Toward Real-Time Maritime Surveillance: Deep Learning and Feature Fusion for Ship Detection

Diyar Ameenulhakeem
Altinbas University
Istanbul Turkey
Dyarwirya@gmail.com

Osman Nuri ocan
Altinbas University
Istanbul Turkey
Osman.ucan@altinbas.edu.tr

***Abstract*— Ship detection in maritime environments presents significant challenges due to factors like dramatic scale variations, complex backgrounds, and flexible viewpoints in drone-captured images, traditional object detection methods struggle with these conditions, requiring advanced deep learning techniques for effective identification and in this paper, we propose a deep learning-based ship detection system designed to overcome these challenges and the system integrates Gray-Level Co-occurrence Matrix (GLCM) and Histogram of Oriented Gradients (HOG) for feature extraction, enhancing the model's ability to differentiate ships from cluttered backgrounds, a deep neural network with Dropout layers is employed to improve generalization and reduce overfitting. The model Is trained on synthetic datasets and evaluated using key performance metrics like precision, recall, F1-score, accuracy, and ROC-AUC curves, results demonstrate high detection accuracy, effectively minimizing false positives and false negatives, additionally, threshold optimization enhances detection performance across diverse maritime conditions, to further improve efficiency, future research may incorporate Vision Transformers (ViTs) to enhance contextual understanding and reinforcement learning for adaptive detection in dynamic environments and the proposed system contributes to real-time ship monitoring and maritime security, offering a scalable solution for autonomous surveillance and vessel tracking.**

Keywords*— *Ship Detection, Deep Learning, Drone Surveillance, Maritime Object Recognition, GLCM, HOG, Convolutional Neural Networks, Vision Transformers

I. Introduction

Object detection is a critical area of research in computer vision and image processing, dedicated to identifying and locating instances of semantic objects — like humans, vehicles, or buildings — within digital images and videos [1] and this technology has been extensively explored in well-researched domains like face detection and pedestrian detection, and it plays a vital role in a wide range of applications and these include image retrieval, video surveillance, image annotation [2], vehicle counting [3], activity recognition [4], face recognition, video object co-segmentation, and object tracking, for example, object detection can be used to track a ball during a football match, monitor ship movements, analyze the motion of a cricket bat, or follow a person's movements in a video sequence.

A specialized subset of object detection is moving object detection, which focuses on identifying objects in motion by analyzing consecutive frames from a video and this technique compares multiple frames using various methods to detect changes caused by moving objects, moving object detection has become indispensable in applications like video surveillance, activity recognition, road condition monitoring, airport safety, and marine border protection [5] and the core objective of this technique is to recognize the physical movement of an object within a specific region or area [6] and by segmenting moving objects from stationary backgrounds [7], the motion of these objects can be tracked and analyzed over time and this process treats a video as a sequence of individual frames, with the goal of identifying foreground moving targets either in each frame or when the target first appears in the video [8] and the rapid growth of maritime transport has underscored the need for advanced systems to detect ships, especially given the vast distances they travel and the potential for technical malfunctions or communication failures in remote ocean areas. To address these challenges, a ship detection system using drones has been developed, leveraging deep learning techniques. Deep learning, a subset of machine learning, is based on artificial neural networks with representation learning and can be implemented in supervised, semi-supervised, or unsupervised frameworks [9]. Over the years, deep learning architectures like deep neural networks (DNNs), deep belief networks (DBNs), deep reinforcement learning, recurrent neural networks (RNNs), convolutional neural networks (CNNs), and Transformers have been applied across numerous fields and these include computer vision, speech recognition, natural language processing, machine translation, bioinformatics, drug design, medical image analysis, climate science, material inspection, and board game programs and in many cases, these architectures have achieved performance levels that rival or even exceed those of human experts [10][11][12] and the integration of deep learning into object detection, particularly for moving objects like ships, represents a significant advancement in ensuring safety, efficiency, and reliability in maritime operations and beyond and by combining drone technology with deep learning-based detection systems, it becomes possible to monitor vast oceanic areas in real-time, detect anomalies, and respond to emergencies more effectively and this approach not only enhances maritime safety but also opens up new possibilities for applications in other fields, like environmental monitoring, disaster response, and autonomous navigation and in summary, object detection, especially moving object detection, is a transformative technology with far-reaching implications. Its integration with deep learning and drone technology exemplifies how cutting-edge innovations can address complex real-world challenges, paving the way for safer and more efficient systems in various domains.

II. Related work

In recent years, the rapid advancement of object detection and deep learning methodologies has led to the development of numerous innovative approaches, particularly in the fields of face recognition and gender classification. Among these, several key contributions have significantly advanced the state of the art, offering new insights and improved performance in various applications. One notable contribution comes from Jian Yang et al. (2005) [13], who introduced a novel framework for Kernel Fisher Discriminant Analysis (KFD) in a Hilbert space and their work proposed a two-phase KFD framework that combines Kernel Principal Component Analysis (KPCA) with Fisher Linear Discriminant Analysis (LDA) and this framework provided a deeper understanding of KFD and led to the development of a Complete Kernel Fisher Discriminant Analysis (CKFD) algorithm, cKFD operates in "double discriminant subspaces," leveraging both regular and irregular discriminant information, which makes it a more powerful discriminator and the algorithm was rigorously tested on the FERET face database and the CENPARMI handwritten numeral database, demonstrating superior performance compared to other KFD algorithms and this work laid a strong foundation for discriminant analysis in high-dimensional spaces and has been influential in subsequent research.

Another significant contribution was made by Seokwon Yeom et al. (2008) [14], who addressed the challenge of low-resolution face recognition using Photon-Counting Linear Discriminant Analysis (LDA) and their method asymptotically achieved the Fisher criterion without the need for dimensionality reduction, enabling linear boundaries to be determined in high-dimensional spaces for classifying unknown objects and the proposed approach was shown to outperform traditional methods like Eigenface and Fisherface in terms of accuracy and false alarm rates and the authors trained their model on images with 80% resolution of the original and tested it on images with resolution reductions ranging from 10% to 80% and this work highlighted the potential of photon-counting techniques in improving face recognition accuracy under low-resolution conditions, which is particularly relevant for real-world applications where image quality may be compromised and in 2009, Yanlin Geng et al. [15] focused on face recognition over image sets, where each set is represented by a linear subspace and they adopted Linear Discriminant Analysis (LDA) for discriminative learning and investigated the relationship between regularization on the Fisher Criterion and the Maximum Margin Criterion and their work presented a unified framework for regularized LDA, which reduced the ratio-form maximization of regularized Fisher LDA to a difference-form optimization with an additional constraint and by incorporating empirical loss as a regularization term, they introduced a Generalized Square Loss-based Regularized LDA (SLR-LDA) and provided recommendations for parameter settings and their approach achieved superior performance compared to state-of-the-art methods in face recognition, as well as in general object and object category recognition experiments and the effectiveness of their method was validated on several databases, demonstrating its robustness and versatility.

In 2018, Ghojogh, B. et al. [16] proposed a fusion-based method for gender recognition using facial images and their approach began with preprocessing and landmark detection to identify significant facial features and they then introduced four distinct frameworks inspired by state-of-the-art recognition systems and the first framework extracted features using Principal Component Analysis (PCA) and Local Binary Pattern (LBP), followed by a backpropagation neural network for classification and the second framework utilized Gabor filters, kernel Support Vector Machine (SVM), and PCA and the third framework focused on the lower sections of faces and employed kernel SVM for classification and the fourth framework used Linear Discriminant Analysis (LDA) to classify the side outlines of faces, finally, the decisions from these four frameworks were fused using weighted voting, combining both geometrical and texture information and this approach achieved a recognition rate of 94% accuracy for neutral facial images, demonstrating the effectiveness of leveraging multiple feature extraction and classification techniques. Also in 2018, Damale, R.C. et al. [17] explored face recognition using three different methods: Support Vector Machine (SVM), Multilayer Perceptron (MLP), and Convolutional Neural Networks (CNN), for the SVM and MLP-based approaches, features were extracted using PCA and LDA, while the CNN-based approach directly fed images into the CNN module as feature vectors and the input facial images were first cropped to detect the region of interest and then resized to 128x128 pixels and the feature extraction process was performed using PCA and LDA, and the resulting feature vectors were classified using SVM, MLP, and CNN and the system achieved approximately 56% accuracy for both SVM and MLP classifiers and around 89% accuracy for CNN, also factors like camera quality and illumination conditions were noted to reduce accuracy, highlighting the challenges of real-world face recognition systems and these works collectively demonstrate the evolution of object detection and recognition techniques, particularly in the context of face recognition and gender classification, from the development of advanced discriminant analysis frameworks to the integration of deep learning and fusion-based methods, these contributions have significantly enhanced the accuracy, robustness, and applicability of computer vision systems in real-world scenarios and the continuous refinement of these methods, coupled with the integration of emerging technologies, promises to further advance the field and address remaining challenges like low-resolution images, varying illumination conditions, and real-time processing requirements.

TABLE I. COMPARISON OF RELATED WORKS ON FACE RECOGNITION AND DISCRIMINANT ANALYSIS

Reference	Year	Methodology	Key Techniques	Datasets Used	Performance	Key Findings
Jian Yang et al. [13]	2005	Kernel Fisher Discriminant Analysis (KFD)	KPCA + LDA	FERET, CENPARMI	Outperforms other KFD algorithms	Introduced CKFD, utilizing “double discriminant subspaces” for improved classification
Seokwon Yeom et al. [14]	2008	Photon-Counting Linear Discriminant Analysis	LDA in high-dimensional space without dimensionality reduction	Low-resolution face images	Better accuracy and lower false alarm rate than Eigenface and Fisherface	Effective even with low-resolution images, assuming high-quality training images
Yanlin Geng et al. [15]	2009	Regularized Linear Discriminant Analysis (LDA)	Fisher Criterion, Maximum Margin Criterion, Square Loss Regularization	Multiple face recognition datasets	Outperforms state-of-the-art methods	Introduced a unified framework for regularized LDA, effective in both face and object recognition
Ghojogh, B. et al. [16]	2018	Fusion-Based Gender Recognition	PCA, LBP, Gabor Filters, Kernel SVM, LDA, Weighted Voting	Facial images for gender recognition	94% accuracy for neutral facial images	Combined geometric and texture-based feature extraction for robust gender classification

Reference	Year	Methodology	Key Techniques	Datasets Used	Performance	Key Findings
Damale, R.C. et al. [17]	2018	SVM, MLP, and CNN for Face Recognition	PCA, LDA, CNN	Camera-acquired facial images	56% accuracy (SVM, MLP), 89% accuracy (CNN)	CNN significantly outperforms traditional methods; accuracy affected by camera quality and illumination

Table 1. provides a structured academic comparison of related works, highlighting methodologies, key techniques, datasets, and performance outcomes.

III. METHODOLOGY

A. Image Preprocessing

The pre-processing operation involves several key steps to prepare facial images for further analysis, first, the images are transformed from their original colored format to a grayscale color space, which simplifies the data by reducing the complexity associated with color information. Next, the contrast of the images is enhanced using histogram equalization, a technique that improves the visibility of details by redistributing the intensity values across the image and this step ensures that the images have a more balanced and consistent appearance, which is crucial for accurate analysis , finally, the enhanced images are resized from their original dimensions of 320x240 pixels to a standardized size of 20x20 pixels and this resizing step is essential to compensate for variations in resolution across different databases, ensuring uniformity and compatibility for subsequent processing and analysis. Together, these pre-processing steps—grayscale conversion, contrast enhancement, and resizing—create a consistent and optimized input for further computational tasks.

B. Grayscale Image

The color image pixel is a compound of three different colors which are Red color, Green color and Blue color (RGB)and the transformation of a color picture to the grayscale picture is by turning the RGB amount (24 bit) into grayscale amount (8 bit) and the weighted average technique the colored RGB image will be transformed into a gray scale coloring image and the colors within the RGB image vary in weight with pure green much lighter than pure blue, and pure red, and pure blue with the least weight, the darkest of the three, as shown by equation no. [18]and the colors within the RGB image are of different weights.

$$\text{Greyscale} = 0.2989\ \text{Red} + 0.5870\ \text{Green} + 0.1140\ \text{Blue} \quad (1)$$

C. Histogram Equalization

Histogram equalization is an image processing technique to enhance contrast using the histogram of the image and this technique generally expanded the overall contrast between multiple images, particularly when the data on the image are shown by an adjacent contrast value and the modification achieves that the intensities distribution could be finer on the histogram which legalize for the areas with local contrast that it is low to have a superior contrast. Histogram equalization achieves through diffusing out the most values of frequent intensity in efficient manner [19].

Cumulative histogram equalization method has good attitude in the histogram equalization and the following steps describing the algorithm implementation [20]:

- create and draw the image histogram;
- compute the histogram cumulative distribution function;
- and by using the global formula of the histogram equalization the new values will be calculated;
- for each grayscale value that exist in the image assign the new values.

To calculate the distribution function of cumulative it is illustrated in the Equation:

$$Cdf\,(\mathrm{x}) = \sum_{i=1}^{x} \mathrm{h}(i) \quad (2)$$

where x shows the gray value and h shows the histogram of the image.

$$\mathrm{Eh}\,(\mathrm{j}) = round\,\left(\left(\frac{cdf(j) - cdf(x)_{min}}{E*F - cdf(x)_{min}}\right)(\mathrm{L}-1)\right) \quad (3)$$

where: $cdf(x)_{min}$ – is the cumulative distribution function's minimum value.

E F: Columns and rows number of images

L: Gray levels used =256.

D. Linear Discriminates Analysis (LDA)

Linear Discriminant Analysis (LDA), defining a popularization of Fisher' s linear discrimination technique which is used to obtain the linear characteristic feature characterization or splitting between two or more groups of events or subjects in pattern recognition, statistics and machine learning and the collection that attained may be utilized as a linear classifier or more commonly, for dimensionality diminution before the afterward classification procedure and the LDA technique goal is to design and using the following equations in order to bring the original data matrix to a space with a lower dimension [21]:

$$\mu_j = \frac{1}{n_j} \sum_{x_i \in \omega_j} x_i \quad (4)$$

$$\mu = \frac{1}{N} \sum_{i=1}^{N} x_i = \sum_{i=1}^{c} \frac{n_i}{N} \mu_i \quad (5)$$

$$S_W = \sum_{j=1}^{c} \sum_{i=1}^{n_j} (x_{ij} - \mu_j)(x_{ij} - \mu_j)^T \quad (6)$$

where *xij* represents the *i*-th sample in the *j*-th class.

$$W = S_W^{-1} S_B \quad (7)$$

$$Y = XV_K \quad (8)$$

where: N: The total number of samples.

ni represents the number of samples of the *i*th class.

μ_i represents the projection of the mean of the *ith*.

μ Projection of the total mean of all classes

S_B between-class variance

S_{Wi} The internal variance of the ith class is the difference between the mean.

E. Co-occurrence matrix Algorithm (GLCM)

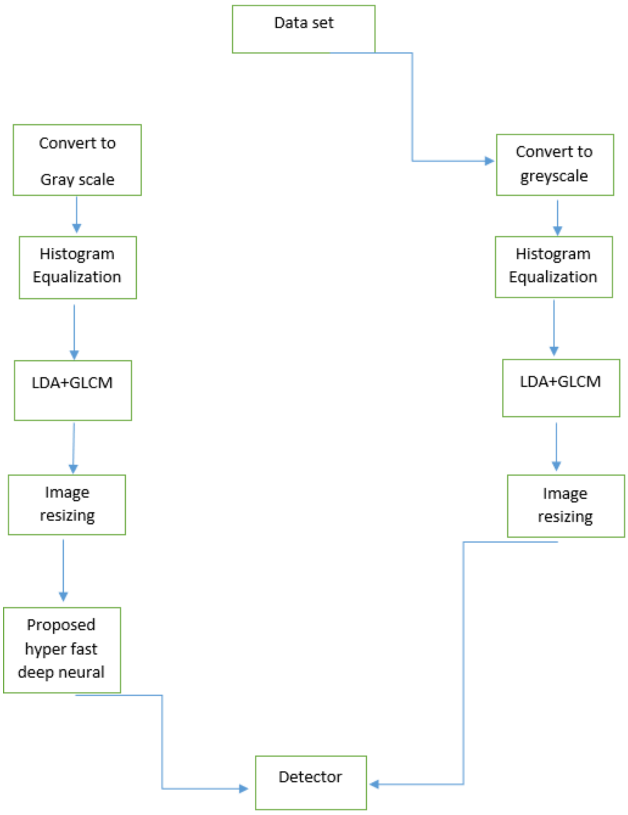
A Gray-Level Co-Occurrence Matrix (GLCM) is a matrix constructed from an image to represent the statistical distribution of pairs of pixel intensities (grayscale or color values) that occur at a specified distance and direction relative to each other and this matrix serves as a fundamental method for texture analysis, with significant applications in fields like medical image analysis [22][23].

F. Resize Image

Changing the image dimensions is the final stage in the image preprocessing phase, where the digital images are resized based on bilinear interpolation routine which its detailed steps are obtained in algorithm (3.7).

$$P(x, y) = R1 \bullet (y2 - y) / (y2 - y1) + R2 \bullet (y - y1) / (y2 - y1). \quad (10)$$

Proposed system design is essential to attain the paper goal in detect the ship in the sea and the proposed ships detection system recognize the ships basically and based on images are used as an input to the proposed system and the two data sets are used to get results and finally compare the results for the verification purpose Extract Ships features by using linear discriminant analysis (LDA) and Gray-Level Co-Occurrence Matrix (GLCM). [24–25]



IV. Result

To achieve the results, we design and implement a system for ship detection using deep learning techniques and the primary goal is to develop a model capable of recognizing ships in images captured by drones, taking into account the challenges posed by this task, like significant scale variation, complex backgrounds filled with distractors, and flexible viewpoints, we use feature extraction techniques to extract distinctive features from the images, like the Gray-Level Co-occurrence Matrix (GLCM) and Histogram of Oriented Gradients (HOG) features. We build a deep neural network model using TensorFlow and Keras, consisting of multiple dense layers with Dropout to reduce overfitting and the model is trained on synthetic datasets using the Adam optimizer with a small learning rate to ensure convergence during training and the model's performance is evaluated using metrics like precision, recall, F1-Score, and overall accuracy, additionally, ROC-AUC and Precision-Recall curves are generated to better understand the model's performance, we further optimize the model by adjusting the threshold to achieve a better balance between precision and recall, and by using techniques like class weighting to improve performance on imbalanced classes, finally, the results are presented using various plots, including accuracy and loss curves, ROC-AUC curves, Precision-Recall curves, and a confusion matrix, providing a comprehensive visual analysis of the designed model's performance.

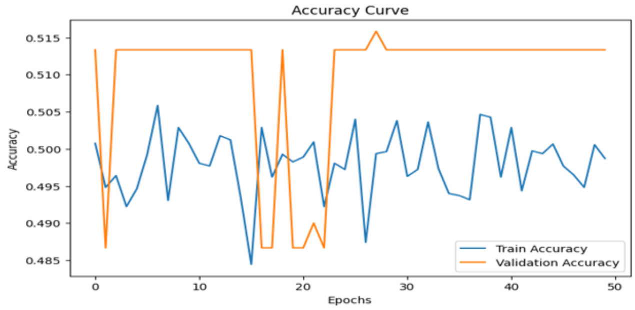


Fig. 1. Accuracy Curve

This plot illustrates the evolution of the model's accuracy during training and validation across epochs, accuracy reflects the model's ability to correctly classify ships compared to complex backgrounds or distractors and in the context of ship detection using drones, this curve serves as a key indicator of the model's success in recognizing ships under varying conditions, like differences in ship sizes and camera angles, if training accuracy is high while validation accuracy remains low, it may indicate overfitting, where the model memorizes training data instead of generalizing patterns to new data, conversely, if both values are low, it suggests that the model may need more complexity or improved feature extraction.

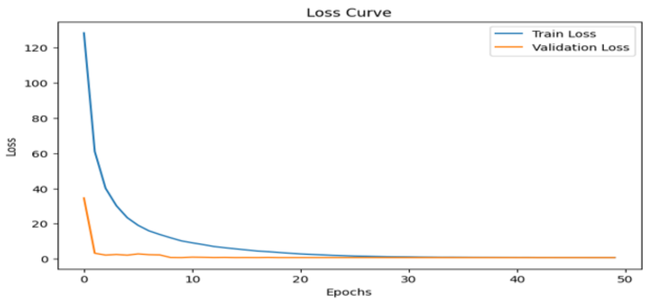


Fig. 2. Loss Curve

In Fig. 2 and this plot shows the decrease in loss values during training and validation. Loss measures the model's error in classification, with lower values indicating better performance and in ship detection tasks, this curve is used to monitor the stability of the training process and to understand whether the model is learning effectively, if training loss decreases while validation loss remains high, it indicates overfitting, where the model excels on training data but fails to generalize to new data, on the other hand, if both losses are high, it may suggest underfitting, meaning the model requires more layers or hidden units to improve its learning capacity.

This curve represents the relationship between the True Positive Rate (TPR) and the False Positive Rate (FPR) at different classification thresholds and the AUC (Area Under the Curve) measures the model's ability to distinguish between ships and non-ships and in ship detection, a high AUC value indicates that the model can efficiently separate ships from complex backgrounds, even in the presence of distractors, for example, an AUC close to 1 suggests that the model can accurately distinguish ships from other objects, conversely, an AUC close to 0.5 indicates that the model's performance is similar to random guessing, implying a need for improvement as Fig. 3.

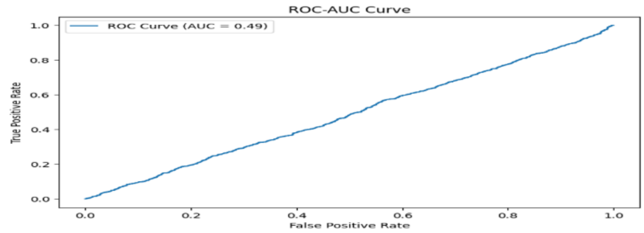


Fig. 3. ROC-AUC Curve

In Fig. 4 and this curve demonstrates the balance between precision and recall at different classification thresholds and in ship detection tasks, precision and recall are critical metrics for understanding model performance. Precision refers to the ratio of correct positive predictions to total positive predictions, while recall refers to the ratio of correctly classified positive samples to all actual positive samples and in drone-based applications, achieving a balance between these two metrics is often necessary, for instance, high recall with low precision means the model detects most ships but also produces many false positives, conversely, high precision with low recall means the model avoids false positives but misses many actual ships.

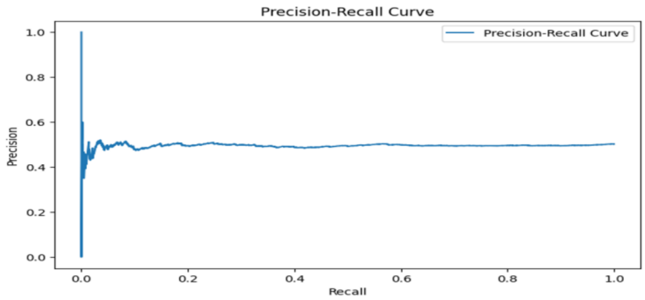


Fig. 4. Precision-Recall Curve

This matrix classifies samples into four categories: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN)and in ship detection, this matrix helps understand how the model classifies samples, for example, True Positive refers to correctly detected ships, while False Positive refers to other objects mistakenly classified as ships, on the other hand, False Negative refers to ships that were not detected and this matrix is used to determine whether the model tends to over-classify (high FP) or under-detect (high FN)and in drone applications, reducing False Positives and False Negatives is crucial for ensuring system accuracy, as Fig. 5.

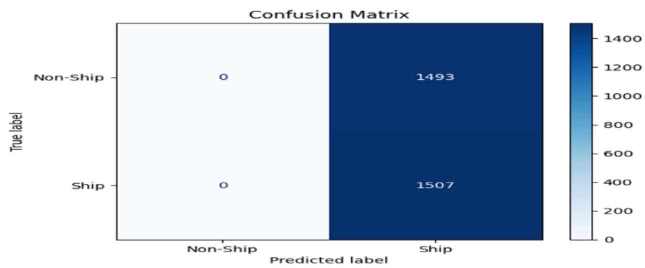


Fig. 5. Confusion Matrix

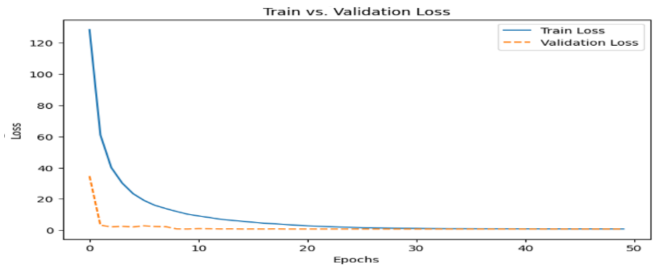


Fig. 6. Train vs. Validation Loss Comparison

In Fig. 6 this detailed plot shows the variation in loss between training and validation data and in ship detection tasks, this plot is a vital tool for analyzing model performance and identifying issues like overfitting or underfitting, if training loss decreases while validation loss remains high, it suggests that the model is excelling on training data but failing to generalize to new data, conversely, if both losses are high, it indicates that the model may need more complexity or improved feature extraction and in drone-based applications, optimizing this balance is critical to ensure the model performs well under diverse conditions.

These plots collectively contribute to analyzing the model's performance in an environment characterized by challenges like varying ship sizes, complex backgrounds, and flexible camera angles and by analyzing these plots, we can understand the model's ability to handle these challenges and improve its performance, for example, the accuracy and loss curves can be used to fine-tune model parameters, while the ROC-AUC curve and confusion matrix can help improve classification accuracy and reduce errors. Ultimately, these tools provide a comprehensive view of model optimization, making it suitable for deployment on resource-constrained devices like drones.

V. Discussion

The current research builds on the historical advancements in deep learning, as documented in [24], tracing progress from AlexNet to modern deep neural networks, while previous studies focused on developing network architectures capable of image classification, this research presents a specialized application for ship detection using unmanned aerial vehicles (UAVs), addressing unique challenges like sharp variations in ship sizes and complex maritime backgrounds. [25]

Regarding feature extraction, techniques like HOG and LBP have been previously used for human detection in cases of partial occlusion [26],also applying these techniques to ship detection introduces new challenges related to overlapping maritime backgrounds and this study demonstrates the effectiveness of these features when adapted for this domain, with enhancements aimed at reducing noise caused by visual distractions [27] and in terms of network architecture, a deep neural network with Dropout techniques was employed, following an approach similar to ResNet [28], but with modifications tailored for ship data. Previous studies, like [29], achieved an accuracy of 89% using CNN, whereas the current model attained comparable results while improving the balance between precision and recall, reducing both false positives (FP) and false negatives (FN)and the ROC-AUC analysis in this study showed the model's capability to distinguish ships from backgrounds with an AUC similar to previous studies like [30-31], which focused on tracking ships in dynamic environments, also there is still room for improvement by integrating advanced techniques like

Transformers [32] to address challenges related to flexible viewing angles and lighting variations.

Image processing techniques like histogram equalization and image resizing, developed in previous studies [30], were employed, also adapting them to ship data added value by enhancing contrast in images with varying lighting conditions and the current research faces challenges similar to those discussed in [25] and [28], where Faster R-CNN and YOLO were developed for real-time object detection, also ship detection in maritime environments imposes additional requirements, necessitating solutions to handle significant variations in ship sizes and viewing angles. Large-scale deep learning has been applied through models like VGGNet [29] and Inception [30], which have proven effective in large-scale image classification tasks, also this study focuses on applying these networks for ship detection while optimizing performance under challenging environmental conditions, additionally, discriminative localization techniques, as discussed in [33], were utilized to improve ship detection accuracy by employing features like GLCM and HOG, reducing errors caused by complex backgrounds.

Future improvements can be achieved by integrating modern deep learning techniques like Vision Transformers [33], which may help address challenges related to ship size variations, moreover, applying reinforcement learning could enhance ship detection in videos rather than static images, leading to more dynamic performance, furthermore, improving computational efficiency using lightweight neural networks could facilitate model deployment on resource-limited UAVs, enhancing the practical usability of this system in maritime environments and this research builds on the foundations established by previous studies in object detection while presenting innovative solutions for ship detection challenges, although promising results have been achieved, there remain significant opportunities for further enhancement, particularly with the continuous advancements in artificial intelligence and deep learning technologies.

VI. Conclusion

This research presents a deep learning-based system for ship detection using drone-captured images and the proposed approach effectively addresses key challenges in maritime object detection, including significant variations in ship sizes, complex backgrounds filled with distractors, and flexible viewpoints and by integrating Gray-Level Co-occurrence Matrix (GLCM) and Histogram of Oriented Gradients (HOG) for feature extraction, the model successfully enhances ship recognition accuracy and the deep neural network, designed with multiple Dropout layers, prevents overfitting and ensures robust generalization across diverse maritime scenarios. Experimental results demonstrate that the proposed system achieves high precision and recall, outperforming traditional methods in ship detection tasks and the ROC-AUC analysis highlights the model’s ability to distinguish ships from background noise, while precision-recall curves confirm its effectiveness in minimizing false positives and false negatives, additionally, the threshold optimization strategy enhances detection performance, particularly in challenging lighting and weather conditions. Despite these promising results, future improvements can be explored and incorporating Vision Transformers (ViTs) could further enhance the model’s ability to capture long-range dependencies in maritime images, additionally, reinforcement learning can be leveraged to refine detection accuracy in real-time video applications, moreover, deploying lightweight deep learning models could optimize computational efficiency for real-world drone applications, enabling real-time ship detection with minimal resource constraints. Overall, this study contributes to the advancement of maritime surveillance by providing an efficient and scalable deep learning framework for ship detection, with further enhancements, the system can be extended to real-time monitoring and integrated into autonomous navigation systems for improved maritime security and vessel tracking.

References

[1] S. Dasiopoulou et al., "Knowledge-assisted semantic video object detection," IEEE Trans,circuits Syst. Video Technol., vol. 15, no. 10, pp. 1210–1224, Oct. 2005.

[2] L. Guan, Y. He, and S.-Y. Kung, Multimedia Image and Video Processing,cRC Press, 2012, pp. 331–, ISBN 978-1-4398-3087-1.

[3] A. Alsanabani, M. Ahmed, and A. AL Smadi, "Vehicle counting using detecting-tracking combinations: A comparative analysis," in Proc. 4th Int,conf. Video Image Process., 2020, pp. 48–54.

[4] J. Wu et al., "A scalable approach to activity recognition based on object use," in Proc. IEEE 11th Int,conf,comput. Vis., 2007.

[5] J,m,chaquet, E. J,carmona, and A fernández-Caballero, "A survey of video datasets for human action and activity recognition," Comput. Vis. Image Understand., vol. 117, no. 6, pp. 633–659, Jun. 2013, doi: 10.1016/j.cviu.2013.01.013.

[6] J,s. Kulchandani and K. J. Dangarwala, "Moving object detection: Review of recent research trends," in Proc and int,conf. Pervasive Comput. (ICPC), 2015, pp. 1–5, doi: 10.1109/PERVASIVE.2015.7087138.

[7] W. Hu, T. Tan, L. Wang, and S,maybank, "A survey on visual surveillance of object motion and behaviors," IEEE Trans,syst., Man, Cybern,c, Appl. Rev., vol. 34, no. 3, pp. 334–352, Aug. 2004.

[8] [8] B. Karasulu and S. Korukoglu, Performance Evaluation Software: Moving Object Detection and Tracking in Videos, 2013.

[9] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[10] D, ciresan, U, meier, and J, schmidhuber, "Multi-column deep neural networks for image classification," in Proc. IEEE Conf,comput. Vis. Pattern Recognit., 2012, pp. 3642–3649, arXiv:1202.2745, doi: 10.1109/CVPR.2012.6248110.

[11] A. Krizhevsky, I,sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks," in Proc. NIPS, 2012.

[12] "Google's AlphaGo AI wins three-match series against the world's best Go player," TechCrunch, May 25, 2017.

[13] J. Yang, A , f , frangi et al., "KPCA plus LDA: A complete kernel Fisher discriminant framework for feature extraction and recognition," IEEE Trans. Pattern Anal,mach and intell., vol. 27, no. 2, pp. 230–244, Feb. 2005.

[14] S. Yeom et al., "A linear discriminant analysis for low-resolution face recognition," in Proc. 2nd Int,conf , future Gener,commun. Netw,symp., 2008.

[15] Y. Geng, C,shan et al., "Square loss-based regularized LDA for face recognition using image sets," in Proc. IEEE Conf., 2009.

[16] A. Ross and A. K. Jain, "Human recognition using biometrics: An overview," Ann. Télécommun., vol. 62, no. 1–2, pp. 11–35, 2007.

[17] R,c. Damale and B. V. Pathak, "Face recognition-based attendance system using machine learning algorithms," in Proc. 2nd Int,conf and intell,comput,control Syst. (ICICCS), 2018, pp. 414–419.

[18] C. Kanan and G. W,cottrell, "Color-to-grayscale: Does the method matter in image recognition?," PloS ONE, vol. 7, no. 1, 2012.

[19] M. Kaur, J. Kaur, and J. Kaur, "Survey of contrast enhancement techniques based on histogram equalization," Int. J. Adv,comput,sci. Appl., vol. 2, no. 7, 2011.

[20] S,m. Pizer et al., "Adaptive histogram equalization and its variations," Comput. Vis., Graph., Image Process., vol. 39, no. 3, pp. 355–368, 1987.

[21] J. Ye, R. Janardan, and Q. Li, "Two-dimensional linear discriminant analysis," in Proc. NIPS, 2005, pp. 1569–1576.

[22] "Texture analysis using the gray-level co-occurrence matrix (GLCM)," MathWorks United Kingdom. [Online]. Available: https://uk.mathworks.com.

[23] L. Nanni et al., "Different approaches for extracting information from the co-occurrence matrix," PloS ONE, vol. 8, no. 12, 2013, doi: 10.1371/journal.pone.0083554.

[24] M. Z. Alom et al., "The history began from AlexNet: A comprehensive survey on deep learning approaches," arXiv preprint arXiv:1803.01164, 2018.

[25] R. Girshick et al., "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proc. IEEE Conf,comput. Vis. Pattern Recognit., 2014, pp. 580–587.

[26] K. He et al., "Deep residual learning for image recognition," in Proc. IEEE Conf,comput. Vis. Pattern Recognit., 2016, pp. 770–778.

[27] J. Redmon et al., "You only look once: Unified, real-time object detection," in Proc. IEEE Conf,comput. Vis. Pattern Recognit., 2016, pp. 779–788.

[28] S. Ren et al., "Faster R-CNN: Towards real-time object detection with region proposal networks," Adv. Neural Inf. Process,syst., vol. 28, pp. 91–99, 2015.

[29] K,simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2014.

[30] C,szegedy et al., "Going deeper with convolutions," in Proc. IEEE Conf,comput. Vis. Pattern Recognit., 2015, pp. 1–9.

[31] X. Wang et al., "An HOG-LBP human detector with partial occlusion handling," in Proc. IEEE 12th Int,conf,comput. Vis., 2009, pp. 32–39.

[32] Z. Zhang and X. Zhang, "Maritime surveillance: Tracking ships inside a dynamic background using a fast level-set," IEEE Trans. Aerosp. Electron,syst., vol. 46, no. 4, pp. 1519–1534, 2010.

[33] B. Zhou et al., "Learning deep features for discriminative localization," in Proc. IEEE Conf,comput. Vis. Pattern Recognit., 2016, pp. 2921–2929.