

IDENTIFICAÇÃO DE CULTURAS CAFEEIRAS EM IMAGENS DE SATÉLITE UTILIZANDO ALGORITMO DE RECONHECIMENTO DE PADRÕES

Jaqueline R. AGUILAR¹; João M. RIBEIRO²; Walbert J. R. dos SANTOS³; Jovani LUIZ⁴

RESUMO

Este artigo tem como finalidade apresentar o resultado da aplicação de um método de Inteligência Artificial (k-NN), desenvolvido em Python (utilizando biblioteca “Sklearn”), utilizando Google Earth, o software QGIS 2.18.3 e Banco de Dados (MySQL), construído a partir de imagens de satélite da série “Landsat8”. Desta forma, foram gerados e utilizados arquivos para treinamento e aplicação do algoritmo de classificação de padrões, visando classificar novas áreas de cultura cafeeira, no município de Muzambinho (MG), utilizando imagens de satélite.

Palavras-chave: Sensoriamento Remoto; Landsat; Processamento Digital de Imagens; Inteligência Artificial.

1. INTRODUÇÃO

Estima-se que metade do processamento cerebral é destinada a processar informações fornecidas pela visão (AZEVEDO; CONCI, 2003). O complexo sistema de visão humana, vem sendo recriado com o uso de recursos computacionais (hardware, software e fundamentos teóricos) onde a captura de imagens, pode ser realizada por câmeras digitais, sensores infravermelhos ou de ultrassom, radares, satélites ou outros dispositivos.

Diferentes complexidades computacionais permitem analisar as características presentes nas imagens, descobrir padrões e tendências dos dados obtidos onde, após estarem organizados, sem ruídos, imperfeições, inconsistências, pode-se utilizá-los usando algoritmos computacionais para fazer, por exemplo, uma identificação de elementos por similaridade. O sensoriamento remoto se baseia na aquisição de informações de um objeto alvo sem que haja contato físico entre alvo e sensor através da medição da interação da radiação eletromagnética com o objeto (MENESSES e

¹ Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais – Campus Muzambinho. Muzambinho/MG - Email: jaguilarodrigues@gmail.com;

² Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais – Campus Muzambinho. Muzambinho/MG - E-mail: joao.ribeiro@muz.ifsuldeminas.edu.br;

³ Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais – Campus Muzambinho. Muzambinho/MG - E-mail: walbert.santos@muz.ifsuldeminas.edu.br;

⁴ Instituto Federal de Educação, Ciência e Tecnologia do Sul de Minas Gerais – Campus Muzambinho. Muzambinho/MG - Email: jovaniluiz@outlook.com;

ALMEIDA,2012). Desta forma ele pode ser uma importante ferramenta no subsídio de informações confiáveis e atualizadas sobre uma determinada região.

Considerando esses conceitos, a proposta deste estudo é utilizar os dados, extraídos e organizados em um banco de dados, para extrair informações, obtidas a partir de imagens de satélite, e aplicar o algoritmo k-NN (*k-Nearest Neighbor*), para reconhecimento de padrões, com o intuito de comparar os valores dos pixels, quanto às diferentes bandas do satélite, visando identificar áreas cafeeiras no município de Muzambinho (MG).

2. MATERIAL E MÉTODOS

O experimento utilizou o método de extração dos dados de imagens de satélite (Aguilar et al, 2017) para a criação e atualização do banco de dados, sendo que a última data da captura da imagem e consequente inserção no banco de dados foi 01 de dezembro de 2017, desde a data inicial da série Landsat8. Posteriormente, foram estudados conceitos de Processamento de Imagens, Visão Computacional e Inteligência Artificial, que forneceram o entendimento mínimo destas áreas e a base para a construção do algoritmo para classificação de padrões. Para essa tarefa, o método de aprendizagem de máquina escolhido foi o k-NN (*k-Nearest Neighbor*), no qual a classificação ocorre por meio de um esquema de votação entre k vizinhos, mais próximos do elemento a ser rotulado. A implementação foi realizada na linguagem Python, por possuir bibliotecas voltadas para a aplicação de algoritmos de aprendizagem de máquina.

O processo teve início com a coleta de dados e/ou a seleção de pontos de “café” e “não café”, demarcados com a ferramenta Google Earth, que possibilitava visualizar lavouras de áreas cafeeiras conhecidas. Estas seleções, identificadas por coordenadas através do software de georreferenciamento QGIS v.2.18.3, gerou um conjunto de dados que, após ser editado, foi inserido em tabelas do banco de dados resultando em 2 arquivos para cada data: o arquivo “Treino.txt”, contendo os valores das 11 bandas de 335 pixels de “café” e 617 pixels de “não café”, e o arquivo “Validacao.txt”, contendo também 11 valores de bandas para 518 pixels de “café” e 617 pixels de “não café”. Cada pixel, ou cada registro destas tabelas, equivalem a uma área rotulada de 30x30m. Um novo arquivo, referente a área do IFSULDEMINAS - Campus Muzambinho, foi gerado contendo também 11 valores de bandas e as coordenadas x e y dos pontos.

3. RESULTADOS E DISCUSSÕES

Para verificar o desempenho do método escolhido (k-NN), foram utilizadas 05 (cinco) diferentes datas de capturas das imagens de satélites, que teve como critério de seleção a “presença” e “ausência” de nuvens, já que estas, talvez pudessem influenciar diretamente nos resultados. Para

cada data, 05 (cinco) combinações de bandas foram realizadas: as 11 bandas, as bandas 2,3,4, as bandas 2,4,5, as bandas 2,4,8 e as bandas 2,3,4,5. Por fim, após o treinamento, a verificação do desempenho do algoritmo foi feita executando o algoritmo – para cada data e em cada uma das combinações de bandas – e alternando o valor de k com o intuito de obter o melhor resultado de acertos em cada execução. Os melhores resultados, encontram-se na tabela 01.

Datas	Combinação de Bandas	Melhor valor para k	Acertos “café”	Acertos “Não Café”	Porcentagem de acertos (0 a 1)
31.07.2013	B11-Todas	$k=3$	510/518	608/617	0.985022026
08.02.2014	B234-RGB	$k=39$	473/518	605/617	0.9497797357
22.08.2015	B234-RGB	$k=5$	494/518	610/617	0.9726872247
08.08.2016	B2345-RGBIP	$k=3$	257/518	408/617	0.5859030837
28.09.2017	B234-RGB	$k=5$	507/518	611/617	0.9850220264

Tabela 01 - Os melhores resultados do método k -NN em diferentes configurações de banda e k vizinhos

Com base nesses procedimentos, procurou-se fazer a classificação de culturas cafeeiras especificando a área pertinente ao Campus Muzambinho. Desta forma, foi gerado um novo arquivo contendo as coordenadas dos pontos classificados como “café” e exportados para a ferramenta Google Earth a fim de validar o seu desempenho, comparando-os ao arquivo gerado anteriormente desta área. Após a classificação, os números foram confrontados com um mapeamento manual das áreas de café referente às datas trabalhadas. Os resultados podem ser visualizados na tabela 02.

Datas	Combinação de Bandas	Pontos indicados como “café” pelo algoritmo	Quantidade de pontos para “café” no Campus	Número de acertos para “café”
31.07.2013	B11-Todas	933	192	159
08.02.2014	B234-RGB	1297	162	42
22.08.2015	B234-RGB	657	162	96
08.08.2016	B2345-RGBIP	257	133	28
28.09.2017	B234-RGB	667	118	81

Tabela 02 - Resultados da classificação de café na área do IFSULDEMINAS - Campus Muzambinho

Alguns pontos nas áreas de café não foram reconhecidos como “café”, pois o cafeeiro possui um comportamento espectral muito variável (MOREIRA et al 2004) e o algoritmo também identificou áreas de mata e eucalipto como “café”, que apresentaram semelhança no comportamento espectral com o cafeeiro, conforme observado por MOREIRA et al (2007). Tal fato, também dificulta sua classificação por meio de métodos automatizados.

4. CONCLUSÕES

Os resultados obtidos, por apresentarem diversas variações, mostram que o uso de algoritmos de aprendizagem de máquina, na classificação de lavouras, ainda precisa ser muito explorado para obtenção de melhores resultados. São muitos os algoritmos e o conjunto de dados, adequados para a criação de um padrão, visando a correta classificação. É possível que os algoritmos utilizados não consigam identificar adequadamente diversas culturas, confundindo áreas de “mata” e “eucalipto” com o “café”, por exemplo, por apresentarem semelhança no comportamento espectral. Porém, um estudo sobre esta temática e a melhor compreensão de como essa questão pode ser melhorada, pode-se obter resultados bastante interessantes e significativos. A presença de nuvens é outro fator que pode influenciar diretamente nos resultados e contribuir para também gerar confusão na identificação de uma cultura agrícola, além da escolha adequada do satélite para obtenção de dados.

AGRADECIMENTOS

Agradecimentos especiais ao NIPE do IFSULDEMINAS pela bolsa de Iniciação Científica.

REFERÊNCIAS

- ANAIS DA JORNADA CIENTÍFICA E TECNOLÓGICA E SIMPÓSIO DE PÓS-GRADUAÇÃO DO IFSULDEMINAS - VOLUME 9 - 2017. **Levantamento e organização de dados de culturas cafeeiras localizadas na cidade de Muzambinho baseadas em imagens de satélite.** Muzambinho, MG, 2017. Disponível em: <https://jornada.ifsuldeminas.edu.br/index.php/jcmch4/jcmch4/paper/viewFile/3394/2353> acesso 05/08/2018.
- AZEVEDO, E.; CONCI, A. **Computação Gráfica: Teoria e Prática.** 1a Edição. São Blucher. São Paulo, 2003.
- FACELI, K. et al; **Inteligência Artificial: Uma Abordagem de Aprendizagem de Máquina.** Rio de Janeiro: LTC, 2011.
- MENESES, P. R.; ALMEIDA, T. D. **Introdução ao Processamento de Imagens de Sensoriamento Remoto.** Universidade de Brasília, Brasília, 2012, 266 p.
- MOREIRA, M. A; ADAMI, M; RUDORFF, B. F. T. **Análise espectral e temporal da cultura do café em imagens Landsat.** Pesquisa Agropecuária Brasileira, v. 39, n. 3, p. 223-231, 2004.
- MOREIRA, M.A.; BARROS, M.A.; FARIA, V.G.C. de; ADAMI, M. **Tecnologia de informação: imagens de satélite para o mapeamento de áreas de café de Minas Gerais.** Informe Agropecuário, v.28, p.27-37, 2007.