Fine Tuning MobileNet Neural Networks for Oil Spill Detection

Caixia Wang
GeoComputing Lab
Department of Geomatics
University of Alaska Anchorage
Anchorage, AK, USA
cwang12@alaska.edu

Andrew Coulson
GeoComputing Lab
Department of Geomatics
University of Alaska Anchorage
Anchorage, AK, USA
avcoulson@alaska.edu

***Abstract*— The monitoring of open water and early identification of oil spills in the Alaska Arctic has become increasingly critical due to the rise in oil and gas exploration and shipping activities, facilitated by the increasing number of ice-free days resulting from global warming. This escalating risk of oil spills is further compounded by potential accidents in offshore operations, illicit oil discharges, and knowledge gaps in Arctic coastlines, rapidly changing due to rising seas, permafrost melting, and coastal erosion. To address these pressing challenges, we propose a deep learning model based on MobileNet neural networks to detect oil spills in remotely sensed images. Compared to traditional pattern recognition methods, the proposed model can learn from examples to map input new data into the design and automatically optimize the training objective without designing rules and specifying critical parameters to solve the inference task. The experiments demonstrated we were able to obtain an overall accuracy of 0.93 with our proposed methods.**

Keywords—oil spill, remote sensing, machine learning, deep learning, imagery

I. INTRODUCTION

A. Background

Oil is a significant threat to sea ecosystems among marine pollution types. In the Arctic, oil and gas exploration operations and shipping traffic have been snowballing, with more ice-free days present in Northern seas attributed to global warming [1]. The risk of oil spills escalates due to potential accidents in oil and gas offshore operations, possible illegal oil discharge, or nautical knowledge gaps along Arctic coastlines due to their rapid changes caused by rising seas, permafrost melting, and coastal erosion. The harmful impacts of oil spills can be pervasive and have long-term consequences on wildlife, habitats, and cultural and subsistence uses of these resources. And the region in the Arctic has very little capacity to respond, according to the US National Research Council (NRC) [2]. Responders need to travel long distances to reach the spill's locations. A lack of ships with appropriate cleanup equipment coupled with rapidly changing weather conditions also interferes with response efforts in a timely and efficient manner. Therefore, it has become increasingly critical to efficiently monitor open water in the Alaska Arctic and identify oil slicks to provide early alerts and valuable information for rapid responses and mitigation efforts.

Funded by the National Science Foundation (award #2119689) and ConocoPhillips Arctic Science and Engineering Endowment.

B. Related Work

Vessels or airplanes, especially if equipped with specialized sensors (such as IR/NIR sensors, UV sensors, and laser fluorosensors) [3]–[8], have been used to detect and map oil spills at sea. The use of unmanned aerial vehicles and autonomous underwater vehicles for oil detection and tracking has grown. However, they are limited to certain observational conditions. Fig. 1 shows a visible image of a small test slick, where the sun glint on the lower left obscures slick edges and puzzles the interpretation. Additionally, each mission can cover a very limited area. Satellite-carried Synthetic Aperture Radar (SAR) provides a unique option for broad area surveillance, independent of sunlight, cloud cover, and weather, and more cost-effective than air patrolling. As the brightness of the captured SAR image represents the properties of the target surface, the oil film decreases the reflection of microwave (radar) signals of the sea surface and appears as dark patches in SAR images compared to the surrounding clean sea areas.

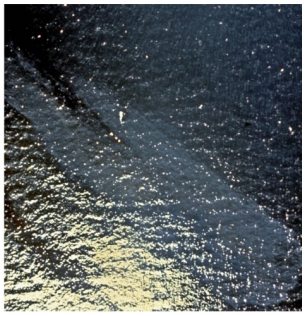


Fig. 1. A visible image of a small test slick. Note the sun glint on the image. [13]

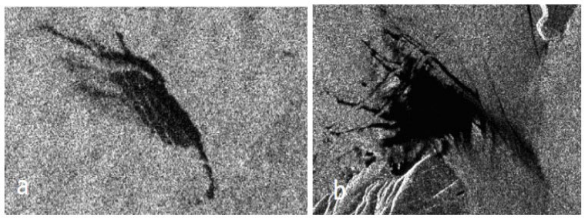


Fig. 2. (a) Verified oil spill on a SAR image; (b) Verified look-alike (not an oil spill) on a SAR image. [10]

While many works on oil slick segmentation and classification using SAR have been developed based on traditional pattern recognition [9], the main challenges remain. One issue is that fresh oil spills are brighter than older spills. They have a weak backscattering contrast relative to their surroundings and thus cannot be easily distinguished. Moreover, dark areas can have various contrast values, depending on local

sea state, oil spill types, image resolution, and incident angle. Third is that look-alike phenomena (e.g. Fig. 2b) can present as dark areas on a SAR image, such as wave shallows behind structures or topographic features, shallow seaweed beds, biogenic oils, and sea-life sperm [10].

II. Proposed Method

A. MobileNet Architecture

In this work, we introduce a deep learning-based architecture specifically designed for oil spill detection. Our method utilizes MobileNet [11], a powerful convolution neural network (CNN) architecture originally developed for object detection in mobile and embedded-based computer vision applications. MobileNet achieves its efficiency through the implementation of depth-wise separable convolutions, combining depth-wise convolutions and pointwise (1x1) convolutions.

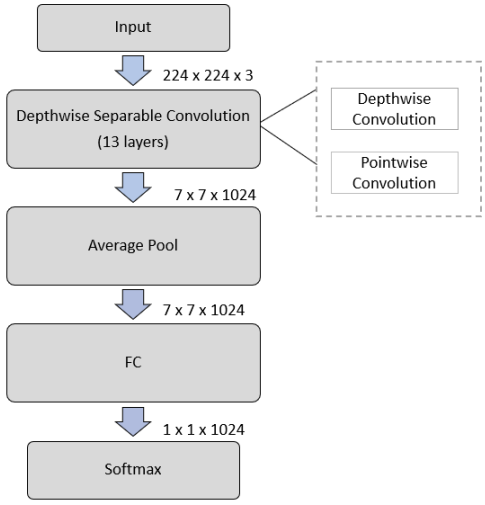


Fig. 3. The basic architecture of the MobileNet.

This unique combination significantly reduces the number of parameters compared to networks employing regular convolutions of the same depth. As such, MobileNet enables the creation of lightweight deep neural networks, offering a balance between optimal latency and reduced model size. Fig. 3 illustrates the basic representation of the model architecture.

In the context of a regular convolution, when given an input of shape $D_F \times D_F \times M$, where D_F represents the input width and height, and M is the number of input channels, and a filter shape of $D_K \times D_K \times N$ (with D_K as the width and height of the filter and N being the number of filters), the convolution process combines the input data with the filters to produce a new set of outputs of shape $D_G \times D_G \times N$, where D_G represents the output width and height, and N is the number of output channels in a single step. Consequently, the computation cost of performing this regular convolution is determined by the following:

$$D_K \cdot D_K \cdot M \cdot N \cdot D_F \cdot D_F \tag{1}$$

In contrast to regular convolutions, depth-wise separable convolution follows a two-step process when applied to the same input of shape $D_F \times D_F \times M$ (Fig. 4). First, it employs depth-wise convolution, where a filter of shape $D_F \times D_F \times 1$ is individually applied to each input channel, generating an output of shape $D_G \times D_G \times M$. The computation cost of this first step is $D_K \cdot D_K \cdot M \cdot N \cdot D_G \cdot D_G$. Next, depth-wise separable convolution involves point-wise convolution, where a filter of shape $1 \times 1 \times N$ is applied to the output obtained from the first step, resulting in a new output of shape $D_G \times D_G \times N$. The computation cost of this second step is $D_G \cdot D_G \cdot M \cdot N$. To determine the total computational cost of depth-wise separable convolution, we combine the costs from both steps:

$$\frac{M \cdot D_G^2 \cdot D_K^2 + N \cdot D_G^2 \cdot M}{D_K^2 \cdot M \cdot N \cdot D_G^2} = \frac{1}{N} + \frac{1}{D_K^2} \tag{2}$$

By prioritizing latency optimization while ensuring small network sizes, MobileNet empowers developers to select a network that aligns with the resource limitations and performance requirements of their specific applications. The use of MobileNet as the foundation for our oil spill detection architecture promises efficient and effective results, contributing to advancements in environmental monitoring and emergency response efforts.

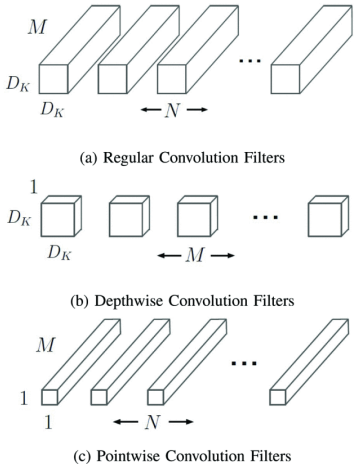


Fig. 4. Regular convolution vs. depth-wise separable convolution [11].

Although there is a trade-off between accuracies and network efficiency with respect of size and speed, MobileNet demonstrates a strong performance compared with other popular models on ImageNet classification. Additionally, it is well-suited for the task of detecting oil spills, particularly on computationally limited platforms where timely analysis is essential.

B. Training the Pre-Trained Model

MobileNets were originally trained for various recognition tasks such as object detection, fine grain classification, face attributes and large scale geolocalization [11]. The source domain MobileNets have been pretrained in differs from the target domain in this work, i.e., oil spill detections. Specifically, MobileNets have been trained using images from the ImageNet library, which houses thousands of images belong to 1000 different categories. Oil spill images were not included in the ImageNet Library from which MobileNets were originally

trained. Therefore, the model has not learned the features of oil spills. However, this already trained model has learned to understand edges, colors, textures and more objectively, features that could be components of oil spills. All these learned generic features are knowledge we could benefit from in the new model for classifying oil spills.

To solve a new but related task without training from scratch, transfer learning techniques [12] were adopted to use generic image features learned in pretrained CNNs. Fig. 5 illustrates proposed transfer learning structure for MobileNets. Like many deep learning systems and models, MobileNets are structured as layered architectures that learn various features at different layers. These layers are eventually connected to a fully connected layer (typically in supervised learning) to produce the final output. Initial lower layers of the network learn very generic features and the higher layers learn very task-specific features. This hierarchical architecture enables us to use the pre-trained network while excluding its final layer, effectively using it as a fixed feature extractor for the specific task at hand. Additionally, we have the flexibility to selectively retrain certain previous layers to fine-tune the model for the target task, thereby achieving better performance and adaptability.

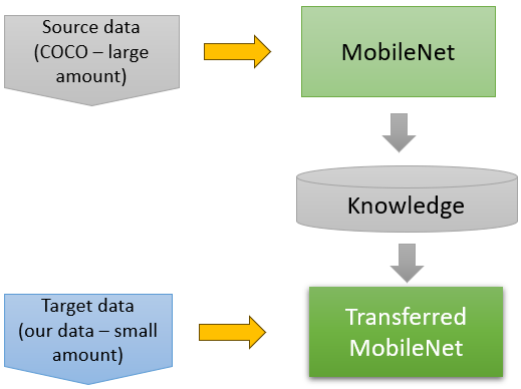


Fig. 5. Schematic of the transfer learning process.

In the proposed structure, we revised the pre-trained MobileNets model to predict two classes in the task, instead of 1000 categories in the original model. Specifically, we removed the last five layers of the network, including the average pooling, fully connected, and SoftMax layers. They are replaced by new output layers responsible for predicting two classes: oil spill or not oil spill. To ensure effectively training for our task, we decided to freeze all layers of the model, except for the last 22 layers. This means that the weights for these layers will remain unchanged during the training process with the new data. The rationale behind this choice is to prevent overfitting and preserve the learned representations from the pre-trained model. By retaining the relevant knowledge in these layers, we enhance the model's capacity to generalize to the new task. Our experimentation involved multiple tests, during which we evaluated different configuration of freezing layers. Through the evaluation, we observed that freezing the layers up to the last 22 yielded the most accurate prediction for our specific task. Compared to traditional pattern recognition methods, the deep learning-based model can learn from examples to map input data into the design output and automatically optimize the training objective without designing rules and specifying critical parameters to solve the inference task.

III. Experiments

A. Experimental Setup

This study employed the TensorFlow framework along with its Keras neural-network library to fine-tune and test the MobileNet model in the application of oil-spill prediction. The process was conducted on a Dell Precision 5820 Tower equipped with an Intel Xeon CPU, and two Nvidia RTX 6GB GPUs. Moreover, a broad of libraries was integrated, notably Fast.ai, PyTorch, and scikit-learn, to facilitate the experiments.

B. Datasets

The pre-trained MobileNet model was fine-tuned on a Kaggle dataset and additional data we manually annotated. This combined dataset comprises two classes, namely "oil spill" and "no oil spill," containing approximately 436 images. They are randomly sampled into 342 training images, 51 validation images, and 43 test images. In Table I, we provide a selection of training images utilized for fine-tuning the pre-trained MobileNets. These images are all in true-color format, featuring blue, green, and red channels, and have slightly varying sizes. The collection includes images representing ground-level views as well as satellite or air-borne perspectives.

C. Results and Discussion

Preprocessing is essential to ensure that images are in the required format for MobileNet architecture. This involves rescaling all images to 224×224 and converting the pixel values from the range [0, 255] to [-1, 1]. During the fine-tuning process, we set the learning rate to 0.000. Each epoch comprises 69 steps, with a batch size of 32. The resulting MobileNet model contains approximately 3 million parameters, resulting in a compact size of 12 megabytes (MB). This is notably smaller than other widely used CNN models (as indicated in Table II) and even smaller than the original MobileNet model, which is 17 MB in size.

We evaluate the accuracy of our method using the common confusion matrix, visually represented through a heatmap in Fig. 6. Among the 43 test images, one non oil-spill image was mistakenly classified as spill image and two oil-spill images were incorrectly detected as non oil-spill images. However, our method demonstrates high precision for both non-spill (0.90) and spill classes (0.957). The overall accuracy (ACC) achieved by our approach is 0.93, demonstrating its effectiveness in accurately identifying oil spills and non-spill images.

TABLE I. EXAMPLES OF TRAINING IMAGES USED FOR FINE-TUNING MOBILENETS.

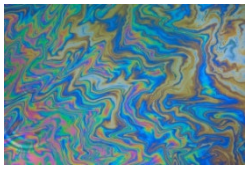
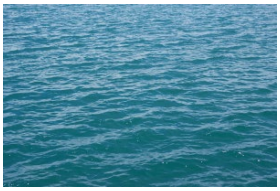
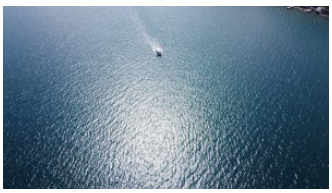
Examples for the oil spill class			
			
Examples for the no oil spill class			
			

TABLE II. COMPARISON OF MODEL SIZE AND NUMVER OF PARAMETERS OF POPULAR CNN MODELS WITH RESPECT TO MOBILENET

Model	Size	Approximate Number of parameters
VGG16	553 MB	138,000,000
AlexNet	240 MB	60,000,000
ResNet-152	240 MB	60,200,000
GoogLeNet	26.4 MB	6,600,000
MobileNet	17 MB	4,200,000
Our MobileNet	12 MB	3,230,000

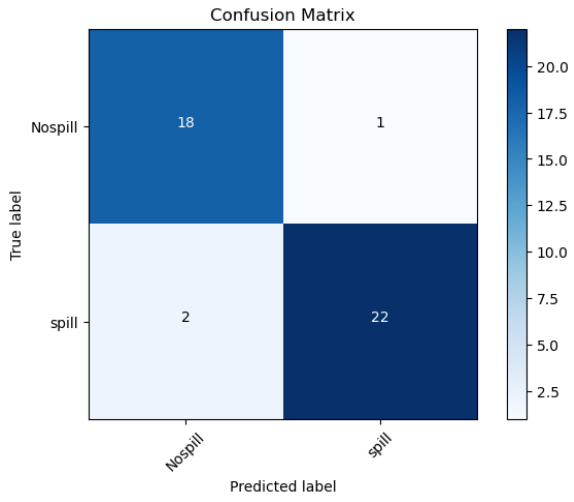


Fig. 6. Confusion matrix in the heatmap representation

IV. CONCLUSION

In this study, we focused on detecting oil spills in remotely sensed images using an efficient CNN model based on MobileNet architecture. To optimize the model's performance for our specific task, we employed transfer learning techniques by fine-tuning the pre-trained MobileNet model with our data. Our study highlights the vital role of transfer learning in CNN model development for custom applications, particularly when dealing with computationally limited platforms. The utilization of transfer learning allowed us to adapt the MobileNet model effectively to our target task, resulting in improved accuracy and efficiency in oil spill detection.

REFERENCES

[1] Climate change: IPCC report warns rapid changes needed to stem catastrophic global warming - CNN.” Accessed: Oct. 06, 2023. [Online]. Available: https://edition.cnn.com/2018/10/07/world/climate-change-new-ipcc-report-wxc/index.html

[2] “Front Matter | Responding to Oil Spills in the U.S. Arctic Marine Environment | The National Academies Press.” Accessed: Oct. 02, 2020. [Online]. Available: https://www.nap.edu/read/18625/chapter/1

[3] I. Leifer *et al.*, “State of the art satellite and airborne marine oil spill remote sensing: Application to the BP Deepwater Horizon oil spill,” *Remote Sensing of Environment*, vol. 124. Elsevier, pp. 185–209, Sep. 01, 2012. doi: 10.1016/j.rse.2012.03.024.

[4] A. L. Iler and P. D. Hamilton, “Detecting oil on water using polarimetric imaging,” in *Ocean Sensing and Monitoring VII*, W. W. Hou and R. A. Arnone, Eds., SPIE, May 2015, p. 94590P. doi: 10.1117/12.2180169.

[5] D. B. Chenault, J. P. Vaden, D. A. Mitchell, and E. D. DeMicco, “Infrared polarimetric sensing of oil on water,” in *Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions 2016*, C. R. Bostater, S. P. Mertikas, X. Neyt, C. Nichol, and O. Aldred, Eds., SPIE, Oct. 2016, p. 99990D. doi: 10.1117/12.2241866.

[6] N. Pinel, G. Monnier, I. Sergievskaya, and C. Bourlier, “Simulation of infrared emissivity and reflectivity of oil films on sea surfaces,” in *Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions 2015*, C. R. Bostater, S. P. Mertikas, and X. Neyt, Eds., SPIE, Oct. 2015, p. 963806. doi: 10.1117/12.2194278.

[7] D. Yin, X. Huang, W. Qian, X. Huang, Y. Li, and Q. Feng, "Airborne validation of a new-style ultraviolet push-broom camera for ocean oil spill pollution surveillance," in *Remote Sensing of the Ocean, Sea Ice, and Large Water Regions 2010*, C. R. Bostater, Jr., S. P. Mertikas, X. Neyt, and M. Velez-Reyes, Eds., SPIE, Oct. 2010, p. 78250I. doi: 10.1117/12.874742.

[8] S. Sun *et al.*, "Oil slick morphology derived from AVIRIS measurements of the Deepwater Horizon oil spill: Implications for spatial resolution requirements of remote sensors," *Mar Pollut Bull*, vol. 103, no. 1–2, pp. 276–285, Feb. 2016, doi: 10.1016/j.marpolbul.2015.12.003.

[9] C. Brekke and A. H. S. Solberg, "Oil spill detection by satellite remote sensing," *Remote Sensing of Environment*, vol. 95, no. 1. Elsevier, pp. 1–13, Mar. 15, 2005. doi: 10.1016/j.rse.2004.11.015.

[10] D. Stathakis, K. Topouzelis, and V. Karathanassi, "Large-scale feature selection using evolved neural networks," in *Image and Signal Processing for Remote Sensing XII*, L. Bruzzone, Ed., SPIE, Sep. 2006, p. 636513. doi: 10.1117/12.688149.

[11] A. G. Howard *et al.*, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications".

[12] F. Zhuang *et al.*, "A Comprehensive Survey on Transfer Learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, Jan. 2021, doi: 10.1109/JPROC.2020.3004555.