Comparison of CNN and SVM for Ship Detection in Satellite Imagery

Nur Jati Lantang Marfu'ah
Department of Informatics
Faculty of Industrial Technology
Yogyakarta
Islamic University of Indonesia
nurjatilantang@gmail.com

Arrie Kurniawardhani
Department of Informatics
Faculty of Industrial Technology
Yogyakarta
Islamic University of Indonesia
arrie.kurniawardhani@uii.ac.id

***Abstract*— . Satellites with optical sensors generate images of the Earth over relatively large areas. Optical satellite image provides unique insights into various markets, including agriculture, defense and intelligence, and energy. Ship detection using satellite images is crucial because it can help manage marine traffic services, defense and intelligence, and fisheries management. In this study, optical satellite images are used for training models for detecting the ship. Machine Learning (ML) algorithms such as deep learning and Support Vector Machine (SVM) have been applied to detect objects in previous studies. Convolution Neural Network (CNN)-based deep learning technology outperformed many algorithms that have existed to some extent [1]. CNN has proven to be able to outperform SVM to detect ships with an average training accuracy is 0,9912 or 99.12% and the validation accuracy is 0,9798 or 97,89%. While SVM gets an accuracy of 0,9438 or 94,38%.**

Keywords— Optical Satellite Imagery; Object Detection; Machine Learning; Convolution Neural Network; Support Vector Machine;

I. INTRODUCTION

Detection of ships in satellite images has been widely applied in maritime security and sea traffic control [2]. Ship detection using satellite images is very important because it can help manage marine traffic services, defense and intelligence, and fisheries management. Remote sensing has a very important role in monitoring ships because of its long operating distance and wide monitoring range[3].

Optical satellite images have a higher image resolution and more content can be displayed than other remote sensing images, which are more suitable for ship detection. However, optical satellite images usually have two main issues: 1) Weather conditions like clouds, fog, and sea waves produce more pseudo targets for ship detection. 2) Optical satellite images with higher resolution naturally produce a more significant amount of data than other remote sensing images [2].

Today, research in the field of computer vision is very popular. In computer vision contained operations starting from capturing object images by a camera system, processing image objects into a more concise and simple form but still representing objects, until the most important is analyzing to determine the type of object[4]

Machine learning algorithms such as deep learning and support vector machine (SVM) have been applied to detect objects in previous studies [2][5][6][7][8][9][10][11][12]. Recently CNN-based deep learning technology outperformed many algorithms that have existed to some extent[1]. SVM was chosen because it is considered as one of the best and uncomplicated initial classification algorithms[5].

Although the Machine Learning algorithm has been widely applied to image classification, an algorithm is not always suitable for every data type or image type. Therefore this study aims to compare which algorithm is better for detecting ship objects between CNN and SVM.

II. LITERATURE REVIEW

A. Optical Satellite Imagery

Satellite imagery became publically available in 1972 and led to the founding of NPA Satellite Mapping (NPA) as a consultancy in the same year. Since then, the evolution in the capabilities of both optical satellites and data processing has been staggering. Satellites with optical sensors generate images of the Earth over relatively large areas.

Recently interest in remote sensing systems using satellite images was growing, such as in maritime security, traffic control, fisheries surveillance, illegal disposal of oil waste, and marine pollution [13]. Optical satellite images provide unique insights into various markets, including agriculture, defense and intelligence, and energy [1] [13] [14][15][16][17][18].

B. Support Vector Machine (SVM)

SVM is one of the best classification algorithms and is not as complicated as Deep Learning [5]. Support vector machine aims to find the hyperplane that maximizes distances between the hyperplane and the support vectors (the closest data points)[19]. In other words, there is labeled training data (supervised learning), the algorithm produces an optimal hyperplane that categorizes new examples. In the two dimensional spaces, this hyperplane is a line separating an airplane into two parts where in each class are located on both sides.

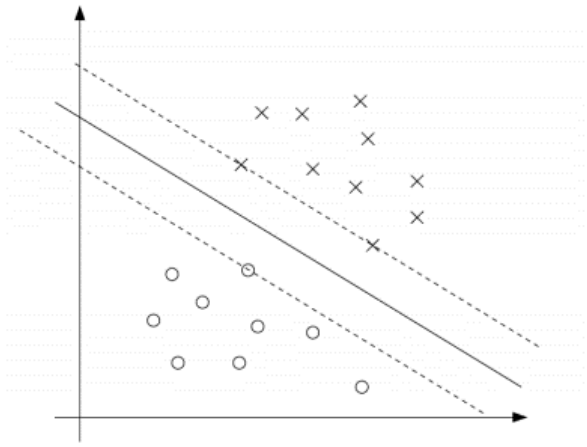


Figure 1. Sample cut to divide into two classes

The problem can be formulated as a quadratic programming that reads.

$$\begin{array}{ll} \underset{x}{\text{minimize}} & \frac{1}{2}||w||^2 \\ \text{subject to} & y^{(i)}(w^T x^{(i)} + b \geq 1),\ i = 1, \ldots, n, \end{array}$$

w,b : parameters of our hypothesis function,

y(i) : represent the label for a specific example,

x(i) : the ith example out of n

γ : the minimum geometric margin of all training examples.

C. Convolution Neural Network (CNN)

Deep learning is used to speed up the learning process of the neural network by using many layers, usually more than 7 layers[20]. One deep learning model that is often used for image classification is Convolution Neural Network (CNN). CNN is one of the variation models in Artificial Neural Network (ANN) algorithm that is widely used in image recognition [23]. In previous studies, many classification algorithms used to detect satellite images were the result of the development of the ANN algorithm, such as Convolution Neural Network[21] [22].

As the name implies, CNN utilizes the convolution process. By moving a certain sized convolution kernel (filter) to an image. The CNN method proved to be able to outperform other Machine Learning methods in the case of object classification in images[23]

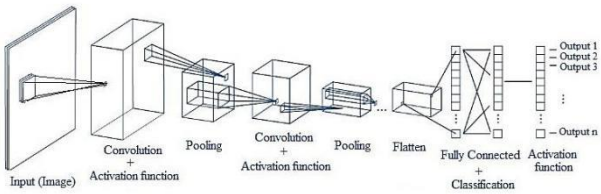


Figure 2. CNN fully connected illustration

D. Keras, TensorFlow, and Scikit-Image

TensorFlow is an open-source for machine learning platform. It has a large ecosystem, libraries, and community, which can help researchers develop advanced technology in machine learning and make it easier for developers to use machine learning in building their systems. Keras is TensorFlow's implementation of the Keras API specification. Keras makes TensorFlow easier to use without sacrificing flexibility and performance. Scikit-Image contains many algorithms for image processing. Scikit-Image is available free and without limitations. [24].

III. METHODOLOGY

A. Data Set

In this study, according to how to obtain the data used can be categorized as secondary data, which is an archive of a provider of commercial satellite imagery called Planet. This institution uses several small satellites to take pictures of the entire Earth every day. The limited equipment resources make it difficult for researchers to get satellite image data directly (primary data), so secondary data is used. Secondary data is data that is not obtained directly by researchers, the data here can be in the form of documents or archives owned by institutions or someone who is the subject of research.

Information about the dataset used:

1. Satellite imagery for the San Francisco Bay and San Pedro Bay regions in California, United States.

2. The satellite imagery used is the capture of port area sightings from above.

3. The type of satellite imagery used is an optical sensor image or cannot go through the cloud.

4. Data consists of 4.000 images with a size of 80 x 80 pixels and 8 optical sensor images in high-resolution ports. From 4.000 images, there are 1.000 images of "ships" (Figure 3.) and 3.000 images of "non-ships" (Figure 4).

5. Image format is PNG or (.png) format.

6. JSON formatted file containing data, labels, scene id's and location metadata.

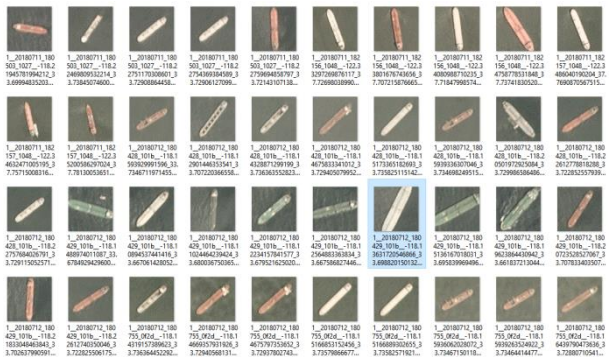


Figure 3. Images of ship

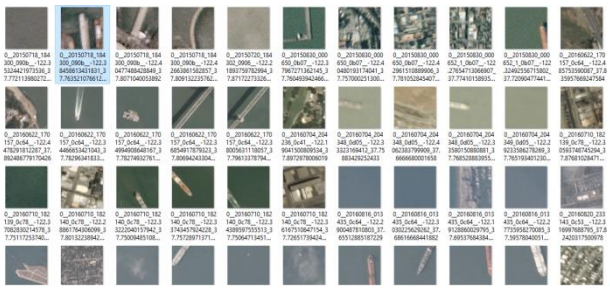


Figure 4. Images of non-ship

B. SVM

We construct the SVM model use scikit-image in Python 3.7 version. We use scikit-image hog function to extract the HOG features. Apart from the HOG features, the color histogram and the raw color feature are also used. The SVM was trained with train/val split of 3.200/800, an image size of 80 x 80. Steps of detection ship using SVM algorithm as shown as Figure 5.

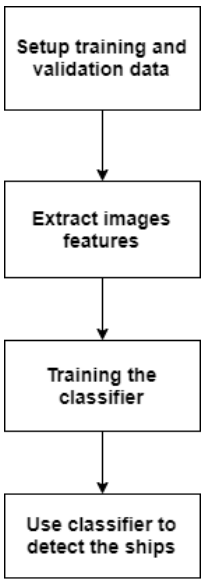


Figure 5. Steps of detection with SVM

C. CNN

We construct the CNN model use Keras from TensorFlow 2 in Python 3.7 version. CNN was trained with train/val split of 3.200/800, an image size of 80 x 80, 18 epochs. Our network consists of 10 layers of CNN as shown in Figure 6.

In this network, there are convolution, max-pooling, and fully connected processes. The convolution process is applied without padding, so it does not change the size of the image both before and after the convolution process. The purpose of max-pooling is to take samples that represent inputs, reduce their dimensions and make it possible to make assumptions about features contained in buried sub-regions. Fully Connected layer connects every neuron in one layer to every neuron in another layer.

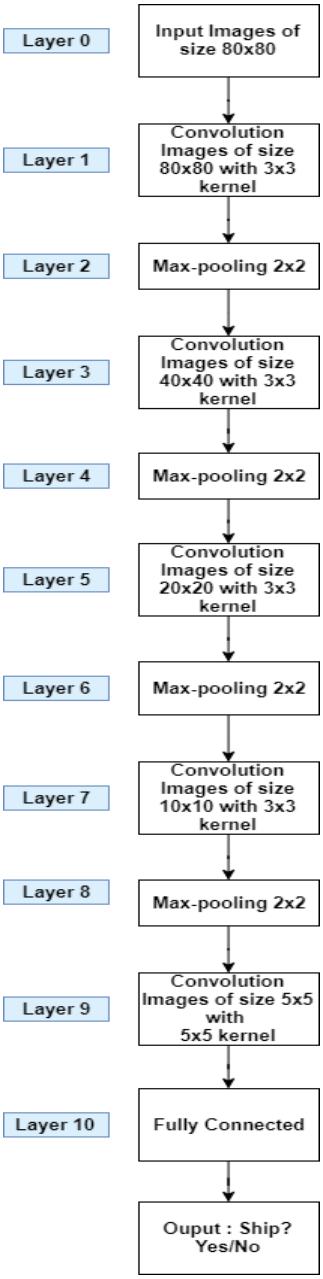


Figure 6. Layers of CNN

IV. RESULT

A. SVM

After organizing the training and test data, we read images from the disc and extract color features, histogram features, and HOG features then package them all in cells using a wrapper function. Our program takes 3.41 seconds to train the classifier. After training the classifier, we test it with validation data that has been previously split and we get an accuracy of 0.9438.

To find ships in the scene images, we use a sliding window. A sliding window approach has been implemented, where overlapping tiles in each test image are classified as ship or non-ship. In the detection process we get duplicated results as in Figure 7. After this we remove duplicate detection and false-positive as in Figure 8. The saved detection is a window with the highest detection value.

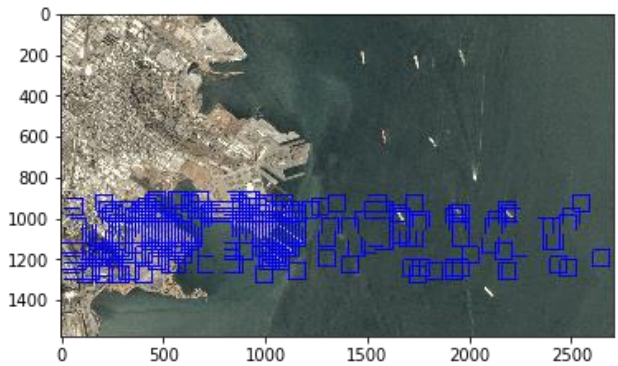


Figure 7. Duplicated detection

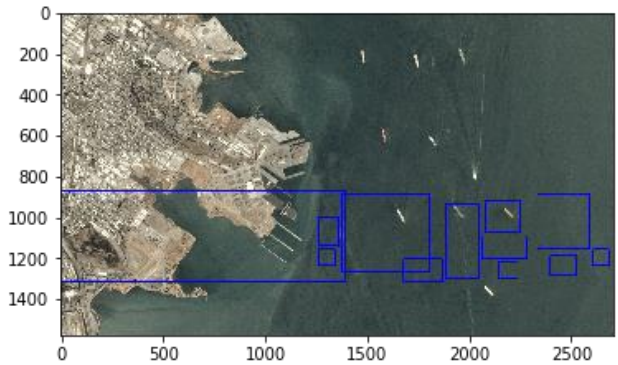


Figure 8. Final detection

B. CNN

CNN gets better results than SVM does when applied to this data set. The CNN was trained with train/val split of 3.200/800, an image size of 80 x 80, 18 epochs. Loss and accuracy are calculated both in training and validation at each epoch as in table 1.

TABEL I. ACCURATION RATE

Epoch	Training		Validation	
	Loss	Accuracy	Loss	Accuracy
1	0.0388	0.9862	0.0778	0.9800
2	0.0342	0.9897	0.667	0.9800
3	0.0265	0.9903	0.0644	0.9812
4	0.0322	0.9884	0.0642	0.9800
5	0.0354	0.9897	0.0619	0.9825
6	0.0544	0.9816	0.0521	0.9850
7	0.0287	0.9903	0.0705	0.9787
8	0.0265	0.9919	0.0656	0.9800
9	0.0262	0.9925	0.0822	0.9750
10	0.0257	0.9897	0.0634	0.9775
11	0.0165	0.9937	0.0694	0.9762
12	0.0211	0.9919	0.0635	0.9837
13	0.0158	0.9947	0.0720	0.9812
14	0.0158	0.9950	0.0655	0.9850
15	0.0202	0.9934	0.0799	0.9812
16	0.0202	0.9916	0.1077	0.9663
17	0.0166	0.9947	0.0657	0.9812
18	0.0114	0.9962	0.0645	0.9825

TABEL II. TIME TO GO THROUGH EACH EPOCH

Epoch	Time (Second)	Cumulative time (Second)
1	24	24
2	24	48
3	29	77
4	29	106
5	29	135
6	29	164
7	29	193
8	29	222
9	29	251
10	29	280
11	29	309
12	29	338
13	30	368
14	29	397
15	29	426
16	29	455
17	29	484
18	29	513

In table 2, it can be seen the time needed for each epoch to train the classifier. The total time needed to train the classifier in 18 epochs is 513 seconds (6 minutes 33 seconds). So CNN is longer than SVM does, which only requires 3.41 seconds to train the classifier.

From the plot in Figure 9, it can be seen that the loss value in training and validation is quite small with an average of 0,0259 for training and 0,1032 for validation. The loss has decreased in almost every epoch for training, but in validation loss value was high at epoch 16 and backed down at epoch 17 and 18. In general, the more the number of epochs that are run, the smaller the loss value obtained. In Figure 10, this model has high accuracy, the average training accuracy is 0.9912 or 99.12% and validation accuracy is 0.9798 or 97.89%.

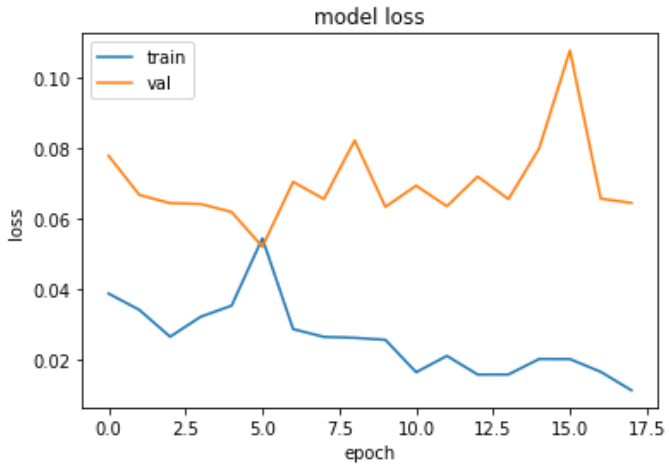


Figure 9. Loss rate

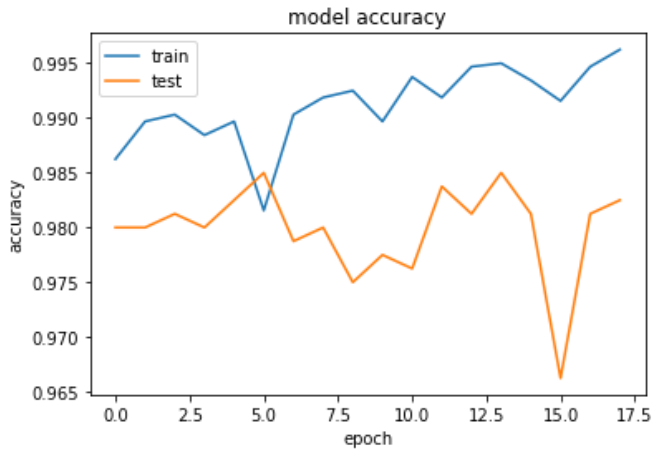
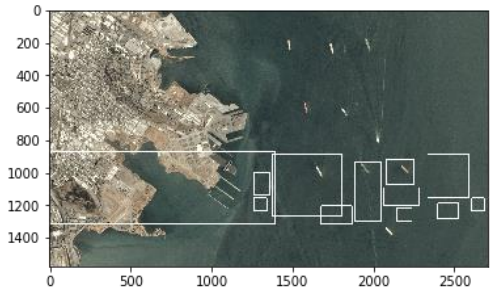
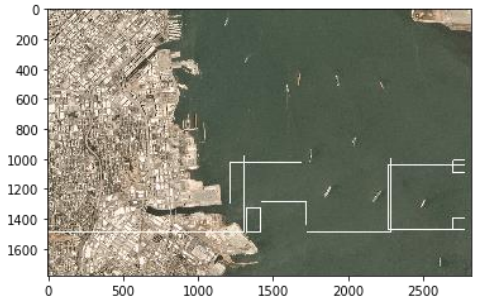


Figure 10. Accuracy rate

C. Test Image

The comparation of test result using high resolution satellite images as bellow:

Method used	Test Images
SVM	
CNN	
SVM	
CNN	

As can be seen above, CNN can detect ship objects better than SVM does when testing port views from high-resolution satellite images. Almost all ship objects in the picture can be detected properly by CNN. Whereas in SVM, there are many error detection.

V. Conclusion and Future Work

Detection of ships in satellite imagery successfully uses machine learning and computer vision algorithms. By comparing SVM and CNN on this data, it can be seen that CNN has higher accuracy and is considered better in detecting ship objects. CNN has more steps so that the time needed to run it is longer than SVM does. We hope that in the future, the detection of ship objects in satellite imagery will be better. The detection process will be faster in the future with a better GPU and CPU, or a more efficient algorithm.

VI. References

[1] G. Sun *et al.*, "Combinational shadow index for building shadow extraction in urban areas6 from Sentine 2A MSI imagery," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 78, no. January, pp. 53–65, 2019.

[2] B. Ganapthy, A. Kavsik, C. Palaniappan, M. Saravanan, and P. S. Prakash, "Detection of Ship Using DNN and ELM," vol. 6, no. 4, pp. 3415–3420, 2016.

[3] C. Zhu, H. Zhou, R. Wang, and J. Guo, "A novel hierarchical method of ship detection from spaceborne optical image basedon shape and texture-features," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 9, pp. 3446–3456, 2010.

[4] H. Amir, S. Gatot, and W. Wisnu, "Klasifikasi Objek Dalam Visi Komputer dengan Analisis Diskriminan," *Makara, Teknol.*, vol. 6, no. 1, pp. 24–32, 2002.

[5] M. Yang and G. Thung, "Classification of Trash for Recyclability Status," pp. 1–6, 2016.

[6] A. Salmador, J. P. Cid, and I. R. Novelle, "Intelligent Garbage Classifier," *Int. J. Interact. Multimed. Artif. Intell.*, vol. 1, no. 1, pp. 31–36, 2008.

[7] M. Ghuge and M. E. Student, "AUTOMATIC WASTE SORTING BASED ON," vol. 7, no. 6, pp. 593–595, 2018.

[8] A. Pon, S. Kumar, and D. Shankar, "Analysis of Invariant Feature Based Multistage Classifier for Compressed-Domain Ship Detection and Recognition," *Int. J. Comput. Sci. Mob. Comput.*, vol. 44, no. 4, pp. 607–612, 2015.

[9] S. G. Paulraj, S. Hait, and A. Thakur, "Automated municipal solid waste sorting for recycling using a mobile manipulator," *Proc. ASME Des. Eng. Tech. Conf.*, vol. 5A-2016, pp. 1–10, 2016.

[10] C. Kyrkou and T. Theocharides, "SCoPE: Towards a systolic array for SVM object detection," *IEEE Embed. Syst. Lett.*, vol. 1, no. 2, pp. 46–49, 2009.

[11] J. Zhang and C. H. Chen, "Moving objects detection and segmentation in dynamic video backgrounds," *2007 IEEE Conf. Technol. Homel. Secur. Enhancing Crit. Infrastruct. Dependability*, pp. 64–69, 2007.

[12] A. Shakeel, W. Sultani, and M. Ali, "Deep built structure counting in satellite imagery using attention based re-weighting," *ISPRS J. Photogramm. Remote Sens.*, vol. 151, no. October 2018, pp. 313–321, 2019.

[13] C. Corbane, F. Marre, and M. Petit, "Using SPOT-5 HRG data in panchromatic mode for operational detection of small ships in tropical area," *Sensors*, 2008.

[14] S. Wang *et al.*, "DEM generation from Worldview-2 stereo imagery and vertical accuracy assessment for its application in active tectonics," *Geomorphology*, vol. 336, pp.107–118, 2019.

[15] Y. Shendryk, Y. Rist, C. Ticehurst, and P. Thorburn, "Deep learning for multi-modal classification of cloud, shadow and land cover scenes in PlanetScope and Sentinel-2 imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 157, no. June 2018, pp. 124–136, 2019.

[16] K. Ahmad *et al.*, "Automatic detection of passable roads after floods in remote sensed and social media data," *Signal Process. Image Commun.*, vol. 74, no. December 2018, pp. 110–118, 2019.

[17] F. Calleja, B. Ondiviela, C. Galván, M. Recio, and J. A. Juanes, "Mapping estuarine vegetation using satellite imagery: The case of the invasive species Baccharis halimifolia at a Natura 2000 site," *Cont. Shelf Res.*, vol. 174, pp. 35–47, 2019.

[18] N. Proia and V. Pagé, "Characterization of a bayesian ship detection method in optical satellite images," *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 226–230, 2010.

[19] M. Akinkunmi, *Introduction to Statistics Using R*, vol. 11, no. 4. 2019.

[20] A. Ahmad, "Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning," *J. Teknol. Indones.*, no. October, p. 3, 2017.

[21] P. Ding, Y. Zhang, W. J. Deng, P. Jia, and A. Kuijper, "A light and faster regional convolutional neural network for object detection in optical remote sensing images," *ISPRS J. Photogramm. Remote Sens.*, vol. 141, no. June 2017, pp. 208–218, 2018.

[22] Z. Deng, H. Sun, S. Zhou, J. Zhao, L. Lei, and H. Zou, "Multi-scale object detection in remote sensing imagery with convolutional neural networks," *ISPRS J. Photogramm. Remote Sens.*, vol. 145, pp. 3–22, 2018.

[23] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Neural Information Processing Systems (NIPS)*, 2012.

[24] [Online]. Available: https://scikit-image.org/