

## Segmentação e Classificação de Imagens SAR Aplicadas à Detecção de Alvos Escuros em Áreas Oceânicas de Exploração e Produção de Petróleo

Patrícia Carneiro Genovez <sup>1</sup>  
Nelson Francisco Favilla Ebecken <sup>2</sup>  
Corina da Costa Freitas <sup>3</sup>  
Cristina Maria Bentz <sup>4</sup>  
Ramon Morais de Freitas <sup>2</sup>  
Luciano Vieira Dutra <sup>2</sup>

<sup>1</sup> Integrated Petroleum Expertise Company - IPEXco  
Rua Dezenove de Fevereiro, 69/71, Botafogo – Rio de Janeiro - RJ, Brasil  
patricia.genovez@ipexco.com.br

<sup>2</sup> Universidade Federal do Rio de Janeiro – UFRJ/COPPE  
Centro de Tecnologia, Sala B100, Ilha do Fundão – 21949-900 – Rio de Janeiro – RJ, Brasil  
nelson@ntt.ufrj.br

<sup>3</sup> Instituto Nacional de Pesquisas Espaciais - INPE  
Caixa Postal 515 – 12245-970 – São José dos Campos - SP, Brasil  
{corina, dutra}@dpi.inpe.br; ramon@dsr.inpe.br

<sup>4</sup> PETROBRAS/CENPES - Centro de Pesquisas  
Cidade Universitária, Q.7, Ilha do Fundão – 21949-900 - Rio de Janeiro - RJ, Brasil  
cris@petrobras.com.br

**Abstract.** Automatic oil detection systems have been developed to improve SAR image interpretation, composed of four principal stages: a) image pre-processing; b) dark spot detection; c) feature extraction, and; d) oil and look-alike classification. The dark spot detection is considered the main step in the processing chain: without the geometry of the spots, the oil and look-alikes classification is unfeasible. In this context, this work aimed to develop an automatic procedure able to detect dark spots in SAR images, by the integration of segmentation and pattern recognition techniques. The results presented are continuity of the studies carried on by Genovez (2010) and consider the tree last stages as follow: a) features extraction, exploratory analyses and feature selection; b) dark spot detection using data clustering, and; c) validation of the proposed method. Considering that in the scientific community there isn't a wide agreement about the operational use of fully automatic methods, the development of an intelligent hybrid system, including decision rules able to conduct the images for one automatic or semi-automatic processing, was an interesting approach. The potential of these rules to improve the automation process was indicated. Nevertheless, more samples to return more robust rules are recommended in order to be widely applied to all SAR images acquired.

**Palavras-chave:** synthetic aperture radar (SAR), dark spot detection, oil detection, offshore exploration and production areas, image segmentation and clusterization, intelligent hybrid system, radares de abertura sintética (SAR), detecção de alvos escuros, detecção de óleo, exploração e produção de petróleo em áreas offshore, segmentação e clusterização de imagens, sistema híbrido inteligente.

### 1 – Introdução:

Sistemas para detecção automática de vazamentos de óleo têm sido desenvolvidos para otimizar e auxiliar a interpretação das imagens SAR (Solberg et al., 2005; Bentz, 2006). A cadeia de processamentos considera prioritariamente as áreas oceânicas, emitindo um alarme quando uma região escura, de baixo retroespalhamento, for identificada como uma suspeita de mancha de óleo (Keramitsoglou et al., 2006). Estes sistemas geralmente são compostos por quatro etapas (Figura 1): a) pré-processamento das imagens; b) individualização de regiões escuras nas imagens, utilizando técnicas de segmentação e de agrupamento; c) cálculo de atributos para cada região escura detectada, e; d) classificação das regiões escuras como óleo e/ou evento assemelhado com base nos atributos calculados (Brekke & Solberg, 2005).

Dentre as etapas envolvidas, a detecção dos alvos escuros é a principal etapa da cadeia de processamentos. Sem a detecção e individualização da geometria dos alvos não há como prosseguir na etapa de classificação dos eventos em óleo ou assemelhados (Kanaa *et al.*, 2003; Pelizzari & Dias, 2003; Karathanassi *et al.*, 2006; Topouzelis *et al.*, 2009). Diversas abordagens para a detecção de alvos escuros têm sido propostas combinando técnicas de segmentação de imagens e agrupamento de regiões.

A pesquisa realizada considerou apenas a segunda etapa integrante de um sistema de detecção automático de óleo com o objetivo de propor um procedimento automático para a detecção de alvos escuros em imagens SAR. Espera-se que os resultados obtidos nas etapas investigadas indiquem parâmetros ótimos que possam ser considerados como padrão para que o sistema seja automatizado em um futuro próximo, acoplando todas as etapas envolvidas.

## 2 – Recursos e Procedimentos Metodológicos:

Doze imagens SAR com diferentes níveis de complexidade provenientes dos satélites RADARSAT-1 e RADARSAT-2 nos modos de imageamento *ScanSAR Narrow* e *ScanSAR Wide* foram cedidas pela PETROBRAS para os estudos desenvolvidos por Genovez *et al.* (2009) e Genovez (2010). Do conjunto de imagens dez foram utilizadas para o desenvolvimento das análises e duas foram utilizadas para a validação do sistema proposto.

As imagens foram pré-processadas, segmentadas e agrupadas considerando os parâmetros padronizados e pré-definidos por Genovez *et al.*, 2009 e Genovez, 2010, sendo estes: a) Espaçamento de *Pixel*: 100m; b) Filtragem: Lee 3x3, e; c) Segmentação: *software* SPRING (Câmara *et al.*, 1996; Bins *et al.*, 1996) com Área Mínima = 25 *pixels* e Limiar de Similaridade = 4. Posteriormente, foram calculados para cada região escura ou clara resultante do processo de segmentação 9 atributos agrupados em 4 categorias: a) Estatísticos de Primeira Ordem: extraídos a partir dos Níveis de Cinza das imagens; b) Geométricos: extraídos da geometria dos polígonos resultantes da segmentação, e; c) Texturais: calculados a partir da matriz de co-ocorrência dos níveis de cinza – GLCM (Haralick *et al.*, 1979). Estes atributos foram previamente selecionados dentre um universo inicial de 36 atributos através de estudos efetuados por Genovez *et al.*, 2009 e Genovez, 2010, sendo estes: Mediana, Média, Moda, Máximo, GLCM da Média, Ajuste Retangular, GLCM da Homogeneidade, Assimetria dos Polígonos e Densidade. Detalhamento técnico sobre o conjunto de 9 atributos pode ser consultado em Genovez (2010).

Estes 9 atributos foram clusterizados considerando diversas combinações utilizando o *software Clustan* no modo *Focal Point (FP)*, o qual utiliza o método de partição *k-means* para efetuar o agrupamento dos dados (Wishart, 2004). Os padrões de similaridade utilizados para separar as regiões em diferentes *clusters* são obtidos a partir de métricas de distância, sendo a Distância Euclidiana utilizada no *software*.

Para validar a implementação operacional de um método plenamente automático uma etapa de comparação dos resultados obtidos pelo *Clustan* com métodos semi-automáticos como o MaxVer a Máquinas de Suporte Vetorial (SVM) foi efetuada. Todos os resultados gerados foram avaliados através de métricas utilizadas para comparar os alvos escuros pelas clusterizações em relação à *Phantons*, imagens de referência elaboradas manualmente considerando a experiência do intérprete. A principal métrica utilizada foi denominada Exatidão de Alvos Escuros (*ExAE*) calculada com base em uma matriz de confusão sendo a divisão do número de *pixels* escuros corretamente agrupados como alvos escuros pelo somatório dos *pixels*: i) escuros corretamente agrupados como alvos escuros; ii) escuros erroneamente agrupados como oceano (Omissão: *Om*), e; iii) *pixels* claros erroneamente

agrupados como alvos escuros (Inclusão: *In*). As métricas estão detalhadas em Genovez (2010).

Finalmente um sistema híbrido inteligente capaz de operar através de regras de decisão que indiquem as imagens aptas ao processamento automático ou semi-automático foi proposto. Nesta etapa, dois novos recortes de cena foram utilizados para validar a regra de decisão obtida.

Em síntese, a metodologia foi composta por 3 etapas: a) reconhecimento de padrões; b) avaliação de desempenho de algoritmos automático e semi-automáticos, e; c) regras de decisão e validação do sistema híbrido inteligente para detecção de alvos escuros.

### 3.2 – Reconhecimento de Padrões em dados SAR:

As clusterizações foram efetuadas no *software Clustan* módulo *Focal Point (FP)* inserindo os 9 melhores atributos um a um, considerando a ordem de importância indicada por um ranking cumulativo (Genovez, 2010) para verificar a parcela de contribuição de cada atributo nos resultados da clusterização em 2 categorias: alvos escuros e oceano. A validação do melhor conjunto de atributos foi efetuada comparando os alvos escuros resultantes das clusterizações, com os alvos escuros definidos como referência nos *Phantoms*, utilizando como principal métrica de avaliação a *ExAE*, sendo o agrupamento perfeito obtido quando *ExAE* for igual a 1. A Tabela 1 apresenta: (a) as *ExAE* obtidas com a inserção simultânea dos 9 atributos para as 10 cenas, e; (b) as máximas *ExAE* obtidas a partir da melhor combinação de atributos resultante da inserção gradual dos 9 atributos um a um.

Tabela 1 - Síntese dos resultados obtidos para as clusterizações efetuadas para cada cena.

Resultados Obtidos pela Clusterização	Cena 1	Cena 2	Cena 3	Cena 4	Cena 5	Cena 6	Cena 7	Cena 8	Cena 9	Cena 10	Média	Mediana	Máximo
(a) <i>ExAE</i> : com todos os 9 Atributos	0,351	0,799	0,349	0,215	0,248	0,335	0,440	0,562	0,240	0,331	0,352	0,342	0,799
Nº de Atributos Necessários para atingir esta <i>ExAE</i>	9												
(b) Máxima <i>ExAE</i> obtida: inserção gradual dos 9 Atributos	0,392	0,799	0,348	0,215	0,248	0,507	0,775	0,612	0,282	0,331	0,410	0,370	0,799
Nº de Atributos Necessários para atingir esta <i>ExAE</i>	4	2	4	9	7	2	3	4	2	9	5	4	9

. A média das *ExAEs* obtida com a clusterização simultânea dos 9 atributos (0,35) foi menor do que a obtida a partir da melhor combinação de atributos resultantes da inserção gradual dos 9 atributos (0,41). Em 6 dentre as 10 cenas estudadas foi possível convergir para uma maior *ExAE* com um menor número de atributos, oscilando entre 2 e 4. Estes resultados validaram a hipótese de que a seleção de um conjunto reduzido de atributos considerados significativos é melhor do que a utilização de um amplo conjunto de dados que inclua atributos não representativos (Hughes, 1968).

Considerando que na média foram indicados 5 e na mediana 4 atributos como suficientes para o processo de clusterização automática, o teste não paramétrico de Wilcoxon foi efetuado para verificar se a diferença entre as *ExAE* obtidas com 4 e 5 atributos seria estatisticamente significativa. Com um *p*-valor de 0,68, o teste evidenciou que não há diferença estatisticamente significativa utilizando 4 ou 5 atributos. Entretanto, como para a cena 10 a retirada do 5º atributo piorou significativamente os resultados da clusterização, optou-se por padronizar o número de atributos em 5 sendo 4 estatísticos de primeira ordem (Mediana, Média, Moda e Máximo) e um textural GLCM da Média. A Figura 1 apresenta, para as 10 cenas, o resultado das clusterizações com maiores *ExAE* obtidas a partir da melhor combinação de atributos resultante da inserção gradual dos 9 atributos selecionados.

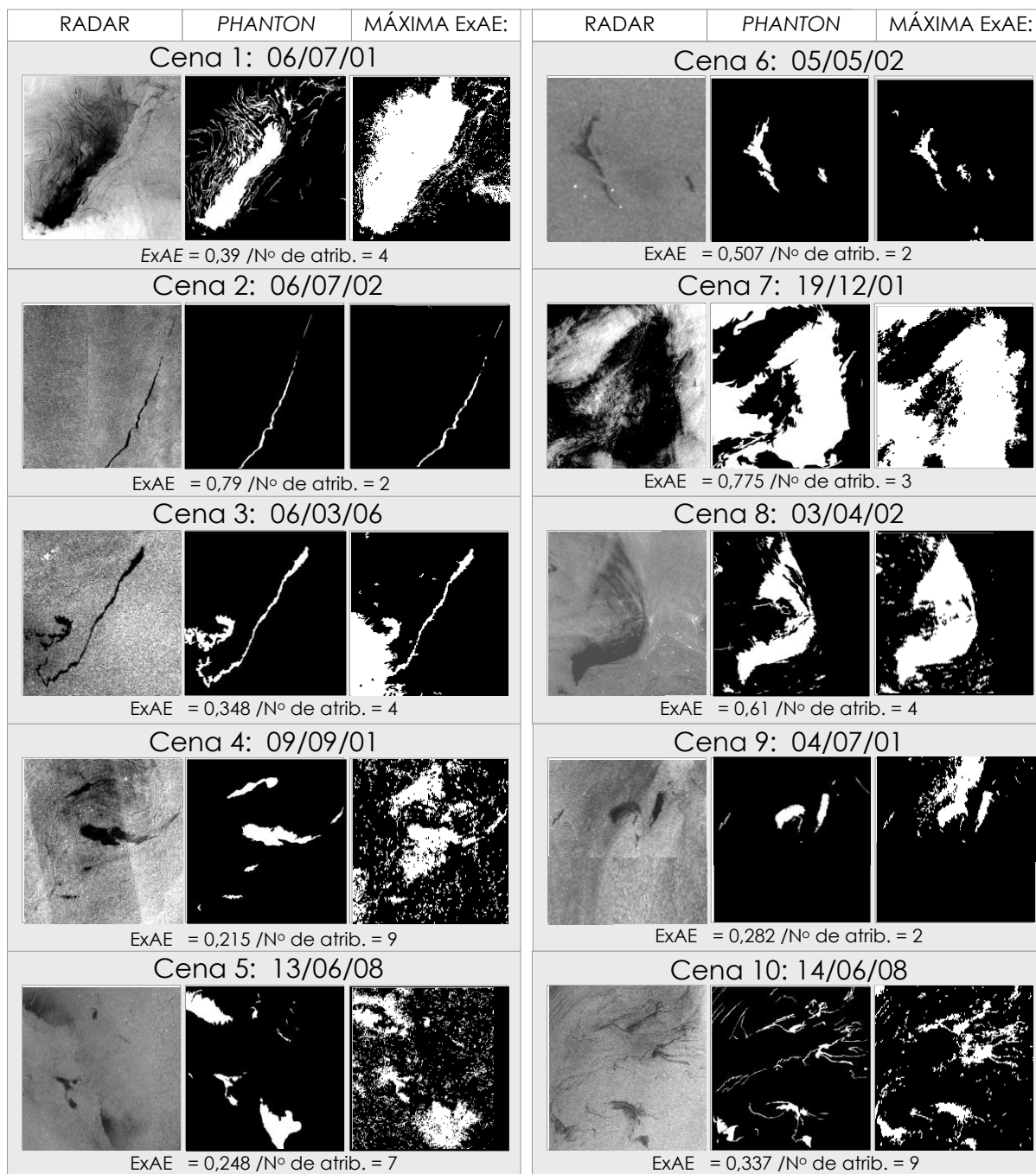


Figura 1 - Espacialização das clusterizações obtidas considerando a melhor combinação de atributos.

Apesar da automatização com um número reduzido de atributos ter sido indicada como factível, a análise global dos resultados evidenciou que, apenas 5 cenas obtiveram uma clusterização considerada satisfatória, sendo estas as cenas 2, 6, 7, 8 e 10. Destas, 4 cenas tiveram *ExAE* maiores ou iguais a 0,5 e 1 cena - com *ExAE* menor do 0,5 - foi considerada como satisfatória a partir de uma análise qualitativa.

Como a detecção de alvos escuros é considerada a etapa mais importante da cadeia de processamentos, definindo a geometria dos alvos escuros que será utilizada nas etapa posterior de classificação em óleo e assemelhados, a acurácia é um fator que não pode ser desconsiderado. Neste contexto, para que o sistema proposto seja passível de ser implementado operacionalmente, uma avaliação comparativa de desempenho entre técnicas completamente automáticas e semi-automáticas foi considerada essencial.

### 3.3 – Desempenho de Algoritmos Automático e Semi-automáticos:

Considerando que na literatura ainda não há um consenso sobre qual abordagem manual, semi-automática ou automática, representa o melhor custo/benefício para operacionalizar a detecção de alvos escuros em imagens SAR, uma comparação entre os métodos automático e semi-automático foi realizada para verificar com maior rigor o procedimento mais eficaz. Para tal, os resultados obtidos pelo *Clustan* (automático/não supervisionado) foram comparados qualitativamente e quantitativamente aos resultados obtidos por Limiarização e pelos classificadores semi-automáticos supervisionados Maxver e SVM através da métrica de Exatidão de Alvos Escuros (*ExAE*).

Os resultados indicaram que os métodos semi-automáticos conseguiram retornar *clusters* mais representativos e, conseqüentemente, geometrias mais precisas para os alvos escuros, considerando que a Média das métricas de *ExAE* obteve melhor desempenho para o classificador SVM (0,67), seguido da Limiarização (0,58). Nos métodos semi-automáticos foram observados mais erros de omissão (*Om*), enquanto que no método automático ocorreram mais erros de inclusão (*In*). Apesar das técnicas semi-automáticas terem tido um melhor desempenho, a performance automática do *Clustan* na identificação dos alvos escuros nas cenas 2 e 6 foi equivalente aos métodos supervisionados. Mesmo considerando uma menor precisão nas bordas, nas cenas 8 e 10 a performance automática também pôde ser considerada equivalente à dos métodos supervisionados, à medida que todos os métodos erraram incluindo ou omitindo regiões. A Figura 2 ilustra os resultados dos processamentos automáticos/semi-automáticos obtidos para os 10 recortes de cena estudados.

Considerando que para determinadas cenas a clusterização automática foi efetiva e para outras não, a pesquisa realizada propôs - de forma inédita - a elaboração de um sistema híbrido inteligente capaz de indicar através de regras de decisão as imagens aptas para o processamento automático ou semi-automático, com o objetivo de produzir geometrias mais precisas para a etapa de classificação automática. Para tal, regras de decisão foram estabelecidas e confrontadas a 2 novas imagens SAR, indicando as cenas a serem processadas automaticamente e as cenas que necessitariam de supervisão.

### 3.4 – Regras de Decisão e Validação do Sistema Híbrido Inteligente para Detecção de Alvos Escuros em dados SAR:

Apesar da diversidade das pesquisas desenvolvidas e dos inúmeros métodos e algoritmos utilizados para a detecção de alvos escuros em imagens SAR (Genovez, 2010), a utilização operacional de métodos plenamente automáticos ainda não foi plenamente validada na comunidade científica (SEASAR, 2010). A proposta de elaboração de um sistema híbrido inteligente capaz de indicar, através de regras de decisão previamente estabelecidas, as imagens aptas para o processamento automático, seria uma alternativa interessante.

Para tal, o método de Árvores de Classificação e Regressão - ACR (*Classification and Regression Trees - CART*), disponível no módulo *Data Mining* do *Software Statistica*, foi utilizado para estabelecer regras de decisão simples e compactas para indicar a aptidão das imagens para o processamento automático, considerando duas categorias: satisfatórias (S) e não satisfatórias (NS). A variável categórica dependente *Y* estava associada ao desempenho do clusterizador, indicando se o mesmo foi satisfatório (S) ou não (NS). As 5 novas variáveis preditoras (*X*) foram calculadas a partir da distância entre os *clusters* de alvos escuros e oceano para as 10 cenas, sendo: a)  $x_1 = \mu_{\text{Mediana Oceano}} - \mu_{\text{Mediana Alvos Escuros}}$ ; b)  $x_2 = \mu_{\text{Média Oceano}} - \mu_{\text{Média Alvos Escuros}}$ ; c)  $x_3 = \mu_{\text{Moda Oceano}} - \mu_{\text{Moda Alvos Escuros}}$ ; d)  $x_4 = \mu_{\text{Máximo Oceano}} - \mu_{\text{Máximo Alvos Escuros}}$ ; e)  $x_5 = \mu_{\text{GLCM Média Oceano}} - \mu_{\text{GLCM Média Alvos Escuros}}$ .

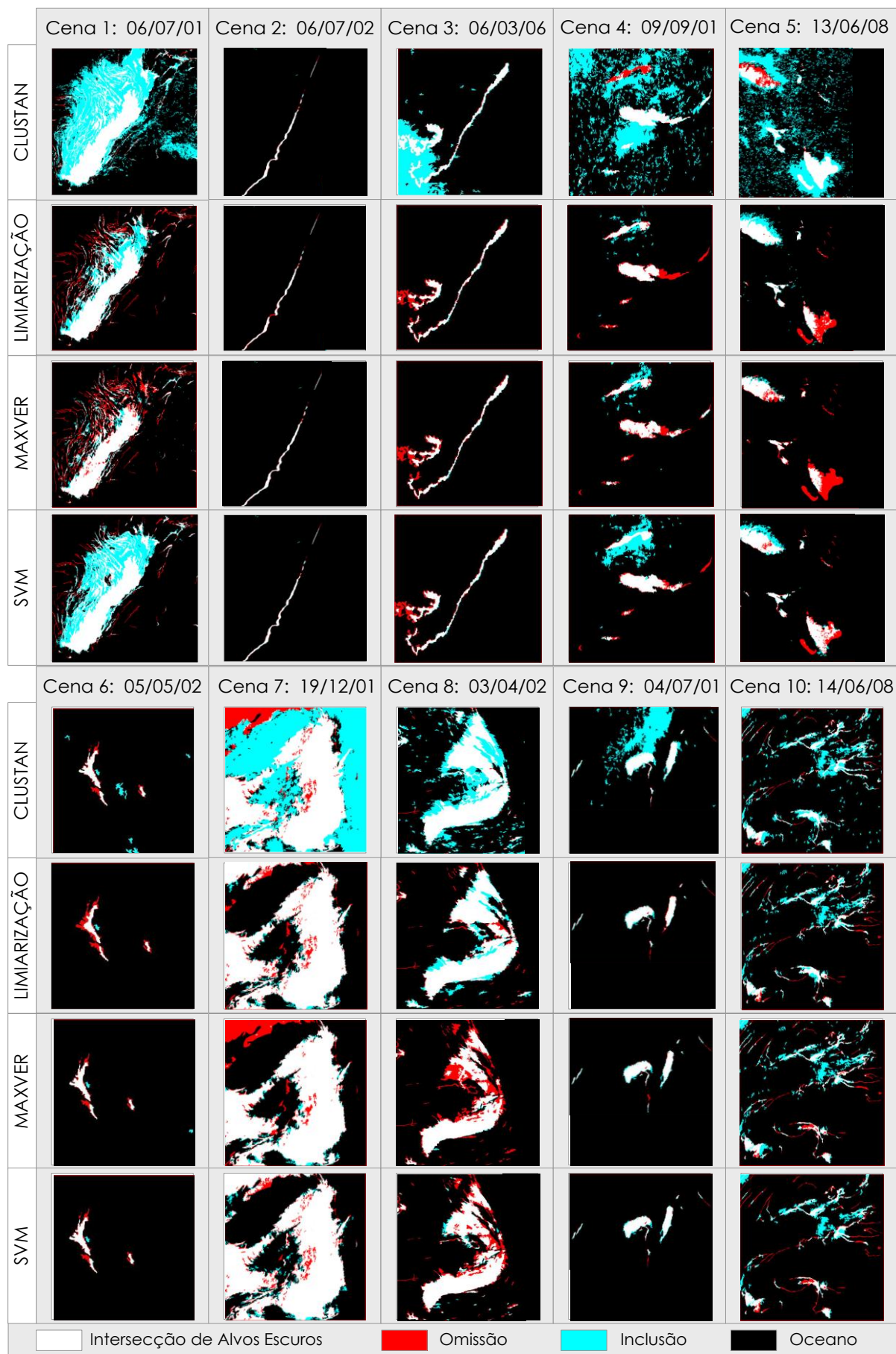


Figura 2 – Resultados da comparação entre os métodos automático e semi-automáticos.

Com base nestes atributos, a ACR retornou uma regra de decisão compacta, considerando como atributo de partição inicial a diferença entre a Moda e como atributo de partição secundário a diferença entre o Máximo nos dois tipos de *clusters*. A Figura 3 disponibiliza o fluxograma de decisão com as regras obtidas a serem aplicadas às novas imagens com objetivo de indicar a aptidão das mesmas ao processamento automático ou semi-automático.

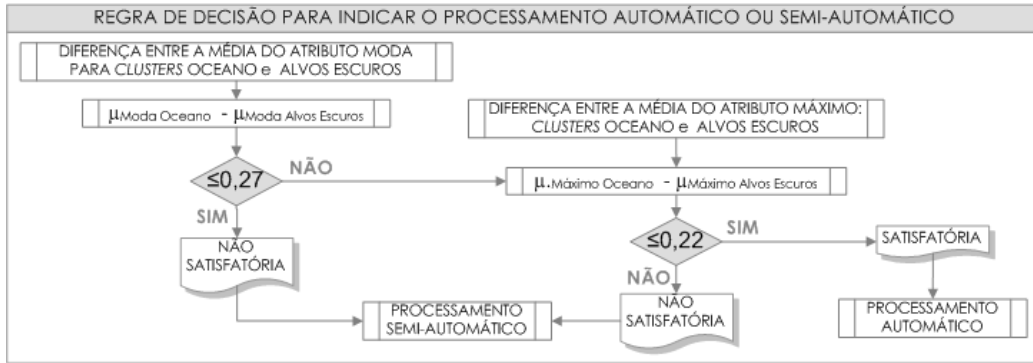


Figura 3 – Regra de decisão para definir o processamento como automático ou semi-automático.

Para validar esta regra, duas novas imagens foram pré-processadas, segmentadas e agrupadas considerando os parâmetros padronizados e pré-definidos nos itens 1 e 2. Após as clusterizações as métricas de avaliação foram calculadas em relação aos *Phantons* - delineados manualmente - para verificar se as imagens seriam aptas ou não ao processamento automático. A Tabela 2 disponibiliza as métricas de avaliação e o desempenho do clusterizador considerando como parâmetro uma *ExAE* maior ou igual 0,5.

Tabela 2 - Métricas desempenho do clusterizador calculados para os 2 novos recortes de cena.

Cenas p/ Validação	Método Utilizado	ExAE	Finterse- ção AE	Erro de Omissão	Erro de Inclusão	Desempenho do Clusterizador
Cena A	Clustan	0,46	0,93	0,07	0,10	<b>Não Satisfatório</b>
Cena B	Clustan	0,50	0,88	0,12	0,08	<b>Satisfatório</b>

Os atributos de Diferença entre a Média dos *clusters* de alvos escuros e oceano para o atributo de Moda e de Máximo foram calculados para serem confrontados com a regra de decisão pesquisada. Os valores obtidos para as duas imagens utilizadas validaram a regra de decisão, sendo o processamento semi-automático indicado para o recorte de cena A e o processamento automático indicado para o recorte de cena B (Tabela 3).

Tabela 3 – Atributos de validação calculados para os 2 recortes de cena A e B.

Cenas de Validação	Dif. Moda	Dif. Máximo	ExAE	Desempenho do Clusterizador	Indicação de Processamento
Cena A	0,26	0,10	0,46	<b>Não Satisfatório</b>	<b>Semi-automático</b>
Cena B	0,35	0,18	0,50	<b>Satisfatório</b>	<b>Automático</b>

Os resultados evidenciaram o potencial da utilização de regras de decisão para auxiliar no processo de automatização, indicando situações em que a intervenção humana é necessária para produzir geometrias mais representativas para os alvos escuros. A Figura 4 apresenta um esquema síntese do fluxo de processamento do sistema híbrido proposto para a detecção de alvos escuros em imagens SAR a ser futuramente acoplado com a etapa de classificação dos alvos escuros em óleo ou assemelhados.

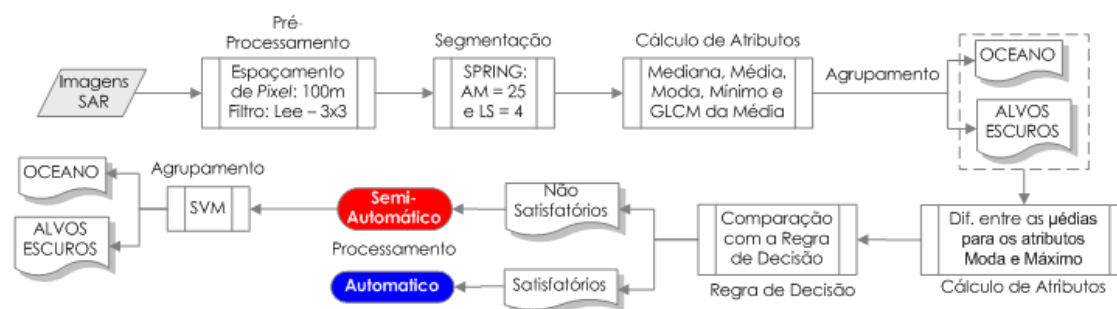


Figura 4 – Sistema híbrido inteligente proposto para a detecção automática de alvos escuros.

#### 4 – Conclusões e Perspectivas Futuras:

Os resultados evidenciaram o potencial das regras de decisão para auxiliar no processo de detecção de alvos escuros, sendo recomendável a utilização de mais recortes de cenas, além dos utilizados, para retornar regras mais robustas aplicáveis a toda nova imagem SAR adquirida. A implementação de um sistema híbrido permitirá o processamento automático nas situações em que as imagens forem favoráveis, bem como a intervenção humana em cenas desfavoráveis. O encontro SEASAR (2010), que reuniu especialistas focados no desenvolvimento científico e na aplicação operacional de sistemas para detecção de óleo, consolidou a conclusão de que: “Sistemas completamente automáticos para a detecção de óleo ainda não são plenamente operacionais.” Esta conclusão reforça o caráter inédito do sistema híbrido proposto e, considerando a relevância do tema e o potencial dos dados SAR, a continuidade e o aprimoramento dos estudos desenvolvidos como perspectiva futura.

#### Referências Bibliográficas:

- Bentz, C. M. Reconhecimento Automático de Eventos Ambientais Costeiros e Oceânicos em Imagens de Radars Orbitais [Rio de Janeiro], XIII, 115 p., COPPE/UFRJ, Tese de Doutorado, Engenharia Civil, 2006.
- Bins, L., L. Fonseca, and G. Erthal. Satellite Imagery Segmentation: a region growing approach. *Anais in VIII Brazilian Symposium on Remote Sensing*. São José dos Campos, BR: INPE, 1996.
- Brekke C., Solberg A. H. S. Oil spill detection by satellite remote sensing – Review. *Remote Sensing of Environment* (95), pp 1-13, 2005.
- Câmara, G., et al., SPRING: “Integrating Remote Sensing and GIS with Object-Oriented Data Modelling”. *Computers and Graphics*, 15(6): p. 13-22, 1996.
- Genovez, P. C.; Freitas, C. C.; Bentz, C. M.; Freitas, R. M.; Dutra, L. V. Detecção de Óleo e Ocorrências Ambientais em Imagens SAR: Avaliação de Diferentes Parâmetros de Pré-processamento e Diferentes Segmentadores. *Anais XIV Simpósio Brasileiro de Sensoriamento Remoto*, Natal - RN, Abril de 2009.
- Genovez, P. C. Segmentação e classificação de imagens SAR aplicadas à detecção de alvos escuros em áreas oceânicas de exploração e produção de petróleo. [Rio de Janeiro], XIII, 230 p., COPPE/UFRJ, Tese de Doutorado, Engenharia Civil, Abril de 2010
- Haralick, R.M. “Statistical and structural approaches to texture”, *Proceedings of the IEEE*, v. 67, n. 5, pp. 786-804, 1979.
- Hughes, G. On the Mean Accuracy of Statistical Pattern Recognizers. *IEEE Transactions On-Information Theory*, January, 1968
- Karathanassi, V.; Topouzelis, K.; Pavlakis, P.; Rokos, D. An object-oriented methodology to detect oil spills. *International Journal. of Remote Sensing*. 27, 5235-5251, 2006.
- Keramitsoglou I., Cartalis C., Kiranoudis C. Automatic Identification of oil spills on satellite images. *Environmental Modelling & Software* (21), pp 640-652, 2006.
- Pelizzari, S.; Dias, J.M.B. Bayesian Adaptive Oil Spill Segmentation of SAR Images Via Graph Cuts, 2003 [online] <[http://earth.esa.int/workshops/seasar2006/participants/256/paper\\_SoniaPelizzari\\_SEASAR2006\\_final.pdf](http://earth.esa.int/workshops/seasar2006/participants/256/paper_SoniaPelizzari_SEASAR2006_final.pdf)>, acesso março 2007.
- Solberg A., Clayton P., Indregard M. D2 – **Report on benchmarking oil spill recognition approaches and best practice. Kongsberg Satellite Services** – Norway Archive No.: 04-10225-A-Doc, 2005.
- Topouzelis, K.; Stathakis, D.; Karathanassi, V. Investigation of genetic algorithms contribution to feature selection for oil spill detection, *International. Journal o Remote Sensing*, Vol. 30, N 3, pp. 611-625, 2009.
- Wishart, D. **Focal Point Clustering User Guide**. Published by Clustan limited, Edinburgh Road, Scotland, Second Edition, 2004.