



A EFICÁCIA DA SEGMENTAÇÃO POR CRESCIMENTO DE REGIÕES NA IDENTIFICAÇÃO DE PISTAS DE POUSO NÃO PAVIMENTADAS MEDIDA PELO ÍNDICE JACCARD

Aluno Esp Fot Tiago Lopes Bezerra¹, Aluno Esp Fot Felipe Americo Amaral Coradesque¹, Cap Esp Fot Marielcio Gonçalves Lacerda², 1º Ten QOCN EST Juliane Venturelli Silva Lima¹

¹ Centro de Instrução e Adaptação da Aeronáutica

² Instituto de Estudos Avançados

RESUMO

Identificar uma pista de pouso não homologada em imagens de satélites da vasta extensão territorial do ambiente amazônico é uma árdua tarefa para os órgãos fiscalizadores, mas que pode se tornar mais viável utilizando-se técnicas de segmentação de imagens. Neste trabalho objetivou-se medir, por meio do índice Jaccard, a eficácia da segmentação por crescimento de regiões na identificação de pistas de pouso de solo não pavimentado em imagens satelitais da região amazônica. Para tanto, foi utilizado o software SPRING para segmentar imagens do satélite CBERS 4 das referidas pistas. Essas segmentações foram medidas no software QGIS e comparadas a verdades terrestres geradas manualmente nesse mesmo programa. Então, os resultados foram analisados em função do índice de similaridade de Jaccard obtido. Análises realizadas indicam que, com 95% de certeza, a segmentação por crescimento de regiões apresenta aproximadamente 67% de eficácia na segmentação de pistas de pouso em condições semelhantes às estudadas, o que demonstra que a metodologia abordada neste trabalho pode ser uma excelente ferramenta de auxílio na busca de pistas de pouso na região amazônica.

Palavras-chave: Segmentação por Crescimento de Regiões. Índice Jaccard. Pistas de pouso não homologadas.

ABSTRACT

Identifying an unapproved airstrip in satellite images from the vast territorial extension of the Amazonian environment is an arduous task for the inspection agencies, but it can become more viable using image segmentation techniques. This study aimed to measure, using the Jaccard index, the effectiveness of segmentation by growing regions in identifying airstrips of unpaved soil in satellite images of the Amazon region. For that, SPRING software was used to segment CBERS 4 satellite images of the referred airstrips. These segmentations were measured in the QGIS software and compared to terrestrial truths manually generated in that same program. Then, the results were analyzed according to the Jaccard similarity index obtained. The performed analyzes indicate that, with 95% certainty, the segmentation by growth of regions presents approximately 67% of efficiency in the segmentation of

airstrips in conditions similar to those studied. This demonstrates that the methodology addressed in this work can be an excellent tool to help in the search for airstrips in the Amazon region.

Keywords: Region Growth Segmentation. Jaccard Index. Non-certified airstrip.

1. INTRODUÇÃO

Pistas de pouso não homologadas são construídas nos vazios demográficos amazônicos para apoiar diversas atividades ilegais, principalmente aquelas relacionadas ao narcotráfico e ao tráfico de armamentos, de modo a possibilitar aos traficantes distribuir seu produto ilegal e fugir da repressão estatal por meio de rotas aéreas que necessitam do apoio de tais pistas de pouso que, por serem construídas na região amazônica, são de difícil localização (COUTO; OLIVEIRA, 2017).

Aeronaves de Reconhecimento e satélites imageadores têm sido amplamente empregados na localização de pistas não homologadas, os quais obtêm imagens em muito alta e altíssima resolução (VHR – *Very High Resolution* e UHR – *Ultra High Resolution*) (EHLERS; JANOWSKY; GÄLHER, 2002). Entretanto os custos inerentes ao seu uso, associados à grande extensão do território amazônico, à pequena dimensão de uma pista e sua distribuição aleatória, fazem com que a identificação de tais pistas, exclusivamente por meio da análise visual, seja uma tarefa árdua, mas que pode ser facilitada com o processamento digital de imagens (SILVA, 2013).

Nesse cenário, técnicas de processamento digital de imagens que se baseiam na segmentação, como análise autônoma de imagens, podem diminuir o tempo de identificação dessas pistas (GONZALEZ E WOODS, 2010), isso porque, na segmentação, a imagem é subdividida em objetos ou regiões (LACERDA, 2019), as quais são definidas a partir do agrupamento de pixels adjacentes em áreas homogêneas (RADOUX et al., 2019), gerando polígonos irregulares identificados como um único objeto (XAUD et al., 2011).

Existem diversas técnicas de segmentação de imagens, mas, de acordo com a literatura pesquisada, entre os algoritmos que as descrevem, o de crescimento de regiões é preferível nas aplicações de Sensoriamento Remoto, pois considera a característica espacial dos dados e garantem a formação de segmentos fechados (ESPINDOLA, 2006). Por esse motivo, estudar sua eficácia, visando à identificação de pistas não homologadas, pode fomentar o uso desses recursos digitais e tornar a localização de tais pistas mais efetiva (GONZALEZ E WOODS, 2010).

É necessário que o resultado seja analisado

para possibilitar a mensuração da precisão obtida. Uma das métricas que pode ser utilizada para medir a eficácia de algoritmos de segmentação é o índice Jaccard, que é um coeficiente utilizado para comparar similaridade de um conjunto de dados (NIWATTANAKUL, 2013).

Em face do exposto, este trabalho realizou uma pesquisa descritiva a fim de analisar a probabilidade de erro em função do Índice de Similaridade de Jaccard, de modo a determinar a eficácia do algoritmo de segmentação por crescimento de regiões em imagens de pistas não pavimentadas em ambiente amazônico, obtidas a partir de sensores satelitais.

2. REFERENCIAL TEÓRICO

Na região amazônica do país, constituída predominantemente de florestas latifoliadas, fechadas, e de vazios demográficos que dificultam a repressão estatal, as ações criminosas que se favorecem desse contexto são de destacada relevância, principalmente no que se refere ao tráfico de drogas e armas (COUTO; OLIVEIRA, 2017).

Segundo Couto e Oliveira (2017), a Amazônia é a rota primária da distribuição de cocaína para a Europa e África. Os autores afirmam ainda que, para atravessar drogas e armas pela região, os narcotraficantes fazem uso de hidrovias, rodovias e de rotas aéreas.

Essa região faz fronteira com sete países e, durante os períodos chuvosos, seu acesso terrestre torna-se impossibilitado devido ao alagamento, fazendo necessário o uso de aviões de pequeno porte, o principal modal logístico (SILVA, 2013).

Em apoio às suas operações aéreas, os criminosos utilizam-se de pistas não homologadas, construídas no meio da floresta, ou de pistas de fazendas. (COUTO; OLIVEIRA, 2017).

Para Silva (2013), essas pistas são elementos de fundamental importância porque, além do narcotráfico, apoiam também outras atividades ilícitas, tais como o tráfico de armas, a biopirataria, os garimpos ilegais e a exploração madeireira.

Diante da importância dessas pistas de pouso para as atividades ilícitas e da dificuldade de identificá-las, estudos como o de Alves et al. (2009) e de Silva (2013) discutem, respectivamente, formas de localizá-las por meio do uso de aeronaves remotamente pilotadas e da aplicação de técnicas de processamento digital de imagens (PDI) em imagem Synthetic Aperture Radar (SAR).

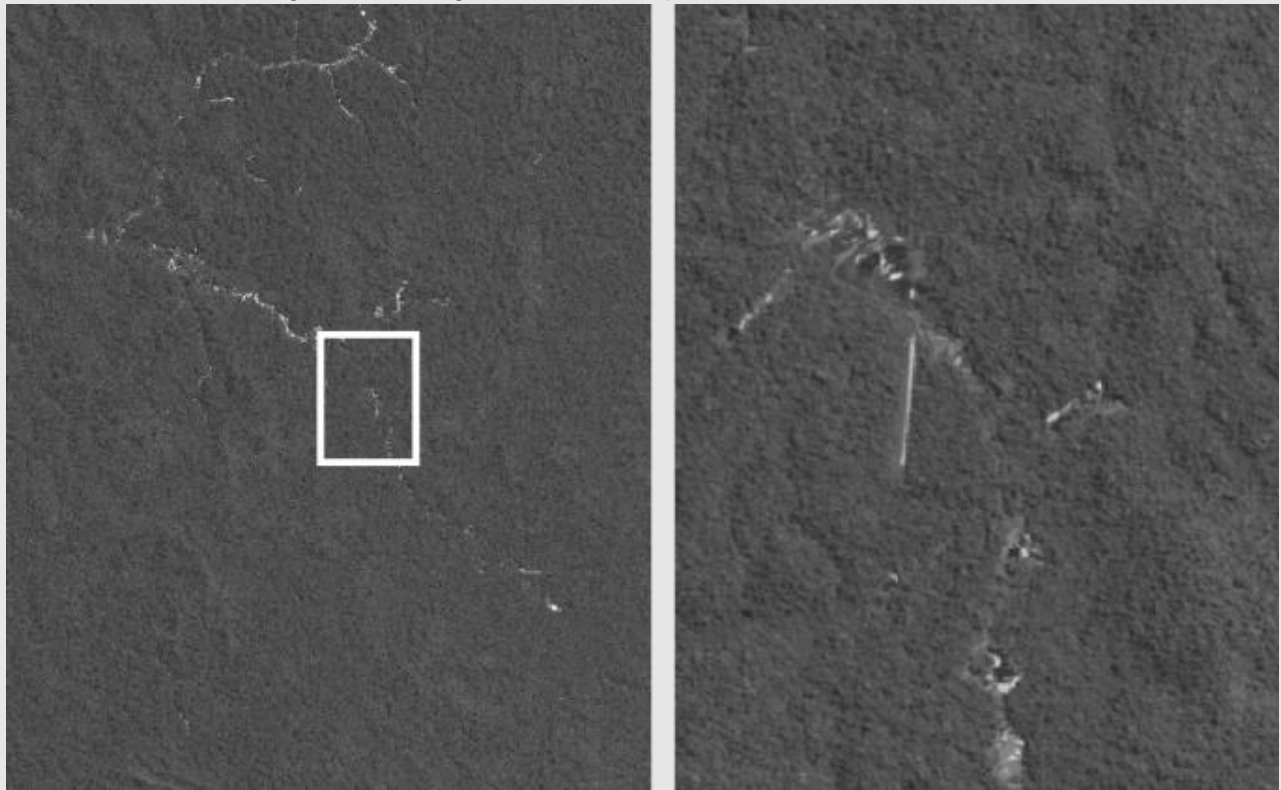
O uso do processamento digital é de especial importância, pois contribui para que as limitações da capacidade de interpretação visual humana sejam superadas. Como demonstrado na imagem satelital CBERS 4, apresentada na

Figura 1, a pista de pouso fica visualmente aparente somente quando a imagem é ampliada sobre sua exata posição, devido às pequenas dimensões que possui em relação à área imageada.

O Programa CBERS é uma parceria entre

Brasil e China no setor técnico-científico espacial para a geração de dados primários de Sensoriamento Remoto. O empreendimento já contempla seis satélites lançados, chamados de CBERS 1, 2, 2B, 3, 4 e 4A (INPE, 2018), os quais produzem imagens, inclusive da região amazônica.

Figura 1 – Imagem satelital de pista em diferentes escalas



Fonte: Os autores

No CBERS 4, a câmera PAN conta com quatro bandas espectrais. A B01 tem 5 m de resolução espacial e sua faixa espectral está delimitada entre 0,51 e 0,85 μm . Já as demais bandas têm resolução espacial de 10 m e suas faixas espectrais estão delimitadas da seguinte forma: B02, de 0,52 a 0,59 μm ; B03, de 0,63 a 0,69 μm ; e B04, de 0,77 a 0,89 μm (INPE, 2019).

Os algoritmos de segmentação têm sido amplamente empregados na extração de informações de imagens de sensores remotos (ESPÍNDOLA, 2006), como os produzidos no CBERS 4, e sua importância decorre do impacto sobre o sucesso final dos procedimentos de análise computadorizada (GONZALEZ; WOODS, 2010).

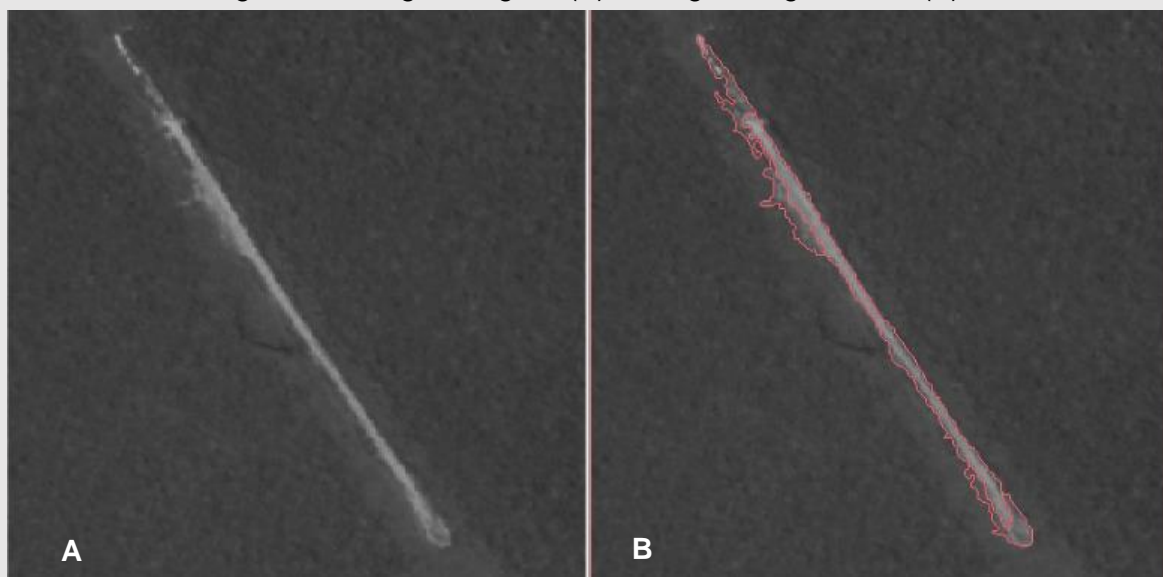
Além disso, devido à vasta extensão do ambiente amazônico, a segmentação ganha especial relevância por ser, conforme Zhang (1996), uma das tarefas mais críticas na análise automática de imagens, caracterizada pela possibilidade de ser executada sem a necessidade da presença de um técnico ou operador.

A segmentação, conforme explica Saldanha e Freitas (2009), é uma técnica de processamento digital que consiste em reduzir a quantidade de informações de imagens, dividindo-a em regiões aproximadamente homogêneas, visando permitir a rápida extração de informações importantes.

Já nas palavras de Gonzalez e Woods (2010), a segmentação subdivide uma imagem em regiões ou objetos que a compõem até que os objetos ou as regiões de interesse de uma aplicação tenham sido detectados.

Segundo Espíndola (2006), dentre os algoritmos de segmentação de imagens conhecidos, os do tipo “crescimento de regiões” são preferíveis em aplicações que empregam sensores imageadores remotos, pois consideram a característica espacial dos dados e garantem a formação de segmentos fechados. Essa característica pode ser observada na Figura 2.

Figura 2 – Imagem original (A) e imagem segmentada (B)



Fonte: Os autores

A segmentação por crescimento de regiões consiste em um algoritmo de processamento de imagens que agrupa os pixels em regiões maiores através da junção de pixels vizinhos, utili-

zando critérios de homogeneidade em propriedades similares como características espectrais e espaciais (CLAYTON; NEVES, 2001).

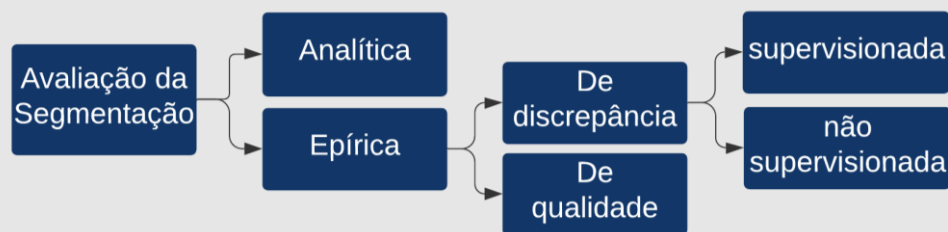
Para tanto, é comum que esse tipo seg-

mentação determine inicialmente um conjunto de pixels “semente” e, a partir deles, promova o surgimento e o crescimento de regiões, unindo a cada semente os pixels vizinhos que tenham uma determinada propriedade predefinida, como intervalos de intensidade ou cor, semelhante às das sementes. (GONZALEZ e WOODS, 2010).

Devido à sua importância, muito esforço foi

empregado no desenvolvimento de técnicas de segmentação (ZHANG; GERBRANDS, 1994). O surgimento dessas diferentes técnicas fomenta o desenvolvimento de procedimentos através dos quais as segmentações possam ser comparadas e avaliadas. Os tipos de avaliação de segmentação descritos por Zhang (1996) estão resumidos na Figura 3.

Figura 3 – Tipos de segmentação



Fonte: Os autores

Tais avaliações, segundo Zhang (1996), podem examinar o algoritmo analítica ou empiricamente. As avaliações analíticas são aquelas aplicadas diretamente sobre os princípios e propriedades dos algoritmos de segmentação, sem levar em conta o produto final. As empíricas, entretanto, julgam os algoritmos através dos resultados que produzem.

As avaliações empíricas ainda podem ser classificadas em de qualidade ou de discrepância. As avaliações de qualidade procuram, no resultado da segmentação, propriedades estabelecidas previamente de maneira arbitrária. Já as de discrepâncias comparam os resultados com referências que representem respostas ide-

ais ou esperadas (ZHANG, 1996).

As técnicas de avaliação de discrepâncias se dividem em dois métodos: avaliação não supervisionada e avaliação supervisionada, também chamada de método relativo de avaliação (CHABRIER *et al.*, 2004). O método relativo de avaliação verifica a eficácia do algoritmo de segmentação, comparando seus resultados com uma segmentação realizada manualmente, através de métricas de validação conhecidas como coeficientes (ou índices) de similaridade (ZHANG, 2008).

Para esse trabalho, a métrica de validação empregada foi o Índice de Similaridade de Jaccard, que consiste na razão entre a interseção e

a união do conjunto de amostras com a segmentação manual, chamada aqui de vetorização,

sendo o valor do coeficiente obtido através da Equação 1.

$$Jaccard(A,B) = |A \cap B| / |A \cup B|, \quad (1)$$

na qual “A” é a segmentação delineada manualmente do contorno da pista e “B” é a segmentação realizada pelo algoritmo de segmentação.

Sendo o índice de Jaccard uma probabilidade, seus possíveis valores variam entre 0 (zero), indicando nenhuma sobreposição dos conjuntos de pixels, e 1 (um), representando sobreposição total dos conjuntos de pixels (THADA, 2013).

3. MATERIAIS E MÉTODOS

A seleção das amostras e o processamento dos dados foram realizados conforme descrito a seguir.

Inicialmente, foram selecionadas imagens pancromáticas de alvos, obtidas pelo satélite CBERS 4 na banda B01, oferecidas no catálogo do INPE - Instituto Nacional de Pesquisas Espaciais. Para o passo seguinte, foi selecionado o software SPRING (CÂMARA et al., 1996), um banco de dados geográfico, desenvolvido pelo Instituto Nacional de Pesquisas Espaciais (INPE), para ambientes Windows e Unix (SANTOS, PELUZIO; SAITO, 2010). Dentre as características do SPRING que o tornaram a ferramenta

selecionada para este trabalho, está o fato do software ser tecnologia nacional de domínio público, podendo ser adquirida gratuitamente no sítio online da Divisão de Processamento de Imagens (DGI) do INPE. Criou-se, então, no SPRING, um banco de dados com as imagens georreferenciadas.

Feito isso, passou-se à definição dos parâmetros de segmentação, a saber, a similaridade e a área mínima. Definidos os parâmetros, passou-se à segmentação propriamente dita das imagens, seguida de sua exportação em arquivo do tipo *shapefile*.

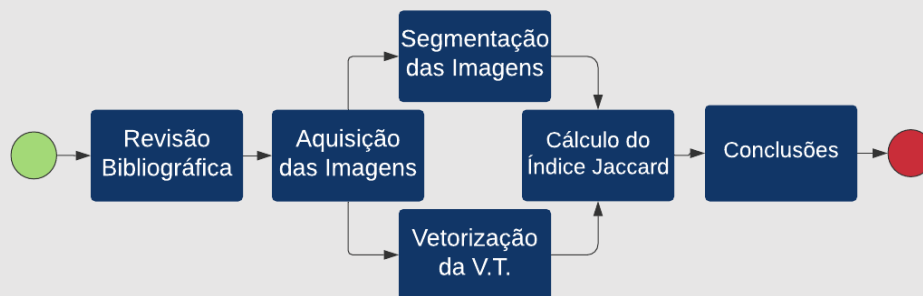
Em seguida, para as operações e medições dos segmentos gerados, os arquivos foram importados no QGIS, Sistema de Informações Geográficas (SIG) que permite análise de dados espaciais e a manipulação de dados matriciais e vetoriais, para que fossem realizadas as vetorizações dos alvos manualmente. As medições e manipulações com segmentos foram realizadas no QGIS pelo fato de o SPRING não realizar operações, como a interseção ou diferença de vetores. Dessa forma, os vetores gerados no SPRING precisaram ser exportados para o QGIS

para que, a partir de então, fossem processados. Assim como o SPRING, o QGIS é um software aberto e de distribuição gratuita (ALMEIDA, 2011).

Por fim, juntou-se no QGIS, o arquivo vetorizado manualmente e a segmentação gerada no

SPRING e compararam-se os dois resultados utilizando, para isso, o índice de similaridade de Jaccard. Um fluxograma que resume as etapas realizadas na metodologia pode ser visualizado na Figura 4.

Figura 4 – Diagrama da metodologia utilizada



Fonte: Os autores

3.1 Imagens CBERS4 PAN

Neste trabalho, foram utilizadas imagens obtidas a partir do sensor PAN (Pancromático) embarcado no satélite CBERS 4.

A B01 foi a banda selecionada para esse trabalho, não apenas por possuir a mais alta resolução espacial disponível na plataforma, mas também pela sua alta correlação com a experiência visual humana, o que favorece o processo de análise visual das imagens.

A fim de manter a clareza e a objetividade do texto, deste ponto em diante, as imagens da banda B01 da câmera PAN do satélite CBERS 4, serão identificadas como CBERS 4 PAN 5M, fazendo-se as distinções de data de aquisição e

órbita/ponto quando necessário.

As imagens foram adquiridas digitalmente, tendo as solicitações sido realizadas por meio do catálogo de imagens do INPE. O serviço é mantido pela Divisão de Geração de Imagens (DGI) e pode ser acessado pelo endereço eletrônico <<http://www.dgi.inpe.br/catalogo/>>.

3.2 Amostragem

Inicialmente foram analisadas pistas de uma lista de alvos da disciplina de Geoprocessamento Aplicado à Inteligência Operacional, do Curso de Formação de Oficiais Especialistas em Fotografia (CFOE-FOT). Posteriormente, por meio da análise visual de imagens CBERS 4 PAN 5M,

outras pistas foram identificadas em posições próximas às anteriores.

Os critérios estabelecidos para a seleção das pistas de pouso que constituíram a amostra foram os seguintes:

1. Localizadas na Região Amazônica;
2. Contorno delimitado por vegetação nativa;
3. Superfície constituída de solo natural ou estabilizado;
4. Dimensões aproximadas de 20 m de largura e 300 m de comprimento;
5. Disponibilidade de imagens pancromáticas da câmera PAN do CBERS 4; e
6. Ausência de interferências atmosféricas significativas na área do alvo no momento da obtenção da imagem.

Considerando esses critérios e a disponibilidade de imagens satélites, foram selecionadas dezoito pistas. Esses alvos foram codificados da seguinte maneira: T01 a T03 para as imagens utilizadas como treino; S01 a S15 para as imagens utilizadas nos testes, conforme descrito no Apêndice A.

3.3 Segmentação

Para iniciar o processo de segmentação, criou-se, no *software* SPRING, um banco de dados do tipo *SQLite* e um projeto no Sistema Coordenado de Referências UTM, Zona 20S e Datum WGS84. Esses parâmetros são os mesmos aplicados pela DGI do INPE na geração das imagens CBERS 4 PAN 5M.

Após a importação das imagens para o sistema, foram selecionados três alvos (T01, T02 e T03) para o cálculo da Similaridade. No SPRING, duas regiões são consideradas distintas se o valor médio dos seus *pixels* for superior à Similaridade estabelecida.

Assim, para cada um dos três alvos, foram medidos aleatoriamente os valores de vinte e cinco pixels da área da pista de pouso e de outros vinte e cinco pixels das adjacências. Então, foram calculados o valor médio para área da pista (MED-P), o valor médio para as adjacências (MED-A) e a diferença desses valores. A Tabela 1 mostra os resultados obtidos em valores de números digitais que, nesse caso, são grandezas adimensionais.

Tabela 1 – Cálculo de Similaridade

Alvo	MED-P	MED-A	DIFERENÇA
T01	128.08	80.56	47.52
T02	84.16	67.88	16.28
T03	82.4	67.92	14.48

Fonte: os autores.

Outro parâmetro necessário à segmentação é a Área, que determina o tamanho mínimo, em pixels, que uma região segmentada precisa ter para que não seja absorvida por outra adjacente. Para o cálculo desse valor, considerou-se inicialmente o proposto por Alves *et al.* (2009). Nessa pesquisa, os autores afirmam que as pistas de pouso utilizadas pelo crime organizado na Amazônia precisam ter, no mínimo, 20 m de largura e 300 m de comprimento. Considerando que as imagens CBERS 4 PAN 5M têm resolução espacial de 5 m, essas dimensões podem ser expressas como 4 x 60 pixels, totalizando uma área de 240 pixels.

Após a definição desses parâmetros, as imagens que continham os alvos S01 a S15 foram recortadas em retângulos de, no mínimo, 10 km de lado. Cada recorte continha, pelo menos, um desses alvos e foi submetido a sucessivas segmentações a fim de aferir os valores de Similaridade e Área.

Durante os testes com as imagens recortadas, a combinação de valores que implicou em mais segmentações satisfatórias foi: Similaridade igual a 13 cm e Área igual a 320 pixels. Sendo assim, esses foram os parâmetros utilizados para segmentar as imagens completas.

Com os parâmetros previamente estabelecidos, e sem qualquer tipo de corte, correção,

tratamento ou processamento, foram segmentadas as seguintes imagens:

CBERS_4_PAN5M_20190812_169_107_L4_BA ND1;
CBERS_4_PAN5M_20200715_169_109_L4_BA ND1; e
CBERS_4_PAN5M_20190812_169_107_L4_BA ND1.

Concluída a segmentação, as regiões obtidas foram extraídas do SPRING através da exportação de arquivos *shapefile* que, posteriormente, foram importados no QGIS.

3.4 Vetorização

No QGIS, foram importados os mesmos arquivos das imagens CBERS_4_PAN5M segmentadas integralmente no SPRING. O projeto criado para isso foi configurado com o mesmo Sistema de Coordenadas de Referências do projeto do SPRING, ou seja, projeção UTM Zona 20S e Datum WGS84.

Também no QGIS, as áreas equivalentes aos alvos S01 a S15 foram vetorizadas manualmente. O objetivo deste procedimento era criar uma referência a ser comparada com a segmentação automatizada realizada pelo SPRING.

Nas regiões criadas, a vetorização manual incluiu não somente a área da pista propriamente dita, mas também outras áreas adjacentes a ela, que apresentavam respostas radiométricas suficientemente semelhantes à da respectiva pista,

por considerar-se que toda a área pertence à pista em análise, incluindo a região de apoio à operação das aeronaves.

As vetorizações realizadas foram armazenadas em arquivos do tipo *shapefile*, sendo um para cada alvo selecionado, totalizando, portanto, quinze arquivos.

3.5 Cálculo do índice Jaccard

Concluída a vetorização manual, os arquivos tipo *shapefile*, contendo as regiões criadas pela segmentação automatizada do SPRING, foram importados para o software QGIS, no mesmo projeto previamente criado para as imagens CBERS 4 PAN 5M.

Para realizar as operações entre as regiões manuais e automatizadas, foi necessário extrair, no QGIS, os segmentos automatizados para arquivos individuais. Esse procedimento também gerou um arquivo *shapefile* para cada alvo, totalizando quinze arquivos.

A etapa seguinte foi a execução das operações de interseção e união dos arquivos, correspondentes à vetorização manual e segmentação automatizada respectiva a cada alvo. Em ambas as operações, o arquivo do SPRING foi inserido no *Input Layer*, e o arquivo da vetorização manual, no *Overlay Layer*, mesmo que não tenha sido verificada diferença no processo inverso.

Especificamente após o processo de união, foi necessário realizar o comando “Dissolver” em cada arquivo, pois o QGIS armazena regiões unidas na forma de conjuntos disjuntos. Isso significa que, dadas as regiões A e B, o resultado da união seriam as regiões $A \cap B^c$, $B \cap A^c$ e $A \cap B$, sendo C o complementar do evento. O comando “Dissolver” resolve a condição provocada pelo processo de união, gerando uma única região que soma a mesma área e tem os mesmos atributos do total das regiões disjuntas previamente mencionadas.

Depois que a união e interseção foram concluídas, os quinze arquivos resultantes de cada processo foram consolidados em um *shapefile* para cada operação. Feito esse ajuste, as tabelas de atributos foram exportadas como planilhas, sendo essas consolidadas em um único arquivo. Com os valores das áreas de interseção e união das regiões, foi calculado o índice Jaccard de cada alvo, conforme apresentado na Equação 1.

3.5 Análise do índice Jaccard

Para a análise inferencial do índice Jaccard, foi utilizado o método *Bootstrap* (CASELLA; BERGER, 2002). Este procedimento estatístico consiste em reamostrar N vezes, com reposição, a amostra original. Então, para cada amostra obtida, calcula-se a estatística de interesse, que

neste trabalho foi a média amostral.

As N médias amostrais obtidas são, então, utilizadas para estimar a distribuição das médias amostrais. Desta forma, utiliza-se a distribuição aproximada obtida para a construção de intervalos de confiança e testes de hipóteses.

Uma das vantagens desse método é que não há suposição de normalidade dos dados, usualmente presente nos testes paramétricos. A principal suposição desse método, porém, é que a amostra original represente bem a população alvo.

As amostras *Bootstrap* do índice Jaccard foram obtidas por meio do software R, sendo utilizada a semente aleatória 061285 para realizar 100.000 reamostragens da amostra original.

Chamam-se sementes os pixels selecionados inicialmente. A partir dessas sementes, as

regiões crescem à medida que vão sendo agregadas a pixels ou a sub-regiões vizinhas que atendam a determinado critério de heterogeneidade. As sementes podem ser selecionadas de maneira aleatória, determinística ou pelo usuário (PEDRINI; SCHWARTZ, 2008).

3.6 Índice de circularidade

O índice de circularidade é um parâmetro matemático que permite inferir sobre a forma de uma determinada figura, pois o valor calculado tende a unidade quando a forma do alvo se aproxima da do círculo e a zero quando se afasta dela (CARDOSO *et al.*, 2006), conforme a Equação 2. Na qual “A” é a área do polígono e “P” é o seu perímetro.

$$Circularidade = 12,57 \times \frac{A}{P^2} \quad (2)$$

4. RESULTADOS

A Tabela 2 exhibe, além do índice Jaccard, os valores, em metros quadrados, das regiões geradas pela segmentação e pela vetorização de cada alvo, bem como os valores das áreas de união e interseção entre elas.

Do processamento estatístico *Bootstrap* realizado sobre os valores do índice Jaccard decorrem os resultados apresentados na Tabela 3.

A média *Bootstrap* para o índice Jaccard foi igual a 0,67, com desvio padrão igual a 0,04, conforme Tabela 3. O intervalo de confiança de 95% estimado foi igual a [0,58; 0,74]. Isto é, com 95% de confiança, espera-se que a média populacional do Índice Jaccard esteja nesse intervalo.

Tabela 2 - Áreas calculadas (m²) e resultados obtidos.

Id	Área Segmentação	Área Vetorização	Área Interseção	Área União	Jaccard
S01	9859,86579	12449,58086	9489,87474	12819,57200	0,74026
S02	9001,44953	18367,65491	8953,22648	18415,87788	0,48617
S03	12351,45709	14571,28402	10788,62909	16134,11454	0,66868
S04	11869,05262	27246,48803	9536,84789	29578,69284	0,32242
S05	9139,36762	9359,60450	7660,74514	10838,24992	0,70682
S06	10533,43387	10061,38888	8754,75824	11840,06529	0,73942
S07	10729,55981	11545,37636	9209,09076	13065,84581	0,70482
S08	20533,76199	21036,82023	18478,84084	23091,73999	0,80024
S09	29878,21614	31475,52604	27505,48097	33848,26592	0,81261
S10	15319,22722	21876,07503	14503,08498	22692,21698	0,63912
S11	15030,14710	11708,22435	9414,86820	17323,50828	0,54347
S12	13673,06660	18869,57449	12519,10737	20023,53488	0,62522
S13	20507,28182	21140,32944	17806,79372	23840,81675	0,74690
S14	43260,40996	39255,85736	35688,03221	46828,23431	0,76211
S15	40514,36903	39429,95186	34621,80714	45322,49434	0,76390

Fonte: Os autores

Tabela 3 - Resultados do *Bootstrap*.

Informação	Valor
Média	0,67
Desvio Padrão das médias	0,04
Intervalo de confiança (95%)	[0,58; 0,74]

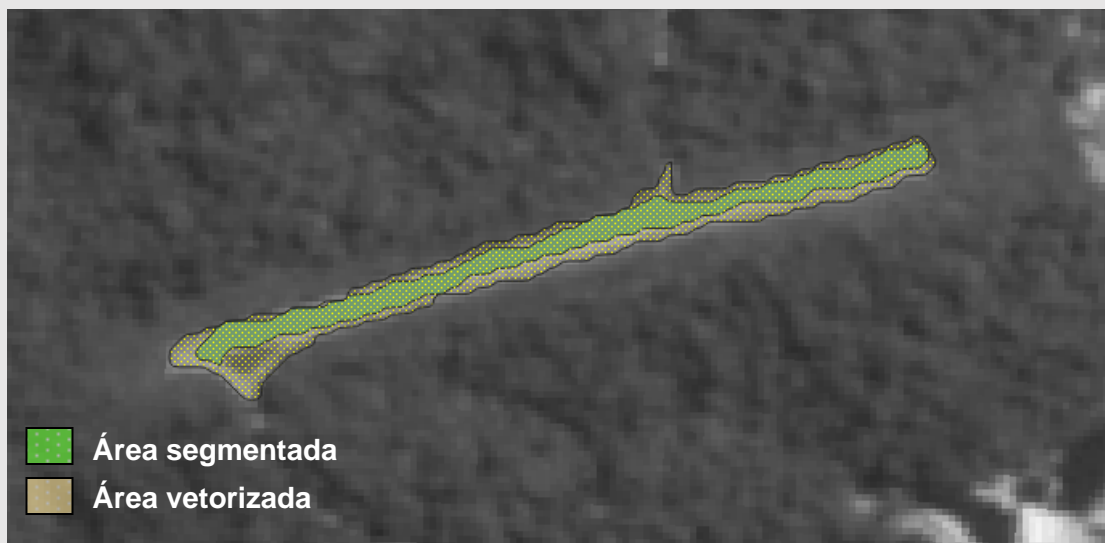
Fonte: Os autores

5. DISCUSSÃO

O primeiro aspecto a ser destacado é que as áreas vetorizadas manualmente são, em média, 34,5% maiores que as áreas segmentadas automaticamente. Dentre os quinze alvos, onze tiveram área vetorizada maior que o objeto segmentado, como exemplificado na Figura 5. Uma provável explicação para esse fato é a relação entre a forma como a vetorização foi realizada e o método empregado para determinar o parâmetro da

similaridade. Enquanto na vetorização foi incluída a área útil do aeródromo, no cálculo foram medidos apenas valores de pixels da área da pista e de fora do aeródromo, ou seja, não foram incluídos no cálculo pixels das regiões de transição. Dessa forma, pode ter sido determinado um valor de similaridade inferior ao necessário para incluir toda a área útil do aeródromo em uma única região.

Figura 5 - Áreas vetorizadas e segmentadas do alvo S02.



Fonte: Os autores

Nos alvos S06, S11, S14 e S15, as áreas segmentadas foram, respectivamente, 4,6%, 28,3%, 10,2% e 2,7% maiores que as das respectivas regiões vetorizadas, ou seja, um comportamento contrário ao previamente discutido.

Por meio da leitura de valores individuais *depixels*, foi possível concluir que, em S11 e S14, esse fenômeno foi causado por áreas nas quais a diferença de tonalidade entre pixels adjacentes era menor do que a similaridade estabelecida, isso fez com que pixels descartados na vetorização fossem incluídos na segmentação.

Já em S06 e S15, verificou-se que existiam diferenças suficientes entre os pixels do alvo e os pixels adjacentes. Além disso, essas segmentações também incluíram áreas descartadas pelas vetorizações correspondentes. Sendo assim, a explicação mais provável está na constituição

das regiões.

Além da Similaridade, a formação de uma região também depende do funcionamento do algoritmo. Como o crescimento das regiões inicia nas sementes e esses pontos determinam quais pixels devem ser avaliados para compor uma mesma região, isso pode influenciar na área e na forma de um segmento, como exemplifica a Figura 6 que traz duas segmentações com Similaridade e Área iguais, mas sementes diferentes.

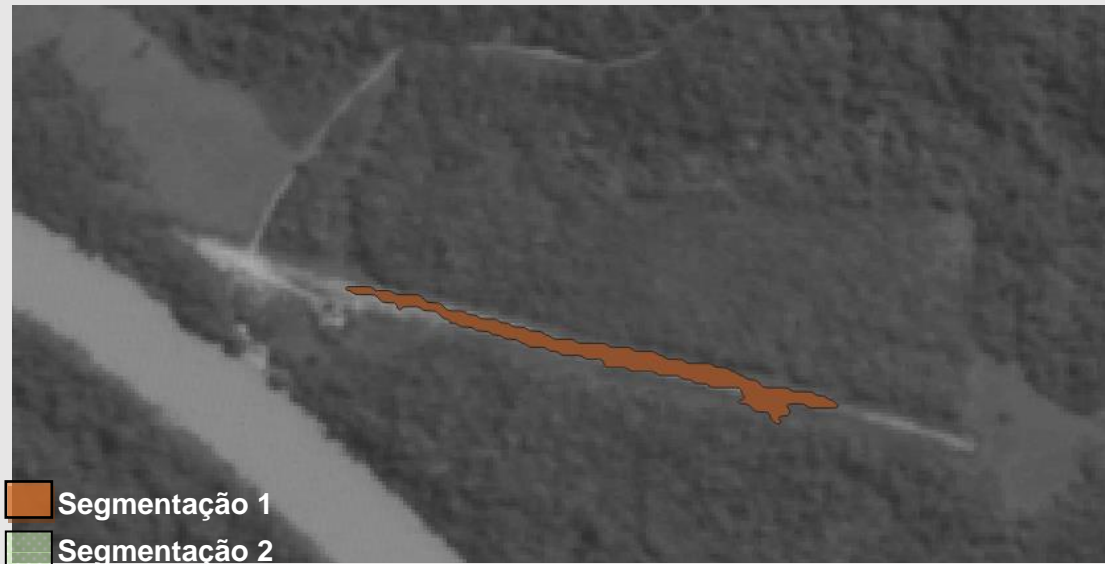
Quanto ao tamanho das áreas, as segmentações são, em média, 16,4% maiores que as interseções correspondentes. Isso significa que elas têm aproximadamente 16,4% de suas áreas disjuntas das vetorizações. A existência de áreas de disjunção é uma combinação dos mesmos fatores que influenciam no tamanho das áreas, ou seja, o valor da Similaridade e a posição das

segmentes.

No exemplo da Figura 7, é possível verificar esses dois fenômenos ocorrendo simultaneamente: a região vetorizada, hachurada em azul, é

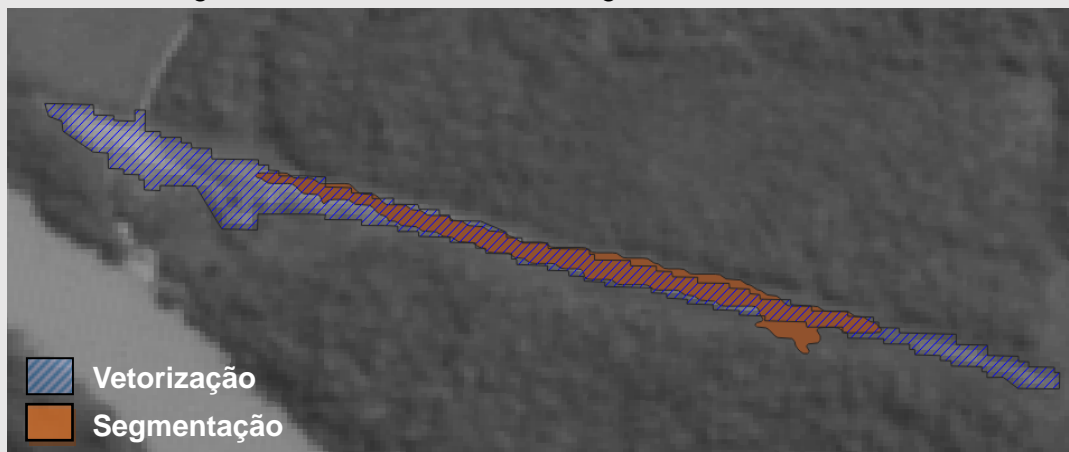
maior que a região segmentada, que está em marrom, mas essa tem áreas não contidas na anterior.

Figura 6 - Duas segmentações do alvo S04 com mesma Similaridade e Área.



Fonte: Os autores

Figura 7 - Áreas vetorizadas e segmentadas do alvo S04.



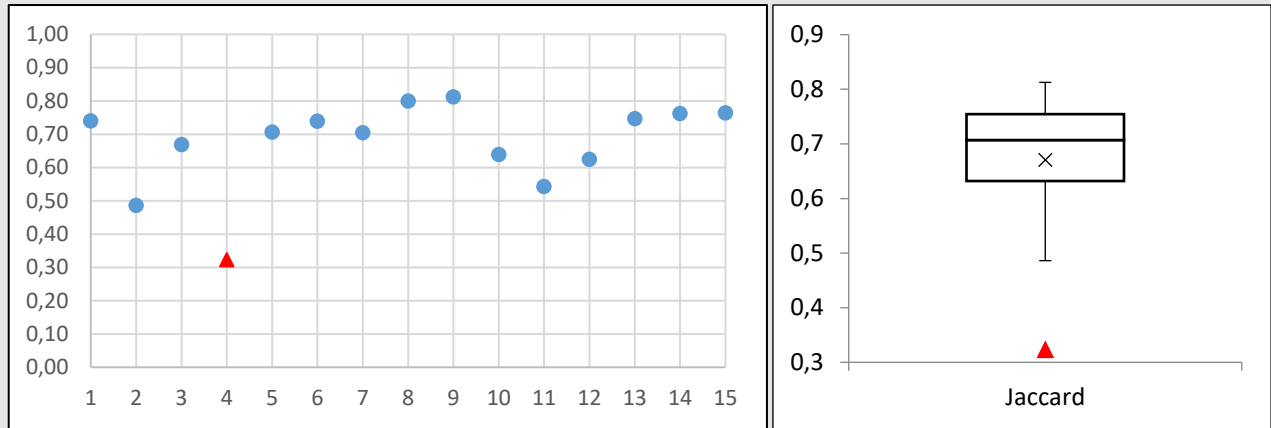
Fonte: Os autores

O tamanho relativo das áreas e as regiões de disjunção são os principais fatores que influenciam o índice Jaccard e, por este motivo,

quanto mais semelhantes os segmentos, maior é o índice obtido. Analisando-se a Figura 8, é possível verificar que, das quinze imagens analisadas, treze obtiveram o índice superior a 0,5, o

que corresponde um valor de 86,67%. Isso significa que suas segmentações e vetorizações possuem mais semelhanças que diferenças.

Figura 8—Índice Jaccard para os quinze alvos e *Boxplot*



Fonte: Os autores

À direita da Figura 8 pode ser observado que o *Boxplot* dos índices Jaccard apresenta linha central elevada (aproximadamente 0,7) e pouca amplitude, o que indica alta concentração de dados nos maiores valores da escala. O mesmo pode ser concluído verificando-se que a média desse conjunto é, aproximadamente, 0,67, e o coeficiente de variação de 6,05%.

Outro aspecto que pode ser extraído do *Boxplot* é a distribuição assimétrica dos dados. Ao se apresentar mais próxima do 3º quartil, a mediana indica que as maiores variações ocorrem nos menores valores do conjunto. O alvo que mais contribui para esse comportamento é o S04, representado como um triângulo vermelho, cujo baixo índice Jaccard caracteriza como um *outlier* nesse conjunto dados.

As segmentações realizadas no S04, de-

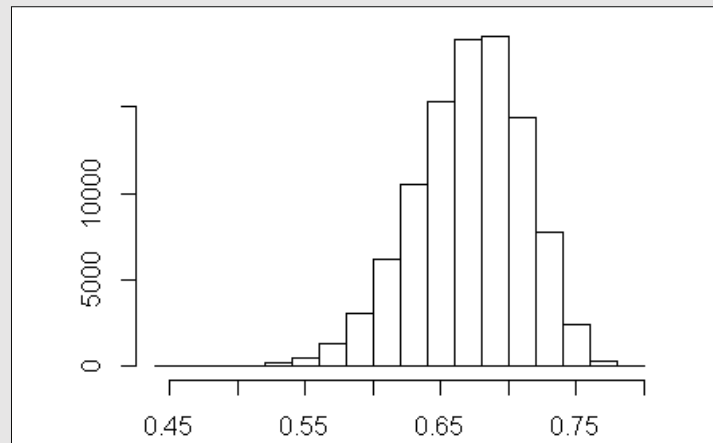
monstradas na Figura 6, indicam que a posição das sementes foi o fator que mais comprometeu o índice desse alvo. Caso as áreas das duas segmentações apresentadas fossem unidas e o resultado fosse utilizado para calcular o índice Jaccard, seria obtido o valor de 63%, que é muito próximo da média do conjunto.

Apesar do índice Jaccard baixo em relação aos demais alvos, essas segmentações preservaram a mesma forma geral da pista, ou seja, mais alongada em uma dimensão do que nas demais. Como pode ser detectado pelo seu índice de circularidade

A circularidade do S04 é de 0,08097, enquanto a do S09, alvo de maior índice Jaccard, é 0,06315. Ou seja, apesar de uma variação de mais de 150% nos índices Jaccard, a circularidade entre eles variou apenas 1,7%.

A Figura 9 mostra a distribuição das médias amostrais aproximadas para o índice Jaccard. O gráfico possui uma linha de distribuição ajustada e ligeiramente assimétrica à esquerda.

Figura 9 – Histograma da distribuição amostral para a média do índice Jaccard



Fonte: Os autores

6. CONCLUSÕES

Como discutido nesse artigo, a identificação automatizada das pistas de pouso é uma tarefa de alta relevância para a segurança nacional, mas também de alta complexidade. Portanto, o desenvolvimento de técnicas de identificação automatizada de feições que possam localizar essas pistas é uma necessidade que cresce juntamente com o aumento da disponibilidade de imagens.

Sabendo que a segmentação é uma etapa crítica da identificação automatizada de feições, verificar que o crescimento de regiões tem, para alvos semelhantes aos testados nesse artigo, uma eficácia média aproximada de 0,67, reforça a hipótese de que dominar esse processo pode

contribuir significativamente para o problema em questão.

A baixa dispersão dos dados, indicada pela pequena amplitude do *Boxplot* e pelo baixo coeficiente de variação, cerca de 6,05%, associada à linha de distribuição bem ajustada, demonstrada pelo histograma, são indícios de que, apesar do número pequeno de amostras, o conjunto de dados é representativo da população de pistas com características semelhantes.

Sendo assim, como demonstra o intervalo de confiança calculado, pode-se afirmar, com 95% de certeza, que reproduzir esse experimento em condições semelhantes permite identificar, aproximadamente, 58% a 74% de pistas de pouso não pavimentadas em ambiente amazônico.

Para tanto, como foi verificado nas discus-

sões acerca do tamanho e da forma das regiões, é necessário que os valores da Similaridade e da Área, assim como a posição das sementes, tenham sido corretamente ajustados para os alvos de interesse.

Essa tarefa pode se tornar um desafio quando se trata de alvos desconhecidos e localizados nos vazios demográficos da Região Amazônica. Logo, investigar o comportamento desses parâmetros para alvos semelhantes, mas

previamente conhecidos, demonstra-se uma problemática relevante para trabalhos futuros.

Uma forma de mitigar o impacto desses parâmetros sobre a identificação das pistas de pouso, como se demonstrou no exemplo do alvo S04, é calcular o índice de circularidade do alvo. Portanto, investigar a circularidade das pistas de pouso também pode ser uma linha de pesquisa a ser perseguida por futuras inquirições acadêmicas.

REFERÊNCIAS

- ALMEIDA, L. C. Análise espacial de dados com o QUANTUM GIS: exercícios realizados durante tópico especial ofertado pelo programa de Pós-Graduação em Geografia da UFSC. **Observatorium: Revista eletrônica de Geografia**, v.3, n.8, p.173-194, 15dez.2011.
- ALVES, S. A. S. *et al.* **Abordagem metodológica baseada em imagens do SAR-R99B para identificar prováveis pistas de pouso não-homologadas na Amazônia.** Acta Amaz., Manaus, v. 39, n. 3, p. 727-729, set. 2009. DOI <https://doi.org/10.1590/S0044-59672009000300031>. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0044-59672009000300031&lng=pt&nrm=iso. Acesso em: 19.fev.2020.
- BASTOS, F. I. P. M. *et al.* (Org.). **III Levantamento Nacional sobre o uso de drogas pela população brasileira.** Rio de Janeiro: FIOCRUZ/ICICT, 2017. 528 p. Disponível em: <https://www.arca.fiocruz.br/handle/icict/34614>. Acesso em: 19.fev.2020
- CAMARA, G; SOUZA, RCM; FREITAS, Um, GARRIDO, J. SPRING: Integrating remote sensing and GIS by object-oriented data modeling **COMPUTERS & GRAPHICS**,, p. 395-403, 1996.
- CARDOSO, C. A. *et al.* Caracterização morfométrica da bacia hidrográfica do rio Debossan. **Revista árvore**, Nova Friburgo, RJ, v. 30, n. 2, p. 241-248, 2006. Disponível em: <https://www.scielo.br/pdf/rarv/v30n2/a11v30n2.pdf>. Acesso em: 13 set.2020.
- CASELLA, G; BERGER, R. L. **Statistical inference.** Pacific Grove, CA, Duxbury, v.2, p. 337-472, 2002.
- CHABRIER, Sébastien et al. A comparative study of supervised evaluation criteria for image segmentation. In: **2004 12th European Signal Processing Conference.** IEEE, 2004. p. 1143-1146. Disponível em: <https://ieeexplore.ieee.org/abstract/document/7079854> Acesso em: set. 2020.
- CLAYTON, S.; NEVES, M..Estudo e implementação de técnicas de segmentação de imagens. **Revista**

Virtual de Iniciação Acadêmica da UFPA, v. 1, n. 2, jul. 2001. Disponível em:

http://www.academia.edu/download/30184437/ed_02_scmn.pdf. Acesso em: 09.mar.2020.

COUTO, A. C.; OLIVEIRA, I. de O. A geografia do narcotráfico na Amazônia. **Revista Geographia Opportuno Tempore**, Londrina, v. 3, n. 1, p. 52 - 64, 2017. Disponível em:

<http://www.uel.br/revistas/uel/index.php/Geographia/article/view/31774>. Acesso em: 19.mar.2020.

EHLERS, M.; JANOWSKY, R.; GÄLHER, M. **New Remote Sensing Concepts for Environmental Monitoring**. Remote Sensing for Environmental Monitoring, GIS Applications, and Geology, v. 4545, 2002.

ESPÍNDOLA, G. M. D. **Ajuste de parâmetros em algoritmos de segmentação de imagens por crescimento de regiões**. 2006. Dissertação (Mestrado) - Curso de Pós-Graduação em Sensoriamento Remoto, Instituto Nacional de Pesquisas Espaciais. São José dos Campos, SP, 2006. Disponível em: <http://mtc-m16b.sid.inpe.br/col/sid.inpe.br/MTC-m13@80/2006/12.21.11.14/doc/publicacao.pdf>. Acesso em: 09mar.2020.

HARALICK, R. M. e SHAPIRO, L. G. (1985). Image segmentation techniques. **Computer Vision, Graphics, and Image Processing**, v. 29, Issue 1, January 1985, p. 100-132. DOI

10.1016/s0734-189x(85)90153-7. Disponível em:

<https://www.sciencedirect.com/science/article/abs/pii/S0734189X85901537>. Acesso em: 19.mar.2020.

GONZALEZ, R. C. e WOODS R. C. **Processamento digital de imagens**. 3ª ed. São Paulo: Pearson Prentice Hall, 2010.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). **Câmeras imageadoras CBERS-3 e 4**. 2019. Disponível em: <http://www.cbears.inpe.br/sobre/cameras/cbers3-4.php>. Acesso em: 23ago.2020.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS (INPE). **Sobre o CBERS**. 2018. Disponível em: <http://www.cbears.inpe.br/sobre/index.php>. Acesso em: 23ago.2020.

ITO, M. H.; FONSECA FILHO, H.; CONTI, L. A.. Uso do software livre QGIS (Quantum GIS) para ensino de Geoprocessamento em nível superior. **Revista Cartográfica**, v. 94, p. 127-148, 2017. Disponível em: <https://repositorio.usp.br/item/002881150>. Acesso em: 24mar.2020.

LACERDA, M. G. **Abordagem GEOBIA para imagens VHR Obtidas por Aeronaves Remotamente Pilotadas e Sensores Satelitais com o Uso de Classificadores Individuais e Ensemble**. 2020. 195f. Dissertação de Mestrado – Instituto Tecnológico da Aeronáutica, São José dos Campos.

LEE, S. U.; CHUNG, S. Y.; PARK, R. H. (1990). A comparative performance study of several global thresholding techniques for segmentation. **Computer Vision, Graphics, and Image Processing**, Volume 52, Issue 2, November 1990, p. 171-190. DOI 10.1016/0734-189x(90)90053-x. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/0734189X9090053X>. Acesso em 23 mar. 2020.

NIWATTANAKUL, S. et al. **Using of Jaccard coefficient for keywords similarity**. In: Proceedings of the international multiconference of engineers and computer scientists. 2013. p. 380-384. Disponível em: https://www.researchgate.net/publication/317248581_Using_of_Jaccard_Coefficient_for_Keywords_Similarity. Acesso em: 02set.2020.

PEDRINI, H., SCHWARTZ, W. **Análise de Imagens Digitais: princípios, algoritmos e aplicações**. São Paulo: Thomson Learning, 2008.

RADOUX, Julien et al. **Improving ecotope segmentation by combining topographic and spectral data**. Remote Sensing, v. 11, n. 3, p. 354, 2019. Disponível em:

<https://www.mdpi.com/2072-4292/11/3/354> Acesso em: 09 mar. 2020.

SALDANHA, M. F. S.; FREITAS, C. C. Segmentação de Imagens Digitais: uma revisão. In: Workshop Dos Cursos De Computação Aplicada Do Inpe, 9. (WORCAP), 2009, São José dos Campos. **Anais...** São José dos Campos: INPE, 2009. On-line. IBI: <8JMKD3MGP8W/37NC86S>. Disponível em: <<http://urlib.net/rep/8JMKD3MGP8W/37NC86S>>. Acesso em: 09 mar. 2020.

SANTOS, A.R. dos; PELUZIO, T.M. de O.; SAITO, N. S. **Spring 5.1. 2**: passo a passo: aplicações práticas. Alegre, ES: CAUFES, 2010.

SILVA, E. Veículos aéreos não tripulados: panorama atual e perspectivas para o monitoramento de atividades ilícitas na Amazônia. In: Simpósio Brasileiro de Sensoriamento Remoto (SBSR), 16, 2013, Foz do Iguaçu. **Anais...** São José dos Campos: INPE, 2013. Artigos, p. 9324-9331. CD-ROM, On-line. Disponível em: <<http://www.dsr.inpe.br/sbsr2013/files/p1457.pdf>>. Acesso em 18 fev. 2020.

XAUD, Haron Abraham Magalhães et al. Análise de imagens orientada a objetos para integração de imagens SAR e ópticas no mapeamento de áreas periurbanas. In: **Embrapa Roraima-Artigo em anais de congresso (ALICE)**. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 15., 2011, Curitiba. Anais... Curitiba, PR: SBSR, 2011. Disponível em: <https://www.alice.cnptia.embrapa.br/bitstream/doc/989156/1/Trabalho7Xaudetal2011.pdf>. Acesso em: set 2020.

ZHANG, Y. J.; GERBRANS, J. J. Objective and quantitative segmentation evaluation and comparison. **SignalProcessing**, Volume 39, Issues 1–2, September 1994, Pages 43-54 DOI10.1016/0165-1684(94)90122-8. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/0165168494901228>. Acesso em 19.mar.2020.

ZHANG, Y. J. A survey on evaluation methods for image segmentation. **Pattern Recognition**, v. 29, Issue 8, August 1996, Pages 1335-1346. DOI:10.1016/0031-3203(95)00169-7 Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/0031320395001697>. Acesso em: 19mar.2020.

ZHANG, Y. J.. Evaluation and comparison of different segmentation algorithms. **Pattern Recognition Letters**, Volume 18, Issue 10, October 1997, Pages 963-974. DOI 10.1016/s0167-8655(97)00083-4. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0167865597000834>. Acesso em: 19mar.2020.

APENDICE A – TABELA DE ALVOS

ALVO	LATITUDE	LONGITUDE	IMAGEM
S01	-5.8074211775	-56.6575279398	CBERS_4_PAN5M_20190812_169_107_L4_BAND1
S02	-5.7808461514	-56.6106648408	CBERS_4_PAN5M_20190812_169_107_L4_BAND1
S03	-5.8375973489	-56.5952375122	CBERS_4_PAN5M_20190812_169_107_L4_BAND1
S04	-5.9888633338	-56.6333499487	CBERS_4_PAN5M_20190812_169_107_L4_BAND1
S05	-5.7682181001	-56.6787298805	CBERS_4_PAN5M_20190812_169_107_L4_BAND1
S06	-5.7951651084	-56.8503901261	CBERS_4_PAN5M_20190812_169_107_L4_BAND1
S07	-7.4763594109	-56.8907451903	CBERS_4_PAN5M_20200715_169_109_L4_BAND1
S08	-7.5337888071	-56.6618525224	CBERS_4_PAN5M_20200715_169_109_L4_BAND1
S09	-5.7943395062	-57.0517222893	CBERS_4_PAN5M_20190812_169_107_L4_BAND1
S10	-5.7174042813	-56.4353497392	CBERS_4_PAN5M_20200715_169_107_L4_BAND1
S11	-5.6914784226	-56.4908846346	CBERS_4_PAN5M_20200715_169_107_L4_BAND1

S12	-6.0841283330	-56.2920670217	CBERS_4_PAN5M_20200715_169_107_L4_BAND1
S13	-7.5120846608	-56.5341594925	CBERS_4_PAN5M_20200715_169_109_L4_BAND1
S14	-5.7452270052	-56.3886204868	CBERS_4_PAN5M_20200715_169_107_L4_BAND1
S15	-7.8997921982	-56.6189313138	CBERS_4_PAN5M_20200715_169_109_L4_BAND1
T01	-1.0238440870	-63.7968538090	CBERS_4_PAN5M_20150405_178_101_L4_BAND1
T02	-0.7718000000	-60.0736000000	CBERS_4_PAN5M_20180916_174_101_L2_BAND1
T03	-6.5344444440	-49.6450000000	CBERS_4_PAN5M_20200710_162_107_L4_BAND1