

PERFORMANCE DE CLASSIFICADORES DE SIG EM IMAGENS LANDSAT-5 DA BACIA HIDROGRÁFICA DO RIO SÃO FRANCISCO VERDADEIRO – OESTE DO PR, EM 1985

Ana Caroline Royer¹, Gabriela Moro Berté², Vanderlei Leopold Magalhães³

¹Universidade Tecnológica Federal do Paraná, Rua do Estado de São Paulo, n.16, Bragança-Portugal, anaroyer@alunos.utfpr.edu.br; ²Universidade Tecnológica Federal do Paraná, Avenida Brasil, n.4232, Medianeira, Paraná, Brasil, gabrielamoroberte@gmail.com e ³Universidade Tecnológica Federal do Paraná, Avenida Brasil, n.4232, Medianeira, Paraná, Brasil, vlmagalhaes@utfpr.edu.br

RESUMO

Com o objetivo de avaliar o desempenho de diferentes classificadores inseridos nos Sistemas de Informações Geográficas (SIG), realizou-se, no SPRING 5.2.6, o processamento e a classificação do tipo supervisionada (Maxver e Bhattacharya) e não supervisionada (Isoseg) de imagens do satélite Landsat-5 da bacia do rio São Francisco Verdadeiro – oeste do PR, datadas de 1985. O classificador Maxver apresentou erro de inclusão e de omissão. Verificou-se conflito entre as classes de solo exposto, área agrícola e vegetação, entretanto a área de ambas é equivalente. Bhattacharya apresentou um resultado satisfatório, contudo o melhor desempenho foi obtido ao empregar a classificação não supervisionada, pois Isoseg foi o único a identificar toda a área da bacia. De acordo com a finalidade do estudo, é possível definir o desempenho requerido dos classificadores e, caso o emprego deles seja insuficiente, a ferramenta de edição matricial é oportuna.

Palavras-chave — Classificação supervisionada, Classificação não supervisionada, Uso e cobertura da terra.

ABSTRACT

In order to evaluate the performance of different classifiers inserted in the Geographic Information Systems (GIS), the processing and classification of the supervised (Maxver and Bhattacharya) and unsupervised type (Isoseg) of images were carried out in SPRING 5.2.6 of the Landsat-5 satellite of the São Francisco River basin True-west of PR, dating from 1985. The Maxver classifier presented an inclusion and omission error. There was conflict between the classes of exposed soil, agricultural area and vegetation, however the area of both is equivalent. Bhattacharya presented a satisfactory result, but the best performance was obtained by using the unsupervised classification, because Isoseg was the only one to identify the entire area of the basin. Depending on the purpose of the study, it is possible to define the required performance of the classifiers and, if their use is insufficient, the matrix editing tool is timely.

Key words — Supervised classification, Unsupervised classification, Use and land cover.

1. INTRODUÇÃO

O mapeamento temático é uma das diversas aplicações do sensoriamento remoto. O benefício de dados, ferramentas e geotecnologias disponibilizados nos dias atuais permite a elaboração de mapeamentos voltados à análise multitemporal de uma determinada região, identificando as mudanças ocorridas no uso e cobertura do terra, avaliando os avanços positivos e negativos no âmbito ambiental. Produtos de satélites com elevada resolução temporal tornam possível o monitoramento dos fenômenos climáticos, além dos desmatamentos e queimadas, auxiliando na gestão e planejamento de ações futuras.

As técnicas de processamento digital de imagens (PDI) consistem em operações ou transformações numéricas aplicadas nas imagens, e podem ser reunidas em três conjuntos: técnicas de pré-processamento, de realce e classificação de imagens [1]. Utilizando os algoritmos classificadores disponíveis no software livre SPRING, o enfoque deste trabalho foi a classificação de imagens, analisando o desempenho dos classificadores supervisionados (Maxver e Bhattacharya) e não supervisionados (Isoseg).

Deseja-se, assim, avaliar a viabilidade dos métodos na classificação para elaboração de mapas temáticos, neste trabalho com imagens Landsat-5 TM (*Thematic Mapper*) da bacia do Rio São Francisco Verdadeiro no ano de 1985, fase de início do imageamento deste satélite, que foi lançado em 1984. O propósito é a obtenção de dados para comparações e análises em estudos posteriores e verificar a eficiência das técnicas de geoprocessamento disponíveis para a análise de bacias hidrográficas, uma vez que seu estudo é fundamental para a gestão e planejamento integrado dos recursos naturais.

2. MATERIAIS E MÉTODOS

2.1 Área de estudo

Com uma área estimada em 2.300 km², a Bacia do Rio São Francisco Verdadeiro – BRSFV, abrange onze municípios da região Oeste do Paraná e está situada no Terceiro Planalto Paranaense [2]. Sua nascente está localizada na área urbana de Cascavel e as águas do Rio São Francisco e seus tributários tem como destino o lago de Itaipu, conforme Figura 1 [3].

A cobertura pedológica é constituída predominantemente por Latossolos, Nitossolos e nas áreas mais dissecadas, Neossolos [4]. De acordo com a classificação de Köppen, a região apresenta clima do tipo Cfa, com precipitação anual oscilando entre 1600 a 2000 mm [5].



Figura 1. Localização da Bacia Hidrográfica do Rio São Francisco Verdadeiro. Fonte: CIH [3]

2.2 Materias

As imagens foram capturadas pelo sensor *Thematic Mapper* do satélite Landsat-5. Para recobrir a área de estudo foram necessárias duas imagens, datadas em 15/08/1985 e 22/08/1985, órbita 223-ponto 077 e órbita 224-ponto 077, respectivamente. Optou-se por realizar o download das imagens georreferenciadas, disponíveis no banco de dados do Departamento de Pesquisa Geológica dos Estados Unidos (USGS/USA) no endereço <http://earthexplorer.usgs.gov>. As imagens possuem resolução espacial e temporal de 30 m e 16 dias, respectivamente. A proximidade das datas contribuiu para o ajuste dos níveis de cinza no mosaico.

2.3 Métodos

No SIG SPRING 5.2.6 criou-se o banco de dados BD_BACIA_RSJV e o projeto BACIA_RSJV, com projeção LatLong e datum Sirgas 2000. As cenas foram mosaicadas e recortadas de acordo com o limite da bacia, definido pelo Grupo Integrado de Monitoramento e Análise de Bacias Hidrográficas (GIMAB – UTFPR Medianeira). Na sequência, uma composição colorida e contraste foram gerados, associando cada banda a um canal: Vermelho, B; Infravermelho próximo, G, Infravermelho médio, R, obtendo o realce desejado para efeitos comparativos e de identificação dos alvos nas posteriores classificações.

2.4 Classificadores

Segundo INPE (2008), “a classificação é o processo de extração de informação em imagens para reconhecer padrões e objetos homogêneos” [6]. A classificação de imagens faz uso de algoritmos e pode ser realizada de maneira supervisionada – quando as classes são definidas *a priori* pelo

analista, no intuito de simular a ação de um foto intérprete – ou não supervisionada, onde os dados se agrupam por meio do fenômeno de *clustering* e as classes são definidas *a posteriori*, a partir do resultado da análise [1].

A classificação pixel a pixel, tem como fundamento a análise dos atributos espectrais (níveis de cinza) de cada pixel para encontrar regiões homogêneas, identificando a probabilidade de um pixel pertencer a uma classe predeterminada. Os classificadores por região, por sua vez, utilizam-se da informação espectral do pixel bem como a de seus vizinhos para que o sistema reconheça áreas homogêneas dentro de uma imagem [8] e [9].

Maxver é um algoritmo supervisionado, definido como classificador por máxima verossimilhança, onde cada pixel será destinado a classe que tem a mais alta probabilidade de pertencer, isto é, a máxima verossimilhança. Maxver considera a ponderação das distâncias entre as médias dos níveis digitais das classes, utilizando parâmetros estatísticos e assumindo que todas as bandas têm distribuição normal [8]. Meneses (2012) enfatiza que o algoritmo necessita ser treinado para poder distinguir as classes uma das outras. Na etapa de treinamento foram selecionadas pequenas áreas que fossem fiéis representativas da classe coletada em questão. É nesse momento que surge a dificuldade em obter amostras de pixels puros que representem uma determinada classe espectral. São vários os fatores que mascaram e interferem na resposta espectral de um objeto contido numa área, impossibilitando que a homogeneidade se estabeleça rigorosamente, além de um mesmo tipo de alvo poder apresentar diferentes comportamentos espectrais em função de diversas variáveis [7] e [8].

Para minimizar as características espectrais similares, sobreposição de classes, é possível variar o rigor da classificação, modificando o limiar de aceitação que indica a porcentagem de pixels de uma classe que será classificada como pertencente a esta classe. De acordo com informações do INPE (2008) um limiar de 99%, engloba 99% dos pixels, sendo que 1% de todos os pixels da classe (os de menor probabilidade) será ignorado. Adotando um limiar de 100%, todos os pixels serão classificados, não havendo, portanto, rejeição. O limiar de aceitação adotado foi de 95% [6].

O classificador por região Iseog é exclusivo do software livre SPRING e se baseia em técnicas de *clustering* a partir da distância de Mahalanobis, calculada entre a classe e as regiões disponíveis, sendo essa a medida de similaridade utilizada entre elas [9]. Os principais procedimentos consistem em realizar a segmentação da imagem e em seguida definir um limiar de aceitação, dado em porcentagem, que elucida a distância máxima de Mahalanobis de acordo com a qual as regiões podem estar afastadas do centro da classe para serem dadas como pertencentes a esta classe. O limiar de aceitação adotado foi de 95%, e os limiares de similaridade e área que melhor se adaptaram em separar os objetos da cena, após vários testes, foram de 5 e 30, respectivamente. Cabe ao analista associar os *clusters* identificados pelo algoritmo às devidas classes

temáticas. A pós-classificação exige um conhecimento prévio, maior tempo e esforço do pesquisador [9].

O classificador por região Bhattacharya, por sua vez, adota como princípio base a distância de Bhattacharya, medindo a separabilidade estatística entre os pares de classes espectrais, calculada através da distância média entre as distribuições de probabilidades das classes. A classificação não é totalmente automática, sendo necessário realizar a etapa de treinamento, fato que o define como classificador supervisionado [9]. As regiões obtidas na fase da segmentação, assim como polígonos representativos da região podem ser utilizados como amostras. Com o intuito de comparação, os valores dos limiares de área e similaridade foram os mesmos adotados para o classificador Isoseg.

3. RESULTADOS E DISCUSSÃO

A definição de limiares e parâmetros para cada classificador requereu uma série de testes, uma vez que não existem valores recomendados pela literatura, pois estes variam de acordo com as características do local e finalidade do estudo. A etapa de segmentação foi decisiva à medida que seccionava a bacia em polígonos menores, o que diminuía a confusão de classes na posterior amostragem e classificação.

Durante a etapa subsequente de treinamento, necessárias para os classificadores supervisionados Maxver e Bhattacharya, houve preocupação em garantir que tal procedimento se apresentasse com confiabilidade aceitável, fixando um valor médio de 30 regiões amostradas para cada classe temática. Ao extrair os atributos das regiões, o classificador não supervisionado Isoseg gerou 29 temas.

O período analisado em questão (1985) dificulta a possibilidade de comparar as áreas catalogadas com um mapa de referência que represente o uso e cobertura do solo da unidade territorial analisada da mesma época. Contudo, é possível estimar de forma satisfatória a área englobada pela maioria das classes temáticas, a partir de comparação com a composição colorida e realçada. O resultado da classificação para os diferentes classificadores utilizados na BRSFV está disposto na Tabela 1.

Classificadores	Bhattacharya	Isoseg	Maxver
Água	38,76	28,85	23,05
Vegetação	416,1	696,54	225,57
Área agrícola	1066,44	672,59	636,67
Solo exposto	841,26	1008,32	1045,33
Área urbana	32,1	29,78	27,47
Afloramento rochoso	22,35	20,37	414,94
Total	2417,01	2456,45	2373,03

Tabela 1. Área em km² das classes temáticas de uso e cobertura da terra.

Apenas o classificador Isoseg identificou toda a área da bacia, que por sua vez pode diferir da área estimada por

outros estudos de acordo com a delimitação e escala utilizada. É interessante destacar que nenhum dos classificadores foi capaz de reconhecer a área urbana, visto que a mesma se confunde com o solo exposto. Para esta classe realizou-se a edição matricial. A resposta espectral das classes área agrícola e solo exposto são similares, uma vez que o solo fica descoberto durante alguns períodos de seu manejo. Na década de 1980, a técnica de cultivo por meio do qual arava-se o solo era habitual e dominante. Na Figura 2, a diferença entre área agrícola cultivada e vegetação é compreensível. Os três classificadores demonstraram incompatibilidade ao categorizarem as duas classes temáticas. Deste modo, a diferença entre as áreas classificadas como vegetação, área agrícola e solo exposto deixam de ser significativas.

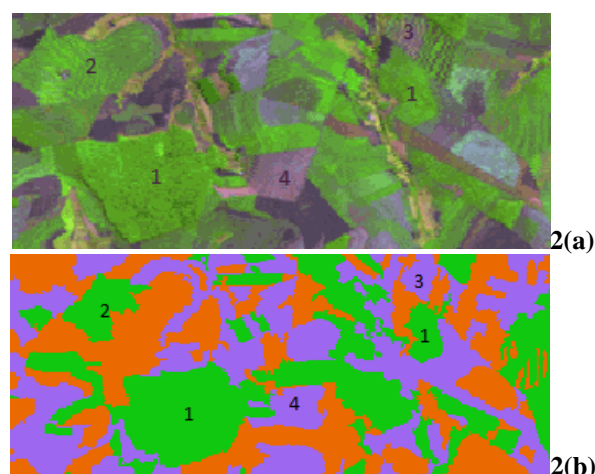


Figura 2. (a) Composição colorida (3B4G5R) onde o número 1 corresponde a vegetação; 2 representa área agrícola; 3 pode ser interpretado como área agrícola e solo exposto e; 4 assemelha-se a solo exposto. (b) Imagem classificada Isoseg; a posição dos números é equivalente à (a).

O classificador Maxver (Figura 3) superestimou a área de afloramento rochoso, apesar de várias tentativas de amostragens. Trata-se de erro de inclusão, ao mesmo tempo em que para a vegetação observou-se erro de omissão. Desta forma, constatou-se que o melhor desempenho no mapeamento das cenas foi obtido ao empregar o classificador Isoseg (Figura 4), uma vez que ele se sobressai em relação ao Bhattacharya (Figura 5), visto que este último deixou de rotular algumas regiões.

A qualidade de resposta espectral da cena é influenciada pela época do ano, clima (irradiação solar, presença de nuvens), dentre outros fatores que podem até comprometer sua utilização para determinados estudos. As resoluções radiométrica e espacial são, também, interferentes do nível de detalhamento e qualidade do resultado. Para o satélite Landsat-5, a resolução radiométrica e espacial equivalem à 256 níveis de cinza (8 bits) e 30m, respectivamente, cujos números atualmente já são superados por outros satélites.

Uma constatação interessante desta época da bacia é a inexistência de mata ciliar ao longo dos cursos hídricos

principais. Isso é explicado pela consolidação do reservatório de Itaipu, ocorrido 3 anos antes da captura da cena, período neste estudo de caso insuficiente para o reestabelecimento da vegetação ripária.

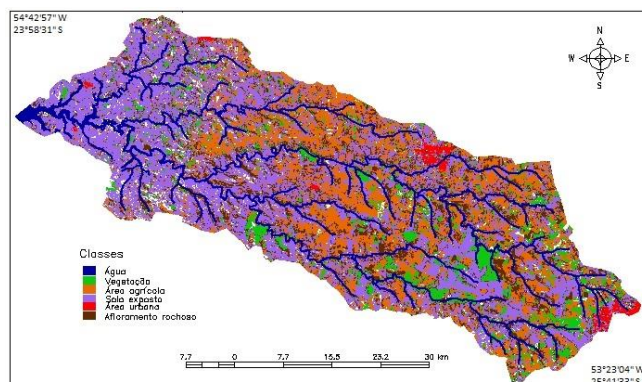


Figura 3. Mapa de uso e ocupação da terra de acordo com o classificador Maxver

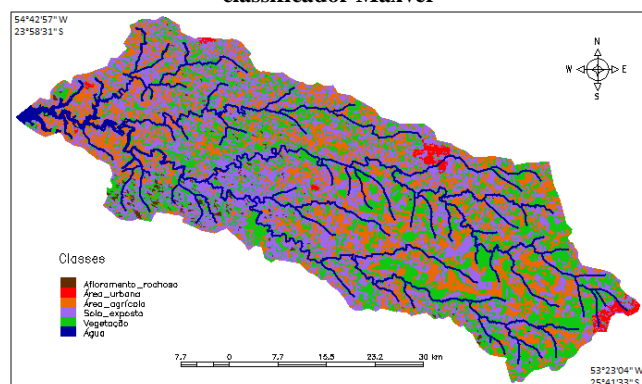


Figura 4. Mapa de uso e ocupação da terra elaborado com o classificador Iseog

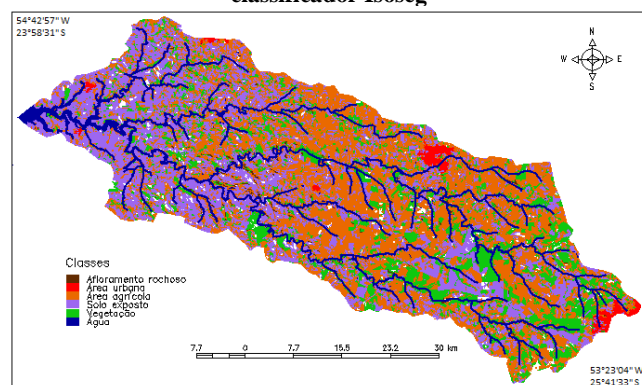


Figura 5. Mapa de uso e ocupação da terra de acordo com o classificador Bhattacharya

4. CONCLUSÕES

Ao final dos procedimentos e análises é possível inferir que, de fato, os elementos que interferem na qualidade da imagem no momento da captura são essenciais para uma interpretação e elaboração de mapeamentos confiáveis. Os classificadores que demonstraram melhor desempenho foram aqueles que

fazem uso da imagem segmentada, Bhattacharya e Iseog, sendo que este último destacou-se por ser o único que classificou toda a área da bacia. A dificuldade ao estabelecer valores aos parâmetros e atributos durante a segmentação foi constatada, ao mesmo tempo que a realização de testes culminou em dígitos adaptáveis à realidade da cena.

A classificação tendo em vista o mapeamento de uso e cobertura da terra pode ser bem produzida ao empregar os SIGs. Uma vez estabelecidos os objetivos do estudo, é possível definir o desempenho exigido dos classificadores e, caso o emprego deles seja insuficiente para atingir a finalidade, a edição matricial para remapear as regiões classificadas equivocadamente, realizada de forma manual, é uma ferramenta oportuna e confiável.

5. REFERÊNCIAS

- [1] Florenzano, T.G., “Geomorfologia: conceitos e tecnologias atuais”, Ed. Oficina de Textos, São Paulo, 318 p., 2008.
- [2] Mineropar, “Atlas geomorfológico do Estado do Paraná”. Escala base 1:250 000, modelos reduzidos 1:500 00/ Minerais do Paraná. Universidade Federal do Paraná, Curitiba, 63p., 2006.
- [3] Centro Internacional de Informática – CIH, “Programa Hidrológico Internacional”, Bacia Hidrográfica do Rio São Francisco Verdadeiro, disponível em <<http://saofrancisco.hidroinformatica.org/br/basin.html>>, Acesso em 21 fev. 2017.
- [4] Da Rocha, A., et al., “Mapeamento da Fragilidade Potencial e Emergente na Bacia Hidrográfica do Paraná 3”, Revista Estudos Geográficos, v.14, n.1, 2016. Disponível em <<http://www.periodicos.rc.biblioteca.unesp.br/index.php/estgeo/article/view/12025/8047>> Acesso em 30 mar. 2017.
- [5] Instituto Agrônomo do Paraná (IAPAR), “Cartas Climáticas do Paraná”. Londrina, 2000.
- [6] Instituto Nacional de Pesquisas Espaciais (INPE), “Manuais: tutorial de geoprocessamento SPRING”, 2008. Disponível em: <<http://www.dpi.inpe.br/spring/portugues/tutorial/classific.html>>. Acesso em 20 fev. 2017.
- [7] Moreira, M.A., “Fundamentos do sensoriamento remoto e metodologias de aplicação”, 4. ed. atual. e ampl. – Viçosa, MG: Ed. UFV, 2011.
- [8] Meneses, P.R., “Classificação Pixel a Pixel de imagens”. In: Paulo R. Meneses; Tati de Almeida. (Org.). Introdução ao processamento de Imagens de Sensoriamento Remoto. Brasília: UnB/CNPq, 2012, v.1, p.191-206. Disponível em: <<http://cnpq.br/documents/10157/56b578c4-0fd5-4b9f-b82a-e9693e4f69d8>> Acesso em 12 fev. 2017.
- [9] Brites, R.S; Bias, E.D; Rosa, A.N.C.S., “Classificação por Regiões”. In: Meneses, P.R; Almeida, T. (Orgs.). Introdução ao processamento de imagens de sensoriamento remoto. Brasília: UnB/CNPq, 2012, v.1, p.209-220. Disponível em: <<http://cnpq.br/documents/10157/56b578c4-0fd5-4b9f-b82a-e9693e4f69d8>> Acesso em 12 fev. 2017.