ENHANCED SHIP DETECTION IN SATELLITE IMAGES USING HYBRID DEEP LEARNING AND MACHINE LEARNING APPROACHES

Smit Patel[*], *Tirth Halvadia*[*], *Nidhi Patel*[*]

[*]Department of Electrical & Computer Engineering, University of Waterloo, Canada

ABSTRACT

An essential function of marine surveillance systems is the automatic identification of ships in remote sensing photographs, particularly for early identification of potential pirate threats. This paper addresses the critical need for highly accurate ship detection systems to enhance national security. Existing solutions often face limitations in accuracy due to small datasets or require extensive computational resources. The approach we propose is to use a combination of machine learning and deep learning techniques to identify ships in port areas. Our approach combines established machine learning algorithms, such as Support Vector Machines (SVM), AdaBoost, XGBoost, Random Forest, and Stochastic Gradient Boosting, with pre-trained convolutional neural networks (CNNs) and transfer learning models. These include VGG-16 trained on ImageNet, ResNet-50, and a combination of VGG-16 and AlexNet. This combined strategy achieves a remarkable accuracy of 99.375%, significantly surpassing or matching the accuracy of existing solutions. The proposed approach offers a promising advance while potentially requiring less computational resources compared to alternative methods.

Index Terms— Hybrid Models, Machine Learning, Deep Learning, Ship Detection, Satellite Images

1. INTRODUCTION

The shipping industry is a major facilitator of international trade, accounting for an amazing 90% of all worldwide goods exchange. Its importance goes beyond only helping businesses engaged in production; it also makes it possible for products to be delivered directly to customers, acting as a cornerstone of the world economy. But there are drawbacks to this quick expansion of maritime travel as well, like increased traffic infractions and an increase in illicit operations like smuggling, sea jacking, and maritime terrorism [1]. Efforts must be made to ensure the security and safety of marine operations due to the considerable risk that these threats pose to national, regional, and international economies [2].

To improve marine security management due to these difficulties, there has been an increasing focus on ocean vessel surveillance, especially in coastal areas. An increasingly important aspect of maritime surveillance is automatic ship detection, which has sparked a lot of computer vision research, especially using Deep Learning techniques. These techniques show potential in detecting illegal operations like oil spills, tracking maritime activity for defense and security, and enhancing marine safety [3].

Managing small and sometimes poorly annotated datasets presents a major difficulty in satellite imagery for ship detection. This is especially clear when using algorithms for unsupervised learning [4]. Notwithstanding this difficulty, during the past few decades, significant advancements have been made in the design and development of methods for identifying ships in satellite imagery.

This research article aims to provide a comprehensive review of recent advancements in ship detection from satellite imagery, with a focus on leveraging deep learning techniques. It proposes hybrid models that combine deep learning with traditional machine learning approaches to address the limitations of small datasets and enhance detection accuracy. The primary objectives of this research project are multifaceted: Firstly, it seeks to address the challenge of eliminating false positives generated by natural phenomena such as clouds and tidal waves in maritime photographs. Additionally, the project aims to enhance image clarity, with a particular focus on capturing the movement of ships and delineating distinctive wave lines.

The subsequent sections of this article delve into literature review, proposed methodology, experimental setups and results, and conclusion.

2. LITERATURE REVIEW

The automatic detection of ships from satellite data has undergone a revolutionary change according to the latest developments in deep learning. CNNs have emerged as the dominant approach in several studies, [5][6][10][11][17][19][20]. Researchers have successfully implemented various CNN architectures for ship detection, including ResNet [7][8][9], DenseNet [12] [19], U-Net [9], and different YOLO (You Only Look Once) versions [16].

Beyond achieving high accuracy, researchers are actively addressing challenges associated with low-resolution satellite images. For instance, studies explore using Generative Adversarial Networks (GANs) for image pre-processing to improve detection accuracy in such scenarios, [14]. Furthermore, domain adaptation strategies have been studied with encouraging outcomes to address the problem of blurry or damaged photos [13].

Some studies delve beyond just achieving high accuracy and explore explainable AI techniques to understand how CNNs make decisions regarding ship detection [18]. Furthermore, the integration of ship detection with Automatic Identification System (AIS) data is being investigated for enhanced maritime threat detection capabilities [15].

Table 1. Analysis of methods applied on same dataset

Ref.	Year	Approach	Performance
[5]	2019	CNN	97% accuracy
[6]	2019	CNN	98.15% accuracy
[7]	2019	ResNet & Inception	98.56% accuracy
[8]	2019	ResNet-50	99.5% accuracy
[9]	2019	ResNet-34 & U-Net	82.5 F-score
[10]	2020	CNN	99.6% accuracy
[11]	2020	SVM CNN	94.38% accuracy 97.89% accuracy
[12]	2020	DenseNet	90% accuracy
[13]	2020	Transferable Attention enhanced Adversarial Adaptation Network	91.15% accuracy
[14]	2021	EEGAN CNN RCNN Faster RCNN	NA 92% accuracy 96% accuracy 98% accuracy
[15]	2022	VGG-16	98% accuracy
[16]	2022	YOLOV3 YOLOV4 YOLOV5	97% accuracy 98% accuracy 99% accuracy
[17]	2022	CNN	97.12% accuracy
[18]	2023	CNN	98.83% accuracy
[19]	2023	CNN DenseNet	98.5% accuracy 98.48% accuracy
[20]	2023	CNN	85.1% accuracy

In conclusion, deep learning has significantly impacted ship detection from satellite images. Table 1 provides details of existing works applied to same dataset, with their approach. A notable gap exists in the exploration of hybrid machine learning and deep learning models for ship detection, as evidenced by the absence of studies integrating Machine Learning algorithms with Deep Learning approaches. As such, our research aims to fill this gap by proposing a hybrid model and investigating its efficacy in ship detection tasks. The exploration of hybrid models presents a new frontier in ship detection research, offering promising avenues for improving detection accuracy and robustness.

3. METHODOLOGY

The core methodology involves a two-stage process illustrated in Fig 1. In the first stage, the labeled dataset is fed into three deep learning/transfer learning models. These models are responsible for extracting informative features from the satellite images. The extracted features essentially capture the essential characteristics that differentiate ship-containing images from those without ships.

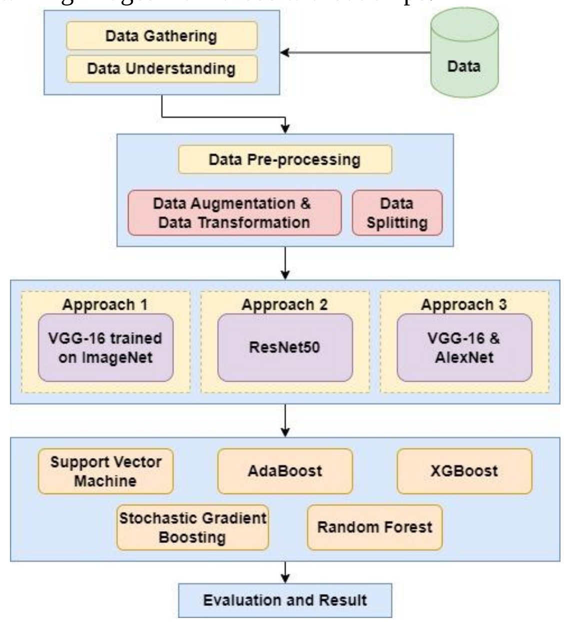

Fig. 1 Proposed Methodology

Five different machine learning models are fed the features that the deep/transfer learning models have extracted in the first step. These machine learning models act as classifiers. They utilize the features to learn the patterns that distinguish ship-containing images from non-ship images. By analyzing these patterns, the machine learning models can then classify new unseen satellite images as containing a ship or not.

The detailed inner workings of the specific deep learning/transfer learning models and the machine learning models used are likely elaborated on in upcoming sections of paper.

3.1 Dataset

The "Ships in Satellite Imagery" dataset [21], a set of 4,000 tagged photos created especially to support the creation of automated ship detection systems, is used in this study. The dataset was created from satellite images from Planet that was taken over the San Pedro Bay and San Francisco Bay areas.

There are two classes for the images in the dataset: "ship" and "no-ship." With one thousand photos, the "ship" class showcases a variety of ships in varying sizes, orientations, and climatic conditions. With 3,000 photos, the "no-ship" class is more varied. It consists of a random selection of land cover elements (water, flora, etc.), partial ship images that don't fit the "ship" class requirements, and even photographs that machine learning models had previously incorrectly classified.

Each image is provided in 80x80 RGB format and comes with a corresponding label ("ship" or "no-ship"). The dataset is also available in JSON format, offering additional information like scene IDs, locations, and the breakdown of pixel value data for each image channel (red, green, blue).

3.2 Data Pre-Processing

Within the JSON file "data" list, there is a single list of 19200 numbers that make up the dataset. We adhere to a particular order in order to obtain the precise channel information (red, green, blue). The red channel values are represented by the first 6400 elements in the list, the green channel values by the next 6400 elements, and the blue channel values by the last 6400 elements. We perform data augmentation on the original RGB satellite images to improve the robustness of the model for ship/non-ship classification and to enrich the training dataset. This entails performing stretching, rotation, and picture shearing operations. With this method, the 4,000-image original dataset is expanded to 12,800 shots for training and 3,200 shots for validation.

A crucial step in our approach involved splitting the dataset into training and validation sets to assess the model's accuracy comprehensively. An 80:20 splitting ratio was used, meaning that 80% of the data were set aside for the model's training. The remaining 20 percent was set aside for verification. During the training phase, this splitting method helps find the best model configuration.

3.3 Proposed Model Architecture

We compare the efficacy of multiple hybrid deep learning models with conventional machine learning techniques through benchmarking. ResNet-50, VGG-16 enhanced on ImageNet, and VGG-16 paired with AlexNet features are examples of the hybrid models. We also assess Random Forest, XGBoost, SVM, AdaBoost, and Stochastic Gradient Boosting.

We exploited these pre-trained models by loading them and configuring them to operate in a feature extraction mode, essentially bypassing the final classification layers. Subsequently, both the training and testing images will be fed through each pre-trained model independently. This process generates feature vectors for each image, encapsulating the essential characteristics learned by the respective models.

3.3.1 VGG-16 pretrained on ImageNet

In the first approach, we leveraged the expertise of a pre-trained VGG-16 model. VGG-16, meticulously trained on the vast ImageNet dataset, has developed exceptional capabilities for extracting informative features from images. To harness this power, we'll utilize a pre-trained VGG-16 model, but crucially exclude the final classification layers. This essentially transforms VGG-16 into a powerful feature extractor. These extracted features are passed to baseline machine models to evaluate performance.

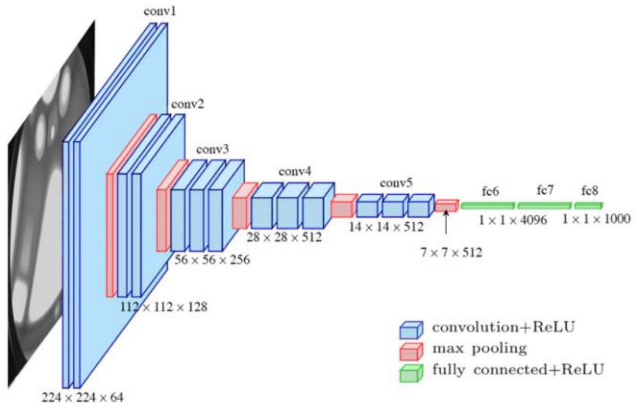

Fig. 2 VGG-16 Architecture. Adapted from [22]

3.3.2 ResNet-50

In second approach, we leveraged the feature extraction capabilities of ResNet-50. We load a pre-trained version of the model, but crucially, we exclude the final classification layer. This essentially transforms ResNet-50 into a powerful feature extractor. We flattened the output from the last convolutional layer of ResNet-50. This process creates feature vectors for each image, where each vector encapsulates the characteristics learned by the model.

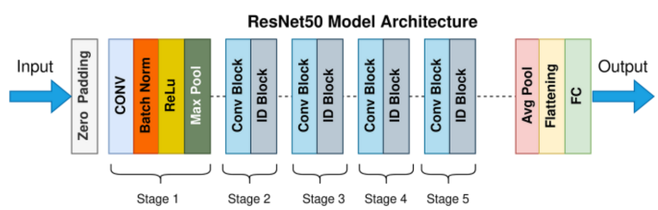

Fig. 3 ResNet-50 Architecture. Adapted from [23]

3.3.3 VGG-16 combined with AlexNet

The third approach involves harnessing pre-trained deep learning architecture AlexNet and VGG. The feature vectors extracted from AlexNet and VGG will have different shapes due to their unique network architectures. To enable their combined utilization in the classification stage, we transformed them into a uniform representation. Then we flattened each feature vector into a single feature vector for each image. This creates a consistent data format suitable for the baseline classifier used for classification of images.

4. RESULTS & DISCUSSION

4.1 Experimental Setup

This research implemented a ship detection model built on a hybrid architecture. Python served as the programming language for this project, with TensorFlow and Keras libraries providing the core functionalities.

4.1.1 Hardware and Software Environment

We used a personal computer with the following specifications to test and implement the suggested methodology: an Intel® CoreTM i5-8250U processor by Intel Corporation, 8 GB of RAM, a 500 GB SSD for storage, and Microsoft Windows 10 (Microsoft, Redmond, WA, USA) as the operating system.

4.1.2 Training Environment

To expedite the training process for the neural networks, we employed Google Colab, a freely available platform.

4.1.3 Machine Learning Hyperparameter Tuning

For the Support Vector Machine (SVM) classifier, we opted for a linear kernel. To ensure consistent results across training runs, a random state of 42 was set for tree-based classifiers like Random Forest. The loss function employed for the SGDClassifier was 'modified_huber'. These hyperparameter settings were meticulously chosen based on both prior research and experimentation. This fine-tuning aimed to optimize the performance of the machine learning algorithms within the hybrid architecture.

4.2 Result and Analysis

The results reveal some key findings. The combination of VGG-16 and XGBoost achieved the highest overall accuracy (99.375%) across all architectures as shown in Table 2. Moreover, SVM gives maximum average accuracy for all approaches which suggests that SVM excelled at classifying images when leveraging features extracted from the pre-trained VGG-16 model. Interestingly, VGG-16 pertained on ImageNet outperformed ResNet-50 & VGG-16 and AlexNet with all the classifiers tested. Consistently higher accuracy scores with VGG-16 indicate it might be a more suitable choice for this specific image classification task.

Furthermore, SVM and XGBoost consistently achieved higher accuracy compared to AdaBoost, Random Forest, and SGB. This implies that SVM and XGBoost were generally more effective classifiers for the image classification problem at hand.

An interesting finding emerged when comparing the pre-trained and non-pretrained versions of VGG-16. Utilizing a VGG-16 model pre-trained on ImageNet resulted in similar or slightly improved performance compared to the non-pretrained version for most classifiers. This suggests that the pre-trained weights provided a strong foundation for feature extraction, even if the improvement wasn't substantial across all classifiers.

In conclusion, these results underscore the critical role of both feature extraction techniques and classifier selection in achieving optimal image classification accuracy. VGG-16 appears to be a well-suited model for this task, and SVM and XGBoost emerge as effective classifiers when combined with VGG-16 features. Further exploration could involve fine-tuning hyperparameters for each classifier and potentially investigating other deep learning models for feature extraction to determine if even higher accuracy is achievable.

Table 2. Performance analysis of the proposed models

Approach	Algorithm	Accuracy
VGG-16 pretrained on Imagenet	SVM	99.25
	XGBoost	99.375
	Adaboost	99.125
	Random Forest	98.125
	SGB	99.0
ResNet-50	SVM	98.75
	XGBoost	96.375
	Adaboost	95.5
	Random Forest	93.875
	SGB	97.75
VGG-16 and AlexNet	SVM	99.25
	XGBoost	98.875
	Adaboost	98.75
	Random Forest	97.62
	SGB	98.25

5. CONCLUSION & FUTURE WORK

The research investigated the efficacy of leveraging a hybrid deep learning and machine learning approach for ship detection in satellite imagery. The proposed methodology achieved promising results, with the combination of VGG-16, a pre-trained deep learning model, and XGBoost, a machine learning classifier, achieving the highest overall accuracy of 99.375%.

Our findings highlight several key takeaways. Firstly, the pre-trained VGG-16 model demonstrably outperformed the other deep learning models explored for feature extraction in this specific image classification task. Secondly, SVM and XGBoost emerged as the most effective classifiers when combined with the features extracted from VGG-16. These results emphasize the significance of both feature extraction techniques and classifier selection in achieving optimal ship detection accuracy.

Future research directions could involve fine-tuning hyperparameters for the various classifiers employed. Additionally, exploring alternative deep learning models for feature extraction alongside potentially incorporating techniques to address false positives caused by natural phenomena remain promising avenues for further investigation. Ultimately, the goal is to achieve even greater accuracy and robustness in ship detection from satellite imagery, thus enhancing maritime security and safety measures.

6. REFERENCES

[1] F. Biljecki, J. Stoter, H. Ledoux, S. Zlatanova, and A. Çöltekin, “Applications of 3D City Models: State of the Art Review,” *ISPRS International Journal of Geo-Information*, vol. 4, no. 4, pp. 2842–2889, Dec. 2015.

[2] C. Kirpes, G. Hu, and D. Sly, “The 3D Product Model Research Evolution and Future Trends: A Systematic Literature Review,” *Applied System Innovation*, vol. 5, no. 2, p. 29, Feb. 2022.

[1] C. Bueger, “What is maritime security?” *Marine Policy*, vol. 53, no. 1, pp. 159–164, 2015

[2] K. Ward, R. Tough and S. Watts, “Sea clutter: Scattering, the K distribution and radar performance,” in *IET: Radar, Sonar and Navigation*. Stevenage: Michael Faraday House, 2013.

[3] K. R. Aldosari. “The maritime commons: digital repository of the World Maritime University”. *World Maritime University*, Malmo, 2019.

[4] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, USA, 2014, pp. 580-587, doi: 10.1109/CVPR.2014.81.

[5] F. Lei, W. Wang and W. Zhang, "Ship Extraction using Post CNN from High Resolution Optical Remotely Sensed Images," *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, Chengdu, China, 2019, pp. 2531-2535, doi: 10.1109/ITNEC.2019.8729145.

[6] H. Xiao, X. Wang, and P. Zhao. “Satellite Image Recognition for Smart Ships Using a Convolutional Neural”. *International Journal of Decision Science*, vol. 10, no. 2, June. 2019, doi: 10.22146/ijccs.37461.

[7] P. Sharma, U. K. Gupta, M. Oza and S. Sharma, "Reception - A Deep Learning Based Hybrid Residual Network," *2019 6th Swiss Conference on Data Science (SDS)*, Bern, Switzerland, 2019, pp. 125-129, doi: 10.1109/SDS.2019.00012.

[8] R., and R. Joshi. “Smart Ship Detection using Transfer Learning with ResNet”. *International Research Journal of Engineering and Technology (IRJET)*, vol. 6, no. 9, Sep. 2019.

[9] R. Venkatesh and A. Metha, “Segmenting ships in satellite imagery with squeeze and excitation U-Net,” *arXiv (Cornell University)*, Oct. 2019. http://arxiv.org/pdf/1910.12206.pdf

[10] F. Uçar and D. Korkmaz, “A ship detector design based on deep convolutional neural networks for satellite images,” *Sakarya University Journal of Science*, vol. 24, no. 1, pp. 197–204, Feb. 2020, doi: 10.16984/saufenbilder.587731.

[11] N. J. L. Marfu’ah and A. Kurniawardhani, “Comparison of CNN and SVM for ship detection in satellite imagery,” *AUTOMATA*, vol. 1, no. 1, Jan. 2020, [Online]. Available: https://journal.uii.ac.id/AUTOMATA/article/view/13973

[12] M. M. Stofa, M. A. Zulkifley, and S. Z. M. Zaki, “A deep learning approach to ship detection using satellite imagery,” *IOP Conference Series. Earth and Environmental Science*, vol. 540, no. 1, p. 012049, Jul. 2020, doi: 10.1088/1755-1315/540/1/012049.

[13] J. Yang et al., "Domain Adaptation for Degraded Remote Scene Classification," *2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV)*, Shenzhen, China, 2020, pp. 111-117, doi: 10.1109/ICARCV50220.2020.9305483.

[14] P. Pushkar, L. Aggarwal, M. Saad, A. Maheshwari, H. Awasthi, and P. Nagrath, “Object identification in satellite imagery and enhancement using generative adversarial networks,” *Advances in intelligent systems and computing*, 2020, pp. 643–657. doi: 10.1007/978-981-15-5148-2_57.

[15] A. Kumar, A. Sugandhi, and Y. Prasad, “Detection and Identification of Threat Potential of Ships using Satellite Images and AIS Data,” *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, Jan. 2022, doi: 10.5220/0010914600003124.

[16] K. Patel, C. Bhatt, and P. L. Mazzeo, “Deep Learning-Based Automatic Detection of Ships: An experimental study using satellite images,” *Journal of Imaging*, vol. 8, no. 7, p. 182, Jun. 2022, doi: 10.3390/jimaging8070182.

[17] S. Joshi, V. Mittal, N. Rawat and R. K. Ranjan, "Ship Detection from Despeckled Satellite Images using Deep Learning," *2022 IEEE International Conference on Data Science and Information System (ICDSIS)*, Hassan, India, 2022, pp. 1-7, doi: 10.1109/ICDSIS55133.2022.9915923.

[18] C. Ieracitano, N. Mammone, and F. C. Morabito, “Towards an Explainable Artificial Intelligence Approach for Ships Detection from Satellite Imagery,” *Studies in computational intelligence*, 2023, pp. 289–297. doi: 10.1007/978-3-031-25755-1_19.

[19] M. F. A. M. Jamal, S. S. Almeer, and S. R. Pulari, “Ship Detection from Satellite Imagery Using Deep Learning Techniques to Control Deep Sea Oil Spills,” *Lecture notes in networks and systems*, 2022, pp. 365–375. doi: 10.1007/978-981-19-2821-5_32.

[20] K. Roy, S. S. Chaudhuri, S. Pramanik, and S. Banerjee, “Deep neural network based detection and segmentation of ships for maritime surveillance,” *Computer Systems Science and Engineering*, vol. 44, no. 1, pp. 647–662, Jan. 2023, doi: 10.32604/csse.2023.024997.

[21] Ships in Satellite Imagery: [Online]. Available: https://www.kaggle.com/rhammell/ships-in-satellite-imagery.

[22] https://medium.com/@siddheshb008/vgg-net-architecture-explained-71179310050f

[23] https://towardsdatascience.com/the-annotated-resnet-50-a6c536034758