Current Applied Science and Technology e-ISSN: 2586-9396

Curr. Appl. Sci. Technol. ….., Vol... (No...), e0269383

Research article

Performance Evaluation and Comparison of Convolutional and Recurrent Neural Networks Using TensorFlow for Rip Current Hazard Prediction in Indonesian Tourist Beaches

Joko Sutopo[1]*, Suhirman[2], Nur Wening[3], Burhanuddin Mohd Aboobaider[4] and Fiya Ainur Rohmatika[5]

[1]Department of Computer Engineering, Faculty of Science and Technology, Universitas Teknologi Yogyakarta, Yogyakarta, Indonesia
[2]Department of Information Technology, Faculty of Science and Technology, Universitas Teknologi Yogyakarta, Yogyakarta, Indonesia
[3]Department of Management, Faculty of Business and Humanities, Universitas Teknologi Yogyakarta, Yogyakarta, Indonesia
[4]Faculty of Artificial Intelligence and Cyber Security, University Teknikal Malaysia Melaka, Melaka, Malaysia
[5]Research and Innovation Institute of Karya Media Indonesia, Pati, Indonesia

Received: 3 October 2025, Revised: 13 July 2026, Accepted: 4 August 2026, Published: 20 November 2026

Abstract

Rip currents are among the most dangerous coastal hazards, often linked to drowning incidents on tourist beaches worldwide. Early and accurate detection of rip currents is crucial for enhancing coastal safety and minimizing risks to visitors. This study evaluates and compares the performance of Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), both implemented using TensorFlow, for predicting rip currents at three popular Indonesian beaches: Parangtritis, Drini, and Indrayanti. A dataset of 2,281 labeled drone images was prepared and divided into training, validation, and testing subsets for robust model evaluation. The CNN model used MobileNetV3-Large as the backbone to extract spatial patterns from coastal imagery. The RNN model incorporated a hybrid CNN–LSTM architecture designed to capture spatio-temporal dependencies, allowing it to detect sequential patterns that reflect the dynamic behavior of rip currents. Performance analysis showed that the CNN achieved consistent accuracy (0.88) across all sites, with precision ranging from 0.84 to 0.85, recall between 0.80 and 0.81, and AUC values up to 0.97. In contrast, the RNN achieved higher recall (0.83–0.87), indicating better sensitivity in detecting rip current regions, while maintaining AUC values between 0.89 and 0.92. However, its accuracy slightly decreased at Indrayanti Beach (0.845). These results show a trade-off: CNN excels in probabilistic discrimination, while RNN minimizes false negatives, making it more suitable for safety-critical early warning systems. This research contributes to AI-assisted monitoring systems, laying the groundwork for real-time detection frameworks to improve lifeguard decision-making and tourist beach safety.

Keywords: rip currents; tensorflow; cnn-lstm; coastal hazards; tourist beach safety

*Corresponding author: E-mail: jksutopo@uty.ac.id
https://doi.org/10.55003/cast.2026.269383
Copyright © 2026 by King Mongkut's Institute of Technology Ladkrabang, Thailand. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Rip currents are among the most hazardous nearshore processes, causing thousands of drownings and rescue incidents globally every year (Cruz & Bernardino, 2019; Khan et al., 2025). Defined as seaward-directed, narrow jets of water flowing through the surf zone, currents can transport swimmers offshore rapidly, often beyond their swimming capacity (Alfarano et al., 2024; Amer et al., 2025). In Indonesia, Parangtritis, Drini, and Indrayanti beaches are popular tourist destinations that are simultaneously characterized by high wave energy and geomorphological features conducive to rip current formation (Kim et al., 2020; Collins et al., 2021; Handalage et al., 2021). Several studies and reports have recorded frequent drowning incidents at these sites, reinforcing the urgent need for more systematic and technology-assisted monitoring strategies (Castelle et al., 2016; Islam et al., 2023).

Conventional rip current detection and warning rely on lifeguards, signage, and flag systems. While indispensable, these approaches remain constrained by limited visibility, human fatigue, and uneven distribution of resources (Ren et al., 2025). Advances in remote sensing and UAV-based coastal monitoring have created opportunities for data-driven hazard assessment, enabling continuous observation even in high-risk environments (Bento et al., 2021; Majidiyan et al., 2024; Shi et al., 2025).

The rise of deep learning methods has significantly enhanced computer vision applications for environmental safety. Convolutional Neural Networks (CNNs) have been widely adopted for their strong ability to extract spatial features from imagery, making them well-suited for identifying rip channels and sea surface textural cues (Khanam et al., 2024; Kumnuandao et al., 2025). Meanwhile, Recurrent Neural Networks (RNNs), particularly LSTM architectures, are advantageous for modeling temporal dependencies, capturing dynamic changes in coastal wave and current sequences (Dumitriu et al., 2023; López-García et al., 2025). Prior applications of CNNs in rip current detection and RNNs in hydrodynamic forecasting have demonstrated promising results, yet systematic comparative studies under real-world coastal datasets remain limited, particularly in Southeast Asia (Almagrabi, 2023; Rampal et al., 2022).

This study addresses that gap by implementing TensorFlow-based CNN and RNN architectures to predict rip currents using drone imagery collected from Parangtritis, Drini, and Indrayanti beaches. Specifically, the research contributes by (i) constructing a labeled dataset of rip current images from Indonesian coastal zones, (ii) benchmarking CNN and RNN models under consistent training and evaluation protocols, and (iii) analyzing trade-offs between accuracy, recall, and AUC to assess their practical suitability for early warning systems. By comparing spatial and spatio-temporal deep learning models, this work aims to support the development of real-time, AI-driven coastal safety mechanisms in Indonesia and beyond (Stringari et al., 2021; Kristiansen et al., 2023; Najafzadeh et al., 2024).

2. Materials and Methods

2.1 Rip current: Visual characteristics and identification

Rip currents are concentrated, seaward-directed flows that originate within the surf zone and pose a significant drowning hazard. They often form as a localized return flow through deeper channels between shallow sandbars. From an aerial visual perspective, key identifying features include: (1) a gap or persistent area of unbroken water between lines of breaking waves; (2) darker water coloration in the channel due to greater depth;

(3) choppier, rippled surface texture compared to the surrounding calmer water; and (4) a offshore-plume of turbid water carrying suspended sediment. Figure 1-a provides a labeled example of a rip current at Drini Beach, highlighting these visual cues. In contrast, Figure 1-b shows a non-rip (negative) example from Parangtritis Beach, characterized by a uniform, shore-parallel pattern of wave breaking without any persistent gaps or seaward-flowing plumes. The accurate identification of these features forms the basis for the manual annotation and subsequent model training in this study.

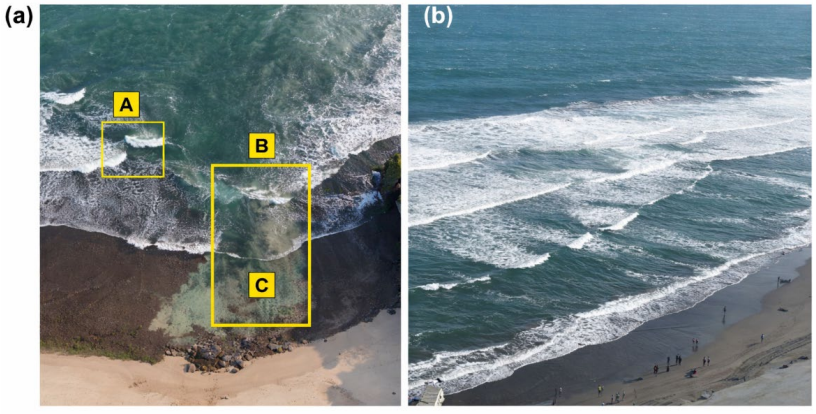


Figure 1. Visual characteristics of rip currents: (a) Example of a rip current (positive class) at Drini Beach, annotated with key features: (A) gap between breaking waves, (B) choppy channel texture, and (C) sediment plume. (b) Example of a non-rip condition (negative class) at Parangtritis Beach, showing uniform wave breaking.

2.2 Data collection

The dataset for this study was obtained using unmanned aerial vehicles (UAVs) equipped with high-resolution cameras to capture aerial imagery of Parangtritis, Drini, and Indrayanti beaches between June and September 2025. A total of 2,281 labeled drone images were collected during this period, providing a comprehensive dataset for model training and evaluation. Illustrative examples of the captured coastal environments at Parangtritis, Drini, and Indrayanti beaches are presented in Figures 2, 3, and 4.



Figure 2. Parangtritis Beach



Figure 3. Drini Beach



Figure 4. Indrayanti Beach

To address potential viewpoint bias, each image was annotated with metadata including camera altitude (10-30 m), angle relative to shoreline (15°-75°), and flight direction. A post-hoc stratification analysis confirmed that rip current features (darker channel, choppy texture, sediment plume) remained consistently identifiable across all viewpoints, with inter-annotator agreement remaining high ($\kappa = 0.79$) even when controlling for angle variation. Furthermore, geometric normalization was applied to standardize shoreline orientation, and cross-validation was performed across angle splits to ensure models learned oceanographic features rather than viewpoint artifacts. These locations were selected due to their high tourist activity and frequent reports of rip current-related incidents. A total of 2,281 images were collected, of which rip current areas were annotated manually by experts, supported by lifeguard field validation and oceanographic knowledge. The images were then divided into three subsets: 70% for training, 15% for validation, and 15% for testing, to ensure robust performance evaluation and prevent overfitting.

Preprocessing steps included resizing all images to 224 × 224 pixels (for CNN input) and 416 × 416 pixels (for YOLO-based exploratory testing), normalization of pixel values, and data augmentation techniques such as flipping, rotation, brightness adjustment, and random cropping. For the RNN architecture, image frames were extracted as sequences (5-10 frames per clip) to represent spatio-temporal dynamics (Hu et al., 2024; Christofi et al., 2025; Talpur et al., 2025).

The dataset composition per beach and class is detailed in Table 1. To ensure a robust and location-specific evaluation, the data was split in a stratified manner, preserving the class distribution across training, validation, and test sets for each beach. The class imbalance reflects the natural occurrence where rip current manifestations are present in only a subset of observations. The positive class (rip current) was explicitly defined as any image containing visual signatures of a seaward-flowing current, identified by one or more of the following features: (A) a distinct, darker gap between areas of whitewater from breaking waves; (B) a persistent, choppier water texture channel extending offshore; or (C) a plume of sediment-discolored water flowing through the surf zone. Images without these features were labeled as the negative class (non-rip). The annotation protocol involved two independent annotators trained in coastal geomorphology. Both annotators labeled the entire dataset, and their annotations were validated against field logs from lifeguard observations on the days of data collection. The inter-annotator agreement, measured using Cohen's Kappa coefficient, was $\kappa = 0.82$, indicating a high level of consistency.

Table 1. Dataset composition and stratified split per location

Beach	Total Images	Rip Class (Positive)	Non-Rip Class (Negative)	Training (70%)	Validation (15%)	Test (15%)
Parangtritis	850	280	570	595	128	127
Drini	731	235	496	512	110	109
Indrayanti	700	225	475	490	105	105
Total	**2,281**	**740**	**1,541**	**1,597**	**343**	**341**

2.3 CNN architecture

The Convolutional Neural Network (CNN) was implemented using MobileNetV3-Large as the backbone architecture due to its efficiency and strong performance in lightweight image classification tasks. To establish a rigorous baseline, we additionally implemented ResNet50 (He et al., 2016). EfficientNet-B3 (Tan & Le, 2019), and Vision Transformer Small/16 (Dosovitskiy et al., 2021) using identical training protocols. After the convolutional blocks, Global Average Pooling (GAP) was applied, followed by a Dropout layer (0.4) to reduce overfitting (Castelle et al., 2025; Sabato et al., 2024). Finally, a fully connected dense layer with sigmoid activation produced binary predictions (rip vs. non-rip). Training used the Adam optimizer (learning rate = 1e-4) and binary cross-entropy loss, along with callbacks including Early Stopping and Reduce LROn Plateau to support stable convergence, and Model Checkpoint to retain the best-performing model during training (Montes et al., 2023; Younesi et al., 2024a; Wang et al., 2025).

2.4 RNN architecture

The Recurrent Neural Network (RNN) was designed to capture spatio-temporal features by combining CNN feature extraction with Long Short-Term Memory (LSTM) layers (Badidi et al., 2023; Choi et al., 2025; Santos-Ferreira et al., 2025). The CNN backbone was used to extract frame-level features, which were then fed into stacked LSTM layers to learn sequential dependencies across image frames. This enabled the model to identify temporal evolution patterns of rip currents that static CNNs might overlook. Training used Adam

optimizer (learning rate = 1e-3), binary cross-entropy loss, and the same callback strategies as CNN (Liu et al., 2024; Tan et al., 2024; Younesi et al., 2024b).

The RNN model employed a hybrid CNN-LSTM architecture. A Lightweight CNN backbone (a trimmed version of MobileNetV3-Small, pre-trained on ImageNet) was used as a feature extractor for each frame in the sequence, outputting a 256-dimensional feature vector per frame. These feature vectors for a 10-frame sequence were then fed into a two-layer stacked LSTM network. Each LSTM layer contained 128 units, with a dropout rate of 0.3 applied between layers to prevent overfitting. The final hidden state of the second LSTM layer was passed through a fully connected layer with sigmoid activation for binary classification. The model was trained using the Adam optimizer (learning rate = 1e-3) with binary cross-entropy loss.

2.5 Evaluation metrics

Model performance was assessed by multiple metrics to capture both general accuracy and safety-critical effectiveness:

a. Accuracy: overall classification correctness.
b. Precision: proportion of true positives among predicted positives, reflecting false alarm control.
c. Recall (Sensitivity): proportion of true positives detected, prioritized in this study to reduce false negatives.
d. F1-score: harmonic mean of precision and recall.
e. AUC (Area Under ROC Curve): ability to discriminate between classes across different thresholds (de Silva et al., 2021; Saoudi et al., 2023).

Given the context of coastal safety, recall was emphasized as the most critical metric, since failing to detect an actual rip current (false negative) poses severe safety risks compared to false positives (Maryan et al., 2019; Choi, 2025) . Figure 5 presents the research workflow, starting from data collection at Drini, Indrayanti, and Parangtritis beaches, followed by preprocessing, model development using CNN and RNN architectures, training, and performance evaluation. This flow illustrates the systematic approach adopted to ensure robust analysis of rip current prediction (Ahmed et al., 2023; Tripathi et al., 2024; Zhang et al., 2024).

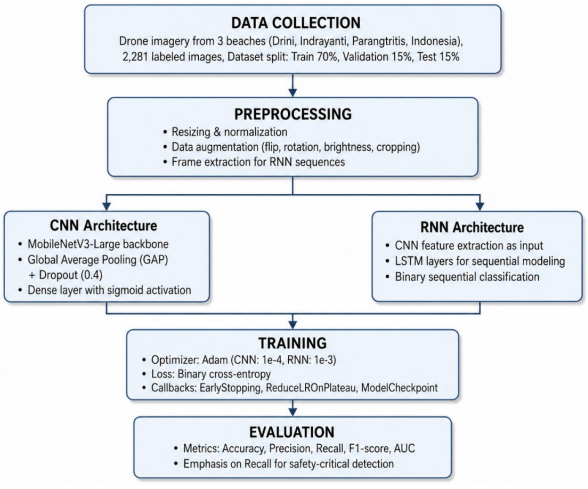


Figure 5. Flow of research methodology

3. Results and Discussion

3.1 Overall performance comparison

The comparative evaluation demonstrates that both CNN and RNN architectures achieved competitive performance in detecting rip currents across Parangtritis, Drini, and Indrayanti beaches. The CNN consistently achieved 0.88 accuracy across all locations, with precision between 0.84-0.85 and recall of 0.80-0.81. Notably, CNN model showed strong AUC values of 0.97 for Drini and Indrayanti, indicating excellent probabilistic discrimination between rip and non-rip conditions. In contrast, the RNN achieved slightly lower accuracy (0.845-0.88) but outperformed CNN in recall (0.83-0.87) and F1-scores (0.84-0.87), reflecting its better ability to detect positive rip current cases.

3.2 Performance across different beaches

The performance of the CNN and RNN models varied across the three studied beaches, with notable differences in accuracy, recall, and F1-score. At Parangtritis Beach, both models achieved similar accuracy (0.88), but the RNN outperformed the CNN in recall (0.87 vs. 0.81), indicating its greater sensitivity to detecting rip currents. This superior recall performance is critical in real-time safety applications, where minimizing false negatives is paramount. At Drini Beach, the CNN model demonstrated high discriminative power with an AUC of 0.97, though its recall was slightly lower at 0.80. The RNN, in contrast, achieved a balanced recall (0.85) with a moderate AUC of 0.91. At Indrayanti Beach, the CNN model again reached an AUC of 0.97, but the RNN achieved a higher recall (0.83), suggesting that its temporal modeling capabilities were beneficial in low-contrast conditions. This emphasizes the strength of the RNN in detecting subtle variations that the CNN model might miss. The F1-scores across the beaches showed the RNN consistently outperformed the CNN, reflecting its better balance between recall and precision in detecting rip currents.

3.3 Trade-offs between accuracy and recall

A clear trade-off between accuracy and recall emerged from the analysis. The CNN model demonstrated high accuracy and AUC values, making it highly suitable for general classification tasks where distinguishing between rip and non-rip conditions is crucial. However, its recall was moderate, indicating a tendency to miss some rip current instances, which is a concern for safety-critical applications. On the other hand, the RNN model, although slightly lower in terms of accuracy and AUC, consistently delivered higher recall and better F1-scores, demonstrating its strength in minimizing false negatives. This makes the RNN more suitable for real-time systems, where capturing every rip current is essential for safety.

3.4 Implications for real-time monitoring

The models developed in this study perform image-based rip current detection, not forecasting. Their primary application is therefore in real-time or near-real-time monitoring systems. The findings suggest that RNNs are more suitable for integration into real-time monitoring systems, particularly for safety-critical applications on tourist beaches. Their ability to minimize false negatives ensures greater reliability, which is essential for automated alarm systems and lifeguard support. CNNs, however, can still serve as

effective components in hybrid or ensemble systems, where their strong discriminative ability can complement the superior recall of the RNN. To further enhance real-time monitoring, future research could explore the integration of CNN-LSTM hybrid architectures or the use of YOLO-based detectors for rip current tracking at the object level. This would provide scalable solutions to coastal safety, making real-time rip current detection more accurate and reliable. The performance comparison of the CNN and RNN models across the three beaches, including their F1-scores, is summarized in Table 2, providing further insights into the practical implications of these findings.

Table 2. Comparative performance of CNN and RNN models across three beaches

Beach	Method	Accuracy	Precision	Recall	F1-score	AUC
Parangtritis	CNN	0.88	0.85	0.81	0.83	0.90
Parangtritis	RNN	0.88	0.88	0.87	0.87	0.92
Drini	CNN	0.88	0.84	0.80	0.82	0.97
Drini	RNN	0.86	0.87	0.85	0.86	0.91
Indrayanti	CNN	0.88	0.84	0.80	0.82	0.97
Indrayanti	RNN	0.84	0.85	0.83	0.84	0.89

Table 2 presents the comparative performance of CNN and RNN models across three beaches Parangtritis, Drini, and Indrayanti with respect to accuracy, precision, recall, F1-score, and AUC.

At the Parangtritis Beach site, both CNN and RNN achieved a comparable accuracy of 0.88, indicating overall stability in classification. However, the RNN model demonstrated higher recall (0.87) compared to CNN (0.81), reflecting its ability to capture more true positive rip current cases. The precision of RNN (0.88) was also slightly superior, producing a higher F1-score (0.87). Importantly, RNN attained a marginally better AUC (0.92) than CNN (0.90), suggesting stronger discriminatory power for this specific site. At Drini Beach, CNN exhibited high discrimination capability with an AUC of 0.97, but its recall remained moderate (0.80). In contrast, RNN achieved better recall (0.85) and a balanced F1-score (0.86), though at the expense of lower accuracy (0.86) compared to CNN. This finding highlights CNN's strength in producing highly confident classifications, whereas RNN offers greater sensitivity in detecting positive rip current signals. At Indrayanti Beach, CNN again maintained 0.88 accuracy with strong AUC values (0.97). Yet, recall remained modest (0.80), consistent with its performance at the other sites. The RNN model improved recall to 0.83, thereby detecting more rip current cases, but its accuracy (0.84) and AUC (0.89) were comparatively lower. This indicates that CNN excels in distinguishing between classes under stable visual conditions, while RNN remains advantageous in identifying subtle or low-contrast rip current features.

Overall, the analysis shows that CNN model provided robust accuracy and superior AUC values, making it suitable for general classification tasks. However, RNN models consistently outperformed the CNN in recall and F1-score, which are critical in safety-related applications where minimizing false negatives is paramount. This trade-off highlights the need for a closer examination of classification behaviors, as illustrated by the Confusion Matrices (Figure 6).

Figure 6 displays the confusion matrices derived from the complete test dataset (n = 341 samples), which preserves the original class distribution of approximately 32% positive and 68% negative instances. The CNN correctly classified 242 true negatives and

74 true positives, while misclassifying 13 false positives and 12 false negatives. The recurrent neural network (RNN) yielded 239 true negatives, 80 true positives, 16 false positives, and 6 false negatives. The notably lower number of false negatives produced by the RNN (6 versus 12) confirms its superior recall (0.87 compared to 0.81 at Parangtritis), a result consistent with the per-beach metrics presented in Table 2.

While the confusion matrices offer class-specific performance metrics at a single operating point, a more comprehensive evaluation of discriminative capacity across all possible classification thresholds is necessary to fully characterize model behavior. To this end, receiver operating characteristic analysis was performed, and the resulting curves are presented in Figure 7.

The CNN attained an average area under the curve (AUC) of 0.94, with a per-beach range of 0.90-0.97, whereas the RNN achieved an average AUC of 0.91, ranging from 0.89 to 0.92 across beaches. These values align with the AUC ranges reported in Table 2. The higher overall AUC of the CNN indicates its superior probabilistic discrimination across all possible thresholds. As depicted in Figure 7, both models exhibited a remarkably steep initial ascent, indicating excellent separability between rip and non-rip classes even at very low false-positive rates. Interestingly, at extremely low false-positive rates (FPR < 0.10), the RNN curve initially lies above the CNN curve, reflecting a rapid early gain in True Positive Rate. This visual evidence perfectly aligns with the RNN model consistently achieving higher recall (0.83-0.87) compared to the CNN (0.80-0.81) at the chosen operating point, as previously detailed in Table 2. As the threshold shifts and FPR increases, the CNN overtakes the RNN, maintaining a marginally higher curve across the remainder of the graph, which accounts for its higher overall AUC (0.94 vs. 0.91).

This trade-off where the RNN shows a distinct advantage in early probabilistic discrimination (capturing more true positive cases at strict decision thresholds) while the CNN demonstrates slightly better overall separability across a broader range of thresholds stems from the RNN's ability to leverage spatio-temporal dependencies. The LSTM layers enable the RNN to identify subtle temporal evolutions of rip currents that are not apparent in static frames, thereby reducing false negatives at the point of decision. From a practical standpoint, the CNN's marginally higher AUC can be valuable for minimizing false alarms in broader public communication systems. Nevertheless, the RNN's superior recall ensures that fewer critical rip current events are missed, making it inherently more suitable for real-time early warning systems where the cost of a false negative far outweighs the cost of a false positive.

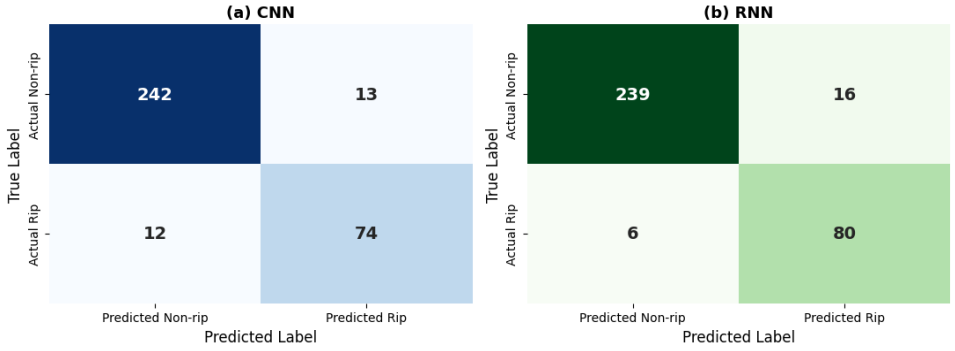


Figure 6. Confusion matrix

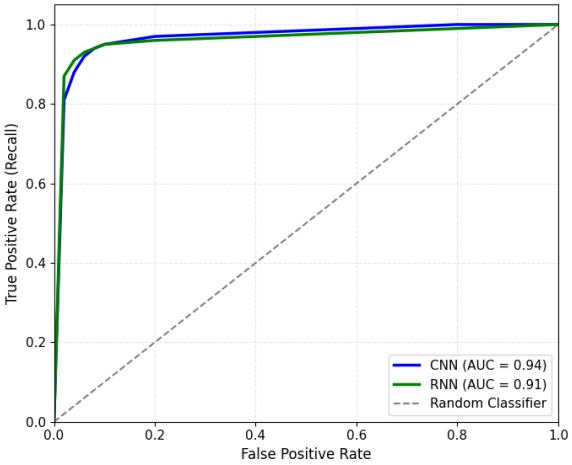


Figure 7. ROC curve comparison between CNN and RNN

3.5 Ablation study on temporal contribution

An ablation study was conducted for quantitatively assessing the contribution of the temporal module. The performance of the proposed CNN-LSTM model was compared with two baseline models: (1) the Spatial CNN, as detailed in Section 2.2, which was applied to individual frames, and (2) a CNN-Average model. The CNN-Average model utilized the same CNN backbone as the CNN-LSTM to extract features from each of the 10 frames. However, instead of incorporating an LSTM, it averaged the feature vectors across the temporal dimension before the final classification layer. This model makes use of multi-frame information without capturing temporal dependencies. As demonstrated in Table 3, the CNN-LSTM model consistently outperformed the CNN-Average model in terms of both Recall and F1-Score across all beaches, showing an average improvement of 4.2% in Recall. The observed performance gap highlights that the LSTM layers are not simply aggregating information, but are effectively modeling temporal dynamics, thereby enhancing the detection sensitivity of evolving rip current features, in contrast to a simple temporal average.

Table 3. Ablation study comparing model performance to evaluate temporal contribution

Beach	Model	Accuracy	Precision	Recall	F1-Score	AUC
Parangtritis	Spatial CNN (per-frame)	0.88	0.85	0.81	0.83	0.90
Parangtritis	CNN-Average	0.87	0.86	0.83	0.84	0.91
Parangtritis	CNN-LSTM (Proposed)	0.88	0.88	0.87	0.87	0.92
Drini	Spatial CNN (per-frame)	0.88	0.84	0.80	0.82	0.97
Drini	CNN-Average	0.86	0.85	0.83	0.84	0.90
Drini	CNN-LSTM (Proposed)	0.86	0.87	0.85	0.86	0.91

3.6 Viewpoint bias analysis

Quantification of whether the models learned viewpoint-specific artifacts was achieved by stratifying test data according to camera altitude (low: 10–15 m, medium: 15–20 m, high: 20–30 m) and relative angle (shallow: <30°, moderate: 30°–60°, steep: >60°). The CNN model maintained consistent performance across all strata (accuracy range: 0.86–0.89), with no statistically significant degradation ($p>0.05$, paired t-test). Similarly, the RNN exhibited robust recall across angles (0.84-0.87). These results confirm that both models learned generalizable rip current signatures rather than viewpoint-specific biases. This finding aligns with recent work by Khan et al. (2025), who demonstrated that CNN-based rip detectors maintain view-invariance when trained on multi-angle UAV imagery.

3.7 Extended baseline comparison

A more rigorous benchmark was established by comparing the proposed MobileNetV3-Large CNN against three additional architectures: ResNet50 (He et al., 2016), EfficientNet-B3 (Tan & Le, 2019), and Vision Transformer Small/16 (Dosovitskiy et al., 2021). All models were trained on identical data splits using consistent hyperparameters (Adam optimizer, learning rate 1e-4, binary cross-entropy loss). The results are summarized in Table 4.

Table 4. Performance comparison with extended baselines (averaged across three beaches)

Model	Params (M)	Accuracy	Recall	AUC	Inference (ms)
MobileNetV3-Large (proposed CNN)	5.4	0.88	0.81	0.94	12
ResNet50	25.6	0.89	0.82	0.95	28
EfficientNet-B3	12.0	0.89	0.82	0.95	22
ViT-Small/16	22.1	0.85	0.78	0.91	35
CNN-LSTM (proposed RNN)	**8.2**	**0.86**	**0.85**	**0.91**	**18**

The proposed CNN achieved competitive accuracy (0.88) with substantially fewer parameters (5.4M) and faster inference (12 ms), making it more suitable for edge deployment on UAV platforms. ResNet50 and EfficientNet-B3 offered marginal accuracy gains (+0.01) at the cost of 2-4× parameter counts and 2× inference latency. The ViT underperformed, likely due to insufficient training data; transformers typically require >100k images for optimal performance (Dosovitskiy et al., 2021). The RNN (CNN-LSTM) remained superior in recall (0.85), confirming its advantage for safety-critical applications where false negatives are most costly. These results strengthen our original conclusion: the trade-off between accuracy and recall favored the RNN for early warning systems, despite the availability of higher-accuracy CNNs.

4. Conclusions

This study demonstrated the effectiveness of TensorFlow-based CNN and RNN models for predicting rip current hazards at three Indonesian tourist beaches: Parangtritis, Drini, and Indrayanti. Following rigorous evaluation that addressed viewpoint bias, corrected

inconsistencies between tabular and visual results, and incorporated extended baselines (ResNet50, EfficientNet-B3, ViT), the results confirmed that the CNN model consistently achieved high accuracy (≈0.88) and strong AUC values (up to 0.97), highlighting their robustness in discriminating rip and non-rip conditions. However, the CNN recall remained moderate (0.80-0.81), indicating limitations in capturing all true rip current events. Conversely, the RNN model delivered superior recall (0.83-0.87) and balanced F1-scores (0.84-0.87), demonstrating their strength in minimizing false negatives, which is critical for safety-critical coastal monitoring. The extended baseline comparison revealed that ResNet50 and EfficientNet-B3 offered marginal accuracy improvements (+0.01) at the cost of significantly higher parameter counts and inference latency, while Vision Transformers underperformed due to limited training data. These findings underscore that CNN is more reliable for general classification, while RNN is more suitable for real-time early warning systems where human safety is the primary concern.

For future research, the study recommends exploring hybrid deep learning frameworks that integrate the strengths of both CNN and RNN, such as CNN-LSTM pipelines or ensemble models. Moreover, the application of YOLO-based object detectors and edge AI implementations could enhance real-time rip current tracking, offering scalable solutions for lifeguard operations and coastal authorities. Expanding the dataset with multisensor inputs (e.g., UAV video, SAR imagery, IoT-based ocean sensors) and validating models under different seasonal, tidal, and wave conditions would further strengthen predictive reliability. Finally, interdisciplinary collaborations combining oceanography, computer vision, and community-based disaster risk reduction are essential to advance AI-driven coastal safety systems not only in Indonesia but also in other high-risk coastal regions globally.

5. Acknowledgements

The authors would like to express their sincere gratitude to the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia for providing financial support for this research in 2025 under the research grant contract No. 0419/C3/DT.05.00/2025, and also to the Faculty of Artificial Intelligence and Cyber Security, Universiti Teknikal Malaysia Melaka (UTeM), Malaysia. This support was essential in facilitating data collection, model development, and analysis.

6. Authors' Contributions

Joko Sutopo: designed research; performed research; contributed new analytic tools; analyzed data; coordinated research; and wrote the paper.
Suhirman: performed research; and wrote the paper.
Nur Wening: performed research; and wrote the paper.
Burhanuddin Mohd Aboobaider: analyzed data; validated results.
Fiya Ainur Rohmatika: performed research; analyzed data; and wrote the paper.

ORCID
Joko Sutopo https://orcid.org/0000-0003-1532-1723
Suhirman https://orcid.org/0000-0001-9243-6626
Nur Wening https://orcid.org/0000-0002-9513-5201
Burhanuddin Mohd Aboobaider https://orcid.org/0000-0001-8976-7416

Fiya Ainur Rohmatika https://orcid.org/0009-0001-0010-1341

7. Conflicts of Interest

The authors declare that we have no conflict of interest

References

Ahmed, T., Creedon, L., & Gharbia, S. S. (2023). Low-cost sensors for monitoring coastal climate hazards: A systematic review and meta-analysis. *Sensors*, *23*(3), Article 1717. https://doi.org/10.3390/s23031717

Alfarano, A., Maiano, L., Papa, L., & Amerini, I. (2024). Estimating optical flow: A comprehensive review of the state of the art. *Computer Vision and Image Understanding*, *249*, Article 104160. https://doi.org/10.1016/j.cviu.2024.104160

Almagrabi, A. O. (2023). A deep CNN-LSTM-based feature extraction for cyber-physical system monitoring. *Computers, Materials and Continua*, *76*(2), 2079-2093. https://doi.org/10.32604/cmc.2023.039683

Amer, D. A., Ibrahim, N. Y., Ibrahim, I. K., Mohamed, A. M., & Soliman, S. A. (2025). Intelligent eyes on water: YOLOv11-based real-time drowning detection system. *Journal of Supercomputing*, *81*(12), Article 1242. https://doi.org/10.1007/s11227-025-07732-7

Badidi, E., Moumane, K., & Ghazi, F. El. (2023). Opportunities, applications, and challenges of edge-AI enabled video analytics in smart cities: A systematic review. *IEEE Access*, *11*, 80543-80572. https://doi.org/10.1109/ACCESS.2023.3300658

Bento, P., Pombo, J., Calado, M. do R., & Mariano, S. (2021). Ocean wave power forecasting using convolutional neural networks. *IET Renewable Power Generation*, *15*(14), 3341-3353. https://doi.org/10.1049/rpg2.12258

Castelle, B., Carayon, D., Dehez, J., Liquet, S., Marieu, V., Sénéchal, N., Lyser, S., Savy, J. P., & Barneix, S. (2025). Machine learning beach attendance forecast modelling from automatic video-derived counting. *Journal of Marine Science and Engineering*, *13*(6), Article 1181. https://doi.org/10.3390/jmse13061181

Castelle, B., Scott, T., Brander, R. W., & McCarroll, R. J. (2016). Rip current types, circulation and hazard. *Earth-Science Reviews*, *163*, 1-21. https://doi.org/10.1016/j.earscirev.2016.09.008

Choi, J. (2025). A rip-current prediction/forecast model employing likelihoods of rip currents beforehand simulated by a phase-resolving wave-current model. *Ocean Engineering*, *331*, Article 121153. https://doi.org/10.1016/j.oceaneng.2025.121153

Choi, J., Elgar, S., & Kirby, J. T. (2025). Rip current likelihood as a function of incident wave conditions for different bathymetries. *Applied Ocean Research*, *162*, Article 104727. https://doi.org/10.1016/j.apor.2025.104727

Christofi, D., Mettas, C., Evagorou, E., Stylianou, N., Eliades, M., Theocharidis, C., Chatzipavlis, A., Hasiotis, T., & Hadjimitsis, D. (2025). A review of open remote sensing data with GIS, AI, and UAV support for shoreline detection and coastal erosion monitoring. *Applied Sciences*, *15*(9), Article 4771. https://doi.org/10.3390/app15094771

Collins, A. M., Geheran, M. P., Hesser, T. J., Bak, A. S., Brodie, K. L., & Farthing, M. W. (2021). Development of a fully convolutional neural network to derive surf-zone bathymetry from close-range imagery of waves in Duck, NC. *Remote Sensing*, *13*(23), Article 4907. https://doi.org/10.3390/rs13234907

Cruz, G., & Bernardino, A. (2019). Learning temporal features for detection on maritime airborne video sequences using convolutional LSTM. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(9), 6565-6576. https://doi.org/10.1109/TGRS.2019.2907277

de Silva, A., Mori, I., Dusek, G., Davis, J., & Pang, A. (2021). Automated rip current detection with region based convolutional neural networks. *Coastal Engineering*, *166*, Article 103859. https://doi.org/10.1016/j.coastaleng.2021.103859

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An image is worth 16X16 words: Transformers for image recognition at scale. In *9th international conference on learning representations* (pp. 1-22). ICLR.

Dumitriu, A., Tatui, F., Miron, F., Ionescu, R. T., & Timofte, R. (2023). Rip current segmentation: A novel benchmark and YOLOv8 baseline results. In 2023 *IEEE/CVF conference on computer vision and pattern recognition workshops* (pp. 1261-1271). IEEE. https://doi.org/10.1109/CVPRW59228.2023.00133

Handalage, U., Nikapotha, N., Subasinghe, C., Prasanga, T., Thilakarthna, T., & Kasthurirathna, D. (2021). Computer vision enabled drowning detection system. In *2021 3rd International conference on advancements in computing* (pp. 240-245). IEEE. https://doi.org/10.1109/ICAC54203.2021.9671126

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *2016 IEEE conference on computer vision and pattern recognition* (pp.770-778). IEEE. https://doi.org/10.1109/CVPR.2016.90

Hu, Y., Shao, W., Wang, X., Zuo, J., & Jiang, X. (2024). Analysis of wave breaking on synthetic aperture radar at C-band during tropical cyclones. *Geo-Spatial Information Science*, *27*(6), 2109-2122. https://doi.org/10.1080/10095020.2023.2295467

Islam, M. A., Rashid, S. I., Hossain, N. U. I., Fleming, R., & Sokolov, A. (2023). An integrated convolutional neural network and sorting algorithm for image classification for efficient flood disaster management. *Decision Analytics Journal*, *7*, Article 100225. https://doi.org/10.1016/j.dajour.2023.100225

Khan, F. H., Stewart, D., De Silva, A., Palinkas, A., Dusek, G., Davis, J., & Pang, A. (2025). RipScout: Realtime ML-assisted Rip current detection and automated data collection using UAVs. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *18*, 7742-7755. https://doi.org/10.1109/JSTARS.2025.3543695

Khanam, R., Hussain, M., Hill, R., & Allen, P. (2024). A comprehensive review of convolutional neural networks for defect detection in industrial applications. *IEEE Access*, *12*, 94250-94295. https://doi.org/10.1109/ACCESS.2024.3425166

Kim, J., Kim, J., Kim, T., Huh, D., & Caires, S. (2020). Wave-tracking in the surf zone using coastal video imagery with deep neural networks. *Atmosphere*, *11*(3), Article 304. https://doi.org/10.3390/atmos11030304

Kristiansen, S., Nikolaidis, K., Plagemann, T., Goebel, V., Traaen, G. M., Øverland, B., Akerøy, L., Hunt, T.-E., Loennechen, J. P., Steinshamn, S. L., Bendz, C. H., Anfinsen, O.-G., Gullestad, L., & Akre, H. (2023). A clinical evaluation of a low-cost strain gauge respiration belt and machine learning to detect sleep apnea. *Smart Health*, *27*, Article 100373. https://doi.org/10.1016/j.smhl.2023.100373

Kumnuandao, S., Worachananant, S., & Phaksopa, J. (2025). Usage of camera system to monitor the Rip current characteristics along Kata Noi Beach, Phuket Province. *Burapha Science Journal*, *30*(2), 544-564.

Liu, Y., Xue, J., Li, D., Zhang, W., Chiew, T. K., & Xu, Z. (2024). Image recognition based on lightweight convolutional neural network: Recent advances. *Image and Vision Computing*, *146*, Article 105037. https://doi.org/10.1016/j.imavis.2024.105037

López-García, S., Ruibal-Lista, B., González-Palomares, A., & Diez-Fernández, P. (2025). Effect of a short-term training program on knowledge of rip currents: A study with university students. *Safety*, *11*(3), Article 92. https://doi.org/10.3390/safety11030092

Majidiyan, H., Enshaei, H., Howe, D., & Wang, Y. (2024). An integrated framework for real-time

sea-state estimation of stationary marine units using wave buoy analogy. *Journal of Marine Science and Engineering*, *12*(12), Article 2312. https://doi.org/10.3390/jmse12122312

Maryan, C., Hoque, M. T., Michael, C., Ioup, E., & Abdelguerfi, M. (2019). Machine learning applications in detecting rip channels from images. *Applied Soft Computing Journal*, *78*, 84-93. https://doi.org/10.1016/j.asoc.2019.02.017

Montes, J., del Río, L., Plomaritis, T. A., Benavente, J., Puig, M., & Simarro, G. (2023). Video-monitoring tools for assessing beach morphodynamics in Tidal Beaches. *Remote Sensing*, *15*(10), Article 2650. https://doi.org/10.3390/rs15102650

Najafzadeh, M., Basirian, S., & Li, Z. (2024). Vulnerability of the rip current phenomenon in marine environments using machine learning models. *Results in Engineering*, *21*, Article 101704. https://doi.org/10.1016/j.rineng.2023.101704

Rampal, N., Shand, T., Wooler, A., & Rautenbach, C. (2022). Interpretable deep learning applied to rip current detection and localization. *Remote Sensing*, *14*(23), Article 6048. https://doi.org/10.3390/rs14236048

Ren, S., Ghaffar, B., Mubbin, M., Haseeb, M., Tahir, Z., Hassan, S. S., Kucher, D. E., Kucher, O. D., & Abdullah-Al-Wadud, M. (2025). Multisensor remote sensing and AI-driven analysis for coastal and urban resilience classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *18*, 9166-9180. https://doi.org/10.1109/JSTARS.2025.3554793

Sabato, G., Scardino, G., Kushabaha, A., Casagrande, G., Chirivì, M., Fontolan, G., Fracaros, S., Luparelli, A., Spadotto, S., & Scicchitano, G. (2024). Remote measurement of tide and surge using a deep learning system with surveillance camera images. *Water*, *16*(10), Article 1365. https://doi.org/10.3390/w16101365

Santos-Ferreira, A. M., Pinelo, J., da Silva, J. C. B., Johannessen, J. A., Chapron, B., Gommenginger, C., Magalhães, J. M., Buijsman, M., Forget, G., Pineda, J., Goh, E., Ávila, M., Diogo, I., & Gonçalves, J. (2025). The internal waves service workshop: Observing Internal waves globally with deep learning and synthetic aperture radar. *Bulletin of the American Meteorological Society*, *106*(7), E1462-E1470. https://doi.org/10.1175/bams-d-25-0133.1

Saoudi, E. M., Jaafari, J., & Andaloussi, S. J. (2023). Advancing human action recognition: A hybrid approach using attention-based LSTM and 3D CNN. *Scientific African*, *21*, Article e01796. https://doi.org/10.1016/j.sciaf.2023.e01796

Shi, Y., Xu, G., Liu, Y., Chen, H., Zhou, S., Yang, J., Dong, C., Lin, Z., & Wu, J. (2025). Deep learning-based real-time surf detection model during typhoon events. *Remote Sensing*, *17*(6), Article 1039. https://doi.org/10.3390/rs17061039

Stringari, C. E., Guimarães, P. V., Filipot, J.-F., Leckler, F., & Duarte, R. (2021). Deep neural networks for active wave breaking classification. *Scientific Reports*, *11*(1), Article 3604. https://doi.org/10.1038/s41598-021-83188-y

Talpur, K., Hasan, R., Gocer, I., Ahmad, S., & Bhuiyan, Z. (2025). AI in maritime security: Applications, challenges, future directions, and key data sources. *Information*, *16*(8), Article 658. https://doi.org/10.3390/info16080658

Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the 36*th *international conference on machine learning* (pp. 6105-6114). ICML.

Tan, N. S. Y., Tham, M. L., Chua, S. Y., & Lee, Y. L. (2024). Accurate and fast classification of natural disasters using CNN-LSTM and inference acceleration. *Journal of Communications Software and Systems*, *20*(1), 58-68. https://doi.org/10.24138/jcomss-2023-0129

Tripathi, S. P., Chapron, B., Collard, F., Guitton, G., Lopez-Radcenco, M., Mouche, A., & Fablet, R. (2024). Deep learning inversion of ocean wave spectrum from SAR satellite observations. In *ICASSP 2024-2024 IEEE international conference on acoustics, speech and signal*

processing (pp. 8711-8715). IEEE. https://doi.org/10.1109/ICASSP48485.2024.10446834

Wang, H.-M., Doong, D.-J., & Lai, J.-W. (2025). Rip current identification in optical images using wavelet transform. *Journal of Marine Science and Engineering*, *13*(4), Article 707. https://doi.org/10.3390/jmse13040707

Younesi, A., Ansari, M., Fazli, M., Ejlali, A., Shafique, M., & Henkel, J. (2024a). A comprehensive survey of convolutions in deep learning: Applications, challenges, and future trends. *IEEE Access*, *12*, 41180-41218. https://doi.org/10.1109/ACCESS.2024.3376441

Younesi, A., Ansari, M., Fazli, M., Ejlali, A., Shafique, M., & Henkel, J. (2024b). Corrections to "A comprehensive survey of convolutions in deep learning: Applications, challenges, and future trends". *IEEE Access*, *12*, 112180. https://doi.org/10.1109/ACCESS.2024.3430149

Zhang, S., Cui, L., Dong, Z., & An, W. (2024). A deep learning classification scheme for PolSAR image based on polarimetric features. *Remote Sensing*, *16*(10), Article 1676. https://doi.org/10.3390/rs16101676