Aboveground Storage Tank Detection Using Faster R-CNN and High-Resolution Aerial Imagery
by

Qianyu Zhao

Department of Civil and Environmental Engineering
Duke University

Date:_____
Approved:

_____
Mark Borsuk, Advisor

_____
Kyle Bradbury

_____
David Carlson

Thesis submitted in partial fulfillment of
the requirements for the degree of Master of Science in the Department of
Civil and Environmental Engineering in the Graduate School
of Duke University

2021

ABSTRACT

Aboveground Storage Tank Detection Using Faster R-CNN and High-Resolution Aerial Imagery
by

Qianyu Zhao

Department of Civil and Environmental Engineering
Duke University

Date:_____
Approved:

_____
Mark Borsuk, Advisor

_____
Kyle Bradbury

_____
David Carlson

An abstract of a thesis submitted in partial
fulfillment of the requirements for the degree
of Master of Science in the Department of
Civil and Environmental Engineering in the Graduate School of
Duke University

2021

Abstract

In recent years, NaTech disasters, which are defined as the technological accidents caused by natural events, have led to huge losses all over the world. To reduce these losses, assessments of the vulnerabilities of industrial facilities are necessary. In this study, an effort was made to locate aboveground storage tanks from remotely sensed imagery. A dataset that identifies different types of tanks was generated. The data were acquired from the National Agriculture Imagery Program (NAIP) and tanks were labeled as closed roof tank, external floating roof tank, spherical tank, water treatment tank, or water tower according to their shapes. After collecting these data, the Faster R-CNN algorithm, an object detection architecture, was applied to test the performance of this algorithm on the prelabeled dataset. Results of testing indicate that the algorithm could well achieve the goal that having a high recall rate for all the classes of tanks. The precision and recall rates were 82.92% and 90.03% for closed roof tanks, 85.85% and 91.68% for external floating roof tanks, 34.81% and 60.26% for spherical tanks, 49.63% and 89.33% water treatment tanks, 9.43% and 38.46% for water towers. For spherical tanks and water towers, although having low precision and recall, the percentage of missed tanks was extremely low, which is 2.08% and 0 respectively. These results suggest that this aboveground storage tank dataset and the pretrained model generated from Faster R-CNN could be further used in future work for tank detection and vulnerability assessment.

Dedication

To my family who gave me unconditional support to pursue my goals. To my advisor Prof. Borsuk for giving me the opportunity to participate in this project. To my colleague Celine Robinson for helping me overcome the difficulties when I was stuck.

Contents

List of Tables

List of Figures

1. Introduction

1.1 Motivation for Study

"Natural" and "Technological" hazards are traditionally treated as independent phenomena. However, so-called "NaTech" disasters have revealed strong interactions between the two hazards in recent years. NaTech events refer to the failures of critical or civil infrastructure and industrial facilities caused by natural hazards and leading to the release of hazardous material into the environment (Gheorghiua et al., 2014). For instance, hurricanes may destroy onshore or offshore aboveground storage tanks (ASTs). The petroleum products or other chemical products in the ASTs may be released into the air, land, and water and be shifted to surrounding ecosystems and communities. Changing environmental conditions have the potential to increase the risk of tank accidents, so it is becoming more critical to mitigate ecosystems and communities' vulnerabilities in extreme weather conditions and assess the fragility of civil infrastructure and industrial facilities (Krausmann et al., 2011) . Locating those aboveground storage tanks is the first step of the assessment.

In recent years, AST detection is especially challenging and has attracted increased attention. ASTs play an important role in energy systems and industry. They are large containers generally made of metal and concrete pads and designed to store a great variety of liquid or gas substances, such as water and petrochemical products. ASTs can be difficult to identify from remote sensing. According to the products they store, ASTs come in different sizes and have different shapes of tops, which makes them challenging to detect. At the same time, ASTs suffer from high risk of vulnerability due

to natural hazards and extreme weather. For example, because of Hurricane Katrina, over 200 releases of contaminants were reported, including hazardous chemicals, oil, and natural gas (Santella et al., 2010).

1.2 Remote Detection

As remote sensing technology has been successfully developed in recent decades, more high-resolution remotely sensed imagery is now available. These datasets facilitate many applications, such as infrastructure surveying, ecosystem recovery from land use, and natural hazard prediction. In these applications, object detection is one of the most important steps. Object detection, including feature extraction and learning methods, is most often applied, and has attracted much research (Leitloff et al., 2010; Shi et al., 2013; Chen et al., 2013; Huang et al., 2007; Huang et al., 2012; Huang., et al 2014).

1.3 Related work

Various methods have been proposed to identify ASTs. In previous literature, unsupervised methods, such as the hierarchical model, Hough transform, and the shape-guide saliency models were applied to detect circular-shaped tanks (Chen et al., 2006; Li et al., 2008; Yi-li et al., 2011; Han & Fu 2012; Zhu et al., 2012; Kushwaha et al., 2013; Cai et al., 2014; Xu et al., 2014; Yao et al., 2014; Ok 2013; Ok & Başeski 2015; Jing et al., 2018). However, these methods have the limitation that they detect tanks according to certain shapes and colors. As industry has developed, new demands have been raised, requiring tanks in various shapes and colors. Therefore, traditional shape- and-color-based models are not able to achieve multi-type tank detection.

Supervised learning of convolutional neural network (CNN)-based algorithms has been implemented for a variety of object detection and classification tasks. The classification task is classifying the class a given image belongs to, while object detection helps to identify the location of an object in an image. In 2012, Krizhevsky et al. (2012) attracted scholars' attention on CNNs by showing high accuracy of image classification. Girshick et al. (2014) bridged the gap between image classification and object detection by showing that a CNN can lead high object detection performance. The CNN-based algorithm Girshick et al. used was called R-CNNs, which yielded high object detection performance. Then R-CNN models have since been improved by increasing speed and accuracy, called Fast R-CNN and Faster R-CNN (Girshick 2015; Ren et al., 2015). Recently, scholars have tried Faster R-CNN in tank detection. In 2020, Zalpour et al. (2020) employed Faster R-CNN to extract regions of interest and then applied a support vector machine to classify storage tanks from the background of an image.

Faster R-CNN performs well for extraction and object detection in regions of interest but is supervised learning, which means that it requires labeled data as input. Only a couple of open access datasets contain bounding box or pixel level annotations. NWPU-VHR-10 is a prelabeled dataset and contains "storage-tank" as one of the labels (Cheng et al., 2014; Cheng & Han 2016; Cheng et al., 2016). However, due to the lack of metadata, NWPU-VHR-10 was not applied in this study. The Oil and Gas Tank dataset is another prelabeled dataset, but the angle of projection is different from other datasets, which leads it to have bad performance when testing images from other datasets (Rabbi

2020). Therefore, in this study, we created a prelabeled and open access dataset that includes various types of tanks for research use.

1.4 Purpose of the thesis

AST detection is a difficult challenge, not only because of the complex background, which contains a great variety of objects, such as buildings and trees, but also due to the differences between imagery datasets. Imagery datasets are collected in different methods, such as by satellite imagery or by aircraft with digital cameras. Therefore, the same tank can have different shapes and colors when presented in different datasets due to the change of resolution. This implies that tools trained to previous datasets do not perform well on more recent datasets.

In addition to the difficulties of data sources, different types of tanks also cause difficulty. ASTs store various liquids and gases. To satisfy different storage requirements, more types of tanks have been designed. As a result, multi-type tank detection is of more significance. Although, in previous literature, scholars have attempted to detect tanks from high-resolution imagery (Zhang et al., 2015; Wang et al., 2016; Özyurt 2019), no research exists on different types of tank detection. Therefore, the first objective of this thesis is to generate a prelabeled dataset using the National Agriculture Imagery Program (NAIP) data, which was acquired between 2018 and 2019. Also, this study makes multi-type tank detection feasible by labeling tanks according to their features. The dataset, which may help in artificial intelligence development, will be shared and published online.

To present the performance of supervised learning algorithms on our prelabeled dataset, the Faster R-CNN algorithm was used for training and testing. Initial training results in a pretrained model with pretrained weights for Faster R-CNN that can be used to test again other images. As deep learning models can take a long time to train, other researchers could benefit by applying our pretrained model to their dataset. For this reason, the second objective of this thesis is to test the performance of the Faster R-CNN on our prelabeled dataset and generate a pretrained model for further use.

2. Data

2.1 Aboveground Storage Tanks

Huge quantities of fuel, water, chemicals, and petroleum products are required for industrial entities and manufacturers, the petrochemical industry in particular. Tanks are therefore designed to fulfill storage needs of the industry. ASTs are usually large containers. They are made of stainless steel, fiberglass, or polyethylene, depending upon the substances being stored.

Different types of industrial ASTs that referenced from GSC Tanks include:

1) Fixed-roof tanks

Fixed-roof tanks are primarily used to store water. They have a cone or dome roof, which is permanently attached to the cylindrical shell.

2) Internal floating roof tanks

Internal floating roof tanks are covered by an external cone or dome roof and also equipped with an internal floating roof. The internal roof rises and falls with the

fluid level, reducing evaporative loss. The external roof protects the fluid inside from extreme weather like heavy rain and strong winds.

3) External floating roof tanks

External floating roof tanks consist of a cylindrical steel shell that is open but contains a floating roof on the surface of the fluid.

4) Variable vapor space tanks

Variable vapor space tanks are typically spherical-shaped in order to store gas at high pressures.

5) Water treatment sedimentation basins

A sedimentation basin is a circular basin, designed to remove settleable solids by gravity from wastewater. There are typically visible sludge scrapers that collect the sludge into a central hopper, which makes them distinguishable in satellite imagery.

6) Water towers

A water tower has a water tank supported by an elevated structure. It is distinguishable from storage tanks because of a large head and a thin pillar beneath the head.

In this study, because it is impossible to tell the difference between fixed-roof tanks and internal floating roof tanks from remotely sensed data, we labeled them together as closed_roof_tanks. There are therefore a total of five labels to represent the above types of tanks. The labels are: closed_roof_tank for fixed-roof tanks and internal floating roof tanks; external_floating_roof_tank for external floating roof tanks; spherical_tank for variable vapor space tanks; water_treatment_tank for sedimentation

basins; and water_tower for water towers. In Appendix A, there are image samples of different types of tanks.

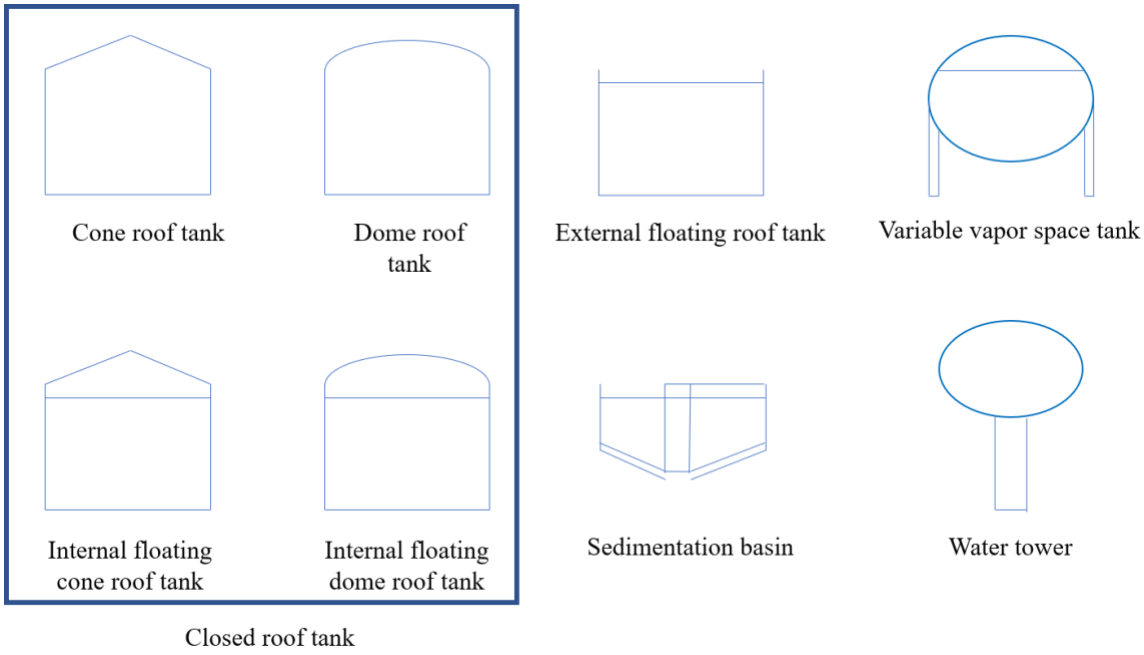

Figure 1: Types of tanks

Table 1: Labels of different types of tanks

Types of tanks	Labels
Fixed-roof tanks	closed_roof_tank
Internal floating roof tanks	closed_roof_tank
External floating roof tanks	external_floating_roof_tank
Variable vapor space tanks	spherical_tank
Sedimentation basins	water_treatment_tank
Water towers	water_tower

2.2 Data Sources

This study was conducted using aerial images of large natural gas plants, petroleum refineries, petroleum terminals, and crude oil storage fields. To collect the aerial images of these sites, the location information of these sites including longitude and latitude is needed. The four sources of the location information are: (1) the Homeland Infrastructure Foundation-Level Data (HIFLD) by the Department of Homeland Security (U.D. of Homeland Security), (2) the U.S Energy Mapping System from the Energy Information Administration (EIA) (U.E.I. Administration), (3) the Greenhouse Gas Reporting Program (GHGRP) by the Environmental Protection Agency (EPA), and (4) large crude oil storage fields across the United States. The information on tank types is not given by these sources. Human review on the labeled data was the way to check whether tanks were labeled as the right types.

The aerial images were acquired from the National Agriculture Imagery Program (NAIP). NAIP (U.-F.A.A.P.F.Office), which is administered by the USDA's Farm Service Agency (FSA), collects aerial imagery of the continental U.S. during the agricultural growing seasons. The datasets are available for governmental agencies and the public. NAIP began in 2003 with a five-year cycle. In 2009, NAIP transferred to a three-year cycle. From 2008, NAIP applied digital sensors instead of film cameras to capture images. Also, the imagery was delivered with four bands of data--RGB (red, green, blue) and near infrared color--in contrast to three bands (RGB) before 2007. In 2019, the spatial resolution of the imagery is 0.6 meter, which means 1 pixel in an image covers 0.36 m^2 of the Earth's surface.

Figure 2: Image sample

2.3 Characteristics of Data

There are a total of 8238 positive images (contain tanks) in our dataset. 5% of these images were divided into a test set (400). We also randomly picked 400 negative images (not containing tanks) for this test set. In the other 95% of positive images, 80% were in training set and 20% were in validation set.

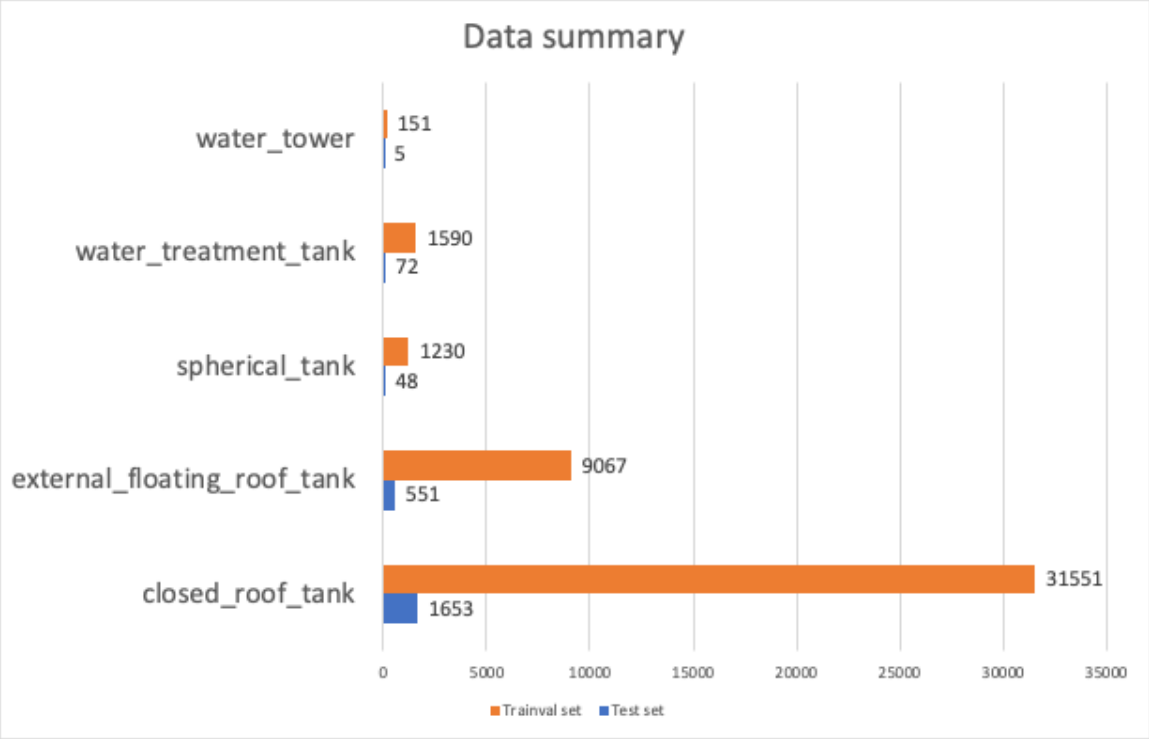

Figure 3: Summary of the dataset

In the whole dataset, there are 45918 tanks including 33204 closed roof tanks, 9618 external floating roof tanks, 1278 spherical tanks, 1662 water treatment tanks, and 156 water towers. In the training and validation sets, there are 43589 tanks containing 31551 closed roof tanks, 9067 external floating roof tanks, 1230 spherical tanks, 1590 water treatment tanks, and 151 water towers.

In the test set, there are 800 images (400 negative and 400 positive), including 2329 tanks. 1653 of them are closed roof tanks, 551 external floating roof tanks, 48 spherical tank, 72 water treatment facility, and 5 water towers.

3. Methods

3.1 Data Preprocessing

In the NAIP dataset, every image tile has a quadrangle name. For instance, "m_4207140_se_19_060_20181016" indicates that the file is in USGS quadrangle 42071, in the SE quarter-quad, in UTM zone 19, and was acquired on 10/16/2018. Because of the images' large size (about 10000*12000 pixels), these tiles are too large for training and testing. Therefore, in this study, they are clipped into images, with a 512*512-pixel size. The process is shown in Figure 4.

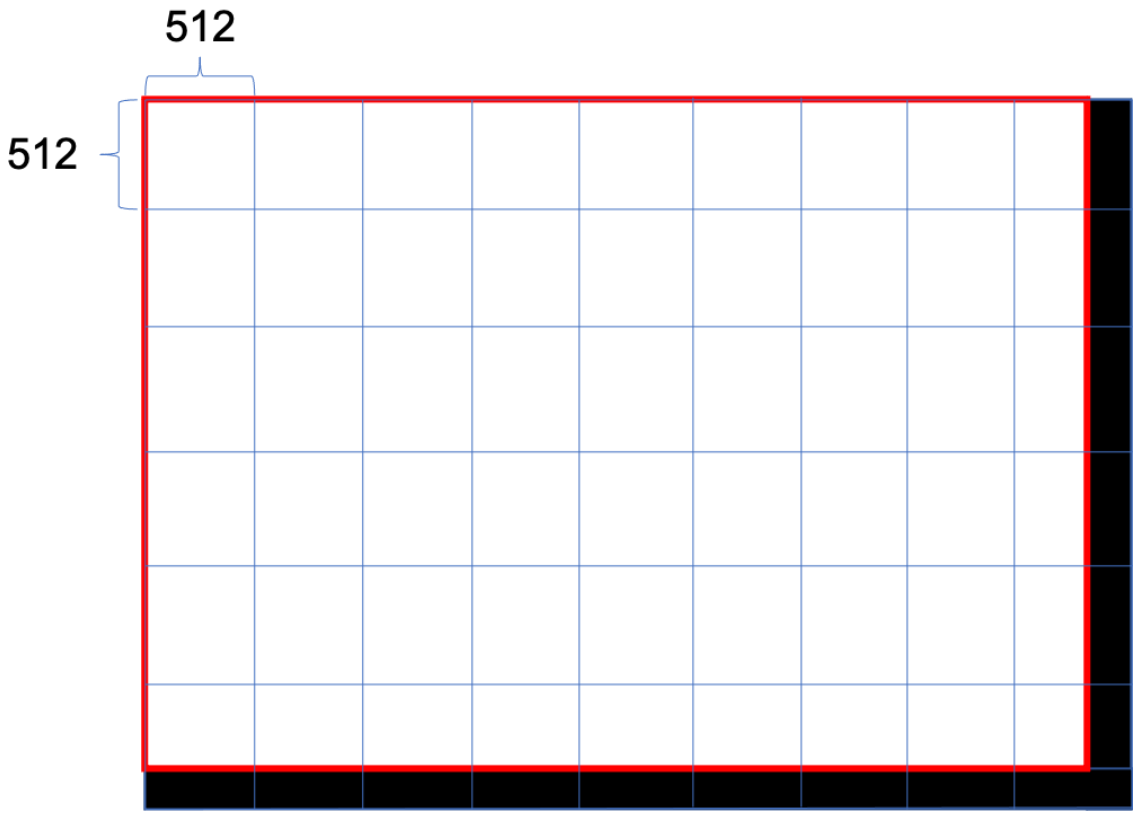

Figure 4: Illustration of data clipping

The white area is the tile, which is divided a grid of 512*512 pixels. If the length or width of a tile is indivisible by 512, then the length or width of the tile is enlarged to a

512-divisible number by filling with black as the black edges shown in Figure 4, which is why some chips have black areas. Such black areas do not influence training results in object detection because only the data of the object area in a chip matter.

3.2 Annotation

LabelImg (https://github.com/tzutalin/labelImg) is an annotation tool that is written in Python. In this study, it was applied to annotate different types of tanks in the images.

In LabelImg, every tank was labeled as closed_roof_tank, external_floating_roof_tank, spherical_tank, water_treatment_tank, or water_tower. Every tank is labeled even though only part of a tank may appear in the chip. Shadows of tanks were not included because the time that the tiles in NAIP were acquired influences the shape and size of the shadows of tanks. After annotation, the annotation files were stored as Extensible Markup Language (XML) files that all have an extension name of ".xml". The training set, validation set, and test set were stored and used in PASCAL VOC2007 format.

3.3 Faster R-CNN Implementation

In 2015, Ren et al. proposed Faster R-CNN, an excellent object detection algorithm (Ren et al., 2015). Compared to Fast R-CNN, Faster R-CNN improved efficiency and accuracy by reducing false positives. Faster R-CNN has the advantages of lower computational burden and no restrictions on the input images' sizes. As shown in Figure 4, after reshaping the input image, the feature maps extracted from a CNN model were passed into two stages: the region proposal network (RPN) and the Fast R-CNN

detector. RPN uses an image to generate a set of object proposals as output. The object proposals are classified by the Fast R-CNN detector. In other words, RPN extracts regions for Fast R-CNN to identify. Both RPN and Fast R-CNN are explained in Figure 5.

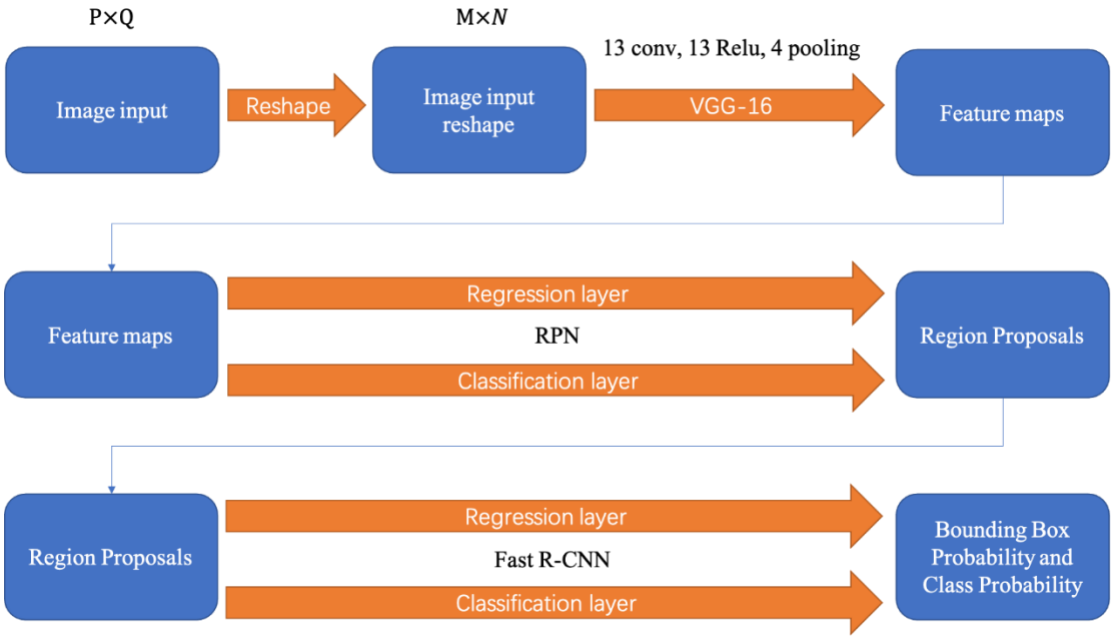

Figure 5: Structure of Faster R-CNN

RPN generates region proposals by a CNN model. In this study, VGG-16 (Simonyan & Zisserman 2014), which has 13 shareable convolutional layers (13 convolutional layers, 13 Relu layers, and 4 pooling layers), is applied to extract feature maps. In the RPN, to generate a set of rectangular region proposals (also called anchors, 9 anchors are generated in this study), a spatial window of 3*3 convolution kernel is applied. Features of each anchor are passed into two fully-connected layers: a regression layer and a classification layer. The regression layer predicts the location of these anchors; while the classification layer gives the probability that these anchors belong to

foreground or background. The outputs of the regression layer and classification layer serve as the region proposals which are passed to the Fast R-CNN network.

After receiving region proposals from RPN, the Fast R-CNN network sends these proposals into a pooling layer. This information then is converted to a vector which is sent to two fully connected layers: regression and classification layers. These two layers calculate the probability of the class and location for each proposal.

The loss function of Faster R-CNN model is given as follows:

$$L_{(tot)} = L_{(rpn\_loss\_cls)} + L_{(rpn\_loss\_box)} + L_{(frc\_loss\_cls)} + L_{(frc\_loss\_box)} \quad (1)$$

where $L_{(tot)}$ is total loss. $L_{(rpn\_loss\_cls)}$ and $L_{(rpn\_loss\_box)}$ are classification and regression loss in the RPN, and $L_{(frc\_loss\_cls)}$ and $L_{(frc\_loss\_box)}$ are classification and regression loss in the Fast R-CNN model.

The detailed classification and regression loss function in RPN and Fast R-CNN is given as follows (Ren et al., 2015):

$$L(\{p_i\},\{t_i\}) = \frac{1}{N_{cls}}\sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}}\sum_i p_i^* L_{reg}(t_i, t_i^*) \quad (2)$$

Here, i is the index of anchor in a mini-batch. N_{cls} is the mini-batch size and N_{reg} is the number of anchor locations. λ is used to balance the classification part and regression part to be equally weighted. For instance, in this study, N_{cls} is 128, N_{reg} is 1024, then λ is set as 8 to balance N_{cls} and N_{reg}. N_{cls}, the mini-batch size, is one of the hyperparameters that are set before training. N_{reg} is 1024 because a 512*512-pixel image will be reduced by 256 times (16 times on width and 16 times on height) to a 32*32*512 feature map, the number of anchor locations is 32*32 (1024).

In classification part, p_i is the predicted probability that anchor i is a tank. p_i^* is 1 or 0. For anchor i, if it has the highest Intersection-over-Union (IoU) with a manually labeled bounding box or if it has an IoU higher than 0.7 with any manually labeled bounding box, then p_i^* of this anchor is 1. If anchor i is not in any conditions mentioned in prior sentence and has IoU lower than 0.3 for all manually labeled bounding box, then p_i^* of this anchor is 0. If anchor i has no attributes as mentioned above, it will not be included in training process. L_{cls} is log loss over object and not object, whose equation is shown in Eq. (3).

$$L_{cls}(p_i, p_i^*) = -\log\left(p_i p_i^* + (1 - p_i)(1 - p_i^*)\right) \tag{3}$$

In regression part, t_i and t_i^* are two vectors containing the 4 parameterized coordinates for predicted bounding box and the manually labeled bounding box associated with anchor i respectively. The parameterization process is shown as follows:

$$\begin{cases} t_x = \frac{x - x_a}{w_a}, t_y = \frac{y - y_a}{h_a}, \\ t_w = \log\left(\frac{w}{w_a}\right), t_h = \log\left(\frac{h}{h_a}\right), \\ t_x^* = \frac{x^* - x_a}{w_a}, t_y^* = \frac{y^* - y_a}{h_a}, \\ t_w^* = \log\left(\frac{w^*}{w_a}\right), t_h^* = \log\left(\frac{h^*}{h_a}\right), \end{cases} \tag{4}$$

where x and y represent the box's center coordinates. w and h are its width, and height. x, x_a, x^* are for predicted bounding box, anchor box, and manually labeled bounding box respectively (likewise for y, w, and h).

$L_{reg}(t_i, t_i^*) = R(t_i - t_i^*)$ where R is Smooth L1 function, whose equation is shown in Eq. (5).

$$Smooth_{L1}(x) = \begin{cases} 0.5 * x^2, |x| < 1 \\ |x| - 0.5, otherwise \end{cases} \tag{5}$$

3.4 Important Training Hyperparameters

Hyperparameters are parameters control the learning process in machine learning. Hyperparameters do not have influence on the training set, but have impact on the speed and quality of models.

1) Weight Decay

Weight decay is the parameter that controls the impact of regularization on the model. Regularization is normally applied for to mitigate overfitting, such as L2 norm regularization. In this study, weight decay is set as 0.0005.

2) Learning Rate

Learning rate controls how much the model weights are updated in the stochastic gradient descent (SGD). If learning rate is too large, the model will be difficult to converge. If learning rate is too small, the solution may be local optimal rather than global optimal. Normally, learning rate is 0.01 to 0.001. In this study, learning rate is set as 0.001.

3) Learning Rate Decay

Learning rate decay is a strategy to balance the tradeoff that large learning rate and small learning rate. Large learning rate is good at training speed, but not good for model to converge. Small learning rate is the opposite. At the start of training, a relatively large learning rate is used to locate a global optimal solution quickly. Then, every fixed-step time, learning rate will be reduced by a factor to find the global optimal solution.

In learning rate decay, two hyperparameters control the whole process: decay rate and decay step size. In this study, decay rate is 0.1 and decay step size is 10,000, which means every 10,000 times training, the learning rate will reduce to 10% (for example, 0.001 to 0.0001).

4) Momentum

In SGD, momentum is used to increase the speed of training. Momentum can accumulate the gradient of the past steps. This hyperparameter control how much the gradient direction of last step affects that of this step. In this study, momentum is 0.9.

5) Max Iteration

Max iteration is how many steps (iterations) in the training process. It is 50,000 in this study.

6) Batch Size

Batch size is the number of training samples in one step (mini-batch). For small dataset, the batch size is equal to the dataset size. However, for large dataset, too large batch size decreases the speed of the model, and too small batch size will lead to underfitting. Therefore, batch size is normally between dozens to hundreds. The batch size is set as 128 in this study.

Table 2: Summary of hyperparameters

Hyperparameters	Value
Weight decay	0.0005
Learning rate	0.001

Decay rate (learning rate decay)	0.1
Decay step size (learning rate decay)	10,000
Momentum	0.9
Max iteration	50000
Batch size	128

3.5 Data Augmentation

In this study, the effectiveness of Faster R-CNN is improved by data augmentation. Faster R-CNN has no rotation capability because features of objects do not rotate accordingly (Cheng et al., 2014). Therefore, the chips are rotated at angles of 0°, 90°, 180°, 270°, horizontally flipped, and vertically flipped. Thus, the detection accuracy of Faster R-CNN is improved based on increasing the input data by rotating and flipping (Zalpour et al., 2020).

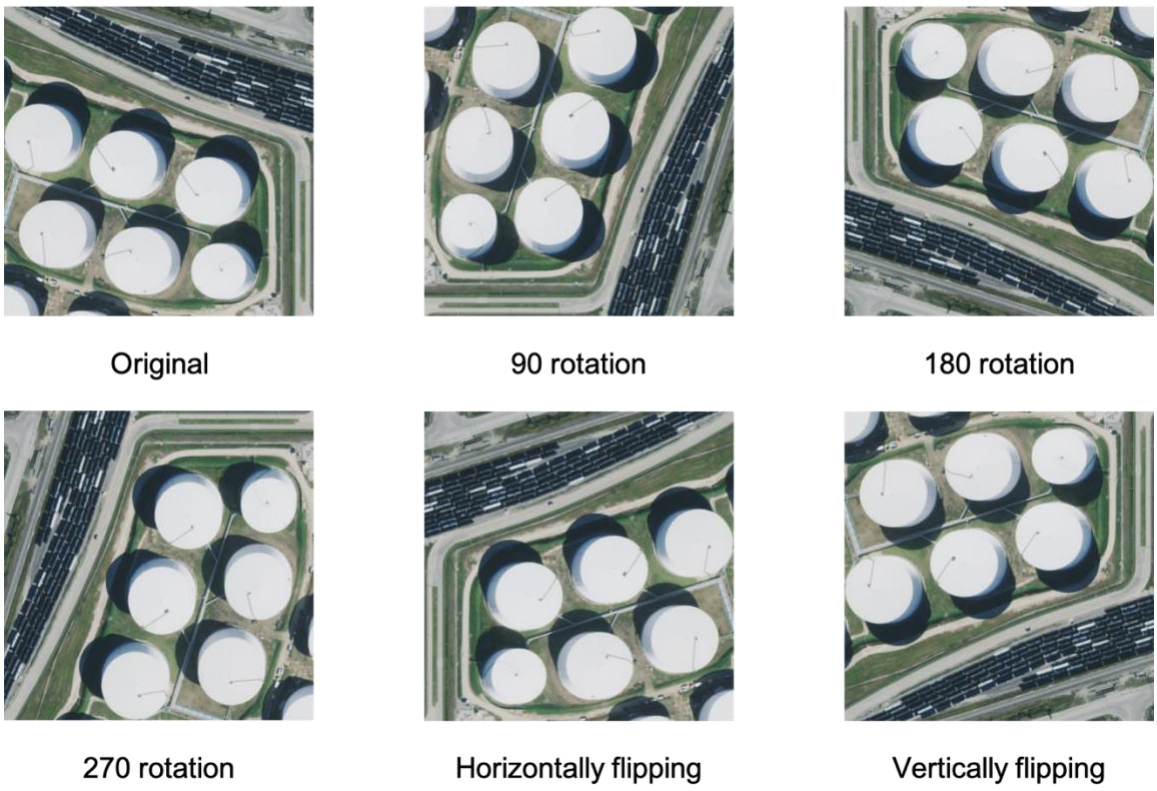

Figure 6: Data augmentation example

3.6 Human Review

The detection results of Faster R-CNN were manually reviewed. In this study, both positive (with tanks) and negative (without tanks) images were tested. We manually identified the number of ASTs in each test image and verified the detected objects and their classes.

3.7 Model Assessment

In this study, confusion matrix, 3 rates (accuracy rate, precision rate, recall rate), and percentage of missed tanks are used to assess the model performance.

Confusion matrix is a table where samples of each row are in a predicted class, while samples of each column are in an actual class. A case that is detected as an object is

Positive, otherwise it is Negative. A case that is in an actual class is Positive, otherwise it is Negative.

Table 3: Confusion matrix

		Actual class	
		Positive (P)	Negative (N)
Predicted class	Positive (P)	TP	FP
	Negative (N)	FN	TN

Accuracy (ACC) is the percentage of correctly detected cases in all testing results, calculated as follows:

$$ACC = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

Precision rate (PPV) is the percentage of correctly detected positive cases in all predicted positive cases, calculated as follows:

$$PPV = \frac{TP}{TP+FP} \quad (7)$$

Recall rate (TPR) is the percentage of correctly detected positive cases in all actual positive cases, calculated as follows:

$$TPR = \frac{TP}{TP+FN} \quad (8)$$

Percentage of missed tanks (PMT) is the percentage of missed cases in all actual positive cases, calculated as follows:

$$PMT = \frac{Number\ of\ missed\ cases}{TP+FN} \quad (9)$$

4. Results and Discussion

4.1 Training Results

After training, the losses are shown in Figure 7.

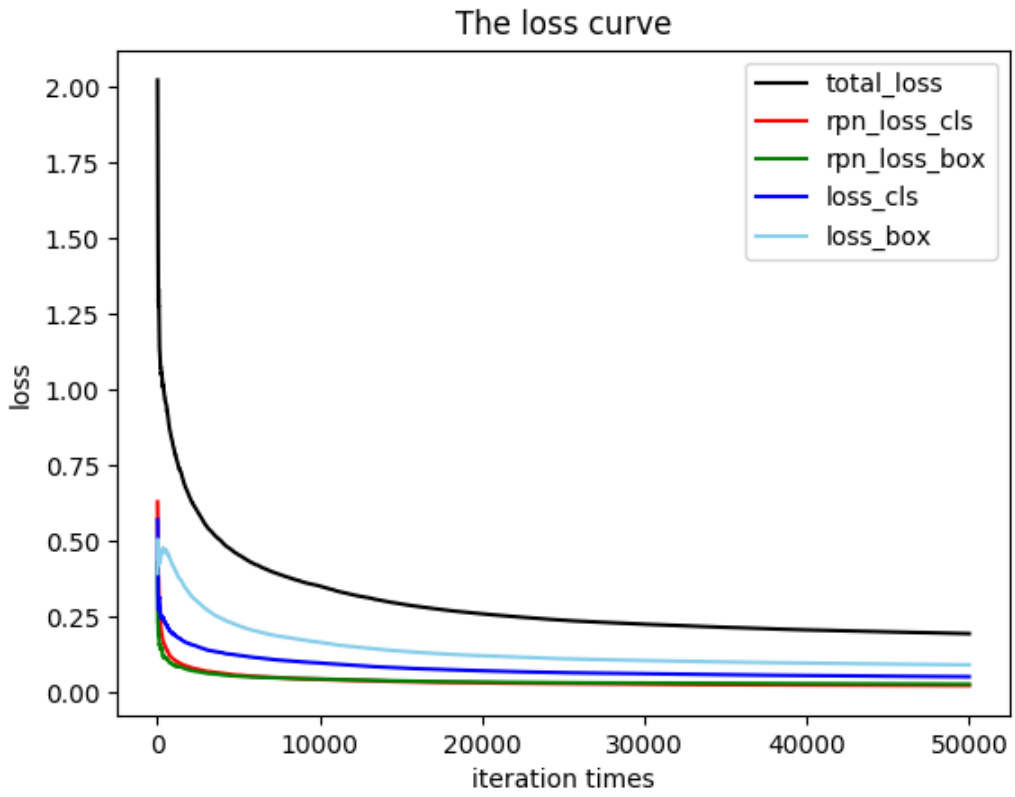

Figure 7: Loss curves

In Figure 7, to better determine whether the model converges, these five curves are smoothed by taking the average every ten values. total_loss curve represents the total losses ($L_{(tot)}$ in Eq (1)). rpn_loss_cls and rpn_loss_box are classification losses ($L_{(rpn\_loss\_cls)}$ in Eq (1)) and regression losses ($L_{(rpn\_loss\_box)}$ of RPN in Eq (1)) respectively. loss_cls and loss_box are classification losses ($L_{(frc\_loss\_cls)}$ in Eq (1)) and regression losses ($L_{(frc\_loss\_box)}$ of Fast R-CNN in Eq (1)) respectively. These five curves

indicate that the model converged after training. Also, the curves suggest that the learning rate is appropriate in this study, because if the learning rate is too high or too low, the model will not converge.

4.2 Testing hyperparameters

In the testing process, two hyperparameters control the results. One is Non-maximum suppression (NMS) threshold, the other one is overlap threshold.

NMS threshold is used to discard overlapping predicted bounding box. If two bounding boxes have IoU larger than NMS threshold, then the bounding box with lower predicted probability will be removed. In this study, some of the storage tanks are cut off because they are at the edge. To avoid that several bounding boxes represent the same tank, the NMS threshold is set as 0.3.

Overlap threshold controls the predicted probability with which the bounding boxes will be kept. In this study, the target is to detect objects as many as possible. In other words, higher Recall rate is needed. Because if the objects are wrong, they can be manually corrected; otherwise, if a tank is not detected, it cannot be checked by a human. As a result, the overlap threshold should be low. In this study, the overlap threshold is set as 0.1.

4.3 Testing Results

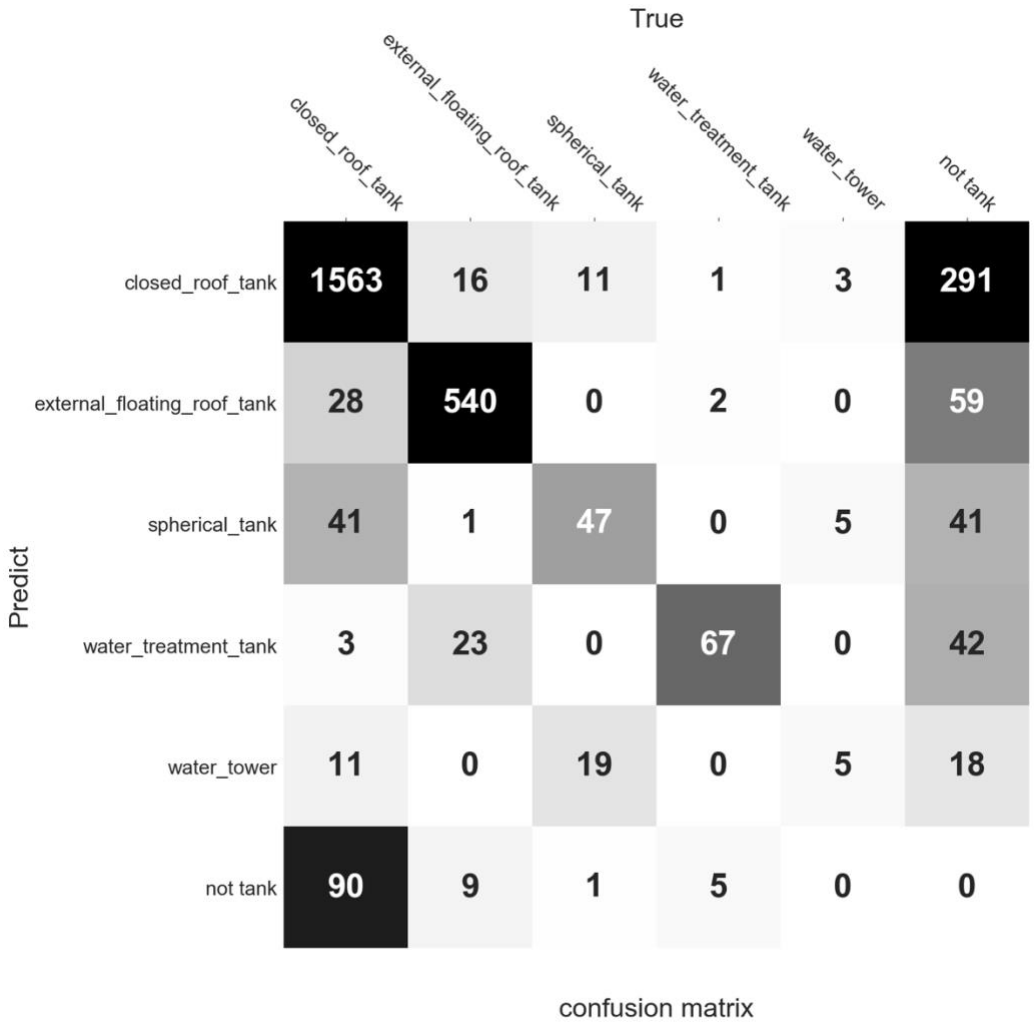

Figure 8: Confusion matrix of testing results

In this confusion matrix, the overall accuracy is 75.53%. A total of 1563 closed roof tanks, 540 external floating roof tanks, 47 spherical tanks, 67 water treatment tanks, and 5 water towers were correctly detected. 90 closed roof tanks, 9 external floating roof tanks, 1 spherical tank, and 5 water treatment tanks were missing. To assess the model performance, analysis of each class is necessary. The target criteria of a good performance in this study are: 1) high TPR and 2) low PMT (percentage of missed tanks).

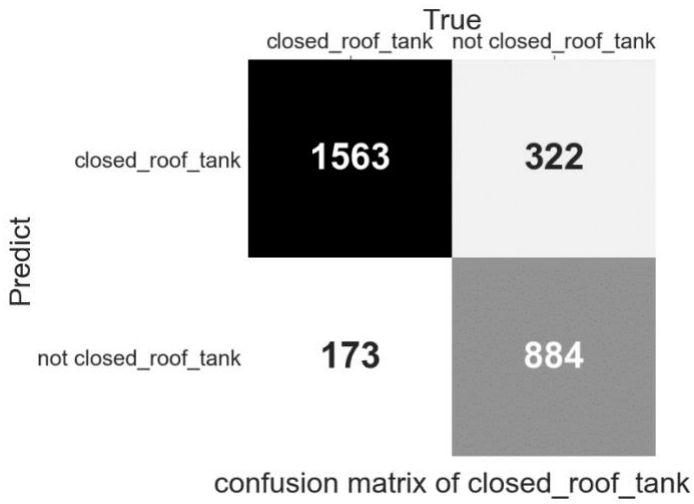

Figure 9: Confusion matrix of closed_roof_tank

For closed roof tanks, the PPV is 82.92%, TPR is 90.03%. From Figure 8, 90 closed roof tanks were not detected, which means the percentage of missed tanks was 5.18%. All the detected closed roof tanks were correctly predicted as closed roof tank. In summary, the performance of the model on closed roof tank was good because of high TPR and low PMT.

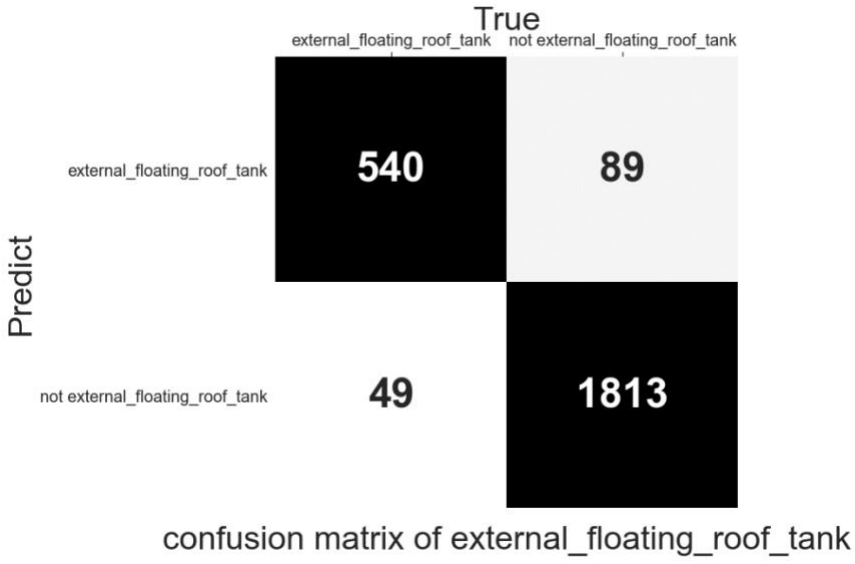

Figure 10: Confusion matrix of external_floating_roof_tank

For external floating roof tanks, the PPV is 85.85%, TPR is 91.68%. From Figure 8, 9 external floating roof tanks were not detected, which means the percentage of missed tanks was 1.53%. Only 2 external floating roof tanks were wrongly predicted. Actually, this was a labeling fault. Two closed roof tanks were labeled as external floating roof tank. Excluding this manual problem, the model could well detect external floating roof tanks with high PPV, TPR, and PMT.

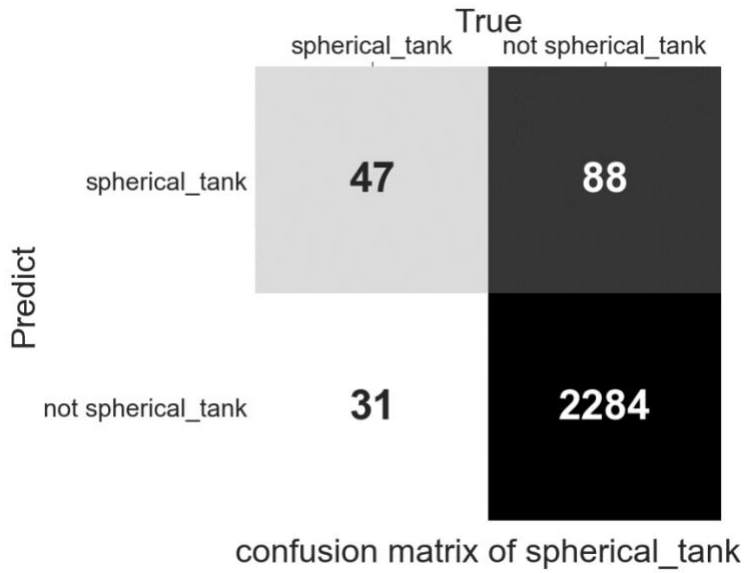

Figure 11: Confusion matrix of spherical_tank

For spherical tanks, the PPV is 34.81%, TPR is 60.26%. From Figure 8, one spherical tank was not detected, which means the percentage of missed tanks was 1.28%. The model's capacity to detect spherical tanks should be further tested. From the view of missed spherical tanks, it performed well for the reason that only one was missed. However, the TPR are relatively low. In some cases, the model predicted closed roof tanks as spherical tank at a low probability, which may be the reason of low PPV and TPR.

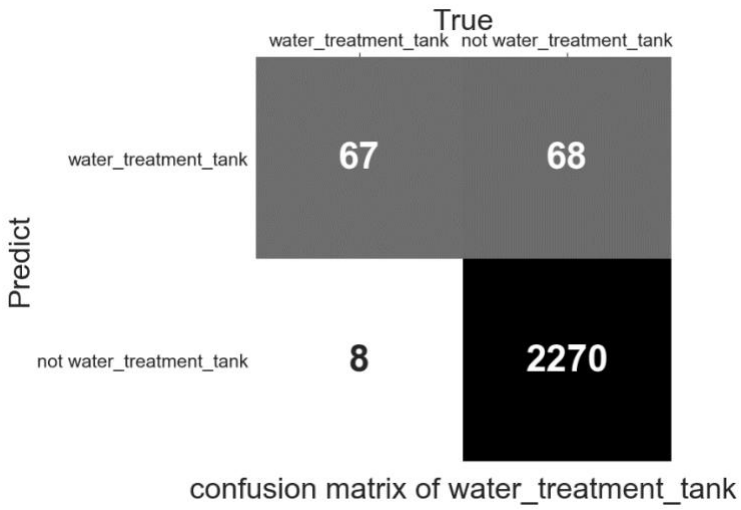

Figure 12: Confusion matrix of water_treatment_tank

For water treatment tanks, the PPV is 49.63%, TPR is 89.33%. From Figure 8, five water treatment tanks were not detected, which means the percentage of missed tanks was 6.67%. The model had good performance on water treatment tanks, because the TPR was high, which achieved our target criteria.

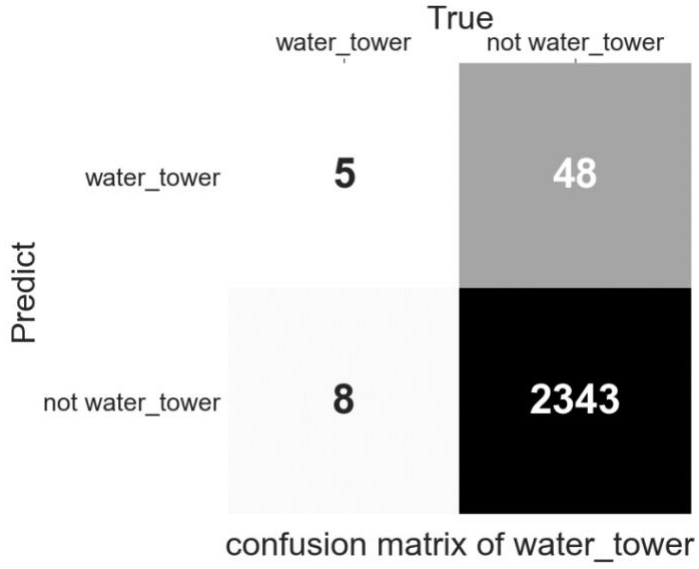

Figure 13: Confusion matrix of water_tower

For water towers, the PPV is 9.43%, TPR is 38.46%. From Figure 8, all water towers were detected, which indicated the model could correctly predict water towers.

However, the PPV and TPR are low. The reason may be that the samples of water towers are much less than other types of tanks.

5. Conclusion

In this study, a high-resolution aerial imagery dataset was established. It contains five labels of different tank types. Faster R-CNN was applied to train and test this dataset. The results show that this dataset can be well trained to complete multi-type tank detection task. Also, the pretrained model's performance on closed roof tanks, external floating roof tanks, and water treatment tanks has been demonstrated to be excellent.

Although, this dataset and the pretrained model were enough to achieve the objectives of this thesis, future work could focus on the hyperparameters of the model. In testing part, the overlap threshold could be adjusted to have a better PPV or TPR. At a low overlap threshold, in some cases, several labels were predicted on the same tank. Therefore, increasing the overlap threshold may lead to a higher rate for both TPR and PPV. In addition to the work of model, more data could be involved in this dataset. More samples contain water towers and spherical tanks could be involved to have a better pretrained model.

6. Implications

To assess the fragility of ASTs in NaTech events, criteria should be different depending on the type of a tank. In this study, Faster R-CNN algorithm was demonstrated to be excellent on multi-type tank detection. The pretrained model trained

from Faster R-CNN will be used to locate tanks and identify their types in more research areas to help NaTech event assessment. In addition, a dataset that contains five tank types was generated. This dataset is flexibly used to be trained by unsupervised or supervised algorithms for object detection tasks, which makes it be a good source for scholars to conduct research on artificial intelligence in the further.

Appendix A: Image samples

Image samples of tanks in different types are shown in LabelImg software.

Figure 14: closed_roof_tank

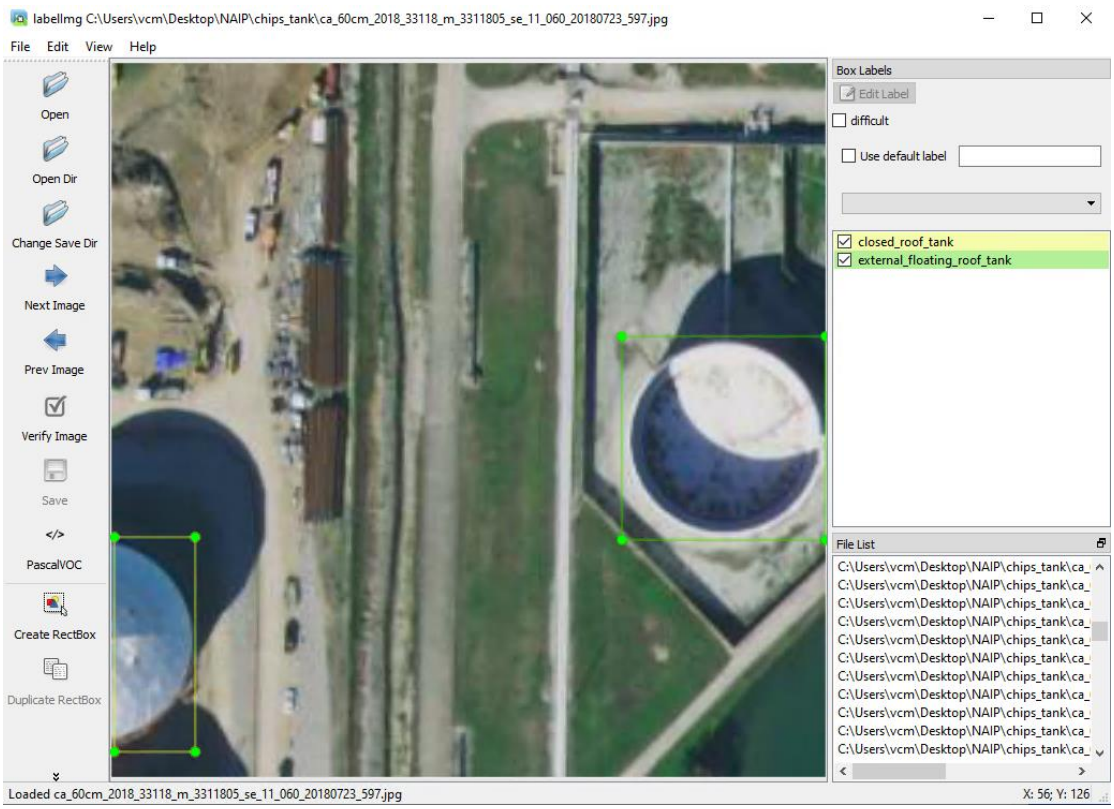

Figure 15: external_floating_roof_tank

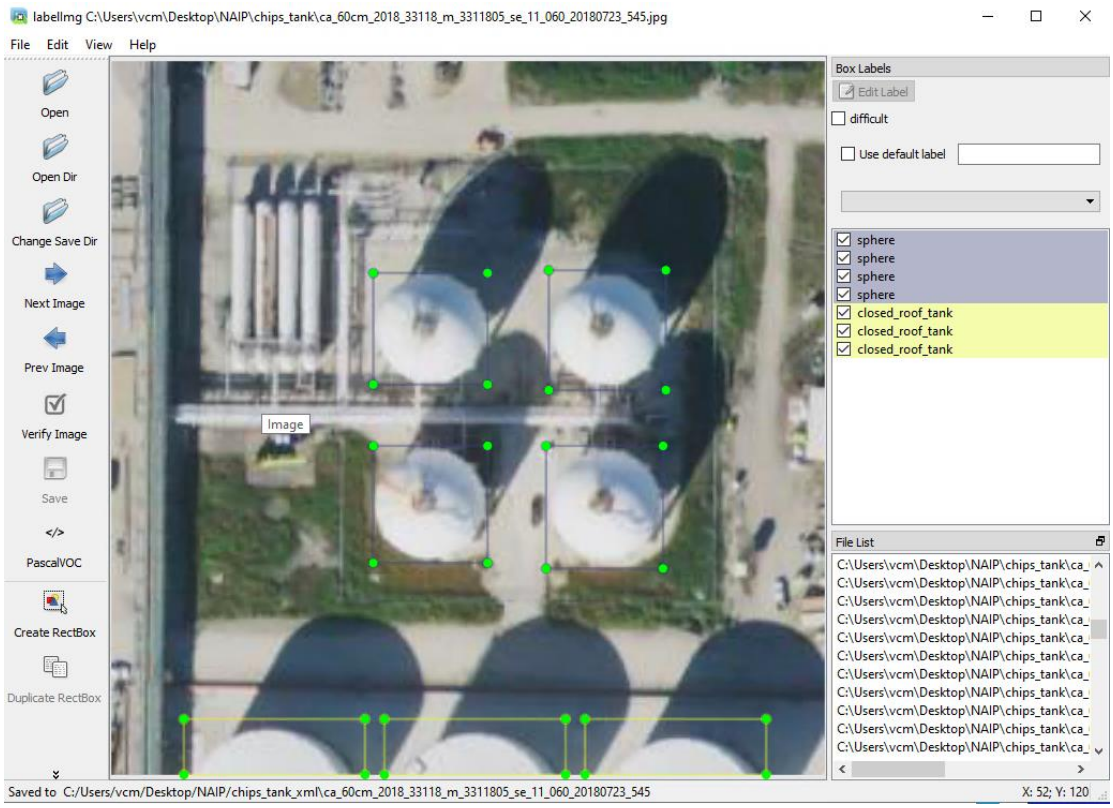

Figure 16: spherical_tank

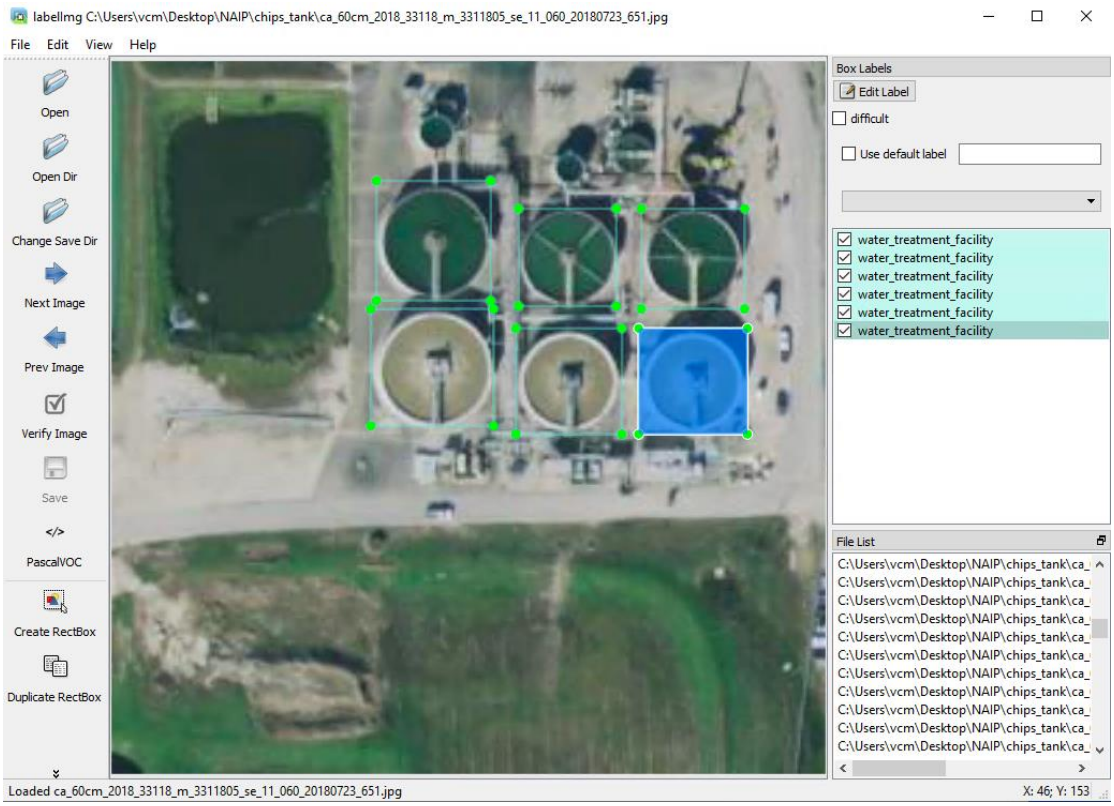

Figure 17: water_treatment_tank

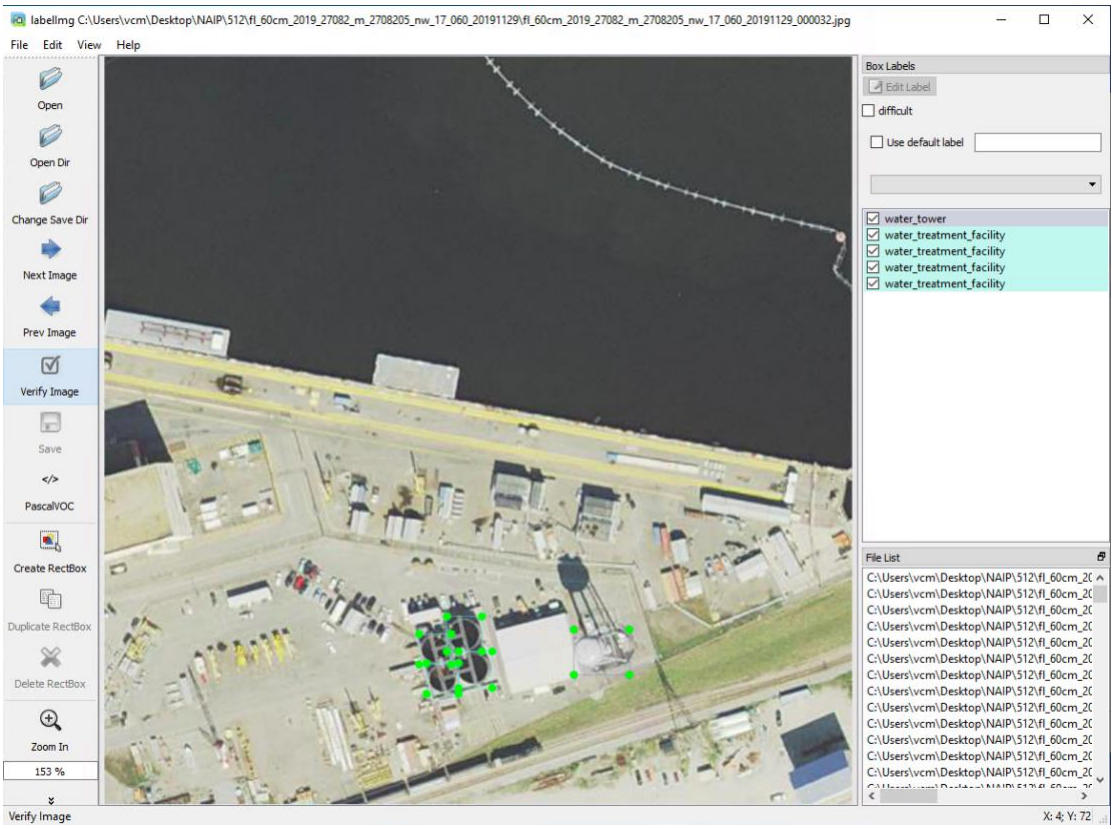

Figure 18: water_tower

Appendix B: Testing result sample

Figure 19: Testing result sample

References

Cai, X., Sui, H., Lv, R., & Song, Z. (2014, May). Automatic circular oil tank detection in high-resolution optical image based on visual saliency and Hough transform. In *2014 IEEE Workshop on Electronics, Computer and Applications* (pp. 408-411). IEEE.

Chen, A. J., & Li, J. Z. (2006). Automatic recognition method for quasi-circular oil depots in satellite remote sensing images. *Opto-Electronic Engineering*, *9*, 021.

Chen, X., Xiang, S., Liu, C. L., & Pan, C. H. (2013, November). Aircraft detection by deep belief nets. In *2013 2nd IAPR Asian Conference on Pattern Recognition* (pp. 54-58). IEEE.

Cheng, G., & Han, J. (2016). A survey on object detection in optical remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *117*, 11-28.

Cheng, G., Han, J., Zhou, P., & Guo, L. (2014). Multi-class geospatial object detection and geographic image classification based on collection of part detectors. *ISPRS Journal of Photogrammetry and Remote Sensing*, *98*, 119-132.

Cheng, G., Zhou, P., & Han, J. (2016). Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, *54*(12), 7405-7415.

Han, X., & Fu, Y. (2012). Circular array targets detection from remote sensing images based on saliency detection. *Optical Engineering*, *51*(2), 026201.

Huang, X., Zhang, L., & Li, P. (2007). Classification and extraction of spatial features in urban areas using high-resolution multispectral imagery. *IEEE Geoscience and Remote Sensing Letters*, *4*(2), 260-264.

Huang, X., & Zhang, L. (2012). An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery. *IEEE transactions on geoscience and remote sensing*, *51*(1), 257-272.

Huang, X., Lu, Q., & Zhang, L. (2014). A multi-index learning approach for classification of high-resolution remotely sensed images over urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, *90*, 36-48.

Gheorghiua, A. D., Töröka, Z., Ozunua, A., Antonionib, G., & Cozzanib, V. (2014). NaTech risk analysis in the context of land use planning. Case study: petroleum products storage tank farm next to a residential area. *CHEMICAL ENGINEERING*, *36*.

Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440-1448).

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580-587).

Greenhouse gas reporting program (ghgrp). https://www.epa.gov/ghgreporting. Accessed: 2021-2-5.

GSC Tanks. Aboveground Storage Tanks: A Complete Guide for Users. https://www.gsctanks.com/aboveground-storage-tanks/. Accessed: 2021-2-5.

Jing, M., Zhao, D., Zhou, M., Gao, Y., Jiang, Z., & Shi, Z. (2018). Unsupervised Oil Tank Detection by Shape-Guide Saliency Model. *IEEE Geoscience and Remote Sensing Letters*, *16*(3), 477-481.

Krausmann, E., Cozzani, V., Salzano, E., & Renni, E. (2011). Industrial accidents triggered by natural hazards: an emerging risk issue. *Natural Hazards and Earth System Sciences*, *11*(3), 921-929.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, *25*, 1097-1105.

Kushwaha, N. K., Chaudhuri, D., & Singh, M. P. (2013). Automatic Bright Circular Type Oil Tank Detection Using Remote Sensing Images. *Defence Science Journal*, *63*(3).

Leitloff, J., Hinz, S., & Stilla, U. (2010). Vehicle detection in very high resolution satellite images of city areas. *IEEE transactions on Geoscience and remote sensing*, *48*(7), 2795-2806.

Li, B., Yin, D., Yuan, X., & Li, G. Q. (2008). Oilcan recognition method based on improved Hough transform. *Opto-Electronic Engineering*, *35*(3), 30-34.

Ok, A. O. (2013). A new approach for the extraction of aboveground circular structures from near-nadir VHR satellite imagery. *IEEE transactions on geoscience and remote sensing*, *52*(6), 3125-3140.

Ok, A. O., & Başeski, E. (2015). Circular oil tank detection from panchromatic satellite images: A new automated approach. *IEEE Geoscience and Remote Sensing Letters*, *12*(6), 1347-1351.

Özyurt, F. (2019). Efficient deep feature selection for remote sensing image recognition with fused deep learning architectures. *The Journal of Supercomputing*, 1-19.

Rabbi, J., Chowdhury, S., & Chao, D. Oil and gas tank dataset. *Mendeley Data*, *3*, 2020.

Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. *arXiv preprint arXiv:1506.01497*.

Santella, N., Steinberg, L. J., & Sengul, H. (2010). Petroleum and hazardous material releases from industrial facilities associated with Hurricane Katrina. *Risk Analysis: An International Journal*, *30*(4), 635-649.

Shi, Z., Yu, X., Jiang, Z., & Li, B. (2013). Ship detection in high-resolution optical imagery based on anomaly detector and local shape feature. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(8), 4511-4523.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

U. D. of Homeland Security. Homeland infrastructure foundation-level data (hifld). https:// hifld-geoplatform.opendata.arcgis.com/. Accessed: 2021-2-5.

U. E. I. Administration. U.s energy mapping system. https://www.eia.gov/state/maps.php. Accessed: 2021-2-5.

U.-F.A.A.P.F.Office, National geospatial data asset (ngda) naip imagery. https://gis.apfo.usda.gov/arcgis/rest/services/NAIP. Accessed: 2021-2-5.

Wang, Q., Zhang, J., Hu, X., & Wang, Y. (2016, May). Automatic detection and classification of oil tanks in optical satellite images based on convolutional neural network. In *International Conference on Image and Signal Processing* (pp. 304-313). Springer, Cham.

Xia, G. S., Hu, J., Hu, F., Shi, B., Bai, X., Zhong, Y., & Lu, X. (2017). AID: A benchmark data set for performance evaluation of aerial scene classification. *IEEE Transactions on Geoscience and Remote Sensing*, *55*(7), 3965-3981.

Xu, H., Chen, W., Sun, B., Chen, Y., & Li, C. (2014). Oil tank detection in synthetic aperture radar images based on quasi-circular shadow and highlighting arcs. *Journal of Applied Remote Sensing*, *8*(1), 083689.

Yang, Y., & Newsam, S. (2010, November). Bag-of-visual-words and spatial extensions for land-use classification. In *Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems* (pp. 270-279).

Yao, Y., Jiang, Z., & Zhang, H. (2014, November). Oil tank detection based on salient region and geometric features. In *Optoelectronic Imaging and Multimedia*

Technology III (Vol. 9273, p. 92731G). International Society for Optics and Photonics.

Yi-li, H. X. W. F., & Gang, L. (2011). Oil Depots Recognition Based on Improved Hough Transform and Graph Search [J]. *Journal of Electronics & Information Technology*, *1*(008).

Zalpour, M., Akbarizadeh, G., & Alaei-Sheini, N. (2020). A new approach for oil tank detection using deep learning features with control false alarm rate in high-resolution satellite imagery. *International Journal of Remote Sensing*, *41*(6), 2239-2262.

Zhang, L., Shi, Z., & Wu, J. (2015). A hierarchical oil tank detector with deep surrounding features for high-resolution optical satellite imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(10), 4895-4909.

Zhu, C., Liu, B., Zhou, Y., Yu, Q., Liu, X., & Yu, W. (2012, July). Framework design and implementation for oil tank detection in optical satellite imagery. In *2012 IEEE International Geoscience and Remote Sensing Symposium* (pp. 6016-6019). IEEE.