

# COMPARAÇÃO CLASSIFICADORES EM IMAGENS DE SATÉLITES COM DIFERENTES RESOLUÇÕES ESPACIAIS

Thiago da Conceição Martins<sup>1</sup>, Everton Luís Poelking<sup>2</sup>

<sup>1</sup>Engenheiro Florestal. Programa de Pós-graduação em Ciência Agrárias, Universidade Federal do Recôncavo da Bahia, Rua Rui Barbosa, 710, Centro Cruz das Almas-Bahia CEP 44380-000. [thiagocmarts@gmail.com](mailto:thiagocmarts@gmail.com); <sup>2</sup>Professor Dr. Universidade Federal do Recôncavo da Bahia. Centro de Ciências Agrárias, Ambientais e Biológicas. Rua Rui Barbosa, 710, Centro Cruz das Almas-Bahia CEP 44380-000. [everton@ufrb.edu.br](mailto:everton@ufrb.edu.br)

## RESUMO

O SR forma imagens através dados dos satélites fazendo a varredura superficial da terra. Com técnicas relacionadas à classificação de imagens reconhecendo padrões espectrais e espaciais. Objetivo do trabalho é avaliar diferentes classificadores em imagens de satélites com diferentes resoluções indicando o mais adequado. O processamento foi feito através dos programas SAGA GIS 7.7.0 ArcGis 10.2. Os classificadores foram: MaxVer, Redes Neurais, *Random Forest*, *Decision Tree*, *Support Vector Machine*. Foram utilizadas amostras de treinamento em diferentes quantidades 50 e 100 amostras para cada classe. SVM, RNA e MAXVER ultrapassaram a faixa dos 0,8 de Índice Kappa para duas quantidades amostrais. Landsat-8 os valores foram menores, com 50 amostras, para MaxVer e para SVM. Para extrair o melhor dos classificadores deve-se coletar mais amostras de treinamento, principalmente nos locais que mais apresentaram confusões entre as classes. Assim, é preciso uma maior pesquisa sobre a quantidade 'ótima' de amostras de treinamento.

**Palavra-chave:** Sensoriamento Remoto, *machine learning*, Índice Kappa.

## ABSTRACT

The SR forms images using satellite data by scanning the surface of the earth. With techniques related to image classification recognizing spectral and spatial patterns. The objective of the work is to evaluate different classifiers in satellite images with different resolutions indicating the most appropriate. The processing was done through the programs SAGA GIS 7.7.0 ArcGis 10.2. The classifiers were: MaxVer, Neural Networks, Random Forest, Decision Tree, Support Vector Machine. Training samples were used in different amounts 50 and 100 samples for each class. SVM, RNA and MAXVER exceeded the range of 0.8 Kappa Index for two sample amounts. Landsat-8 values were lower, with 50 samples, for MaxVer and for SVM. In order to extract the best from the classifiers, more training samples should be collected, especially in the places that presented the most confusion between the classes. Thus, further research is needed on the 'optimal' amount of training samples.

**Key words:** Remote Sensing, machine learning, Kappa Index.

## 1. INTRODUÇÃO

O Sensoriamento Remoto tem início na década de 1960, tornando-se uma das tecnologias que melhor possibilita fazer de coleta automática de dados, bem como o levantamento e monitoramento de recursos terrestres em escala global [1].

Existem, no entanto, algumas dificuldades para interpretar de formar eficiente as imagens de satélite, com efeito, os classificadores são as ferramentas que melhoram a compreensão dos usuários do Sistema de Informações Geográficas (SIG) que estão relacionadas ao processo de classificação de imagens.

Além disso, importantes técnicas relacionadas à classificação de imagens consistem no reconhecimento de padrões espectrais e espaciais possibilitando a produção de classes informacionais. Tendo em vista que, os classificadores chamados de *machine learning* são considerados mais adequados para reconhecimento de padrões espectrais, pois seus sistemas constroem modelos analíticos mais automatizados.

Os classificadores conhecidos como paramétricos possuem distribuição normal dos dados. Levando em consideração este parâmetro estatístico, um dos algoritmos mais utilizados é o de máxima verossimilhança (MAXVER). No entanto, os classificadores não paramétricos podem ser usados com base em dados que não se comportam com as características da curva normal [2].

Além de também serem chamados de classificadores *machine learning* por construírem modelos analíticos mais automatizados, são eles: *Decision Tree* (DT), *Support Vector Machine* (SVM), *Random Forest* (RF) e Rede Neural Artificial (RNA) [3]. Esses classificadores não paramétricos não fazem suposições sobre a distribuição de frequência, por isso, cada vez mais estão tornando-se populares para classificar dados remotos, que dificilmente dispõem de distribuições normais. Dentre eles os algoritmos RF e SVM têm se destacado, por apresentarem superioridade no tratamento das imagens [4].

Objetivo é avaliar o desempenho de diferentes classificadores supervisionados em imagens de satélites com diferentes resoluções espaciais e indicar o mais adequado. Para a análise do desempenho dos classificadores foi utilizado o Índice Kappa.

## 2. MATERIAL E MÉTODOS

Os dados necessários para o desenvolvimento do projeto, como imagens de satélite, foram adquiridos dos sítios de internet (*Planet e Earth Explore*). O processamento das imagens foi feito com programas de computador SAGA GIS 7.7.0. ©, Quantum GIS e ArcGis 10.2. ©.

O mapeamento do uso e cobertura das terras foi feito com base na classificação supervisionada de imagens matriciais dos satélites Landsat-8, Sentinel-2 e PlanetScope, às imagens são da mesma data 02/12/2019, com resolução espacial de 30, 10 e 3m respectivamente.

A classificação foi gerada no programa de computador SAGA GIS 7.7.0 ©. Os classificadores utilizados na classificação das imagens foram: (MAXVER) e os algoritmos *machine learning* Redes Neurais Artificiais (RNA), *Random Forest* (RF), *Decision Tree* (DT) e *Support Vector Machine* (SVM). Foram coletadas 20 amostras de treinamento para cada classe de uso da terra. Foi utilizada uma técnica para gerar pontos aleatórios dentro das amostras de treinamento, em diferentes quantidades 50 e 100 pontos para cada tipo de classe. Sendo assim, ao final do processamento das três imagens foram gerados 30 mapas temáticos.

Foi adotado o coeficiente Kappa para verificação da acurácia entre os diferentes usos da terra. Após à classificação foi calculado o índice Kappa para todas as imagens, comparando e analisando qual obteve melhor desempenho.

### 3. RESULTADOS

Observando os resultados dos classificadores para a imagem do Landsat-8, foram os menores valores de Índice Kappa em comparação aos outros sensores para as 50 amostras coletadas para cada classe de uso da terra, um dos fatores que pode ter interferido no resultado é a resolução da imagem, visto que a resolução espacial deste sensor é de 30m e foi feito o buffer de 50m raio para cada ponto. Porém, ao ter inserido as 100 amostras houve um aumento no desempenho dos valores de índice Kappa.

Classificadores	Kappa 50 amostras	Kappa 100 amostras
MAXVER	0,56	0,82
SVM	0,74	0,88
RNA	0,66	0,81
RF	0,67	0,80
DT	0,67	0,74

**Tabela 1. Resultado do desempenho dos classificadores para o satélite Landsat-8.**

Os resultados dos algoritmos de classificação apresentados para o satélite S-2/MSI apontam uma similaridade nos resultados, para as coletas feitas para 50 amostras de treinamento o Índice Kappa obteve os valores entre 0,73 para DT e 0,80 para RNA. Por outro lado, os valores do Kappa obtido da classificação para as 100 amostras de treinamento, os algoritmos forneceram valores índice Kappa entre 0,75, também para DT e 0,88 para SVM

de acordo com o desempenho da classificação estabelecido por Congalton e Green o desempenho dos algoritmos foi ‘muito bom’ e ‘excelente’, respectivamente (Tabela 2) [5].

Classificadores	Kappa 50 amostras	Kappa 100 amostras
MAXVER	0,77	0,82
SVM	0,78	0,88
RNA	0,80	0,81
RF	0,78	0,80
DT	0,73	0,75

**Tabela 2. Resultado do desempenho dos classificadores para o satélite Sentinel-2.**

Os intervalos dos valores do Índice Kappa quando a faixa for de 0,61 a 0,80, à alta precisão, por outro lado, quando o valor do Kappa for acima de 0,81 a classificação é muito alta. Com isso, seguindo a classificação dos altores os valores obtidos pelos classificadores ficaram entre as faixas de alta precisão e precisão muito alta (Tabela 3) [6].

Classificadores	Kappa 50 amostras	Kappa 100 amostras
MAXVER	0,80	0,84
SVM	0,84	0,80
RNA	0,80	0,80
RF	0,72	0,73
DT	0,71	0,72

**Tabela 3. Resultado do desempenho dos classificadores para o satélite PlanetScope.**

O mapeamento foi realizado na bacia do rio Capivari, com os diferentes sensores, como visto não apresentaram diferenças significativas. No entanto, para melhor representar a área com os determinados usos da terra e para fins de exemplificação, a figura 1 apresenta o mapa gerado pelo algoritmo SVM, para imagem do Sentinel-2 que obteve o maior valor de Índice Kappa dentre os três sensores dos satélites com 0,88, no entanto, poderia ser apresentado qualquer um dos mapas gerados pelos outros quartos classificadores, com os valores de Kappa aproximados.

Como visto nos resultados, os dados da Tabela 1 para a coleta das 50 amostras treinamento foram as que mais apresentaram erros de omissão e comissão, as classes que mais apontam esses erros foram Agricultura x Pastagem apresentaram pixels de mistura com 49% de acerto. Em estudos semelhantes evidenciaram também essa confusão, os autores, por sua vez, avaliaram as imagens do satélite S-2/MSI [7].

Esses pixels de mistura também estão presente nas imagens RapidEye para áreas urbanas verdes de classe, 25% dos pixels estão mal colocados ou classificados como

outra classe, consequentemente o kappa estimado é de 0,72 [8].

RNA e MAXVER para a classificação do uso e ocupação da terra, evidenciando os níveis de degradação das pastagens, os autores concluíram que o algoritmo RNA alcançou resultados levemente superior em comparação ao MAXVER.

#### 4. CONCLUSÃO

Desta maneira, os classificadores apresentaram desempenho ‘alta precisão’ e ‘precisão muito alta’ para as imagens do Sentinel-2 e PlanetScope, não apresentaram grandes destaques individuais, não possibilitando indicar qual dos classificadores podem ser indicados para cada imagem de satélite nas diferentes resoluções espaciais. Uma das possibilidades para extrair o melhor dos classificadores é coletar mais amostras de treinamento, principalmente nos locais que mais apresentaram confusões entre as classes de treinamento. Além disso, é preciso uma maior pesquisa sobre a quantidade ótima de amostras de entrada dos algoritmos *machine learning*.

#### 5. REFERÊNCIAS BIBLIOGRÁFICAS

[4] ADAM, E.; MUTANGA, O.; ODINDI, J.; ABDEL-RAHMAN, E. M. Land-use/cover classification in a heterogeneous coastal landscape using RapidEye imagery: evaluating the performance of random forest and support vector machines classifiers, **International Journal of Remote Sensing**, 2014. 35:10, 3440-3458, DOI: 10.1080/01431161.2014.903435.

[2] ADREOLA, R. Support Vector Machines na Classificação de Imagens Hiperespectrais, 130f. Dissertação (Mestrado em Sensoriamento Remoto) – **Centro Estadual de Pesquisa em Sensoriamento Remoto e Meteorologia**. Universidade Federal do Rio Grande do Sul. 2009.

[5] CONGALTON, R. G., GREEN, K. Assessing the accuracy of remotely sensed data: Principles and practices. **New York: Lewis Publishers**, 1999. 137p.

[1] Meneses, P. R.; Sano, E. E. (2012). Classificação pixel a pixel de imagem. in: Meneses, P. R. e Almeida, T. introdução ao Processamento de Imagens de Sensoriamento Remoto. **Unb: Brasília**. Cap. 12.

[3] MAS, J.F., FLORES, J.J., 2007. The application of artificial neural networks to the analysis of remotely sensed data. **Int. J. Remote Sens.** 29, 617–663.

[8] N. Kranjčić i D. Medak, "Evaluating Different *Machine Learning* Methods on RapidEye and PlanetScope Satellite Imagery", **Geodetski list**, vol.74 (97), br. 1, str. 1-18, 2020. [Online]. Dostupno na: <https://hrcak.srce.hr/237682>. [Citirano: 21.11.2020.].

[7] REX, F.E.; KÄFER, P.S.; DEBASTIANI, A.B.; KAZAMA, V.S.; Potencial de imagens MSI (Sentinel-2) para classificação do uso e cobertura da terra. **Enciclopédia Biosfera**, v.15, n.27, p.219-231, 2018. DOI: 10.18677/EnciBio\_2018A67.

[6] VIERA, A. J., GARRETT, J. M. (2005): Understanding interobserver agreement: the kappa statistic, **Fam Med**, 37.

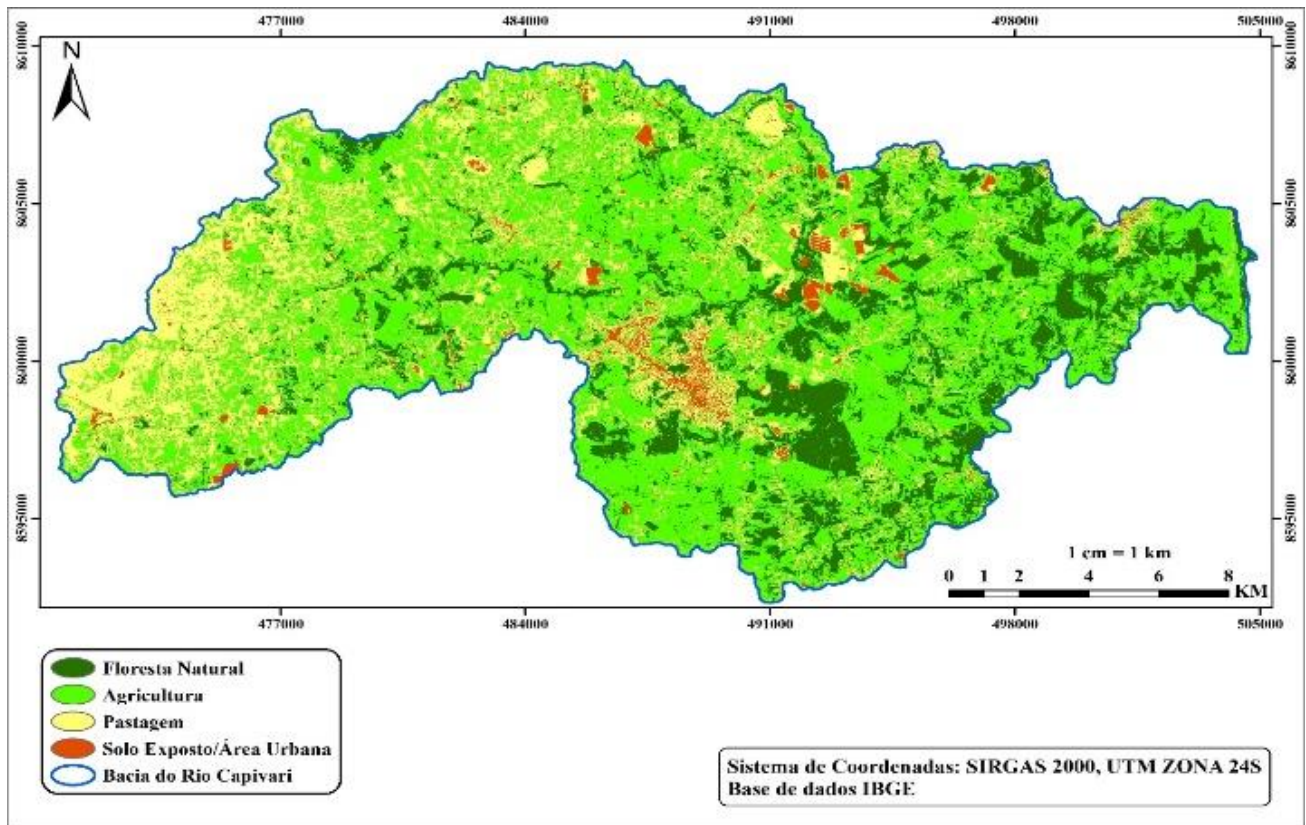


Figura 1. Classificação do uso e cobertura da terra da bacia do rio Capivari pelo algoritmo SVM.