Smart Ship Detection using Transfer Learning with ResNet

Richa[1], Richa Joshi[2]

[1,2]*Student, Banasthali Vidyapith, Rajasthan, 304022, India*

---***---

Abstract - *Satellite image processing is one of the most growing fields in research today. There is so much to explore and so many ways to do it that it seems full of endless opportunities and possibilities. There are several features which can be extracted like buildings, roads etc. from land satellite images and ships, boats etc. from satellite images of sea and ocean. In this paper we will be concentrating on detecting ships automatically from the images obtained by various satellites. This is one of the major challenging tasks due to various disturbances and noises in these kinds of images. Ships can be found in different sizes as well as shapes which make it more difficult to find a pattern or some regularity in these images. It is comparatively easier in homogeneous environment consisting of just ships of different types in water. But when it comes to heterogeneous environment consisting of other elements like coasts, harbour, vessel, rocks, islands etc. the challenge increases tenfold. There are various statistical and image processing approaches which can do this manually then again this won't be that efficient, changing the parameters again and again with different images and all can be a time consuming and tedious process. That's why we choose one of the modern approaches which provide us with the opportunity of being automatic or at least semi-automatic, Deep learning. Deep learning together with computer vision opens doors to the possibilities which we couldn't even have thought of. In this paper we will be exploring those possibilities with the help of a certain convolutional neural network known as ResNet and implementing it using transfer learning. This algorithm gives very high accuracy in classification of satellite images without doing any manual extraction and works with complex heterogeneous backgrounds too.*

Key Words: Convolution Neural Network, ResNet, Transfer Learning, Deep learning, Neural Network, Remote imagery, Satellite images.

1. INTRODUCTION

Images have been a way to share information from quite a time. For humans it is natural to identify elements of an image, distinguish between them but the same can't be said in case of computers as they understand them in the form of a number matrix and hence can't process it as humans. This is why image processing came to light so that computers can work with images too. Many of the techniques we know today of digital image processing were developed around 1960s Massachusetts Institute of Technology, Bell Laboratories, University of Maryland, and a few other research facilities, with application to satellite imagery, medical imaging, character recognition, and photograph enhancement [1].

Now came the challenge of extracting information from images like different features so that even computers can make sense of images and understand what they depict not just humans and just like humans. This is how computer vision came into existence, beginning at universities that were pioneering artificial intelligence [2]. Deep learning is the alternative approach, since it learns the feature by itself and predicts the features based on its application. Using an artificial neural network such as CNN gives an entirely new look and direction to this research field. It makes such a complex classification problem easier as it is able to learn features from the given training set of images.

Detecting ships in high resolution satellite images has number of applications in the arena of maritime surveillance, military, safe and effective transportation in the harbours, illegal fishing etc. [2]. Various research work has been done and is still continuing in this field from the 70s-80s. Several methods have been proposed in this field so that the detection accuracy increases. Even different hybrid methods are proposed but they are not consistent with extreme weather as it may lead to false detection. These state of the art methods were though not optimal in several cases but they were robust up to a certain extent. This is why deep learning was considered. CNN models can be trained with hundreds of examples with all the worst case scenarios and hence are able to give result with great accuracy. The neural network in CNN gives excellent result in object detection which is our prime goal and here the object is a ship. There are two major sections of these layers first for extracting features and the second one for classifying them accordingly. It also contained a hidden and an output layer.

Today remote sensing plays a vital role in detection of objects like ships from various images and hence helps in locating the position of ships. Satellite image processing has various applications, some of them being surveillance of traffic on harbours during busy hours, monitoring of coasts and harbours for smuggling activities, rescuing various sea machines like vessels and ships, military operations, pollution control in oceans, monitoring of oil discharge and sea pollution, naval and maritime welfare etc.

For extraction of ship all they need is synthetic aparture radar (SAR) images or panchromatic images as they are of high resolution. Monitoring of ships from satellite images helps in timely and periodic checking for anonymous movements in their territory region. Though in real time application it is much more complex and computation should be done considering this real time factor while detection and recognition. Remote sensing plays a very important role in monitoring ships as it operates from some distance and has a wide monitoring range. At the same time the sea surface gives better valid information than the appearance of ship.

1.1 OBJECTIVE:

The main objective of this paper on ship detection using satellite images is to detect ships with high accuracy. Since ship detection methods are used in extremely sensitive matters needing very precise result, accuracy is one of the most crucial factors to consider. With the help of ResNet and transfer learning technique we tried to achieve best possible and accurate results.

1.2 THEORETICAL BACKGROUND:

Convolutional neural network is one of the main neural nets used for images classification and images recognition. The main areas where CNN is widely used are object detection, face recognition etc. In CNN image classification, an image is taken as an input, processed and then classified under certain pre decided categories. CNNs are particularly used in finding patterns in images to recognize faces, objects etc. These nets learn directly from images as it uses pattern to classify data and hence eliminates the use of feature extraction manually [3].

Transfer learning is a machine learning approach where a model developed and trained for a certain task is reused as a starting point for another task [4]. It is a popular technique used in various computer vision and natural language processing tasks. It allows for a model to attain improved accuracy and efficiency for a particular use case. Here, we use transfer learning on satellite imagery by using a pre-trained Convolutional Neural Network, ResNet-50.

ResNet is nothing but a deep residual network that was introduced as an effort to not only increases the performance efficiency of a network but to also make it less prone to overfitting. It was established, as a part of the common trend of the research community that network architectures need to go deeper, in the year 2015 [5].

It is to be noted, however, that increasing the layers of a network should not be done by simply stacking the layers on top of one another since it may lead to the problem of vanishing gradients in deep networks. The problem is, as the gradient is back-propagated to earlier layers, repeated multiplication may make the gradient infinitively small. As a result of this as the network goes deeper the performance may get easily saturated or even degrade rapidly.

ResNet essentially brings up the idea of an "identity shortcut" such that one or more layers are skipped as depicted in Figure 1.
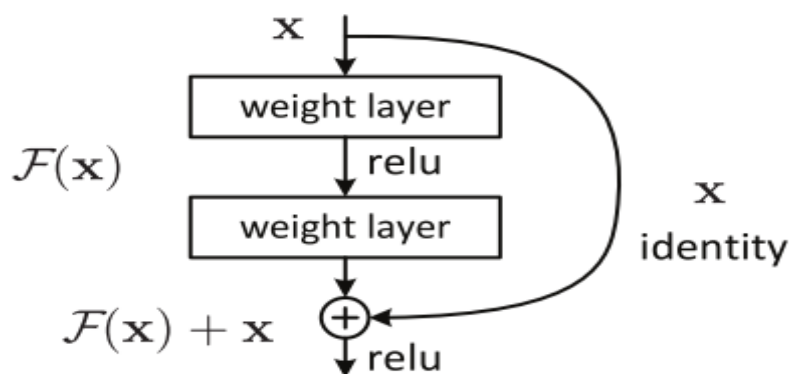

Figure 1: Shows a residual block of ResNet that helps create an identity shortcut connection, without giving rise to the vanishing gradient problem.

ResNet argues over the fact that stacking of layers should not decrease the performance efficiency of a deep network as one could simply stack identity mappings (a layer that does not perform any function) upon the current network and it would still perform the same. This makes way for the argument that fitting the stacked layers in a residual mapping is

easier than directly fitting them in a network. A residual block fundamentally controls how much information should be allowed to flow across the shortcut [6].

ResNet allows its multiple layer implementations in the form of ResNet-18, ResNet-34, ResNet-50, and ResNet-101 and so on. The ResNet-50 is a 50-layer implementation of the ResNet model which can be used in various training problems and transfer learning applications [10].

Fastai is a library built upon the platform of Pytorch that is used for simplification of training neural networks in a fast and accurate manner [7]. It is used in various applications for computer vision tasks, natural language processing tasks, collaborative filtering tasks etc. In our case we use the library to access our pre-trained model for transfer learning. The ResNet-50 model used here for transfer learning has been pre-trained on the ImageNet dataset [8]. Under this approach we use the Ships in Satellite Imagery from the Planet Satellite Imagery dataset [9]. Within this dataset each image is labelled as containing a ship or not containing a ship.

2. PRACTICAL APPROACH AND RESULTS

We start by importing the required library and then extracting the required dataset. Once the images are obtained, an ImageDataBunch is created. An ImageDataBunch is nothing but an object used to specifically handle Images for data augmentation [7]. Once the ImageDataBunch is created the augmented data can be viewed, it is shown as in Figure 2.
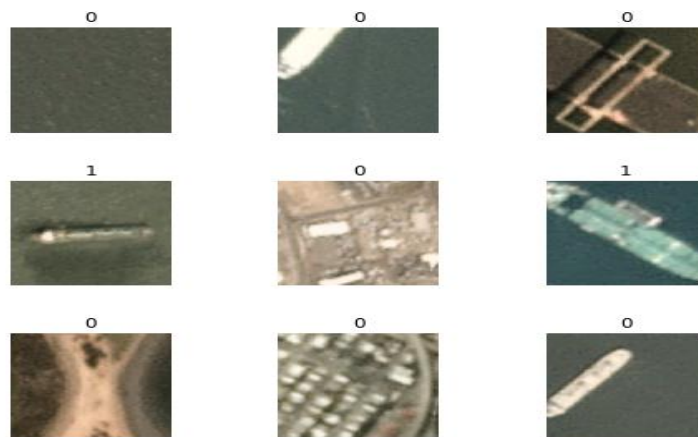

Figure 2: Shows an augmented form of the dataset required for the model's training along with its label where 0 identifies as not containing a ship and 1 as containing a ship.

Here each class has been labelled as 0 for not having a ship and 1 for having a ship. After the data has been augmented as per the requirement, the pre-trained model is downloaded. Consequently, an optimum learning rate is found for training post which the first training cycle is commenced. Thus, the pre-trained model is now trained on our dataset. The following results are obtained in the first training cycle as shown in Figure 3.

epoch	train_loss	valid_loss	error_rate	time
0	0.029670	0.040743	0.016250	01:06
1	0.032692	0.050607	0.015000	01:06
2	0.036503	0.039161	0.011250	01:07
3	0.029998	0.033349	0.013750	01:07
4	0.031400	0.035978	0.010000	01:07

Figure 3: The first training cycle with a final error rate of 1% in training

As it is, we obtain a very well trained model with an error rate of close to 1%. However, to further improve the performance we unfreeze the layers of the trained model and re-train them for a second training cycle. The final results are obtained as shown in Figure 4.

epoch	train_loss	valid_loss	error_rate	time
0	0.007969	0.015180	0.005000	01:26
1	0.006488	0.012538	0.002500	01:27
2	0.029535	0.051979	0.016250	01:27
3	0.024623	0.010481	0.001250	01:27
4	0.013846	0.007747	0.002500	01:27
5	0.009254	0.008634	0.002500	01:26
6	0.007517	0.013707	0.005000	01:27
7	0.004527	0.011681	0.005000	01:25

Figure 4: The second training cycle with a final error rate of 0.5% in training

We now have a reduced error rate of 0.5% and an accuracy close to 99.5% which makes the model highly efficient. As a result to check where the model is lacking we have also plotted the confusion matrix for the second and the final training cycle as shown in Figure 5.
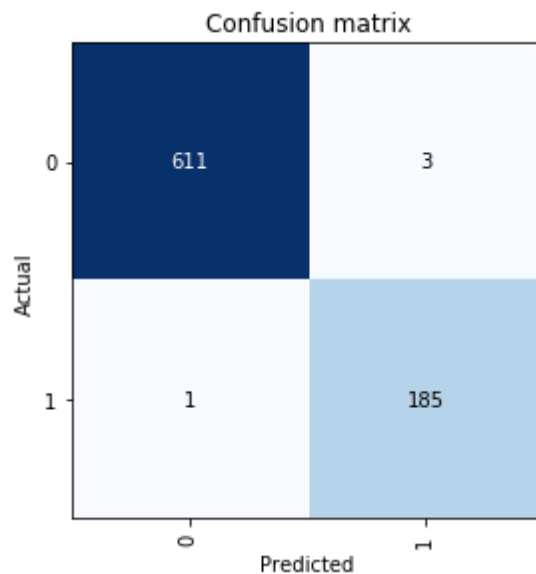

Figure 5: The confusion matrix for the final training cycle.

As is clear from the confusion matrix the model successfully predicts 611 images as not containing ships while 3 images are wrongly predicted as not containing ships. On the contrary, the model successfully predicts 185 images as containing ships while 1 image is wrongly predicted as not containing a ship. We therefore have a very accurate model now; we do not train it further to avoid its over-fitting on our dataset.

3. CONCLUSION

Various hybrid approaches can be developed which can increase the accuracy of ship detection in satellite images. The method proposed by us in this paper outperforms several state of the art approaches when it comes to accuracy and other essential factors. This has been obtained due to the usage of a pre trained model under the transfer learning approach and latest libraries. The experimental results on high resolution satellite images explain that the proposed method is not only efficient in terms of its computation and accuracy but also has a vast scope for further research.

REFERENCES

[1]Azriel Rosenfeld, Picture Processing by Computer, New York: Academic Press, 1969

[2] J. S. Iwin Thanakumar, R. S. Benson Edwin, J. Sasikala and J. D. Sujitha, "Smart Vessel Detection using Deep Convolutional Neural Network," 2018 Fifth HCT Information Technology Trends (ITT), Dubai, United Arab Emirates, 2018, pp. 28-32.

[3] https://www.mathworks.com/discovery/convolutional-neural-network.html

[4] A Gentle Introduction to Transfer Learning for Deep Learning by Jason Brownlee on December 20, 2017.

[5] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. arXiv preprint arXiv:1512.03385,2015.

[6] An Overview of ResNet and its Variants by Vincent Fung.

[7] Welcome to Fastai: https://docs.fast.ai/

[8] The ImageNet Dataset: http://www.image-net.org/

[9] Ships in Satellite Imagery Classify ships in San Francisco Bay using Planet satellite imagery https://www.kaggle.com/rhammell/ships-in-satellite-imagery

[10] Understanding and Implementing Architectures of ResNet and ResNeXt for state-of-the-art Image Classification: From Microsoft to Facebook [Part 1] by Prakash Jay.