

# 11ª Jornada Científica e Tecnológica do IFSULDEMINAS & 8º Simpósio de Pós-Graduação

## USO DO SOFTWARE WEKA NA AVALIAÇÃO DE MÉTODOS DE CLASSIFICAÇÃO DE IMAGENS

**Luzia H. SILVA<sup>1</sup>; Raquel A. A. D. PEREIRA<sup>2</sup>; Camila S. A. LACERDA<sup>3</sup>**

### RESUMO

Devido a constante evolução dos sensores, pesquisadores se sentem motivados a desenvolverem novos métodos de classificação avançada, com o objetivo de facilitar os procedimentos para realização da mesma, o que possibilita substituir a visão humana pela automação da tarefa com uso de equipamentos para captura de imagens. Os métodos de classificação por regiões têm como princípio a segmentação da imagem, tais segmentos representam objetos que são classificados de forma independente dos pixels individuais dentro da cena. Os métodos de classificação, ainda buscam amenizar as dificuldades encontradas durante o processamento das imagens devido à grande dimensão de dados. No presente trabalho é apresentada uma comparação das técnicas de classificação orientadas a objetos disponíveis no *software* livre *Weka* desenvolvido por um grupo de pesquisadores da Universidade de Waikato, Nova Zelândia.

**Palavras-chave:** Segmentação; classificadores; amostras.

### 1. INTRODUÇÃO

Segundo Abhir (1991) segmentação é uma parte fundamental dos aspectos de "baixo nível" da visão computacional e tem muitas aplicações práticas, como em imagens médicas, industriais e imagens de satélite.

A segmentação consiste na primeira etapa para o processamento de imagens, pois a partir dela são gerados os objetos representativos, nos quais o processo de classificação será executado. Já a classificação consiste no processo de extração de informação de imagens para reconhecer objetos com características semelhantes. Como citado anteriormente, a classificação utilizada no presente trabalho é baseada em regras orientadas a objetos.

Segundo Martin (1999) uma aplicação de análise de imagem orientada a objeto é a fusão de dados de múltiplas fontes. Integrações de diferentes tipos de dados que desempenham um papel importante para o processo de classificação de imagens, que devido ao elevado grau de detalhes, vêm

---

<sup>1</sup>Graduada em Engenharia de Agrimensura e Cartográfica, IFSULDEMINAS – Campus Inconfidentes. E-mail: luziahsilva01@gmail.com.

<sup>2</sup>Graduada em Engenharia de Agrimensura e Cartográfica, IFSULDEMINAS – Campus Inconfidentes. E-mail: raquelaadp@gmail.com.

<sup>3</sup>Orientadora, IFSULDEMINAS – Campus Inconfidentes. E-mail: camila.lacerda@ifsuldeminas.edu.br.

sendo utilizado por diversas técnicas para as classificações. Segundo Lacerda (2018), por exemplo, é importante ressaltar a existência de métodos de classificação avançada, como a Árvores de Decisão (*Decision Tree*), Floresta Randômica (*Random Forest*), Máquinas Vetores de Suporte (*Support Vector Machine – SVM*), Redes Neurais Artificiais (*Artificial Neural Networks*), Redes Neurais Artificiais Profundas (*Deep Learning*), entre outros.

Neste trabalho foram utilizados três dos métodos de classificação, sendo eles Árvore de Decisão, Floresta Randômica e Máquinas Vetores de Suporte, que serão detalhados posteriormente.

## 2. MATERIAIS E MÉTODOS

Os materiais utilizados para realização do estudo foram as imagens fornecidas pelo Instituto Nacional de Pesquisas Espaciais (INPE) do satélite *Landsat 8*, da região de Inconfidentes/MG, caracterizado por conter uma resolução de 15 metros no pancromático e 30 metros da multiespectral. Ainda, foram utilizados os softwares comerciais, que apresentam recursos necessários para a análise de processamento de imagens, segmentando e classificando-as, bem como softwares para auxílio dos mesmos. Os *softwares* utilizados foram: *ENVI Classic 5.3*, *Ecognition Developer 64*, *Weka 3.8*, *Notepad++*, *LibreOffice calc* e o *QGIS 2.18.26*.

Inicialmente foi utilizado o *software ENVI* para o pré-processamento das imagens do satélite *Landsat 8*. A área de estudo está situada na região de Inconfidentes/MG (Figura 1).

Figura 1- Região de estudo, Inconfidentes/MG



Fonte: INPE

Para melhor resolução espacial efetuou-se então a fusão das imagens, processo pelo qual é realizada a união de diferentes características espectrais e espaciais. A imagem fusionada foi inserida no *software* eCognition e o processo de segmentação foi efetuado. O método escolhido foi o de algoritmo de segmentação de multiresolução, uma técnica de crescimento de regiões, em que, para obter resultados satisfatórios consistem em critérios de decisão onde é preciso alterar, por exemplo, parâmetros como o fator de escala, responsável por controlar a quantidade de variação espectral nos objetos. A imagem resultou em 844 segmentos, que por meio destes foram coletadas amostras para o

processo de classificação (Tabela 1). As amostras coletadas para área urbana e via foram menores, devido a proporção desta ser bem menor em relação a área total em estudo.

Tabela 1 - Número de Amostras coletadas para cada classe.

ID	Número de Amostras
Mata	44
Solo Exposto	35
Pasto	40
Via	8
Área Urbana	13

Fonte: os autores

A partir da segmentação e coleta de amostras foi efetuada a importação dos dados para o software *weka*, onde se realizou modificações trabalhadas nos programas *notepad++* e *LibreOffice calc*. Com as devidas transformações, foram feitos diferentes métodos de classificação. O primeiro método consiste na Árvore de Decisões, responsável por classificar o segmento a partir do seu ponto inicial, seguindo até que uma classe seja detectada. O segundo refere-se ao método da Floresta Randômica, que trabalha com uma grande coleção de árvores de decisão, já o terceiro consiste, na Máquinas Vetores de Suporte, utilizado para representar distâncias como pontos no espaço, mapeados de forma que cada classe seja dividida por um espaço claro que seja tão amplo quanto possível.

Os resultados das classificações decorreram na relação das porcentagens dos segmentos corretamente classificados, categorizando cada imagem conforme a veracidade à imagem original, e na importação das imagens classificadas para o software *QGIS*, as quais foram categorizadas por meio da composição de cores que permitiram o mapeamento dos dados das classes devidamente classificada, além de facilitar a visibilidade da imagem.

#### 4. RESULTADOS E DISCUSSÕES

Os resultados gerados estão sendo apresentados nas Figuras 2, 3 e 4, categorizadas no software *QGIS*. Estas representam as classificações pelos diferentes métodos citados anteriormente da imagem segmentada.

Figura 2-Classificação pelo método de árvore de decisão



Figura 3-Classificação pelo método de Floresta Randômica





Figura 4 - Classificação pelo método SVM / SMO

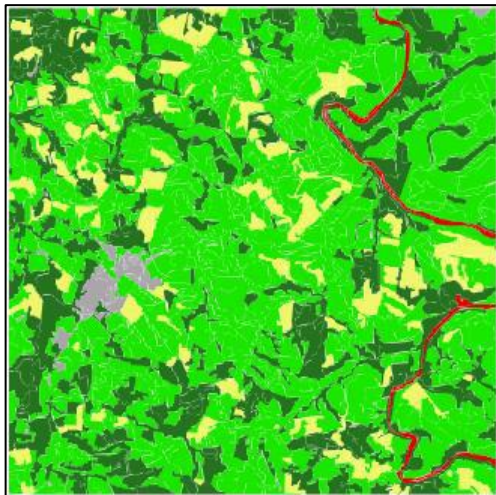
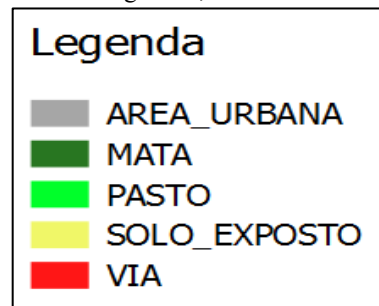


Figura 5 – Legenda das classificações apresentadas nas Figuras 2, 3 e 4.



Por meio da classificação realizada no *software Weka*, obteve-se também a relação das porcentagens dos segmentos corretamente classificados pelos métodos de classificação, que mediante análise, foi possível visualizar nas figuras que o método de Máquinas Vetores de Suporte - SVM (Figura 4) foi o mais acurado, classificando corretamente 93,57% dos segmentos, seguido pelo método da Floresta Randômica (Figura 3), acertando na classificação de 91,43% dos segmentos, e o método de Árvore de Decisões (Figura 2), com 87,85% dos segmentos corretamente classificados. A análise visual da eficácia dos métodos de classificação, pode ser realizada ao comparar as classificações com a área de estudo (Figura 1).

## 5. CONCLUSÕES

A classificação de imagens é de grande relevância para reconhecer objetos com características homogêneas. Pelos métodos avaliados, foi possível a identificação dos alvos de interesse da imagem, o que possibilitou a análise dos resultados que apresentaram maior eficácia e precisão pelos métodos de classificação, dentre esses o que resultou maior número de segmentos classificados foi o método de Máquinas Vetores de Suporte (SVM).

Para obter melhores resultados, sugere-se a coleta de uma quantidade maior de amostras no processo de coleta de dados e ainda, novos testes com diferentes classificadores.

## REFERÊNCIAS

B.SC., Abhir H. Bhalerao. **Multiresolution Image Segmentation**. 1991. Disponível em: <[https://warwick.ac.uk/fac/sci/dcs/people/abhir\\_bhalerao/publications/thesis.pdf](https://warwick.ac.uk/fac/sci/dcs/people/abhir_bhalerao/publications/thesis.pdf)>.

LACERDA, Camila Souza Dos Anjos. **Metodos avancados de classificao de imagem**. 20 oct. 2018, 14 dec. 2018. 53 p. Notas de Aula.

SCHÄPE, Martin Baatz Und Arno. **Multiresolution Segmentation: an optimization approach for high quality multi-scale image segmentation**. 1999. Disponível em: <[http://www.ecognition.com/sites/default/files/405\\_baatz\\_f\\_p\\_12.pdf](http://www.ecognition.com/sites/default/files/405_baatz_f_p_12.pdf)>.