T.C.

MANİSA CELAL BAYAR ÜNİVERSİTESİ

MÜHENDİSLİK FAKÜLTESİ

BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Deep Learning Based Ship Classificaiton from Multi-Modal Satellite Images

Tasarım Projesi / Lisans Bitirme Tezi

HAZIRLAYANLAR

Ali Osman BEKER 180315010

Sümeyra POLAT 180315055

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 "Deep Learning Based Ship Classificaiton 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 : Sümeyra POLAT 180315055

Projenin Savunulduğu Tarih : 09.12.202

Bilgisayar Mühendisliği Bölüm Başkanı

DOÇ. DR. TUĞBA ÖZACAR ÖZTÜRK

CONTENTS

GRAPH LIST

Sayfa

FIGURE LIST

ABSTRACT

Agriculture, defense and intelligence, energy, and the financial sector are just a few of the industries that might benefit from satellite imagery's unique insights. The entire Earth is being photographed every day by constellations of small satellites used by new commercial imaging suppliers like Planet. Machine learning and computer vision algorithms are needed to aid automate the analysis process as the volume of new imagery is outpacing businesses' capacity to manually review each image that is produced. The purpose of this project is to contribute to the challenging task of locating huge ships in satellite photos. Numerous problems, such as supply chain analysis and port activity monitoring, can be solved by automating this procedure.
Keywords: CNN ,R-CNN, DINOv2, Forwar&Forward

INTRODUCTION

The project Deep Learning Based Ship Detection from Multi-Modal Satellite photos intends to develop a deep learning-based system for autonomously identifying ships in satellite photos. A sizable collection of multi-modal satellite images, including images from several sensors and wavelengths, will be used to train the system. The system will be able to identify ships in a variety of circumstances and situations by utilizing deep learning algorithms, which will be an improvement over current ship detection techniques. The system will be judged based on how well it can precisely and consistently spot ships in satellite photos, and its potential uses will be investigated.

PROBLEM DEFINITION

The difficulty of properly and reliably detecting ships in satellite photos is the issue that the Deep Learning Based Ship Detection from Multi-Modal Satellite photos project seeks to address. The accuracy and robustness of traditional ship identification techniques are frequently constrained, and they are also susceptible to false positives and false negatives. This might be an issue for applications like marine surveillance and ship movement tracking, where having accurate and trustworthy information regarding the presence of ships is crucial. The project's goal is to create a system that can get over these constraints and enhance the precision and dependability of ship detection in satellite photos using deep learning techniques.

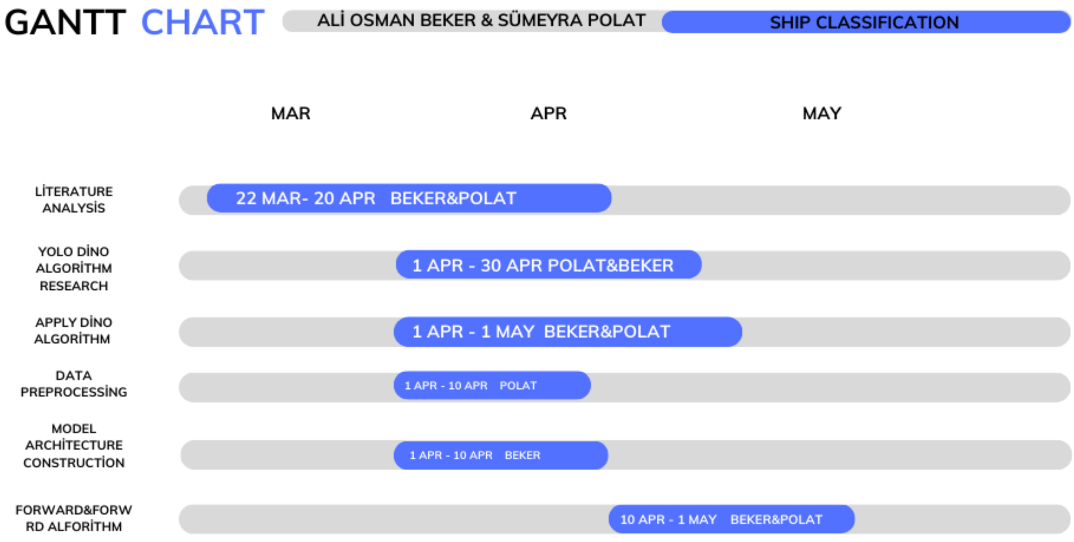
LITERATURE ANALYSIS

The literature was checked for project-related articles, and codes from various projects were rerun and examined. The team worked together to identify the initial optimal model after changing the parameters in the code and putting various models through various stages. The codes were reorganized taking into account the models described and shown in the literature. Depending on the results, the initial model was improved. No image processing techniques were applied to the image. In the course of the project's research, this pre-process will be put to the test.

METHODS AND TECHNOLOGIES TO BE USED

A system for automatically identifying ships in satellite photos will be developed as part of the Deep Learning Based Ship Detection from Multi-Modal Satellite photos project. Deep learning, a sort of machine learning that entails training a sizable neural network on a sizable dataset in order to discover patterns and make predictions, will be the main technique used. The system will be trained using this to identify ships in satellite photos. An extensive collection of multi-modal satellite photos will be gathered and categorized in order to train the deep learning model. photos from several sensors and wavelengths will be included, giving the system access to a variety of sources of data on the ships seen in the photos. The deep learning model will be trained using the labeled dataset to discover ships in the pictures. After the model has been trained, its accuracy and dependability will be tested using a different dataset of satellite photos. The system's performance will be compared to that of more conventional methods of ship recognition in order to gauge how well it can reliably and accurately identify ships in the photos. We used Pandas and Numpy for shape our data and Pytorch for making, training and testing for our model. These configuration modifications were made mainly on the foundations that we currently created in the first period. Towards the final delivery, the project will be completed using the DINOv2 Algorithm that Facebook has just published and which we will talk about in the Works section.

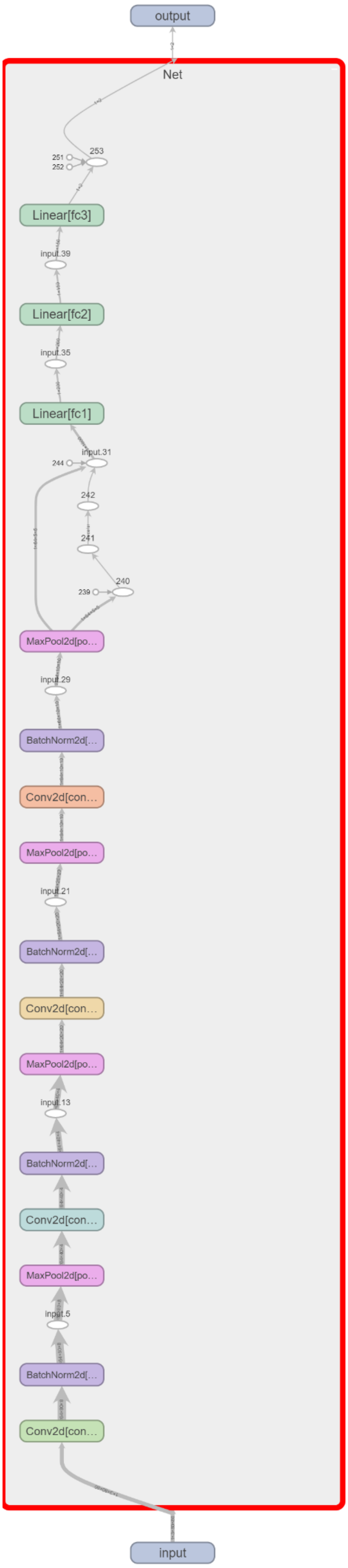
PROJECT TEAM AND TASK SHARING



STUDIES

First, we used the pandas package to transform 80x80 RGB data in.json format containing 4000 photos into a dataframe. Next, we examined the columns that our data contained. The necessary location and seen_ids columns were then removed. Only two data and label columns remain. The data we have is then pre-processed. 3000 of these 4000 photos have ships, while the remaining 1000 show ships without. Then, employing the Numpy module, we assigned the data and labels data to two variables. The data variable we assigned was then brought into the proper format using the reshape function. We first divided the data we formed into train and test data in order to train our model that we will construct. We set aside &80 of data for the test_size 20% train and validation set when we separated these sets. The remaining train and validation data, 25% validation, and 55% train set were then distributed. In Keras, a well-liked open source deep learning package, we created a convolutional neural network (CNN). The CNN is a Sequential model that consists of three dense layers, a smoothing layer, and five 2D convolutional layers.

In the first Graduation Project, we used the Tensorflow library for deep learning. However, we had to switch to Pytorch because we couldn't use both the training time and the GPU effectively, and at the same time we couldn't fully create the experimental models. This has helped us a lot. We made it more convenient both when adding the model to a website and downloading the model. Apart from that, we were able to train the Forward&Forward algorithm, which I will talk about later, more comfortably and effectively. In the later stages of the project, we will segment the picture with the DINOv2 algorithm and then scan the picture with our own model. Then we will detect and categorize the ships point by point with the SAM algorithm.



Dataset and Preprocessing

We start by gathering a dataset including pictures of both ships and non-ship items. Training, validation, and test sets are separated out from the dataset. We preprocess the data by scaling the photos to a fixed size, leveling the pixel values, and applying data augmentation methods like rotation, flipping, and brightness modifications to improve the diversity of the training set before feeding the images into the model.

Model Architecture and Implementation

The PyTorch nn.Module is a subclass of the model that is being provided, entitled Net. The following layers make up the Net class's architecture:

Four convolutional layers make up the customized CNN model, which are each followed by a batch normalization layer, a max-pooling layer, and a dropout layer. Convolutional layers are followed by three fully connected layers, which are then flattened. The last layer has two output nodes that match to the classifications "Ship" and "Not A Ship." To provide probability ratings for each class, the final layer employs the softmax activation function.

The PyTorch nn.Module is a subclass of the model that is being provided, entitled Net. The following layers make up the Net class's architecture:

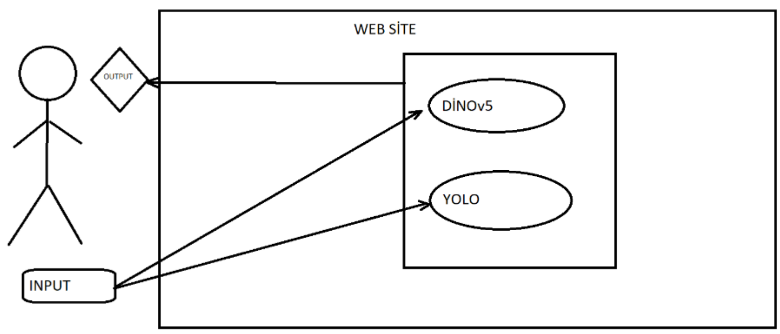
Conv1, Conv2, Conv3, and Conv4 are four convolutional layers, each with 64 filters, a 3-kernel size, and a 1-padding. The local feature representations from the input images are learned by these layers.

Each convolutional layer has its own batch normalization layer (bn1, bn2, bn3, and bn4). By normalizing the layer's input features, batch normalization increases training stability and speed.

Pool1, Pool2, Pool3, and Pool4 are four max-pooling layers with kernel sizes of 2 and strides of 2. These layers lower the feature maps' spatial dimensions, strengthening the model's resistance to slight input translations. With a dropout rate of 0.25, there are four dropout layers (dropout1, dropout2, dropout3, and dropout4). By randomly setting a portion of the input units to 0 during training, the regularization approach known as dropout aids in preventing overfitting. Flatten layer that transforms the last pooling layer's output into a 1D tensor. fc1, fc2, and fc3 are three completely connected (dense) layers with 200, 150, and 2 output units, respectively. The final class probabilities are mapped to the high-level characteristics that the convolutional layers extracted.

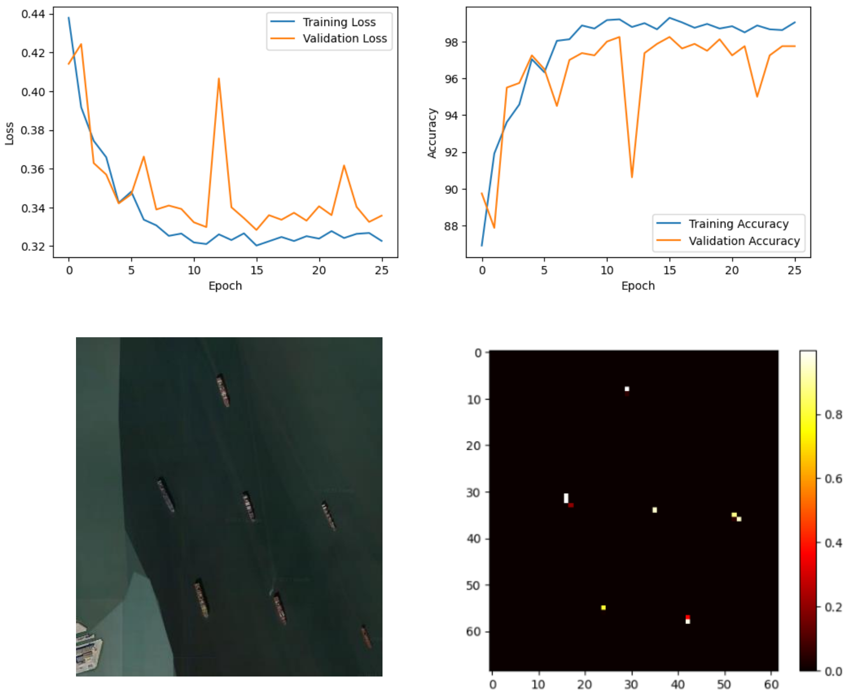
Use Case Model

The algorithms we use will be in a format that users can present in a web site and scan the photo they want, and the user will be able to choose the algorithm himself. The current use case model is demo and aims to explain simply. A more detailed form will be added when prosoned.



Training and Evaluation

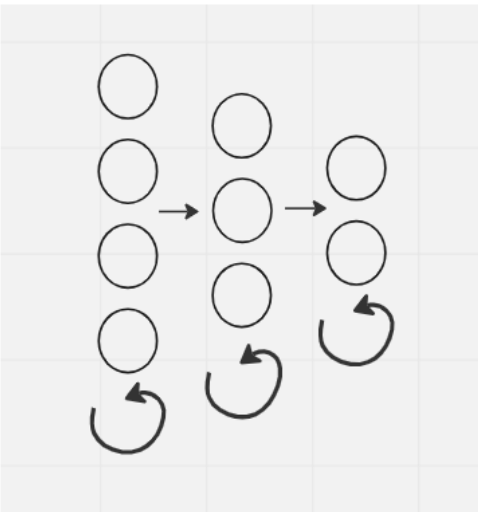
The Adam optimization algorithm and the cross-entropy loss function are used to train the model. To identify the ideal hyperparameters, such as learning rate, batch size, and dropout rate, we do experiments. During training, the model's performance is assessed on a different validation set, and its ultimate performance is evaluated on a test set.



Experimental Model

The suggested model, known as Net, is a multi-layer feedforward neural network. Each layer is a subclass of the nn.Linear class from the torch.nn module and is referred to as Layer. The Adam optimizer is used to generate the layers using ReLU activation functions.

The model is trained independently for each layer using a layer-by-layer method. The mean squared error between the positive and negative instances for each class is used to compute the loss function, which is minimized during the training process for each layer. For a predetermined number of epochs, the training procedure is run, and the loss values are stored for later analysis.



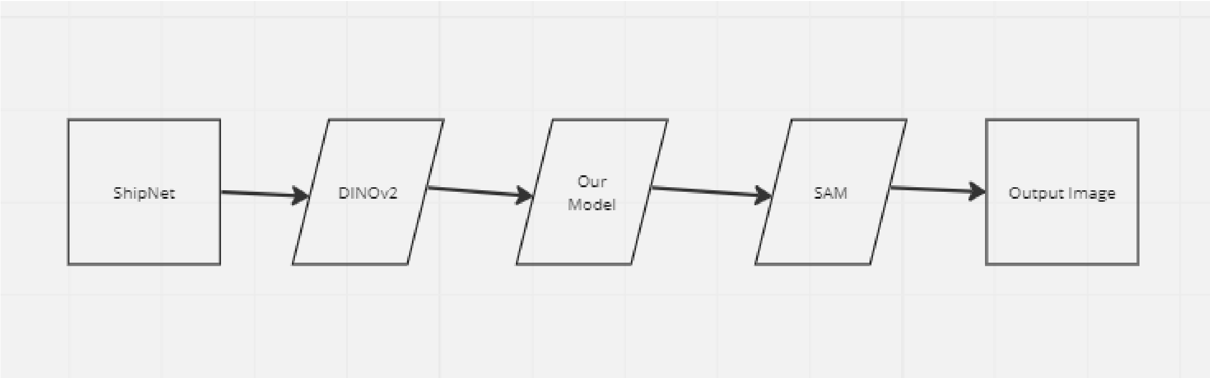
```
training layer 0 ...
training layer 1 ...
training layer 2 ...
train accuracy: 0.734375
test accuracy: 1.0
```

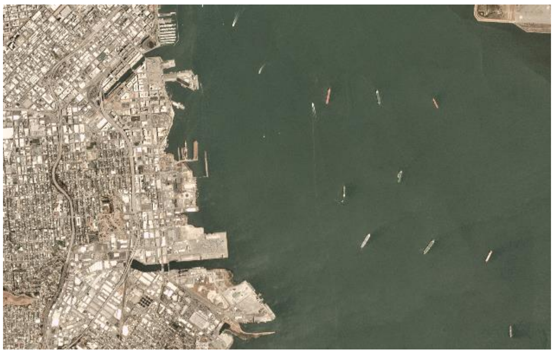
Model Analysis

By examining the learning curves, accuracy, and loss numbers, we examine the model's performance. We also examine the model's resistance to various types of noise, its ability to generalize to other datasets with a comparable structure, and its effectiveness at recognizing ships in varied environments (such as those with various illumination, occlusions, or views).

What Is Next

We are going to use two segmentation model for categorization of shi. Firstly DINOv2 and the other SAM(Segment of Anything) but we are working on it.





Results and Discussion

We determine the advantages and disadvantages of the specialized CNN model for ship detection in photos through our studies. We offer insights into the model's performance in various scenarios and suggest tactics for enhancing its efficiency. The findings help to clarify the difficulties in ship detecting tasks and the potential applications of deep learning methods.

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

- Geoffrey Hinton "The Forward-Forward Algorithm: Some Preliminary Investigations"

- Meta AI Research∗ "DINOv2: Learning Robust Visual Features without Supervision"

- Meta AI Research "Segment Anything"

INTERDİSCİPLİNARY WORKSPACE

- Environment / Energy
- Logistics / Carriage/ Transportation