T.C.

MANİSA CELAL BAYAR ÜNİVERSİTESİ

MÜHENDİSLİK FAKÜLTESİ

BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Deep Learning Based Ship Detection from Multi-Modal Satellite Images

Tasarım Projesi / Lisans Bitirme Tezi

HAZIRLAYANLAR

Ali Osman BEKER 180315010

Mehmet Batuhan DUMAN 180315036

DANIŞMAN

DOÇ. DR. Övünç ÖZTÜRK

MANİSA 2023

T.C.

MANİSA CELAL BAYAR ÜNİVERSİTESİ

MÜHENDİSLİK FAKÜLTESİ

BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Tasarım Projesi / Lisans Bitirme Tezi

KABUL VE ONAY BELGESİ

Ali Osman BEKER ve Mehmet Batuhan DUMAN'ın "Deep Learning Based Ship Detection from Multi-Modal Satellite Images" isimli lisans projesi çalışması, aşağıda oluşturulan jüri tarafından değerlendirilmiş ve kabul edilmiştir.

Danışman : Doç. Dr. Övünç Öztürk

Üye :Ali Osman BEKER

Üye :Mehmet Batuhan DUMAN

Projenin Savunulduğu Tarih : 13.01.2023

Bilgisayar Mühendisliği Bölüm Başkanı

CONTENTS

Page

ABSTRACT

Satellite imagery provides unique insights into various markets, including agriculture, defense and intelligence, energy, and finance. New commercial imagery providers, such as Planet, are using constellations of small satellites to capture images of the entire Earth every day. This flood of new imagery is outgrowing the ability for organizations to manually look at each image that gets captured, and there is a need for machine learning and computer vision algorithms to help automate the analysis process. The aim of this project is to help address the difficult task of detecting the location of large ships in satellite images. Automating this process can be applied to many issues including monitoring port activity levels and supply chain analysis.

Keywords: CNN ,R-CNN,

INTRODUCTION

The Deep Learning Based Ship Detection from Multi-Modal Satellite Images project aims to develop a system for automatically detecting ships in satellite images using deep learning techniques. The system will be trained on a large dataset of multi-modal satellite images, which will include images from multiple sensors and wavelengths. By using deep learning algorithms, the system will be able to recognize ships in a wide range of conditions and settings, improving upon traditional methods of ship detection. The system will be evaluated on its ability to accurately and reliably detect ships in satellite images, and its potential applications will be explored.

PROBLEM DEFINITION

The problem that the Deep Learning Based Ship Detection from Multi-Modal Satellite Images project aims to address is the challenge of accurately and reliably detecting ships in satellite images. Traditional methods of ship detection are often limited in their accuracy and robustness, and can be prone to false positives and false negatives. This can be problematic for applications such as maritime surveillance and tracking of ship movements, where it is important to have reliable and accurate information about the presence and location of ships. By using deep learning techniques, the project aims to develop a system that can overcome these limitations and improve the accuracy and reliability of ship detection in satellite images.

LITERATURE ANALYSIS

Project-like articles were reviewed in the literature and the codes from different projects were rerun and analyzed. The values were changed in the code and different models were subjected to different stages and the team worked together to find the original optimum model. The codes were rearranged considering the models written and experienced in the literature. The original model was optimized depending on the performance. The image was not subjected to any image processing methods. This pre-process will be tested within the scope of researches in the continuation of the project.

METHODS AND TECHNOLOGIES TO BE USED

The Deep Learning Based Ship Detection from Multi-Modal Satellite Images project will utilize a range of methods and technologies to develop a system for automatically detecting ships in satellite images. The core technology that will be used is deep learning, which is a type of machine learning that involves training a large neural network on a large dataset in order to recognize patterns and make predictions. This will be used to train the system to recognize ships in satellite images. In order to train the deep learning model, a large dataset of multi-modal satellite images will be collected and labeled. This will include images from multiple sensors and wavelengths, which will provide the system with multiple sources of information about the ships in the images. The labeled dataset will be used to train the deep learning model, which will learn to recognize ships in the images. Once the model has been trained, it will be tested on a separate dataset of satellite images to evaluate its accuracy and reliability. The system will be evaluated on its ability to accurately and reliably detect ships in the images, and its performance will be compared to traditional methods of ship detection. Data augmentation is a method that can be used to artificially increase the size of the training dataset by applying random transformations to the images, such as rotation, scaling, and flipping. This can help the model to generalize better to new data and improve its performance. Also we are working on the new training algorithm Forward-Forward(FF) but we haven’t successful yet.

PROJECT TEAM AND TASK SHARING

Mehmet Batuhan Duman took part in the research of technologies thought to be used by Ali Osman Beker in the literature review of similar projects in the distribution of tasks in the project. it was seen that marine vehicles were divided into categories in similar projects. It has been determined that image processing techniques are frequently used in the literature and have a positive effect on prediction. Image processing techniques will be used in the continuation process of the project. In this direction, the use of CNN and R-CNN from similar projects and its effect on the forecast were observed jointly as a team.

STUDIES

First, we converted 80x80 RGB data in .json format with 4000 images into a dataframe using the pandas library. Then we looked at which columns our data has. Then we discarded the locaion and seen_ids columns that we needed. We are left with 2 data and label columns. Afterwards, we pre-process the data we have. 3000 images are images with ships, the remaining 1000 images are images without ships. Then we assigned the data and labels data to 2 variables using the Numpy library. Then we used the reshape method to bring the data variable we assigned to the appropriate format. In order to train our model we will create, we first separated the data we shaped into train and test data. While separating these sets, we reserved a data of 80% for test_size 20% train and validation set. Then, we allocated the remaining 80% train and validation data, 25% validation, and 55% train set. We built a convolutional neural network (CNN) in Keras, a popular open source deep learning library. The CNN is defined as a Sequential model and consists of five 2D convolutional layers, a smoothing layer, and three dense layers.

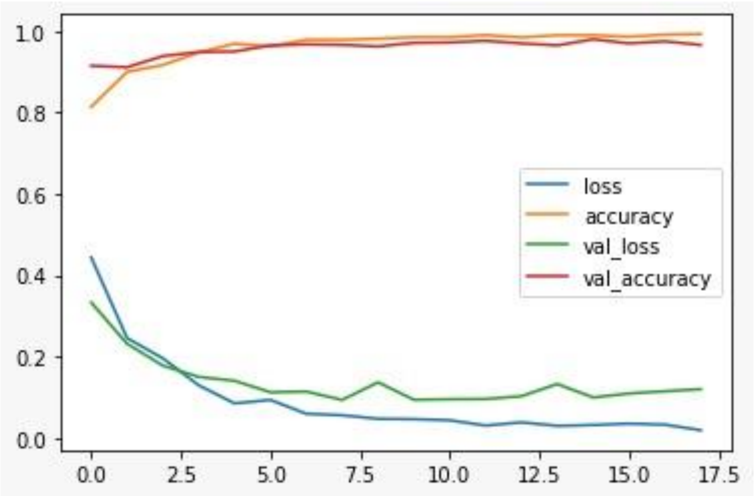
First Model

Our first model starts with a 2D convolutional layer with 32 filters, a kernel size of (3,3), same padding, and a ReLU activation function. It is followed by a max pooling layer with a pool size of (2,2), which reduces the dimensions of the output from the convolutional layer. The output is also passed through a dropout layer with a rate of 0.25. This is a regularization technique that helps prevent overfitting by randomly setting some of the output values to zero.

The model then includes several more convolutional, max pooling and dropout layers, all have 32 filters and a stride of 3. Each of these layers extracts more complex features from the input image.Next, the output from the final pooling layer is flattened and passed through three dense layers with 200, 150 units respectively. Both layers have a ReLU activation function. The dense layers are fully connected layers that are used to make the final prediction. Finally, the output of the second dense layer is passed through a final dense layer with 2 units and a sigmoid activation function. The output is a binary classification problem.

The model is then compiled using categorical cross-entropy loss, the Adam optimizer, and accuracy as the evaluation metric. The model is trained using the fit() method on the x_train and y_train data, with a validation data of (x_val, y_val) and the early stopping callback. The model stops training when the performance on a validation set stops improving and restore the best weights.

Overall, this architecture can be a good starting point for detecting ships in satellite images. However, as with any deep learning model, the specific architecture and set of techniques used will depend on the problem you're trying to solve, the dataset you have available, and the computational resources you have at your disposal. The performance of the model can be improved by using more advanced architectures, incorporating transfer learning, data augmentation and fine tuning the hyperparameters.Finally, we trained the model for 100 epochs using x_train and y_train data for training data and x_val and y_val for validation data. We can then use the trained model to make predictions on the new data.



Graph 1. Model Graph

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 80, 80, 32)        896

 max_pooling2d (MaxPooling2D  (None, 40, 40, 32)       0
 )

 dropout (Dropout)           (None, 40, 40, 32)        0

 conv2d_1 (Conv2D)           (None, 40, 40, 32)        9248

 max_pooling2d_1 (MaxPooling  (None, 20, 20, 32)       0
 2D)

 dropout_1 (Dropout)         (None, 20, 20, 32)        0

 conv2d_2 (Conv2D)           (None, 20, 20, 32)        9248

 max_pooling2d_2 (MaxPooling  (None, 10, 10, 32)       0
 2D)

 dropout_2 (Dropout)         (None, 10, 10, 32)        0

 conv2d_3 (Conv2D)           (None, 10, 10, 32)        9248

 max_pooling2d_3 (MaxPooling  (None, 5, 5, 32)         0
 2D)

 dropout_3 (Dropout)         (None, 5, 5, 32)          0

 flatten (Flatten)           (None, 800)               0

 dense (Dense)               (None, 200)               160200

 dense_1 (Dense)             (None, 150)               30150

 dense_2 (Dense)             (None, 2)                 302

=================================================================
Total params: 219,292
Trainable params: 219,292
Non-trainable params: 0
```

Figure 1. Model Summary

Second Model

We changed or second model because of the it was very slow and unefficient. In this model is a convolutional neural network (CNN), which is a type of neural network that is particularly well-suited for image recognition tasks. A CNN is made up of multiple layers, each of which learns a hierarchical representation of the input image.

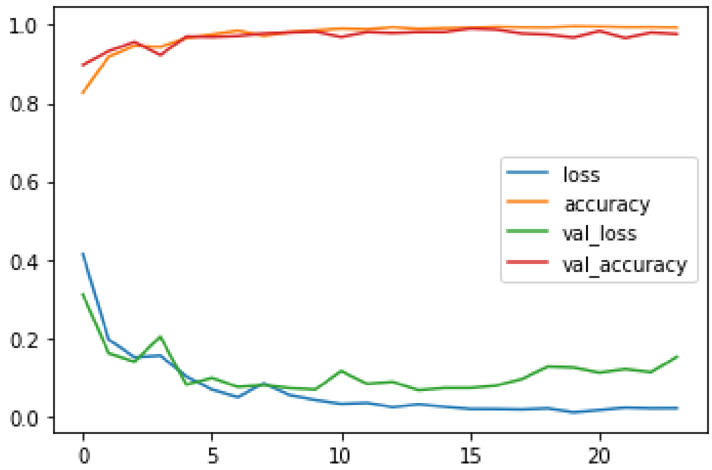
The model starts with a 2D convolutional layer with 64 filters, a kernel size of (3,3), same padding, and a ReLU activation function. This layer extracts the features from the input image. The output is then passed through a max pooling layer with a pool size of (2,2), which reduces the dimensions of the output from the convolutional layer. The output is also passed through a dropout layer with a rate of 0.25. This is a regularization technique that helps prevent overfitting by randomly setting some of the output values to zero.

The model includes several more convolutional, max pooling and dropout layers, with increasing number of filters as well as stride of 3. Each of these layers extracts more complex features from the input image, with the number of filters increasing with each subsequent layer.

Next, the output from the final pooling layer is flattened and passed through two dense layers with 512 and 256 units respectively. Both layers have a ReLU activation function. The dense layers are fully connected layers that are used to make the final prediction. Finally, the output of the second dense layer is passed through a final dense layer with len(y_train[0]) units and a softmax activation function.

The model is then compiled using binary cross-entropy loss, the Adam optimizer, and accuracy as the evaluation metric. The model is trained using the fit() method on the x_train and y_train data, with a validation data of (x_val, y_val) and the early stopping callback. The model stops training when the performance on a validation set stops improving and restore the best weights.

In overall, this architecture is a good starting point for detecting ships in satellite images, however it may not necessarily be the best architecture for the specific problem and dataset at hand. The performance of the model can be improved by using more advanced architectures, incorporating transfer learning, data augmentation and fine tuning hyperparameters.



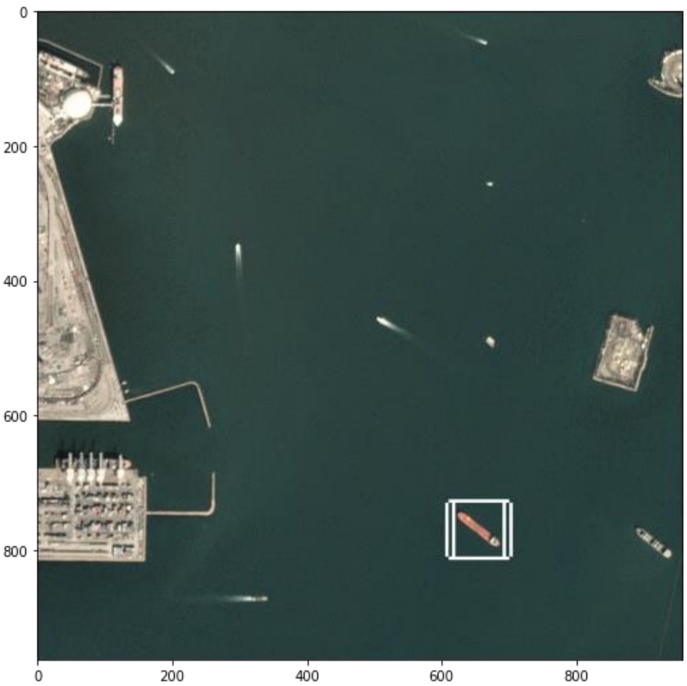
Graph 2. Model Graph

```
Model: "sequential_2"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d_8 (Conv2D)           (None, 80, 80, 64)        1792

 max_pooling2d_8 (MaxPooling  (None, 40, 40, 64)       0
 2D)

 dropout_9 (Dropout)         (None, 40, 40, 64)        0

 conv2d_9 (Conv2D)           (None, 40, 40, 128)       73856

 max_pooling2d_9 (MaxPooling  (None, 20, 20, 128)      0
 2D)

 dropout_10 (Dropout)        (None, 20, 20, 128)       0

 conv2d_10 (Conv2D)          (None, 20, 20, 256)       295168

 max_pooling2d_10 (MaxPoolin  (None, 10, 10, 256)      0
 g2D)

 dropout_11 (Dropout)        (None, 10, 10, 256)       0

 flatten_2 (Flatten)         (None, 25600)             0

 dense_5 (Dense)             (None, 512)               13107712

 dropout_12 (Dropout)        (None, 512)               0

 dense_6 (Dense)             (None, 256)               131328

 dense_7 (Dense)             (None, 2)                 514

=================================================================
Total params: 13,610,370
Trainable params: 13,610,370
Non-trainable params: 0
_________________________________________________________________
```

Figure 2. Model Summary

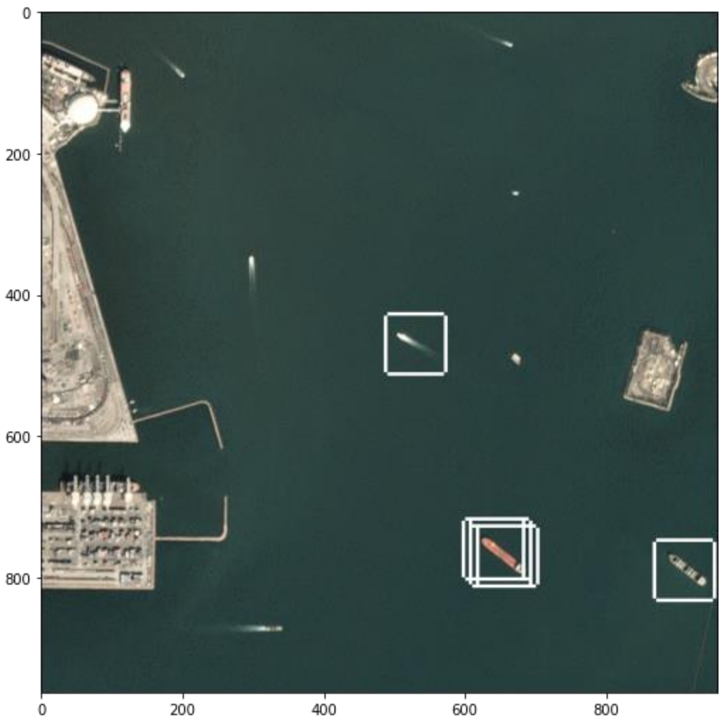
Test Results on Big Scale Map

Our first model's test in Picture 1



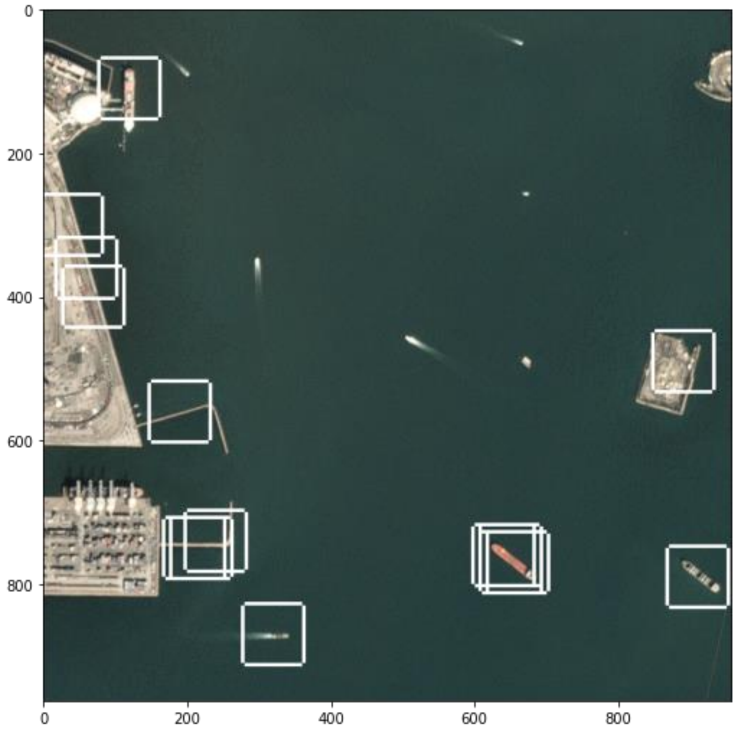
Picture 1. First Model Test.

This the our second model's test.



Picture 1. Second Model Test

This is the our second model’s test with optimized scanning and GPU accelerated in Picture 3.



Picture 1. Second Model Test

CONCLUSION

In conclusion, this article describes the process of building and training a convolutional neural network (CNN) for the task of detecting ships in satellite images using the Keras library. The first model consisted of five 2D convolutional layers, a smoothing layer, and three dense layers. The model was trained using the fit() method on the x_train and y_train data, with a validation data of (x_val, y_val) and the early stopping callback. The Second Model was designed to improve the performance and efficiency of the first model, by increasing the number of filters, reducing the dimensions and changing the kernel size. Our works concludes that while this architecture can be a good starting point for detecting ships in satellite images, it can be improved by using more advanced architectures, incorporating transfer learning, data augmentation and fine tuning the hyperparameters.

REFERENCES

- Kıran E., Karasulu B.,Borandağ E., "Gemi Çeşitlerinin Derin Öğrenme Tabanlı Sınıflandırılmasında Farklı Ölçeklerdeki Görüntülerin Kullanımı ", Zeki Sistemler Teori ve Uygulamaları Dergisi ,5(2) (2022) 161-167

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G., Davis, A., Dean, J. et al., 2016, TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. Distributed, Parallel, and Cluster Computing (cs.DC); Machine Learning (cs.LG) arXiv Preprint, arXiv: 1603.04467.

- Dolapcı, B., Özcan, C., 2021. Automatic Ship Detection and Classification using Machine Learning from Remote Sensing Images on Apache Spark, Journal of Intelligent Systems: Theory and Applications, 4(2), 94-102.

- K. Eldhuset, “An automatic ship and ship wake detection system for spaceborne SAR images in coastal regions,” *IEEE Trans. Geosci. Remote Sens.,*vol. 34, no. 4, pp. 1010–1019, Jul. 1996.

- Beşbinar, B. and Alatan, A. (2015). Inshore ship detection in high resolution satellite images: approximation of harbours using sea-land segmentation, *Proc. of SPIE,* 9643(2D), 2-11.

- Hong, Z.; Jiang, Q.; Dong, H. Fast ship detection based on multi-threshold image segmentation. Comput. Sci. 2006, 33, 273–275.

- Shaoming Zhang, Ruize Wu, Kunyuan Xu, Jianmei Wang, Weiwei Sun “R-CNN-Based Ship Detection from High Resolution Remote Sensing Imagery” Remote Sensing, 2019.

- www.kaggle.com/datasets/rhammell/ships-in-satellite-imagery

INTERDISCIPLINARY WORKSPACE

- Environment / Energy
- Logistics / Carriage/ Transportation