



X Conferência Nacional de Cartografia e Geodesia

Informação Geoespacial para os objetivos do desenvolvimento sustentável

02-03 Nov. 2023 Instituto Politécnico da Guarda

Inovação Agrícola Através da Detecção Remota

Jorge PEREIRA*¹, Ana MARGARIDA², Álvaro PINHEIRO¹, Tiago AFONSO¹, Machiel BOS¹, João NEVES² e André SÁ^{1,3}

¹ TeroMovigo – Earth Innovation, Portugal

² Universidade da Beira Interior, Portugal

³ Politécnico da Guarda, Portugal

(machiel.bos@teromovigo.com; ana.margarida.dias@ubi.pt; andre_sa@ipg.pt)

Palavras-chave: Detecção Remota, Agricultura de precisão, Inteligência Artificial, Deep Learning, Super Resolução, Viticultura

Resumo:

A produção de vinhos de qualidade produzidos de forma sustentável levou ao aumento da procura de estratégias baseadas em inovação digital para melhorar a gestão agrícola e resiliência às alterações climáticas. Neste âmbito, a deteção remota, através de sensores acoplados a Veículos Aéreos Não Tripulados (VANT) e Satélites tem ganho destaque devido à capacidade de avaliarem, embora a diferentes resoluções espaciais e espectrais, o estado nutricional das plantas, medir o stress hídrico entre outras informações relevantes e essenciais para a correta monitorização e avaliação das culturas. No entanto, ainda existem várias limitações associadas a estas tecnologias. Os VANT, incluindo os sensores e software de processamento apresentam elevados custos de aquisição, além da necessidade de operadores técnicos capacitados e dos custos fixos associados às especificidades de cada campanha. Em contraste, os dados de satélite (e.g., Sentinel-1 e Sentinel-2) são fornecidos gratuitamente e com revisitações periódicas a cada 5 dias pelas principais agências espaciais (e.g., ESA, NASA). No entanto, a baixa resolução espacial (10 m/px) destes dados tem restringido a sua aplicação em amplas áreas de cultivo (e.g., criação de mapas de distribuição de terrenos agrícolas), sendo raramente utilizados para tarefas de agricultura de precisão. O desafio é ainda mais pronunciado em culturas em linha, como as vinhas, devido à presença de entrelinhas (espaços entre as fileiras de vinha cultivada). Apesar desta limitação, além da possível captura deficiente devido a barreiras atmosféricas (nuvens), o uso de dados de satélite apresenta vantagens em termos de custo e escalabilidade, sendo possivelmente a solução tecnológica mais democrática no que refere ao acesso e uso deste tipo de soluções.

Este artigo apresenta uma estratégia para resolver esta limitação através do uso de métodos de super-resolução que permitem aumentar a resolução de imagens de muito baixa resolução. Assim, as imagens de alta resolução (3-5cm/px) obtidas por VANT e em combinação com as imagens obtidas por satélite (Sentinel-2) são usadas para treinar métodos que representam o estado-da-arte de super-resolução com vista a melhorar a resolução espacial de imagens de satélite de 10 m até 0.625 m (i.e., um fator de 16) e por sua vez a estimativa de parâmetros de fitossanidade das plantações agrícolas. Esta técnica está a ser implementada com dados obtidos numa vinha de 5,5 ha pertencente a um dos associados da Adega Cooperativa de Pinhel, localizada na cidade de Pinhel (distrito da Guarda), onde 7 voos já foram realizados, acompanhando as diferentes fases da vinha. Os resultados apresentados mostram que o desempenho atual das abordagens de super-resolução permitem obter uma melhoria considerável na resolução das imagens de satélite.



1. Introdução:

Os gestores de sistemas agrícolas enfrentam desafios importantes devido, entre outros aspectos, ao crescimento populacional que tem levado à necessidade de expansão das áreas de cultivo e ao aumento da produtividade agrícola para suprimir as crescentes necessidades alimentares globais. Além disso, é preocupante observar uma estagnação e até mesmo uma diminuição das áreas de cultivo em países com influência relevante na produção agrícola mundial.

Segundo a Organização das Nações Unidas para a Agricultura (Food and Agriculture Organization - FAO) durante o período compreendido entre 1990-2019, a Índia e a China verificaram uma estagnação parcial ou total do crescimento das áreas de cultivo, e alguns dos maiores produtores mundiais têm testemunhado uma diminuição dessas áreas de cultivo, como é o caso dos EUA (-15%), Rússia (-8%), Canadá (-6%) (FAO, 2021^a).

Neste seguimento, a própria FAO, juntamente com outros organismos mundiais dedicados à problemática agrícola, lançou um plano de ação para o período de 2022 até 2025 (FAO, 2023) com o objetivo de delinear uma estratégia de inovação científica para o horizonte de 2022 até 2030, em resposta às alterações climáticas e para a manutenção da biodiversidade em todos os sectores agrícolas com o envolvimento do sector privado. Este plano baseia-se em alguns pilares que incluem o reforço das tomadas de decisão com base em evidências e dados científicos e o apoio à tecnologia e inovação à escala regional e nacional. Em particular, a FAO pretende com estes pilares aumentar a produção dos produtores de pequena e média escala mantendo a área de cultivo, reforçar a resiliência climática e a segurança alimentar através de plataformas geoespaciais e de mapeamento do território (FAO, 2022).

Neste contexto, a TeroMovigo, uma empresa dedicada a mapeamento e deteção remota, procura criar soluções alinhadas com as propostas da FAO, especialmente para o sector vitivinícola, onde Portugal assume a décima primeira posição de maior produtor mundial de vinho (IVV, 2023).

Em parceria com a Adega Cooperativa de Pinhel e a Universidade da Beira Interior, a TeroMovigo iniciou um projeto de estudo baseado em agricultura de precisão propondo avaliar o estado nutricional das videiras, medir stress hídrico, entre outras informações relevantes para a monitorização e avaliação das culturas, com a ajuda da deteção remota através de sensores acoplados a Veículos Aéreos Não Tripulados (VANT) e Satélites.

Esta forma de monitorização já é uma realidade com alguns operadores a fornecerem soluções de agricultura de precisão através de deteção remota com satélites, VANT e até mesmo avião. No entanto, as aeronaves tripuladas e os VANT, incluindo os sensores e software de processamento apresentam elevados custos de aquisição, além da necessidade de operadores técnicos capacitados e dos custos fixos associados às especificidades de cada campanha de observação. Em contraste, os dados de satélite (e.g., Sentinel-1 e Sentinel-2) são fornecidos gratuitamente e com revisitações periódicas a cada 5 dias pelas principais agências espaciais (e.g., ESA, NASA). No entanto, a baixa resolução espacial (10 m/px) destes dados tem restringido a sua aplicação em amplas áreas de cultivo (e.g., criação de mapas de distribuição de terrenos agrícolas), sendo raramente utilizados para tarefas de agricultura de precisão. Neste aspecto, é ainda de referir que a aplicação destas tecnologias em culturas em linha, como as vinhas, é ainda incipiente devido à presença de entrelinhas (espaços não cultivados entre as fileiras de vinha) que causam de forma significativa o problema "mixed-pixel value". Apesar das limitações de resolução e da possível captura deficiente devido a barreiras atmosféricas (nuvens), o uso de dados de satélite apresenta vantagens em termos de custo e escalabilidade, sendo possivelmente a solução tecnológica mais democrática.

Este artigo apresenta uma estratégia para superar esta limitação através do uso de modelos de super-resolução, como o SwinIR (Liang et al., 2021), que permitem aumentar a resolução de imagens de muito baixa resolução. Assim, as imagens de alta resolução (3-5cm/px) obtidas por VANT, em combinação com as imagens obtidas por satélite (Sentinel-2), são usadas para treinar esse mesmo modelo de super-resolução, visando melhorar a resolução espacial das imagens de satélite de 10 m até 0.625 m e, conseqüentemente, a estimativa de parâmetros de fitossanidade das plantações agrícolas.

2. Metodologia

2.1 Área de Estudo

A área em estudo, situada em Pinhel, faz parte da região vitivinícola da Beira Interior (Figura 1), abrangendo a vinha que está sob monitorização e que ocupa uma área aproximada de 5,5 hectares. Esta vinha é reconhecida com a Denominação de Origem Protegida (DOP) "Beira Interior", uma certificação concedida pela Comissão Vitivinícola Regional da Beira Interior (CVRBI).

A monitorização ao longo da campanha anual (2022/2023) teve como principais objetivos criar ortofotomapas de alta resolução (3-5cm/px) da vinha através da captura de imagens por sensores acoplados a um VANT e obter para cada mesma data ou data equivalente (data mais próxima), os respetivos ortofotomapas de baixa resolução (10 m/px) capturados por Sentinel-2 com a finalidade de criar os "pares de imagens" que alimentam o modelo de super-resolução a testar, neste caso o modelo SwinIR.

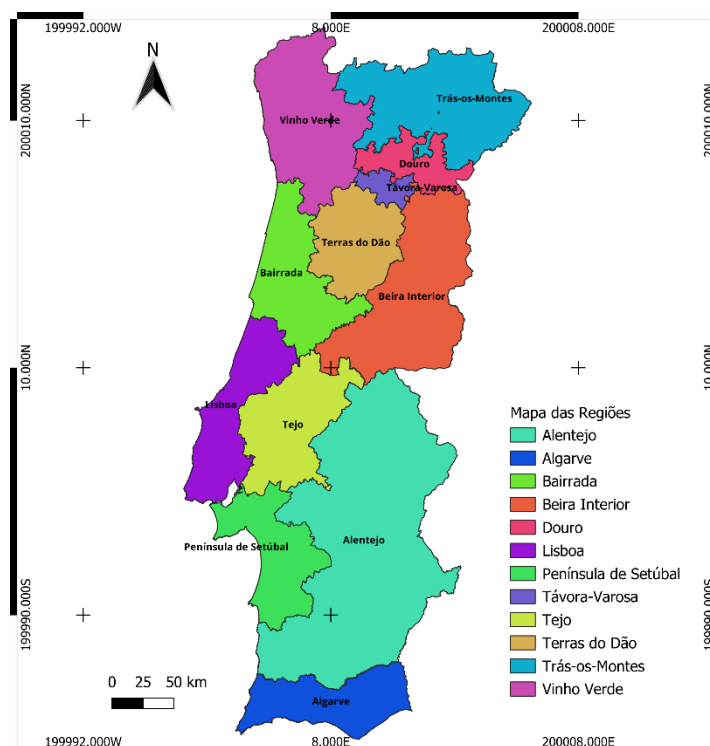


Figura 1 - Mapa das Denominações de Origem Protegida (DOP) identificando cada uma das diferentes regiões vitivinícolas de Portugal.

2.2 Criação de Ortofotomapas através de imagens capturadas por um VANT

Os ortofotomapas foram gerados através do software de processamento Agisoft PhotoScan® (Agisoft©, 2023) que processou aproximadamente 7182 imagens aéreas capturadas por uma câmara multiespectral Micasense© ALTUM-PT (Micasense, 2023) com capacidade de capturar imagens em 6 bandas distintas do espectro eletromagnético, embora neste estudo se faça apenas uso de 5, banda azul com centro de comprimento de onda 475 nanómetros (nm) e uma largura de 20 nm, banda verde com centro de comprimento de onda 560 nm e uma largura de 20 nm, banda vermelha com centro de comprimento de onda 668 nm e uma largura de 10 nm, banda rededge com centro de comprimento de onda 717 nm e uma largura de 10 nm e banda Near Infrared (NIR) com centro de comprimento de onda 840 nm e uma largura de 40 nm, para cada voo. O plano de voo elaborado no software proprietário da Quantum Systems©, Q-Base 3D© (Quantum Systems, 2023) foi realizado a uma altitude de voo de 120 metros, uma sobreposição lateral de 71% e uma sobreposição longitudinal de 66% que resultou numa distância de amostra do solo (GSD) de (5,17cm/px), Figura 2.



X Conferência Nacional de Cartografia e Geodesia

Informação Geoespacial para os objetivos do desenvolvimento sustentável

02-03 Nov. 2023 Instituto Politécnico da Guarda

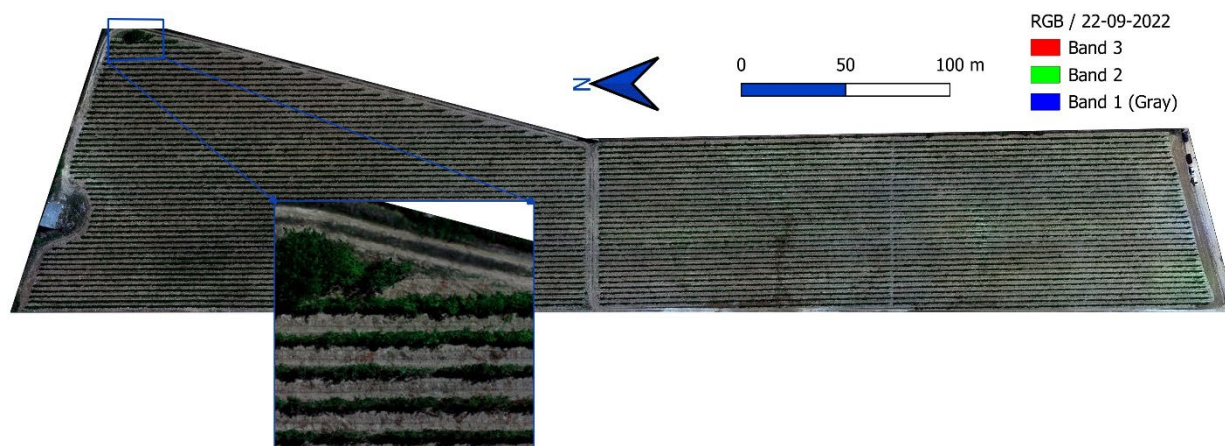


Figura 2 – Ortofotomapa da vinha em estudo com um GSD de aproximadamente (5cm/px), exibindo a combinação de bandas vermelha, verde e azul para a campanha de observação de 22-09-2022 (após a vindima da campanha anterior).

2.3 Obtenção de Ortofotomapas Sentinel-2

O Sentinel-2 lançado como parte do Programa Copérnicos da Agência Espacial Europeia (ESA) é uma constelação de dois satélites equipados com sensores multiespectrais que medem a radiância refletida pelo topo da atmosfera (TOA) em 13 bandas distintas e à imagem da câmara multiespectral Micasense utilizada no VANT, faremos apenas uso, para as imagens de satélite, das bandas azul com centro de comprimento de onda 443 nanómetros (nm) e uma largura de 28 nm, banda verde com centro de comprimento de onda 561 nm e uma largura de 47 nm, banda vermelha com centro de comprimento de onda 665 nm e uma largura de 48 nm, banda rededge1 com centro de comprimento de onda 705 nm e uma largura de 20 nm e banda Near Infrared (NIR) com centro de comprimento de onda 842 nm e uma largura de 116 nm.

O Sentinel-2 apresenta uma resolução espacial de (10m/px) para as bandas supracitadas excetuando a banda rededge1 que apresenta uma resolução de (20m/px). As imagens podem ser descarregadas através do seu site oficial (Copernicus, 2024).

Quadro 1 – Informação sobre as campanhas de observação (CdO) realizadas por VANT e as observações correspondentes do Sentinel-2 e respetiva diferença temporal entre os dois métodos de observação.

| Nome dos Dados | Data de Captura | Fonte do Dados | Diferença temporal (Dias) |
|----------------|-------------------|----------------|---------------------------|
| $VANT_{CdO1}$ | 22 Setembro 2022 | VANT | +1 |
| S_{CdO1} | 21 Setembro 2022 | Sentinel-2 | -1 |
| $VANT_{CdO2}$ | 03 Março 2023 | VANT | +3 |
| S_{CdO2} | 28 Fevereiro 2023 | Sentinel-2 | -3 |
| $VANT_{CdO3}$ | 04 Abril 2023 | VANT | 0 |
| S_{CdO3} | 04 Abril 2023 | Sentinel-2 | 0 |
| $VANT_{CdO4}$ | 16 Maio 2023 | VANT | +3 |
| S_{CdO4} | 19 Maio 2023 | Sentinel-2 | -3 |
| $VANT_{CdO5}$ | 10 Agosto 2023 | VANT | +2 |
| S_{CdO5} | 12 Agosto 2023 | Sentinel-2 | -2 |
| $VANT_{CdO6}$ | 17 Agosto 2023 | VANT | 0 |
| S_{CdO6} | 17 Agosto 2023 | Sentinel-2 | 0 |

Os produtos de nível-2A derivados dos produtos de nível-1, disponibilizados pelo Copernicus já contemplam imagens com correção atmosférica efetuada. Cada produto de nível-2A é composto por imagens com sobreposição lateral e longitudinal de $110 \times 110 \text{ km}^2$ no sistema de coordenadas projetadas (UTM/WGS84).

É possível para cada imagem selecionar através do desenho de polígonos apenas a área que se tenciona estudar, neste caso a área onde a vinha está inserida.

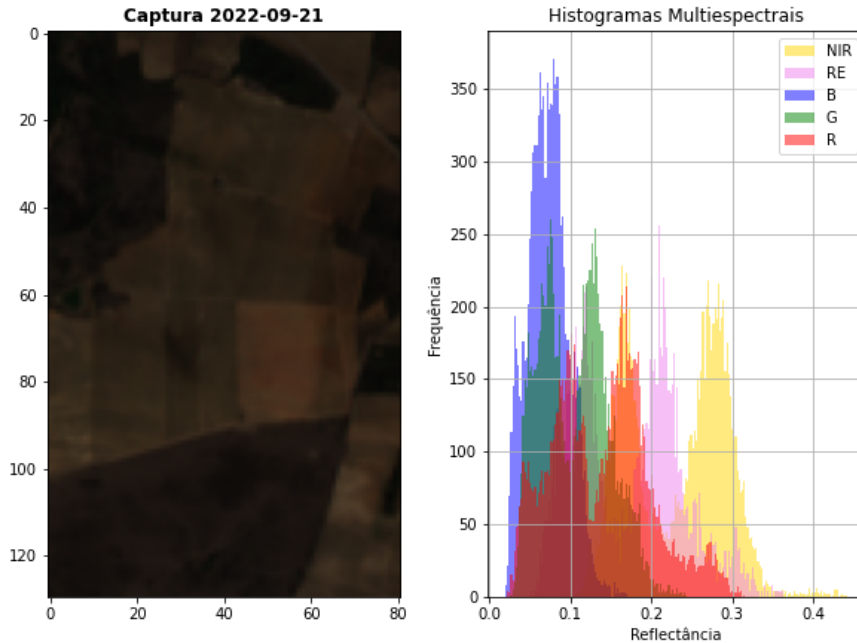


Figura 3 – Área de estudo onde está inserida a vinha a monitorizar e o respetivo histograma das bandas retiradas das imagens de nível-2A do Sentinel-2.

2.4 Treino do modelo de super-resolução SwinIR

O modelo SwinIR é baseado no Swin Transformer (Liang et al., 2021). A Figura 4 apresenta a arquitetura deste modelo em forma de diagrama.

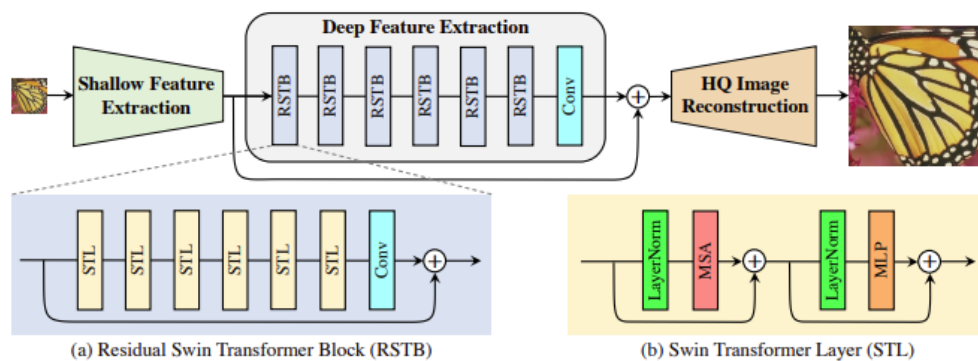


Figura 4 - Diagrama da arquitetura do modelo SwinIR (Liang, J et al, 2021)

As imagens por VANT (5cm/px) são reamostradas para (0.625m/px) e as imagens com a data correspondente do satélite Sentinel-2 (10m/px) são recortadas para obter um conjunto de várias imagens mais pequenas de $256 \times 256 \text{ px}$ de resolução, para os recortes das imagens reamostradas, e de $16 \times 16 \text{ px}$ para os recortes de baixa resolução.

Após este processo o modelo SwinIR recebe como entrada, cada par de imagens (reamostragem, baixa resolução) correspondentes para iniciar o treino necessário, o processo resulta numa imagem de saída a que chamamos de super-resolução.

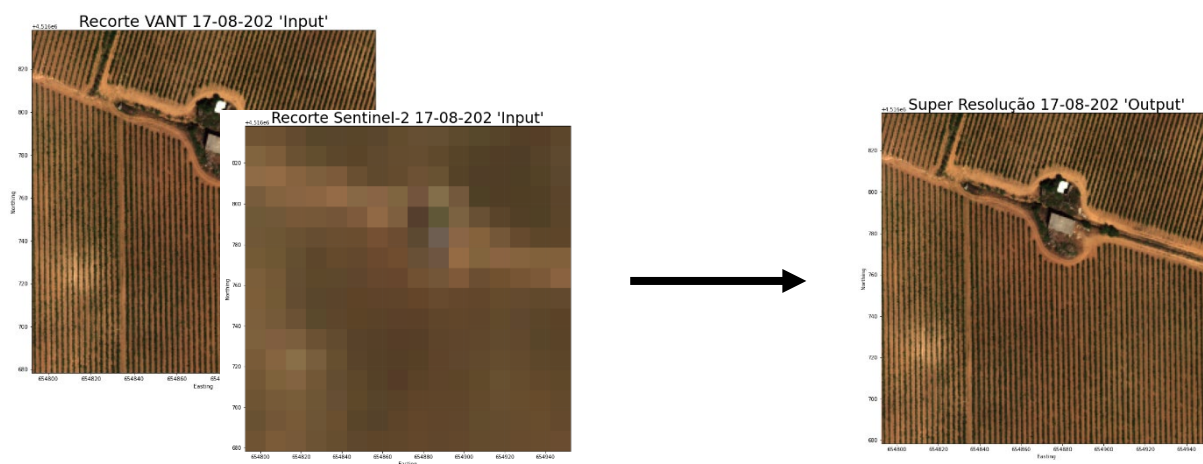


Figura 5 – Recortes de entrada do par correspondente de imagens (reamostragem, baixa resolução) e imagem de saída super-resolução, para a data de 17-08-2022

O modelo é constituído por três módulos principais: o módulo da extração de características com filtros definidos manualmente, o módulo de extração de características profundas e o módulo de reconstrução da imagem. Durante o treino o modelo otimiza os seus parâmetros minimizando a “L1 Pixel Loss”.

O primeiro módulo extrai uma característica de uma imagem de baixa resolução a partir de uma camada de convolução, transformando o espaço da imagem de entrada num espaço de características de maior dimensionalidade. O resultado é transmitido ao módulo de extração de características profundas para processamento adicional e para o módulo de reconstrução para garantir a preservação das informações de baixa frequência.

Na segunda fase, correspondente ao módulo de extração de características profundas, o objetivo geral é recuperar dados de alta frequência. Este módulo contém um conjunto de blocos “Residual Swin Transformer” (RSTB) e uma camada convolucional. Cada RSTB extrai uma característica intermediária, bloco por bloco, resultando na característica profunda final após a camada convolucional. O RSTB é principalmente composto por camadas “Swin Transformer” (STL), baseadas no mecanismo de atenção do Transformer original (Dosovitskiy et al. 2021), diferenciando-se pela implementação da atenção local e do mecanismo de translação de janela.

Em vez de dividir a imagem de entrada em blocos, considerando cada bloco um token para calcular a atenção, o Swin Transformer divide a imagem em janelas locais não sobrepostas. A atenção é então calculada independentemente para cada janela, considerando cada pixel como um token, introduzindo assim a atenção local. O uso da divisão regular e da divisão com o mecanismo de translação das janelas é alternado para possibilitar conexões entre janelas. A segunda divisão envolve a translação da característica por $(M/2, M/2)$ pixels antes da divisão, onde M representa o tamanho da janela.

No final, o módulo de reconstrução combina as características obtidas dos dois primeiros módulos para criar a representação de alta qualidade da imagem de entrada. Para o “upsampling” é utilizada uma camada de convolução sub-pixel, o Pixel Shuffle (W. Shi, 2016).

3. Discussão

Os resultados preliminares são neste caso apresentados sem quaisquer dados quantitativos ou qualitativos, pois à data da comunicação (X CNCG) os resultados provenientes do modelo SwinIR tinham apenas um par de dias.

Realizou-se, no entanto, uma combinação da banda vermelha e NIR, neste caso para calcular o índice de vegetação de diferença normalizada (NDVI) para cada pixel da imagem reamostrada (0.625m/px) e da imagem de super-resolução de saída (0.625m/px) correspondente, por forma a identificar visualmente se poderíamos à primeira vista estimar parâmetros de fitossanidade das zonas com correspondência entre ambas.



$$NDVI = \frac{NIR-VERMELHO}{NIR+VERMELHO} \quad (\text{Rouse et al., 1973})$$

Este é um passo para o desenvolvimento de uma solução que possa servir para democratizar o acesso e uso deste tipo de soluções tecnológicas por parte dos pequenos e médios produtores, uma vez que na sua génese, usa imagens de livre acesso como são as imagens de Sentinel-2, reduzindo consideravelmente os custos de operações como levantamentos por avião, ou até satélites de maior resolução e de maior custo. Esta solução permite afirmar que a geodesia pode desempenhar um papel importante até mesmo na produção de vinhos.

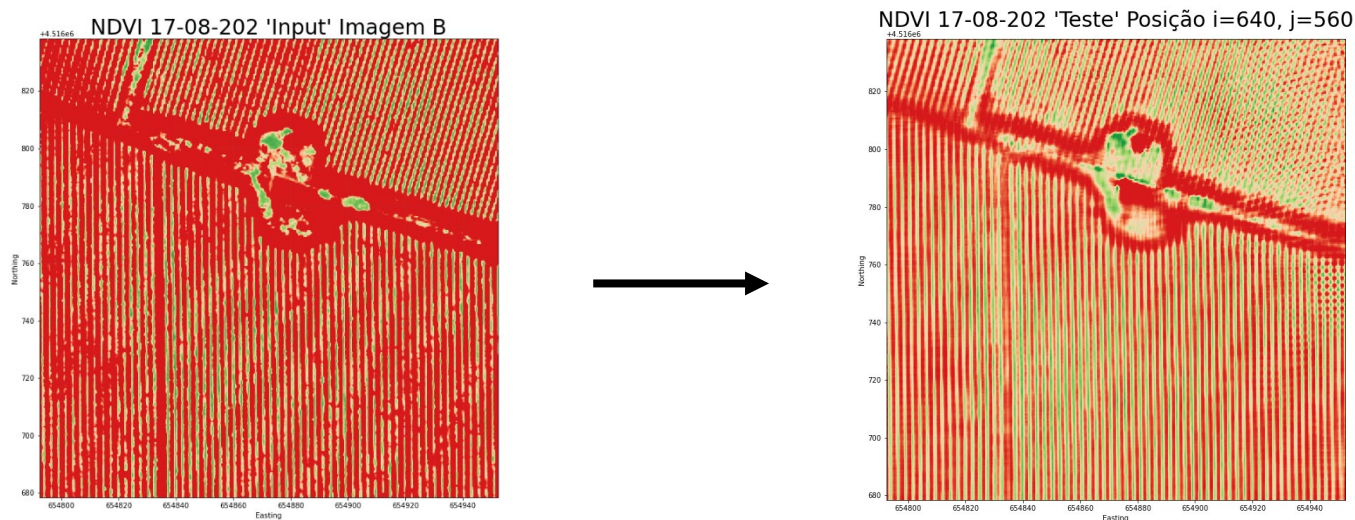


Figura 6 – Esquema que compara o recorte reamostrado (0.625m/px) com a correspondente imagem de super-resolução (0.625m/px), imagens compostas como combinação da banda Vermelha e NIR datadas do 17-08-2022.

3.1 Portal

A solução culminará num portal que possibilitará ao utilizador/agricultor entre outras funcionalidades, registar as suas vinhas através do desenho de polígonos, enquanto pode consultar o estado de fitossanidade das vinhas e com isso atempadamente tomar decisões, como a de alocar recursos de maneira mais eficiente, por forma a obter uma maior produtividade, evitando desperdícios de recursos onde não seriam inicialmente necessários. A figura seguinte (Figura 6) mostra o protótipo do que será a plataforma.

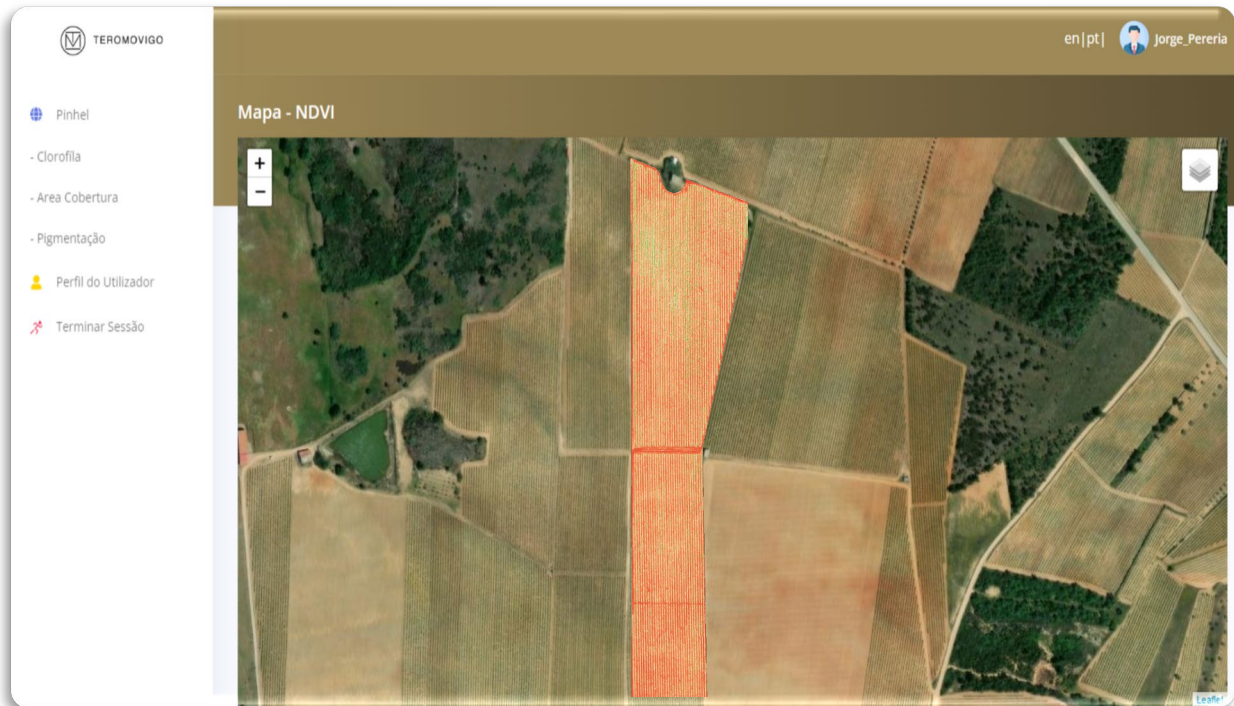


Figura 6 – Protótipo da plataforma desenvolvida pela TeroMovigo.

Referências Bibliográficas

FAO, 2021a. FAOSTAT: Land use. In: FAO.org [online]. <http://www.fao.org/faostat/en/#data/EL>, consultado em 01 de Março de 2024.

FAO, 2023. The Action Plan 2022–2025 for the implementation of the FAO Science and Innovation Strategy. Rome.

FAO, 2022. FAO Science and Innovation Strategy. Rome, <https://www.fao.org/3/cc2273en/cc2273en.pdf>, consultado em 01 de Março de 2024.

IVV, 2023. Apuramento da Produção de Vinho na Campanha 2023/2024, <https://www.ivv.gov.pt/> consultado em 01 de Março de 2024.

Micasense, 2023. Micasense Knowledge Base. <https://support.micasense.com/hc/en-us/articles/214878778-What-is-the-center-wavelength-and-bandwidth-of-each-filter-for-MicaSense-sensors> consultado em 01 de Março de 2024.

Quantum Systems, 2023. Mission Planning, Flight Monitoring, Data Post-Processing, <https://quantum-systems.com/qbase-3d/> consultado em 01 de Março de 2024.

Copernicus, 2024 Product Types. <https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi/product-types/level-2a> consultado em 01 de Março de 2024.

Copernicus, 2024 Copernicus Browser. <https://browser.dataspace.copernicus.eu/> consultado em 01 de Março de 2024.

Liang. J, Cao. J, Sun. G, Zhang. K, Van Gool. L, e Timofte. R. Swinir: Image restoration using swin transformer. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops, 2021, pp. 1833–1844.



X Conferência Nacional de Cartografia e Geodesia

Informação Geoespacial para os objetivos do desenvolvimento sustentável

02-03 Nov. 2023 Instituto Politécnico da Guarda

Dosovitskiy. A, Beyer. L, Kolesnikov. A, Weissenborn. D, Zhai. X, Unterthiner. T, Dehghani. M, Minderer. M, Heigold. G, Gelly. S, Uszkoreit. J, e Hounsby. N. An image is worth 16x16 words: Transformers for image recognition at scale. International Conference on Learning Representations, 2021.

Rouse. J.W, Haas. R.H, Schell. J.A, Deering. D.W. (1973): Monitoring vegetation systems in the great plains with ERTS. Third 80 ERTS Symposium, NASA SP-351, pp. 309-317.

Shi. W, Caballero. J, Huszar. F, Totz. J, Aitken. A. P, R. Bishop, Rueckert. D, e Wang. Z. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 1874–1883.