

# APLICAÇÃO DE MÉTODOS DE INTELIGÊNCIA ARTIFICIAL EM IMAGENS DE SATÉLITES PARA A ELABORAÇÃO DE PLANOS MUNICIPAIS DE ARBORIZAÇÃO URBANA

Gabrielli Beatriz dos Santos<sup>1</sup>

gabrielli.santos01@fatec.sp.gov.br

CST em Ciência de Dados – Faculdade de Tecnologia de Adamantina – FATEC

Paulo R. S. Ruiz

paulo.ruiz2@fatec.sp.gov.br

CST em Ciência de Dados – Faculdade de Tecnologia de Adamantina – FATEC

## 1. Introdução

Nos ambientes urbanos são encontrados diversos desafios no planejamento para atender às demandas crescentes de qualidade de vida e distribuição de serviços, devido ao rápido crescimento de sua população, a qual superou o índice rural globalmente [1]. A arborização urbana auxilia na melhoria das condições ambientais da população equilibrando o ambiente natural e urbano [2].

O objetivo deste trabalho é analisar a acurácia da classificação supervisionada a partir de imagens do satélite CBERS 4A, identificando suas potencialidades e limitações para auxílios a planos municipais de arborização urbana.

## 2. Metodologia

A área de estudo localiza-se na cidade de Adamantina – SP. Essa área foi escolhida por possuir uma grande diversidade de área verde, com a presença do Parque Caldeira, compreendendo uma parte da região noroeste da cidade. O trabalho foi realizado a partir de uma imagem orbital do satélite CBERS-4A, obtida em setembro de 2023, com resolução espacial de até 2 metros em 4 bandas espectrais.

O pré-processamento consiste na conversão dos dados de nível de cinza para radiância, sendo uma transformação física, logo em seguida é aplicada a correção atmosférica, que considera o quanto a contaminação da atmosfera, com gases e partículas, afetou os dados captados pelos sensores.

Foi adotada a classificação por regiões, onde a

imagem foi segmentada por meio do Algoritmo Multirresolução (Multiresolution Segmentation), disponível no software eCognition Developer 8.7. A imagem foi ajustada para falsa cor, onde os canais RGB receberam respectivamente as bandas infravermelha próxima, vermelha e verde, a vegetação se destaca em cor vermelha [3].

Foram selecionados 714 segmentos para serem utilizados como amostras de treinamento para as classes, sendo elas: Árvore (147), Grama (295) e Outro (272). Foram extraídos 15 atributos, sendo a média, desvio padrão, atributos customizáveis (ACs), brilho e dos índices de vegetação: Normalized Difference Vegetation Index (NDVI) e o Soil-Adjusted Vegetation Index (SAVI).

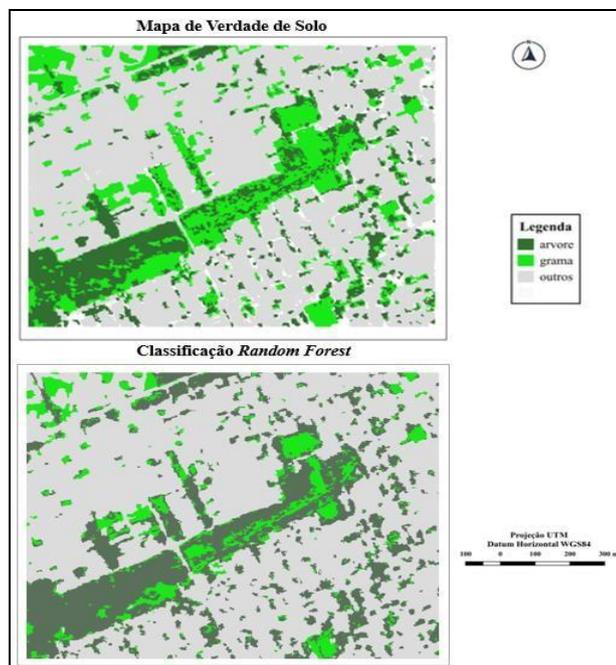
Após o recolhimento das amostras, a etapa de treinamento ocorre a partir do algoritmo Random Forest para a construção do modelo de classificação dos dados. Esse algoritmo é um método de aprendizado de máquina baseado em árvores de decisão que combina múltiplas árvores para melhorar a precisão da previsão e controlar o problema do overfitting, reduzindo a variância do modelo e tende a ser mais robusto e preciso do que uma única árvore de decisão [4].

A avaliação da qualidade das classificações foi realizada por meio de comparações com o mapa de verdade de solo desenvolvido, utilizando 2500 a 3000 pontos aleatórios da imagem, criando assim uma tabela associando cada ponto à classe e a matriz de confusão com as métricas de acurácia, sendo possível obter o cálculo da exatidão global, índice kappa, índice kappa condicional, erros de omissão e comissão de cada classe [5].

### 3. Resultados e Discussões

A Figura 1 apresenta o mapa de verdade de solo e o resultado da classificação utilizando o algoritmo Random Forest, elaborado empiricamente com o auxílio da imagem de satélite. Sendo possível verificar que há muitos erros nas classes *Árvore* e *Gramma*, visíveis confusões nessas classes.

Figura 1 – Mapa de verdade de solo e Classificação Random Forest.



Fonte: Autoria Própria (2024)

Para a validação dos resultados foram utilizados 2771 pontos distribuídos aleatoriamente nas imagens, a fim de verificar a correspondência entre as classes. O índice de exatidão global alcançado foi de 83,22%, com índice kappa de 0,68. A matriz de confusão apresenta erros significativos entre as classes *arvore* e *grama*.

A partir da análise da avaliação por classe (Tabela 1) é possível verificar que as classes *Árvore* e *Outro* apresentam as maiores exatidões do produtor, o modelo classificou corretamente as amostras, onde a classe *arvore* apresenta uma baixa exatidão do usuário, mostrando que a classificação não é confiável para essa classe.

Tabela 2 – Avaliação da qualidade por classe.

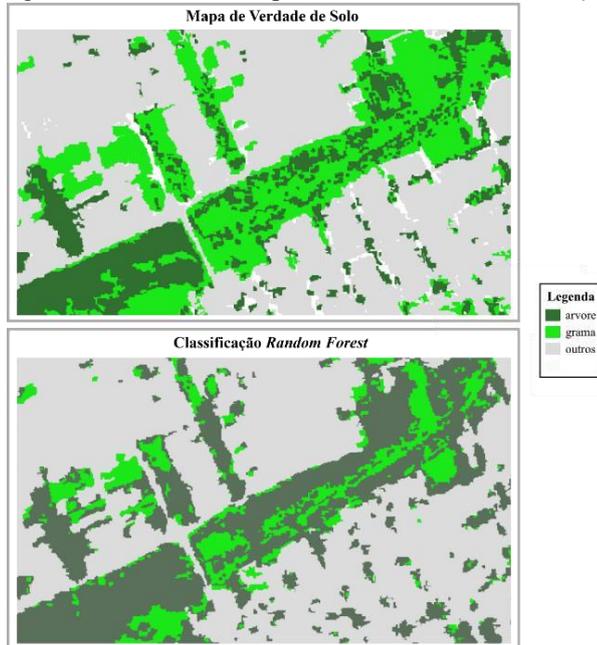
Avaliação	Árvore	Gramma	Outro
Exatidão do produtor da classe	0,91	0,40	0,95
Erro de omissão da classe	0,09	0,60	0,05
Exatidão do usuário da classe	0,53	0,92	0,94
Erro de comissão da classe	0,47	0,08	0,06
Índice <i>Kappa</i> usuário	0,42	0,89	0,75
Índice <i>Kappa</i> produtor	0,87	0,32	0,79

Fonte: Autoria Própria (2024)

A matriz de confusão ajuda a entender os índices condicionais por classe, que das 577 amostras só 229 foram realmente classificadas como *grama*. Quando analisamos a distribuição espacial da classificação no mapeamento (Figura 6), observa-se que a classe *árvore* tomou o lugar da classe *grama*, evidenciando-se uma grande confusão entre as classes.

As confusões verificadas são justificadas pelo comportamento espectral dos dois alvos [6]. Onde a vegetação apresenta maior reflectância na região do infravermelho e para ser possível a diferenciação das classes é preciso aumentar as bandas na região do espectro eletromagnético. Essa característica é limitada no sensor utilizado. Nesse sentido, são necessários maiores esforços para o refinamento de amostras e configurar os parâmetros dos algoritmos de Aprendizado de Máquina a serem utilizados. Além disso, para alcançar resultados com acurácias elevadas é necessário adotar múltiplos algoritmos, configurando em uma classificação híbrida.

Figura 6 – Detalhe do mapa de verdade e da classificação na região do Parque Caldeira.



Fonte: Autoria Própria (2024)

#### 4. Conclusões

O presente trabalho realizou uma classificação supervisionada por meio do algoritmo Random Forest de uma região da cidade de Adamantina utilizando dados do satélite CBERS 4A. Foram adotadas três classes, sendo: Árvore, Grama e Outros, a fim de avaliar o potencial dos dados para o mapeamento da cobertura arbórea.

O modelo de classificação desenvolvido foi capaz de classificar toda a imagem, obtendo índice de exatidão global de 83,22%, com índice Kappa de 0,68, sendo verificadas confusões entre as classes espectralmente semelhantes: Árvore e Grama. Mesmo com as confusões verificadas é possível ser adotado esse satélite em mapeamentos urbanos, mas são necessários esforços em aliar técnicas híbridas de classificação de imagens com a adoção consorciada de diferentes algoritmos de Aprendizado e Máquina a fim de alcançar acurácias elevadas. Dessa forma, a hipótese é respondida de maneira afirmativa, mas ressalvadas as limitações do sensor utilizado.

#### Referências

- [1] UNITED NATIONS POPULATION DIVISION (UNPD). World Urbanization Prospects: The 2018 Revision. Nova York, 2018.
- [2] LONDE, V.; MENDES, T. Arborização urbana: uma análise de sua importância na cidade de Teresina, Piauí. GeoTextos, Salvador, v. 10, n. 2, p. 350-370, 2014.
- [3] MYINT, S. W.; GOBER, P.; BRAZEL, A.; GROSSMAN-CLARKE, S.; WENG, Q. Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. Remote Sensing of Environment, Nova York, v. 115, n. 5, p. 1145-1161, 2011.
- [4] BREIMAN, L. Random forests. Machine Learning, v. 45, n. 1, p. 5-32, 2001
- [5] LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical Data. Biometrics, v. 33, n. 1, p.159-174, 1977. DOI 10.2307/2529310. Disponível em: <<http://www.jstor.org/stable/pdf/2529310.pdf?acceptTC=true>>. Acesso em: 05 mai. 2024.

- [6] PONZONI, F. J.; SHIMABUKURO, Y. E. Sensoriamento Remoto no Estudo da Vegetação. São José dos Campos: Parêntese, 2010, 127p.

### *Agradecimentos*

À Fatec de Adamantina pela disponibilização dos laboratórios. Ao CNPq pela concessão da bolsa de Iniciação Científica Tecnológica.

<sup>1</sup> bolsista de Iniciação Tecnológica do CNPq (ICT).