

Applicability Analysis Of VGGNET Based Transfer Learning For Oil Spill Classification On SAR Data

¹Mr. Naishadh Mehta ²Prof. Pooja Shah , ³Dr. Vijay Ukani , ⁴Mr. Aayush Gohil

¹Institute of Technology
Nirma University
Ahmedabad, India
19mced07@nirmauni.ac.in

²Institute of Technology
Nirma University
Ahmedabad, India
poojashah@nirmauni.ac.in

³Institute of Technology
Nirma University
Ahmedabad, India
vijay.ukani@nirmauni.ac.in

⁴Institute of Technology
Nirma University
Ahmedabad, India
19bce069@nirmauni.ac.in

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 16 April 2021

Abstract : Oil Spill detection serves a significant application of marine pollution monitoring. Timely rescue in case of ocean oil spills due to accidents or vessel discharges is very important for the marine ecosystem. This paper discusses the experiment done in this line using VGGNet based transfer learning. The input to the proposed system is decomposed polarimetric Synthetic Aperture Radar data. The decomposition is done through ESA PolSARpro 5.1. The paper discusses the analysis of empirical study of transfer learning based VGGNet for classification of oil spill on ocean surface.

Keywords—SAR; Oil Spill; Transfer Learning; Classification; VGGNet

I. INTRODUCTION

Oil spills on the ocean surface are a serious threat to marine ecosystem. Automation of oil spill detection through Synthetic Aperture Radar images is considered to be a good aid for oil spill disaster management. Marine oil slicks cause serious waste of energy and deeply damage the marine ecosystem. The marine pollution of this kind occurs due to crude oil extraction and/or transportation and accidental or intentional discharge of oil. Also, with the developing ocean technologies, marine transportation has increased which has increased the possibility of accidents and illegal oil discharges. These accidents again lead to dangerous amount of oil slicks especially if a oil transporting vehicle is victimized. Half of the total oil spills in the marine environment comes from operative discharges by shipping and in most of these cases the discharges are illegal. These illegal discharges are not just by the oil tankers, there are many classes of ship suspected of being responsible [7]. It has been estimated that 457,000 tons of oil are released by shipping into the ocean every year [7]. This fact leads to serious concerns about contingency planning, mitigation and remediation in regards to protect marine ecosystem form toxic oils. Developing a cost-effective oil spill detection system has been the subject of research since past two decades.

The development of all weather, day and night Synthetic Aperture Radar (SAR) sensors has made it possible to remotely locate oil spill disasters with the high-resolution image offerings. The oil spill detection has improved with the launch of polarimetric SAR missions. SAR oil slick detection is possible due to the damping effect of oil slick over the ocean surface. The Bragg scattering produce dark spots on SAR images. Thus, the first step of oil spill detection from SAR image is dark spot detection. However, there are several other phenomena beyond oil slicks that introduce dark formations in SAR images which appear to be quite similar to oil spills. These dark formations are called look alikes and creates hurdles in correctly locating slicks as they appear to be false positives. The methods exploited towards classification of oil spill and look alikes can be broadly segregated

into two categories. First methods that use specific classification feature and the methods in second category work with multi-feature fusion. Considering a single feature limits the accuracy [3]. The texture features of single polarimetric SAR images cannot describe the physical characteristics of the targets completely. This may cause misjudgement during the oil slick detection [5].

Table 1: Statistics of Classification

| | Oil spills | Lookalikes | Total |
|-----------------------------|-------------------|-------------------|--------------|
| Predicted oil spills | 598 | 62 | 660 |
| Predicted lookalikes | 27 | 413 | 440 |
| Total | 625 | 475 | 1100 |

In a few recent studies, look-alikes and oil spills were classified using decision trees while, dark spots have been segmented using hysteresis algorithms. Decision trees [1] with an information gain as parameter, is usually biased with features valued higher- introducing over-fitting. Neural networks [8] with an efficient distributed processing ability, a higher accuracy rate combined with a potential to fit the relationship and allow non-linearity to feature in between the input parameters and the output is a very advantageous paradigm. Feature classification and segmentation tasks in sequence are generally basis for the artificial neural networks [8]. Network training is considered a challenging task owing to the direct relation of the number of parameters with the features that are to be considered for the classification, quite losing the efficiency of the learning process undertaking, while further increasing the learning time. Convolutional neural networks are a multilayer neural network-based method for classifying images and recognizing objects [4] [9] which is a good paradigm for 2-D recognition and structure-based learning while learning features from the pixel-based recognition. The training improves the accuracy of recognition with a shorter time for training.

In this paper, the methodology presented uses VGGNet [6] for oil spill and look alike classification from quad polarimetric ALOS PALSAR data. With the fact that full polarimetric SAR data is costly and the specificity of the oil spill detection application the available data is significantly less to train a neural network, the benefits of transfer learning have been exploited. A pre-trained inception model from Google which was built on ImageNet data to identify images in those pictures has been used. With a data set consisting of limited target specific images, classification accuracy of 95% has been achieved. VGGNet [6] is an inductive transfer where the initial training is done on a very generic dataset that allows us to further fine tune the model to the corresponding processed SAR images in a way that the features learnt would be of the form of edges and spatiotemporal in nature. Section 3 covers the information related to data used in the study and what treatment the data has been given before it is given as input to the classification phase especially in terms of polarimetric processing. The results and validation are discussed in section 4 followed by conclusion and acknowledgement. The next section describes the methodology along with basic details of transfer learning and also elaborates on how data augmentation is performed.

II. METHODOLOGY

For oil spill and look alike classification task, the available SAR data does not allow appropriate training. Approaching a neural network to small sets causes over-fitting. Augmenting the dataset further causes over-fitting of the model. Here, we use a transfer learning approach to get by the problem of over-fitting, yet using the augmentation as a further help for the learning model. A basic assumption for taking the approach is the information that is transfer amongst the domains is used to undertake the selection of the initial parameters. This allows a more optimized and improved model that undergoes a substantial regularizing impact. Our experiments found that augmenting data increase the efficiency of the model owing to the reduction in irrelevant parameters being learnt. Hence, we undertake the augmentation in mini batches and increase the dataset size by a factor of 2, based on a probability of choosing different functions for the augmentation task.

Features obtained from transfer learning are better capable to introduce a sense of generalization by improving performance than just initializing the parameters randomly. The approach utilized brings together the images in their feature representations of the SAR and 2-D image formats. It also finds a line of communication between fine-tuning of various operations and their hard coding, which allows domain independence for our model. There is a good chance of similarity among the layer differentiated SAR images and the generic images undertaken for training. Hence during the fine-tuning training, we try to retain as many characteristics we can by sharing the lower dimensional representations with the SAR polarimetry domain.

The basic learning process is described in Figure 1 where the tasks share the type of representations. We define the input variables, output variables, the set of neurons. We define the hidden lower layers as shared amongst the neurons and task-specific upper layers. For an efficient representation in the polarimetry task, the input of the computer vision layer is used for the initialization of the lower layer. The shared layers are hardcoded, which mean no changes occur on their weights and biases. The upper layers are updated as and when they are fine-tuned where the bias tilts towards the polarimetry domain. This allows the implementation of transfer learning from the ImageNet dataset consisting of 10 million images using parameters already stored in JSON format.

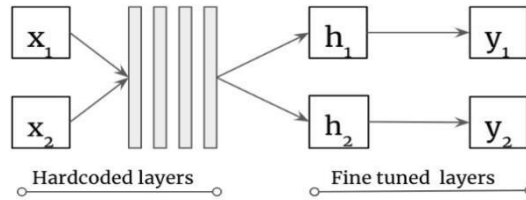


Figure 1: Approach using VGGNet

A convolutional neural network, trained over 14 million images-based model has been popular for transfer learning which is 16 layers deep called VGGNet [8]. The architecture is described in Figure 2. In total there are 13 convolutional, 4 pooling layers which occur behind the group of convolutional layers, 2 fully connected layers and at the bottom a prediction layer. Each convolutional layer further houses batch normalization for an appropriate reconstruction of the architecture to reduce the gradient vanishing that occurs during the training process. The trained VGGNet provides the initial input weights from its npy file. 28 x 28 size images are created from our curated dataset. As described earlier, the fine tuning is undertaken on the labelled dataset. The initial layers have been hardcoded while the higher layers are flexibly re-trainable.

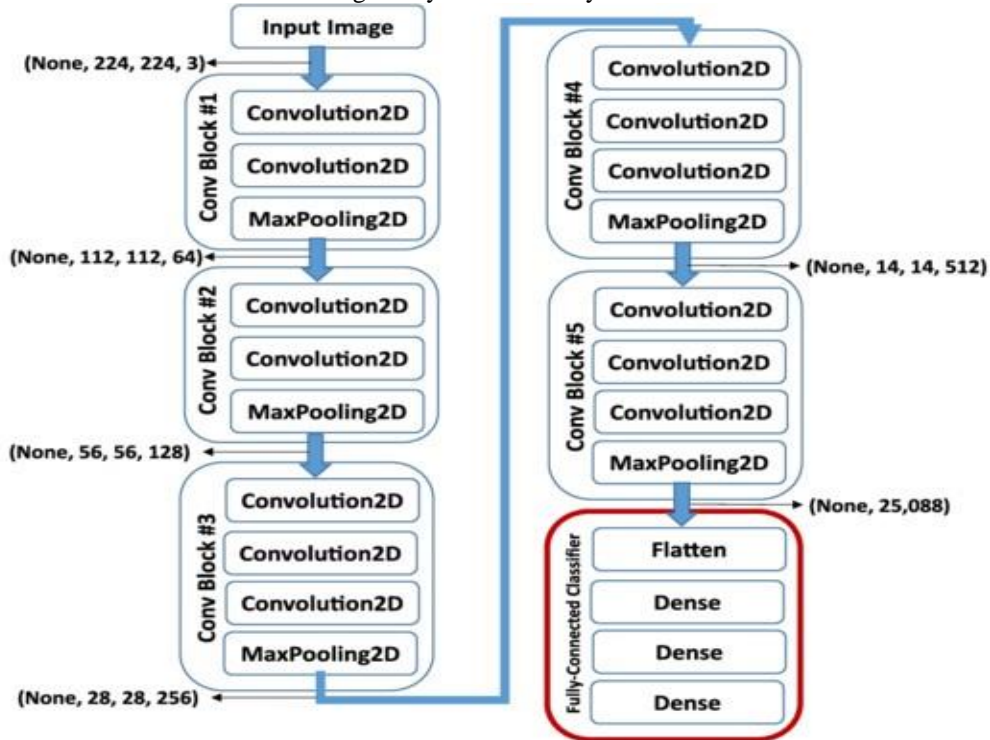


Figure 2: VGGNet Layer Architecture

A. Data Augmentation

We use various augmentation approaches to get sufficient training examples for the neural network with a random selection of each to generate 1000 images of each. We use the filter, flip, extract, noise addition methods for the augmentation tasks. The images and the sub images from the other approaches can be rotated, flipped, mirrored for training set generation. Various oil spills and look alikes exist in the images which are manually cropped out for training the images. Figure 3 depicts the process undertaken for a sample image. Data augmentation process undertaken during the research is depicted in Algorithm 2.

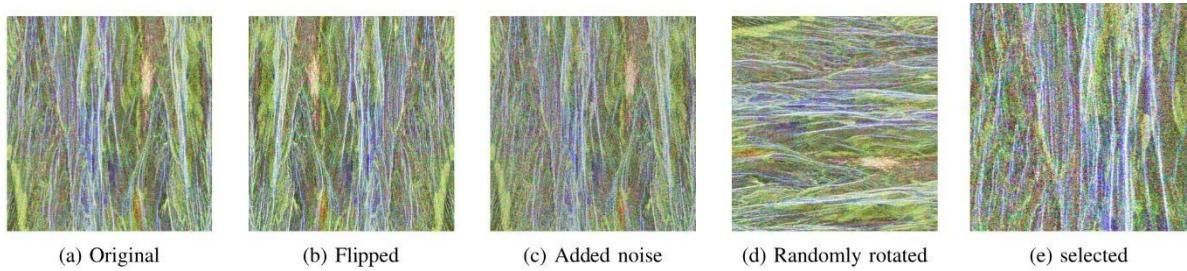


Figure 3: Data Augmentation

Procedure augment (seed, image): Switch seed

```
do: {
    when 1: flip by 180_o
    when 2: flip by 90_o
    when 3: crop image with random size.
    when 4: add noise to the image.
}
add new_image to db
```

Algorithm 1: Method for Augmenting Images

```
While flag is True:
do {
    seed= random_generation (1 to 4)}
    new_image=augment (image, seed)
    flag=random_generation (True, False)
}
```

Algorithm 2: Flow of Random Augmentation

III. DATA PROCESSING

The SAR data used in this study is captured by ALOS PALSAR sensor and is freely available through Alaska Satellite Facility's (ASF) online portal "vertex". PALSAR was an L-band SAR mission by JAXA operational from 2006 to 2011 with repeat coverage of 48 days and a global spatial extent. From the available variants the data products used in this study are L1.5 quad polarimetric data from various ocean surface regions with probable oil slicks and other oceanic features.

A. Data Preparation:

Oil slicks appear much darker than the surrounding region in SAR imagery. For detection it is important to study these darker regions in further detail. Thus, the first step in oil spill detection is dark spot detection. The quad-polarimetric SAR data used in this study is susceptible to noise. Thus, the decomposition chosen here for data preparation is Pauli decomposition as it is anti-interference and has high adaptability in general. The Pauli decomposition imagery is clearer than the original SAR image. This benefits the dark spot detection process. The Pauli decomposition expresses the measured scattering matrix [S] in the so-called Pauli basis. If we considered the conventional orthogonal linear (h, v) basis, in a general case, the Pauli basis is given by the following four 2 X 2 matrices.

$$[S_a] = 1/\sqrt{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{1}$$

$$[S_b] = 1/\sqrt{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2}$$

$$[S_c] = 1/\sqrt{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{3}$$

$$[S_d] = 1/\sqrt{2} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \tag{4}$$

$$Shv = Sv h \tag{5}$$

Which in turn reduce the pauli basis to

$$\{[S_a], [S_b], [S_c]\} \quad (6)$$

and the given scattering matrix can be expressed as

$$[S]=\begin{bmatrix} S^{hh} & S^{hv} \\ S^{vh} & S^{vv} \end{bmatrix} = \alpha [S_a] + \beta [S_b] + \gamma [S_c] \quad (7)$$

Where

$$\alpha = S_{hh} + S_{vv} / \sqrt{2} \quad (8)$$

$$\beta = S_{hh} - S_{vv} / \sqrt{2} \quad (9)$$

$$\gamma = \sqrt{2} S_{hv} \quad (10)$$

The interpretation of the decomposition is based on the basis and the coefficient $[S_a]$ is referred to single bounce (odd bounce) scattering. The intensity of the coefficient α determines the power scattered by targets characterized by single bounce. $[S_b]$ represents double bounce (even bounce) scattering. $[S_c]$ represents scattering from the scatterers which are able to return orthogonal polarization.

The data decomposition is performed through the free and open-source software PolSARpro made available by European Space Agency (ESA). The RGB channels for creating SAR image from Pauli decomposition are

$$|\alpha|^2 \rightarrow \text{Red} \quad (11) \quad |\beta|^2 \rightarrow \text{Green} \quad (12) \quad |\gamma|^2 \rightarrow \text{Blue} \quad (13)$$

IV. RESULTS AND VALIDATION

Our model was validated using the SAR data collected by the ALOS PALSAR sensor. TensorFlow has been used as a deep learning framework on a Macintosh macOS Mojave 2GHz Intel Core i5 system. Quantifications are undertaken based on precision, recall for the understanding of the system's accuracy.

Precision is the ratio of the true positive observations and the total positive observations. Recall is the ratio of the true positive observations and true actual positive observations. The training data possesses uneven class distribution hence, we introduce F1- scores defined as the harmonic mean of the precision and recall.

$$\text{Precision} = \text{True Positive} / \text{True Positive} + \text{False Positive} \quad (14)$$

$$\text{Recall} = \text{True Positive} / \text{True Positive} + \text{False Negative} \quad (15)$$

$$\text{F1-Score} = 2 * \text{Precision} * \text{Recall} / \text{Precision} + \text{Recall} \quad (16)$$

Hence, the Precision here is **0.9568**, Recall is **0.9966** and a F1- Score of **0.97627**.

We investigate the effects of training size on performance using our augmented dataset in Table 2. We identified the accuracy and training times for the different sets of training data with variant sizes. We can see the time taken for training and the size of the freezed model increase drastically, but the accuracy increases as well. Hence a trade-off between the two can be specified and undertaken.

Table 2: Variance in the size of training set fed to the VGGNet

| Size of training set | 10% | 20% | 50% |
|---------------------------------|--------|---------|---------|
| Time taken for training | 53153s | 134244s | 413422s |
| Accuracy attained while testing | 93.12% | 96.2% | 98.9% |
| Size of freezed model | 12.6M | 19.3M | 49M |

The model trained is compared with other pixel-based techniques- random forest, decision trees, logistic regression and pixel-batch based neural network trained only on the SAR dataset. Each with a pre-trained 16 layered VGGNet implementations for a basis of comparison. Overall accuracies have been depicted in Table 3. We found out that the deep neural network-based techniques garnered better classification results and accuracies than other algorithms. We converge the reason for the phenomenon to the potential of fine tuning the deeper layers of the neural network while existing parameters already trained which allow a prior probability in the weights of the network.

Table 3: Accuracy comparison

| Methods | | Accuracy (%) |
|-------------------|--------------------------|--------------|
| Pixel based | Random Forest | 56.01 |
| | Logistic Regression | 45.53 |
| | Decision Trees | 47.75 |
| Pixel Batch based | Convnets | 65.53 |
| | VGGNet transfer learning | 98.7 |

The model predicts the oil spill and look alike by classifying them and providing us with a confidence value for the prediction.

V. CONCLUSION

With this research we showed the applicability of VGGNet for a model to learn various features to classify various images comprising of an object of interest in the frame and compare it with other pixel based and batch-based approaches. We discussed the VGGNet transfer learning and then the importance of various data set sizes for transfer learning on the polarimetric SAR data and achieved comparable results to existing methods.

REFERENCES

1. Fingas, Mervin. Oil spill science and technology. Gulf professional publishing, 2016.
2. Gopalakrishnan, Kasthurirangan, et al. "Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection." *Construction and Building Materials* 157 (2017): 322-330.
3. Guo, Hao, Danni Wu, and Jubai An. "Discrimination of oil slicks and lookalikes in polarimetric SAR images using CNN." *Sensors* 17.8 (2017): 1837.
4. Koushik, Jayanth. "Understanding convolutional neural networks." arXiv preprint arXiv:1605.09081 (2016).
5. Masjedi, Ali, Mohammad Javad Valadan Zoej, and Yasser Maghsoudi. "Classification of polarimetric SAR images based on modeling contextual information and using texture features." *IEEE Transactions on Geoscience and Remote Sensing* 54.2 (2015): 932-943.
6. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).
7. Singha, Suman, Tim J. Bellerby, and Olaf Trieschmann. "Detection and classification of oil spill and look-alike spots from SAR imagery using an artificial neural network." 2012 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2012.
8. Topouzelis, Konstantinos, et al. "Potentiality of feed-forward neural networks for classifying dark formations to oil spills and lookalikes." *Geocarto International* 24.3 (2009): 179-191.
9. Zeiler, Matthew D., and Rob Fergus. "Visualizing and understanding convolutional networks." European conference on computer vision. Springer, Cham, 2014.