

Imagens de satélite, geologia e IA

CITAÇÃO

Fernandes, J. C. *et al.* (2022)
Imagens de satélite, geologia e IA,
Rev. Ciência Elem., V10(04):053.
doi.org/10.24927/rce2022.053

EDITOR

João Nuno Tavares
Universidade do Porto

EDITOR CONVIDADO

Alexandre Lopes Magalhães
Universidade do Porto

RECEBIDO EM

23 de novembro de 2022

ACEITE EM

23 de novembro de 2022

PUBLICADO EM

20 de dezembro de 2022

COPYRIGHT

© Casa das Ciências 2022.
Este artigo é de acesso livre,
distribuído sob licença Creative
Commons com a designação
[CC-BY-NC-SA 4.0](https://creativecommons.org/licenses/by-nc-sa/4.0/), que permite
a utilização e a partilha para fins
não comerciais, desde que citado
o autor e a fonte original do artigo.

rce.casadasciencias.org



Joana Cardoso-Fernandes, Alexandre Lima, Ana Cláudia Teodoro
FCUP/ICT/ Universidade do Porto

A geologia, enquanto Ciência, procura fundamentalmente compreender os processos que levaram à formação do planeta Terra e à sua evolução contínua ao longo de milhões de anos, incluindo o ciclo de formação das rochas e a sua relação com a atmosfera, a biosfera e a hidrosfera, a colisão e rutura dos continentes como resultado da tectónica de placas, mas também compreender a formação de recursos geológicos e como explorá-los.

A guerra na Ucrânia trouxe-nos uma nova perspetiva sobre as políticas energéticas, a transição energética e até sobre a soberania energética com os países que tentam encontrar soluções para a crise energética sem dar um passo atrás nos objetivos de desenvolvimento sustentável para alcançar a descarbonização global e combater as alterações climáticas.

Embora, como cientistas, possamos ter pouca influência sobre os decisores, podemos, no entanto, fazer a nossa parte no processo de transição energética. As energias renováveis, como a energia solar ou eólica, requerem uma multiplicidade de metais nos processos de fabrico e, para a maioria deles, a recuperação e reciclagem de peças antigas ainda não é possível, e pode não ser suficiente num futuro próximo. Portanto, a transição energética coloca uma grande pressão e cria uma grande procura de metais que ocorrem naturalmente na crosta terrestre, muitas vezes em conjuntos minerais estáveis em depósitos minerais. Estes depósitos minerais podem encontrar-se numa grande variedade de rochas e configurações. Um exemplo deste tipo de rocha são os pegmatitos, que podem ser explorados para minerais e elementos estratégicos para a digitalização e transição energética, como o quartzo de alta pureza para ser usado em ótica, fibras, semicondutores, energia solar e o lítio para mobilidade elétrica e armazenamento de energias renováveis. Se existirem na Europa, menor pegada carbónica terão ao serem explorados e convertidos no velho continente.

Durante as últimas quatro décadas, os dados de deteção remota, nomeadamente imagens de satélite, têm sido uma importante ferramenta na prospeção geológica para encontrar diferentes tipos de depósitos minerais, especialmente em áreas remotas e/ou de difícil acesso. Recentemente, dados de deteção remota e, conseqüentemente, imagens de satélite, têm sido usados para desenvolver novas ferramentas de prospeção de baixo impacto, menos invasivas para as populações, flora e fauna. Estes instrumentos foram melhorados ao longo dos anos para a identificação de pegmatitos, uma importante fonte de recursos minerais.

Assim, é fundamental saber quais os dados (imagens de satélite) disponíveis para o desenvolvimento de tais abordagens e compreender as diferenças, vantagens e desvantagens de cada produto de satélite (FIGURA 1). Os sensores dos satélites medem a radia-

ção eletromagnética em gamas específicas de comprimento de onda, que é quantificada e convertida numa imagem digital (*raster*), chamada banda, composta por uma matriz bidimensional de elementos de imagem (*pixels*). A intensidade de cada *pixel* corresponde a um valor numérico discreto armazenado como níveis cinzentos distintos, com níveis mais brilhantes para as mais altas intensidades. Cada banda tem uma determinada dimensão (*frame*), correspondente à área na superfície da Terra capturada por uma imagem, e cada *pixel* corresponderá a uma área menor (geralmente quadrada) dentro da *frame* da imagem. Com esta informação podemos comparar imagens de satélite de acordo com (i) a resolução espectral (quanto mais estreita for a faixa de comprimento de onda da banda e o maior número de bandas, melhor a resolução); (ii) a resolução radiométrica, quanto maior for o número de *bits*, maior o número de níveis cinzentos disponíveis e a resolução; e (iii) a resolução espacial (quanto menor a área representada por um *pixel*, maior é a resolução). Há também um outro tipo de resolução, específica aos dados de satélite, que é a resolução temporal, ou seja, o tempo necessário para revisitar a mesma região da Terra.

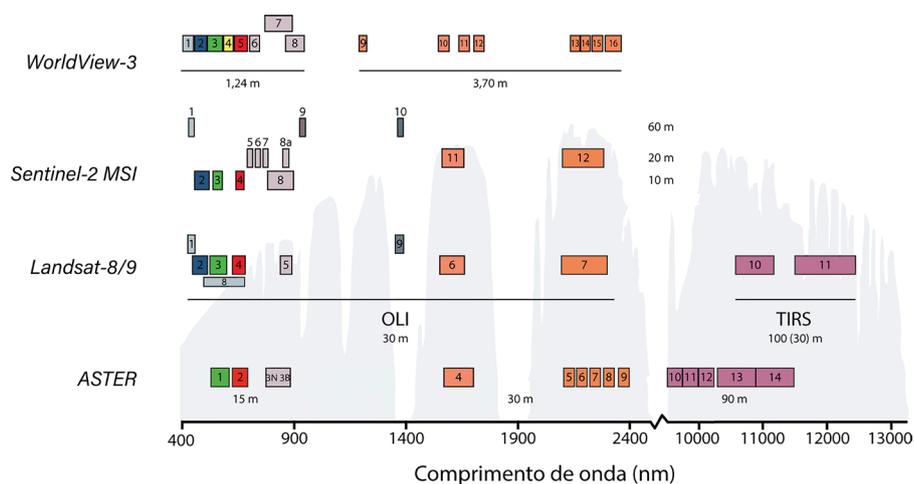


FIGURA 1. Comparação da resolução espectral e espacial de sensores de satélite distintos. O satélite *Landsat-8* contém dois sensores a bordo: *Operational Land Imager* (OLI) e *Thermal Infrared Sensor* (TIRS). O *Sentinel-2* transporta o sensor *Multispectral Instrument* (MSI).

O sensor *ASTER* a bordo do sistema de observação da Terra (*EOS*) da *NASA*, lançado em dezembro de 1999, fornece dados cruciais para os geólogos delinear depósitos minerais ao longo dos anos devido à sua resolução espectral média-alta, com dados armazenados em 14 bandas cuja resolução espacial e radiométrica varia com o comprimento de onda (FIGURA 1). As bandas 1–3 têm uma resolução espacial de 15 m com quantização em 8 *bits* (256 níveis de cinza). As bandas 4–9, também com quantização em 8 *bits*, têm uma resolução de 30 m. As bandas 10–14 têm uma resolução espacial de 90 m com quantização em 12 *bits* (níveis cinzentos de 0 a 4 095). Infelizmente, houve uma falha num dos módulos do *ASTER* em 2008, afetando as bandas 4–9 para futuras utilizações.

O *Landsat-8*, lançado em 2013, e o seu sucessor *Landsat-9*, lançado recentemente, em 2021, apresentam nove bandas espectrais com uma resolução espacial de 30 m para bandas 1–7 e 9, 15 m para a banda 8 (pancromática), e duas bandas térmicas (10 e 11) que são recolhidas a 100 m mas reamostradas para 30 m de resolução espacial (FIGURA 1). Enquanto as bandas do *Landsat-8* apresentam os dados em 12 *bits* (ou seja, 4 096 tons de cinza), o *Landsat-9* tem uma resolução radiométrica melhorada de 14 *bits* (16 384 níveis de cinza distintos para um dado comprimento de onda).

A missão *Sentinel-2*, composta pelos satélites gêmeos *Sentinel-2A* e *Sentinel-2B*, no âmbito do *Programa Copernicus da União Europeia*, fornece produtos de satélite com 13 bandas espectrais (FIGURA 1): quatro bandas a 10 m de resolução espacial, seis bandas a 20 m e três bandas a 60 m de resolução espacial, todas com uma resolução radiométrica de 12 *bits* (4 096 níveis de cinzento).

Todos os satélites mencionados fornecem dados abertos a pedido, uma vez que as missões correspondentes são lideradas por agências espaciais governamentais. No entanto, os satélites comerciais estão a ser cada vez mais utilizados devido às suas melhores especificações/resoluções. Um exemplo é o satélite *WorldView-3* da *Maxar*, lançado em 2014, um sensor multiespectral de muito alta resolução espacial que adquire dados em 11 *bits* (2 048 níveis de cinzento) em oito bandas recolhidas a 1,24 m de resolução espacial e dados de 14 *bits* (16 384 níveis de cinzento) em oito bandas a 3,70 m de resolução (FIGURA 1), embora os dados sejam reamostrados para uma resolução inferior para distribuição comercial. As imagens são armazenadas como inteiros de 16 *bits* (65 536 níveis de cinzento) ou são reduzidas a 8 *bits* (256 níveis de cinzento) para reduzir os tamanhos dos ficheiros e melhorar o desempenho do processamento de imagem.

Portanto, se considerarmos um produto *Sentinel-2* (290 km²) onde todas as bandas foram reamostradas para uma resolução de 10 m, cada banda será composta por 5 490 x 5 490 (ou seja, 30 140 100 *pixels* por banda) totalizando 391 821 300 *pixels* nas 13 bandas. Enquanto isso, se considerarmos uma imagem *Worldview-3* cobrindo 100 km², com todas as bandas cobrindo a mesma gama de comprimentos de onda de *Sentinel-2* (FIGURA 1) reamostrada para uma resolução de 2 m, então cada banda será composta por 26 403 836 *pixels*, totalizando 422 461 376 *pixels* em todas as 16 bandas. A FIGURA 2 apresenta uma comparação entre a resolução espacial e os respetivos cubos de dados para o campo pegmatítico de Tysfjord na Noruega, uma das áreas de estudo do projeto GREENPEG.

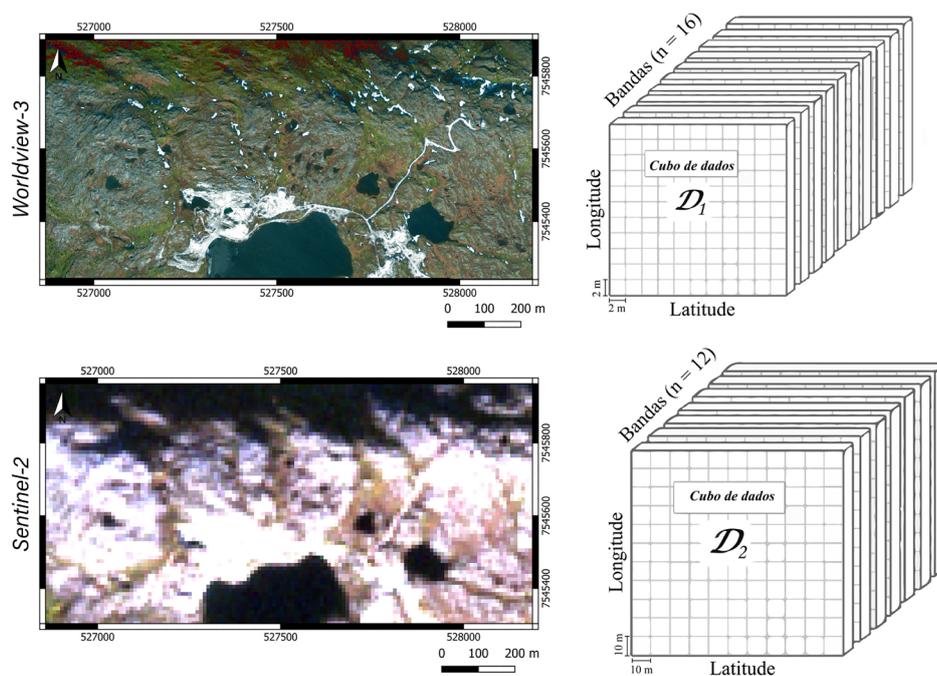


FIGURA 2. A) Comparação da resolução espacial dos satélites *Worldview-3* e *Sentinel-2* sobre um pegmatito conhecido em Tysfjord, Noruega. B) Comparação dos respetivos cubos de dados.

Assim, é claro como apenas dois produtos diferentes com resolução espacial média-alta e muito alta podem criar conjuntos de dados gigantescos com milhões de entradas de dados. No entanto, há outro tipo de sensor caracterizado por uma resolução espectral muito elevada: o sensor hiperespectral. Em 2019, a Agência Espacial Italiana lançou o satélite hiperespectral *PRISMA* que possui 238 bandas na região de comprimento de onda entre os 400 e os 2 500 nm com uma resolução espacial de 30 m. Com uma largura de 1 254 *pixels* por uma altura de 1 239 *pixels*, cada banda é composta por 1 553 706 *pixels*, e o total das 238 bandas resulta em 369 782 028 *pixels*. A FIGURA 3 apresenta uma pequena comparação entre o número de bandas na região de 400–1 000 nm para os satélites *PRISMA* e *Sentinel-2*, bem como uma comparação da resolução espacial de ambos os produtos de satélite na região de St. Austell na Cornualha (Reino Unido), muito semelhante geologicamente à maior mina de lítio a céu aberto a abrir ainda nesta década, que foi recentemente anunciada no Maciço Central Francês. A FIGURA 3B) apresenta o cubo de dados tridimensional para o produto *PRISMA* que cobre a área de St. Austell¹.

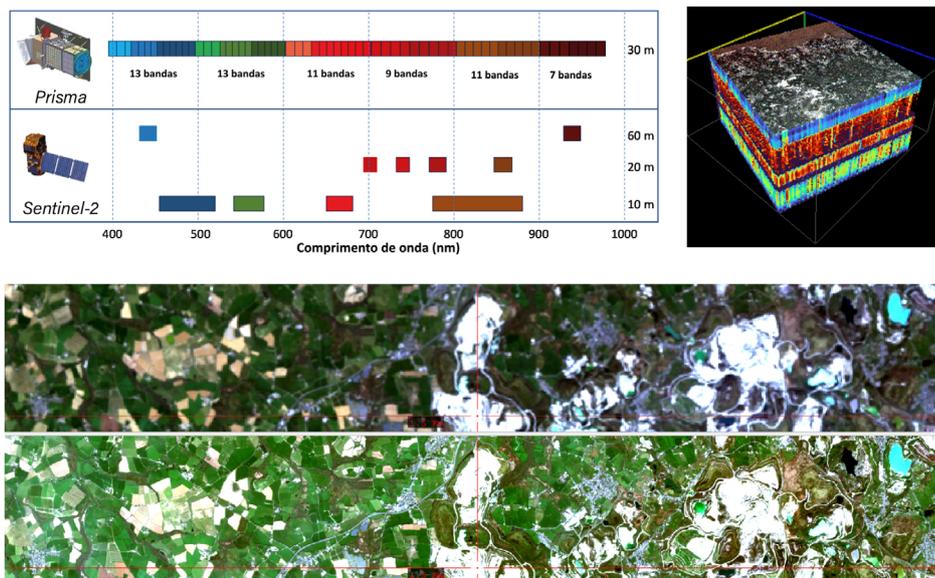


FIGURA 3. A) Comparação da resolução espectral dos satélites *PRISMA* e *Sentinel-2* na região de 400–1000 nm². B) Cubo de dados *PRISMA* para uma área de estudo em St. Austell, Cornwall (Reino Unido)³. C) Comparação da resolução espacial dos satélites *PRISMA* (topo) e *Sentinel-2* (base) na região de St. Austell³.

É evidente que não só é um desafio escolher o produto satélite certo, tendo em conta as características de uma área de estudo e os resultados esperados, mas também encontrar formas rápidas e precisas de lidar com esses grandes conjuntos de dados. Os algoritmos clássicos de processamento de imagem utilizados em aplicações geológicas são: (i) combinações RGB; (ii) rácios de bandas; (iii) análise principal dos componentes; (iv) *matched filtering* (MF); (v) *linear spectral unmixing* (LSU), entre outros. No entanto, a maioria destes algoritmos não são adequados para trabalhar com conjuntos de dados muito grandes. Recentemente, novos desenvolvimentos foram feitos através do uso de inteligência artificial, isto é, métodos como algoritmos de *machine* e *deep learning*. Alguns dos algoritmos de *machine learning* mais utilizados são Máquinas de Suporte Vetorial (SVMs)⁴, Redes Neurais (ANNs)^{5,6} e *Random Forests* (RF)⁷. A FIGURA 4 resume esquematicamente o funcionamento destes três algoritmos como classificadores para separar duas ou mais classes.

Os algoritmos de *machine learning* têm uma grande capacidade para lidar com dados de alta dimensão (ou seja, muitas variáveis, neste caso, bandas de satélite) e modelar classes distintas com assinaturas espectrais complexas, sem fazer suposições sobre a distribuição de dados. Portanto, tais métodos tendem a produzir maior precisão em comparação com os classificadores tradicionais.

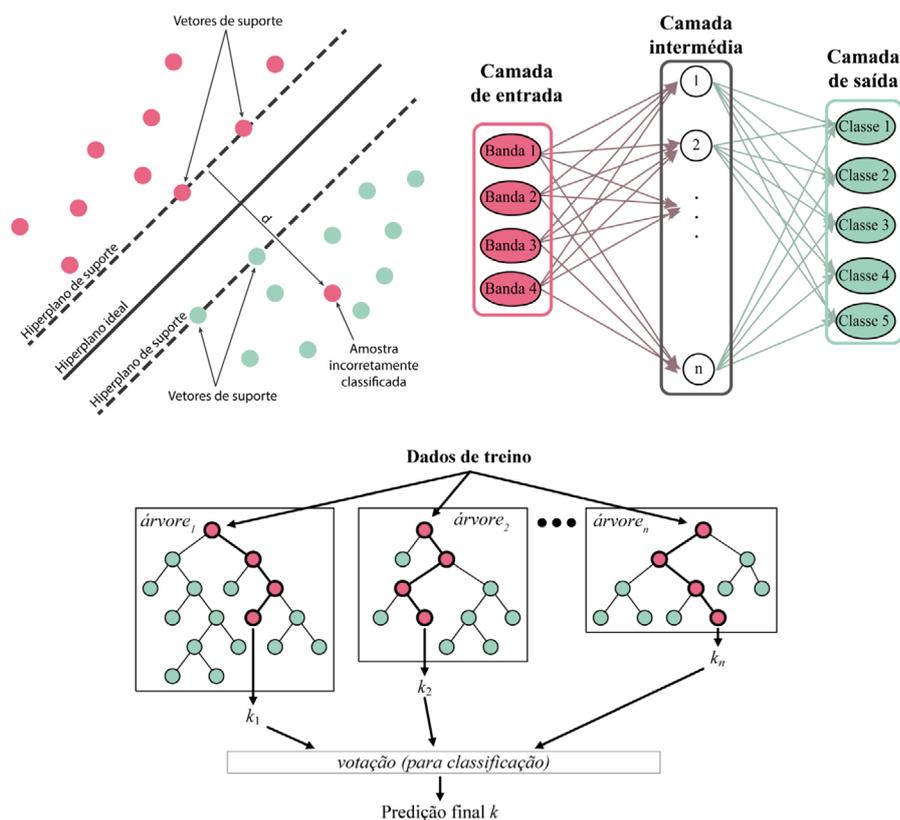


FIGURA 4. Representação esquemática de como funcionam algoritmos distintos de machine learning: A) SVM tenta encontrar um hiperplano ideal que separa duas classes^{4, 8}. B) A ANN armazena as informações em nós distribuídos em camadas paralelas consecutivas, em que cada nó numa camada está ligado a todos os nós das camadas adjacentes⁵. C) RF composto por um conjunto de árvores, cada uma votando para obter uma previsão final⁹.

As SVMs tentam encontrar um hiperplano ideal que separa duas classes, maximizando a distância entre elas (FIGURA 4A)). São uma boa opção quando o conjunto de dados de treino (isto é, os pixels da imagem de satélite cuja classe é previamente conhecida) é pequeno. As ANNs correspondem a uma classe de modelos que tentam imitar o cérebro humano, uma vez que adquirem conhecimento através de um processo de aprendizagem paralela (FIGURA 4B)), sendo métodos bastante versáteis para resolver tarefas altamente complexas. O desafio é especificar corretamente os parâmetros, uma vez que as ANNs facilmente sobretreinam (isto é, produzem uma classificação com *overfitting*). Para evitar este fenómeno de *overfitting*, o algoritmo RF é uma alternativa muito robusta, uma vez que corresponde a um classificador de conjunto, ou seja, combina um grande número de preditores menores (árvores de decisão) que treinam em porções diferentes e aleatórias do conjunto de dados. A previsão de todas as árvores individuais é combinada num processo de votação para definir a predição final da classe (FIGURA 4C)).

Já algoritmos de *deep learning* como *convolutional neural networks* (CNNs), *stacked auto-encoders* (SAEs), and *deep belief networks* (DBNs) podem ser ferramentas muito poderosas para analisar conjuntos de dados de séries temporais, por exemplo.

Considerações finais

É evidente que nós, como cientistas, temos de sair da nossa “zona de conforto” na procura de formas e métodos inovadores para lidar com os desafios atuais que enfrentamos, explorando e gerindo conjuntos de dados cada vez maiores como resultado da evolução tecnológica e da digitalização das sociedades. Dessa forma poderemos arranjar soluções que nos permitam combater as alterações climáticas, como por exemplo acelerar a descarbonização da economia, por exemplo através da aceleração da transição energética, que forçosamente se vai fazer à custa de encontrar mais metais, principalmente a menor distância de onde eles vão ser fabricados e utilizados, de forma a reduzirmos a sua pegada carbónica.

REFERÊNCIAS

- ¹ CARDOSO-FERNANDES, J. et al., [Detecting Lithium \(Li\) Mineralizations from Space: Current Research and Future Perspectives](#), *Applied Sciences*, 10, 5, 1785. 2020.
- ² NIROUMAND-JADIDI, M. et al., [Water Quality Retrieval from PRISMA Hyperspectral Images: First Experience in a Turbid Lake and Comparison with Sentinel-2](#), *Remote Sensing*, 12, 23, 3984. 2020.
- ³ CARDOSO-FERNANDES, J. et al., [Assessing the PRISMA Potential for Mineral Exploration to Vector Low-Grade Lithium Deposits](#), *IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium*, 5985-5988. 2022.
- ⁴ CARDOSO-FERNANDES, J. et al., [Semi-Automatization of Support Vector Machines to Map Lithium \(Li\) Bearing Pegmatites](#), *Remote Sensing*, 12, 14, 2319. 2020.
- ⁵ TEODORO, A. C., [Applicability of data mining algorithms in the identification of beach features/patterns on high-resolution satellite data](#), *Journal of Applied Remote Sensing*, 9, 1, 1-19. 2015.
- ⁶ CARDOSO-FERNANDES, J. et al., [Lithium \(Li\) pegmatite mapping using Artificial Neural Networks \(ANNs\): preliminary results](#), *IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium*, Waikoloa, HI, USA, 557-560. 2020.
- ⁷ SANTOS, D. et al., [Spectral Analysis to Improve Inputs to Random Forest and other Boosted Ensemble Tree-Based Algorithms for Detecting NYF Pegmatites in Tysfjord, Norway](#), *Remote Sensing*, 14, 15, 3532. 2022.
- ⁸ YU, L. et al., [Towards automatic lithological classification from remote sensing data using support vector machines](#), *Computers & Geosciences*, 45, 229-239. 2012.
- ⁹ VERIKAS, A. et al., [Electromyographic Patterns during Golf Swing: Activation Sequence Profiling and Prediction of Shot Effectiveness](#), *Sensors*, 16, 4, 592. 2016.