Satellite Oil Spill Detection Using Artificial Neural Networks

Article *in* IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing · December 2013

DOI: 10.1109/JSTARS.2013.2251864

3 authors:

Suman Singha
Danish Meteorological Institute
127 PUBLICATIONS **1,095** CITATIONS

SEE PROFILE

Tim Bellerby
University of Hull
34 PUBLICATIONS **1,412** CITATIONS

SEE PROFILE

Olaf Trieschmann
44 PUBLICATIONS **808** CITATIONS

SEE PROFILE

Satellite Oil Spill Detection Using Artificial Neural Networks

Suman Singha, Tim J. Bellerby, and Olaf Trieschmann

***Abstract*—Oil spills represent a major threat to ocean ecosystems and their health. Illicit pollution requires continuous monitoring and satellite remote sensing technology represents an attractive option for operational oil spill detection. Previous studies have shown that active microwave satellite sensors, particularly Synthetic Aperture Radar (SAR) can be effectively used for the detection and classification of oil spills. Oil spills appear as dark spots in SAR images. However, similar dark spots may arise from a range of unrelated meteorological and oceanographic phenomena, resulting in misidentification. A major focus of research in this area is the development of algorithms to distinguish oil spills from 'look-alikes'. This paper describes the development of a new approach to SAR oil spill detection employing two different Artificial Neural Networks (ANN), used in sequence. The first ANN segments a SAR image to identify pixels belonging to candidate oil spill features. A set of statistical feature parameters are then extracted and used to drive a second ANN which classifies objects into oil spills or look-alikes. The proposed algorithm was trained using 97 ERS-2 SAR and ENVSAT ASAR images of individual verified oil spills or/and look-alikes. The algorithm was validated using a large dataset comprising full-swath images and correctly identified 91.6% of reported oil spills and 98.3% of look-alike phenomena. The segmentation stage of the new technique outperformed the established edge detection and adaptive thresholding approaches. An analysis of feature descriptors highlighted the importance of image gradient information in the classification stage.**

***Index Terms*—Artificial neural network, 'CleanSeaNet', image segmentation, marine pollution, oil spill detection, synthetic aperture radar.**

I. Introduction

OIL spills resulting from intentional or unintentional releases into oceanic or coastal waters represent a major threat to marine ecosystems and the adverse effect of oil spills on these ecosystems is subject of considerable political, environmental and scientific concern. The oil spill in the Gulf of Mexico, resulting from an explosion on the Deepwater Horizon oil rig was widely reported, with an estimated five million barrels of crude oil relished into the Gulf's ecosystem [1]. It was found the that large quantities of oil spilled during the earlier 1989 Exxon Valdez disaster can still be found beneath gravel beaches in Alaska [2]. However, about half of the total oil spills in the marine environment come from operative discharges by shipping and in most of these cases the discharges are illicit. These illegal discharges are not limited to oil tankers, and many classes of ship are suspected of being responsible [3]. During the last few decades, maritime transportation has grown steadily and the quantity of illicit oil discharges has grown with the volume of traffic. Europe provides the world's largest market for crude oil imports, representing about one third of the world's total. 90% of oil and refined products are transported to and from the continent by sea; unfortunately, some of this oil makes its final way into the sea [3]. It has been estimated that 457,000 tonnes of oil are released by shipping into the ocean every year [3]. European Maritime Safety Agency's (EMSA) oil spill monitoring and vessel detection service, 'CleanSeaNet' is continuously monitoring the EU waters since 2007 and decline in number of oil spill in EU waters has been observed through CleanSeaNet service from 10.77 possible spills identified per million km^2 in 2008 to 5.68 per million km^2 in 2010 [4] (Fig. 1).

The construction of a cost effective oil spill detection system has been the subject of research for approximately two decades [5]–[11] preparing the ground for operational services responding to the maritime legislation. Methods for the automatic detection and tracking of oil spills and illegal oil discharges are of fundamental importance to improve the efficiency of surveillance systems. Satellite remote sensing techniques offer the advantage of being able to observe events in remote and inaccessible areas on a large scale and thus represent an attractive source of information for detecting the location and extent of oil spills [5], [12]. Moreover, the rate and direction of movement of oil may be obtained through multi-temporal imagery thereby assisting in drift prediction modeling that may be used to facilitate clean-up operations and warning systems [6].

Early applications of remote sensing for oil spill detection only used data from airborne visible (VIS) and infrared (IR) sensors. However, these sensors suffer a number of disadvantages in this context, including the absence of any clear discriminating feature between oil spills and the surrounding sea surface and the unavailability of data during the night (VIS) or in bad weather conditions [12]. These disadvantages do not apply to active radar sensors, which are capable of providing data at night and under adverse weather conditions. Sideward Looking Airborne Radars (SLAR) are usually mounted on aircraft while Synthetic Aperture Radar (SAR) sensors can be operated on both aircraft and satellites with comparable resolution. Satellite remote sensing allows forming the basis of an extremely viable

Manuscript received September 30, 2012; revised January 22, 2013; accepted March 04, 2013. Date of publication March 27, 2013; date of current version November 21, 2013. This work was supported in part by the Commonwealth Scholarship Commission and in part by University of Hull.

S. Singha and T. J. Bellerby are with Earth Observation Science Laboratory, Department of Geography, Environment and Earth Sciences, University of Hull, Hull HU6 7RX, U.K. (e-mail: S.Singha@2009.hull.ac.uk).

O. Trieschmann is with Satellite Based Monitoring Services, European Maritime Safety Agency, Lisbon 1249-206, Portugal.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2013.2251864

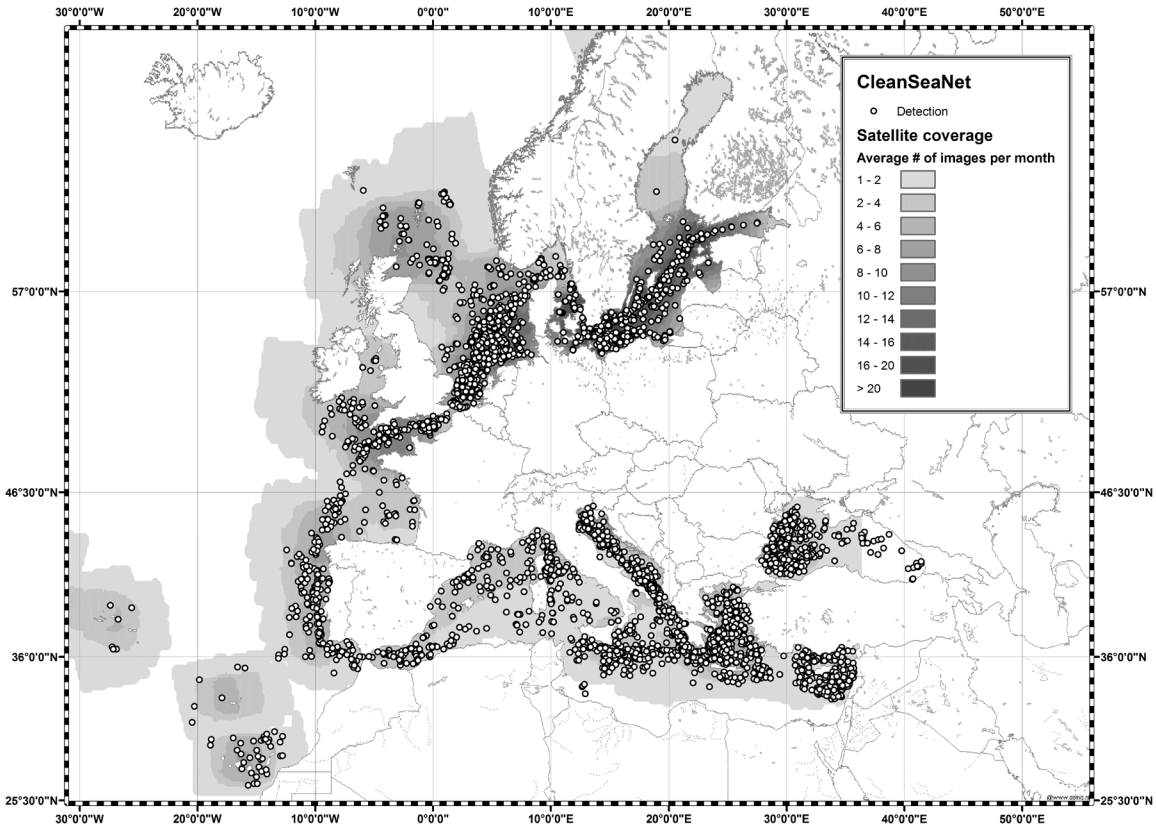

Fig. 1. 'CleanSeaNet' oil spill location map in EU waters occurred from Jan 2009 to Jan 2011 (311 checked and 154 Conformed within 3 hours of satellite acquisition). Dark patches showing operational coverage area.

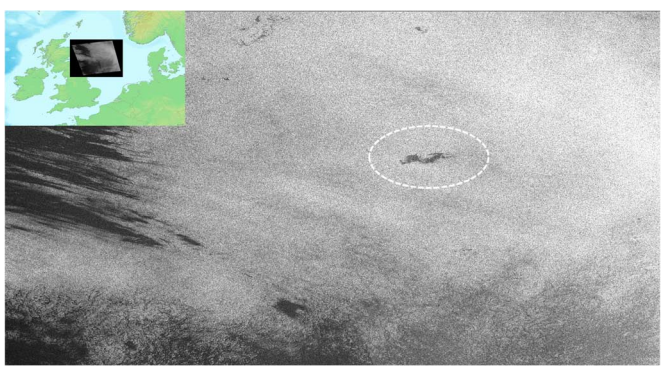

Fig. 2. Example of dark spots on a RADARSAT-2 ScanSARNarrow Beam (SNB-1F) image acquired on 27th of March 2012 at 17:28:59 over North Sea near the coast north-east England and Scotland (Right hand side inset). The bright white circle showing the potential oil spill spots reported by EMSA along with other look-alike spots. (CleanSeaNet Image ID: 21283, Radarsat-2 image CSA/MDA/EMSA 2012).

and useful tool for detecting oil spills, particularly over large areas in a marine environment.

The radar backscatter energy from a spill free area is governed mainly by Bragg scattering of surface gravity-capillary waves. Oil films decrease these waves and hence decrease the backscatter of the radar microwave radiation, causing spills to appear as dark spots in radar images while the surrounding spill-free sea remains relatively bright [13]. The best compromise between sensitivity, resolution and independence of weather effects is achieved in the radar C frequency band. In this frequency range, a minimum wind speed of 2–3 m/s creates sufficient backscatter from the oil free surface, to contrast with the reduced backscatter of oil films of thicknesses down to 10 microns [13]. Radar backscatter is also dependent on the local wind speed. The higher wind speed results in higher intensity value thus making the background bright and vice versa. All of these factors affect the statistics of background backscatter pixel values. However, if the wind speed is too high ($>$10–12 m/s) it causes the spill to disappear since the short waves receive enough energy to counterbalance the damping effect of the oil film on the ocean surface [14]. It has been observed that the wind direction relative to the plane of the incident radar wave also affects the backscatter level in a scene [13]. A crosswind (wind blowing perpendicular to the range direction) produces lower backscattering than an upwind or downwind (wind blowing along the range direction). Wind speeds ranging between 3–7 m/s have been found to be optimal oil spill detection [14]. Unfortunately, dark spots may also occur in radar images due to other meteorological or oceanographic effects, and a major challenge in the implementation of radar oil spill detection approaches is to discriminate between oil spills look-alike spots [15]–[18]. The background backscatter value within a SAR image varies significantly according to the relative position of the satellite to the backscatter location: for near range pixels the backscatter is significantly higher than for far range pixels. Moreover, a recent study [19] demonstrates the potential of polarimetry to enhance the contrast of man-made objects and oil spills against the surroundings using different combinations of SAR polarimetric channels. Fig. 1 shows an

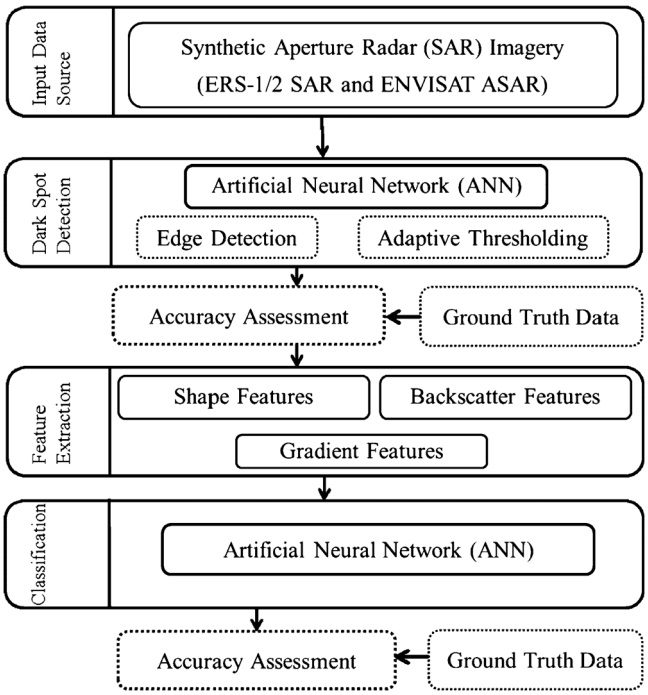

Fig. 3. A flowchart of the methodology for oil spill detection using SAR imagery. The dotted boxes indicate the workflow elements only used during training and quality assessment.

example RADARSAT-2 SAR image containing a verified oil spill where the image is already brightness corrected for the incidence angle.

A typical SAR-based oil-spill detection process consists of three stages: image segmentation, feature extraction and classification [16], [17]. The segmentation stage identifies candidate features within the imagery though a binary classification of image pixels. A feature dataset is then formed by extracting contiguous features from the segmented image and deriving a quantitative description (feature vector) of the shape and form of each feature. The final classification stage uses the feature vector information to segregate oil spills from look-alikes.

Segmentation techniques range from the relatively simple and straightforward, such as Adaptive Thresholding [18], Adaptive Level Set Thresholding [20], and Spatial Density thresholding [21] to more complex fuzzy logic based methods [22], region growing techniques [23] and multi-scale segmentation [24]. Single-input neural networks have also been employed in this context [25], [26].

The feature extraction stage is critical to obtaining accurate classifications. A wide range of feature descriptors (usually referred as "features") are reviewed in [16], [27]. These may be broadly grouped into the following categories:

i) Features based on the geometry and shape of the segmented region
ii) Features based on the backscatter values of the spot and its surroundings
iii) Textural features
iv) Contextual features

The perimeter and the area of the object are usually considered as the most common geometrical and shape features. If the oil slick is a fresh oil spill released from a moving ship (e.g., a tanker cleaning its tank), *elongatedness* appears to be an appropriate feature. It is expressed as a ratio between the width and length of the slick. Several researchers [16], [23] have shown the utility of statistical descriptors of pixel backscatter values, primarily the mean and standard deviation. Border gradients of backscattering values between the dark spot and image background also provide useful information [23], as do the respective power-to-mean ratios of oil spill spots and their surroundings [18]. Textural features are primarily based on spatial correlations among the neighbouring pixels. Contextual information includes coincident weather data, such as the speed of the wind, and 'probable cause' indicators such as distances to possible sources such as ships or offshore oil rigs [22], [28]–[31]. A range of classification methods have been employed to identify oil spills and look-alikes from vectors of feature descriptors, including statistical classifiers, including: Mahalanobis classifiers [22], neural networks [23], [25], [30], [32] and artificial intelligence fuzzy modeling systems [31]. SAR image quality has been shown to have a significant impact on the effectiveness of feature classification [33].

This paper introduces a neural network based satellite SAR oil spill detection system that employs separate neural networks for both the segmentation and classification stages. The segmentation stage uses local textural information to yield improved accuracy over established edge detection techniques, including adaptive thresholding and single input neural network approaches. The feature classification stage incorporates a range of established feature descriptors and introduces a set of textural features based on image gradients. Gradient-based features are relatively untested in this context with only gradient magnitude having been previously studied [23]. The new algorithm has been validated using a substantial dataset, comprising full-swath SAR images.

II. Methodology

A. Dataset

The new technique was trained and calibrated using 97 C-Band ERS-2 SAR and ENVISAT ASAR images obtained from the European Space Agency (ESA) and European Maritime Safety Agency (EMSA). This training and calibration dataset consisted of 183 reported oil spill spots and 720 'look-alike' features. The methodology was validated using a separate and mutually exclusive dataset incorporating 82 ENVISAT ASAR and RADARSAT-2 SAR images obtained from European Maritime Safety Agency (EMSA). This validation dataset consisted of 226 reported oil spill spots and 4670 'look-alike' features. Fig. 1 shows an example RADARSAT-2 SAR image containing a verified oil spill. All images in two datasets went through an automated image pre-processing step. This pre-processing was designed to mask out land areas and to compensate for the trend of radar backscatter from the sea induced by the increase in incidence angle at far range, which typically yields a progressive brightness reduction across the image. The preprocessing step also incorporated a histogram stretching transformation to enhance the dynamic range of pixel values that can be fully exploited by the subsequent segmentation procedure. Radar speckle reduction was also

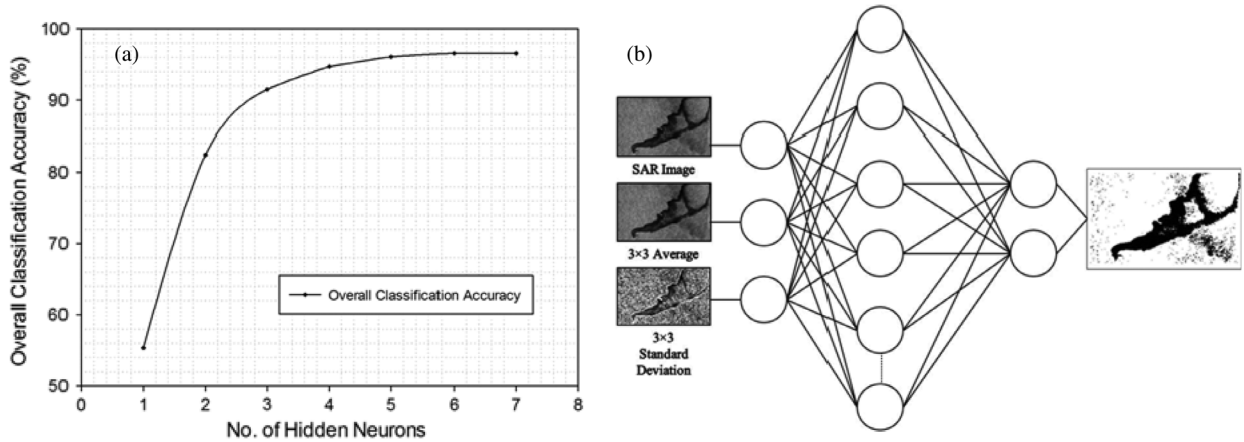

Fig. 4. Accuracy assessment and ANN architecture for segmentation.

applied using 2-D edge preserving filters with variable kernel sizes depending on the image resolution [34].

B. Overview

Fig. 3 shows a flowchart of the methodology used in this study. Pre-processed, speckle-free SAR images are converted to binary images to segregate dark spots from the background using neural network. From this segmented binary image, a total number of 14 parameters (features) are extracted for each dark spot. These feature vectors define the geometry, shape, backscatter statistics and gradient parameters for their respective spot. This vector is input to a second neural network to classify between oil spills and look alike spots.

C. Image Segmentation

Image segmentation in the new methodology is performed using a new Artificial Neural Network (ANN). This was compared against two established methodologies: i) Edge detection segmentation and ii) Adaptive Thresholding.

1) Segmentation Using ANN: Fig. 3 shows the architecture of the feed-forward back propagation neural network used for segmentation [31], [32]. The input layer consists of three neurons, corresponding to the intensity value of a given SAR image pixel, the 3×3 pixel-centered mean intensity and the 3×3 pixel-centered standard deviation. The output layer consists of two neurons corresponding to the two classes (i.e., dark spot and background). The numbers of neurons in the hidden layer was determined by trial and error [30], [35]. For each SAR image, training areas pertaining to the dark formation and background were selected. The neural network architecture was then varied by adding hidden neurons and assessing the resulting effect on the accuracy of image segmentation. Fig. 4(a) suggests that no significant changes in accuracy were observed after increasing the number of hidden neuron from six to seven and higher. Using the test data set described above, a hidden layer of six neurons was selected as optimal (Fig. 4(b)).

2) Edge Detection Segmentation: Two alternative algorithms were implemented to assess the relative performance of the ANN approach with respect to established techniques. The first of these was Edge Detection [36]. In this type of segmentation, two types of parameters must be specified: edge detection parameters and segmentation parameters. The three edge detection parameters are: *pre-smoothing* factor, *threshold* and *minimal length*. The pre-smoothing factor specifies the number of times the image is smoothed before edge detection. A differencing threshold is then employed to detect edge pixels: if the difference between intensity value of a pixel and the intensity of one of its neighboring pixels is higher than the threshold, then the pixel is considered as a candidate for edge pixel. The appropriate threshold value is selected arbitrarily and varies from image to image. Specifying a higher threshold will result in smaller number of edges, which in turn may reduce the number of segments. As the pre-processed images were not affected by the reduced backscatter trend caused by increased incidence angles, a (different) static threshold could be applied on each image. In this study, threshold values varied from 45 to 70 DN for different SAR images in training and calibration dataset with an 8 bit dynamic range.

The minimal length parameter refines the number of edges to be used for further analysis. Any edge lengths less than a minimum value, measured in pixel units, are eliminated. The segmentation parameter used for segmenting the image after edge detection is minimum value difference. *Minimal Value Difference*: Specifies the minimum value difference between neighboring segments. Specifying the higher minimum value difference results smaller number of segments. If the value difference between neighboring pixels is less than minimum value, then these are considered as a part of the same segment. Both of these parameters were manually optimized on an image by image basis.

3) Adaptive Thresholding: The second established segmentation technique employed was Adaptive Thresholding [18]. Here the detection threshold for a given pixel is set to a fixed dB below the mean value of an encompassing 5×5-pixel moving window. This technique has proven to be both robust and simple to implement. The optimal threshold separation is dependent on wind speed. In this present study, as wind speed data was unavailable, the threshold was set to 50 (manually optimized for 8-bit dynamic range) to give an optimal result.

TABLE I
SET OF FEATURE USED IN VARIOUS STUDIES

Feature Category	Features	Descriptions
Shape Features	**1.Area (A)**	The area of the dark spot measured in pixel units from a vectorized image
	2.Perimeter (P)	The total boundary length of the dark spot measured in pixel units from a vectorized image.
	3.Shape complexity: Perimeter to Area Ratio (C)	The perimeter of the image object divided by two times the square root of its area multiplied by π . C = $\frac{P}{2\sqrt{\pi A}}$
Backscatter Features	**4. Dark spot mean (μ_d)**	Mean value of the pixels belonging to dark spot.
	5. Dark spot standard deviation (σ_d)	Standard deviation value of the pixels belonging to dark spot.
	6. Backgrounds mean (μ_b)	Mean value of the pixels belonging to local background area around the dark spots.
	7. Backgrounds standard deviation (σ_b):	Standard Deviation value of the pixels belonging to local background area around the dark spots.
	8. Dark spot power to mean ratio $\left(\frac{\sigma_d}{\mu_d}\right)$	Ratio between standard deviation and mean value of the pixels belonging to the dark spot.
	9. Background power to mean ratio $\left(\frac{\sigma_b}{\mu_b}\right)$**:**	Ratio between standard deviation and mean value of the pixels belonging to the background.
Gradient Features (gradient calculation	**10. Gradient mean (μ_g),**	Mean value of the gradient.
	11. Gradient standard deviation (σ_g),	Standard deviation value of the gradient.
	12. Gradient max (G_{max})	Maximum value of the gradient.
see text)	**13.Gradient min (G_{min})**	Minimum value of the gradient.
	14. Gradient power to mean ratio $\left(\frac{\sigma_g}{\mu_g}\right)$	Ratio between gradient standard deviation and gradient mean value.

Fig. 6 shows an image segmentation using the edge detection and adaptive thresholding approaches.

D. Feature Extraction

The output of the segmentation stage is a binary image depicting dark objects and the background. The next stage of the algorithm involves the generation of a vector of features that quantitatively describe relevant characteristics of the object, including: backscattering attenuation, shape, size, texture, and boundary characteristics. The resulting vector depicts the characteristics of each object and forms the basis for computing feature descriptors.

A range of features have been suggested in the literature [18], [22]–[24], [27], [28]. Based on a review of previous studies, 14 feature parameters were selected. These are listed in Table I.

Three categories of feature were taken into account: shape, backscatter and gradient. Gradient values were calculated using a SOBEL operator [37]. This is a discrete differentiation operator, computing an approximation of the gradient of the image intensity function. At each point in the image, the result of the SOBEL operator is either the corresponding gradient vector or the norm of this vector. The SOBEL operator is based on convolving the image with a small, separable, and integer valued filter in the horizontal and vertical directions and is therefore relatively inexpensive to compute. However, this gradient approximation is relatively crude, in particular for high frequency variations in the image. The operator uses two 3 × 3 kernels which are convolved with the original image—one for horizontal changes G_x, and another for vertical G_y [31]. If $\mathbf{A}$ is the source image, and $\mathbf{G}_x$ and $\mathbf{G}_y$ are the horizontal and vertical derivative approximations, then:

$$G_Y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} \otimes A \text{ and } G_X = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} \otimes A \tag{1}$$

Where $\otimes$ denotes the 2-dimensional convolution operation.

At each point in the image, the resulting gradient approximations can be combined to give the gradient magnitude, using:

$$G = \sqrt{G_x^2 + G_y^2} \tag{2}$$

This information can also be used to calculate the gradient's direction:

$$\Theta = \arctan 2(G_y, G_x) \quad with\ \Theta\ in\ [-\pi, \pi] \tag{3}$$

For example, Θ is 0 for a vertical edge which is darker on the left side.

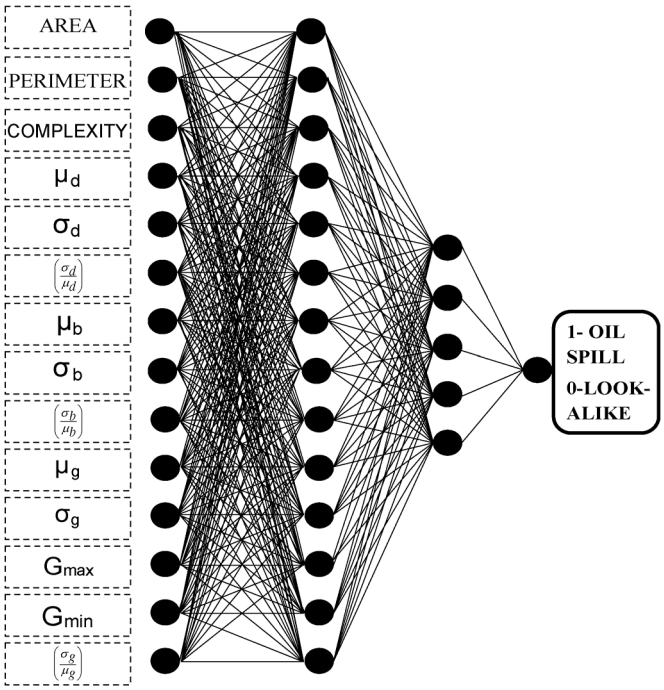

Fig. 5. Schematic diagram of the Neural Network implemented for classification between oil spills and look-alikes.

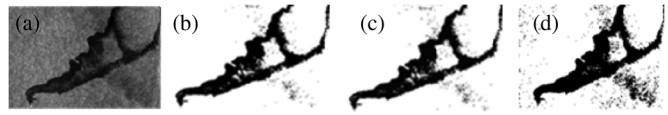

Fig. 6. Segmentation of an example SAR image (ENVISAT ASAR Wide Swath image acquisition date: 17 November 2002) (a) SAR image containing dark spot. (b) Segmentation using ANN. (c) Segmentation using adaptive thresholding. (d) Segmentation using edge detection.

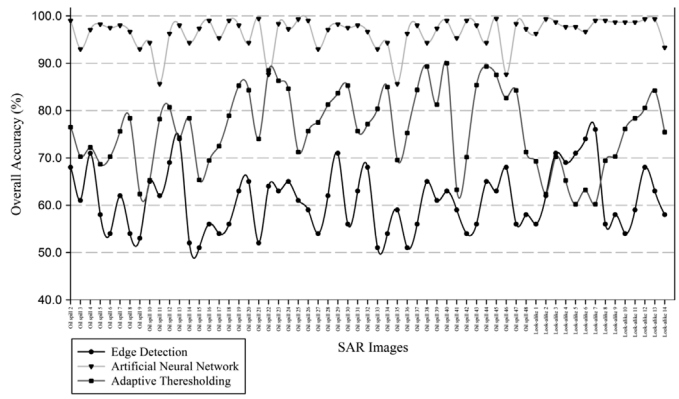

Fig. 7. Comparison of the segmentation accuracy obtained using edge detection, adaptive thresholding technique and the proposed ANN.

The gradient power to mean ration is used to represent the contrast in the gradient image in quantitative terms [37].

E. Feature Classification

In the final processing stage, a neural-network based classifier is used to distinguish oil spills from look-alikes based on the fourteen-element feature vector describing each dark object. A range of feed-forward network architectures were investigated, with a final topology of 14-14-5-1 (two hidden layers; Fig. 5) found to be the most accurate as well as efficient (see the Results section below).

Several different combinations of training set, testing set and neural network training parameters (including momentum factor, learning rate and the number of epochs) were tested in order to optimize classification accuracy.

III. Results and Discussion

The new ANN based segmentation and classification technique was implemented using the 97-image training and calibration dataset described above. To evaluate the initial, image segmentation, stage of the algorithm, two alternative segmentation techniques (edge detection and adaptive thresholding) were also applied to the calibration dataset. Fig. 6(b) shows an example output from the neural network image-segmentation stage for a dark spot. Fig. 6(c) shows the same output for adaptive theshholding and Fig. 6(d) for edge detection. The neural network approach, while not generating a perfect classification, does appear to be relatively robust against noise.

In order to assess the image segmentation stage, a manual image classification was performed on each image. Fig. 7 compares the overall segmentation accuracies using edge detection, adaptive thresholding and neural network, validated against manually segmented imagery for a subset of training and calibration dataset. The average accuracy (accuracy assessment was carried out using confusion matrix for each image) obtained using Edge Detection was 59.27% and for Adaptive Thresholding was 74.13%, whereas the average overall accuracy achieved using the neural network was 96.57%.

Feature vectors were divided into training and calibration sets. The training set contained 100 oil spills and 400 look-alike features, ensuring a balanced representation of object types as occurrence of look-alikes in an image is significantly higher than that of oil spills. The training set was constructed to represent a range of oil spill morphologies (e.g., linear, diffused) and types of look-alike spot (e.g., low wind, algae etc.). The calibration set was used to evaluate the generalization capability of the network. It contained 83 and 320 feature vector for oil spill and look-alike respectively. Training and calibration set were mutually exclusive [38]. In addition, the classification was tested by using different combinations of feature parameters to form the feature vectors. Table II and Fig. 8 compare the classification accuracies of different feature vector compositions, as determined using their respective calibration sets. As discussed below, these dependent accuracy values are higher than those obtained for the full-swath validation dataset, which is understandable since the latter contains a greater variety of artifacts.

Given that values for area and perimeter varied considerably between the available example oil spill and look-alike spots, it was assumed that those features may be unduly influencing the output for neural network classifier. It was also observed that one of the gradient feature, $G_{\min}$, was statistically very similar for both types of spots. Therefore, these three features were discarded from the training feature set and then similar training and classification stage were carried out using this reduced feature data set. In order to assess the relative contributions of various classes of feature, three further reduced data sets were

TABLE II
CLASSIFICATION ACCURACY OF NEURAL NETWORK CLASSIFIER USING DIFFERENT FEATURE SETS

Feature Set	Accuracy
All	**92.16%**
Reduced Feature Set which does not include the area, perimeter, or the gradient min	**89.70%**
Reduced Feature Set but Shape complexity excluded	**93.54%**
Reduced Feature Set but Backscatter Features excluded	**70.96%**
Reduced Feature Set but Gradient Features excluded	**67.74%**

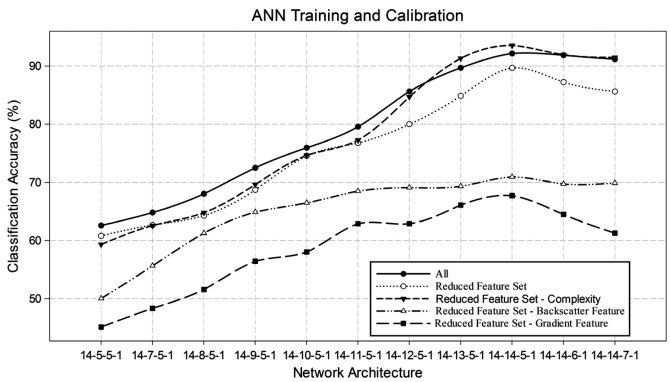

Fig. 8. Classification accuracy of the neural network classifier using different combinations of feature parameters.

employed for network training and validation, respectively excluding shape complexity, backscatter and gradient feature sets. Exclusion of complexity features actually improves the accuracy of the classification for all ANN configurations The gradient and backscatter features clearly make a strong contribution to the overall classification accuracy, indicating that these inputs are worthy of further study. In all cases the X-14-5-1 (where X is number of input neuron) network architecture proved optimal among the architectures studied and the reduced feature set with shape complexity excluded shows the best stability when varying the ANN architecture (see Fig. 8).

The complete new methodology, including segmentation, feature extraction and classification stages, was validated using the full swath validation dataset. The basis of reported oil spills in the validation dataset is manual interpretation and a brief review of this technique can be found in [27]. Fig. 9 shows one example of ENVISAT ASAR WideSwath image and 7 reported oil spill by EMSA via manual classification (Fig. 9 right-hand side inset). Similar kinds of comparison were carried out in some previous studies [14].

As these test images includes land areas, those were removed in a pre-processing stage that included land masking, radiometric normalization and filtering. The pre-processed images were then segmented followed by feature extraction using the reduced feature set (see Table II). After obtaining feature vectors for each spot, the second-stage ANN was used to classify detected dark spots. Altogether 207 out of 226 reported spills in the validation dataset were correctly identified by the new algorithm, a small percentage of reported oil spills were missed.

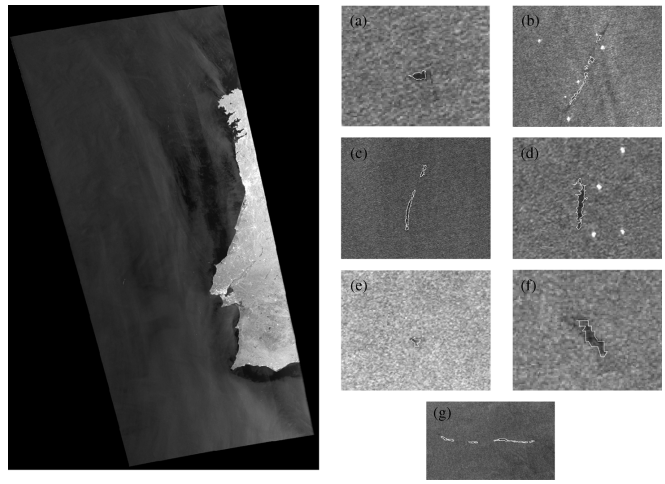

Fig. 9. Example of dark spots detected on single ENVISAT ASAR WideSwath image acquired on 12th June 2012 at 17:31:51 over Atlantic Ocean near the coast of Portugal. Image windows on the right-hand side inset showing potential oil spill spots reported by EMSA. (CleanSeaNet Image ID: 5793, ENVISAT image ESA/EMSA 2012).

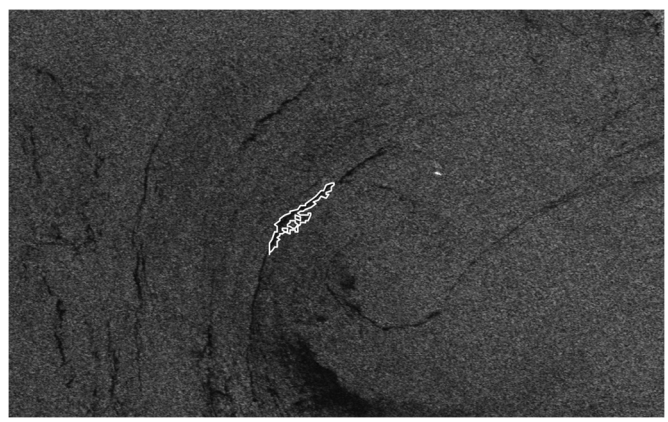

Fig. 10. Example of a look-alike identified as an oil spill (shown in bright white outline).

The potential oil spills shown in Fig. 9(a) and Fig. 9(e) was missed due to its very low contrast, although the sizes of those spills were significantly smaller compared to other linear spills present in the image. Some of the look-alike spots were identified as oil spills (false positive) and main reason was identified for this type of misclassification, presence of Algae bloom (with linear feature mimicking oil spill spots).

The presence of Algae is a very common phenomenon in the Mediterranean and Baltic seas from May to September and some of the images in validation set are from mentioned time frame. Fig. 10 shows an example of false positive where part of a linear feature due to algae bloom identified as oil spill. The error matrix shown in Table III evaluates the performance level of the proposed methodology.

There are 19 cases where the proposed algorithm disagrees with the verified detection (false negative) and 79 cases where dark spots were identified as oil spill which are not confirmed by EMSA (potential false positive). It can be observed from Table III that the proposed algorithm is successfully able to identify most of the oil spills (91.6%) and look-alikes (98.3%)

TABLE III
ERROR MATRIX USED FOR QUANTITATIVE ASSESSMENT OF THE PROPOSED ALGORITHM

	Algorithm Oil Spill	Algorithm Look-Alike
Reported Oil Spill	**207**	**19**
Reported Look-Alike	**79**	**4591**

present in the large validation dataset, this indicates the robustness of the proposed algorithm.

While it is problematic to compare different methodologies without using exactly the same data set, a non-exhaustive comparison with previous work is given to provide context. The neural network classification proposed in [23] yielded accuracies of 82% for oil spills and 90% as look-alikes while [25] yielded 91% for oil spill and 87% of look-alike. The probabilistic method of [22] gave an overall accuracy close to 80% while the Probabilistic method (statistical modelling with a rule based approach) of [11] achieved 78% for oil spill and 99% for lookalikes. The final classification accuracy achieved by the proposed methodology is thus higher than some previous related studies particularly for oil spills Moreover the proposed methodology has the capability to address operational/commercial NRT requirements and has been validated against a robust and proven operational method (manual classification).

IV. CONCLUSIONS

This paper has demonstrated a new oil spill classification system using two neural networks in sequence for image segmentation and feature classification respectively. The Performance of the neural network segmentation stage was found to better than that of either edge detection or adaptive thresholding. The classification stage correctly identified 91.6% of oil spills and 98.3% of look-alikes in a large dataset containing full-swath test images. The contribution of gradient features to the classification stage is significant and these features clearly deserve attention in future studies. Test results additionally demonstrate the ability of the proposed algorithm to segregate oil spills from 'look-alikes' in full-swath SAR imagery. Processing time for a full swath image (10–14 min on a 2.1 GHz Quad Core windows based system) is well below CleanSeaNet's critical near real time requirement which is 30 minute for a standard image swath. Future work will concentrate on widening the validation of the new technique in particular for different SAR sensors and on incorporating contextual parameters (particularly wind speed) into the analysis.

ACKNOWLEDGMENT

The authors would like to thank EMSA for the provision of data. The views, opinions, and findings contained in this report are those of the authors and should not be construed as an official EMSA or European Commission's position, policy, or decision. The authors would also like to thank Dr. Konstantinos N. Topouzelis (Department of Marine Sciences, University of the Aegean) for inspiring discussion and the 'CleanSeaNet' team for their immense support.

REFERENCES

[1] BBC News. [Online]. Available: http://www.bbc.co.uk/, Accessed During Gulf of Mexico (BP) Oil Spill Incident (May–August 2010)
[2] H. Li and M. C. Boufadel, "Long-term persistence of oil from the *Exxon Valdez* spill in two-layer beaches," *Nature Geosci.*, Jan. 2010, Published online.
[3] GESAMP, 2007, "Joint Group of Experts on the Scientific Aspects of Marine Environmental Protection (GESAMP) 2007—Report No. 75," *Estimates of Oil Entering the Marine Environment from Sea-Based Activities*, p. 61 [Online]. Available: www.imo.org
[4] O. Trieschmann and D. Reichenbach, "The integrated european satellite-based and aerial oil spill surveillance and vessel detection services," in *Interspill*, London, U.K., Mar. 2012.
[5] G. Ferraro, B. Baschek, G. Montpellier, O. Njoten, M. Perkovic, and M. Vespe, "On the SAR derived alert in the detection of oil spills according to the analysis of the EGEMP," *Marine Pollution Bull.*, vol. 60, no. 1, pp. 91–102, Jan. 2010.
[6] K. Karantzalos and D. Argialas, "Automatic detection and tracking of oil spills with level set segmentation from SAR imagery," *Int. J. Remote Sens.*, vol. 29, no. 21, pp. 6281–6296., Nov. 2008.
[7] T. I. Bern, T. Wahl, T. Anderssen, and R. Olsen, "Oil spill detection using satellite based SAR: Experience from a field experiment," in *Proc. 1st ERS-1 Symp.*, Cannes, France, Nov. 1992, pp. 829–834.
[8] A. Skøelv and T. Wahl, "Oil Spill Detection Using Satellite Based SAR," Phase 1B Competition Report Norwegian Defence Research Establishment, Tech. Rep., 1993.
[9] T. Wahl, A. Skøelv, and J. H. S. Andersen, "Practical use of ERS-1 SAR images in pollution monitoring," in *Proc. IEEE Int. Geosci. Remote Sensing Symp.*, Aug. 1994, vol. 4, pp. 1954–1956.
[10] A. H. S. Solberg, S. T. Dokken, and R. Solberg, "Automatic detection of oil spills in ENVISAT, RADARSAT and ERS SAR images," in *Proc. IEEE Int. Geosci. Remote Sensing Symp.*, Jul. 2003, vol. 4, pp. 2747–2749.
[11] A. H. S. Solberg, C. Brekke, and P. O. Husøy, "Oil spill detection in RADARSAT and ENVISAT SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, pp. 746–755, Mar. 2007.
[12] M. F. Fingas and C. E. Brown, "Review of oil spill remote sensing," *Spill Sci. Technol. Bull.*, vol. 4, no. 4, pp. 199–208, 1997.
[13] W. Alpers and H. Huhnerfuss, "Radar signatures of oil films floating on the sea surface and the Marangoni effect," *J. Geophys. Res.*, vol. 93, pp. 3642–3648, Apr. 1988.
[14] O. Garcia-Pineda, B. Zimmer, M. Howard, W. Pichel, X. Li, and I. R. MacDonald, "Using SAR images to delineate ocean oil slicks with a texture-classifying neural network algorithm (TCNNA)," *Can. J. Remote Sens.*, vol. 35, no. 5, pp. 411–421, 2009.
[15] C. Brekke, "Automatic Screening of Synthetic Aperture Radar Imagery for Detection of Oil Pollution in the Marine Environment," Unpublished Ph.D. dissertation, Faculty of Mathematics and Natural Science, Univ. Oslo, , Dec. 2007, 189pp.
[16] C. Brekke and A. H. S. Solberg, "Oil spill detection by satellite remote sensing," *Remote Sens. Environ.*, vol. 95, pp. 1–13, Jan. 2005.
[17] C. Brekke and A. H. S. Solberg, "Feature extraction for oil spill detection based on SAR images," in *Lecture Notes in Computer Science, 14th Scandinavian Conf.*, H. Kalviainen, Ed., Jun. 2005, vol. 3540, pp. 75–84, et al..
[18] A. Solberg, G. Storvik, R. Solberg, and E. Volden, "Automatic detection of oil spills in ERS SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 4, pp. 1916–1924., 1999.
[19] R. Shirvany, M. Chabert, and J.-Y. Tourneret, "Ship and oil-spill detection using the degree of polarization in linear and hybrid/compact dual-pol SAR," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 3, pp. 885–892, 2012.
[20] R. R. Ganta, S. Zaheeruddin, N. Baddiri, and R. R. Rao, "Segmentation of oil spill images with illumination-reflectance based adaptive level set model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 5, pp. 1394–1402, 2012.
[21] Y. Shu, J. Li, H. Yousif, and G. Gomes, "Dark-spot detection from SAR intensity imagery with spatial density thresholding for oil-spill monitoring," *Remote Sens. Environ.*, vol. 114, no. 9, pp. 2026–2035, Sep. 2010.

[22] B. Fiscella, A. Giancaspro, F. Nirchio, P. Pavese, and P. Trivero, "Oil spill detection using marine SAR images," *Int. J. Remote Sens.*, vol. 21, no. 18, pp. 3561–3566, Dec. 2000.
[23] F. Del Frate, A. Petrocchi, J. Lichtenegger, and G. Calabresi, "Neural networks for oil spill detection using ERS-SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2282–2287, Sep. 2000.
[24] K. Topouzelis, V. Karathanassi, P. Pavlakis, and D. Rokos, "Oil spill detection: SAR multi-scale segmentation & object features evaluation," in *Proc. 9th Int. Symp. Remote Sensing (SPIE)*, 2003, pp. 77–87.
[25] K. Topouzelis, V. Karathanassi, P. Pavlakis, and D. Rokos, "Detection and discrimination between oil spills and look-alike phenomena through neural networks," *ISPRS J. Photogramm. Remote Sensing*, vol. 62, pp. 264–270, Sep. 2007.
[26] K. Topouzelis, V. Karathanassi, P. Pavlakis, and D. Rokos, "Dark formation detection using neural networks," *Int. J. Remote Sens.*, vol. 29, no. 16, pp. 4705–4720, Sep. 2008.
[27] K. Topouzelis, "Oil spill detection by SAR images: Dark formation detection, feature extraction and classification algorithms," *Sensors*, vol. 8, pp. 6642–6659, Oct. 2008.
[28] A. Solberg, M. Indregard, and P. Clayton, "D2-report on benchmarking oil spill recognition approaches and best practice," Tech. Rep., Oceanides Project, European Commission, Archive No. 04-10225-A-Doc Contract No: EVK2-CT-2003-00177, May 2005.
[29] H. A. Espedal and T. Wahl, "Satellite SAR oil spill detection using wind history information," *Int. J. Remote Sens.*, vol. 20, no. 1, pp. 49–65, Jan. 1999.
[30] G. Calabressi, F. Del Frate, J. Lichtenegger, and A. Petrocchi, "Neural networks for the oil spill detection using ERS-SAR data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Hamburg, Germany, 28 June–2 Jul. 1999, pp. 215–217.
[31] I. Keramitsoglou, C. Cartalis, and C. Kiranoudis, "Automatic identification of oil spills on satellite images," *Environmental Modeling Softw.,* vol. 21, p. 640, 2006.
[32] S. Singha, T. J. Bellerby, and O. Trieschmann, "Detection and classification of oil spill and look- alike spots from SAR imagery using an artificial neural network," in *Proc. IEEE Int, Geosci. Remote Sens. Symp.*, Munich, Germany, pp. 4403–4406.
[33] M. Vespe and H. Greidanus, "SAR image quality assessment and indicators for vessel and oil spill detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 10, Apr. 2012.
[34] M. Vespe, M. Posada, G. Ferraro, and H. Greidanus, "Data fusion of SAR derived features and ancillary information for automatic oil spill detection," *Fresenius Environmental Bull.*, vol. 20, pp. 36–43, 2011.
[35] P. M. Atkinson and A. R. L. Tatnall, "Neural networks in remote sensing," *Int. J. Remote Sens.*, vol. 18, no. 4, pp. 699–709, Apr. 1997.
[36] M. Pathegama and Ö. Göl, "Edge-end pixel extraction for edge-based image segmentation," *Trans. Eng., Comput. Technol.*, vol. 2, pp. 213–216, 2005.
[37] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis, an Introduction*. Berlin, Germany: Springer-Verlag, 1999, .
[38] G. M. Foody and M. K. Arora, "An evaluation of some factors affecting the accuracy of classification by an artificial neural network," *Int. J. Remote Sens.*, vol. 18, no. 4, pp. 799–810, Mar. 1997.

Suman Singha received the M.Tech. degree from Indian Institute of Technology, Roorkee, India, in 2009, and the M.Sc. degree in remote sensing from the University of Hull, Hull, U.K., in 2010, respectively.

In September 2010, he joined the NERC Earth Observation Science Laboratory, Department of Geography Environment and Earth Sciences, University of Hull, U.K., as a doctoral candidate, where he developing models for identifying oil spills from Synthetic Aperture Radar. Main objective of his work is to distinguish oil spill from other meteorologically produced 'look-alike' spots. He has been closely working with European Space Agency (ESA) and European Maritime Safety Agency (EMSA) since early 2011.

Tim J. Bellerby received the B.Sc. degree in applied mathematics from the University of Warwick, Warwick, U.K., in 1986 and the Ph.D. degree in geophysics from the University of Sheffield, Sheffield, U.K., in 1991.

He is Co-Director of the Earth Observation Science Laboratory and a Senior Lecturer at the Department of Geography, University Of Hull, U.K. He has formerly been a Research Associate at the University of Bristol Centre for Remote Sensing, U.K., and a contractor to the United States National Weather Service. He has served as a consultant to the United Nations Food and Agricultural Organization. His research interests include the application of artificial intelligence and machine vision approaches to satellite cloud monitoring and precipitation estimation and the development of ensemble and probabilistic representations of uncertainty in precipitation estimates. He has an interest in monitoring and forecasting the River Nile.

Olaf Trieschmann is presently with the European Maritime Safety Agency, where he is heading the 'CleanSeaSet' service.

He is a senior expert on remote sensing and oil spill detection, was founding chairman of EGEMP (European Group of Experts on Satellite Monitoring of Sea-based Oil Pollution), member of the German delegation to the international GEO/GEOSS (Global Earth Observation System of Systems) initiative and member of the GMES (Global Monitoring of the Environment and Security) "Marine Core Service Implementation Group". Dr. Trieschmann received his diploma and PhD in Physics from the University Karlsruhe, Germany. He started his professional career in remote sensing of the upper atmosphere in the framework of the ozone hole and global warming research activities. Since 2000 he is working on earth observation systems and in particular on oil spill detection technologies.