ISEIS

Journal of Environmental Informatics 42(1) 1-12 (2023)

Journal of Environmental Informatics

www.iseis.org/jei

Assessing Environmental Oil Spill Based on Fluorescence Images of Water Samples and Deep Learning

D. P. Liu[1 *], M. Liu[2 *], G. Y. Sun[1], Z. Q. Zhou[1], D. L. Wang[1], F. He[1], J. X. Li[1], J. C. Xie[1], R. Gettler[3], E. Brunson[3], J. Steevens[3 **], and D. Xu[1 **]

[1] *Department of Electrical Engineering and Computer Science, Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, Missouri 65211, USA*

[2] *School of Mathematics and Statistics, Changchun University of Technology, Changchun 130024, China*

[3] *U.S. Geological Survey Columbia Environmental Research Center, Columbia, Missouri 65201, USA*

Received 21 September 2021; revised 17 December 2021; accepted 26 July 2022; published online 15 February 2023

ABSTRACT. Measuring oil concentration in the aquatic environment is essential for determining the potential exposure, risk, or injury for oil spill response and natural resource damage assessment. Conventional analytical chemistry methods require samples to be collected in the field, shipped, and processed in the laboratory, which is also rather time-consuming, laborious, and costly. For rapid field response immediately after a spill, there is a need to estimate oil concentration in near real time. To make the oil analysis more portable, fast, and cost effective, we developed a plug-and-play device and a deep learning model to assess oil levels in water using fluorescent images of water samples. We constructed a 3D-printed device to collect fluorescent images of solvent-extracted water samples using an iPhone. We prepared approximately 1,300 samples of oil at different concentrations to train and test the deep learning model. The model comprises a convolutional neural network and a novel module of histogram bottleneck block with an attention mechanism to exploit the spectral features found in low-contrast images. This model predicts the oil concentration in weight per volume based on fluorescence image. We devised a confidence interval estimator by combining gradient boosting and polymodal regressor to provide a confidence assessment of our results. Our model achieved sufficient accuracy to predict oil levels for most environmental applications. We plan to improve the device and iPhone application as a near-real-time tool for oil spill responders to measure oil in water.

Keywords: water pollution, oil spill assessment, plug-and-play device, fluorescence image, deep learning, convolutional neural network, confidence assessment

1. Introduction

Quantifying oil during a spill response is a significant challenge due to limits in the current detection methods (Fingas, 2014). Conventional laboratory analytical methods follow the procedure: samples are collected in the field and shipped to a laboratory, and then benchtop instruments are used to determine the contaminant concentrations, often long after the oil spill has occurred. These standard methods generate high-quality data because of their reliability, accuracy, and precision. However, there is a need for rapid screening of samples to support the spill response or real-time decision making in the field.

There are numerous methods and technologies to analyze oil during a spill. A recent publication outlined the various available technologies and those used during the Deepwater Horizon Oil Spill, and summarized them in the adapted Table 1 (White et al., 2016), where over 50,000 oil analysis samples and over 52% of the water samples collected and analyzed were below detection limits for total polycyclic aromatic hydrocarbons (PAH) (NOAA, 2020). These technologies, ranging from field-deployed fluorometers or spectrophotometers to laboratory technology, have been adapted or ruggedized for use in the field. Among them, portable mass spectrometry or fluorometry can provide immediate results but has a high cost for instrumentation (Nixon and Michel, 2015). Remote sensing is also employed but has less accuracy in measuring pollution levels (Fingas, 2014). Simple methods for estimating oil levels in the field can be valuable for disaster responders and resource managers to evaluate chemical contamination in natural systems. However, several fundamental challenges must be overcome to implement these technologies in rapidly measuring oil in aquatic samples as part of the oil spill response, including measurement accuracy, confidence assessment (error estimate) of measurement results, and usability.

Among the oil detection methods, the fluorescence approach, either field-portable or laboratory-based, detects, and provides a semiquantitative analysis of oil from spills through the absorption of ultraviolet light and emission as fluorescence (Fantasia

* Co-first authors.

** Corresponding author. Tel.: +573-876-1819.

E-mail address: jsteevens@usgs.gov (J. Steevens).

** Corresponding author. Tel.: +573-882-2299.

E-mail address: xudong@missouri.edu (D. Xu).

ISSN: 1726-2135 print/1684-8799 online

 doi:10.3808/jei.202300491.

Table 1. Methods Used to Analyze Oil During a Spill and Their Performance

Instrument	Sensitivity	Cost (USD)	Usability	Availability	Portability
Mass Spectrometer	High	160,000	Specialized	Limited	Med
Fluorometer	High	6 ~ 13,000	Non-expert	Wide	Med
Beam Transmissometer	High	7 ~ 8,000	Non-expert	Wide	Low
Holocam	High	Unavailable	Specialized	Limited	Low
Current Device	Med	< 200	Non-expert	In Development	High

Note: Cost is listed in United States dollar (USD).

et al., 1971; Keizer and Gordon Jr, 1973). Alternative methods have also been developed for oil detection in the field using qualitative means by visually comparing fluorescence images of samples to a set of reference oil standards (Russell et al., 2016). In addition, mobile platforms were adopted for detecting and quantitating water quality in the field to assess oil in the environmental water faster, more portably, and more cost-effectively (Seo et al., 2019). Some mobile systems employed artificial intelligence-based models (Gunda et al., 2019), and some mobile phone cameras detected wavelength beyond visible light (Granica and Tymecki, 2019). However, the accuracies of these methods are not high or robust enough for practical usage. In addition, they did not adopt cutting-edge computational methods such as deep learning to quantify oil concentration from images.

Previous research describing the use of optical or fluorescent data to estimate oil has relied on a linear regression standard calibration curve to predict the level of oil in the sample. The development of a fluorescence-based screening tool shows the challenges in calibration using visible images (Russell et al., 2016). Using this approach, visible differences in the oil levels are limited to between 1 to 100 mg/L; no visual predictions are possible at levels above 100 mg/L. Using a laboratory-based fluorometer and gas chromatograph, the detection range of those same samples can be extended from 100 mg/L to nearly 1,000 mg/L. One challenge with using standard linear regression calibration curves is that they rely on a standard used to calibrate the system and often deviate from linearity. Furthermore, oil is composed of a complex mixture of aromatic and aliphatic compounds (USEPA, 2003). The composition of these chemicals varies according to the source of crude material, aging, and processing. Traditional linear regression methods will fail to accommodate for the variability of the composition of PAH and aliphatic contribution to the total amount of oil. Statistical models, such as linear regression, are based on determining the statistical relations between two parameters. In contrast, the focus of machine learning is the accuracy of the prediction based on the input parameters. Coupling sensors with machine learning models offer the benefit of being able to interpret and address variations in the data input to improve the model predictions (Han et al., 2019), including the variability of the sample composition (oil type), degradation, fluorescence, system used to measure fluorescence, and the visible image produced by the associated hardware.

In this work, we developed new hardware to support fluorescent image collection of environmental water samples and devised our deep learning model for assessing oil content in the samples. Our hardware was built upon previous techniques for oil detection in the field using qualitative means by visually comparing samples to a set of reference oil standards (Russell et al., 2016). The water sample is collected, extracted, transferred to a cuvette and inserted into a device attached to a mobile phone. The phone is used to capture images of oil fluorescence upon illumination with a 380 nm light source. Such usage of cell phone or small computer to measure chemicals in a sample has been published previously for spectrophotometers and agriculture applications (McGonigle et al., 2018; Ayaz et al., 2019; Dutta, 2019; Pramanik et al., 2019). Other devices are small fluorometers but with the same problem that they rely on statistically based linear regression calibration curves not amenable for field applications of complex oil samples (McCracken et al., 2017; Hossain et al., 2019). The main novelty of this work is to apply deep learning to quantify oil concentration. Following the attention mechanism paradigm in the residual channel attention block (Zhang et al., 2018), our model integrates histogram features with attention as a histogram bottleneck block (HBB). We further employed a confidence interval estimator to assess the performance. Our model is the first method to adopt deep learning on oil spill estimation, to the best of our knowledge.

2. Methods

2.1. Data Preparations for Training the Machine Learning Model

2.1.1. Sample Generation

We used the source oil from the Macondo Prospect well MC 252, located approximately 80 km southeast of the confluence of the Mississippi River with the Gulf of Mexico, to train and test the model. We chose MC 252 riser oil as a standard material because it has a relatively high aromatic content (approximately 16% aromatic hydrocarbons) and has been intensively studied (Reddy et al., 2012; Incardona et al., 2013; Forth et al., 2017). As part of the spill response, an archived oil sample obtained from the riser insertion tube on the Enterprise Discoverer on May 20, 2010 (Joye et al., 2011) was selected for use. The percentage of aromatic-containing compounds in the source oil is approximately 25% by mass. The source oil is known to contain Nalco EC9323A defoamer (injected topsides), methanol (via subsea injection at 10,000 ppm), and VX9831 oxygen scavenger/catalysts solution (subsea injection), as well as Corexit™ 9500 and 9527 dispersants. However, the dispersants do not contain aromatic chemicals and are therefore not expected to interfere in the model's training (Choyke and Ferguson, 2019). The MC 252 oil has a large 100 nm bandwidth with a maximum of 424 nm in the fluorescence spectrum.

Further data acquisition was conducted by measuring the emission intensity at this wavelength (424 nm). Certified ACS

hexane was purchased from Fisher Scientific (Cat # H292-500, Waltham, Massachusetts) to prepare the oil samples. Oil was weighed gravimetrically to create a concentrated stock solution and dissolved by directly adding hexane in a volumetric flask. The different concentrations of the oil standards were created by serial dilution in duplicate using the concentrated stock solution. Spectral analysis of the oil sample was conducted on a Horiba Fluorolog-3 spectrofluorometer with a xenon lamp light source and PMT detector (Kyoto, Japan). Before the data acquisition, the xenon lamp signal was verified via background scan calibration. The signal-to-noise ratio was determined by a water-Raman scan using laboratory 18 MΩ deionized water. The emission spectrum of the oil sample was obtained by scanning a 10 mg/L oil in hexane solution from 350 ~ 600 nm using a 380 nm excitation wavelength and 5 nm bandpass. A standard curve was generated by diluting a 500 mg/L oil in hexane by the appropriate volume to make a 0 ~ 50 mg/L oil in hexane series and measuring fluorescence at the emission maximum.

2.1.2. Image Collection

A three-dimensional (3D) printed cell phone adaptor (Figure 1) was used to illuminate and collect images for the analysis using an iPhone model 7, 8 or SE. The device was designed using the Tinkercad (tinkercad.com) 3D modeling program. The 3D model of the device was exported as a .stl file to MatterControl software (matterhackers.com) for slicing and printing. The device was printed on a Makeit Pro-M 3D printer (makeit-3d.com) using the MatterHackers PRO series, with black, PTEG, 3D printing filament. It was designed to be used with an iPhone 7 (Model Number MN8L2LL/A) to capture images. A Mastiff brand, A2, 3W, 380 nm, ultraviolet, LED flashlight was used as the UV light source (Hong Kong, China). Samples to be analyzed were placed in a 2.5 mL disposable polystyrene cuvette (12.5 × 12.5 × 45 mm) with a cap.

The capped cuvette containing the sample was slid into a chamber on top of the device. A separate cover was placed over the top of the cuvette to block out ambient light. The printed device positioned the UV flashlight directly below the cuvette to illuminate the sample. The device held the iPhone in place with the lens of the primary camera positioned parallel to a long side of the cuvette and in the approximate center from top to bottom and side to side. Once in place, the iPhone camera could view the illuminated cuvette through a circular hole in the chamber wall that held the cuvette. The cuvette was placed into the device to photograph a sample, covered with the cuvette cap, and illuminated with a UV flashlight. A standard photograph (not square) with no filters or effects was taken for each sample illuminated using the prototype hardware with a 380 nm LED light source. The fluorescence of each sample was photographed in duplicate to produce 1,647 images to train and test the model.

Oil spill residues are complex and can occur in various forms, including sheens, tarballs, mousse, mats, and coating physical or biological surfaces (NOAA Hazmat, 2016). This method aims to provide an initial support tool for screening suspected samples that trigger additional chemical analysis. Therefore, we developed a set of oil samples that target higher levels of crude oil in water for fluorescence detection. First, we created three series of oil samples to train or test the model ranging from 2 to 500 mg/L. The low-range training series was composed of MC 252 from 0 ~ 200 mg/L in hexane in 2 mg/L increments with five replicates. Next, a high-range training series of 200 ~ 300 mg/L in hexane was prepared in 5 mg/L increments with five replicates. Finally, to test the upper end of the model, we developed a training series from 300 ~ 500 mg/L in 25 mg/L increments with five replicates. Six repeated sample images as independent measurements were collected in each incremental concentration level (2, 5, or 25 mg/L). During image collection, we minimized the influence of incidental sample cuvette reflectance by using different cuvettes for each independent measurement and measuring the sample twice with the same light source. The concentration level in our dataset thereby follows a nearly uniform distribution, which supports our model to estimate the concentration without handling imbalanced data.

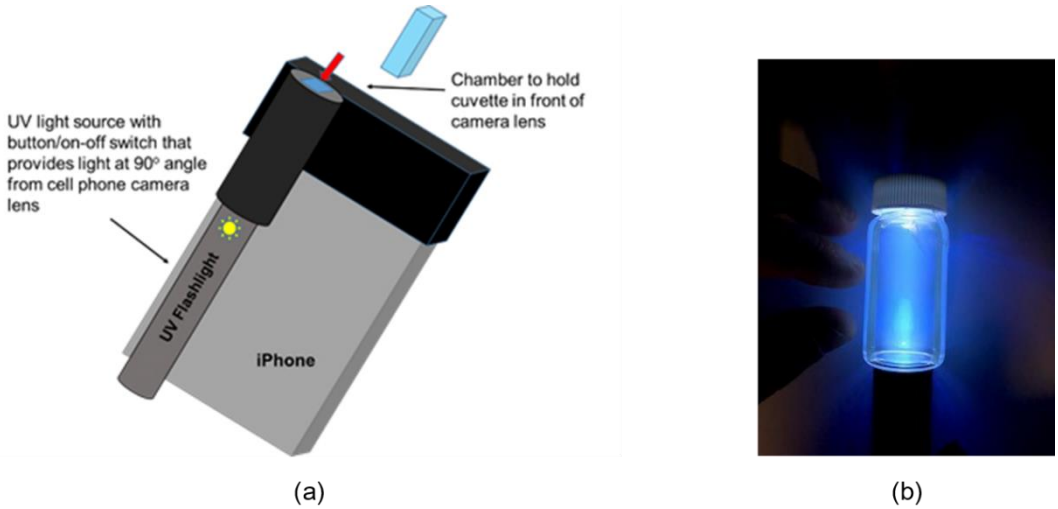

Figure 1. Data collection device: (a) 3D printed iPhone adaptor for image capture; (b) image of Macondo MC252 oil fluorescence. Photo by Ryan Gettler, USGS.

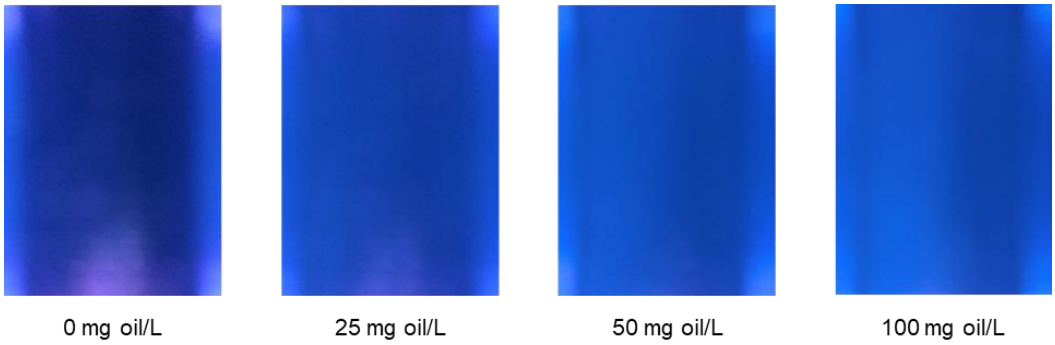

Figure 2. Calibration images of MC 252 crude oil in hexane at increasing concentrations from left to right.

2.1.3. Data Preprocessing and Feature Analysis

Figure 2 illustrates the relation between image intensity and crude oil concentration in our oil spill sample dataset (https://github.com/minoriwww/waterquality). A clear trend of increasing brightness and intensity is due to the fluorescence emission observed as concentration increases. This trend is visible from 2 to 50 mg/L; however, beyond 50 mg/L, the difference in image intensity becomes challenging to distinguish, consistent with results from previous investigators using visual inspection to estimate oil concentrations (Russell et al., 2016). There are some image patterns, such as the cuvette on the sides of the image and the light source on the bottom of the image, but overall, the images have low contrasts. As a common practice, we reduced image resolution to 512 × 512 pixels for input into deep learning models, saving computing resources while maintaining the information content. The red, green, and blue (RGB) channels were used to represent the color spectra. The 1,647 images in our dataset were split 80:20 into a training set and a test set: 1,317 samples in the training set and 330 samples in the test set. The training sample images were augmented by horizontal flipping, shift, and distortion to reduce the variance of deep-learning models during the training.

2.2. Deep-Learning Model with Histogram Features

The estimation of oil concentrations from images in the OilSS is different from most computer vision tasks because of the low contrast within an image and high similarity among images. Unlike face recognition, OilSS images are flat, with most features hidden from visual inspection. The explicit features of OilSS images that human eyes can capture are merely color and slight color variations, which are insufficient to estimate the oil concentration. Inspired by histogram extraction in digital image processing, we analyzed the features in the spectral space (Clark et al., 2010) and color distributions. We added histogram features with location information for the deep-learning model (see Supporting Information). Part or whole histograms of a figure were collected and used as extra histogram features on machine learning models. Several deep learning models have adopted such histogram features (Zhang et al., 2005; Gustafsson et al., 2019; Latif et al., 2019).

Different combinations of histogram features were investigated in other studies, such as color channel histogram (calculating histogram in RGB channels separately) and histogram of oriented gradient (HOG) (Rebetez et al., 2016; Nam and Kim, 2017; Sedighi and Fridrich, 2017; Hussain et al., 2019). However, these studies did not examine integrating the histogram extractor with our attention mechanism. In this work, we employed a deep neural network and the attention mechanism to model histogram features from an input image and use the features in the prediction task. Attention mechanism is a technique that mimics cognitive attention. It can enhance some parts of the input data or features in a neural network layer while diminishing other parts. This can help the model focus on the small but important part of the data by using a weighted combination of input vectors, with the most relevant vectors having the highest weights. The weights can be visualized to help interpret the model and learn using the training data.

We assume the target (the concentration level in this case) can be determined by the input image, and the difference between two input images is highly related to their differences in image features and histograms. To extract the histogram feature, we devised a histogram pooling layer with attention, encapsulated as a histogram attention block (HAB), and multiple HABs are further integrated as histogram bottleneck block (HBB) on the top of a lightweight ResNet50 (He et al., 2016). Histogram pooling extracts local histogram features by convolution filter, then combines them to obtain a histogram of the whole image. Compared with direct whole-image histogram calculation, the histogram pooling-based method converges faster, making the result more stable. Moreover, using the histogram pooling with the attention structure, the regional information (especially corners and edges, as shown in Figure 4) is included in more detail. The residual network (ResNet) is a deep convolutional neural network with 50 residual layers, with a strong representation power because of its deeper network. To study the network, we followed the notation in a network architecture report (Lathuilière et al., 2019): let CB^i denote the i^{th} convolutional block (also called ResBlock), GAP denote a global average pooling layer, and SM denote a soft-max layer. Therefore, Equation (1) can be used to describe the ResNet50 architecture. Equation (2) can be used to describe the model ResNet50 with 2 HBBs. Our overall network architecture is shown in Figure 3.

$$CB^1 - CB^2 - CB^3 - CB^4 - CB^5 - GAP - SM \quad (1)$$

$$HBB^1 - CB^1 - HBB^2 - CB^2 - CB^3 - CB^4 - CB^5 - GAP - SM \quad (2)$$

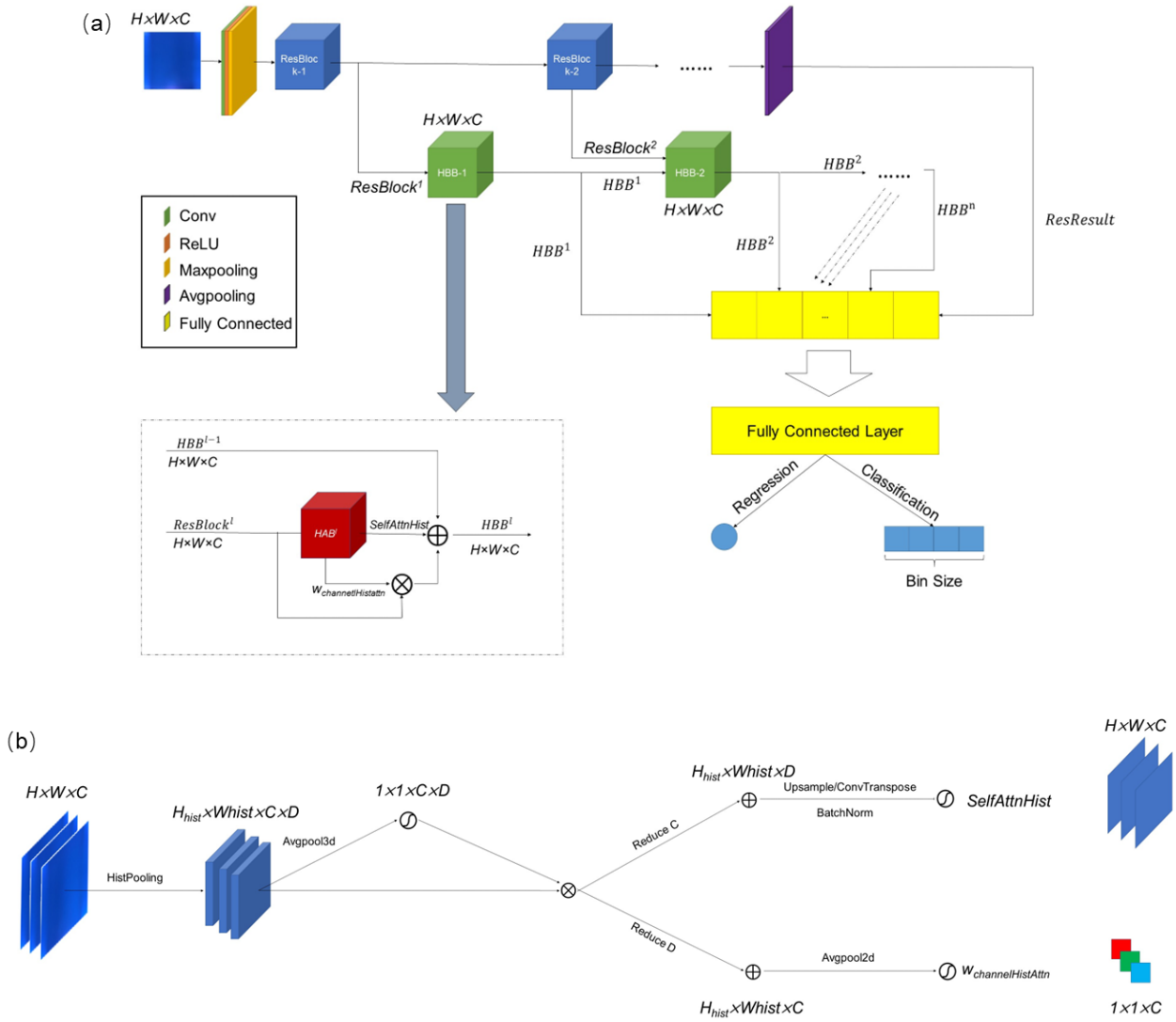

Figure 3. Architecture of the deep learning model (a) including histogram bottleneck block (HBB, shown in the green box) and the histogram attention block (HAB, shown in red box). The detail of HAB in (a) is illustrated in (b). C, D, H, W denote channel, depth, height, and width, respectively. Other abbreviations are convolution (Conv), rectified linear unit (ReLU), pooling operation that selects the maximum value (Maxpooling), pooling operation that selects the average value (Avgpooling), width histogram (Whist), self-attention of the histogram (SelfAttnHist), and channel attention histogram (ChannelHistAttn).

2.2.1. Histogram Attention Block (HAB)

Each histogram-related block extracts the histogram information separately. The histogram attention block (HAB) adopts a histogram pooling layer called HistPool for histogram extraction (Ghosh et al., 2015; Yin et al., 2016), as shown in Figure 3(b). The extraction yields a 3D feature map with a size of Hhist × Whist × D, and then a two-branch dimension reduction is processed. The upper branch keeps the histogram with channel integration, while the lower branch encloses the histogram information in each channel. These two branches are designed to catch different features and can be tested separately.

Inspired by the principal component analysis (PCA) filter convolution of PCANet (Wu et al., 2017), HistPool adopts the regional maximum activation of convolutions (RMAC) pooling (Tolias et al., 2015) augmented histogram filter as an enhanced histogram extractor for the position information. Using random block sampling in the RMAC pooling, the receptive field is enlarged for more informative histogram features. The input feature maps first pass through the HistPool, with an additional histogram dimension as depth. Then, the resulting 3D feature maps attend themselves on channel and depth dimensions. The output of HAB has two output branches, as shown in Figure 3(b), one branch for $W_{channelHistAttn}{}^{l}$ and the other branch for self-attention of the histogram. The left will keep the origin channels, with the histogram dimension reduced by averaging, while the right one will reduce the original channel and the histogram dimension. The HAB block before branching is represented as:

$$HAB = HistPool(ResBlock^{l-1}) \otimes Sigmoid(GAP_{H,W}^{C,D}(HistPool(ResBlock^{l-1}))) \quad (3)$$

$$GAP_{H,W}^{C,D}(x) = \frac{1}{H \times W} \sum_{h}^{H} \sum_{w}^{W} x_{c,d}(h,w) \quad (4)$$

where $GAP_{H,W}^{C,D}$ is a 3D average pooling; C, D, H, W denote channel, depth, height, and width, respectively.

The self-attention histogram $SelfAttnHist^l$ along depth dimension is obtained by Equation (5).

$$SelfAttnHist^l = W_{ConvTrans}(\sum_{c}^{C} HAB) \quad (5)$$

where $W_{ConvTrans}$ represents a learnable weight set of a transposed 2D convolution layer as an upsampling operation on the channel dimension, as shown in the right branch. The sum operation first squeezes (reduces) the channel dimension to 1, and then $W_{ConvTrans}$ projects depth D (with bin size as its depth) into C, considered as a new channel. This channel squeeze-then-switch trick aims to align histogram feature depth to the input channel. The upsampling projects the depth dimension from the bin size to the input channel size, facilitating further combinations.

2.2.2. Histogram Bottleneck Block (HBB)

We investigated a pooling mechanism of RMAC (Tolias et al., 2015) for the filter convolution (Conv) on the feature map. Here the Conv kernel is substituted with a histogram operation kernel. Finally, we integrated and stacked the components, named histogram bottleneck block (HBB), inspired by the residue design of bottleneck block *ResBlock* in ResNet (He et al., 2016) and residual channel attention block (Zhang et al., 2018). HBB takes advantage of the recurrent structure and uses the previous layer's output by the ResNet bottleneck block and HBB.

Upon extracting histogram feature maps, our next step was to study the relations between the feature space and the original image space (Rebetez et al., 2016; Wang et al., 2016; Sedighi and Fridrich, 2017; Hussain et al., 2019). First, we examined the attention mechanism, which can fulfill the need of focusing feature importance using feature interactions. The attention mechanism selects and intensifies a focused location by its weight. It is naturally spatial, channel-wise, and multi-layer (Chen et al., 2017). The channel level attention and spatial level attention generally have different focuses of an image or its features (Chen et al., 2017), and those attention heads can be stacked as blocks for tackling image super-resolution tasks (Zhang et al., 2018). Based on studies of histogram features, we assume the filter with a histogram kernel can preserve spatial information (Ghosh et al., 2015; Yin et al., 2016). Inside HBB, a unit called histogram attention block (HAB) is proposed to exploit the histogram information while preserving image relative spatial information. The inner construction of HBB is shown in Figure 3(a).

Let CB^i denote the i^{th} convolution block, *GAP* denote a global average pooling layer, and *SM* denote a soft-max layer. According to the notation, the ResNet50 architecture can be described as follows: $CB^1 - CB^2 - CB^3 - CB^4 - CB^5 - GAP - SM$ (Lathuilière et al., 2019). An HBB in the l^{th} layer harnesses the image and its histogram information, with residual learning via Equation (6).

$$HBB^l = f_{act}[\mathbf{W}_{channelHistAttn}^{l} ? ResBlock^{l-1} \oplus SelfAttnHist^l + f_{act}(HBB^l - 1)] \quad (6)$$

where f_{act} represents non-linear transformation, such as ReLU, sigmoid, or max; $ResBlock^{l-1}$ is the output of the $(l-1)^{th}$ bottleneck block of ResNet, with an element-wise sum $\oplus$ by $SelfAttnHist^l$, one of the HAB outputs; another output $\boldsymbol{W}_{channelHistAttn}^{l}$ is an attention weight for the input channel, determined by the attention function of histogram itself, formulated as:

$$\mathbf{W}_{channelHistAttn}^{l} = Sigmoid(GAP_{HW}^{C}(HAB(ResBlock^{l-1})) \otimes \sum HAB(ResBlock^{l-1})) \quad (7)$$

where the GAP_{HW}^{C} denotes the 2D global average pooling function along the height and width dimensions, aligning the channel number to C. The $\otimes$ represents the mark of element-wise multiplication. Here, the input depends on the layer position of HBB. Every HBB block's output passes through a Conv layer, a normalization layer, and an activation layer. All HBB blocks are concatenated before a fully connected layer, as shown in Figure 3(a). The Bottleneck blocks and histogram blocks are finally connected via block composition. The first HBB blocks HBB^0 takes the original image as input and functions as a standard block color histogram operation.

2.3. Model Training and Fine-Tuning

Our deep neural net model is trained by the ADAM optimizer (Kingma and Ba, 2014) with a learning rate set to 0.0001. We used PyTorch (Paszke et al., 2017) to implement our models with a Nvidia 1080Ti GPU. In addition to OilSS, we also applied the CIFAR-10 dataset (Krizhevsky and Hinton, 2009) to test the generality of our models, which demonstrated accurate and robust performance.

The input image is normalized and resized to 256 × 256 with a batch sample size of 32. For the parameters of RMAC method in HAB, we tested the settings with a changeable size of convolution kernel from 2 × 2 to 7 × 7. Finally, 7 × 7 was selected by balancing the computational cost and accuracy. In each channel of the feature map, the histogram was generated using a bin size of 8 (i.e., every single 2D channel will be upsampled by pooling the features to 3D with a depth of 8, whose granularity is sufficient for our dataset). Adaptive minimum and maximum limits of each channel were used in this pooling. The default up-sampling kernel size was 3 × 3 and the default convolution kernel size was 5 × 5. Since the original parameter setting on ResNet with RMAC took a long time to converge, we optimized the process by trimming the image size and histogram bin number.

2.4. Confidence Interval Estimation

Since predicting oil levels in water samples has uncertainties associated with sample quality and model accuracy, it is useful to show a confidence interval combined with a probability statement to express the degree of uncertainty instead of just a single value prediction. The actual value of a concentration level should fall into a defined interval within a stated probability. Such confidence interval estimations act as a better indicator to enhance the reliability and usability of the model. Furthermore, the incorporation of uncertainty associated with oil concentration measurement during a spill is essential to understand the limitations of the screening tool and resulting data when making decisions about a further evaluation or determining risk from exposure (Suter II, 2016).

We further split all the training samples into the training subsets and the validation subsets on 5-fold cross-validation to investigate the confidence intervals. For real-world applications, we used a combination of the five models by taking the average of the five outputs. We employed a gradient boosting regressor (Nixon and Michel, 2015) to estimate a confidence interval conditional upon the error boundaries from the observation data, i.e., the predicted and actual oil concentrations of our 5-fold validation samples. The regressor assembled several individual decision trees in boosting the ensemble strategy to minimize the quantile loss function (Yu et al., 2003) as follows:

$$quantile\ loss = (\alpha - 1) * (actual - predicted)$$
$$if \quad (actual - predicted) < 0 \tag{8}$$

where α is the quantile. In this work, we set α as the 90th quantile so that the loss function would drive the regressor to fit an error upper boundary with a 90% probability that the actual oil concentration level would fall below its prediction. While we also estimated another 10th quantile regressor to define the lower error boundary with a 10% probability that the actual value would be above its prediction. Since the high oil concentration levels above 300 would be difficult to distinguish by humans and were not the model's focus for practical purposes, we combined all high prediction values above 300 in the validation samples into one class for confidence assessment. Furthermore, to enhance the generalization of our confidence interval, we smoothed the fitting upper and lower curves as quadratic curves by further applying polynomial regression (Ostertagová, 2012). Eventually, we can send a prediction of any query sample to the pre-built confidence interval estimator and get its corresponding upper and lower limits to form the 90% confidence interval. This approach can be used to obtain other confidence interval levels based on the requirements of the analysis.

3. Results and Discussion

3.1. Overall Performance

We tested various machine learning models, including linear model, k-nearest neighbor (KNN) model, random forest model, ResNet50, and ResNet50 + HBB, as shown in Table 2. Among them, the ResNet50 model with two HBBs (ResNet-50 + HBB) performed the best for both regression and classifycation, with a bin size of 10 mg/L for the oil content values. All the samples in each bin, e.g., 100 ~ 110 mg/L, will be given the same label. Hence, it is set as the default model in this study, with root mean square error (RMSE) as a metric. We also tested our model for classification on the CIFAR-10 dataset, a collection of images widely used to train machine learning models and computer vision algorithms (Krizhevsky and Hinton, 2009). We observe that Resnet with HBB achieved higher accuracy than plain Resnet50, with R^2 equal to 0.78 and 0.75, respectively. Since typical multi-class classification losses ignore the ordered information between the discrete labels, we develop an ordinal loss to learn our network parameters (Niu et al., 2016). The ordinal loss can improve the accuracy by 0.09. Both ResNet and ResNet + HBB converge rapidly, while the HBB embedded model has higher accuracy.

We also tested different networks' performance with and without image generation as data augmentation. Table 2 shows that in both situations, with and without the generated input, our default model consistently achieves the best result with an RMSE of around 11 and prediction accuracy of approximately 80% for the classification of the values in 10 bins. A possible reason for the success of our model may be the translational invariance property of the inception block (Larsson et al., 2017), which helps reduce the variance by increasing the data flow path and filtering out the noises.

Table 2. Comparison of Different Models in Different Configurations

Metric&Tasks	RMSE (Regression)		Accuracy (Classification)	
Dataset	OilSS-color	OilSS-gray	OilSS-color	OilSS-gray
Linear	15.89	15.89	0.12	0.15
KNN	20.28	20.28	0.37	0.37
Random Forest	32.52	32.28	0.48	0.47
ResNet50	13.14	13.56	0.61	0.60
ResNet50 + HBB	11.49	11.32	0.81	0.80
ResNet50 (pretrain)	13.37	13.42	0.66	0.65
ResNet50 (pretrain) + HBB	12.48	12.02	0.79	0.78

3.2. Impact of Pretraining and Color Input

Most image deep-learning models are pre-trained on a large dataset before training the model using the defined training set. Surprisingly, in our case, the pre-trained weights using ImageNet (http://image-net.org) weaken the model in converging time and performance, as shown in Table 2. Furthermore, pre-training the model by ImageNet may mislead the model since the OilSS is dramatically different from the images in ImageNet with simple color and low contrasts. Therefore, the performance of our default model without pretraining outperforms the one with pretraining. We also tested whether color is the most important feature of our data using OilSS-color with the original image's color and OilSS-gray with its grayscale. As shown in Table 2, the color input can give more information to HBB so that the performance on OilSS-color is better. For other models, the result is close.

Classification accuracy is the percentage of predictions that fall into the correct ranges of 10 bins. All units are in mg/mL. Comparison is based on the root mean square error (RMSE) and accuracy classification is based on the models using the oil spill sample dataset (OilSS). Different models considered are linear, KNN (k-nearest neighbor), random forest, ResNet50 (convolution neural network with 50 layers), and ResNet50 + HBB (histogram bottleneck block).

3.3. Attention Mechanism

We analyzed the middle layer of our model, as shown in Figure 4. The darker colors (green and blue) indicate that those regions played more important roles in the predictions. We can observe that the feature maps are able to tell the difference, which means our network captured useful features. Also, we analyzed some failed cases where the prediction concentration level diverted significantly from the ground truth labels. In Figure 4(b), the model focuses on the corner of the input image, where the light source is positioned. It suggests a possibility of light leaking in the measurement for this sample.

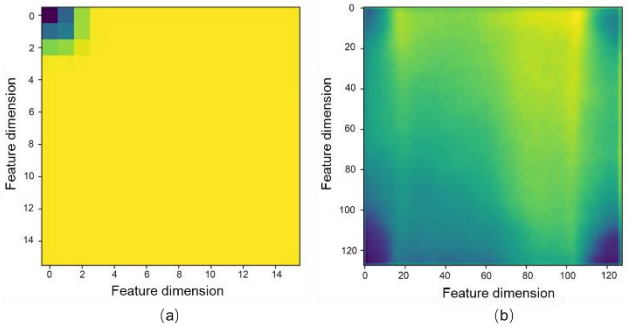

Figure 4. Feature maps that show the network values of the convolution (Conv) layers: (a) feature map of the 10th Conv layer; (b) feature map of the third-last Conv layer, which is a failed case. Each feature map shows the attention value of each neural network parameter in Conv layer. Yellow indicates light weights, and green or blue indicates heavy weights of the feature map. A heavier weight indicates more importance of the corresponding neural network parameter in the prediction model.

3.4. Final Model Performance

We conducted a 5-fold cross-validation for the final model (ResNet50 + HBB). We divided our data into six folds evenly. Five folds were used to perform the 5-fold cross-validation, and the last fold was reserved for testing the combined model of the five folds. Figure 5 shows the result of an example fold, indicating the model training converges well in its loss function. Figure 6 and Table 3 show the validation results for each fold and the combined model. It can be seen that the performance variations among the five folds are small. The combined model outperforms most folds (except fold 3). The combined model using more data is probably more robust; hence, it is expected to achieve better results in practice.

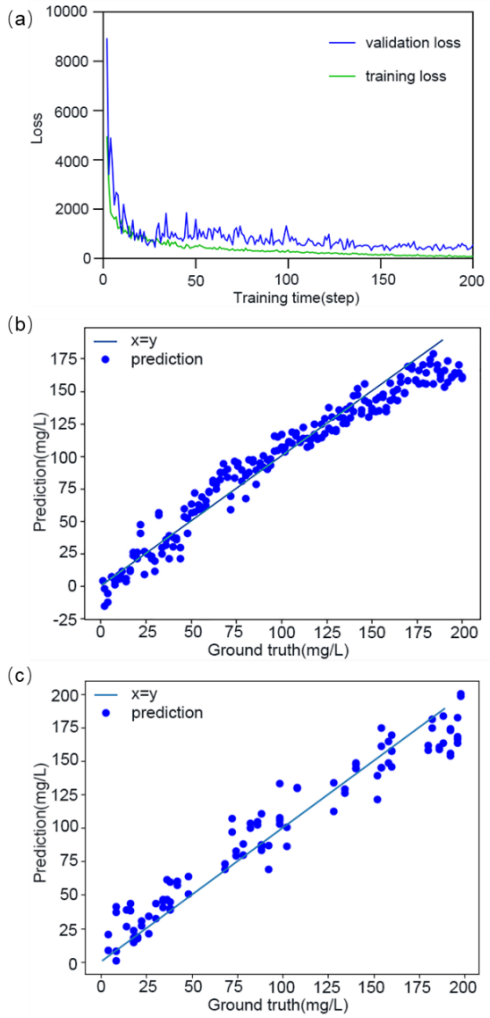

Figure 5. An example (fold) of the 5-fold cross-validation result: (a) the loss curve of the training and validation, showing that the model gradually converges as the number of training steps increases; (b) training result; (c) validation result. Each dot in (b) and (c) represents one sample, showing the predicted value vs. ground-truth value (experimentally measured value) of oil concentration in mg/mL. Note the different y-axis in (b) and (c).

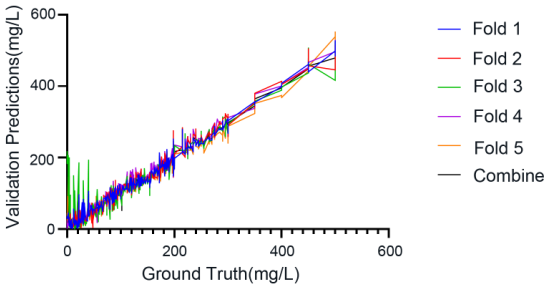

Figure 6. Validation results for each fold and the combined model.

Table 3. Validation Performance of Each Fold and Combined Model

Model	RMSE	MAE	MAPE	R^2
Fold 1	23.2	14.6	0.37	0.954
Fold 2	35.5	16.8	0.42	0.895
Fold 3	20.2	14.0	0.29	0.965
Fold 4	27.5	15.8	0.38	0.935
Fold 5	22.9	17.8	0.24	0.945
Combined Model	20.5	12.3	0.39	0.962

3.5. Confidence Interval Estimation

We divided the validation data into several bins according to their predictions from our deep model. We then produced the boxplot as Figure 7(a) to display the distribution of their predictions using bins. It shows a few significant outliers. Next, we estimated the error boundaries based on a gradient boosting regressor with 90th and 10th quantile loss functions, as shown in Figure 7(b). Finally, we further smoothed the boundaries by a polynomial regression as shown in Figure 7(c) and retrieved the confidence intervals for testing samples according to these boundaries (Figure 7(d)).

To evaluate the performance of the confidence interval estimator, we measured the hit rate and average range of the predictive interval. The hit rate is the percentage of ground truth falling into the predictive confidence intervals. The average range denotes the mean confidence intervals from all testing samples. The higher target rate indicates that our confidence interval estimator is more precise, while the narrower confidence interval implies smaller uncertainty. In practice, we obtained their trade-off by setting the 90th quantile loss in a gradient boosting regressor and setting a two-degree polynomial in polynomial regression. The confidence intervals of our model achieved an 80.06% hit rate and an average interval width of 37.62 on the testing samples. Thus, our users can be more than 80% certain that the actual concentration level of any testing sample lies in its predictive interval, which is useful for understanding the confidence of data generated using the model and limitations when making decisions about the oil level during a spill.

4. Discussion

This study demonstrates that using a cost-effective mobile device can perform useful scientific measurements. Similar methods have been developed for medical applications (Wei et al., 2013) but have not been significantly used in field measurements. For oil assessment, field-based technologies provide real-

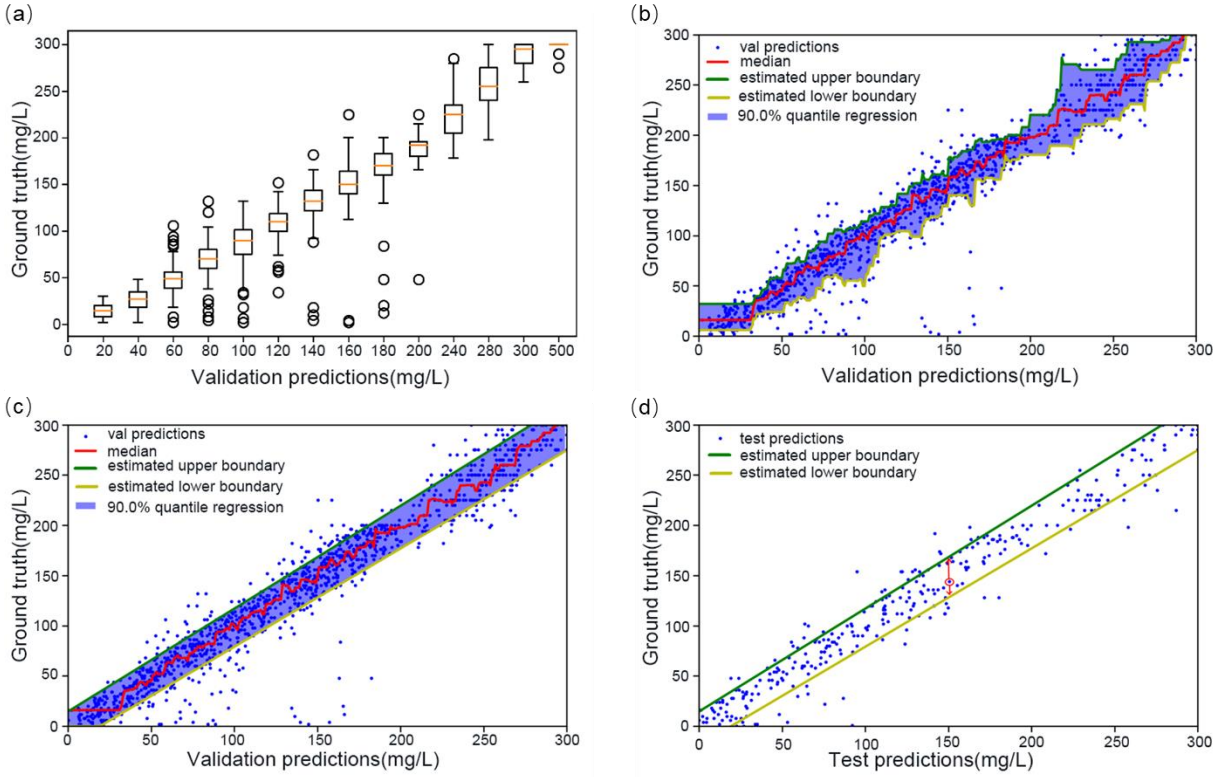

Figure 7. Confidence interval estimation: (a) boxplot of predictions of our deep architecture on validation samples (a combination of 5 validation results from the 5 corresponding models in the 5-fold cross-validation), where the box denotes the data of the first quartile (Q1) to the third quartile (Q3), the red line is the median, the bar represents the data in the range of [Q1 – 1.5 × (Q3 – Q1), Q3 + 1.5 × (Q3 – Q1)], and the data out of the range are outliers; (b) error boundaries based on gradient boosting regressor of 90% quartile; (c) boundaries by a polynomial regressor of 90% quartile; (d) query confidence interval following the red arrows for a testing sample shown in the red circle. All units are in mg/L.

time information but often use large devices deployed from ships or require more extensive power sources. While it is cheaper to use laboratory-based methods for oil assessment, it usually takes weeks to months to get results after collecting samples from the field. A cell phone device, as reported here, can fill a gap. Although the data quality cannot reach the measurements from state-of-art technologies, the results from our device are useful for practical purposes. The device is affordable for many to monitor the environment at large. Potentially, it can be used for citizen science. For example, high-school students may use the device to monitor their neighboring open water pollution.

The key novelty of this work is the application of deep learning in mobile device-based oil assessment. We witnessed more and more applications of deep learning in many engineering areas, but in each domain, it requires some special considerations to make the deep-learning application effective (LeCun et al., 2015). In this case, the images are flat and have a low contrast, which are dramatically different from typical images like human photos and natural scenes, where the mainstream deep learning applies. Hence, we developed a new general framework to handle our images, which includes (1) using both the real space images and the histograms in the phase space; (2) applying an attention mechanism to enhance the performance and provide some model interpretability; (3) developing a confident assessment to give a confidence interval of the prediction result. This framework integrates the histogram bottleneck block and the histogram attention block, and it may be used for other industrial measurements and engineering assessments, such as hydrological image analysis that previous studies addressed (Song, 2020; Li et al., 2022).

A limitation of the approach described here is that a limited clean image dataset (OilSS and CIFAR-10) was used to train the network. The model will be further tested using a broader range of petroleum hydrocarbon types with field-collected samples with potential matrix interferences, enabling our model to be transferable to diverse field test cases. In addition, we are developing a mobile app for field use and data collection and further enhancing our model based on the new data. Another issue is that smartphones have built-in artificial intelligence to modify images (Morikawa et al., 2021). Although our deep-learning framework can most handle this factor in the prediction, it adds some complexity that may reduce the robustness of the prediction. Hence, we are building new hardware independent of mobile devices, with fixed fluorescence light and camera lens. This also allows fluorescence images of water samples to be collected and transmitted to a wide range of mobile devices, such as Android phones, iPad, etc., improving usability. Future efforts to develop the model for application in the field should include demonstration and use by the oil spill community in real-world applications. Our goal is to make the device and the mobile app widely used in the oil spill responder community.

5. Conclusions

This paper formulated the oil spill estimation task, built a standard image dataset (OilSS), and proposed a histogram feature enhanced model based on a novel deep learning model. We designed HBB to enclose learnable weights for the attention mechanism by HAB on both histogram features and image's spatial features. In addition, a confidence interval-based metric has been applied to assess the model performance for estimating oil concentrations. Our model outperforms several other models, including ResNet-50, while keeping the running time low. The model shows an excellent overall accuracy performance and a reliable confidence assessment, especially in the most relevant range, i.e., 50 ~ 300 mg/L. When the oil content level is very high or very low, the prediction accuracy is relatively low. Still, the qualitative prediction (whether below the threshold of 50 mg/L or above the threshold of 300 mg/L) is reliable. Hence, our model is ready for practical applications as a screening tool. Our approach can also apply to analyze other images with narrow color spectra and low contrasts.

Acknowledgments. This work has been supported in part by the U.S. Geological Survey (Grant #G18AC00327), the Department of Interior Inland Oil Spill Preparedness Program (IOSPP), and the U.S. Geological Survey Environmental Health Program. In addition, we like to thank Xinrui Yang for their technical assistance. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Data Availability. Data and source code of this work are available at https://github.com/ minoriwww/waterquality.

References

Ayaz, M., Ammad-Uddin, M., Sharif, Z., Mansour, A. and Aggoune, E.H.M. (2019). Internet-of-Things (IoT)-based smart agriculture: toward making the fields talk. *IEEE Acess*. 7, 129551-129583. https://doi.org/10.1109/ACCESS.2019.2932609

Chen, L., Zhang, H.W., Xiao, J., Nie, L.Q., Shao, J., Liu, W. and Chua, T-S. (2017). Sca-cnn: Spatial and channel-wise attention in convolutional networks for image captioning. *IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 6298-6306. https://doi.org/10.1109/CVPR.2017.667

Choyke, S. and Ferguson, P.L. (2019). Molecular characterization of nonionic surfactant components of Corexit® 9500 oil spill dispersant by high-resolution mass spectrometry. *Rapid Commun. Mass Spectrom.* 33(22), 1683-1694. https://doi.org/10.1002/rcm.8512

USGS (2010). *A method for quantitative mapping of thick oil spills using imaging spectroscopy*. United States Geological Survey (USGS) Open-File Report 2010-1167.

Dutta, S. (2019). Point of care sensing and biosensing using ambient light sensor of smartphone: Critical review. *Trends Analyt. Chem*. 110, 393-400. https://doi.org/10.1016/j.trac.2018.11.014

Fantasia, J.F., Hard, T.M., and Ingrao, H.C. (1971). *Investigation of oil fluorescence as a technique for the remote sensing of oil spills* (No. DOT-TSC-CG-71-07). United States. Coast Guard..

Fingas, M. (2014). *Handbook of oil spill science and technology*. John Wiley & Sons, 1-693 https://doi.org/10.1002/9781118989982

Forth, H.P., Mitchelmore, C.L., Morris, J.M. and Lipton, J. (2017). Characterization of oil and water accommodated fractions used to conduct aquatic toxicity testing in support of the Deepwater Horizon oil spill natural resource damage assessment: *Environ. Toxicol. Chem*. 36(6), 1450-1459. https://doi.org/10.1002/etc.3672

Ghosh, T., Fattah, S.A., Shahnaz, C., Kundu, A.K. and Rizve, M.N. (2015). Block based histogram feature extraction method for bleeding detection in wireless capsule endoscopy. *TENCON 2015-2015 IEEE Region 10 Conference*, Macao, 1-4. https://doi.org/10.1109/TENCON.2015.7373186

Granica, M. and Tymecki, Ł. (2019). Analytical aspects of smart phone fluorometric measurements. *Talanta*. 197, 319-325. https://doi.org/10.1016/j.talanta.2019.01.032

Gunda, N.S.K., Gautam, S.H. and Mitra, S.K. (2019). Artificial intelligence based mobile application for water quality monitoring. *J. Electrochem. Soc*. 166(9), B3031. https://doi.org/10.1149/2.0081909jes

Gustafsson, F.K., Danelljan, M., Bhat, G. and Schön, T.B. (2019). DC TD: Deep conditional target densities for accurate regression. *Computer Vision-ECCV 2020, European Conference on Computer Vision (ECCV)*, online, 325-343. https://doi.org/10.48550/arXiv.1909.12297

Han, L., Yang, G.J., Dai, H.Y., Xu, B., Yang, H., Feng, H.K., Li, Z.H. and Yang, X.D. (2019). Modeling maize above-ground biomass based on machine learning approaches using UAV remote-sensing data. *Plant Methods*. 15(1), 1-19. https://doi.org/10.1186/s13007-019-0394-z

He, K.M., Zhang, X.Y., Ren, S.Q. and Sun, J. (2016). Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 770-778. https://doi.org/10.1109/CVPR.2016.90

Hossain, M.A., Canning, J. and Yu, Z.K. (2019). Fluorescence-based determination of olive oil quality using an endoscopic smart mobile spectrofluorimeter. *IEEE Sens. J*. 20(8), 4156-4163. https://doi.org/10.1109/JSEN.2019.2961419

Hussain, M.A., Hamarneh, G. and Garbi, R. (2019). ImHistNet: Learnable image histogram based DNN with application to noninvasive determination of carcinoma grades in CT scans. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Shenzhen, 11769, 130-138. https://doi.org/10.1007/978-3-030-32226-7-15

Incardona, J.P., Swarts, T.L., Edmunds, R.C., Linbo, T.L., Aquilina-Beck, A., Sloan, C.A., Gardner, L.D., Block, B.A. and Scholz, N.L. (2013). Exxon Valdez to Deepwater Horizon: Comparable toxicity of both crude oils to fish early life stages. *Aquat. Toxicol*. 142, 303-316. https://doi.org/10.1016/j.aquatox.2013.08.011

Joye, S.B., MacDonald, I.R., Leifer, I. and Asper, V. (2011). Magnitude and oxidation potential of hydrocarbon gases released from the BP oil well blowout. *Nat. Geosci*. 4(3), 160-164. https://doi.org/10.1038/ngeo1067

Keizer, P.D. and Gordon Jr, D.C. (1973). Detection of trace amounts of oil in sea water by fluorescence spectroscopy. *J. Fish. Res. Board Can*. 30(8), 1039-1046. https://doi.org/10.1139/f73-175

Kingma, D.P. and Ba, J. (2014). Adam: A method for stochastic optimization. *3rd International Conference for Learning Representations*, San Diego. https://doi.org/10.48550/arXiv.1412.6980

Krizhevsky, A. and Hinton, G. (2009). *Learning multiple layers of features from tiny images*. University of Toronto.

Larsson, G., Maire, M. and Shakhnarovich, G. (2017). Colorization as a proxy task for visual understanding. *IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, 6874-6833. https://doi.org/10.1109/CVPR.2017.96

Lathuilière, S., Mesejo, P., Alameda-Pineda, X. and Horaud, R. (2019). A comprehensive analysis of deep regression. *IEEE Trans. Pattern Anal. Mach. Intell*. 42(9), 2065-2081. https://doi.org/10.1109/TPAMI.2019.291 0523

Latif, A., Rasheed, A., Sajid, U., Ahmed, J., Ali, N., Ratyal, N. I., Zafar, B., Dar, S.H. and Khalil, T. (2019). Content-based image retrieval and feature extraction: a comprehensive review. *Math. Probl. Eng*. 2019, 9658350. https://doi.org/10.1155/2019/9658350

LeCun, Y., Bengio, Y. and Hinton, G. (2015). Deep learning. *Nature*. 521, 436-444. https://doi.org/10.1038/nature14539

Li, K.L., Huang, G.H., Wang, S., Baetz, B. and Xu, W.H. (2022). A stepwise clustered hydrological model for addressing the temporal autocorrelation of daily streamflow in irrigated watersheds. *Water Resour. Res*. 58(2), e2021WR031065. https://doi.org/10.1029/2021WR031065

McCracken, K.E., Tat, T., Paz, V. and Yoon, J.Y. (2017). Smartphone-based fluorescence detection of bisphenol A from water samples. *RSC Adv*. 7(15), 9237-9243. https://doi.org/10.1039/C6RA27726H

McGonigle, A.J.S., Wilkes, T.C., Pering, T.D., Willmott, J.R., Cook, J.M., Mims, F.M. and Parisi, A.V. (2018). Smartphone spectrometers. *Sensors*. 18(1), 223. https://doi.org/10.3390/s18010223

Morikawa, C., Kobayashi, M., Satoh, M., Kuroda, Y., Inomata, T., Matsuo, H., Miura, T. and Hilaga, M. (2021). Image and video processing on mobile devices: a survey. *Vis Comput*. 37(12), 2931-2949. https://doi.org/10.1007/s00371-021-02200-8

Nam, S. and Joo Kim, S. (2017). Modelling the scene dependent imaging in cameras with a deep neural network. *IEEE International Conference on Computer Vision*, Venice. 1717-1725. https://doi.org/10.1109/ICCV.2017.190

Niu, Z.X., Zhou, M., Wang, L., Gao, X.B. and Hua, G. (2016). Ordinal regression with multiple output cnn for age estimation. *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, 4920-4928. https://doi.org/10.1109/CVPR.2016.532

Nixon, Z. and Michel, J. (2015). Predictive modeling of subsurface shoreline oil encounter probability from the Exxon Valdez oil spill in Prince William Sound, Alaska. *Environ. Sci. Technol*. 49(7), 4354-4361. https://doi.org/10.1021/es502579u

NOAA (2020). Data Integration Visualization Exploration and Reporting (DIVER). Deepwater horizon natural resource and damage assessment data. https://www.diver.orr.noaa.gov/deepwater-horizon-nrda-data (accessed October 1, 2002).

NOAA Hazmat (2016). *Open water oil identification job aid for aerial observation. Office of Response and Restoration Job Aid. Version 3*, https://response.restoration.noaa.gov/sites/default/files/OWJA-2016.pdf (accessed January 23, 2022).

Ostertagová, E. (2012). Modelling using polynomial regression. *Procedia Eng*. 48, 500-506. https://doi.org/10.1016/j.proeng.2012.09.545

Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Li, Z.M., Desmaison, A., Antiga, L. and Lerer, A. (2017). Automatic differentiation in pytorch. *31st Conference on Neural Information Processing Systems (NIPS 2017)*, Long Beach.

Pramanik, P.K.D., Sinhababu, N., Mukherjee, B., Padmanaban, S., Maity, A., Upadhyaya, B.K., Holm-Nielsen, J.B. and Choudhury, P. (2019). Power consumption analysis, measurement, management, and issues: a state-of-the-art review of smartphone battery and energy usage. *IEEE Acess*. 7, 182113-182172. https://doi.org/10.1109/ACCESS.2019.2958684

Rebetez, J., Satizábal, H.F., Mota, M., Noll, D., Büchi, L., Wendling, M., Cannelle, B., Pérez-Uribe, A. and Burgos, S. (2016). Augmenting a convolutional neural network with local histograms - a case study in crop classification from high-resolution UAV imagery. *24th European Symposium on Artificial Neural Networks*, Bruges, 515-520.

Reddy, C.M., Arey, J.S., Seewald, J.S., Sylva, S.P., Lemkau, K.L., Nelson, R.K., Carmichael, C.A., McIntyre, C.P., Fenwick, J., Ventura, G.T., Van Mooy, B.A.S. and Camilli, R., (2012). Composition and fate of gas and oil released to the water column during the Deepwater Horizon oil spill. *PNAS*. 109(50), 20229-20234. https://doi.org/10.1073/pnas.1101242108

Russell, A.L., Martin, D.P., Cuddy, M.F. and Bednar, A.J. (2016). A rapid, fluorescence-based field screening technique for organic species in soil and water matrices. *Bull. Environ. Contam. Toxicol*. 96(6), 773-778. https://doi.org/10.1007/s00128-016-1771-1

Sedighi, V. and Fridrich, J. (2017). Histogram layer, moving convolutional neural networks towards feature-based steganalysis. *Electronic. Imaging*. 2017(7), 50-55. https://doi.org/10.2352/ISSN.2470-1173.2017.7.MWSF-325

Seo, D., Oh, S., Shin, S., Lee, M., Hwang, Y. and Seo, S. (2019). Smart phone compatible on-site fluorescence analyzer for spilled crude oil

based on CMOS image sensor. *Sens. Actuators B Chem*. 289, 93-99. https://doi.org/10.1016/j.snb.2019.03.086

Song, C.M. (2020) Hydrological image building using curve number and prediction and evaluation of runoff through convolution neural network. *Water*. 12(8), 2292. https://doi.org/10.3390/w12082292

Suter II, G.W. (2016). *Ecological risk assessment*. CRC press, https://doi.org/10.1201/9781420012569

Tolias, G., Sicre, R. and Jégou, H. (2015). Particular object retrieval with integral max-pooling of CNN activations. arXiv preprint arXiv: 1511.05879. https://doi.org/10.48550/arXiv.1511.05879

USEPA. (2003). *Characteristics of spilled oils, fuels, and petroleum products: composition and properties of selected oils*. EPA/600/R-03/072 (NTIS PB2004-101169).

Wang, Z., Li, H.S., Ouyang, W.L. and Wang, X.G. (2016). Learnable histogram: Statistical context features for deep neural networks. *European Conference on Computer Vision*, Amsterdam. 9905, 246-262. https://doi.org/10.1007/978-3-319-46448-0-15

Wei, Q.S., Qi, H.F., Luo, W., Tseng, D., Ki, S.J., Wan, Z., Göröcs, Z., Bentolila, L.A., Wu, T.T., Sun, R. and Ozcan, A. (2013). Fluorescent imaging of single nanoparticles and viruses on a smart phone. *ACS Nano*. 7(10), 9147-9155. https://doi.org/10.1021/nn4037706

White, H.K., Conmy, R.N., MacDonald, I.R. and Reddy, C.M. (2016). Methods of oil detection in response to the Deepwater Horizon oil spill. *Oceanography*. 29(3), 76-87. https://doi.org/10.5670/oceanog.2016.72

Wu, J.S., Qiu, S.J., Zeng, R., Senhadji, L. and Shu, H.Z. (2017). PCA-Net for color image classification in various color spaces. *International Conference on Cloud Computing and Security*, Nanjing, 10603, 494-505. https://doi.org/10.1007/978-3-319-68542-7_42

Yin, Z.X, Abel, A., Zhang, X.P. and Luo, B. (2016). Reversible data hiding in encrypted image based on block histogram shifting. *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai, 2129-2133. https://doi.org/10.1109/ICASSP.2016.7472053

Yu, K.M., Lu, Z.D. and Stander, J. (2003). Quantile regression: applications and current research areas. *J. R. Stat. Soc. Ser. D-Stat*. 52(3), 331-350. https://doi.org/10.1111/1467-9884.00363

Zhang, H.M., Gao, W., Chen, X.L. and Zhao, D.B. (2005). Learning informative features for spatial histogram-based object detection. *2005 IEEE International Joint Conference on Neural Networks*, Montreal, 3, 1806-1811. https://doi.org/10.1109/IJCNN.2005.1556154

Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., and Fu, Y. (2018). Image super-resolution using very deep residual channel attention networks. *European Conference on Computer Vision*, Munich, 294-310. https://doi.org/10.48550/arXiv.1807.02758