Department of Electrical & Electronic Engineering

Brac University

Spring-2022

Course Code: EEE385

Course Title: Machine Learning Laboratory Project

Project Title: Boat Detection by YOLO V5 Object Detection Architecture

Prepared by:

Name: Md. Jobaar Hossain ID: 19321018

Other Group members:

Sl.	***ID***	***Name***
1.	Shams Fardous Arnab	19321030
2.	Muntasir Ahad	19321012

Date of Submission: 10 th May, 2023

Introduction

Bangladesh is a riverine country, where the river network is immensely important for transportation and plays a vital role in national life as well. Water transports are comparatively inexpensive which results in it being eminent among general people. Since it is affordable, water vehicles get overcrowded and exceed capacity limits during holidays. According to a survey done by BIWTA (Bangladesh Inland Water Transport Authority), 238 motor launch accidents were recorded with almost a loss of 2,309 1ives, 374 people were injured and 208 were missing [1]. The effect of a marine accident is far worse, as the rescue mission by our safety guards are mostly manual. Most of the time, Manual resources lack work efficiency which causes delays in rescuing. Moreover, currently available measures are not efficient enough to figure out an appropriate number of victims who can be saved with immediate aid.

In Bangladesh, other sectors have better examples of utilizing modern technology which is also advancing day by day. But when it comes to the SAR (search and rescue) works, the involvement of modern technologies is inadequate yet. SAR is an essential part of reducing dangers to people and the environment in catastrophes as well as harsh situations.

Currently, the advances in AI(Artificial Intelligent) and machine learning have opened a path to the object detection models that can analyze satellite and UAV images and identify specific objects. The goal of this project is to incorporate boat detection model training in YOLO v5 to improve the effectiveness of rescue operations and reinforce the search and rescue field. This paper offers an adequate discussion about the current situation of oat detection models in this field and presents our methodology for training a model including the dataset used and the process of training. Finally, we will showcase the performance of our trained model in specific criterias such as accuracy, precision, recall and f1-score comparisons of different models.

Background Study

Boat detection model based on convolutional neural network (CNNs) using VGG19 that is trained using various boat pictures. The goal was to create a model to detect the type of boat passing through the canal using images collected through surveillance video cameras. By collecting various boat images via video camera a dataset of 3152 images was created andCNN model was used to classify boat types based on video data achieving a practical result as F1-score of 0.70.[2]

Ship detection and recognition using satellite remote sensing technology for potentially military and civilian fields, focusing on detection of small targets in remote sensing images.[4]

It can help civilians find lost ships at sea, monitor marine fisheries, and detect illicit ships like pirate ships or ferry boats. When utilized in a military setting, it may be used to detect enemy ships, direct long-range missiles for precision attacks, and monitor the waters during warfare for unlawful infiltration by warships. However, because of the long-distance shooting involved, the majority of objects detected in remote sensing images are small targets.[3]

Currently, there are two types of approaches for recognizing ships in remote sensing images: conventional methods and deep learning-based methods. Deep learning approaches are now the most commonly employed for spotting ships in optical remote sensing photos. Researchers have made significant progress in two-stage and single-stage target identification systems. Gu Jiaojiao and colleagues, for example, employed Faster RCNN to solve the problem of repeated object recognition in remote sensing pictures by changing the number and size of anchor frames. Furthermore, Ma Junjie and colleagues used the YOLO network to establish a nonlinear observation equation of the remote sensing ship target, which was converted into a linear equation to determine the specific position coordinates of the ship and improve its accuracy.[5]

Our project is different from the mentioned versions by the accuracy increase as one of the challenges in boat detection is detecting small boats in the distance. YOLOv5 is better at detecting small objects than previous YOLO versions because of the use of anchor boxes. This model is designed to be easier to train and customize than previous versions, making it more accessible.

Methodology

Detail explanation of the models we are using along with the parameters used Steps of yolov5:

- Data collection and annotation: Collect the images that you want to use for object detection and label them with bounding boxes around the objects of interest. You can use tools like LabelIng or CVAT to annotate the images.

- Data preprocessing: Resize the images to a uniform size and convert them to the format that Yolov5 can understand. Yolo v5 requires the images to be in the format of JPG or PNG and the annotations to be in YOLO format (txt files).

- Splitting the dataset: Split the dataset into training, validation, and testing sets. The training set is used to train the Yolov5 model, the validation set is used to tune the hyperparameters and to evaluate the model during training, and the testing set is used to evaluate the final performance of the model.

- Configuration and training of Yolov5: Configure the Yolov5 model architecture by modifying the configuration file. You need to adjust the number of classes in the dataset, the anchor boxes based on the object sizes, and other hyperparameters such as learning rate, batch size, and epoch.

- Training the model: Train the Yolov5 model using the custom dataset. The training process involves feeding the images into the network, calculating the loss function, and adjusting the weights of the network using backpropagation. The training can be performed on a local machine or on cloud-based services like AWS or Google Cloud.

- Evaluation and optimization: Evaluate the model's performance on the validation set and optimize the hyperparameters based on the results. The common evaluation metrics for object detection include precision, recall, and F1-score.

- Testing the model: Test the final trained Yolov5 model on the testing set to evaluate its performance on unseen data. You can visualize the predicted bounding boxes and check whether the model is accurately detecting the objects in the images.

The boat detection dataset in YOLOv5 code is likely a collection of images or videos that have been annotated with labels indicating the location of boats within them. The YOLOv5 code uses a deep neural network to detect and classify boats in these images or videos.

Results

The performance of the boat detection model can be evaluated using a number of performance measures such as accuracy, precision, recall, and F1-score.

- Accuracy: measures the overall correctness of the model's predictions, and is calculated as the number of correctly predicted boat locations divided by the total number of predictions made.

- Precision: measures the proportion of true positives (correctly predicted boat locations) out of all predicted boat locations. It is calculated as the number of true positives divided by the sum of true positives and false positives (incorrectly predicted boat locations).

- Recall: measures the proportion of true positives out of all actual boat locations in the dataset. It is calculated as the number of true positives divided by the sum of true positives and false negatives (boat locations that were missed by the model).

- F1-score: is a measure that combines both precision and recall into a single value, and is useful for evaluating models when both measures are important. It is calculated as the harmonic mean of precision and recall, and ranges from 0 to 1, with higher values indicating better performance.

In general, a good boat detection model should have high values for accuracy, precision, recall, and F1-score. However, the optimal values for each measure may vary depending on the specific use case and the relative importance of false positives (detecting boats that are not there) versus false negatives (missing boats that are there).

In this model, the dataset consists of 965 images based on 3 classifications which are bigboat ,boat and small boat. the dataset was collected from kaggle and then annotated with roboflow inorder to train this in yolo v5. Two versions of datasets were created with the same images but with different prepositions and arguments such as flip bounding boxes ,rotations hue and saturations which latter helped to create different variations in the images we selected .the first version has a lower mAP 8.7% ,pression 12.5% and recoil of 21.7% .it can be a reason of Low quality annotations, Insufficient training data,Ineffective data augmentation .as a result we had to rebuild the dataset using different augmentation and increased the images which latter increased the mAP .precision,F1-score and recall to 70.1% ,74.8% and 66.6%.

Model	**Accuracy**	**Recall**	**F1-score**	**Precision**
YOLO V5	0.91	0.66	0.748	0.71
CNN VGG19	0.725	0.71	0.70	0.70

Figure: training data 1

Figure: training data 2

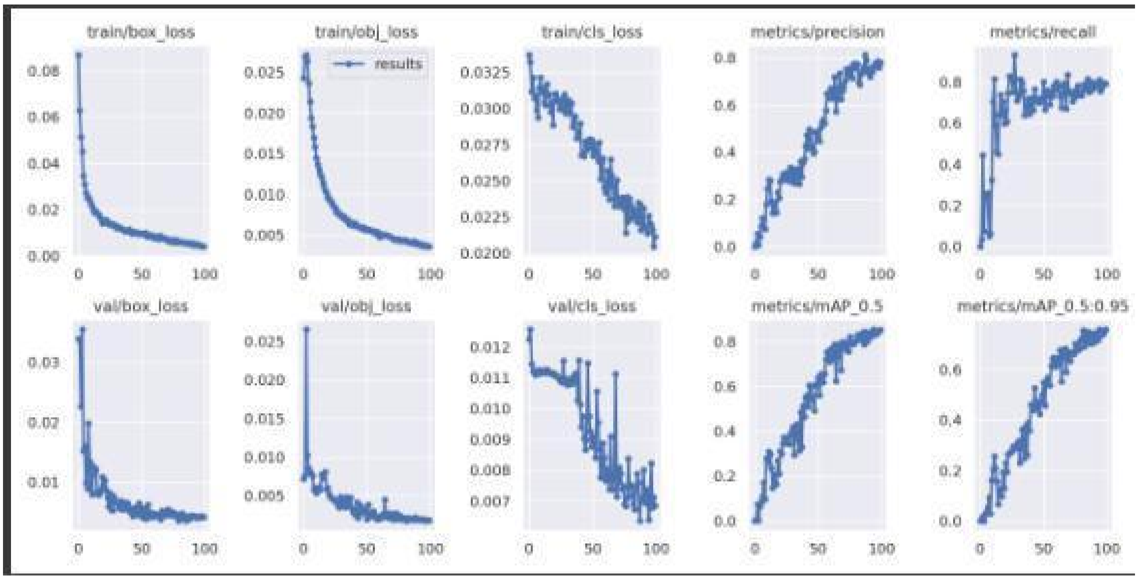

Figure: result curves

Conclusion

We proposed a YOLO model to detect boat type using boat pictures obtained from a surveillance camera attached to a UAV. The results of the proposed model show an average F1-score of 0.748, indicating its potential for practical use as it showcases more accuracy than its CNN model.

Future work

While this model is already quite accurate, there is always room for improvement. Researchers could explore ways to improve the accuracy of the model, perhaps by incorporating more data or tweaking the architecture. it can detect boats in a single image, but it cannot track boats across multiple images. Adding object tracking capabilities would allow the system to follow boats as they move through a scene, which could be useful for applications like maritime surveillance. YOLOv5 is already quite fast, but real-time detection of boats could be valuable in many applications, such as maritime search and rescue. Researchers could explore ways to speed up the model or develop more efficient hardware to run it on. Object detection is widely used in video surveillance systems to identify potential threats or suspicious behavior. YOLOv5 could be used to detect people or objects in real-time, enabling faster response times to potential security incidents.

Code link-
https://colab.research.google.com/drive/1vve2vLbdvxLdL7snNjQSUx7K19QaPBHV?usp=sharing&fbclid=IwAR1ojjyWTSz_SmixNVTVblfpLrHVWlMQwBJ4by-0QOjxRJAJvs9girkMDnk#scrollTo=uDxebz13RdRA

References

1. B. Mishra, D. Garg, P. Narang, and V. Mishra, “Drone-surveillance for search and rescue in natural disasters c.” Computer Communications, vol. 156, pp. 1-10, 2020, doi: 10.1016/j.comcom.2020.03.012.

2. Akiyama, T., Kobayashi, Y., Kishigami, J., & Muto, K. (2018, October). CNN-based boat detection model for alert system using surveillance video camera. In *2018 IEEE 7th Global Conference on Consumer Electronics (GCCE)* (pp. 669-670). IEEE.

3. Guo Liqiao, Lin Jianyuan, Wang Wenying. The application of refined urban management based on high-scoring remote sensing data[J]. Urban Development Research, 2012, 19(11): 57- 63+69

4. Sun, X., Zhang, Y., Wang, H., & Du, Y. (2022). Research on ship detection of optical remote sensing images based on Yolo V5. Journal of Physics, 2215(1), 012027. https://doi.org/10.1088/1742-6596/2215/1/012027

5. Ma Junjie, Huang Daqing, Liu Yaohui, et al. Ship target positioning based on improved ship yolo [J]. Electronic Measurement Technology, 2020, 43(5): 112-115.