vol. 31, no. 3, pp. 1174-1184, 2021, doi: 10.1002/IMA.22532.

[17] A. Maghami, M. Salehi, and M. Khoshdarregi, “Automated vision-based inspection of drilled CFRP composites using multi-light imaging and deep learning,” CIRP Journal of Manufacturing Science and Technology, pp. 35441-35453, 2021, doi: 10.1016/J.CIRPJ.2021.07.015.

[18] S. D. Yang, Y. Q. Zhao, F. Zhang, M. Liao, Z. Yang, Y. Wang, et al., “An efficient two-step multi-organ registration on abdominal CT via deep-learning based segmentation,” Biomedical Signal Processing and Control, pp. 70, 2021, doi: 10.1016/J.BSPC.2021.103027.

[19] Y. Liu, Z. Zhang, X. Liu, L. Wang, and X. Xia, “Performance evaluation of a deep learning based wet coal image classification,” Minerals Engineering, pp. 171, 2021, doi: 10.1016/J.MINENG.2021.107126.

[20] Q. Zhang, F. Lee, Y. G. Wang, D. Ding, S. Yang, C. Lin, et al., “CJC-net: A cyclical training method with joint loss and co-teaching strategy net for deep learning under noisy labels,” Information Sciences, pp. 579186-579198, 2021, doi: 10.1016/J.INS.2021.08.008.

[21] H. C. Dan, G. W. Bai, and Z. H. Zhu, “Application of deep learning-based image recognition technology to asphalt– aggregate mixtures: Methodology,” Construction and Building Materials, vol. 297, Art. no. 123770, 2021, doi: 10.1016/J.CONBUILDMAT.2021.123770.

[22] R. Vandaele, S. L. Dance, and V. Ojha, “Deep learning for automated river-level monitoring through river-camera images: An approach based on water segmentation and transfer learning,” Hydrology and Earth System Sciences, vol. 25, no. 8, pp. -4453, 2021, doi: 10.5194/hess-25-4435-2021.

[23] H. G. Doherty, R. A. Burgueño, R. P. Trommel, V. Papanastasiou, and R. I. Harmanny, “Attention-based deep learning networks for identification of human gait using radar micro-Doppler spectrograms,” International Journal of Microwave and Wireless Technologies, vol. 13, no. 7, pp. 734-739, 2021, doi: 10.1017/S1759078721000830.

[24] A. Rahman, Z. Y. Wu, and R. Kalfarisi, “Semantic Deep Learning Integrated with RGB Feature-Based Rule Optimization for Facility Surface Corrosion Detection and Evaluation,” Journal of Computing in Civil Engineering, vol. 35, no. 6, Art. no. 04021018, 2021, doi: 10.1061/(ASCE)CP.1943-5487.0000982.

[25] X. Chen, Z. Chen, J. Li, Y. D. Zhang, X. Lin, and X. Qian, “Model-driven Deep Learning Method for Pancreatic Cancer Segmentation Based on Spiral-transformation,” IEEE Transactions on Medical Imaging, vol. 41, no. 1, pp. 75-87, 2021, doi: 10.1109/TMI.2021.3104460.

[26] Y. Pu, J. Li, J. Tang, and F. Guo, “DeepFusionDTA: Drug-target binding affinity prediction with information fusion and hybrid deep-learning ensemble model,” IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 19, no. 5, pp. 2760-2769, 2021, doi: 10.1109/TCBB.2021.3103966.

[27] S. Sharma, V. Rana, and V. Kumar, “Deep learning based semantic personalized recommendation system,” International Journal of Information Management Data Insights, vol. 1, no. 2, Art. no. 100028, 2021, doi: 10.1016/J.JJIMEI.2021.100028.

[28] Y. Ma and M. Gan, “DeepAssociate: A deep learning model exploring sequential influence and historycandidate association for sequence recommendation,” Expert Systems with Applications, pp. 185, 2021, doi: 10.1016/J.ESWA.2021.115587.

[29] M. Sharma, I. Kandasamy, and V. Kandasamy, “Deep Learning for predicting neutralities in Offensive Language Identification Dataset,” Expert Systems with Applications, pp. 185, 2021, doi: 10.1016/J.ESWA.2021.115458.

[30] M. Shen, G. Li, D. Wu, Y. Liu, J. R. Greaves, W. Hao, et al., “Multi defect detection and analysis of electron microscopy images with deep learning,” Computational Materials Science, pp. 199, 2021, doi: 10.1016/J.COMMATSCI.2021.110576.