Journal of Neonatal Surgery
ISSN(Online): 2226-0439
Vol. 14, Issue 12s (2025)
https://www.jneonatalsurg.com

OPEN ACCESS

Identify Vessels in UAV Data by Dynamic Multi-Label Image Classification

Vithya Ganesan[1], Kirubakaran N[*2], Deepak R[3], Muruganandam S[4], Dr.T.Vengatesh[5], Dr.Viswanathan Ramasamy Reddy[6]

[1]Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Veddeswaram, Guntur-522302, Andhra Pradesh, India

Email ID: vithyamtech@gmail.com

[2*]Associate Professor, Department of Computer Science and Engineering, Chennai Institute of Technology, Chennai.

Email ID: drkiru70@gmail.com

[3] Assistant Professor Department of Computer Science and Engineering ,Nitte Meenakshi Institute of Technology,(NMIT) , Bangalore , Karnataka -560064, India

Email ID:aristonetram@gmail.com

[4]Department of Computer Science and Business Systems, Panimalar Engineering College, Varadharajapuram, Poonamallee, Chennai , Tamilnadu, India.

Email ID: murugan1004@gmail.com

[5]Assistant Professor, Department of Computer Science, Government Arts and Science College, Veerapandi, Theni, Tamilnadu, India.

Email ID: venkibiotinix@gmail.com

[6]Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Veddeswaram, Guntur-522302, Andhra Pradesh , India.

Email ID: rvnathan06@gmail.com

Cite this paper as: Vithya Ganesan, Kirubakaran N, Deepak R, Muruganandam S, Dr.T.Vengatesh, Dr.Viswanathan Ramasamy Reddy, (2025) Identify Vessels in UAV Data by Dynamic Multi-Label Image Classification. *Journal of Neonatal Surgery*, 14 (12s), 597-608.

ABSTRACT

To improve the monitoring and operational control of maritime traffic and nearby marine environments, this study investigates the use of Unmanned Aerial Vehicles (UAVs) and drones for improved surveillance, mapping, and remote sensing. UAVs offer real-time, high-resolution data that can help with more precise tracking of traffic patterns and vessel movements across large sea areas. The study suggests employing Convolutional Neural Networks (CNNs), which are very good at handling and decipheringcomplicatedpicturedata,toprocessthedatagatheredbytheseUAVs.This method overcomes the shortcomings of current tracking and monitoring systems by utilizing CNNs to improve object detection skills, particularly in accurately detecting andcategorizingshipfeatures.Identifyingshipsfromaerialphotosisacrucial problem in UAV-basedsurveillance,whichfrequentlycallsonanalyzingminutecharacteristics and patterns. Because it facilitates more effective monitoring of marine activity and helps manage the ecological impact of sea traffic, the project focuses on large-scale object detection in aerial photography, a field with important commercial and environmental ramifications.

The study presents sampling equivariant algorithms and other optimization-based techniques designed to enhance detection for the particular requirements of aerial images. The sampling equivariant technique is especially helpful for detecting small ship items, which can appear on different scale sand orientations and frequently have a hazy or deteriorating appearance. This method improves ship detection accuracy in difficult situations by reliably identifying these objects despite scale and viewpoint alterations. Furthermore, optimization-based tactics improve feature extraction accuracy by using grayscale sampling approaches to distinguish ships from their backgrounds in high-noise or low-contrast environments. Together, these methods enable the efficient tracking, classification, and prediction of ship motions and directions by extracting important ship properties from UAV photos. By monitoring traffic in environmentally sensitive locations, this predictive capability helps manage maritime safety, expedite emergency response times, and promote environmental conservation.

The results of this study demonstrate the value of UAV Surveillance when paired with cutting-edge CNN algorithms. The suggested methods greatly improve operational safety, environmental monitoring, and the commercial management of marine traffic by enabling more reliable detection of small, far-off objects in complicated, variable situations. This work provides a scalable method to meet the increasing demand for precise and efficient marine surveillance by improving airborne object detection capabilities.

1. INTRODUCTION

The excellent quality, readiness, adaptability, affordability, and ease of deployment of UAV photographs have made the mind in both military and civilian contexts. With the ability to acquire various data for tasks like intelligence collection and border surveillance, UAVs is seen as vital tools in military applications. UAV applications in ship traffic management, such as security, search and rescue, weather and disaster monitoring, product delivery, communication, infrastructure inspection, and traffic surveillance, are becoming increasingly popular, which emphasizes their importance. Environmental data is often collected and processed using remote sensing techniques that give each pixel a different intensity. These techniques must use scalar and vector- valued random fields to detect and track maritime traffic.

However, the traditional image classification techniques commonly employed for UAV imagery analysis provide a significant obstacle. Since these methods are primarily single-label, more than they are needed for deciphering intricate UAV photos with significant variety and in tricacy over maritime environments. In contrast to land-based images, UAV-captured marine photo graphs frequently shows ever a separate vessels in a single frame, each with distinctive characteristics, including size, kind, and functional categorization. Because marine photos are multifaceted, a single-label classification is inadequate because it needs to capture the variety of pertinent information included in the image.

Furthermore, because maritime activities are dynamic, traditional static classification models cannot keep up with changing circumstances, particularly as new boats and activities are added over time. These drawbacks highlight a situation where static and single-label classification methods are no longer appropriate for contemporary UAV- based maritime surveillance.

Dynamic multi-label picture classification techniques present a viable way to close this gap. These methods improve the system's capacity to recognize different vessel kinds and operational states in real-time by allowing the model to assign multiple, contextually relevant labels to each image. The model can continually adjust by processing photos using sophisticated machine learning techniques, supporting a variety of nautical applications and improving situational awareness. This method guarantees a reliable and flexible monitoring solution for marine environments by enabling the system to collect a variety of image properties, including vessel orientation and functional context.

Challenges and Research Gaps

The research gap in this dissertation is in overcoming constraints posed by existing approaches to processing and analyzing UAV photos for maritime surveillance, particularly detection, and tracking of small and far ship targets in adverse weather. Although object detection and tracking traditionally utilize some methods, these have yet to be proved inadequate for maritime purposes due to extreme conditions, such as scale, orientation, and the environmental appearance of most ships. This has made detecting and recognizing most ships, which would be otherwise clear in a background picture, very difficult because ships undergo different changes with an often poly distract or canvas containing many unrelated objects.

Moreover, few studies are specifically devoted to advanced algorithms targeting screening improvement in UAV imagery towards object detection, where objects' sizes and quality differ due to high altitudinal views, which is the nature of aerial imaging. However, the present framework does not fully use better approaches, such as algorithms designed to be sampling equivariant and optimization methods, which deal well with the low contrast and high disorder of images taken from the sea.

In addition, growing demand for advanced techniques that go beyond passive detection of targets and, instead, real-time predict their position and orientation, as a limitation of the current methods is that they need to be more capable of continuous tracking and forecasting with optimization. This study seeks to address those short comings by using convolutional neural networks (CNNs) and enhancing detection efficiency, flexibility, and robustness with suitable algorithms, which is crucial in improving maritime domain awareness and controlling vessel traffic in the sea.

2. LITERATURE SURVEY

The objective for ship detection and recognition using aerial images is to try to find an object in a UAV image and identify which one it is. An image can be represented with a matrix where the scalar values of rows and columns represent the image's height and breadth as used in remote sensing imaging acquisition [2].The bounding box for the object should be estimated at the time of localization [3].A common convention is that the bounding box describes the upper left corner pixel coordinate and lower right corner pixel coordinate [4] by detection and localization to determine what an object is from its localization belongs to a known object or the backdrop. [5].

Single-stagedetectorsdominateobjectdetectionalgorithmsinaerialimagedatasets.The accuracy with two-phase detectors has narrowed or even closed. Object detection in aerial pictures is by single-phase detectors, computer chips, and edge computing Equipment. Ming et al. introduce a "dynamic anchor learning" method for detecting objects with arbitrary orientations to improve accuracy in complex scenes. Their approach uses adaptive anchor boxes that adjust based on object size and orientation, enhancing detection in images not aligned to standard axes. This method is particularly useful for remote sensing and aerial images with common orientations. The study demonstrates significant improvements in object detection precision over traditional anchor-based methods in challenging datasets [6]. Rama krishnan et al. present a computer vision framework for detecting visually sharp objects using deep learning to enhance safety in environments with potential hazards. Their approach utilizes convolutional neural networks to accurately identify sharp objects, improving detection performance over traditional computer vision techniques. This framework has applications in industrial safety and automation, where recognizing and managing sharp objects is essential for reducing risks and enhancing operational safety. The study demonstrates the effectiveness of deep learning in visual object recognition for real-time applications [7].

Cao et al. introduce D2Det, which enhances object detection and instance segmentation by focusing on detection quality and mask precision. The model leverages a new dual-scale feature extraction and a cascaded head structure to improve the accuracy of bounding boxes and segmentation masks, particularly for complex or overlapping objects.

This approach achieves high-quality detection, making it highly effective in applicationsrequiringpreciseobjectlocalizationandsegmentation [8].Lin etal. propose the *focal loss* function to improve the training of dense object detectors on imbalanced datasets, where there are many easier (background) examples than challenging (foreground) ones. Focal loss reduces the influence of easy examples, focusing the model on hard-to-classify objects, which improves detection accuracy for small and less-prominent objects. This approach has been instrumental in enhancing the performance of object detectors, especially in tasks where class imbalance is significant [9]. Bochkovskiy et al. introduce *YOLOv4*, an advanced object detection model designed to achieve optimal speed and accuracy. The paper discusses various improvements over its predecessor, including enhanced backbone networks and new trainingtechniquesthatallowittorunefficientlyonvarioushardwarewhilemaintaining high detection performance. YOLOv4 excels in real-time applications, making it suitable for scenarios where both speed and precision are critical, such as in surveillance and autonomous systems [10].

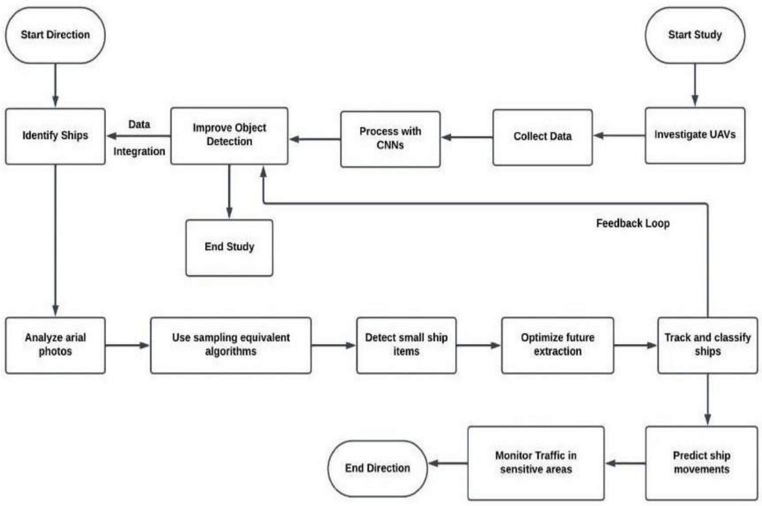
This survey by Liu et al. provides a comprehensive review of deep learning techniques applied to object detection and tracking in UAV-based systems challenges of aerial data, including varying scales, occlusion, and complex backgrounds, and discusses how deep learning models address these issues. The paper categorizes UAV applications in areas like surveillance, traffic monitoring, and environmental tracking and examines various architectures, from CNNs to RNNs.

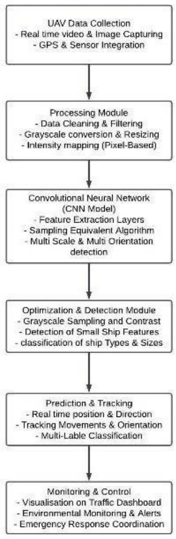
The authors also present future directions for enhancing UAV-based object tracking in complex environments using deep learning[11].Orientations. Their approach addresses challenges in satellite-based ship detection, such as scale variation and cluttered environments, enhancing detection performance in remote sensing applications [12]. Li et al. present a deep learning approach for detecting ships in aerial images, focusing on improving detection in diverse maritime environments. Their model leverages advanced neural networks to capture distinct ship features, handling variations in ship size, orientation, and back ground. This method shows improved accuracy in recognizing small and partially obscured ships, making it effective for real-world applications in monitoring and maritime safety [13]. Zhang et al. propose a multi-scale ship detection approach for remote sensing images, combining CNN-based feature extraction with saliency analysis. This method effectively identifies ships of various sizes and orientations by highlighting critical features and suppressing background noise. The saliency analysis improves ship detection in complex maritime environments, enhancing accuracy and reliability. Their approach is particularly beneficial for applications in surveillance and maritime traffic monitoring [14].Peng and Zhang explore advanced UAV-based remote sensing methods specifically designed for maritime surveillance. They discuss techniques to enhance UAV data processing for ship detection, tracking, and environmental monitoring. Their approach leverages deep learning and image processing methods to handle challenging conditions like varying sea states and weather. The study emphasizes the role of UAVs in improving maritime safety and operational efficiency by providing high-resolution, real-time surveillance capabilities [15]. This study proves how crucial and important deep convolutional neural networks become in transforming the image classification process; such systems get impressive results using Input data from ImageNet, setting new records and performance standards for various tasks in computer vision applications.[16].FasterR-CNN employs region proposal networks to help make object detection real-time, thus greatly speeding up and improving the accuracy there of. It brings a tremendous change in almost every task of object recognition in computer vision applications [17].Indispensableimprovementhasbeenachievedintheaccuracy and reliability level of ship detection using aerial images using convolutional neural networks, as this study demonstrates. This holds true especially for smaller or blocked ships in harsh environments[18].Optimization-based methods applied in UAV imagery greatly develop maritime monitoring through improved accuracy in ship detection in dynamic and complex environments[19]. The combined application of UAVs with CNNs permits accurate detection of objects in aquatic environments, alleviating scale differences and other factors that adversely affect environments [20]. UAV-enabled remote sensing allows collecting a real-time large-scale maritime traffic data set and provides a wide scope of views into the movement of vessels, supporting efficient maritime operations[21].Deep learning technologies to ship detection prevent detection in complex marine settings from being disturbed by variability[22].UAV-based maritime surveillance becomes more efficient in detection accuracy and coordination among UAVs through cooperative reinforcement learning with hybrid rewards[23]. A deep learning framework for multi-scale object detection alone in aerial images provides a performance increase in these applications when compared with others for maritime monitoring[24].Noise, occlusion, and environmental variability were improved by adaptive sampling methods and

CNNs to detect maritime

Objects better in difficult marine conditions [25]. Equivariant algorithms enable reliable detection of small objects in UAV imagery, thereby increasing the possibility of ship identification despite differences in scale, rotation and view point [26]. Advanced extraction features improve UAV surveillance, which will contribute to maritime safety as well through better tracking and monitoring of vessel movements [27]. Techniques incorporating grayscale sampling drastically enhance the detection of noisy ships in systems dependent on UAVs while ensuring effective working in tough conditions [28].Deep reinforcement learning models, thus, could predict the face of maritime using UAVs and enhance their traffic management and operational safety [29].UAV-basedobjectdetectioninenvironmentalconservation, detecting vessels and lent for sustainability in sensitive marine eco systems [30].

Convolution Neural Networks





(1a) (1b)

Fig. 1a. Architecture diagram for the UAV-based maritime surveillance system,

1b. work flow

CNN is very efficient in processing grid-like topology data by using the mathematical linear operation known as convolution [7].The first parameter in a typical convolution network is the input. The second parameter is the kernel. The output of a convolution network is a feature map.Aerial Image CNN architecture has three layers ,including the Input, hidden, and Output Layers.

Object Detection in Aerial Image using CNN Hidden layer includes

One dimensional / 2D (Two Dimensional Signals) /3D (Three Dimensional Signals) Convolution Layer Pooling Layers and Fully Connected Layers

Pooling is an important operation in the convolution network. The inclusion and use of multi-scale in Convolutional Neural Networks can greatly increase an algorithm's

Identification performance in target detection tasks and dramatic size change conditions [8].

2.3ImageTransformation

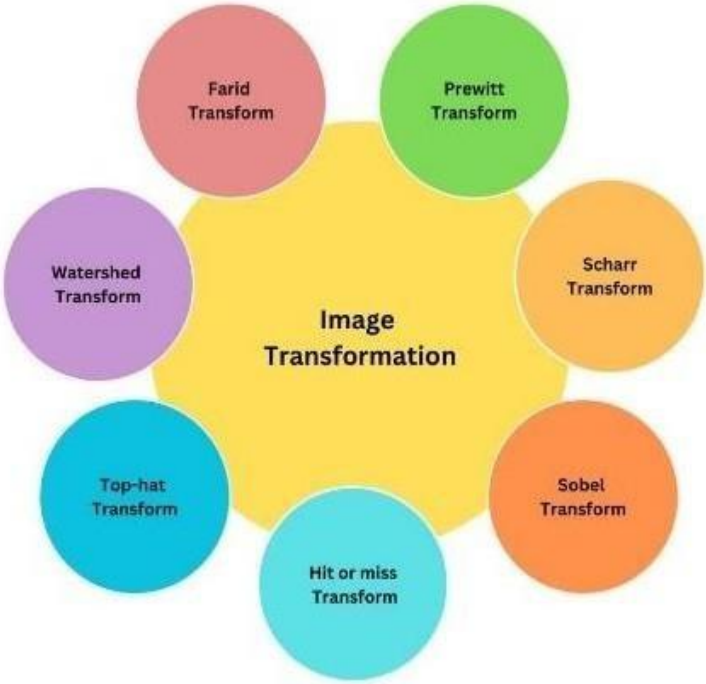


Fig.2. Image transformation model

Fig. 2: explains the type of image transformation model to identify the image classification.

The Sobel filter uses the [121]form, paying more attention to the central pixel. It works very well for edges when the noise in the aerial images is given in the form [9].

The second derivatives gain another way to edge detection. The sum of the second derivatives is known as the Laplacian and known as an isotropic filter, and it is invariant under rotation, as given below.

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial xy^2}$$

Hit-and-miss transform is a morphological operator for finding simple features in images. It is founded on erosion; this happens to be natural since S only erodes A in those pixels (locations) where S is contained within A or matches the set pixels in a tiny section of A. However, it has places where the background pixels in that region do not match and would not usually be categorized as matches. From the above study, manipulation is required to set both S's foreground and background pixels in A.

3. METHODOLOGY

Pre-processing means that after one algorithm is applied to your image, the output of that algorithm is fed to the input of some other algorithms. In contrast, an enhanced image should result from a particular algorithm. Of course, an image enhancement algorithm invokes pre-processing steps, while an enhanced image scans till be used an input for other algorithms. Convert color images to grayscale to reduce computation complexity in specific issues, to lose pointless data from pictures, and to lessen computational space. This is because, in numerous objects, coloring isn't important to perceive and decipher a picture. Gray scale is acceptable for perceiving specific items. Shading pictures contain more data than high-contrast pictures, add point less intricacy, and occupy more room in memory. Preprocess and scale all pictures to abound together size that the data set infer for it, ensuring they are indistinguishable widths and statures for learning, which means pivots, among other relative changes, are standard.

Feature Extraction

Feature extraction can simplify and present information clearly based on the feature extraction method and principal component analysis. Data Gain(IG)calculates the gain inentropyprovidedbythefeatureversusifitweremissing.Relationship-based Feature selection scans feature subsets among features. The evaluation cycle can find subsets of features that are solely superbly correlated with the class; however, it comes with a low connecting accumulation of features that increases the coupling between features and class and decreases the relationship. For example, forward choice in the reverse end, bidirectional inquisition, best first hunting, and hereditary hunt. The channels method uses different scoring functions and selects top-N features with the highest scores.

Details of Proposed System

Theaerialimageistreatedastwo-dimensionalgridtopologydata.CNN'sarchitecture, or related filters, can capture temporal and spatial dependencies in Aerial Images. One of the important concepts of CNN is parameter sharing. It allows us to decrease the number of unique classifier parameters extremely. Without increasing the training aerial image samples, this concept increases the network size significantly. The other two important concepts are sparse interactions and equivariant representations [4]. The same network may detect many instances of the same object, such as two or more Ships from Aerial Images.

$model.add(Conv2D($32,(3,3),$padding='same'$, input_shape = (80, 80, 3), $activation$= $'relu'$))

Inthe abovecode snippet,we created a newconvolution2Dlayerwiththe size of 3*

3. This slides a 3 * 3-pixel matrix over the aerial image, which convolves into fewer outputs.

Pooling Operation

This window extraction function extracts windows of the input feature map s to show the maximum value the channel can take. Maximum pooling differs from a convolution because it typically has two windows and strides two to down-sample the features by a factor of two [6]. Figure 3 shows an image retrieval model.

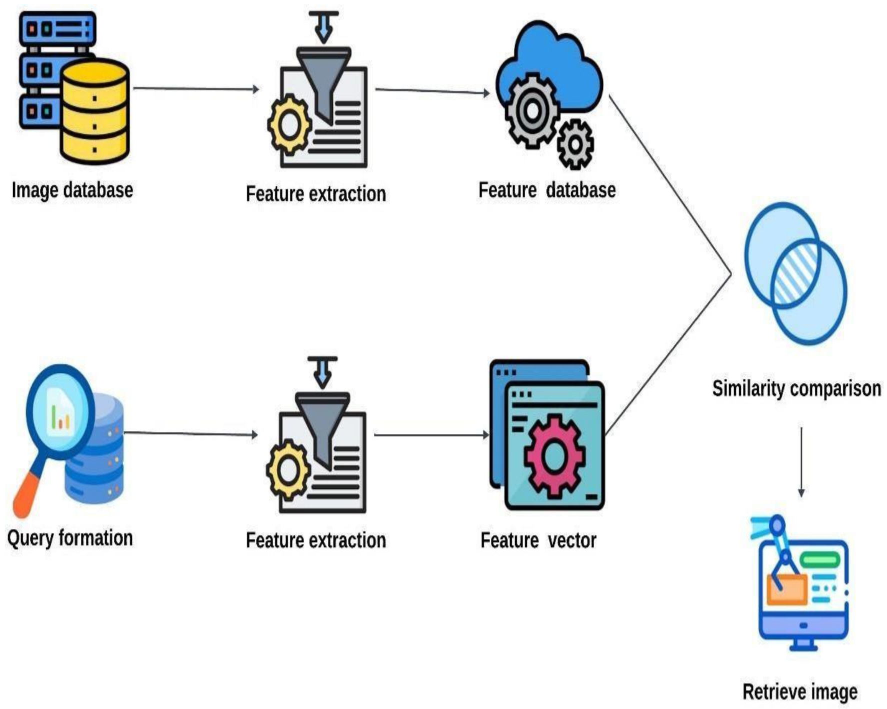


Fig.3.Imageretrievemodel

We can use strides in the previous convolutional layer and average pooling rather than max pooling. Here, each local input patch is transformed by the mean value of each channel taken over the patch instead of the maximum value. However, max pooling dominates these alternatives. In every case, pooling helps the representation be approximately invariant to small input translations. Translation invariance means that if a small amount translates the input, the values of pooled outputs are not to be changed [2].

Drop out Layer

This produces a convolution kernel convolved with layer input to output a tensor of outputs. Dropout Layer: "Drop out values obtained during training so you don't get caught hook, line, and sinker to a problem aren't quite ready for"; max pooling is suitable to this project and works well on spatial data. Dropout is an amazing regularization that is straight forward to incorporate and support seven more models and training algorithms since there is noise in estimating the statistics used for normalizing each variable, batch normalization occasionally reduces generalization errors, and dropout becomes unnecessary.

4. RESULT

All the packages are imported, and the nitreads the ships net.jsonfile.This data set has planet satellite imagery collected in theUSA's BayAreas.Itincorporates400080x80 RGB pictures marked with one or the other a"ship" or"no ship" order. Picture chips were acquired from Plannet Scope full-outline visual scene items, which are rectified to a 3-meter pixel size, as shown in Figure 4.

%%time

withopen('shipsnet.json')asdata_file:

dataset json.load(data_file) shipsnetpd.DataFrame(dataset) Wall time: 12.3 s print(shipsnet.shape) print(shipsnet.columns) print(shipsnet['data'][0:5]) (4000, 4)

Index(['data','labels','locations','scene_ids'],dtype='object')

0[82,89,91,87,89,87,86,86,86,86,84,8...

1[76,75,67,62,68,72,73,73,68,69,69,6...

2[125,127,129,130,126,125,129,133,132,...

3[102,99,113,106,96,102,105,105,103,10....

4[78,76,74,78,79,79,79,82,86,85,83,8...

Name: data,dtype:object

Fig.4 Reading the data from the JSON file

A compressed index is given as shipsnet.zip containing the full data set as.png picture chips. Every filename of the picture follows a special ordering: {label, scene id, longitude,latitude}.png.Mark:valued1or0,addressingthe"ship"classand"no-ship" class in turn. Scene id: The unique identifier of the Planet Scope visual scene from whichthepicturechipwascropped.ItusedwiththeplanetAPI,whichwouldhelpfind and download the whole scene: Scene id:Longitude and latitude are the longitude and scope directions from the position with values of this type separated by just one spotlight: In this dataset, the data are spread in the text file format through a json- organized transport document-ships net.json- It is also holding the information in stacked information, mark of information together with scene Ids as well as location- records:.Thedataarraycontainsthepixelvalueinformationforeachofthe80x80RGB images as a list of 19200 whole numbers. The first 6400 entries are the values of the red channel, the following 6400arevaluesofthegreen,and the last 6400arevaluesof the blue. This establishes an important column demand; all the red channel upsides of the central line of the image shoot up to the end by the first 80 rows in the exhibition. The next three lists, along with the index, carry estimates for the provided lists I in the names I, scene_ids, and names/areas. Each one aligns with the I-th image of the data list. Around 1000 images from ship class.

Pictures in this class are highly zoomed in on a single boat's body. Boats of different sizes, directions, and climatic assortment conditions are included in Figure 2(a).

No- ship class has about 3000images.33%ofthesearearbitrarytestsofdifferentlandcover features- water, vegetation, bare earth, constructions, and so forth that eliminate any

Part of a boat. The third below are "partial boats" that consist only of some part of a boatbutnotenoughtomeetthefullmeaningofthe"boat"class.Thelastthirdofthese images AI models have recently mislabeled are primarily due to massive pixels or prominent straight features in Figures 5a and 5b.

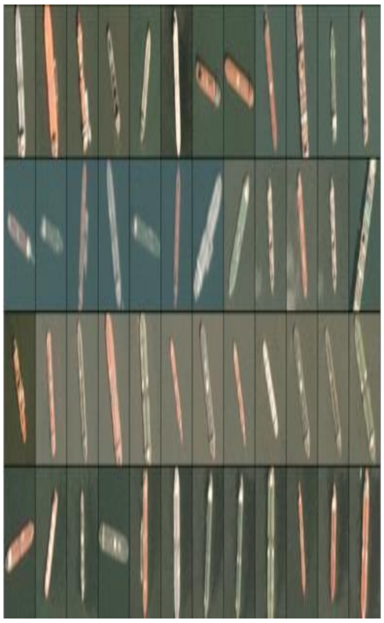


Fig. 5 a Pictures present in the ship class. **Fig. 5 b Pictures present in the no- ship class.**

Converted the RGB image in to the gray-scale image and using mathematical operations such a s min,max;applied mat plot lib to plot all the subplots (1,6)inFig6.

Fig.6:RGB to gray scale conversion.

These images are shuffled through the shuffle function. Thereafter, we have trained as wellastestedfordatainwhich80%willbeusedfortrainingpurposeand20%fortest purposes. After train as well as testing, this is split according to its labels. All these works are done by a random seed that generates random number with the help of some seed value. Fig 6 depicts the length for the prepared, shuffled rain as well as test data.

print("PREPARED ({})".format(len(data_prepared)))

print("SHUFFLED({})".format(len(data_shuffled))) print("TRAIN (())".format(len(data_train)))

print("TEST ({))".format(len(data_test)))

print("TRAIN TRUE ({})".format(len(data_train_true))) print("TRAINFALSE({})".format(len(data_train_false))) print("TEST TRUE ({})".format(len(data_test_true))) print("TEST FALSE ({})".format(len(data_test_false))) PREPARED (4000)

SHUFFLED(4000)

TRAIN(3200)

TEST (800)

TRAINTRUE(796)

TRAINFALSE (2404)

TEST TRUE(204)

TESTFALSE(596)

Fig.7: Length of prepared, shuffled, train and test data.

Red, blue, and green spectrums are applied to a single ship, as shown in Fig 7. Planet Satellite image collected in the Bay Area explained in Fig 8.

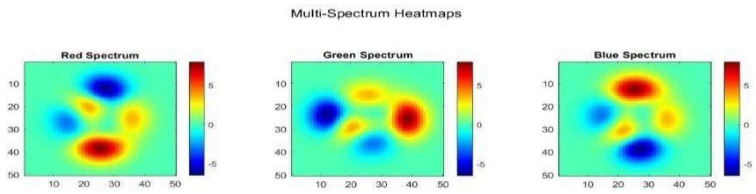


Fig.8:Red, blue, and green spectrum is applied for a single ship.

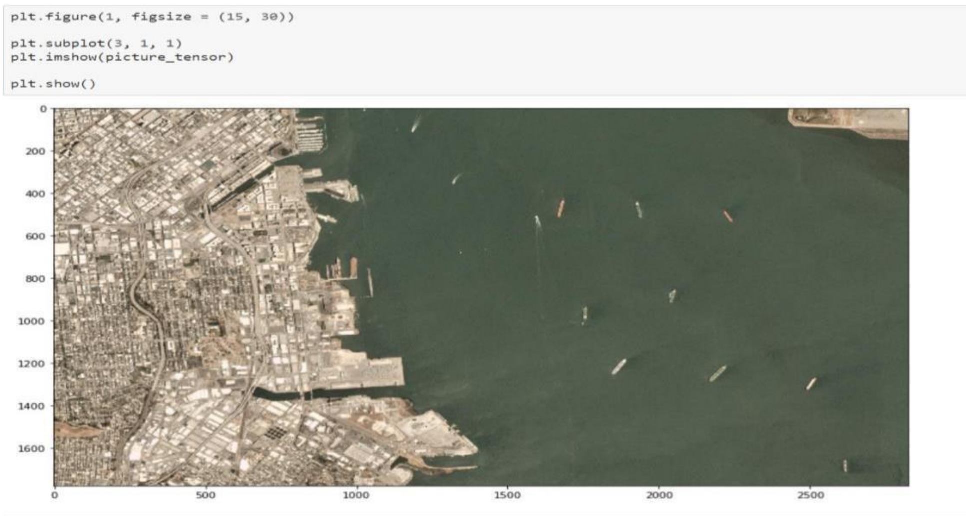


Fig.9:Planet satellite image,whichiscollectedinaBayarea,USA.

import sys

step10;coordinates= []

foryinrange(int((height-(80-step))/step)):forxinrange(int((width-(80-step))/step)): area cutting(x*step, y*step) result = model.predict(area)

if result[0] [1] > 0.90 and not_near(x*step, y*step, 88, coordinates):

coordinates.append([[x*step, y step], result])

print(result) plt.imshow(area[0]) plt.show()

Fig.10Fromtheplanetsatelliteimage,sub-images are analyzed

```
plt.figure(1, figsize = (15, 30))

plt.subplot(3,1,1)
plt.imshow(picture_tensor)

plt.show()
```

Fig.11Multiple ships are classified/identified from the given bay area image.

5. CONCLUSION

The sampling equivariant algorithms integrated with optimization-based strategies were highly effective in improving the network performance, especially when considering Convolutional Neural Networks (CNNs) trained with aerial images. Such methods make use of oriented and horizontal scaling to detect scattered ship image sin sea traffic. The proposed approach, which maintains sampling equivariance, ensures robustness and accuracy in ship detection under such conditions. Unlike many other traditional approaches, it does this improvement without increasing model parameters or computational costs; it is large-scaled efficiency.

The ability of the algorithm to accommodate various orientation sand scales guarantees very high precision within complicated marine environments. It surpasses other methods in being able to deal with multiple ship orientations and plays an important role in the management of traffic at sea. Such development is important to maintain and automatize the monitoring of artificial waters.

The potential integration of Quantum Computing into place might take this area a step further. The quantum algorithms would process faster with precise classifications, facilitating real-time analysis of sea traffic. This fusion would change the face of maritime monitoring systems in ensuring the safety, efficiency, and sustainability required in the future.

REFERENCES

[1] Seung-Taek Kim and Hyo Jong Lee. Lightweight stacked hourglass network for human pose estimation. Applied Sciences, 10(18):6497, 2020.

[2] [2]TongHe,ZhiZhang,HangZhang,ZhongyueZhang,JunyuanXie,andMuLi.Bag of tricks for image classification with convolutional neural networks. CoRR, abs/1812.01187, 2018.

[3] [3]Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28:91–99, 2015.

[4] [4]GowriShankarV,ArunKumarS,Dr.BalikaJChelliah,"MultipleObjectDetection Using Mask R-CNN in Deep Neural Network".

[5] [5]Gui-Song Xia, Xiang Bai, JianDing, Zhen Zhu, SergeBelongie, Jiebo Luo, Mihai

Datcu,MarcelloPelillo,andLiangpeiZhang.Dota:Alarge-scaledatasetforobject detectioninaerialimages.InTheIEEEConferenceonComputerVisionandPattern Recognition (CVPR), June 2018.

[6] [6]Q. Ming, Z. Zhou, L. Miao, H. Zhang, and L. Li, "Dynamic anchor learning for arbitrary-oriented object detection," in Association for the Advancement of Artificial Intelligence (AAAI), 2021, pp. 2355–2363.

[7] [7]Ramakrishnan,Nitesh,AnandhanarayananKamalakannan,BalikaJ.Chelliah,and GovindarajRajamanickam."ComputerVisionFrameworkforVisualSharpObject Detection using Deep Learning Model."International Journal of Engineering and Advanced Technology (IJEAT)ISSN: 2249-8958, Volume-8 Issue-4, April 2019

[8] [8]J. Cao, H. Cholakkal, R. M. Anwer, F. S. Khan, Y. Pang, and L. Shao. D2det: Towards high quality object detection and instance segmentation. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 11482–11491, 2020

[9] [9]T.-Y.Lin,P.Goyal,R.Girshick,K.He,andP.Dollar,"Focallossfor´denseobject detection," in International Conference on Computer Vision (ICCV), 2017, pp. 2980–2988.

[10] [10]A. Bochkovskiy, C. Wang, and H. M. Liao. Yolov4: Optimal speed and accuracy of object detection. CoRR, abs/2004.10934, 2020.

[11] [11]H.Liuetal.,"DeepLearningforUnmannedAerialVehicle-BasedObjectDetection and Tracking: A Survey,"IEEE Access, vol. 9, pp. 11924–11942, 2021. doi: 10.1109/ACCESS.2021.3051125.

[12] [12]Z. Shao et al., "Ship Detection in High-Resolution Satellite Images via Shape and ContextualFeatures,"IEEEGeoscienceandRemoteSensingLetters,vol.19,2022. doi: 10.1109/LGRS.2021.3098676.

[13] [13]F.Lietal.,"AerialImageAnalysisUsingDeepLearningforShipDetection,"IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1–14, 2022. doi: 10.1109/TGRS.2021.3086875.

[14] [14]Y.Zhangetal.,"Multi-ScaleShipDetectionforRemoteSensingImagesBasedon CNN Features and Saliency Analysis,"IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 2, pp. 876-889, 2022. doi: 10.1109/TITS.2021.3051032.

[15] [15]C. Peng and R. Zhang, "Advanced UAV-Based Remote Sensing Techniques for Maritime Surveillance,"IEEE Access, vol. 10, pp. 3440–3452, 2022. doi: 10.1109/ACCESS.2022.3145461.

[16] [16]A.Krizhevsky, I. Sutskever,and G. E.Hinton, "ImageNet classification with deep convolutional neural networks,"Commun. ACM, vol. 60, no. 6, pp. 84–90, Jun. 2017.

[17] [17]S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detectionwithregionproposalnetworks,"IEEETrans.PatternAnal.Mach.Intell., vol. 39, no. 6, pp. 1137–1149, Jun. 2017.

[18] [18]H.C.Goh,A.F.Lim,andD.Hsu,"Convolutionalneuralnetworkforshipdetection fromaerialimages,"in Proc.Int.Conf.Adv.Robot., HongKong,China,2020,pp. 34–39.

[19] [19]K. K. Mantripragada and P. A. Dhebar, "Optimization-based ship detection using UAV imagery for maritime monitoring," in Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS), Jul. 2021, pp. 4585–4590.

[20] [20]Y. Zou, Z. Xiong, and C. Shi, "Object detection in maritime environments using UAVs and CNNs,"IEEE Access, vol. 9, pp. 46755–46765, 2021.

[21] [21]S.S.KhanandS.Qayyum,"UAV-enabledremotesensingforlarge-scalemaritime traffic monitoring," in Proc. IEEE Int. Conf. Remote Sens. GIS (ICRSGIS), Mar. 2022, pp. 125–132.

[22] [22]W.Liuetal.,"Shipdetectionincomplexmarineenvironmentsusingdeeplearning techniques,"IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 13, pp. 5331–5342, 2020.

[23] [23]Z. Li, X. Zhang, and S. Zhang, "Hybrid reward-based cooperative reinforcement learningformaritimeUAVsurveillance,"IEEE Trans.Cybern.,vol.52, no.2,pp. 991–1003, Feb. 2022.

[24] [24]L. Pan, W. Zhang, and X. Wang, "Multi-scale object detection in aerial images using a deep learning framework," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2021, pp. 245–253.

[25] [25]C. Lu,S. Zhu,and S. Duan, "Enhancing maritime object detection with CNNs and adaptive sampling strategies,"IEEE Trans. Intell. Transp. Syst., vol. 23, no. 8, pp. 11422–11432, Aug. 2022.

[26] [26]Q.Xuetal.,"RobustdetectionofsmallobjectsinUAVimageryusingequivariant algorithms,"IEEE Trans. Image Process., vol. 30, pp. 3323–3335, Apr. 2021.

[27] [27]J.Shi,J.Zhang,andJ.Tan,"ImprovingUAVsurveillanceformaritimesafetyusing feature extraction

techniques," in Proc. IEEE Global Ocean Observ. Syst. Symp. (GOOS), Sep. 2020, pp. 89–95.

[28] [28]T.NguyenandT.Le,"Grayscalesamplingtechniquesforhigh-noiseshipdetection in UAV-based systems,"IEEE Access, vol. 8, pp. 62147–62158, 2020.

[29] [29]A.KumarandP.S.Rao,"PredictivemodelingforUAV-basedmaritimemonitoring using deep reinforcement learning,"IEEE Trans. Neural Netw. Learn. Syst., vol. 34, no. 1, pp. 17–26, Jan. 2023.

[30] [30]R. Gupta and V. Jain, "Environmental conservation in marine areas using UAV- enabled object detection,"IEEE Trans. Geosci. Remote Sens., vol. 60, pp. 1–13, 2022.