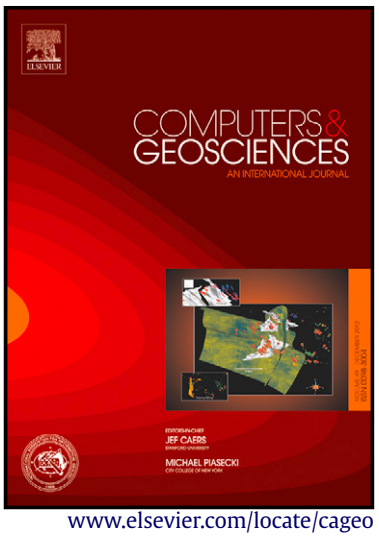

www.elsevier.com/locate/cageo

Automatic decision support system based on sar data for oil spill detection

David Mera, José M. Cotos, José Varela-Pet, Pablo G. Rodríguez, Andrés Caro

PII: S0098-3004(14)00181-2
DOI: http://dx.doi.org/10.1016/j.cageo.2014.07.015
Reference: CAGEO3413

To appear in: *Computers & Geosciences*

Received date: 28 November 2013
Revised date: 22 July 2014
Accepted date: 23 July 2014

Cite this article as: David Mera, José M. Cotos, José Varela-Pet, Pablo G. Rodríguez, Andrés Caro, Automatic decision support system based on sar data for oil spill detection, *Computers & Geosciences,* http://dx.doi.org/10.1016/j.cageo.2014.07.015

Automatic decision support system based on SAR data for oil spill detection

David Mera[a,b], José M. Cotos[b], José Varela-Pet[c], Pablo G. Rodríguez[d], Andrés Caro[d]

[a] *Laboratory of Data Intensive Systems and Applications, Faculty of Informatics, Masaryk University, Brno, Czech Republic.*
[b] *Computer Graphics and Data Engineering, Centro de Investigación en Tecnoloxías da Información, University of Santiago de Compostela, Santiago de Compostela, Spain.*
[c] *Systems Laboratory, Technological Research Institute, University of Santiago de Compostela, Santiago de Compostela, Spain.*
[d] *Department Engineering of Informatic and Telematic Systems, University of Extremadura, Escuela Politécnica, Cáceres, Spain.*

Abstract

Global trade is mainly supported by maritime transport, which generates important pollution problems. Thus, effective surveillance and intervention means are necessary to ensure proper response to environmental emergencies. Synthetic Aperture Radar (SAR) has been established as a useful tool for detecting hydrocarbon spillages on the oceans surface. Several Decision Support Systems have been based on this technology. This paper presents an automatic oil spill detection system based on SAR data which was developed on the basis of confirmed spillages and it was adapted to an important international shipping route off the Galician coast (northwest Iberian Peninsula).

Email addresses: davidmeraperez@gmail.com (David Mera), manel.cotos@usc.es (José M. Cotos), jose.varela.pet@usc.es (José Varela-Pet), pablogr@unex.es (Pablo G. Rodríguez), andresc@unex.es (Andrés Caro)

The system was supported by an adaptive segmentation process based on wind data as well as a shape oriented characterization algorithm. Moreover, two classifiers were developed and compared. Thus, image testing revealed up to 95.1% candidate labeling accuracy. Shared-memory parallel programming techniques were used to develop algorithms in order to improve above a 25% of the system processing time.

Keywords:

Decision Support System, Oil spills, SAR, Adaptive threshold, Shape characterization

1. Introduction

Oceans are essential both for life and for every society's economy. International trade is mainly supported by maritime transport which represents around 80% of global trade by volume and over 70% by value [1]. This intensive traffic along Exclusive Economic Zones (EEZ) of countries generates important pollution problems such as oil spills. Contrary to what is mainly accepted, only 7% of oil spills come from catastrophes like tanker and oil platform accidents. Half of the total oil pollution can be attributed to operational discharges from vessels (usually, the cleaning of ship bilges) [2].

Oil spills affect coasts and marine life generating ecological as well as economic losses. Thus, surveillance agencies should have adequate Decision Support Systems (DSSs) to manage intervention means as well as to ensure proper response to environmental emergencies.

Radar has proved to be a useful tool for detecting hydrocarbon spillages. This technology consists in an active detection system assembled in a platform which basically sends microwave beams to the surface. Beams are reflected and backscattered from the surface so part of the energy is collected again by the radar antenna. The intensity of the signal received by the antenna is measured and recorded for later use in the image construction of the studied area. Due to their backscattering specular behavior, oil spills appear as regions with less brightness in radar images, just like other natural phenomena [3] such as low wind, grease ice and upwelling.

Traditionally, ships and aircrafts equipped with specific tools, like radar systems, have been widely used as surveillance means. Their disadvantages, such as local coverage, high cost and dependency on meteorological conditions make them a limited solution.

Radar devices aboard satellites improve coverage, reduce cost and provide day and night imaging capability regardless of weather conditions. Due to radar fundamentals [4], unfeasible antenna dimensions would be necessary to get useful resolutions. This handicap is solved by using a special type of radar called Synthetic Aperture Radar (SAR). Its operational principle is based on collecting several measurements of the same target location from different points during the satellite pass. This process simulates an antenna aperture equivalent to the distance between the first and last measured point, and can produce image resolutions of a few meters.

Traditionally, SAR images have been analyzed by human operators. Ba-

sically, they localize dark areas and try to distinguish natural phenomena, called look-alikes, from hydrocarbons. Literature describes many efforts to develop automated and semi-automated DSSs which avoid depending on the operator's experience and improve the system's accuracy as well as its processing time. A concise state-of-the-art is offered by Brekke and Solberg [5] and Topouzelis [6]. Apart from pre-processing steps, most documented oil spill detection systems have 3 common phases:

- Dark spot segmentation: since backscattering is lower in oil spills than in clean areas, several threshold techniques based on simple [7] as well as adaptive thresholding [8] [9] have been widely used. Also, other techniques have been traditionally applied such as Artificial Neural Networks (ANN) [10] [11] [12], methods based on textures [13] [14] and techniques focused on edge detection [15] [16].

- Feature extraction: though there are some research works focused on finding a valuable set of characteristics to describe the segmented spots [17] [18], there are not systematic studies about characteristics and their influence on the classification phase.

- Spot classification: relevant literature describes many options and techniques to develop classifiers which analyze candidate characteristics to distinguish look-alikes from oil spills. ANNs [9] [11] [12] as well as statistical classifiers [19] [8] [20] [7] are some of the approaches most largely used.

A reliable detection procedure is the basis for a complete surveillance system since its outputs can be used to forecast the spillage's evolution [21] [22] and to backtrack it to find its source.

Regarding the study location, Galicia, located in the northwest of the Iberian Peninsula, is a region whose economy and way of life are closely related to the ocean. Due to its geographic location, Galicia bears an intensive maritime traffic and it has suffered important oil tanker catastrophes such as Urquiola in 1976, Aegean Sea in 1992 and Prestige in 2002. Besides these fairly infrequent events, oil spills derived from routine maintenance operations are regularly discovered. Specifically, most of them are located around the Finisterre Traffic Separation Scheme (FTSS), shown in Figure 1, which is used to regulate the maritime traffic.

Spanish Maritime Safety and Rescue Agency (SASEMAR) is the public authority responsible for monitoring the Spain EEZ including the Galician coast. SASEMAR's surveillance procedures are based on flight missions as well as limited SAR detection reports supplied by European Maritime Safety Agency (EMSA) via an agreement with the Spanish government. According to the 2012 SASEMAR report, 146 oil spills were detected by flight missions, whereas 102 were revealed by satellite reports.

In this paper, we describe an automatic oil spill detection system based on SAR imaging which makes the best of a previously published oil spill segmentation algorithm (refered to here as the "previous paper" [23]). The present work shows several improvements which were done to the segmentation pro-

cess and describes the remaining system phases, including the development and comparison of two classifiers. We encourage readers to read the previous paper for a deeper context.

2. Datasets

A database with 47 Advanced Synthetic Aperture Radar (ASAR) products from the Envisat was used to develop the detection system. Main product characteristics are shown in Table 1. Specifically, the Wide Swath Mode (WSM) product was selected for oil spill detection because it provides large coverage and enough resolution for finding small-medium spillages. Furthermore, the ocean backscatter is expected to be above the noise floor under most conditions [24].

Table 1: Main characteristics of the dataset products.

Satellite	Sensor	Mode	Angle range	Band	Polarization	Coverage	Resolution
Envisat	ASAR	Width Swath	$15° - 45°$	C-Band	Vertical-Vertical	405 km	150 m (75 m resampled)

Database ASAR images cover the Galician coast from 2007 to 2011 and most of them are centered on the FTSS, which is the core of the Region Of Interest (ROI). This collection was obtained by way of an agreement with the European Space Agency (ESA). Images can be split into two subsets: on the one hand, data from 2007 to 2008 contain oil spill candidates located by EMSA experts. These candidates were labeled with a pollution probability (low, medium, high). On the other hand, images from 2009 to 2011 have confirmed oil spills detected by SASEMAR aircrafts missions.

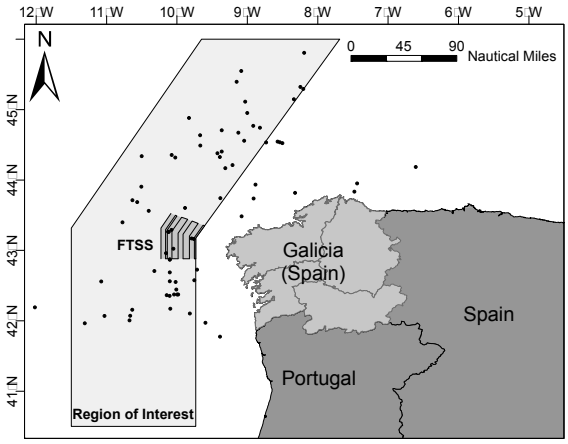

Figure 1: ROI which contains the FTSS. Black dots symbolize spillages of the database.

A subset of images captured in 2011 was saved for testing purposes, while the remainder were used to develop the oil spill detection system.

3. Methodology

The described oil spill detection system comprises three main phases: initially, a segmentation process highlights all the pollution candidates over the image background; after that, a characterization process is applied to calculate a set of characteristics for each segmented candidate; finally, a classification process analyzes every set to classify the associated candidate either as an oil spill or a false positive (look-alike). Moreover, both pre-processing and post-processing steps are carried out. On the one hand, the pre-processing step is used to tailor the dataset for detection algorithms. On the other hand, a post-processing process is applied to system outputs in order to improve the shape of detected oil spills. A detail of these processes

is shown next.

3.1. Segmentation

The segmentation process is based on a previously published Adaptive Threshold (AT) algorithm [23], which exposed the relationship between SAR image intensity values and both wind intensity and satellite Incidence Angles (IAs). Wind data were estimated by the CMOD5 model [25], which is a C-band geophysical model function. It relates the radar backscatter from roughened sea surface to wind speed and direction.

One thousand six hundred and fifty-two pixels were carefully sampled from the oil spill dataset to get measurements affected by different IAs and different wind speeds. Thus, both IA and wind speed values were collected for every sampled pixel. The resulting dataset was clustered according to both values and an intensity upper value was calculated for each clustered subset. The current version of the segmentation process is more conservative than the previous one and it calculates the upper value within every subset as the mean intensity plus the double of the standard deviation. After applying a regression analysis to calculate upper values, a combination of two functions was selected to establish an AT which takes into account both wind data and IA. This combination is comprised by a quartic function to deal with the smaller IAs and a negative exponential function for the bigger ones (Table 2). Finally, the current version of the segmentation algorithm has a shorter wind speed range than the previous one. We decided to gather together pixels with

wind speed values bigger than 7 m/s since only 8% of them had bigger values.

Table 2: Coefficients and cutoff points for adaptive thresholding functions.

Wind speed	Quartic function	Cutoff point	Negative exponential function
$<= 4m/s$	$6.3662 * 10^{-6}x^4 - 0.00083671x^3 + 0.041262x^2 - 0.90719x + 7.5353$	37.8°	$5.9169e^{-0.17881x}$
$<= 5m/s$	$3.0874 * 10^{-6}x^4 - 0.00043845x^3 + 0.023424x^2 - 0.55931x + 5.0588$	36.25°	$6.6689e^{-0.17944x}$
$<= 6m/s$	$2.1022 * 10^{-6}x^4 - 0.00031131x^3 + 0.017414x^2 - 0.4365x + 4.1503$	36.82°	$5.7770e^{-0.17217x}$
$<= 7m/s$	$2.7742 * 10^{-6}x^4 - 0.00039094x^3 + 0.02086x^2 - 0.50088x + 4.5883$	37.52°	$5.4477e^{-0.16858x}$
$> 7m/s$	$1.2416 * 10^{-6}x^4 - 0.00019466x^3 + 0.011604x^2 - 0.31143x + 3.182$	36.8°	$6.2053e^{-0.17101x}$

3.2. Characterization

Previous studies have shown that oil spills can be clustered according to their shape [26]. Several factors, such as the ship course, meteorological conditions and the age of oil spills affect to the slick shape. According to the cited study the vast majority of oil spills are of linear shape, either straight or angular. Literature also shows studies which claim that the unique shapes of the oil spills can be used to discriminate spillages over look-alikes [27] [28]. Consequently, a vector of characteristics, mainly based on shape, was used to characterize oil spill candidates. Furthermore, the vector was also filled with both physical and contextual characteristics. Initially, the vector, which is shown in Table 3, consisted of 21 components where 17 of them were related to shape. In order to reduce the vector dimensionality, a Principal Component Analysis (PCA) was applied to shape characteristics. Finally, five principal components were selected since they contained more than 90% of the shape information. Thus, the final established feature vector was composed by 9 components: five principal components related to the candidate shape, two physical characteristics, and two contextual characteristics.

Table 3: Initial vector of characteristics.

Type	Name	Description
Shape	APR	Ratio between $area(candidate)$ and $perimeter(candidate)$
	Elongation	Ratio between major and minor axis of the candidate.
	MPR	Ratio between $major\_axis(candidate)$ and $perimeter(candidate)$.
	Rectangularity	Ratio between $area(candidate)$ and $area(minimum.enclosing.rectangle)$.
	Circularity	Ratio between $perimeter(candidate)^2$ and $4\pi * area(candidate)$.
	Thickness	Number of erosions that are necessary to completely erode the candidate.
	Hu's moment invariants [29], η_{ij} symbolizes normalized central moments.	
	ϕ_1	$(\eta_{20} + \eta_{02})$
	ϕ_2	$(\eta_{20} - \eta_{02}) + 4\eta_{11}^2$
	ϕ_3	$(\eta_{30} + 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$
	ϕ_4	$(\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$
	ϕ_5	$(\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$
	ϕ_6	$(\eta_{20} - \eta_{02})[(\eta_{30} - \eta_{12})^2 - (\eta_{21}\eta_{03})^2] + 4\eta_{21}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$
	ϕ_7	$(3\eta_{21} - \eta_{30})(\eta_{00} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$
	Flusser and Suk affine moment invariants [30], η_{ij} symbolizes normalized central moments.	
	I_1	$(\eta_{20}\eta_{02} - \eta_{11}^2)/\eta_{00}^4$
	I_2	$(-\eta_{30}^2\eta_{03}^2 + 6\eta_{30}\eta_{21}\eta_{12}\eta_{03} - 4\eta_{30}\eta_{12}^3 - 4\eta_{21}^3\eta_{03} + 3\eta_{21}^2\eta_{12}^2)/\eta_{00}^{10}$
	I_3	$(\eta_{20}\eta_{21}\eta_{03} - \eta_{20}\eta_{12}^2 - \eta_{11}\eta_{30}\eta_{03} + \eta_{11}\eta_{21}\eta_{12} + \eta_{02}\eta_{30}\eta_{12} - \eta_{02}\eta_{21}^2)/\nu_{00}^7$
	I_4	$(-\eta_{20}^3\eta_{03}^2 + 6\eta_{20}^2\eta_{11}\eta_{12}\eta_{03} - 3\eta_{20}^2\eta_{02}\eta_{12}^2 - 6\eta_{20}\eta_{11}^2\eta_{21}\eta_{03} - 6\eta_{20}\eta_{11}^2\eta_{12}^2 + 12\eta_{20}\eta_{11}\eta_{02}\eta_{21}\eta_{12} - 3\eta_{20}\eta_{02}^2\eta_{21}^2 + 2\eta_{11}^3\eta_{30}\eta_{03} + 6\eta_{11}^3\eta_{21}\eta_{12} - 6\eta_{11}^2\eta_{02}\eta_{30}\eta_{12} - 6\eta_{11}^2\eta_{02}\eta_{21}^2 + 6\eta_{11}\eta_{02}^2\eta_{30}\eta_{21} - \eta_{02}^3\eta_{30}^2)/\eta_{00}^{11}$
Contextual	IA	IA of the candidate centroid.
	WM	Mean of candidate wind speed.
Physical	IM	Mean of candidate intensity.
	IR	Ratio between 'IM' and the intensity mean of clean pixels belonging to the window centered at the candidate region.

3.3. Classification

Segmented and characterized candidates must be classified in order to detect the spillages. A segmented candidate database was fixed to develop several classifiers. This database was made up of 155 look-alikes and 80 oil spills. Candidates were labeled using the EMSA and SASEMAR reports. At this point it is important to remark that some confirmed oil spills were broken during the segmentation process but they were dealt as different candidates by the classifiers. The database was split into three subsets: the biggest one, with 70% of elements, was saved for training purposes; the remainder was further split into two subsets with the same number of candidates, one of

them to evaluate the different classifier configurations and the other one to test the selected classifier.

Developed classifiers, which are shown in the Section 4.1, prioritize the minimizing of false negatives to avoid environmental consequences even if the number of false positives is increased.

3.4. Post-processing

The aggressiveness of the detection process greatly influences the number of look-alikes but, as a consequence, fragmentation of oil spills often happens. Thus, a post-processing step, which is summarized in Figure 2.a, is applied in order to recover the original shape of the detected oil spills.

The post-processing procedure analyzes each oil spill. For each slick, the bounding box is retrieved and extended. We have used a 3x3 matrix for extending it, where all the matrix cells have exactly the same shape and area than the original bounding box, and the center is focused on the slick. The established target window, symbolized by the matrix area, is used to retrieve a sub-image from the original SAR product (SARSI) as well as from the classified binary image (CSI). The noise of SARSI is reduced via a Gaussian filter. A threshold, which is established as the mean intensity minus the standard deviation, is applied to the filtered image. Pixels with intensity values lower than the threshold are selected, and a new binary image is obtained (TSAR). This new binary image as well as the sub-image from the classified binary image are analyzed through an iterative process: a dilation

operation is executed over the CSI and a bitwise conjunction operation is applied between both the dilated output (DCSI) and TSAR. After that, the bitwise conjunction operation result (CRDO) replaces the CSI as the dilation operation input. The process ends as soon as the result does not significantly differ from the one yielded by the previous iteration.

The post-processing function does not reduce the original classified oil-spill since a per-element bitwise disjunction operation is applied between the original classified candidate and the output given by the shape improvement process. An example of the oil spill enhancement is shown in Figure 2.b.

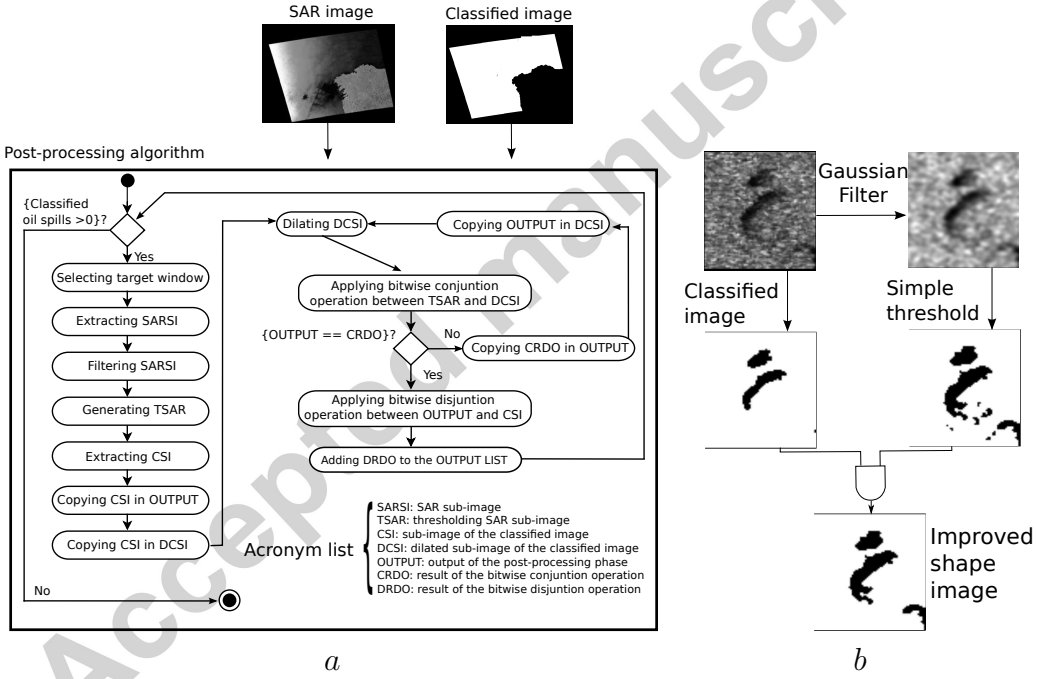

Figure 2: a) Post-processing algorithm steps. b) Example of the oil spill enhancement.

4. Software prototype

Previously shown algorithms were assembled into an operational prototype supported by public licenses. The main goal was to provide the Galician surveillance authorities with a useful tool, since they do not possess their own automatic detection system.

4.1. Classifiers

Two classifiers were implemented and compared in order to state the classification process: on the one hand, a classifier based on ANN and, on the other hand, a classifier based on a decision tree. ANN [31] [32], as well as decision trees [33], have proved to be useful classifiers to develop expert systems focused on remote sensing data. Specific details of the classifiers are shown below.

4.1.1. Artificial neural network

In order to develop a classifier based on a Multilayer Perceptron (MLP) neural network, several network architectures using different learning rates and different momentum terms were tested. The selected architecture comprised three layers. Concretely, the input layer contained 9 neurons to match components in the characteristic vectors, the hidden layer held 11 neurons, and finally, the output layer enclosed two neurons which symbolized each of the possible candidate classes.

The number of intermediate layer neurons was set through empirical tests based on the 'ad hoc' rule that this number should not be higher than the

double of neurons of the input data [34]. Thus, the intermediate layer was tested with different number of neurons from 1 to 18. Finally, a layer with 11 neurons was set since it achieved the best performance. After the empirical tests, a learning rate of 0.3 and a momentum factor of 0.9 were also selected. Moreover, The transfer function for each neuron was set as the weighted sum of all its inputs, and the hyperbolic tangent function was selected as the neuron activation function.

The training set was used to configure the ANN weights, while the validation set was used to test the ANN configuration and to prevent overfitting. Finally, the testing set allowed us to check the final configuration with an independent set and to confirm the ANN predictive power.

4.1.2. Decision Trees

The second implemented classifier was based on a decision tree supported by a Classification and Regression Tree (CART) algorithm [35]. Initially, the training subset was used to develop a decision tree with pure nodes. The Gini method [35] was applied to calculate the node impurity. Later, it was pruned using the validation subset and finally, the performance of the resulting decision tree was checked with the testing subset. One of the main advantages of the decision tree is its easy understanding and interpretation. Thus, the decision tree, which is shown in Figure 3, highlights essential characteristics used to classify candidates. Characteristics with gray background were removed from the decision tree during the pruning process. Furthermore,

Table 4 displays the linear composition of the principal component labeled as PCA3, which has proved to be relevant in the decision tree.

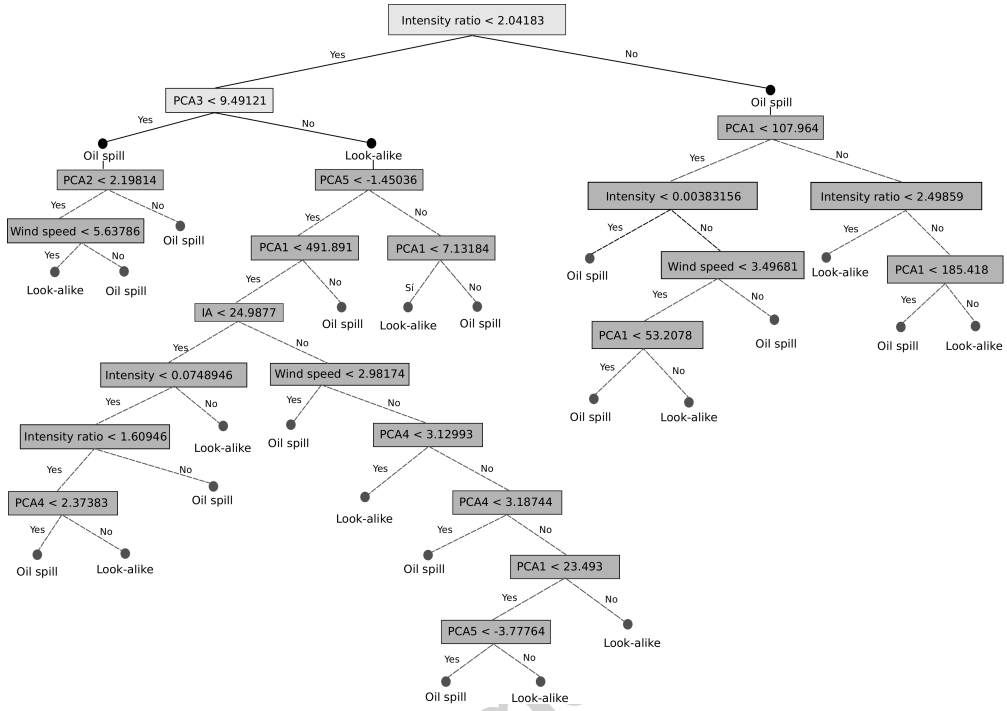

Figure 3: Decision tree where highlighted branches symbolize essential characteristics.

Table 4: Linear combination of the principal component labeled as PCA3.

Name	Value	Name	Value	Name	Value
APR	0.6160	ϕ_1	-0.0735	I_1	0.0373
Elongation	-0.1469	ϕ_2	-0.0683	I_2	-0.1329
MPR	0.3499	ϕ_3	-0.0237	I_3	-0.1053
Rectangularity	0.2473	ϕ_4	-0.0167	I_4	-0.0020
Circularity	0.0457	ϕ_5	-0.0049		
Thickness	0.6099	ϕ_6	0		
		ϕ_7	-0.0151		

5. Results

The fact that previously shown classifiers were developed and tested using the same datasets allows us to compare them from the accuracy point of view.

	Validation		Test		
	Oil Spills	Look-alikes	Oil Spills	Look-alikes	Global efficiency
ANN	85.7% (12/14)	85.2% (23/27)	92.9% (13/14)	96.3% (26/27)	95.1%
Decision tree	92.9% (13/14)	85.2% (23/27)	92.9% (13/14)	92.6% (25/27)	92.6%

Table 5: Classification accuracy of Test and Validation subsets using developed classifiers.

Usually, testing subsets are reduced, since spillage database availability is limited and a balance between the number of oil spills and look-alikes should be met in order to achieve an appropriate development of classifiers. Despite Table 5 shows promising results, classifiers must also be tested over entire SAR images, which usually contain a much larger number of look-alikes.

Classifier behaviors were directly tested over a subset of SAR images which was saved for that purpose. Figure 4 shows an example of the system execution. The top-left frame shows a composition of two images captured by Envisat where gray levels were tuned to improve visualization. Sub-image '1' was collected on November 14th, 2011 while sub-image '2' was collected on March 28th, 2011. Readers are encouraged to check the evolution of the algorithms as described in the previous paper since the displayed composition was also shown there. Three oil spills were located in this composition and two of them are inside the ROI. The top-right frame shows the classification system output using the ANN classifier, while the bottom-left frame shows the output after applying the classifier based on the decision tree. Although

both classifiers detect the three spillages, the ANN shows less look-alikes, which agrees to Table 5. Finally, the bottom right frame shows oil spill SAR sections and their system outputs. Image sections 'a' and 'c' show a common output since classifier results were identical in these regions. However, the classifier outputs were different in Section 'c'. The first output, from the left side, shows the ANN result, while the second one shows the decision tree output.

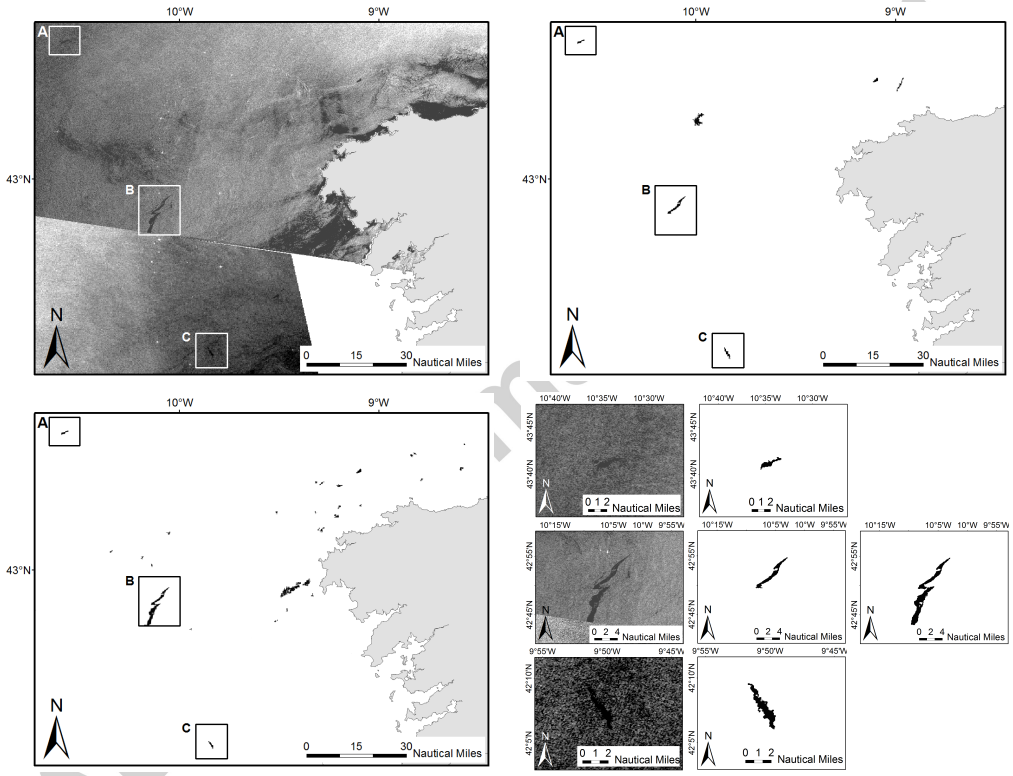

Figure 4: Envisat ASAR image composition and classification system outputs using both classifiers.

5.1. Processing time

Processing time is critical in DSSs focused on oil spill detection. Nevertheless, it is left in the background by most previous studies. Following the experience acquired in the development of the segmentation process, system algorithms were developed using shared-memory parallel programming techniques in order to take advantage of contemporary multicore architectures. Specifically, the OpenMP API [36] was used to parallelize some code sections, such as the loops focused on the analysis of the image pixels.

Both simple and improved parallelized algorithms were tested at a workstation (Core 2 Duo - 2.13 GHZ, 3 GB RAM) with images of 8,088 x 6,481 pixels. Processing time is directly related to the number of segmented candidates. Table 6 shows several time measurements where the parallel improvement average is around 25%.

Table 6: Processing time comparison between sequential and parallel system execution.

			Processing time		
SAR product	Segmented pixels	Oil spill candidates	Sequential	Parallel processing time	Improvement
SAR-P1[1]	210,064	670	62.16 sec.	45.33 sec.	27.07%
SAR-P2[2]	36,366	108	24.4 sec.	18.57 sec.	23.89%
SAR-P3[3]	27,211	65	20.89 sec.	15.29 sec.	26.81%

[1]ASA_WSM_1PNDPA20070608_104715_000000732058_00452_27560_0360
[2]ASA_WSM_1PNIPA20111014_105250_000000733107_00310_50324_4066
[3]ASA_WSM_1PNIPA20110328_223900_000000733101_00030_47458_4070

6. Discussion

There are several sources for oil spills and look-alikes. Fortunately, they are limited in the studied area. Regarding look-alikes, orography and meteorological conditions mainly restrict them to low wind areas. The developed

system filters most of them in the segmentation phase due to the fact that AT is based on both wind speed and satellite IA. Regrettably, slicks are usually eroded and even fragmented in the segmentation process, specially when they are partially located in low wind areas. This issue is mainly solved by the shape enhancement post-process.

In relation to oil spills, there are neither oil platforms nor natural oil seeps in the Galician coast. Thus, maritime traffic is the main oil spill source. Since previous studies have shown than oil spill shape is related to its source, our classifiers mainly analyze the shape to distinguish oil spills from look-alikes. Reported classification accuracies are not easily compared, mainly because oil spill detection approaches are based on different datasets. The classifier accuracy of some previously published relevant classifiers are shown in Table 7. Although the results can not be directly compared, the developed system shows promising results since its accuracy is similar or even better than data shown in Table 7. Concretely, given that the system was developed to minimize false negatives, better results are accomplished in detecting oil spills.

Ref.	Dataset composition	Classifier	Accuracy
[9]	71 oil spills, 68 look-alikes	ANN - MLP	90% oil spills, 82% look-alikes
[11]	34 oil spills, 45 look-alikes	ANN - MLP	87% oil spills, 91% look-alikes
[19]	71 oil spills, 6,980 look-alikes	statistical classifier + rule-based system	94% oil spills,99% Look-alikes
[8]	37 oil spills, 12,110 look-alikes	statistical classifier + rule-based system	78.4% oil spills,99.4% look-alikes
[20]	41 oil spills 12,245, look-alikes	regularized statistical classifier + automatic confidence estimation	92.07% oil spills (without confidence estimation),89.7% look-alikes
[7]	11 oil spills, 6 look-alikes, 4 unknown	mahalanobis distance	82% oil spills, 100% look-alikes, 0% unknown
		probabilistic model	90.9% oil spills, 67% look-alikes, 50% unknown

Table 7: Other oil spill detection system accuracies

Surprisingly, a simple decision tree classifier shows results similar to more complex classifiers. Probably, the reason is that look-alikes are mainly filtered in the segmentation phase. Thus, the classification process is simplified. Moreover, most remaining candidates can be classified using the IR component of the vector of characteristics, and only few of them require more information.

From the processing time viewpoint, there are several techniques in the literature to accelerate the process, such as to analyze sub-images [11], to use skip factors [8], and to reduce the radiometric resolution [37] [12]. Nevertheless, shared memory parallelization techniques have shown promising results since they allow the processing of high resolution SAR data in near real-time.

7. Conclusions and ongoing work

This paper has presented an expert system to find oil spills through SAR images. The system is based on algorithms adapted to Envisat ASAR data. Envisat is still in orbit but sadly, it is inoperative. However, the Sentinel-1 is the ESA space mission designed to take the place of the Envisat satellite and thereby assure the continuity of radar-based Earth observation services. The Sentinel-1 was recently launched on April 3th, 2014 and during the following months it will be fully operative. Thus, a future oil spill detection system based on Sentinel-1 data should be developed on the basis of the previously presented system.

Despite the system accuracy shows a high rate, there is also a minimum

percentage of false positives as well as false negatives. To improve the algorithms we plan to extend the vector of characteristics with contextual data such as the proximity of potential sources. Concretely, ships detected from SAR images can be confronted with data provided by the ship automatic tracking system called Automatic Identification System (AIS). Moreover, this tracking information could be used to identify polluters.

Acknowledgements

The authors wish to thank the financial support provided by the 'Deputación da Coruña' under the 'Bolsas de Investigación 2013' programme. We also wish to express gratitude for the help and data provided by SASEMAR.

[1] R. Asariotis, H. Benamara, H. Finkenbrink, J. Hoffmann, A. Jalmurzina, A. Premti, V. Valentine, F. Youssef, Review of maritime transport, 2012, Tech. rep. (2012).

[2] Esa, Oil pollution monitoring, in ERS and its applications: Marine, Tech. rep. (1998).

[3] H. A. Hovland, J. A. Johannessen, G. Digranes, Slick detection in SAR images, in: Proceedings of IGARSS '94 - 1994 IEEE International Geoscience and Remote Sensing Symposium, IEEE, 1994, pp. 2038–2040.

[4] C. Jackson, J. Apel (Eds.), Synthetic Aperture Radar Marine User's Manual, 1st Edition, U.S. Department of Commerce : National

Oceanic and Atmospheric Administration, 2005, also available online: http://www.sarusersmanual.com (Retrieved July 2014).

[5] C. Brekke, A. H. S. Solberg, Oil spill detection by satellite remote sensing, Remote Sensing of Environment 95 (1) (2005) 1–13.

[6] K. Topouzelis, Oil Spill Detection by SAR Images: Dark Formation Detection, Feature Extraction and Classification Algorithms, Sensors 8 (10) (2008) 6642–6659.

[7] B. Fiscella, A. Giancaspro, F. Nirchio, P. Pavese, P. Trivero, Oil spill detection using marine SAR images, International Journal of Remote Sensing 21 (18) (2000) 3561–3566.

[8] A. H. S. Solberg, C. Brekke, P. O. Husoy, Oil Spill Detection in Radarsat and Envisat SAR Images, IEEE Transactions on Geoscience and Remote Sensing 45 (3) (2007) 746–755.

[9] F. Del Frate, A. Petrocchi, J. Lichtenegger, G. Calabresi, Neural networks for oil spill detection using ERS-SAR data, IEEE Transactions on Geoscience and Remote Sensing 38 (5) (2000) 2282–2287.

[10] S. Singha, T. J. Bellerby, O. Trieschmann, Satellite oil spill detection using artificial neural networks, Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 6 (6) (2013) 2355–2363.

[11] K. Topouzelis, V. Karathanassi, P. Pavlakis, D. Rokos, Detection and

discrimination between oil spills and look-alike phenomena through neural networks, ISPRS Journal of Photogrammetry and Remote Sensing 62 (4) (2007) 264–270.

[12] O. Garcia-Pineda, B. Zimmer, M. Howard, W. Pichel, L. XiaoFeng, I. R. MacDonald, Using SAR images to delineate ocean oil slicks with a texture-classifying neural network algorithm (TCNNA), Canadian Journal of Remote Sensing 35 (5) (2009) 411–421.

[13] G. Benelli, A. Garzelli, Oil-spills detection in SAR images by fractal dimension estimation, in: Geoscience and Remote Sensing Symposium, 1999. IGARSS Proceedings. IEEE 1999 International, Vol. 1, IEEE, 1999, pp. 218–220 vol.1.

[14] M. Marghany, RADARSAT for oil spill trajectory model, Environmental Modelling and Software 19 (5) (2004) 473–483.

[15] C. F. Chen, K. S. Chen, L. Y. Chang, A. J. Chen, The use of satellite imagery for monitoring coastal environment in Taiwan, in: Geoscience and Remote Sensing, 1997. IGARSS '97. Remote Sensing - A Scientific Vision for Sustainable Development., 1997 IEEE International, Vol. 3, IEEE, 1997, pp. 1424–1426 vol.3.

[16] M. Migliaccio, M. Tranfaglia, S. A. Ermakov, A physical approach for the observation of oil spills in SAR images, Oceanic Engineering, IEEE Journal of 30 (3) (2005) 496–507.

[17] K. Topouzelis, D. Stathakis, V. Karathanassi, Investigation of genetic algorithms contribution to feature selection for oil spill detection, Int. J. Remote Sens. 30 (3) (2009) 611–625.

[18] K. Topouzelis, A. Psyllos, Oil spill feature selection and classification using decision tree forest on SAR image data, ISPRS Journal of Photogrammetry and Remote Sensing 68 (2012) 135–143.

[19] A. H. S. Solberg, G. Storvik, R. Solberg, E. Volden, Automatic detection of oil spills in ERS SAR images, IEEE Transactions on Geoscience and Remote Sensing 37 (4) (1999) 1916–1924.

[20] C. Brekke, A. H. S. Solberg, Classifiers and Confidence Estimation for Oil Spill Detection in ENVISAT ASAR Images, IEEE Geoscience and Remote Sensing Letters 5 (1) (2008) 65–69.

[21] R. Periáñez, A. Pascual-Granged, Modelling surface radioactive, chemical and oil spills in the strait of gibraltar, Computers & Geosciences 34 (2) (2008) 163 – 180.

[22] M. Kulawiak, A. Prospathopoulos, L. Perivoliotis, M. uba, S. Kioroglou, A. Stepnowski, Interactive visualization of marine pollution monitoring and forecasting data via a web-based {GIS}, Computers & Geosciences 36 (8) (2010) 1069 – 1080.

[23] D. Mera, J. M. Cotos, J. Varela-Pet, O. Garcia-Pineda, Adaptive thresholding algorithm based on SAR images and wind data to segment oil

spills along the northwest coast of the Iberian Peninsula, Marine Pollution Bulletin 64 (10) (2012) 2090–2096.

[24] R. B. Olsen, T. Wahl, The ship detection capability of ENVISAT's ASAR, in: IGARSS 2003. 2003 IEEE International Geoscience and Remote Sensing Symposium. Proceedings (IEEE Cat. No.03CH37477), IEEE, 2003, pp. 3108–3110.

[25] H. Hersbach, A. Stoffelen, S. de Haan, An improved C-band scatterometer ocean geophysical model function: CMOD5, Journal of Geophysical Research 112 (C3) (2007) C03006.

[26] P. Pavlakis, D. Tarchi, A. Sieber, On the monitoring of illicit vessel discharges using spaceborne sar remote sensing - a reconnaissance study in the Mediterranean sea, Annals of Telecommunications 56 (11) (2001) 700–718.

[27] Y. Guo, H. Z. Zhang, Oil spill detection using synthetic aperture radar images and feature selection in shape space, International Journal of Applied Earth Observation and Geoinformation 30 (0) (2014) 146–157.

[28] D. Chaudhuri, A. Samal, A. Agrawal, A. Mishra, V. Gohri, R. Agarwal, et al., A statistical approach for automatic detection of ocean disturbance features from sar images, Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 5 (4) (2012) 1231–1242.

[29] M.-K. Hu, Visual pattern recognition by moment invariants, Information Theory, IRE Transactions on 8 (2) (1962) 179–187.

[30] J. Flusser, B. Zitova, T. Suk, Moments and moment invariants in pattern recognition, John Wiley & Sons, 2009, Ch. 3.

[31] A. Iglesias Nuno, B. Arcay, J. M. Cotos, J. Varela, Optimisation of fishing predictions by means of artificial neural networks, anfis, functional networks and remote sensing images, Expert Systems with Applications 29 (2) (2005) 356–363.

[32] M. J. Cracknell, A. M. Reading, Geological mapping using remote sensing data: A comparison of five machine learning algorithms, their response to variations in the spatial distribution of training data and the use of explicit spatial information, Computers & Geosciences 63 (0) (2014) 22 – 33.

[33] E. Quirós, A. M. Felicísimo, A. Cuartero, Testing Multivariate Adaptive Regression Splines (MARS) as a method of land cover classification of TERRA-ASTER satellite images, Sensors 9 (11) (2009) 9011–9028.

[34] K. Swingler, Applying Neural Networks: A Practical Guide, pap/dsk Edition, Morgan Kaufmann, 1996, Ch. 3.2.3, pp. 53–56.

[35] L. Breiman, J. H. Friedman, R. A. Olshen, C. J. Stone, Classification and Regression Trees, Wadsworth, 1984.

[36] L. Dagum, R. Menon, OpenMP: an industry standard API for shared-memory programming, IEEE Computational Science and Engineering 5 (1) (1998) 46–55.

[37] J. M. T. Palenzuela, L. G. Vilas, M. S. Cuadrado, Use of ASAR images to study the evolution of the Prestige oil spill off the Galician coast, International Journal of Remote Sensing 27 (10) (2006) 1931–1950.