



Mapeamento da evolução de vegetação usando imagens de alta resolução: uma comparação entre os classificadores por máxima verossimilhança, mineração de dados baseada em objetos e classificadores random forest

Yasmin Gelli, David Costa, Andréa Puzzi, Janie Garcia

Universidade Federal Fluminense, Rio de Janeiro/RJ. ykgelli@id.uff.br

Instituto Federal Fluminense, São João da Barra, RJ david.costa@iff.edu.br

Independent Researcher, Niterói, RJ. apuzzi@id.uff.br

Universidade Federal Fluminense, Niterói, RJ. janie55@terra.com.br

A evolução das tecnologias espaciais, o aumento da produção de imagens orbitais e aéreas e o melhoramento dos algoritmos de classificação, vem contribuindo para popularizar a aplicação de sensoriamento remoto no mapeamento do uso e cobertura da terra. Mapas temáticos produzidos a partir da classificação dessas imagens são uma das aplicações mais comuns do sensoriamento remoto e, nas últimas décadas, cientistas têm trabalhado para mapear estágios de sucessão vegetal através dessa tecnologia.

A análise da sucessão vegetal é um passo básico para realizar estudos e ações de manejo. Através dela são conhecidas quantitativa e qualitativamente as áreas estudadas, informações que podem contribuir para recuperação ambiental, educação ambiental, proteção ambiental, estabelecimento de áreas prioritárias para conservação, entre outros fins. A presente área de estudo, o Morro do Gragoatá em Niterói/RJ, situa-se no bioma da Mata Atlântica e tem um histórico de destruição antrópica. Reconhecido internacionalmente como um dos 36 *hotspots* mundiais, o bioma possui alta prioridade de conservação em virtude da sua enorme diversidade e do alto nível de ameaça em que se encontra.

Na classificação, foram utilizadas ortofotos disponíveis no Sistema de Geoinformação Geográfica da Prefeitura de Niterói. As imagens, de altíssima resolução espacial, possuem 10 cm de pixel em escala de 1:1000 e são já georreferenciadas e ortoretificadas. O aerolevanteamento referente ao ano de 2014 foi feito entre agosto e setembro e o de 2019, entre outubro e novembro. O pré-processamento das imagens envolveu a normalização radiométrica das mesmas, que busca minimizar as variações temporais que os detectores sofrem. Para isso, devem ser selecionados alvos caracterizados pela sua baixa ou nenhuma variação

VIII SRHIDRO

Seminário Regional sobre Gestão de Recursos Hídricos

para que seja estabelecida uma relação entre os níveis digitais das imagens analisadas. A normalização radiométrica é uma correção relativa, já que não exige parâmetros atmosféricos. Sendo assim, o conjunto de dados de controle foi retirado de amostras claras e escuras (*pseudo-invariant features* - PIFs) de acordo com dados da própria imagem. A imagem de 2014 foi utilizada como referência para normalizar a de 2019. Foram coletados 30 PIFs considerados como pouco ou nada variantes em ambas as imagens. Através da ferramenta Calculadora Raster, do software QGIS 3.16.2, as equações lineares foram aplicadas na imagem de 2019, obtendo como resultado, a imagem radiometricamente normalizada.

Os três métodos utilizados para mapear a área em análise foram supervisionados. O software TerraView 5.6.1. foi utilizado para a classificação com o algoritmo Maxver e para a orientada a objeto, com o plugin GeoDMA. Depois do treinamento das amostras, a classificação através do Maxver funciona a partir da cor e do brilho emitido pelos pixels. São diferenciados e, posteriormente, assimilados às diferentes classes através da ponderação das distâncias entre médias dos níveis digitais das mesmas, utilizando parâmetros estatísticos. O sistema GeoDMA (Geographical Data Mining Analyst) tem seu funcionamento acoplado ao programa TerraView, desenvolvido pelo INPE. A entrada do GeoDMA é composta de imagens e objetos (segmentos) resultantes da segmentação, feita em etapa anterior à classificação. O algoritmo utilizado para segmentar a imagem foi por crescimento de regiões, uma vez que esse método considera a característica espacial dos dados e garante a formação de segmentos fechados. Os parâmetros de tamanho máximo de segmento e limiar de similaridade foram ajustados, de acordo com a literatura respectivamente, em 1000 pixels e em 0.20, evitando assim que pixels muito distintos fossem agregados em um mesmo segmento. A classificação do sistema GeoDMA pode ser feita a partir de dois algoritmos, o supervisionado de árvore de decisão e o não-supervisionado de Mapas Auto-Organizáveis (SOM), nesse trabalho foi utilizado o de árvores de decisão. A análise feita integra informação espacial, contida nos segmentos, e informação espectral extraída das bandas de entrada, proporcionando ao usuário uma vasta gama de atributos e estatísticas para serem utilizados na classificação.

Por último, temos a classificação pelo algoritmo do Random forest (RF), feita pela plataforma em nuvem Google Earth Engine, também funciona baseado em pixel e utiliza o algoritmo de árvores de decisão. Cada árvore gerada é treinada individualmente a partir da seleção aleatória (por isso o seu nome, "floresta aleatória") de subconjuntos das amostras de treinamento coletadas e seus respectivos atributos. Um conjunto grande de árvores individuais é produzido e os pixels são atribuídos, por cada uma, às classes existentes. A classe mais popular presente nesse processo é então escolhida através do consenso entre as árvores.

VIII SRHIDRO

Seminário Regional sobre Gestão de Recursos Hídricos

Nas três classificações, foram consideradas quatro classes: arbórea, arbustiva, herbácea e solo exposto, adaptadas ao presente estudo partir da Resolução CONAMA 06/94. A classe arbórea é composta por vegetação arbórea, com idade entre 8 até mais de 30 anos, altura média de 4 a 15 m. Dossel com porte variável conforme a localização. Possui diversidade variável, com espécies nativas pioneiras frutíferas (embaúba, carrapeta, camboatá, capororoca, capororocão, arco de pipa, crindiúva, murici) que possuem mudas em seu sub-bosque, além de arbustos, ervas, lianas, sem epífitas. Serapilheira com desenvolvimento variável. A classe arbustiva foi definida pela presença de vegetação arbustiva, de crescimento rápido, pouco diversificada, com idade entre 6 e 20 anos, altura média entre 1 e 2 m (alecrim do campo, quaresmeira, cambará). A serapilheira é reduzida ou ausente. Mais comum em parte do topo e encostas íngremes do terço médio e superior onde o solo é raso (13). Já a classe herbácea, é composta por vegetação com idade entre 0 e mais de 25 anos e até 1 m de altura (especialmente gramíneas, ciperáceas, leguminosas, orquidáceas, cactáceas). Ocorre em parte do topo e encostas onde o solo é raso, pobre em matéria orgânica.

Foram coletadas cerca de 20 amostras para cada classe e cada algoritmo em ambas as imagens, selecionadas cuidadosamente a fim de garantir representatividade das características de cada alvo. Para análise da evolução da cobertura vegetal do Morro do Gragoatá foram utilizadas duas imagens de alta resolução dos anos de 2014 e 2019. A avaliação da acurácia dos classificadores foi feita através da imagem de 2019. Para avaliar as métricas de acurácia das classificações foram geradas matrizes de confusão cruzando os resultados das classificações com dados de referências, obtidos por cuidadosa interpretação visual. Para cada classificação, utilizou-se 100 pontos, 25 para cada uma das quatro classes, distribuídos de forma estratificada, aleatória e independente, uma vez que são diferentes das amostras de treinamento, garantindo assim a representação significativa e distribuída de todas as classes da imagem. A partir da matriz de confusão, foram calculados a exatidão global, exatidão do produtor, exatidão do usuário. Posteriormente, aplicou-se aquele com melhor desempenho na imagem de 2014. A classificação foi vetorizada para a obtenção de polígonos com áreas mensuráveis.

Segundo a análise da matriz de confusão dos três classificadores para 2019, a melhor separação entre classes, foi resultado do algoritmo *Random forest*, com uma exatidão global de 92%. O GeoDMA e o Maxver tiveram pouca diferença na exatidão com 76% e 74% respectivamente. Todos os algoritmos apresentaram a mesma porcentagem de acerto de usuário para classe arbórea (88%). No GeoDMA, entretanto, houve confusão dessa com a classe de solo exposto. Já para a classe arbustiva houve grande diferença na exatidão do usuário entre os métodos. O

VIII SRHIDRO

Seminário Regional sobre Gestão de Recursos Hídricos

classificador por máxima verossimilhança foi o que fez a maior confusão, com mais da metade das amostras sendo atribuídas erroneamente à classe arbustiva. Para a classe herbácea, no *Random forest* houve amostras de solo exposto atribuídas a ela, já no GeoDMA e Maxver, essa confusão aconteceu com maior frequência para com a classe arbustiva. Dentre as classificações, a classe de solo exposto só não foi bem separada da classe herbácea.

O algoritmo random forest foi aplicado à imagem de 2014 para análise comparativa e foi percebida uma grande evolução na cobertura vegetal entre os anos de 2014 e 2019. A única classe que expandiu sua área de cobertura foi a arbórea, com um aumento de quase 25% em relação a sua área em 2014. As classes arbustiva e herbácea tiveram uma redução percentual bastante próximas de cerca de 30%, a classe arbustiva com a maior redução em números absolutos. Já a classe de solo exposto foi a que teve maior redução percentual de pouco mais de 75%. Com esses resultados é possível afirmar que a região de estudo teve boa parte de sua área regenerada através da sucessão vegetal em seus estratos.

O classificador *Random forest* destacou-se como o melhor método ao fornecer informações consistentes sobre a evolução da cobertura vegetal no Morro do Gragoatá, podendo ser aplicado a outros trabalhos com objetivos semelhantes. Para próximos estudos, é indicado o uso de imagens com melhor resolução espectral a fim de aperfeiçoar a distinção de classes e as acurácias dos métodos Maxver e GeoDMA.

As técnicas para recuperar áreas degradadas, foram eficientes na redução das classes: solo exposto, vegetação herbácea e arbustiva. A classe arbórea aumentou em quase 25% no período analisado. Apontar e avaliar mudanças no uso e cobertura da terra são de suma importância, principalmente no tão ameaçado bioma da Mata Atlântica para aumentar o uso sustentável do ambiente. O apoio a projetos que possibilitem tal ação, como o financiamento do Ministério do Meio Ambiente na área de estudo e a avaliação de sua efetividade como a presente pesquisa são essenciais. Além disso, a metodologia empregada pode ser replicada em estudos similares.

Sensoriamento remoto; Random forest; Morro do Gragoatá.

Nível de Ensino: Graduação