

UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE CIÊNCIAS DA TERRA E ENERGIA



**Super-resolução de imagens Sentinel-2 para melhoria da
classificação da ocupação do solo utilizando redes generativas
adversárias (GAN)**

Sara Daniela da Costa Pires

Mestrado em Engenharia Geoespacial

Trabalho de Projeto orientado por:
Prof. Doutor João Catalão Fernandes

Agradecimentos

A concretização desta dissertação marca o culminar de uma etapa significativa da minha vida académica e pessoal, e não teria sido possível sem o apoio, a dedicação e a inspiração de muitas pessoas a quem dirijo a minha mais profunda gratidão.

Em primeiro lugar, agradeço a todos os professores que me acompanharam ao longo do meu percurso. Cada um, à sua maneira, contribuiu para a construção do meu conhecimento e para o meu crescimento académico, transmitindo não só saber técnico, mas também valores de rigor e de resiliência.

De forma muito especial, expresso a minha sincera gratidão ao meu orientador, Professor João Catalão, pelo acompanhamento próximo, pela exigência científica e pela confiança que depositou em mim. O seu apoio constante, a sua disponibilidade e os seus ensinamentos foram determinantes para a realização deste trabalho.

Aos meus pais, Sara Pires e Carlos Pires, deixo o mais profundo agradecimento pelo amor incondicional, pela motivação diária e pela força transmitida nos momentos mais desafiantes. À minha avó, Cremilda Costa, que pensava sempre que já não estaria cá para me ver entrar na faculdade: a tua sardanisa agora é engenheira. Ao meu companheiro, Rui Castro, agradeço a paciência, a compreensão e o carinho, fundamentais para manter o equilíbrio e a determinação ao longo desta jornada, quando crescer quero ser como tu.

Quero ainda agradecer a todos os meus amigos e colegas de curso, com quem partilhei momentos que tornaram este percurso mais leve e enriquecedor. Um agradecimento especial aos meus melhores amigos, que eles sabem quem são, somos todos uma família (de adotados). Obrigada por alinharem em todas as festas e convívios comigo, tornaram este percurso muito mais leve.

Não poderia deixar de mencionar os meus animais, que encheram os meus dias de alegria e ternura, acompanhando-me em cada etapa deste caminho. Obrigada por me distraírem.

Por fim, dedico esta dissertação à minha querida cadela Daisy, que partiu no último verão. A sua companhia e alegria foram únicas e marcaram profundamente a minha vida. Este trabalho é, também, para ela.

Resumo

A resolução espacial de 10 metros das imagens Sentinel-2 pode ser um fator limitativo na análise detalhada da ocupação e uso do solo. Esta dissertação investiga a aplicação de técnicas de super-resolução baseadas em redes generativas adversárias (GAN) para melhorar a qualidade espacial de imagens Sentinel-2 e avaliar o impacto na classificação da ocupação do solo. Um conjunto de Ortoimagens de altíssima resolução (25 cm) foi utilizado como referência para treinar um modelo GAN capaz de produzir imagens de super-resolução a 2,5 m de resolução, correspondendo a um aumento de detalhe de 4×. A avaliação dos resultados incluiu métricas quantitativas de qualidade de imagem (Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), Mean Squared Error (MSE), Mean Absolute Error (MAE)) e comparações visuais face às imagens originais. Adicionalmente, a influência das imagens de super-resolução na classificação temática foi avaliada utilizando um classificador Random Forest, comparando-se o desempenho na discriminação de classes de ocupação do solo urbano (edifícios, vias, vegetação e água) entre imagens originais e de super-resolução. Os resultados demonstram melhorias expressivas na qualidade das imagens: o produto de super-resolução atingiu valores de PSNR aproximadamente 3 dB superiores e SSIM cerca de três vezes maiores do que nas imagens originais, aproximando-se significativamente da referência de alta resolução. Visualmente, os resultados apresentam maior nível de detalhe e clareza na identificação de estruturas urbanas. Na classificação da ocupação do solo, observaram-se ganhos de desempenho com o uso de imagens de super-resolução, traduzidos em aumentos nos valores de F1 em classes críticas como edifícios e vias, aproximando a exatidão de classificação à obtida com dados de resolução muito elevada. Este trabalho evidencia o potencial das redes generativas adversárias para superar as limitações de resolução das imagens de detecção remota de acesso livre, contribuindo para melhorias na cartografia urbana e na monitorização territorial.

Palavras-chave: Super-resolução, Sentinel-2, Redes Generativas Adversárias, classificação da ocupação do solo, detecção remota.

Abstract

The 10-meter spatial resolution of Sentinel-2 images can be a limiting factor in the detailed analysis of land use and land cover. This dissertation examines the application of super-resolution techniques based on Generative Adversarial Networks (GANs) to improve the spatial resolution of Sentinel-2 images and to evaluate their impact on land cover classification. A set of very high-resolution orthophotos (25 cm) was used as reference to train a GAN model capable of producing super-resolved images at 2.5 m resolution, corresponding to a fourfold increase in detail. The evaluation of the results included quantitative image quality metrics (Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), Mean Squared Error (MSE), Mean Absolute Error (MAE)) and visual comparisons with the original images. In addition, the influence of the super-resolved images on thematic classification was assessed using a Random Forest classifier, comparing performance in discriminating against urban land cover classes (buildings, roads, vegetation, and water) between the original and super-resolved images. The results demonstrate substantial improvements in image quality: the super-resolved product achieved PSNR values approximately 3 dB higher and SSIM values nearly three times greater than those of the original images, closely approaching the very high-resolution reference. Visually, the results exhibit a higher level of detail and greater clarity in the identification of urban structures. In land cover classification, performance gains were observed with the use of super-resolved images, reflected in increases in F1 values for critical classes such as buildings and roads, bringing classification accuracy closer to that obtained with very high-resolution data. This work highlights the potential of Generative Adversarial Networks to overcome the resolution limitations of freely available remote sensing imagery, contributing to improvements in urban mapping and territorial monitoring.

Keywords: Image super-resolution, Sentinel-2, Generative Adversarial Networks, land cover classification, remote sensing.

Índice

1	Introdução	1
1.1	ENQUADRAMENTO GERAL.....	1
1.2	MOTIVAÇÃO E OBJETIVOS	2
1.3	ESTRUTURA DO TRABALHO.....	3
2	Estado da Arte.....	4
2.1	SUPER-RESOLUÇÃO DE IMAGENS	4
2.2	APRENDIZAGEM AUTOMÁTICA E PROFUNDA	5
2.3	REDES NEURONAIS E ARQUITETURAS DE PROCESSAMENTO DE IMAGEM.....	6
2.3.1	<i>Arquitetura das Redes Neurais</i>	6
2.3.2	<i>Processo de Aprendizagem das Redes Neurais</i>	7
2.3.3	<i>Classificação temática com Random Forest (RF)</i>	8
2.4	REDES GENERATIVAS ADVERSÁRIAS (GANs).....	9
2.4.1	<i>Super-Resolution GAN (SRGAN)</i>	10
2.4.2	<i>Enhanced Super-Resolution GAN</i>	11
2.4.3	<i>Limitações e Desafios</i>	12
2.5	AVALIAÇÃO DA QUALIDADE DAS IMAGENS DE SUPER-RESOLUÇÃO	12
2.5.1	<i>Métricas Quantitativas</i>	12
2.5.2	<i>Avaliação Perceptual e Novas Métricas</i>	13
2.6	SÍNTESE DO CAPÍTULO.....	14
3	Dados e Métodos	16
3.1	ÁREA DE ESTUDO	16
3.2	DADOS SENTINEL-2.....	17
3.3	ORTOIMAGENS	18
3.4	METODOLOGIA	19
3.4.1	<i>Arquitetura da Rede</i>	20
3.4.2	<i>Tratamento e Pré-Processamento dos Dados</i>	21
3.4.3	<i>Dados de Treino e Validação (Geração de Tiles)</i>	23
3.4.4	<i>Treino e Predição da Rede</i>	24
3.4.5	<i>Métricas de Avaliação</i>	24
3.4.6	<i>Avaliação Temática</i>	25
3.4.7	<i>Testes Realizados</i>	26
4	Resultados e Discussão	28
4.1	QUALIDADE VISUAL	28
4.2	LEITURA COM POLÍGONOS DE REFERÊNCIA.....	33
4.3	AVALIAÇÃO TEMÁTICA COM RANDOM FOREST	37
4.4	MÉTRICAS DE AVALIAÇÃO DE RECONSTRUÇÃO.....	46
4.5	DISCUSSÃO INTEGRADA POR OBJETIVOS.....	48
	Conclusão	52
	Referência Bibliográfica	55

Índice de Figuras

Figura 2.1 Redes Neurais Feedforward Neural Network vs. Recorrentes. Reproduzido de Fakhfakh et. al (2017).....	7
Figura 2.2 Workflow do Random Forest. Reproduzido de Batool <i>et. al.</i> (2017).....	8
Figura 2.3 Arquitetura geral do modelo GAN utilizado para super-resolução. Reproduzido de Mäkinen, O. (2020).....	10
Figura 3.1 Área de estudo (Lisboa e Cascais). Extensão dos mosaicos de ortoimagens usados (2023) e limites administrativos municipais.....	16
Figura 3.2 Cobertura espectral do MSI do Sentinel-2 e resoluções espaciais (10, 20 e 60 m). Reproduzido de Farris et al. (2022).	17
Figura 3.3 Alcântara (frente ribeirinha, Hospital CUF Tejo). Excerto comparativo ortoimagem 2018 (a) vs ortoimagem 2023 (b).....	19
Figura 3.4 Workflow do estudo.	20
Figura 3.5 Grelhas equivalentes LR/HR no mesmo tile. Janela com grelha LR 32×32 px a 10 m (Sentinel-2 TCI) (a); e grelha HR 128×128 px a 2,5 m (b).....	22
Figura 3.6 Infográfico dos cenários experimentais (overview).	27
Figura 4.1 Painel comparativo (a) LR 2023 → (b) HR 2023 → (c) SR (visão geral).	28
Figura 4.2 “Zooms” temáticos por classe (small multiples). (a) Edifícios; (b) Vias; (c) Interface água-terra.	29
Figura 4.3 Repetição do painel em contextos distintos, Alameda em Lisboa. (a) Validação Interna (Lisboa); (b) Lisboa→Cascais; (c) 2018→2023.....	30
Figura 4.4 Casos-limite e artefactos localizados.....	31
Figura 4.5 Comparativo contextual Lisboa vs. Cascais (água-terra). (a) Lisboa→Cascais; (b) Cascais→Cascais.	32
Figura 4.6 Par temporal controlado (2018 ↔ 2023) no Empreendimento SkyCity, Amadora. (a) 2018; (b) 2023.....	33
Figura 4.7 Enquadramento da subárea: Braço de Prata / PRO, 2023.	34
Figura 4.8 Painel LR → HR → SR com polígonos sobrepostos (visão geral).....	34
Figura 4.9 Repetição controlada nas três abordagens. (a) Validação interna; (b) Cascais→Lisboa; (c) 2018→2023.....	36
Figura 4.10 "Zooms" dos mapas RF nos diferentes cenários com ênfase no edificado.....	38
Figura 4.11 "Zooms" dos mapas RF nos diferentes cenários com ênfase nas vias.....	39
Figura 4.12 "Zooms" dos mapas RF nos diferentes cenários com ênfase na vegetação.....	40
Figura 4.13 “Zooms” dos mapas RF nos diferentes cenários com ênfase na água.	41
Figura 4.14 Gráfico de ganhos F1-score por classe vs LR.	44
Figura 4.15 Ganhos Δ (PSNR/SSIM) - SR vs LR.	47
Figura 4.16 Ganhos Δ (MSE/MAE) - SR vs LR.....	48

Figura 4.17 Comparação entre a imagem original em baixa resolução (LR) e a imagem de super-resolução (SR) no cenário de validação interna.....	50
Figura 4.18 Comparação visual entre a imagem HR e a imagem SR, apresentada em recortes localizados na zona do Estoril (a) e do Terreiro do Paço (b) (cenário de validação interna).....	51

Índice de Tabelas

Tabela 3.1 Esquema de tiling e partições (treino vs. validação).....	23
Tabela 4.1 Legenda das classes: edifícios (rosa), vias (laranja), espaços verdes (verde) e água (azul).	35
Tabela 4.2 Inventário de amostras por classe.	37
Tabela 4.3 Legenda das classes: edifícios (vermelho), vias (cinzento), espaços verdes (verde) e água (azul).	38
Tabela 4.4 Exatidão de LR, de HR e de SR nas três abordagens. (a) LR; (b) HR; (c) SR: Validação interna; (d) SR: Lisboa→Cascais; (e) SR: Cascais→Lisboa; (f) SR: 2018→2023.	42
Tabela 4.5 Ganhos F1-score por classe vs LR.	44
Tabela 4.6 Valores das métricas em todos os cenários.	46
Tabela 4.7 Ganhos Δ de SR face a LR por cenário e métrica.	47

Lista de Abreviaturas

ANN	Artificial Neural Network (Rede Neuronal Artificial)
CNN	Convolutional Neural Network (Rede Neuronal Convolutacional)
DL	Deep Learning (Aprendizagem Profunda)
DNN	Deep Neural Network (Rede Neuronal Profunda)
ESRGAN	Enhanced Super-Resolution Generative Adversarial Network
FFNN	Feedforward Neural Network (Rede Neuronal Feedforward)
GAN	Generative Adversarial Network (Rede Adversária Generativa)
GPU	Graphics Processing Unit (Unidade de Processamento Gráfico)
HR	High Resolution (Alta Resolução)
IA	Inteligência Artificial
LR	Low Resolution (Baixa Resolução)
MAE	Mean Absolute Error (Erro Absoluto Médio)
ML	Machine Learning (Aprendizagem Automática)
MSE	Mean Squared Error (Erro Quadrático Médio)
MSI	Multispectral Instrument (Instrumento Multiespectral - Sentinel-2)
PSNR	Peak Signal-to-Noise Ratio (Rácio Sinal-Ruído de Pico)
RF	Random Forest
RGB	Red, Green, Blue (Vermelho, Verde, Azul)
RNN	Recurrent Neural Network (Rede Neuronal Recorrente)
RRDB	Residual-in-Residual Dense Block
SRCNN	Super Resolution Convolutional Neural Network
SR	Super-Resolution (Super-Resolução)
SRGAN	Super-Resolution Generative Adversarial Network
SRResNet	Super-Resolution Residual Network
SSIM	Structural Similarity Index Measure (Índice de Similaridade Estrutural)
TCI	True Color Image (Imagem a Cores Verdadeiras)
VGG	Visual Geometry Group (rede convolutacional desenvolvida pela Univ. de Oxford)
VIS	Visible (Bandas Visíveis do Espectro Eletromagnético)

1 Introdução

1.1 Enquadramento Geral

A deteção remota por satélite tem-se afirmado como uma técnica indispensável para a monitorização e análise da superfície terrestre, permitindo uma observação sistemática, precisa e de larga escala do planeta. Com o avanço contínuo das tecnologias espaciais e o aumento da disponibilidade de dados multiespectrais e hiperespectrais, com elevada granularidade temporal e espacial, é possível obter informação detalhada sobre fenómenos naturais e antropogénicos, contribuindo significativamente para a investigação e disponibilização de serviços numa vasta gama e áreas do conhecimento como a cartografia topográfica, o ordenamento do território, a gestão ambiental, a agricultura de precisão, o estudo das alterações climáticas ou a resposta a desastres naturais (Chuvieco & Huete, 2010).

A qualidade da informação obtida a partir de satélites depende, em grande medida, da resolução espacial das imagens, a qual determina o nível de detalhe com que os objetos na superfície terrestre podem ser observados (Campbell & Wynne, 2011). Nos últimos anos, a evolução tecnológica dos sensores a bordo dos satélites resultou na melhoria significativa da resolução espacial das imagens captadas, de que são exemplo os satélites WorldView-3 e WorldView-4 com 8 bandas espectrais e uma resolução espacial de 0.32 m na banda do pancromático e 1.2m nas bandas espectrais do visível (Jensen, 2015). No entanto, o elevado custo associado a estas imagens de altíssima resolução (VHR) pode constituir uma limitação à sua utilização em estudos que requerem a cobertura de grandes áreas geográficas ou análises multi-temporais frequentes (Colomina & Molina, 2014).

Em contrapartida, as missões de agências espaciais nacionais, como a missão Landsat da NASA (National Aeronautics and Space Administration) ou a missão Sentinel-2 da ESA (European Space Agency), fornecem dados de acesso livre com qualidade espacial adequada para a grande maioria das aplicações de monitorização ambiental. A missão Copernicus Sentinel-2, por exemplo, disponibiliza imagens multiespectrais em 13 bandas, abrangendo desde o visível até ao infravermelho de ondas curtas, com uma resolução espacial de 10 metros nas bandas do visível (Drusch et al., 2012). No entanto, em situações que requerem um nível de detalhe superior, a resolução espacial de 10 metros das imagens Sentinel-2 pode não ser suficiente para a extração precisa de informações, como a identificação de pequenos objetos ou a classificação detalhada do uso do solo (Li & Roy, 2017).

Para mitigar essas limitações, a técnica da super-resolução (Super Resolution, SR), técnica que visa aumentar artificialmente a resolução de uma imagem, tem sido amplamente usada como abordagem eficaz para aumentar digitalmente a resolução espacial de imagens de satélite (Haut et al., 2018). As técnicas de SR incluem métodos tradicionais, como a interpolação espectral e as abordagens mais recentes baseadas em aprendizagem profunda, como Redes Neurais Convolucionais (CNNs) e as Redes Generativas Adversárias (GANs) (Dong et al., 2016; Ledig et al., 2017). As GANs, em particular, têm demonstrado um elevado potencial para a super-resolução de imagens de satélite, proporcionando melhorias significativas na qualidade espacial das imagens e preservando simultaneamente as características espectrais originais (Haut et al., 2019).

1.2 Motivação e objetivos

A crescente procura de imagens de satélite de alta resolução é impulsionada pela necessidade de informação atualizada e detalhada para diversas aplicações, como a atualização cartográfica, a monitorização ambiental ou a resposta a desastres naturais (Mishra et al., 2017). No entanto, a disponibilidade limitada de imagens de alta resolução de acesso público e o custo elevado das imagens comerciais representam um desafio significativo para a realização de estudos abrangentes e de longo prazo, especialmente em regiões menos favorecidas (Zhu et al., 2017).

Neste contexto, a super-resolução tem vindo a ganhar destaque como uma solução promissora para mitigar estas limitações (Aiazzi et al., 2007). A aplicação destas técnicas pode preencher a lacuna entre a necessidade de dados detalhados e a disponibilidade limitada de imagens de alta resolução, oferecendo uma solução economicamente viável para diversas aplicações (Lanaras et al., 2018). Entre as várias técnicas de super-resolução, as redes generativas adversárias destacam-se pelo seu potencial em gerar imagens com elevado realismo e fidelidade espacial. A aplicação de GANs à deteção remota representa, assim, um campo emergente e altamente relevante, com potencial para transformar a forma como interpretamos e utilizamos dados de satélite.

O principal objetivo desta tese é investigar, implementar e avaliar algoritmos de super-resolução baseados em redes generativas adversárias aplicados a imagens de satélite. Pretende-se analisar o desempenho desses modelos num contexto geográfico urbano, com vista a melhorar a qualidade das imagens e, conseqüentemente, aumentar a eficácia na atualização cartográfica de zonas urbanas. A abordagem proposta explora a sinergia espectral entre as missões Sentinel-2 e ortoimagens produzidas com imagens aéreas (com 0,25 m de resolução espacial) para gerar imagens com uma resolução espacial superior a 10 metros, recorrendo a métodos de aprendizagem profunda para a reconstrução de imagens super-resolução (Ma et al., 2019).

Os objetivos específicos deste estudo são:

1. Desenvolver e avaliar um modelo de super-resolução baseado em GAN para melhorar a resolução espacial das imagens Sentinel-2 para 2,5 metros, preservando a consistência espectral das imagens originais.
2. Validar a eficácia do modelo proposto em diferentes cenários urbanos, comparando os resultados obtidos com imagens de referência de alta resolução.
3. Analisar o impacto da melhoria da resolução espacial na classificação da ocupação do solo, avaliando a precisão das classificações antes e depois da aplicação da super-resolução.

Desta forma, espera-se que este estudo contribua para o avanço das técnicas de processamento de imagens de satélite, melhorando a qualidade dos dados disponíveis e ampliando as possibilidades de utilização das imagens Sentinel-2 em aplicações onde a informação detalhada é crucial, como a atualização cartográfica e o planeamento urbano (Haut et al., 2018).

1.3 Estrutura do trabalho

O presente trabalho está organizado em cinco capítulos, articulando enquadramento, revisão crítica, dados e métodos, resultados e síntese final, de forma a responder aos objetivos definidos para a melhoria cartográfica de imagens Sentinel-2 por super-resolução. O Capítulo 1 apresenta o contexto científico e aplicado da dissertação, a motivação e os objetivos, culminando nesta descrição da organização do documento. O Capítulo 2 corresponde ao estado da arte e estrutura-se em torno de três eixos: fundamentos e evolução da super-resolução em detecção remota; noções essenciais de aprendizagem automática e profunda e suas arquiteturas para processamento de imagem; e redes generativas adversárias, com destaque para SRGAN (Super-Resolution Generative Adversarial Network) e ESRGAN (Enhanced Super-Resolution Generative Adversarial Network), bem como limitações e desafios. Inclui ainda a discussão das abordagens de avaliação da qualidade de imagens de super-resolução, quer por métricas de distorção, quer por critérios perceptuais, e termina com uma síntese que relaciona estas métricas com o problema cartográfico estudado.

O Capítulo 3 descreve os dados e a metodologia. Caracteriza a área de estudo (Lisboa e Cascais), documenta as fontes e especificações das imagens Sentinel-2 e das ortoimagens, e detalha o pipeline metodológico adotado: arquitetura de rede baseada no repositório Satlas-SR, tratamento e pré-processamento, geração de tiles emparelhados LR–HR, método de treino do modelo, métricas de avaliação e desenho das experiências. Inclui, ainda, o protocolo de avaliação temática que sustenta a análise de utilidade cartográfica, com classificação Random Forest aplicada de forma comparável aos domínios LR, SR e HR, bem como a definição das três abordagens realizadas (verificação interna, transferência Lisboa↔Cascais e robustez temporal 2018→2023).

O Capítulo 4 apresenta e discute os resultados. A leitura visual co-registada evidencia os ganhos de nitidez e continuidade; a leitura com polígonos de referência na subárea Braço de Prata/Parque Ribeirinho Oriente demonstra o impacto direto na vetorização; a avaliação temática com Random Forest quantifica diferenças por classe e cenário; e os registos de treino do modelo GAN são analisados para contextualizar a estabilidade e o equilíbrio adversarial. A discussão integrada organiza-se por objetivos, sintetizando implicações e limitações. O Capítulo 5 reúne as conclusões principais, responde às questões de investigação, identifica limitações e propõe linhas de trabalho futuro.

2 Estado da Arte

A super-resolução de imagens tem ganho relevância em múltiplos domínios, em particular na detecção remota, pelo seu potencial para ultrapassar as limitações de resolução espacial dos sensores multiespectrais a bordo de satélites e, assim, aumentar a utilidade cartográfica dos produtos derivados. Este capítulo apresenta uma revisão crítica e sintética da literatura, cobrindo conceitos fundamentais, avanços metodológicos e desafios na super-resolução aplicada a dados Sentinel-2, com ênfase nas Redes Generativas Adversárias (GANs) enquanto abordagem de referência para recuperar detalhe espacial preservando coerência radiométrica.

O capítulo organiza-se em secções que: i) introduzem o conceito de super-resolução e a sua relevância em detecção remota, incluindo métodos clássicos e abordagens de aprendizagem profunda; ii) enquadram as redes neurais e as arquiteturas usadas no processamento de imagem; iii) discutem, de forma focalizada, as GANs, com destaque para SRGAN e ESRGAN, e as suas limitações; e iv) apresentam os critérios de avaliação da qualidade em super-resolução, distinguindo métricas de distorção de abordagens percutuais. Em conjunto, estas secções estabelecem a base científica necessária para os capítulos seguintes, onde a arquitetura adversarial é aplicada a Sentinel-2 e avaliada quanto ao valor cartográfico em contexto urbano.

2.1 Super-resolução de Imagens

A super-resolução de imagens é uma técnica computacional que visa melhorar a qualidade visual e a resolução espacial de imagens de baixa qualidade, criando versões de alta resolução a partir de uma ou mais imagens de entrada. Essa técnica é amplamente aplicada em diversas áreas, como processamento de imagens médicas, vigilância, restauração de vídeos e, mais recentemente, detecção remota espacial / observação da terra (Yang et al., 2019; Wang et al., 2020).

Super-resolução é frequentemente definida como o processo de reconstrução de informação de alta frequência que foi perdida devido a limitações do sensor ou durante o processamento da imagem. No domínio de detecção remota, a importância da super-resolução está associada ao aumento da precisão em tarefas como a classificação da ocupação do solo, análise de mudanças ambientais e monitorização de desastres naturais (Dong et al., 2016). A capacidade de obter imagens com maior nível de detalhe espacial é crucial para aplicações que dependem de dados visuais refinados.

Recentemente, as técnicas de super-resolução evoluíram significativamente, podendo ser classificadas em duas categorias principais. A primeira reúne os métodos baseados em interpolação matemática, como a interpolação bicúbica ou bilinear. Estes métodos criam novos dados espectrais com base em píxeis adjacentes através de operações matemáticas simples. Embora sejam fáceis de implementar, tais abordagens apresentam limitações claras, frequentemente resultando em imagens com artefactos e falta de realismo (Kim et al., 2008).

A segunda categoria reúne as redes neurais e a aprendizagem profunda, com métodos robustos e de elevado realismo, como as SRCNN (Super Resolution Convolutional Neural Networks) e as GANs. Estas técnicas inovadoras destacam-se pela capacidade de aprender padrões complexos diretamente a partir de grandes volumes de dados de treino, o que permite gerar imagens de alta resolução com maior realismo e detalhes mais precisos. Estudos realizados por Ledig et al. (2017) e Goodfellow et al. (2014) demonstram a eficácia destas abordagens, que redefiniram os limites do que é possível alcançar em termos de qualidade visual e precisão espacial em imagens com baixa resolução.

Apesar dos avanços significativos, a aplicação da super-resolução em imagens de detecção remota enfrenta diversos desafios que limitam a sua eficácia e aplicação. Um dos principais desafios é a

complexidade espectral das imagens multi-espectrais, que exige métodos robustos capazes de preservar a informação específica de cada banda, mantendo simultaneamente a coerência espacial e espectral dos dados processados. Zhu et al. (2018) destacam que a incapacidade de preservar essas características pode comprometer a qualidade final das imagens de super-resolução, afetando negativamente as análises subsequentes.

Outro desafio relevante é a necessidade de conjuntos de dados de treino diversificados e espacialmente bem registados (isto é, com correspondência precisa entre as imagens de baixa e alta resolução). A criação de modelos eficazes de super-resolução depende de dados de alta qualidade que capturem diferentes condições ambientais e geográficas. No entanto, a obtenção de tais dados é dificultada por fatores como a disponibilidade limitada de pares de imagens de alta e baixa resolução com alinhamento perfeito, além de problemas relacionados ao ruído e à inconsistência temporal nas aquisições. Um outro desafio atual consiste na integração multissensorial, em que os dados do Sentinel-2 podem ser combinados com dados de outros sensores, como os dados SAR do Sentinel-1. Esta abordagem tem o potencial de melhorar significativamente a resolução e a robustez dos resultados, conforme discutido por Zhu et al. (2019).

Outro foco importante está na validação da qualidade das imagens de super-resolução geradas. A combinação de métricas quantitativas como PSNR e SSIM com avaliações qualitativas realizadas por especialistas em detecção remota é fundamental para garantir a precisão e a utilidade prática dos modelos. Wang et al. (2019) enfatizam que essa abordagem mista oferece uma visão mais abrangente sobre a eficácia das tecnologias de super-resolução, promovendo maior confiança nos resultados obtidos.

Adicionalmente, o custo computacional elevado é um fator limitante significativo. Os modelos avançados, especialmente os baseados em GANs, requerem grandes volumes de dados e processamento intensivo em hardware especializado, como GPUs de alto desempenho. Este requisito não apenas aumenta o custo financeiro associado, mas também pode introduzir instabilidades no processo de treino, como observado por Isola et al. (2017). Estes desafios enfatizam a necessidade de mais investigação para tornar a super-resolução mais acessível e eficiente no contexto de detecção remota.

2.2 Aprendizagem Automática e Profunda

A Inteligência Artificial (IA) tem desempenhado um papel transformador na ciência e tecnologia, impulsionando avanços em diversas áreas do conhecimento. No contexto do processamento de imagens e, em particular, da super-resolução, a IA tem-se revelado essencial através de algoritmos avançados de aprendizagem automática (ML) (Machine Learning) e aprendizagem profunda (DL) (Deep Learning).

A conceção de sistemas artificiais capazes de reproduzir aspetos da inteligência humana remonta a tempos antigos, figurando em mitos e relatos. No entanto, a formalização moderna da IA surgiu apenas com o desenvolvimento dos primeiros computadores digitais programáveis na década de 1940. O termo “Inteligência Artificial” foi introduzido em 1956 durante o Dartmouth Workshop, evento organizado por Marvin Minsky e John McCarthy, onde se consolidaram as bases desta disciplina (McCarthy et al., 2006).

A evolução da IA foi marcada por ciclos de avanço e retração, conhecidos como “verões” e “invernos” da IA. Os períodos de elevado entusiasmo foram seguidos por fases de estagnação devido a limitações computacionais e a expectativas irrealistas. No entanto, nas últimas décadas, a ampliação da capacidade de processamento e a disponibilidade massiva de dados revolucionaram o campo, surgindo aplicações avançadas na área das imagens e em particular na detecção remota (Russel & Norvig, 2020).

A aprendizagem automática constitui um subcampo da IA que se dedica à conceção de algoritmos que permitem a extração de padrões a partir de dados. Segundo Mitchell (1997), “um programa de computador diz-se que aprende a partir da experiência E, relativamente a uma classe de tarefas T e uma métrica de desempenho P, se o seu desempenho em tarefas de T, medido por P, melhora com a experiência E”.

Os modelos de aprendizagem automática são construídos através de um pipeline estruturado que inclui a recolha, tratamento e divisão de dados em conjuntos de treino, validação e teste. Este processo visa minimizar enviesamentos e assegurar a capacidade do modelo de generalizar para novos dados. Durante o treino, o modelo ajusta os seus parâmetros internos para otimizar o desempenho na tarefa definida, mitigando problemas como o sobreajustamento (*overfitting*) e o subajustamento (*underfitting*) (Goodfellow et al., 2016).

A aprendizagem profunda é um subconjunto da aprendizagem automática que recorre a redes neuronais profundas (Deep Neural Networks, DNNs) para modelar relações complexas nos dados. Estas redes são compostas por múltiplas camadas hierárquicas, permitindo a extração progressiva de características abstratas (LeCun et al., 2015). No processamento de imagens, destacam-se as Redes Neuronais Convolucionais (CNNs), que demonstram elevado desempenho na análise de imagens devido à sua capacidade de capturar relações espaciais. As CNNs são amplamente utilizadas em super-resolução de imagens, como na reconstrução de imagens Sentinel-2 de baixa resolução (Krizhevsky et al. 2012). Recentemente, as Redes Generativas Adversativas (GANs) revolucionaram a super-resolução ao introduzirem um modelo dual composto por um gerador, que cria imagens sintéticas, e um discriminador, que avalia a sua autenticidade. Modelos como SRGAN (Ledig et al., 2017) e ESRGAN (Wang et al., 2018) permitiram avanços significativos na resolução de imagens de satélite, superando técnicas convencionais baseadas em CNNs.

2.3 Redes Neuronais e Arquiteturas de Processamento de Imagem

As Redes Neuronais Artificiais (ANNs) são modelos computacionais inspirados na organização e no funcionamento do sistema nervoso, cuja génese remonta ao perceptrão de Rosenblatt (1958). Uma ANN é composta por neurónios artificiais interligados por pesos sinápticos e organizados em camadas (entrada, ocultas e saída). Cada neurónio combina linearmente as entradas e aplica uma função de ativação, permitindo modelar relações não lineares entre variáveis. Inseridas no domínio da aprendizagem automática, as ANNs, e, em particular, as suas variantes profundas, têm desempenhado um papel central no processamento de imagem e noutras áreas, pela capacidade de aprender representações hierárquicas diretamente a partir dos dados (LeCun et al., 2015).

2.3.1 Arquitetura das Redes Neuronais

As redes neuronais artificiais (ANNs) podem ser classificadas conforme a forma como os neurónios se organizam e interligam. Entre as estruturas mais comuns encontram-se as Redes Neuronais Feedforward (FFNNs) e as Redes Neuronais Recorrentes (RNNs).

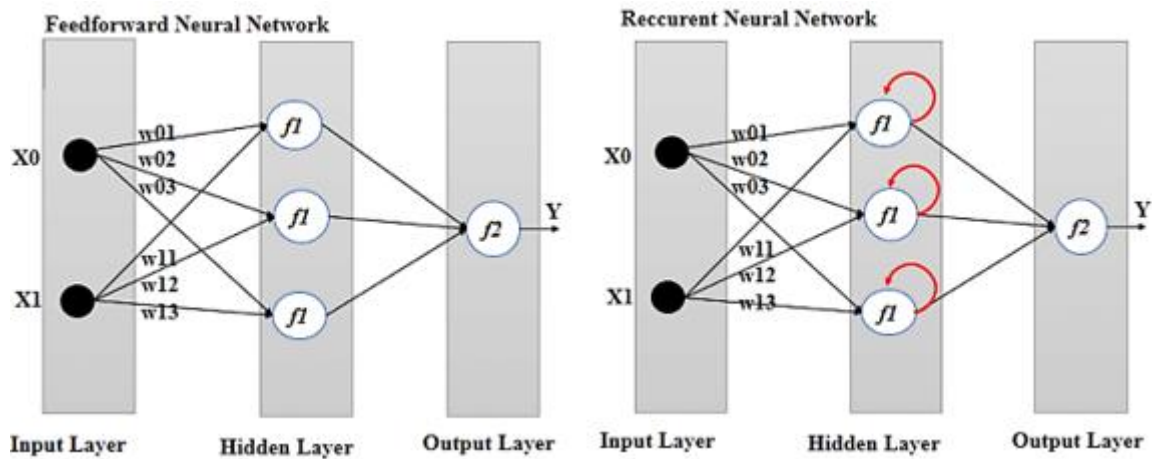


Figura 2.1 Redes Neurais Feedforward Neural Network vs. Recorrentes. Reproduzido de Fakhfakh et. al (2017).

Nas FFNNs, a informação propaga-se apenas no sentido direto, da camada de entrada para a(s) camada(s) oculta(s) e, por fim, para a camada de saída, sem ciclos nem realimentação, sendo treinadas por retropropagação do erro (Figura 2.1) (Goodfellow et al., 2016). Esta família inclui arquiteturas especializadas como as Redes Convolucionais (CNNs), particularmente eficazes no processamento de imagens pela capacidade de explorar a estrutura espacial local através de filtros partilhados, razão pela qual são amplamente usadas em super-resolução e noutras tarefas de visão por computador aplicadas a dados de deteção remota.

Por contraste, as RNNs introduzem ligações temporais que permitem reter e transmitir informação de um passo para o seguinte ao longo de uma sequência (Figura 2.1), tornando-as adequadas ao tratamento de séries temporais e de dados sequenciais (por exemplo, cadeias multitemporais de observação da Terra).

Internamente, os dados são representados sob a forma de tensores, isto é, arranjos multidimensionais que generalizam vetores (1D) e matrizes (2D). Em imagem, é comum usar tensores 3D (altura×largura×canais) (por exemplo, RGB) e, durante o treino, tensores 4D com a dimensão batch. Esta representação permite implementar eficientemente as operações nucleares da aprendizagem profunda (convoluções, multiplicações matriciais, ativações e normalizações) e a diferenciação automática necessária à retropropagação. A execução é acelerada por GPUs e TPUs (Goodfellow et al., 2016; Jouppi et al., 2017).

2.3.2 Processo de Aprendizagem das Redes Neurais

O treino de uma rede neuronal consiste no ajuste iterativo dos pesos e vieses para minimizar uma função de custo que quantifica a discrepância entre as previsões do modelo e as observações de referência (Goodfellow et al., 2016). Em termos gerais, procura-se $\min_{\theta} [\mathcal{L}(f_{\theta}(x), y)]$, onde \mathcal{L} é a perda escolhida para a tarefa. Na super-resolução (SR), os problemas de regressão de imagem para imagem, a perda *pixel-wise* mais comum é o Erro Quadrático Médio (MSE) ou a L1, pela sua relação direta com a qualidade radiométrica local (Bishop, 2006). Para promover realismo perceptivo, é frequente combinarem-se perdas adicionais: a perda perceptual calculada em características de uma rede VGG pré-treinada (Visual Geometry Group, rede convolucional desenvolvida pela Universidade de Oxford), que privilegia estrutura e textura (Johnson, Alahi, & Fei-Fei, 2016), e um termo adversarial quando se treina com GANs, criando saídas visualmente plausíveis (Ledig et al., 2017; Wang et al., 2018).

A minimização da função de custo faz-se por descida do gradiente e variantes estocásticas, com gradientes calculados via retropropagação (*backpropagation*) usando a regra da cadeia (Rumelhart, Hinton, & Williams, 1986). Na prática, utiliza-se *mini-batch* SGD ou otimizadores adaptativos. A taxa de aprendizagem é um hiperparâmetro crítico: valores excessivos dificultam a convergência; valores demasiado baixos tornam o treino moroso e propenso a ótimos pouco profundos (Sutskever et al., 2013). Para gerir esta sensibilidade, recorrem-se a agendadores (p. ex., *step decay*, *cosine annealing* ou ciclos), que ajustam dinamicamente a taxa ao longo das épocas e ajudam a ultrapassar regiões de gradiente reduzido (plateaus) no espaço de otimização (Smith, 2017). Mecanismos como momentum amortecem oscilações e estabilizam a direção de descida (Polyak, 1964).

No contexto específico deste projeto, SR adversarial de dados Sentinel-2 com geradores tipo ESRGAN, a prática consolidada é combinar uma perda *pixel-wise* (L1/MSE) com uma perceptual (VGG) e uma adversarial, recorrendo ao treino com *mini-batches* compostos por pequenos blocos (*tiles*) de imagens de baixa e alta resolução espacialmente co-registradas. Esta formulação equilibra fidelidade geométrica/radiométrica e realismo visual, requisitos essenciais para que o produto SR seja útil em interpretação cartográfica e não apenas esteticamente mais nítido.

2.3.3 Classificação temática com Random Forest (RF)

O Random Forest (RF) é um método de aprendizagem supervisionada amplamente adotado em deteção remota para classificação de uso e ocupação do solo, pela sua robustez, interpretabilidade operacional e baixo custo de parametrização. O algoritmo constrói um conjunto de árvores de decisão sobre amostras *bootstrap* do conjunto de treino, introduzindo aleatoriedade adicional ao seleccionar, em cada divisão, apenas um subconjunto de variáveis candidatas (Figura 2.2). A combinação por votação das árvores reduz a variância do estimador e mitiga o sobreajuste, assegurando boa generalização em cenários com relações não lineares, como sucede em imagens multiespectrais (Breiman, 2001; Belgiu & Drăguț, 2016; Gislason, Benediktsson, & Sveinsson, 2006).

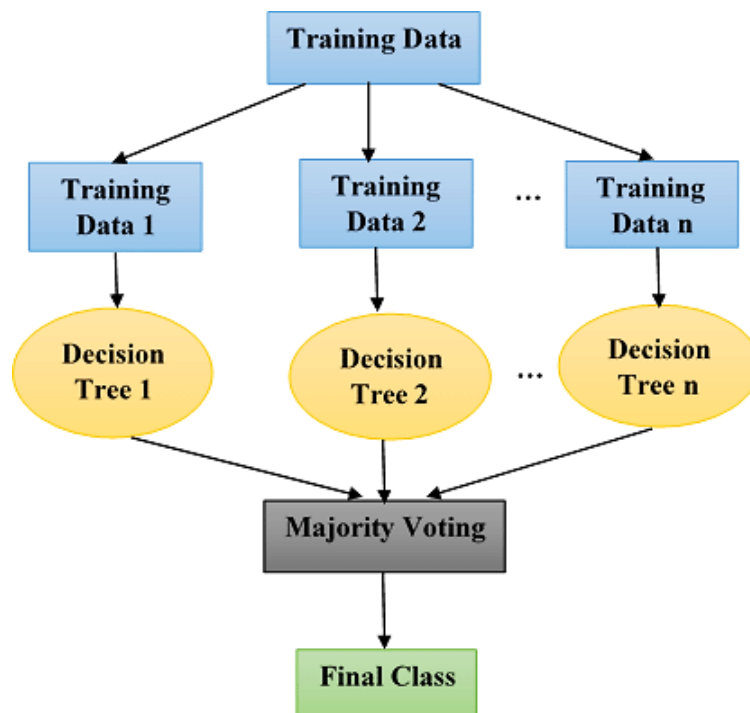


Figura 2.2 Workflow do Random Forest. Reproduzido de Batool *et. al.* (2017).

No contexto geoespacial, o RF destaca-se por três propriedades particularmente úteis. Primeiro, a estimativa *out-of-bag* (OOB) fornece uma medida de erro quase isenta de viés sem necessidade de reservar dados exclusivamente para validação, o que é vantajoso quando as amostras anotadas são dispendiosas. Segundo, as medidas de importância de variáveis por permutação permitem aferir a contribuição relativa de bandas e derivados (por exemplo, componentes RGB/TCI do Sentinel-2, índices de vegetação, texturas), facilitando a interpretação. Terceiro, as probabilidades de classe produzidas pelo ensemble suportam análise de incerteza temática e limiares ajustáveis, úteis em fluxos cartográficos que exigem controle de omissões e comissões por classe (Rodríguez-Galiano et al., 2012; Maxwell, Warner, & Fang, 2018).

A parametrização do RF é frugal: o número de árvores deve ser suficientemente elevado para estabilizar a OOB e a importância de variáveis; o número de preditores considerados em cada divisão (parâmetro *mtry*, que controla o *bias-variance trade-off*); e o tamanho mínimo do nó estabelece a granularidade das fronteiras de decisão. Em imagens de média resolução, as fronteiras entre classes urbanas, por exemplo, vias estreitas adjacentes a edifícios, tendem a apresentar mistura espectral; aqui, o RF beneficia de atributos que realcem arestas e continuidade linear, pelo que os produtos de super-resolução podem aumentar a separabilidade temática sem alterar o modelo, ao disponibilizar ao classificador sinais espaciais mais informativos à escala do píxel (Belgiu & Drăguț, 2016; Maxwell et al., 2018).

Do ponto de vista metodológico, duas precauções são centrais em estudos geoespaciais. A autocorrelação espacial pode inflacionar artificialmente estimativas de desempenho quando treino e validação partilham vizinhanças; por isso, as estratégias de validação espacialmente explícitas, como *spatial/block cross-validation*, são recomendadas quando se pretende inferir capacidade de transferência espacial. Em paralelo, a desproporção entre classes, frequente em contextos urbanos em que “vias” são sub-representadas, deve ser tratada com amostragem estratificada, pesos de classe ou variantes balanceadas RF, sob pena de a precisão global ocultar um fraco desempenho nas classes minoritárias (Roberts et al., 2017; Belgiu & Drăguț, 2016).

Neste projeto, o RF é utilizado como instrumento de avaliação temática para comparar, de forma controlada, o impacto da resolução espacial nas classificações derivadas de Sentinel-2. Mantendo o desenho de amostragem e as mesmas geometrias de referência em LR, SR e HR, isola-se o efeito do produto de imagem na discriminação de edifícios, vias, espaços verdes e água. Assim, o RF funciona como elo metodológico entre os ganhos visuais trazidos pelas GANs de super-resolução e a utilidade cartográfica mensurável, expressa em matrizes de confusão, F1-score por classe e recall, no quadro rigoroso da engenharia geoespacial.

2.4 Redes Generativas Adversárias (GANs)

As Redes Generativas Adversárias (GANs) foram introduzidas por Goodfellow et al. (2014) como uma arquitetura generativa treinada por competição entre duas redes neuronais, um gerador e um discriminador, que jogam um “jogo” de soma zero até o gerador produzir amostras indistinguíveis dos dados reais. Desde então, as GANs transformaram a visão computacional e consolidaram-se como uma solução de referência para síntese e melhoria de imagens, incluindo super-resolução em deteção remota, onde a necessidade de detalhe espacial e coerência radiométrica é crítica para o uso cartográfico.

No contexto geoespacial, as GANs têm permitido ganhos tangíveis na reconstrução de padrões urbanos finos (contornos de edifícios, continuidade de vias) e na preservação de características espectrais relevantes, quando comparadas com abordagens convencionais. Ao aprenderem diretamente a distribuição dos dados, evitam as suavizações típicas da interpolação (por exemplo, bicúbica) e

superam redes convolucionais “simples” orientadas apenas ao erro pixel-a-pixel, produzindo texturas plausíveis e arestas mais nítidas, aspetos essenciais para discriminação temática e vetorização. Esta capacidade é particularmente pertinente para imagens Sentinel-2, em que a resolução nativa limita a leitura de classes urbanas finas e onde se procura, nesta dissertação, elevar a resolução aparente para 2,5 m de forma cartograficamente útil.

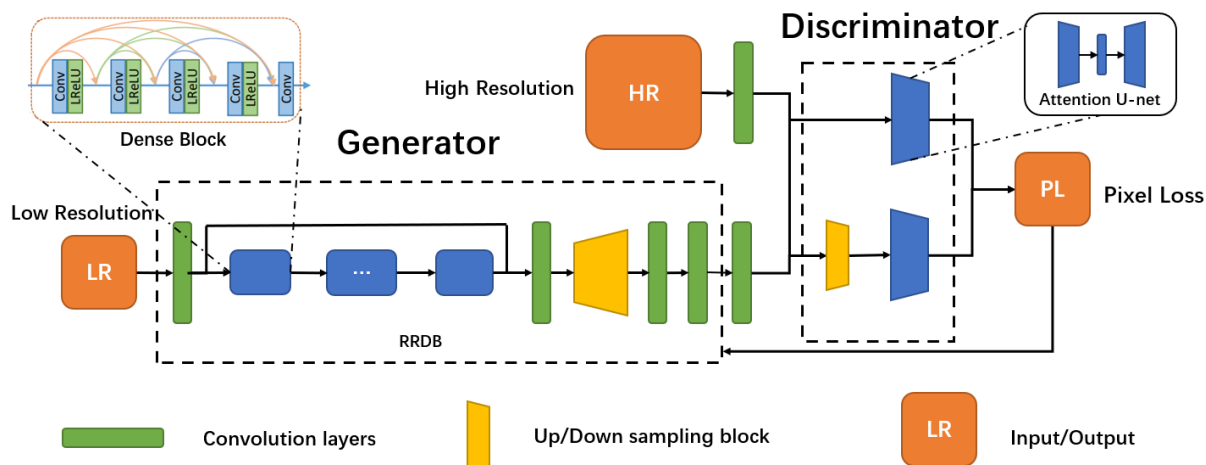


Figura 2.3 Arquitetura geral do modelo GAN utilizado para super-resolução. Reproduzido de Mäkinen, O. (2020).

Formalmente, o funcionamento das GANs (Figura 2.3) assenta na interação entre o gerador G , que mapeia amostras de uma distribuição latente $p_z(z)$ para o espaço dos dados, e o discriminador D , que procura distinguir amostras reais $x \sim p_{data}(x)$ de amostras sintéticas $G(z)$. O treino resolve o problema min-max clássico:

$$\min_G \max_D E_{x \sim p_{data}(x)} [\log \sim D(x)] + E_{z \sim p_z(z)} [\log (1 - D(G(z)))] \quad (2.1)$$

em que G procura aproximar p_{data} e D maximizar a sua capacidade de discriminação (Goodfellow et al., 2014). Quando o equilíbrio é alcançado, D deixa de distinguir com fiabilidade entre amostras reais e geradas, e as saídas de G apresentam realismo estrutural compatível com a referência (Goodfellow et al., 2014; Goodfellow, 2016; Goodfellow, Bengio, & Courville, 2016; Creswell et al., 2018).

Quando aplicadas à super-resolução, as variantes SRGAN e ESRGAN tornaram-se marcos: a SRGAN introduziu a perda perceptual para incentivar estruturas de alto nível e texturas realistas, enquanto a ESRGAN refinou o gerador com blocos residuais-densos (RRDB) e uma formulação adversarial mais estável, elevando a fidelidade visual sem sacrificar a coerência geométrica (Ledig et al., 2017; Wang et al., 2018). Esta linha de desenvolvimento é a base metodológica adotada neste projeto para “super-resolver” imagens Sentinel-2 e avaliar, com rigor geoespacial, o seu impacto na utilidade cartográfica.

2.4.1 Super-Resolution GAN (SRGAN)

O SRGAN, proposto por Ledig et al. (2017), foi o primeiro a articular a estrutura adversária de uma GAN com um gerador de super-resolução profundo, permitindo reconstruir detalhe de alta frequência de forma perceptualmente convincente. A inovação central reside na perda perceptual, calculada no espaço característico de uma rede pré-treinada (VGG-19), aproximando a otimização ao juízo humano

de qualidade visual, em vez de penalizar apenas erros radiométricos ponto-a-ponto (Simonyan & Zisserman, 2015; Johnson et al., 2016). No SRGAN, a função objetivo do gerador combina um termo perceptual com o termo adversarial:

$$\mathcal{L}_G = \mathcal{L}_{\text{perceptual}} + \lambda \mathcal{L}_{\text{adversarial}} \quad (2.2)$$

em que $\mathcal{L}_{\text{perceptual}}$ mede a distância entre as características (*features*) de alto nível de $G(I_{LR})$ e I_{HR} (extraídas da VGG-19), $\mathcal{L}_{\text{adversarial}}$ encoraja amostras indistinguíveis do domínio HR e λ pondera a contribuição adversarial. Esta formulação evita as soluções excessivamente suaves típicas de perdas pixel-wise puras (e.g., MSE), reforçando arestas, texturas e micro-padrões críticos para análise urbana (Ledig et al., 2017; Johnson et al., 2016).

Do ponto de vista arquitetural, o gerador segue um *backbone* residual (SRResNet) com *upsampling* aprendido, enquanto o discriminador é uma CNN binária que avalia a “realidade” das reconstruções. Em detecção remota, esta combinação tem sido aplicada à melhoria de resolução de sensores óticos, reportando ganhos visuais e funcionais face a métodos convencionais (e.g., interpolação bicúbica) precisamente onde o Sentinel-2 é mais penalizado: contornos de edifícios, continuidade viária e interfaces água-terra. Estudos de adaptação do SRGAN a dados de satélite mostram que o aumento de 30 m para 15 m no Landsat-8 pode traduzir-se em melhor discriminação temática quando comparada com métodos puramente interpolativos, sobretudo em classes urbanas de geometria fina (Ledig et al., 2017; Wang et al., 2021). No âmbito deste projeto, aumentar a resolução de S2 (10 m) para 2,5 m com utilidade cartográfica, a lógica do SRGAN é particularmente pertinente: a perda perceptual ajuda a recuperar detalhe estrutural relevante para vetorização e classificação urbana, desde que se mantenha o controlo de fidelidade geométrica e espectral (Simonyan & Zisserman, 2015; Johnson et al., 2016; Ledig et al., 2017).

2.4.2 Enhanced Super-Resolution GAN

O ESRGAN, proposto por Wang et al. (2018), refinou o SRGAN com três contributos decisivos para melhorar detalhe e estabilidade. Em primeiro lugar, removeu o *batch normalization* dos blocos do gerador, evitando artefactos e variações indesejadas de contraste e, simultaneamente, reduzindo o custo computacional. Em segundo, substituiu os blocos residuais convencionais pelos Residual-in-Residual Dense Blocks (RRDB), que combinam ligações residuais e densas em múltiplos níveis, preservando frequências altas e reforçando a passagem de gradientes em arquiteturas profundas. Em terceiro, aperfeiçoou a perda perceptual calculando-a em *features* antes da ativação (em vez do pós-ativação típico), o que se traduz em bordos mais consistentes e texturas mais naturais, aspetos especialmente relevantes para contornos de edifícios, continuidade viária e interfaces água-terra em contexto urbano.

Uma inovação central do ESRGAN é o discriminador relativístico (RaGAN), no qual o discriminador deixa de avaliar amostras de forma binária (real vs. gerada) e passa a estimar “o quão mais real” uma imagem é relativamente a outra. Esta formulação relativística melhora a estabilidade do treino adversarial e induz texturas mais fidedignas ao penalizar diretamente diferenças de realismo entre pares real-gerado (Jolicœur-Martineau, 2019), mitigando oscilações e reduzindo o risco de mode collapse.

No plano da avaliação, o ESRGAN popularizou o Perceptual Index (PI), um índice que combina medidas “sem referência” para aproximar a qualidade perceptual à forma como os humanos julgam detalhe e naturalismo, desencorajando otimizações excessivas em métricas de distorção como PSNR/SSIM que, por si só, tendem a favorecer soluções demasiado suaves. Este enquadramento é coerente com o objetivo deste projeto: aumentar a utilidade cartográfica das imagens Sentinel-2 ao nível de 2,5 m, priorizando nitidez estrutural e legibilidade de padrões urbanos (edifícios e vias) sem comprometer a coerência geométrica e espectral.

2.4.3 Limitações e Desafios

Apesar do elevado potencial das GANs, a sua aplicação na área da observação da Terra tem demonstrado algumas limitações. Uma das principais é a instabilidade durante o treino, causada pela competição entre o gerador e o discriminador. Além disso, a necessidade de utilização de conjuntos de dados diversificados pode limitar a generalização do modelo para cenários não representados no treino. Outra limitação à sua implementação é o elevado custo computacional. O treino de GANs requer um hardware com grande poder de cálculo e requer um elevado período de tempo, especialmente ao trabalhar com imagens multiespectrais de alta resolução. Por fim, o risco de sobreajustamento é uma preocupação constante, particularmente em aplicações que envolvem dados escassos ou pouco diversificados.

A evolução das GANs e das suas variantes mais recentes, como o Hi-ResGAN (High-Resolution GAN) e o F-SRGAN (Feature-Supervised GAN), promete continuar a melhorar a qualidade das imagens de deteção remota, tornando-as uma ferramenta fundamental para aplicações de geociências e análise espacial. O Hi-ResGAN, por exemplo, foi projetado para lidar com a escassez de dados de alta resolução para treino. Neste modelo, técnicas de aprendizagem semi supervisionada foram integradas, permitindo que o Hi-ResGAN mantivesse as características espectrais essenciais enquanto melhorava a resolução espacial das imagens Sentinel-2.

Outro modelo notável é o F-SRGAN, que combina supervisão específica com redes adversárias para criar imagens otimizadas para tarefas de análise, como a identificação de florestas ou corpos d'água. Esses estudos ilustram a versatilidade das GANs e o seu potencial para superar limitações técnicas em deteção remota.

2.5 Avaliação da Qualidade das Imagens de Super-resolução

A avaliação da qualidade das imagens geradas por modelos de super-resolução é um desafio fundamental na área da visão computacional. As diferentes métricas podem ser utilizadas para quantificar a fidelidade da imagem reconstruída em relação à original, sendo essencial escolher as métricas adequadas ao tipo de análise pretendida. De forma geral, estas métricas podem ser divididas em quantitativas, que avaliam a diferença numérica entre as imagens, e perceptuais, que procuram medir a qualidade percebida pelo olho humano. Embora as métricas quantitativas como o Peak Signal-to-Noise Ratio (PSNR) e o Structural Similarity Index Measure (SSIM) sejam amplamente utilizadas, diversas investigações têm demonstrado que estas podem ser insuficientes para captar todas as características visuais que tornam uma imagem realista. Assim, métodos perceptuais, como o Perceptual Index (PI) e o Naturalness Image Quality Evaluator (NIQE), têm sido incorporados para melhorar a avaliação da qualidade das imagens de super-resolução.

2.5.1 Métricas Quantitativas

Uma das métricas mais utilizadas na avaliação da qualidade de reconstrução de imagens é o Peak Signal-to-Noise Ratio (PSNR). O PSNR mede a razão entre a potência máxima de um sinal e a potência do ruído que afeta a sua qualidade. Como muitas das imagens possuem uma elevada dinâmica radiométrica, o PSNR é expresso numa escala logarítmica em decibéis (dB), sendo definido da seguinte forma:

$$PSNR = 20 \log_{10} \left(\frac{MAX_I}{\sqrt{MSE}} \right) \quad (2.3)$$

onde MAX_I representa o valor máximo possível de um píxel na imagem original, e MSE é o erro quadrático médio, definido por:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} |I(i, j) - J(i, j)|^2 \quad (2.4)$$

neste caso, $I(i, j)$ e $J(i, j)$ correspondem aos valores de píxel da imagem original e da imagem de super-resolução, respetivamente, e $m \times n$ representa a dimensão da imagem.

Os valores de PSNR mais elevados indicam uma reconstrução mais próxima da imagem original. No contexto da super-resolução baseada em *aprendizagem profunda*, os valores típicos de PSNR situam-se entre 25 e 35 dB, dependendo do método utilizado. Um PSNR infinito ocorre apenas quando $MSE = 0$, ou seja, quando a reconstrução é idêntica à imagem original. No entanto, apesar da sua popularidade, diversos estudos têm demonstrado que o PSNR nem sempre está correlacionado com a perceção humana da qualidade visual. Em particular, pode falhar na avaliação de texturas e detalhes finos, levando à necessidade de métricas mais sofisticadas.

Para ultrapassar algumas dessas limitações, o Structural Similarity Index Measure (SSIM) foi proposto como uma alternativa mais eficaz. Enquanto o PSNR mede erros absolutos entre os valores dos píxeis, o SSIM tenta prever a qualidade percebida das imagens, avaliando a sua estrutura, luminância e contraste. O SSIM é baseado no princípio de que a perceção visual de uma imagem depende da preservação das suas relações estruturais, em vez da mera minimização do erro píxel a píxel. A sua formulação matemática é dada por:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (2.5)$$

onde μ_x e μ_y representam as médias das intensidades das imagens x e y , σ_x^2 e σ_y^2 correspondem às suas variâncias, e σ_{xy} denota a covariância entre as duas imagens. As constantes C_1 e C_2 são utilizadas para evitar instabilidades numéricas. O SSIM varia entre 0 e 1, onde valores próximos de 1 indicam maior similaridade estrutural entre as imagens comparadas. Esta métrica leva em consideração fenómenos perceptuais, como máscaras de luminosidade e de contraste, que influenciam a forma como os seres humanos percebem variações na imagem. Diferentemente do PSNR, o SSIM tem demonstrado melhor desempenho na avaliação da qualidade das imagens geradas por modelos de super-resolução, sobretudo na preservação de texturas e padrões visuais.

2.5.2 Avaliação Perceptual e Novas Métricas

Embora o SSIM represente um avanço face ao PSNR, a necessidade de métricas que se aproximem ainda mais da perceção humana levou ao desenvolvimento de métodos perceptuais mais avançados. Um desses métodos é o Perceptual Index (PI), utilizado por Wang et al. (2018) no modelo ESRGAN. Este índice combina duas métricas distintas: o NRQM (No-Reference Quality Metric) e o Naturalness Image Quality Evaluator (NIQE), proporcionando uma avaliação mais precisa da qualidade perceptual da imagem de super-resolução.

O NIQE é um avaliador cego da qualidade das imagens, ou seja, não depende de imagens de referência para medir a qualidade. Este método analisa estatísticas naturais das imagens, comparando características estatísticas extraídas da imagem de super-resolução com um modelo treinado em imagens naturais de alta qualidade. Diferente do PSNR e SSIM, que exigem uma imagem de referência para

comparação, o NIQE baseia-se exclusivamente na análise estatística da imagem de entrada. Quanto menor o valor de NIQE, melhor a qualidade percetual da imagem. No entanto, por ser uma métrica estatística, pode apresentar limitações na avaliação de artefactos gerados por redes neurais, exigindo a combinação com outras métricas percetuais, como o LPIPS (Learned Perceptual Image Patch Similarity), que mede a distância percetual entre imagens usando redes pré-treinadas.

Outro avanço na avaliação percetual é o uso de redes neuronais profundas, como a VGG-19, para extrair características de alto nível das imagens geradas. Em vez de comparar diretamente os valores dos píxeis, este método avalia a similaridade entre mapas de características extraídos de diferentes camadas da rede. Essa abordagem permite capturar detalhes estruturais e texturas de forma mais próxima à percepção humana. No ESRGAN, Wang et al. (2018) demonstraram que substituir a função de perda baseada no erro quadrático médio (MSE) por uma perda percetual baseada na VGG-19 resulta em imagens mais realistas e detalhadas.

A utilização combinada de métricas quantitativas e percetuais tornou-se essencial para avaliar a qualidade das imagens geradas por modelos de super-resolução. Enquanto o PSNR e o SSIM continuam a ser amplamente utilizados, as métricas como o Perceptual Index e o NIQE oferecem uma avaliação mais alinhada com a percepção humana. A integração de redes neuronais na avaliação da qualidade das imagens representa um avanço significativo, permitindo que a análise não se limite apenas a diferenças numéricas, mas também a aspetos estruturais e texturais essenciais para a reconstrução realista de imagens de super-resolução.

2.6 Síntese do Capítulo

Este capítulo apresentou uma visão abrangente sobre os principais conceitos e desenvolvimentos na área de super-resolução de imagens, com foco na aplicação de redes generativas adversárias para melhorar a classificação da ocupação do solo utilizando imagens Sentinel-2. Foi evidenciado que a super-resolução desempenha um papel crucial na melhoria da análise de imagens de deteção remota, permitindo superar limitações da resolução espacial dos sensores existente, como o Sentinel-2. Trabalhos anteriores (Dong et al. (2016); Ledig et al. (2017)) demonstraram a eficácia das GANs para gerar detalhes espaciais mais precisos em imagens de baixa resolução, superando métodos tradicionais.

No contexto específico das imagens Sentinel-2, observou-se que a sua resolução espacial limitada (10-60 metros, dependendo da banda espectral) pode comprometer a precisão de classificações detalhadas da ocupação do solo, especialmente em áreas com alta heterogeneidade espacial (Gao et al., 2022). Estudos recentes que aplicaram GANs para super-resolução de imagens satélite confirmam o potencial dessas redes para resolver este desafio (Zhu et al. (2021); Choi et al. (2023)).

No entanto, identificaram-se algumas lacunas. Primeiramente, muitos dos estudos analisados concentram-se em imagens de alta resolução espacial (por exemplo, QuickBird e/ou WorldView), enquanto há relativamente poucos trabalhos focados em melhorar a resolução de imagens Sentinel-2 especificamente para fins de classificação da ocupação do solo. Além disso, embora as GANs demonstrem grande potencial, desafios como a necessidade de grandes conjuntos de dados de treino e a dificuldade em generalizar os modelos para diferentes cenários geográficos ainda precisam ser abordados (Goodfellow et al. (2014); Huang et al. (2022)).

O presente trabalho posiciona-se neste contexto, procurar explorar o uso de GANs para melhorar a resolução espacial das imagens Sentinel-2 e avaliar o impacto dessa melhoria na precisão da classificação da ocupação do solo. Este esforço visa não apenas contribuir para a literatura existente, mas também fornecer uma abordagem prática para aplicações em monitorização ambiental e gestão do

território. O próximo capítulo detalhará a metodologia proposta para alcançar esses objetivos, incluindo o processo de pré-processamento de dados, bem como o treino do modelo GAN.

3 Dados e Métodos

3.1 Área de Estudo

A área de estudo selecionada corresponde aos concelhos de Lisboa e Cascais, situados na região metropolitana de Lisboa, Portugal. Esta área foi escolhida devido à sua diversidade de usos e coberturas do solo, bem como pela relevância socioeconómica e ambiental associada à sua localização estratégica no território nacional. A conjugação de áreas densamente urbanizadas, zonas residenciais periféricas, áreas verdes estruturadas (como parques e jardins) e espaços naturais com forte valor ecológico proporciona um contexto heterogêneo e desafiante para a aplicação de métodos de super-resolução de imagens de satélite.

O concelho de Lisboa, capital do país, é caracterizado por uma elevada densidade populacional e por uma complexa malha urbana, marcada pela presença de edifícios residenciais, infraestruturas de transportes, áreas comerciais e zonas industriais. Já o concelho de Cascais, localizado a oeste de Lisboa, apresenta uma configuração territorial distinta, marcada pela coexistência de áreas urbanas consolidadas, zonas costeiras de elevado valor paisagístico e áreas de vegetação natural, incluindo espaços verdes protegidos. Esta combinação confere-lhe uma dinâmica territorial própria, distinta de Lisboa, mas complementar do ponto de vista metodológico, permitindo avaliar a robustez dos modelos em diferentes cenários geoespaciais.

A escolha destes dois concelhos apresenta várias vantagens para o estudo da super-resolução. Em primeiro lugar, a diversidade interna de coberturas do solo (edifícios, vias, vegetação e massas de água) permite testar a capacidade dos modelos de reconstruir padrões espaciais muito distintos, desde superfícies contínuas e homogêneas até áreas altamente fragmentadas (Figura 3.1). Em segundo lugar, o facto de Lisboa e Cascais se encontrarem lado a lado, mas apresentarem configurações urbanas diferenciadas, possibilita uma análise comparativa que vai além da simples replicação de resultados. Por fim, a existência de dados oficiais de ortoimagens de elevada resolução (25 cm) e a disponibilidade regular de imagens Sentinel-2 constituem uma ótima oportunidade para o estudo com dados reais e atualizados.

Do ponto de vista temporal, estes concelhos apresentam ainda um interesse acrescido: ambos têm sofrido alterações significativas no uso do solo entre 2018 e 2023, em particular no que se refere à requalificação de espaços públicos e transformação da cobertura vegetal em resposta a intervenções humanas. Assim, a análise comparativa entre diferentes momentos temporais contribui para a avaliação da estabilidade e da aplicabilidade prática das metodologias de super-resolução.

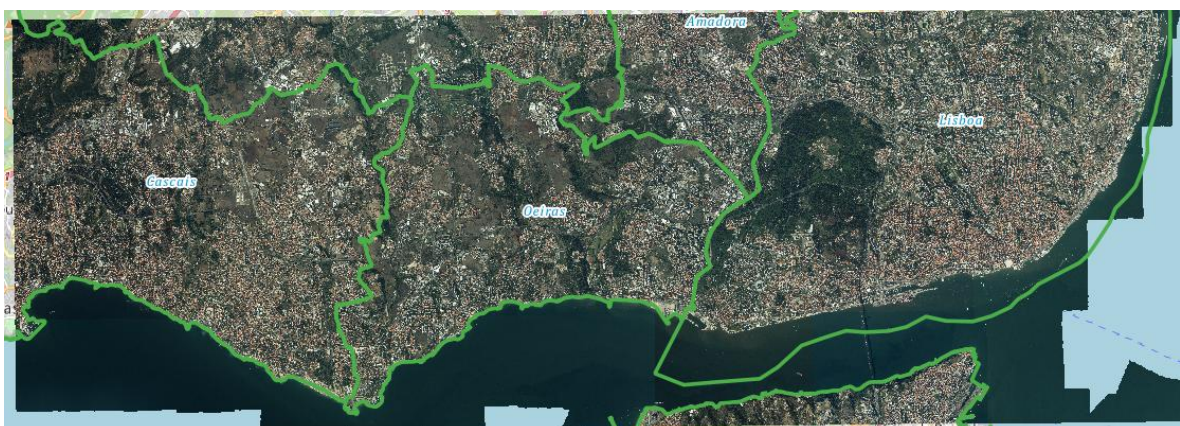


Figura 3.1 Área de estudo (Lisboa e Cascais). Extensão dos mosaicos de ortoimagens usados (2023) e limites administrativos municipais.

3.2 Dados Sentinel-2

A missão Sentinel-2, integrada no programa Copernicus da Agência Espacial Europeia (ESA), foi concebida com o objetivo de fornecer dados multiespectrais de elevada qualidade para aplicações ambientais, agrícolas, florestais, urbanas e de ordenamento do território (Drusch et al., 2012; ESA, 2021). A constelação é composta por dois satélites, Sentinel-2A (lançado em 2015) e Sentinel-2B (lançado em 2017), que operam em órbitas heliosíncronas a cerca de 786 km de altitude. Em conjunto, estes dois satélites asseguram uma frequência de revisita de aproximadamente cinco dias no equador, permitindo uma cobertura global sistemática e consistente (Li & Roy, 2017).

O instrumento a bordo, o MSI (Multispectral Instrument), capta imagens em 13 bandas espectrais, que abrangem desde o visível (VIS) até ao infravermelho de ondas curtas (SWIR). Estas bandas apresentam diferentes resoluções espaciais: 10 metros para as bandas no visível e no infravermelho próximo, 20 metros para bandas no infravermelho de média resolução e 60 metros para as bandas destinadas à caracterização atmosférica (Figura 3.2) (ESA, 2021). Esta diversidade confere ao Sentinel-2 uma grande versatilidade, mas também limitações quando se pretende realizar análises em contextos que exigem elevado detalhe espacial, como o mapeamento de áreas urbanas densas ou a caracterização de pequenas parcelas agrícolas (Belgiu & Drăguț, 2016; Zhu et al., 2018).

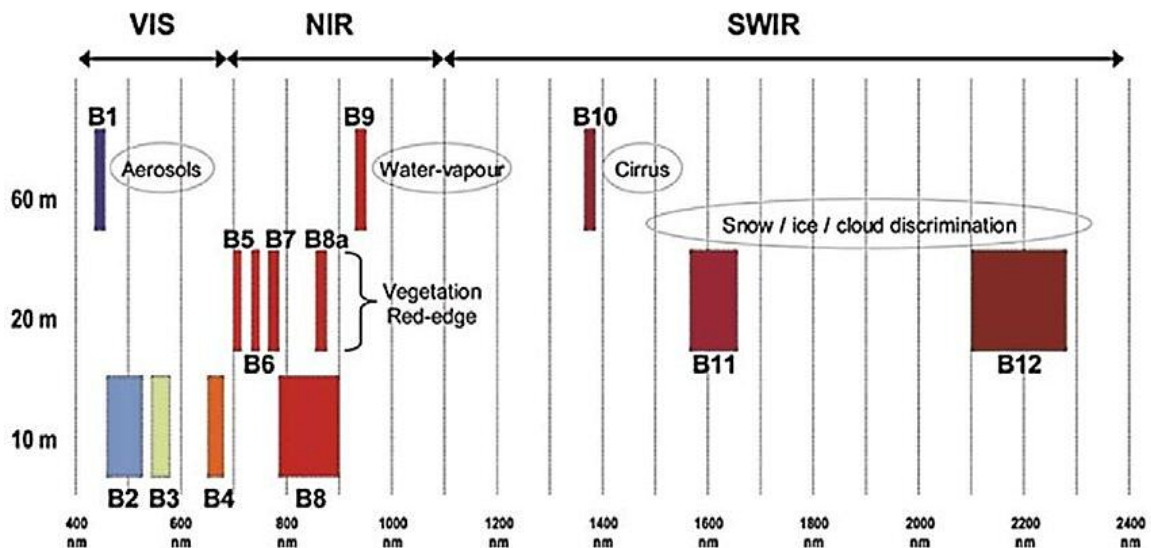


Figura 3.2 Cobertura espectral do MSI do Sentinel-2 e resoluções espaciais (10, 20 e 60 m). Reproduzido de Farris et al. (2022).

As imagens Sentinel-2 utilizadas neste estudo já incluem a correção atmosférica realizada pelo algoritmo Sen2Cor, assegurando uma maior consistência radiométrica e geométrica dos dados (ESA, 2023). Todos os produtos foram obtidos a partir do Copernicus Open Access Hub, garantindo a sua proveniência de uma fonte oficial, pública e gratuita.

No que respeita à seleção temporal, para o ano de 2023 foram escolhidas oito imagens adquiridas entre 27 de agosto e 6 de outubro, todas com percentagem de nebulosidade inferior a 20%, de forma a minimizar o impacto da cobertura atmosférica. Para assegurar a comparabilidade direta com as ortoimagens de referência, foi utilizada apenas a composição RGB (True Color Image - TCI), com resolução espacial de 10 metros. De igual modo, para o ano de 2018 foram recolhidas oito imagens, adquiridas entre 1 e 31 de agosto, respeitando os mesmos critérios de seleção: composição TCI e limiar máximo de 20% de cobertura de nuvens.

A escolha dos períodos de aquisição das imagens Sentinel-2 não foi arbitrária, mas fundamentada na necessidade de assegurar a comparabilidade com as ortoimagens de referência. As ortoimagens de 2018 foram adquiridas a 9 de julho e as de 2023 a 23 de setembro, pelo que a seleção das imagens Sentinel-2 entre agosto e início de outubro representa um compromisso adequado entre disponibilidade de dados e proximidade temporal. Esta estratégia garante que as características sazonais da vegetação e do ambiente urbano se mantêm consistentes entre conjuntos de imagens, reduzindo potenciais discrepâncias decorrentes de diferenças fenológicas ou de variações interanuais. O critério adicional de manter uma cobertura máxima de 20% de nuvens permitiu ainda assegurar a qualidade radiométrica das imagens e minimizar a interferência atmosférica. Desta forma, a seleção temporal adotada maximiza a coerência entre os diferentes conjuntos de dados (ortoimagens e Sentinel-2), tanto no domínio espacial (Lisboa e Cascais) como no domínio temporal (2018 vs. 2023), constituindo uma base sólida para a avaliação da super-resolução no contexto da classificação da ocupação do solo.

3.3 Ortoimagens

As ortoimagens constituem a fonte de dados de referência de maior resolução espacial utilizada neste estudo, funcionando como imagens de alta-qualidade com as quais foram comparados os resultados de super-resolução derivados das imagens Sentinel-2 foram comparados. Uma ortoimagem corresponde a uma imagem aérea corrigida geometricamente, de forma a eliminar distorções provocadas pela inclinação da câmara, relevo do terreno e variações na altitude de voo, permitindo que cada píxel seja representado em coordenadas geográficas precisas (Toutin, 2004). Assim, as ortoimagens combinam a riqueza visual das imagens aéreas com a precisão geométrica dos mapas cartográficos, sendo largamente utilizadas em aplicações de monitorização territorial, planeamento urbano e estudos ambientais (Novák & Tóth, 2016).

Neste trabalho foram utilizadas ortoimagens relativas a dois períodos temporais distintos: 2018 e 2023. As ortoimagens de 2018 foram disponibilizadas pela Direção-Geral do Território (DGT), constituindo o último levantamento nacional com cobertura total para Portugal continental. As imagens apresentam resolução espacial de 25 cm, o que assegura um elevado nível de detalhe para a caracterização da paisagem urbana. Já as ortoimagens de 2023 foram disponibilizadas pelo IFAP (Instituto de Financiamento da Agricultura e Pescas), resultando de uma campanha mais recente, também com resolução espacial de 25 cm, assegurando uma base comparativa atualizada e de elevada qualidade.

A escolha das ortoimagens como dado de referência apresenta várias vantagens. Em primeiro lugar, a sua resolução de 25 cm garante a presença de elementos urbanos naturais bem definidos, como edifícios, estradas e manchas de vegetação, permitindo aferir com elevado rigor a qualidade das imagens de super-resolução geradas. Em segundo lugar, a disponibilidade de ortoimagens em dois momentos temporais distintos (2018 e 2023) possibilitou a realização de experiências não apenas espaciais (Lisboa vs. Cascais), mas também temporais, comparando a eficácia dos modelos em contextos de mudança da paisagem ao longo de cinco anos (Figura 3.3).



Figura 3.3 Alcântara (frente ribeirinha, Hospital CUF Tejo). Excerto comparativo ortoimagem 2018 (a) vs ortoimagem 2023 (b).

3.4 Metodologia

Esta secção descreve o desenho metodológico adotado para avaliar o potencial das técnicas de super-resolução baseadas em GANs na melhoria da utilidade cartográfica de imagens Sentinel-2 em contexto urbano. O objetivo central é verificar se o aumento de detalhe espacial obtido por super-resolução se traduz em ganhos operacionais, i.e., em produtos mais informativos para classificação temática e análise de mudança, quando comparado com o produto de baixa resolução (LR) e com uma referência de alta resolução (HR) derivada de ortoimagens.

A abordagem segue uma lógica pipeline com quatro blocos principais: (i) preparação de dados, garantindo compatibilidade geométrica e radiométrica entre Sentinel-2 TCI (10 m) e ortoimagens (2.5 m); (ii) geração de pares LR-HR por tiles com razão espacial 4x ($32 \times 32 \leftrightarrow 128 \times 128$ píxeis), assegurando o registo 1:1; (iii) treino e inferência com a arquitetura Satlas-Super-Resolution (variante ESRGAN com séries temporais S2), produzindo saídas SR à escala alvo de 2,5 m; e (iv) avaliação da qualidade e utilidade, combinando métricas quantitativas de reconstrução e um ensaio temático com Random Forest. Esta estrutura minimiza fontes de erro (tais como desalinhamento, diferenças de projeção ou discrepâncias de amostragem) e permite atribuir os efeitos observados ao processo de super-resolução.

Do ponto de vista espacial e temporal, o plano experimental foi concebido para testar consistência interna, generalização espacial e robustez temporal. Concretamente: (a) uma verificação de coerência com treino e validação no mesmo universo de dados; (b) um ensaio cruzado Lisboa \leftrightarrow Cascais para aferir transferência entre malhas urbanas contíguas, mas morfologicamente distintas; e (c) um ensaio 2018 \rightarrow 2023 para avaliar comportamento sob mudança real da paisagem. Em todas as abordagens, as partições foram construídas de modo a evitar contaminação dos dados (data leakage) espacial/temporal e a manter a comparabilidade geoespacial entre produtos LR, SR e HR. Por contaminação dos dados entende-se a presença de dados errados no processo de treino que não é detetado nos processos de teste e validação e conduz a predições erradas.

A avaliação temática foi concebida em dois níveis complementares. Numa subárea de referência Braço de Prata/Parque Ribeirinho Oriente, polígonos delineados sobre HR (2,5 m) são sobrepostos a LR e SR para inspeção direta de ganhos de nitidez, continuidade e separabilidade entre classes urbanas (água, edifícios, vias, espaços verdes). Em paralelo, uma avaliação extensiva em toda a área de estudo usa ~300 polígonos de referência distribuídos pelas quatro classes para treinar/testar um classificador

Random Forest em LR, SR e HR com as mesmas amostras emparelhadas, permitindo comparar efeitos do aumento de resolução em métricas padronizadas (precisão, F1-score, recall).

A implementação do algoritmo baseia-se no repositório Satlas-Super-Resolution (Allen Institute for AI), utilizado tal como publicado, com a exceção de um menor número de iterações de treino por limitações computacionais, e no uso de Sentinel-2 TCI (True Color Image) alinhado temporalmente com as ortoimagens, garantindo coerência fenológica. As saídas SR são recompostas em mosaico georreferenciado (GeoTIFF) no sistema cartográfico adotado, assegurando integração direta em fluxos SIG e comparação transparente com LR e HR. A combinação entre um pipeline reprodutível e um plano de experiências estratificado permite responder, com rigor geoespacial, à questão de fundo: quando e quanto a super-resolução acrescenta valor cartográfico em cenários urbanos reais.

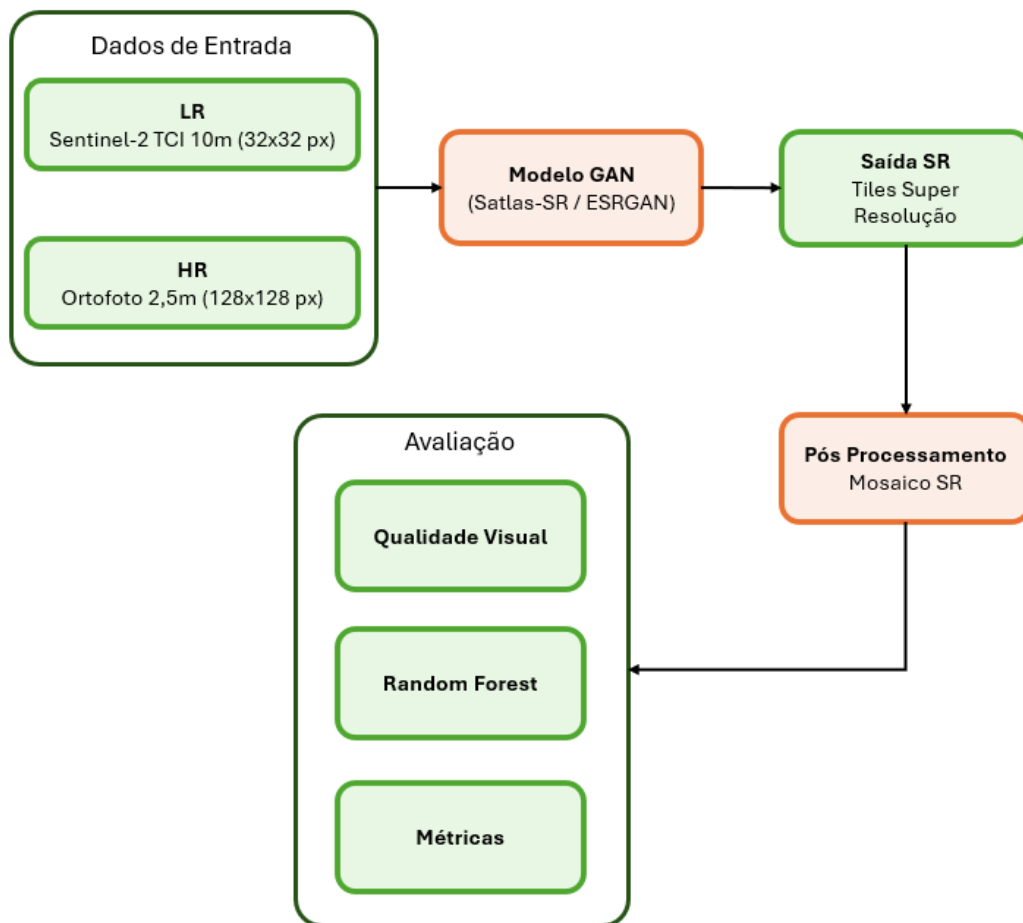


Figura 3.4 Workflow do estudo.

3.4.1 Arquitetura da Rede

A arquitetura adotada nesta dissertação assenta no Framework Satlas-Super-Resolution do Allen Institute for AI (AI2), concebido especificamente para super-resolução em detecção remota com dados multisensor (Sentinel-2 como baixa resolução e ortoimagem aérea como alta resolução). O núcleo do método é adversarial (GAN), com um gerador que reconstrói detalhe espacial a partir de entradas de baixa resolução e um discriminador que avalia a verosimilhança visual das saídas face a dados de referência. No repositório oficial, a família de modelos disponibilizada inclui variantes ESRGAN-like

(baseadas em RRDBNet), bem como sequências não-GAN; os pesos e configurações associadas são fornecidos para diferentes números de imagens Sentinel-2 de entrada e para diferentes subconjuntos espectrais (10 m, 20 m, 60 m) (Wolters, Bastani, & Kembhavi, 2023; Ledig et al., 2017; Wang et al., 2018).

Do ponto de vista operacional, o Satlas-SR organiza os dados em pares LR-HR com escala 4x, refletida no desenho dos tiles: 32x32 píxeis para a entrada Sentinel-2 (LR) e 128x128 píxeis para a referência de alta resolução (HR). No conjunto de dados de origem do projeto (S2), as ortoimagens (HR) são RGB e são fornecidas a 25% da resolução nativa, recortadas em 128x128 píxeis; para cada tile HR existe uma série temporal de entradas Sentinel-2 alinhadas espacial e temporalmente, guardadas à dimensão 32x32 píxeis. Para entradas RGB, o TCI de Sentinel-2 (imagem 8-bit normalizada pela ESA em [0,255]) é renormalizado para [0,1] por divisão por 255 antes de entrar no modelo. Estes detalhes estruturantes: tamanhos de tile, normalização e emparelhamento LR-HR são parte integrante do pipeline e foram seguidos neste trabalho (AI2, 2023).

O gerador segue a linha SRGAN/ESRGAN e usa blocos residuais-densos (RRDB), que ajudam a passar a informação entre camadas e, por isso, recuperam melhor texturas e contornos finos. A mudança de resolução é feita com upsampling de fator 4x, que cria píxeis de forma coerente com o conteúdo, em vez de apenas interpolar. No treino, a rede otimiza uma perda composta: (i) perceptual para manter o aspeto natural; (ii) *pixel-wise* (L1/L2) para garantir proximidade ao alvo de alta resolução; e (iii) adversarial, que orienta o gerador a produzir imagens com aparência mais realista. Na prática, esta combinação dá mais nitidez e melhor definição de arestas do que métodos clássicos, conservando a coerência geométrica. O discriminador é treinado para distinguir saídas SR de verdadeiros alvos HR, conduzindo a um jogo minimax que promove realismo visual sem sacrificar a fidelidade geométrica. No repositório Satlas, esta combinação “adversarial + perceptual + *pixel-wise*” é fornecida de base nas configurações ESRGAN; neste trabalho a implementação foi utilizada tal como publicada, alterando-se apenas o número de iterações (reduzido por constrangimentos computacionais), sem mudanças na arquitetura ou nas perdas (Goodfellow et al., 2014; Ledig et al., 2017; Wang et al., 2018; AI2, 2023).

A integração geoespacial ocorre fora da rede: o treino e a inferência operam estritamente no domínio dos tiles, e o posterior mosaico das saídas SR é realizada com base no índice de grelha dos tiles, reconstituindo produtos raster georreferenciados (GeoTIFF) à resolução alvo (2,5 metros) no sistema de referência cartográfico do estudo (UTM), com programa desenvolvido para o efeito. O repositório disponibiliza os scripts de treino e inferência sobre um tile Sentinel-2 completo, o que se alinha com o procedimento adotado neste trabalho (AI2, 2023).

Em síntese, a arquitetura utilizada combina: (i) GAN com gerador do tipo ESRGAN, (ii) emparelhamento LR-HR consistente (Sentinel-2 TCI 10 m → 32x32 píxeis; orto HR degradada → 128x128 píxeis), (iii) normalização apropriada das entradas, e (iv) mosaico georreferenciado pós inferência. Este stack permite que a avaliação realizada nesta tese incida na questão geoespacial central: em que medida o ganho de detalhe produzido pela GAN se traduz em valor cartográfico (visual e quantitativo) para a classificação temática e análise da ocupação do solo. (Wolters et al., 2023; Ledig et al., 2017; Wang et al., 2018).

3.4.2 Tratamento e Pré-Processamento dos Dados

O tratamento dos dados teve como objetivo garantir a comparabilidade geométrica e radiométrica entre as fontes utilizadas, Sentinel-2 e ortoimagens, e preparar os dados de entrada no formato requerido pela arquitetura Satlas-SR. Todas as operações foram concebidas para reduzir fontes de viés não relacionadas com a super-resolução (desalinhamento espacial, diferenças de projeção, artefactos de

reamostragem, etc), assegurando que os ganhos observados se devem ao modelo e não a inconsistências possíveis no pré-processamento.

Num primeiro passo, procedeu-se à harmonização cartográfica. As ortoimagens de 2018 e 2023 foram reprojatadas para o sistema de referência WGS84 / UTM zona 29N (EPSG:32629), sistema adotado no estudo, e para as ortoimagens de 2018, recortadas aos limites de referência das ortoimagens de 2023 (concelhos de Lisboa e Cascais). As imagens Sentinel-2 foram selecionadas privilegiando as janelas temporais próximas das ortoimagens de referência. Para compatibilizar escalas, as ortoimagens (resolução nativa de 25 cm) foram degradadas para 2,5 m antes do treino, seguindo a prática comum em super-resolução supervisionada: a imagem de alta resolução serve de alvo e a versão degradada atua como ground truth na escala de saída. Este procedimento reduz a discrepância entre o domínio de treino e o de inferência sobre Sentinel-2 e está alinhado com o pipeline do Satlas-SR (AI2, 2023; Ledig et al., 2017; Wang et al., 2018). A operação de degradação foi aplicada de forma a mitigar aliasing, preservando a consistência radiométrica média dos alvos.

O co-registo espacial entre as ortoimagens degradadas e Sentinel-2 foi assegurado por alinhamento ao mesmo grid cartográfico e verificação visual de deslocamentos residuais, garantindo sobreposição píxel a píxel nas áreas de interesse. Após a harmonização geométrica, ambos os conjuntos foram recortados de modo a cobrir exatamente as mesmas extensões. No domínio radiométrico, manteve-se a composição RGB (TCI) para Sentinel-2, conforme a opção metodológica deste estudo. Em consonância com as recomendações do repositório Satlas, os valores 8-bit [0–255] das imagens TCI foram normalizados para [0–1] por divisão por 255 antes da entrada no modelo; as ortoimagens (utilizadas como alvo HR) foram tratadas de forma equivalente (AI2, 2023). Não foram utilizadas bandas não-TCI, evitando, assim, a necessidade de normalizações diferenciadas (e.g., escalas 16-bit) e possíveis assimetrias espectrais entre as fontes.

Concluídas estas etapas, os dados foram preparados para a fase de tiling, requerida pela arquitetura: *tiles* LR de 32×32 píxeis (Sentinel-2 TCI, 10 m) emparelhados com *tiles* HR de 128×128 píxeis (orto degradada, 2,5 m), mantendo a razão espacial 4× definida pelo método (Figura 3.5). A organização por grelha e a indexação espacial coerente garantiram a rastreabilidade dos pares e a prevenção de contaminação entre partições (treino/validação). A produção e partilha destas partições são detalhadas na Secção 3.4.3. Por fim, as saídas do modelo foram posteriormente recompostas em mosaicos georreferenciados (GeoTIFF) com resolução alvo de 2,5 m/píxel e no mesmo sistema de referência, assegurando a integridade cartográfica para análise e comparação subsequente com os produtos LR e HR.

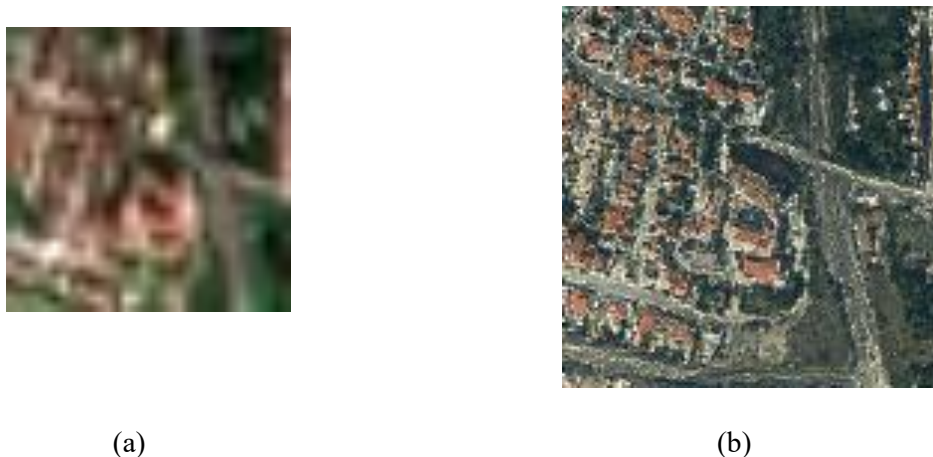


Figura 3.5 Grelhas equivalentes LR/HR no mesmo tile. Janela com grelha LR 32×32 px a 10 m (Sentinel-2 TCI) (a); e grelha HR 128×128 px a 2,5 m (b)

Este fluxo de pré-processamento, harmonização cartográfica, degradação controlada da ortoimagem, normalização radiométrica e recorte espacialmente consistente, segue as melhores práticas de detecção remota e os requisitos específicos de aprendizagem profunda para super-resolução, permitindo uma avaliação justa do contributo do modelo para a melhoria do detalhe espacial com relevância cartográfica (Drusch et al., 2012; ESA, 2021; AI2, 2023).

3.4.3 Dados de Treino e Validação (Geração de Tiles)

A preparação dos dados para treino e validação seguiu a estratégia de tiling definidos na estrutura Satlas-SR, adaptando-a ao contexto Lisboa/Cascais e às épocas de 2018 e 2023. O objetivo foi produzir pares estritamente correspondentes LR–HR com razão espacial 4x, garantindo correspondência geométrica pixel-a-pixel entre as entradas Sentinel-2 e os alvos derivados das ortofotografias.

Em cada época (2018 e 2023), os tiles HR foram gerados por varrimento regular da área de estudo e emparelhados 1:1 com os tiles LR obtidos a partir das imagens Sentinel-2, assegurando a co-localização e mitigando erros de registo. Seguiu-se a convenção de organização em pastas do repositório Satlas (pastas por tile e por sensor), o que facilita a rastreabilidade e a reprodutibilidade (AI2, 2023). Como para cada tile de HR usaram-se 8 imagens LR, as entradas LR foram agregadas conforme o pipeline do repositório (séries temporais empilhadas por tile), não exigindo ordenação cronológica explícita. Não se aplicou nenhum filtro ou máscara adicional a nível de tile assim, a qualidade dos dados de entrada reflete diretamente os critérios de seleção Sentinel-2. A produção dos tiles foi automatizada, resultando em coleções equilibradas por área e época, e armazenadas com metadados de localização e época, o que permitiu reconstruir posteriormente os mosaicos SR georreferenciados e proceder a análises de qualidade e utilidade cartográfica.

A organização dos dados seguiu o desenho experimental já descrito (Sec. 3.4.7; Tabela 3.1): três cenários mutuamente exclusivos: verificação interna, generalização espacial Lisboa↔Cascais e robustez temporal 2018→2023. As partições de treino/validação foram não sobrepostas para evitar *data leakage* e preservar a comparabilidade entre LR, SR e HR.

Resumindo, a geração de tiles assegurou: (i) correspondência espacial exata LR-HR à escala 4x; (ii) normalização consistente das entradas RGB; (iii) partições sem contaminação entre treino e validação nos três cenários realizados (interna, espacial e temporal); e (iv) rastreabilidade geográfica por indexação e metadados (Tabela 3.1). Esta estrutura é coerente com as melhores práticas do Satlas-SR e crucial em avaliar, em condições controladas, o impacto da super-resolução na qualidade visual e no valor cartográfico subsequente (AI2, 2023; Wolters et al., 2023).

Tabela 3.1 Esquema de tiling e partições (treino vs. validação).

	TREINO	VALIDAÇÃO
VALIDAÇÃO INTERNA	LR e HR, 2023	LR e HR, 2023
LISBOA ↔ CASCAIS	LR e HR, Lisboa, 2023	LR e HR, Cascais, 2023
	LR e HR, Cascais, 2023	LR e HR, Lisboa, 2023
2018→2023	LR e HR, 2018	LR e HR, 2023

3.4.4 Treino e Predição da Rede

O treino foi realizado com a arquitetura Satlas-SR (AI2, 2023), baseada no ESRGAN, utilizando séries temporais de oito aquisições Sentinel-2 por tile (experimento “8S2”). A implementação foi usada sem alterações, exceto pela redução do número de iterações por limitações computacionais. Os tiles LR correspondem a janelas Sentinel-2 TCI, enquanto os HR derivam de ortoimagens degradadas, assegurando equivalência métrica numa razão espacial $4\times$. O gerador (RRDBNet) foi otimizado em regime adversarial clássico, combinando perdas *pixel-wise*, *perceptual* e adversarial, de modo a preservar fidelidade geométrica e a reforçar texturas e contornos urbanos relevantes.

Toda a orquestração do treino assentou nos ficheiros de configuração YAML da família *esrgan_baseline_8S2*, onde se definem os caminhos das partições e os parâmetros de execução. A invocação foi efetuada através dos wrappers de alto nível do Satlas-SR, tanto para treino como para inferência, garantindo reprodutibilidade e rastreabilidade. Em cada cenário experimental, validação interna, generalização espacial Lisboa↔Cascais e robustez temporal 2018↔2023, foram executadas experiências independentes, respeitando as partições sem interseção definidas previamente, de modo a evitar *data leakage* espacial e temporal e a medir a verdadeira capacidade de generalização do modelo nas condições de uso previstas.

As predições foram geradas por tile e, numa fase de pós-processamento, recompostas em mosaico por grelha, utilizando o índice espacial dos tiles para reconstituir cenas contínuas. A georreferenciação foi herdada do esquema de grelha aplicado durante a geração dos tiles, produzindo produtos finais GeoTIFF (e imagens auxiliares PNG) à resolução alvo de 2,5 m, no sistema WGS 84 / UTM zona 29N (EPSG:32629), compatível com os fluxos SIG e com a comparação direta face aos produtos LR e HR. Este desenho assegura que qualquer melhoria observada nas métricas (e.g., PSNR e SSIM) decorre do processo de super-resolução e não de desalinhamentos geométricos ou de discrepâncias de amostragem. No contexto deste estudo, o papel do treino não é meramente estético: ao otimizar o gerador com perdas *perceptuais* e adversárias em conjunto com uma referência geoespacial consistente, pretende-se que o ganho de detalhe seja operacionalmente útil, i.e., que se traduza em melhor discriminação temática e em produtos mais informativos para tarefas de cartografia urbana e monitorização de alterações. Assim, o método de treino adotado cumpre a dupla exigência da engenharia geoespacial: rigor na preservação das propriedades espaciais e ganho *perceptual* suficiente para sustentar decisões cartográficas em áreas complexas como Lisboa e Cascais.

3.4.5 Métricas de Avaliação

A avaliação quantitativa da super-resolução incide na fidelidade de reconstrução do produto SR face à referência HR, tendo o LR como linha de base. Para garantir comparabilidade, as imagens foram co-registradas *pixel-a-pixel* e recortadas à mesma área espacial. As métricas foram calculadas por banda (R, G, B) e reportadas como média RGB.

Adotaram-se quatro métricas complementares: MSE e MAE quantificam o erro radiométrico médio (com maior resistência a outliers no MAE), sendo úteis para detetar discrepâncias locais e artefactos; PSNR sintetiza a relação sinal-ruído e é amplamente usada na literatura de super-resolução, facilitando comparações; SSIM foca-se em brilho, contraste e estrutura, aproximando-se da percepção humana da qualidade, o que é particularmente informativo em contexto urbano (contornos de edifícios, continuidade de vias, interfaces água-terra) (Wang, Bovik, Sheikh, & Simoncelli, 2004; Gonzalez & Woods, 2018; Ma et al., 2019; Yang, Zhang, Tian, Wang, & Xue, 2019). A interpretação segue as convenções: MSE/MAE menores e PSNR/SSIM maiores indicam melhor fidelidade.

Operacionalmente, cada métrica foi calculada entre a imagem SR e a imagem HR (objetivo principal) e entre a imagem LR e a imagem HR (referência), permitindo quantificar o ganho SR face a LR sob o mesmo protocolo. Os cálculos foram realizados com um código em Python, utilizando bibliotecas padrão de processamento de imagem, e executados de forma automática sobre todos os cenários. Contudo, reconhece-se o trade-off percepção-distorção: otimizações estritas de MSE/PSNR tendem a produzir imagens demasiado suaves, enquanto abordagens adversárias (como as usadas neste estudo) reforçam a nitidez percebida sem maximizar necessariamente essas métricas (Blau & Michaeli, 2018). Por isso, os indicadores numéricos são interpretados em conjunto com a leitura visual co-registada e com a avaliação temática (RF), alinhando a análise com o objetivo: melhorar a legibilidade cartográfica e a utilidade dos produtos para classificação e análise de mudança.

3.4.6 Avaliação Temática

A avaliação temática foi desenhada para aferir se o ganho de detalhe produzido pela super-resolução tem utilidade cartográfica efetiva em contexto urbano, para além de melhorias puramente visuais. Para tal, foram conduzidas duas avaliações complementares e metodologicamente coerentes com a engenharia geoespacial: uma avaliação visual e localizada, com sobreposição de polígonos de referência numa subárea problemática do tecido urbano; e uma avaliação quantitativa e extensiva, baseada em classificação supervisionada em toda a área de estudo.

Na primeira vertente, centrada na subárea de Braço de Prata e Parque Ribeirinho Oriente, foram delineados polígonos de referência diretamente sobre as ortoimagens degradadas. Estes polígonos representam quatro classes cartograficamente relevantes, água, edifícios, vias e espaços verdes, e foram mantidos invariantes ao comparar os três produtos: LR, HR e SR. A opção por rotular à escala de 2,5 m evita desfasamentos de escala entre a geometria vetorial dos polígonos e o raster subjacente, focando a análise no que importa para o uso cartográfico: nitidez de contornos, continuidade linear de vias, definição de arestas água-terra e separabilidade entre coberturas adjacentes em quarteirões densos. A comparação lado-a-lado (HR vs. LR vs. SR), com os polígonos sobrepostos, permite observar de forma controlada se a SR reduz mistura espectral em bordos (efeito mais penalizador em LR), melhora a coerência geométrica das vias e edifícios e mantém a leitura dos espaços verdes sem introduzir artefactos. Esta avaliação foi repetida em todos os testes experimentais (validação interna, generalização espacial Lisboa↔Cascais e robustez temporal 2018→2023), preservando a mesma geometria de referência para garantir comparabilidade.

A segunda vertente consistiu numa avaliação quantitativa extensiva assente em classificação Random Forest (RF) aplicada separadamente a imagens LR, SH e HR em toda a área de estudo (Lisboa e Cascais). Para construir o conjunto de referência, foram vetorizados ~300 polígonos distribuídos pelas quatro classes, com maior incidência em edifícios para capturar a diversidade morfológica urbana. Sobre cada polígono, amostraram-se píxeis rotulados ao nível do píxel, usando apenas atributos RGB (em coerência com o desenho do estudo), e adotou-se uma partição aleatória de 70% para treino e 30% para validação, mantendo índices espaciais emparelhados entre LR, SR e HR. Deste modo, o classificador aprende e é testado em cada domínio, isolando o efeito da resolução/qualidade do produto na capacidade discriminativa. Não foi aplicado nenhum *buffer* aos polígonos, assim, a amostragem inclui zonas de fronteira suscetíveis à mistura de classes em LR, o que torna a avaliação mais exigente e realista para uso cartográfico urbano. O RF foi escolhido pela robustez ao ruído, capacidade de modelar relações não lineares e baixo custo de parametrização, qualidades que explicam a sua adoção extensiva em deteção remota (Breiman, 2001; Belgiu & Drăguț, 2016). Em cada domínio (LR, SR, HR) o RF foi aplicado para gerar mapas temáticos contínuos, permitindo comparar diferenças de interpretação cartográfica e ganhos quantitativos.

A leitura dos resultados baseou-se na matriz de confusão, reportando-se precisão global e o F1-score por classe. A precisão sintetiza o desempenho agregado; o F1-score por classe equilibra precisão e revocação e é sensível a assimetrias (Congalton & Green, 2009; Foody, 2002). Para interpretação cartográfica, deu-se particular atenção às confusões estruturais típicas de malhas urbanas, vias vs. edifícios e aresta água-terra, e à sua eventual atenuação em SR relativamente a LR, tanto nos mapas temáticos como nas métricas. O protocolo foi replicado em 2018 e 2023 com o mesmo esquema de polígonos e amostragem, assegurando coerência temporal; deste modo, diferenças entre épocas refletem quer a evolução real da paisagem, quer a capacidade do produto SR de sustentar decisões cartográficas consistentes ao longo do tempo.

Em conjunto, a avaliação localizada (qualitativa) e a avaliação extensiva (quantitativa) permitem concluir se e onde a super-resolução acrescenta valor operacional: melhoria de legibilidade de redes viárias e quarteirões, maior separabilidade de classes e ganhos de desempenho estatístico em métricas padronizadas, tudo sob um desenho que controla a comparabilidade geoespacial entre produtos.

3.4.7 Testes Realizados

O desenho experimental deste estudo foi estruturado para isolar, de forma progressiva, três dimensões cruciais do desempenho em super-resolução com relevância geoespacial: consistência interna do pipeline, transferência espacial entre regiões com morfologias distintas e robustez temporal perante alterações reais da paisagem. Para tal, implementaram-se três testes complementares, todos ancorados no mesmo esquema de dados, no mesmo processo de tiling co registado e no mesmo regime de treino/inferência, garantindo comparabilidade metodológica (Ledig et al., 2017; Wang et al., 2018).

O primeiro teste correspondeu a uma validação interna do fluxo de dados. O modelo foi treinado e avaliado sobre o mesmo universo de tiles, preservando a separação formal entre treino e validação, mas sem imposição de deslocamento espacial ou temporal entre partições. Esta etapa não pretende sustentar conclusões de generalização; a sua função é confirmar que a cadeia de pré-processamento, treino, inferência e mosaico está corretamente implementada e que a rede é capaz de produzir saídas SR estáveis e coerentes com as referências HR (Goodfellow et al., 2014). Em termos práticos, esta verificação preliminar permite identificar de forma expedita desalinhamentos residuais, artefactos de costura ou degradações radiométricas não desejadas antes de passar a testes mais exigentes. É expectável que os indicadores de qualidade e a classificação temática alcancem, aqui, valores superiores aos das restantes abordagens, precisamente por não envolverem transferência espacial ou temporal.

O segundo teste avaliou a capacidade de generalização espacial. O treino foi realizado com tiles de um concelho (Lisboa) e a validação com tiles do concelho adjacente (Cascais), e vice-versa, sem partilha de amostras entre as partições. Este desenho confronta a rede com variações de densidade construída, textura urbana e composição de coberturas que caracterizam os dois municípios vizinhos, mas distintos, mitigando assim a influência da autocorrelação espacial de curto alcance que poderia inflacionar métricas quando o treino é co localizado (Tobler, 1970; Anselin, 1995). Do ponto de vista cartográfico, este teste é particularmente informativo: a melhoria esperada com SR deve traduzir-se numa maior separabilidade entre edifícios e vias e numa melhor definição das interfaces vegetação-edificado, precisamente nas áreas onde a malha urbana de Cascais difere da de Lisboa, e onde um modelo treinado num município poderia, de outra forma, perder acuidade.

O terceiro teste examinou a robustez temporal, treinando com os dados de 2018 e validando com os de 2023. Esta configuração introduz uma diferença temporal em que o domínio de validação incorpora alterações reais da paisagem, novas construções, requalificação das redes viárias, mudanças de cobertura vegetal, potenciando um cenário de dataset shift no sentido clássico: a distribuição dos dados no teste

difere da distribuição vista no treino (Moreno-Torres et al., 2012). Ao manter inalterado o pipeline e as fontes (S2 TIC e orto HR degradada), esta abordagem permite atribuir diferenças de desempenho à capacidade da SR em sustentar decisões cartográficas em contextos evolutivos, e não a fatores de preparação de dados. Em termos de análise temática, a repetição do protocolo de classificação com os mesmos polígonos de referência e a mesma amostragem (70/30) nas duas épocas fornece um quadro coerente para aferir ganhos operacionais: uma melhoria consistente das métricas em SR relativamente a LR, em 2018 e em 2023, constitui uma evidência de utilidade prática em monitorização urbana.

Em todos os testes, foram rigorosamente preservadas a coerência geométrica dos pares LR-HR, a normalização radiométrica das entradas RGB e o esquema de partições sem interseção, evitando fugas de informação entre treino e validação. A validação interna funciona como prova de sanidade do processo; a transferência Lisboa↔Cascais responde à questão da generalização espacial em malhas urbanas com assinaturas morfológicas distintas; e o ensaio 2018→2023 enfrenta explicitamente a variabilidade temporal das cidades (Figura 3.6). Em conjunto, o triplo desenho experimental alinha-se com as boas práticas de avaliação em deteção remota e aprendizagem automática geoespacial, permitindo interpretar ganhos em SR tanto sob métricas de qualidade de imagem como sob métricas de utilidade cartográfica (AI2, 2023; Congalton & Green, 2009; Foody, 2002).

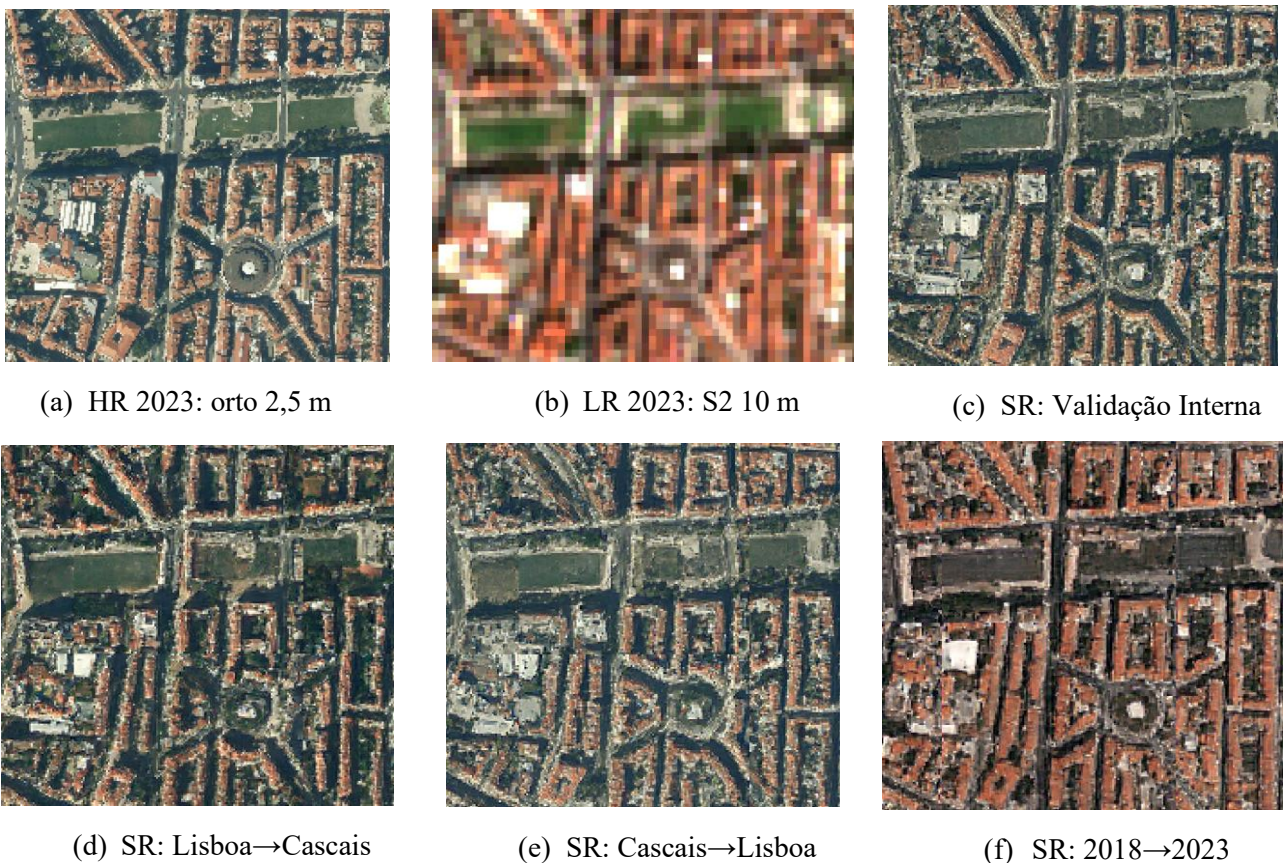


Figura 3.6 Infográfico dos cenários experimentais (overview).

4 Resultados e Discussão

Este capítulo apresenta e interpreta os resultados obtidos nos três testes considerados, verificação interna, generalização espacial Lisboa ↔ Cascais e robustez temporal 2018 → 2023, com foco nos quatro eixos de análise: qualidade visual das reconstruções, leitura com polígonos de referência na subárea Braço de Prata/Parque Ribeirinho Oriente, avaliação temática com Random Forest em toda a área de estudo e análise dos registos de treino do modelo GAN. A discussão é conduzida de modo a responder explicitamente aos objetivos definidos: (1) demonstrar a viabilidade de super-resolução LR → SR preservando a consistência espectral; (2) avaliar a eficácia do modelo em cenários urbanos distintos face à referência HR; e (3) quantificar o impacto da SR na classificação da ocupação do solo.

4.1 Qualidade visual

A avaliação da qualidade das imagens de super-resolução (SR) foi conduzida de forma exclusivamente visual, com ênfase na legibilidade cartográfica e na coerência geoespacial dos produtos. O objetivo é determinar se o aumento de detalhe espacial obtido pela arquitetura ESRGAN (implementação Satlas-SR) se traduz em ganhos operacionais face ao produto LR (Sentinel-2 TCI a 10 m) e à referência HR (ortoimagem degradada para 2,5 m) (Figura 4.1), evitando interpretações condicionadas por métricas numéricas. Para tal, todas as comparações foram realizadas em escala cartográfica constante, sobre recortes corrigistados e com polígonos de referência sobrepostos, garantindo que as diferenças observadas derivam do processo de super-resolução e não de desalinhamentos ou de variações de escala.



Figura 4.1 Painel comparativo (a) LR 2023 → (b) HR 2023 → (c) SR (visão geral).

O protocolo de leitura privilegiou três eixos. Primeiro, a definição de contornos e arestas em classes urbanas críticas, nomeadamente edifícios e vias, por serem elementos estruturantes da cartografia de base. Segundo, a continuidade geométrica de traçados estreitos (ruas, rotundas de pequena dimensão, separadores), onde a mistura espectral do LR tende a quebrar a conectividade. Terceiro, a estabilidade de interfaces com forte contraste, como a separação água-terra e as transições vegetação-edificado, relevantes para mapeamento de margens, taludes e manchas verdes. A inspeção realizou-se em multiescala: painéis lado a lado LR → SR → HR para leitura de contexto e “zooms” locais para análise fina de bordos, sempre com a mesma janela espacial e projeção (Figura 4.2).

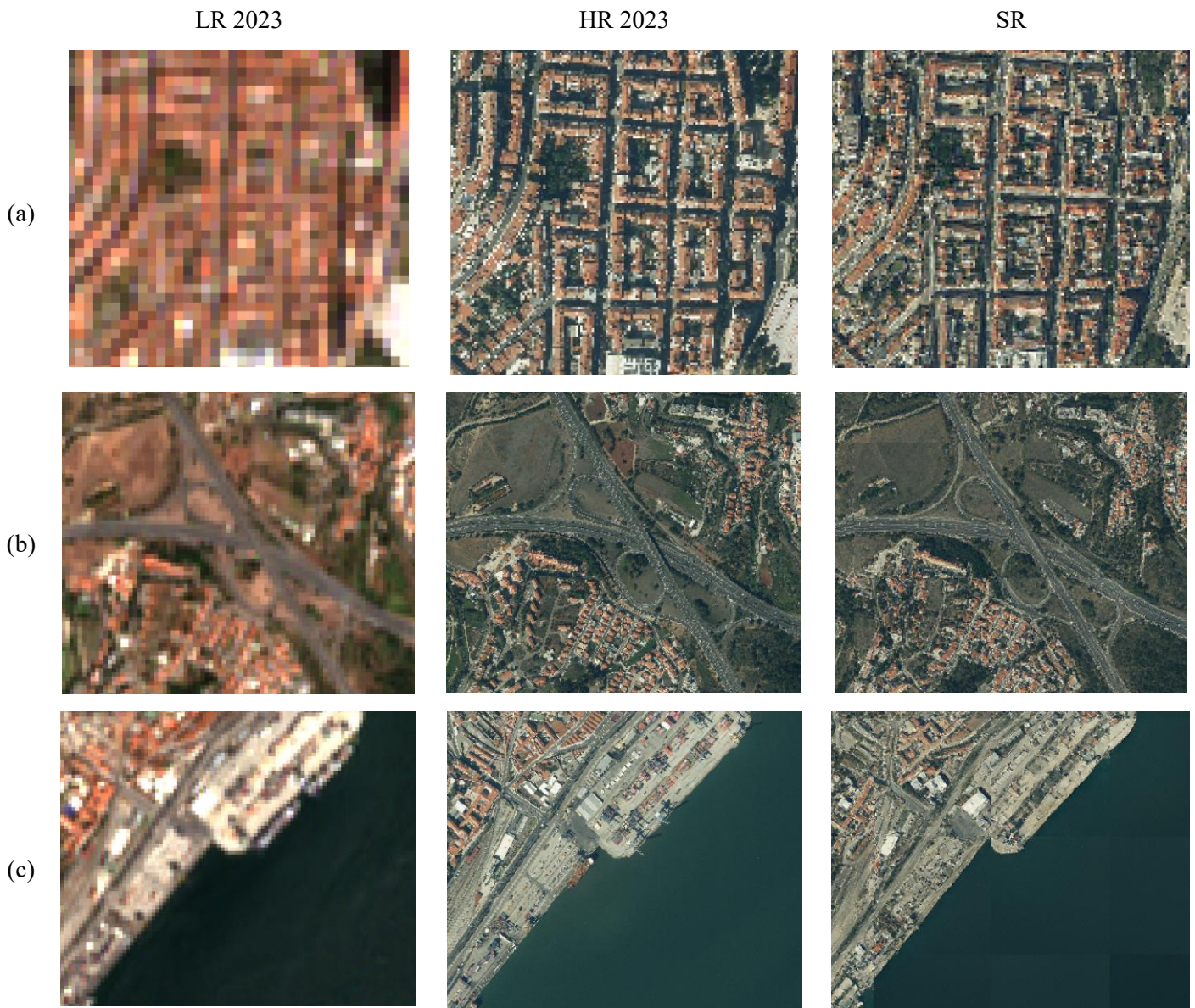


Figura 4.2 “Zooms” temáticos por classe (small multiples). (a) Edifícios; (b) Vias; (c) Interface água-terra.

De forma consistente nas três abordagens experimentais, a SR produz ganhos visuais claros face à LR. Em áreas densas de Lisboa, os limites de bairros exibem linhas mais contínuas, as coberturas e telhados surgem com nitidez superior e as discontinuidades entre blocos tornam-se discerníveis. Em vias estreitas, observa-se redução de discontinuidades em cruzamentos e maior regularidade nos alinhamentos, aproximando-se da leitura oferecida pela HR. Na classe água, as margens apresentam-se mais lisas e contínuas, com menor contaminação para os pixels adjacentes de edificado (redução da mistura em bordos). Em espaços verdes, o desenho de jardins e canteiros e a morfologia de taludes aparecem mais definidos, preservando o cromatismo esperado (Figura 4.3). Em Cascais, onde a textura natural e a frente costeira têm maior peso, a SR atenua a aparência esbatida/pouco contrastada típica do LR, ainda que os ganhos sejam mais subtis em superfícies homogêneas.

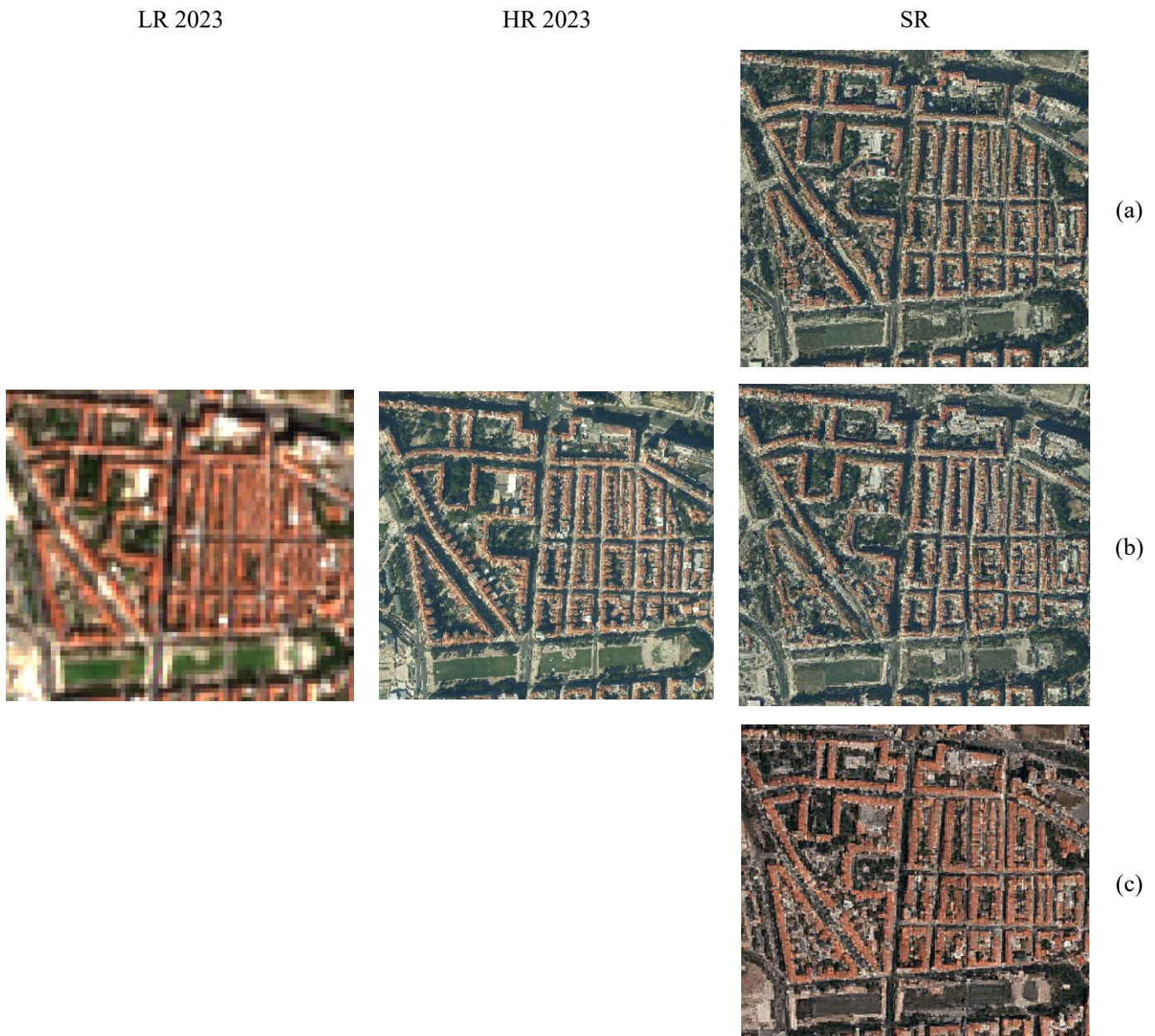


Figura 4.3 Repetição do painel em contextos distintos, Alameda em Lisboa. (a) Validação Interna (Lisboa); (b) Lisboa→Cascais; (c) 2018→2023.

A leitura qualitativa evidenciou também condições limite e modos de falha característicos de modelos adversários de super-resolução. Em zonas de sombra e superfícies com brilho especular (p. ex., água ondulada ou coberturas metálicas), a combinação de perdas perceptuais e adversárias pode realçar microtexturas que aumentam a impressão de nitidez sem corresponder integralmente à microgeometria real, um trade-off conhecido na literatura entre percepção e distorção (Ledig et al., 2017; Wang et al., 2018; Blau & Michaeli, 2018). Em raras situações, surgiram pequenos “halos” à volta de contornos muito contrastados e marcas em “quadrícula” nas transições muito abruptas, artefactos compatíveis com processos de *upsampling* aprendidos quando a regularização é insuficiente (Figura 4.4) (Odena, Dumoulin, & Olah, 2016). Estes efeitos são localizados e não comprometem a leitura global, mas recomendam prudência em análises que dependam exclusivamente de textura fina em zonas adversas (sombra forte ou reflexos).

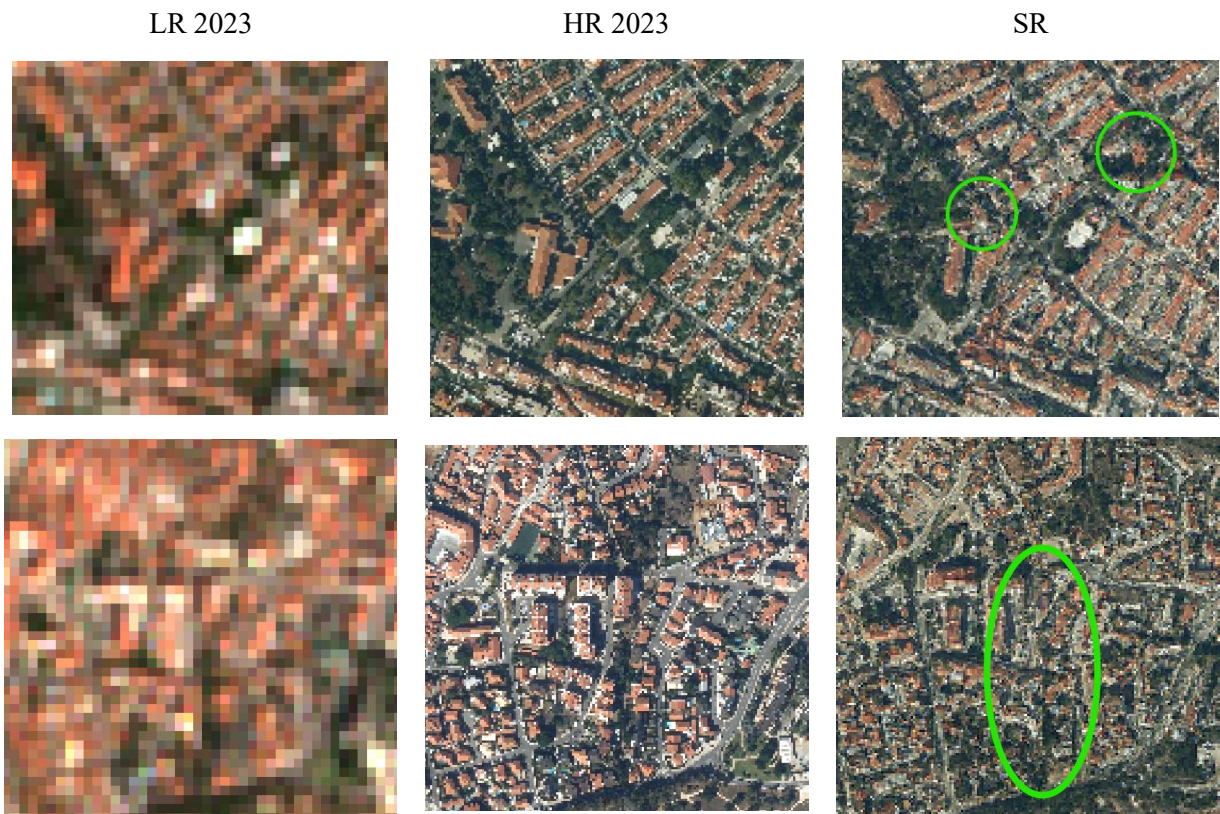


Figura 4.4 Casos-limite e artefactos localizados.

Especialmente, os ganhos visuais são heterogêneos e contexto-dependentes. Em malhas urbanas compactas de Lisboa, a SR aproxima-se substancialmente da imagem HR, sobretudo em vias e edifícios. Em zonas periurbanas e costeiras de Cascais, a melhoria é mais modesta e manifesta-se sobretudo nas interfaces (ex.: duna/edificado) e na clareza de limites (Figura 4.5).



Figura 4.5 Comparativo contextual Lisboa vs. Cascais (água-terra). (a) Lisboa→Cascais; (b) Cascais→Cascais.

Temporalmente, os resultados de 2018→2023 mantêm os benefícios relativos da SR, mas exigem cautela em áreas de obras recentes: discrepâncias localizadas entre o estado captado pela ortoimagem de referência e o conjunto de datas Sentinel-2 podem produzir pequenas incoerências visuais que não decorrem do modelo em si, mas do desfasamento temporal entre fontes. Um exemplo destas incoerências é a representação, em algumas zonas, de telhados brancos como se fossem telhados de tijolo, efeito associado à diferença temporal entre os dados e não a falhas do processo de super-resolução (Figura 4.6).

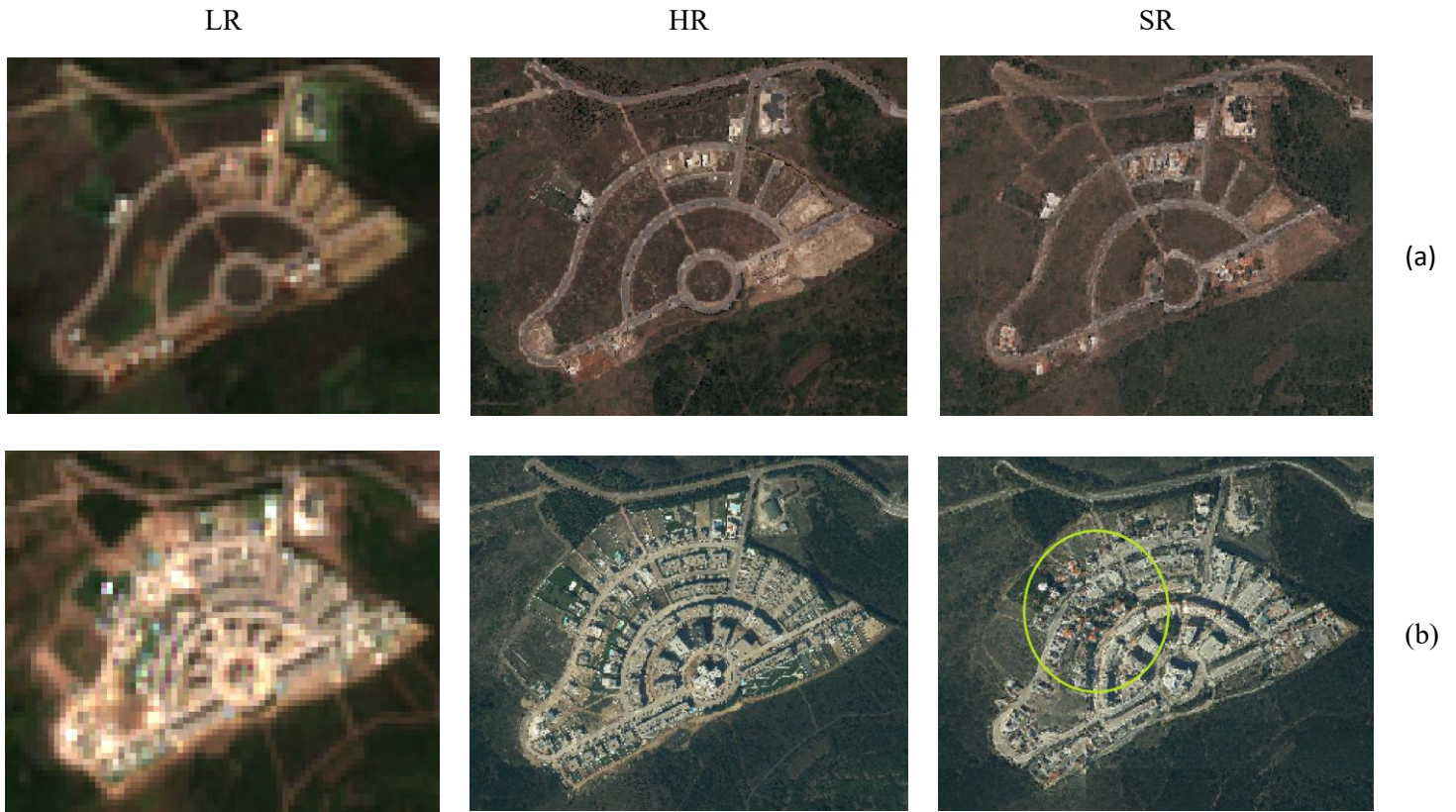


Figura 4.6 Par temporal controlado (2018 ↔ 2023) no Empreendimento SkyCity, Amadora. (a) 2018; (b) 2023.

Do ponto de vista cartográfico, a leitura georreferenciada permite afirmar que a super-resolução acrescenta valor operacional nas classes que estruturam a interpretação urbana: maior definição e continuidade da rede viária, contornos de edifícios mais fiéis e interfaces mais estáveis. Estes efeitos facilitam a vetorização manual, reduzem ambiguidades na delimitação de quarteirões e diminuem a mistura espectral aparente nos bordos – precisamente onde o LR é mais penalizado. Assim, mesmo sem recorrer a métricas numéricas, a evidência visual sustentada por painéis co registados e por sobreposição de polígonos é suficiente para demonstrar que a ESRGAN (Satlas-SR) produz melhorias consistentes e úteis para a leitura cartográfica em contexto urbano, respeitando a coerência espacial exigida pela engenharia geoespacial.

4.2 Leitura com polígonos de referência

A leitura apoiada em polígonos de referência foi concebida para responder, de forma controlada, à questão central desta dissertação: até que ponto o ganho de detalhe espacial produzido pela super-resolução melhora a legibilidade cartográfica de classes urbanas quando comparada com o produto original de baixa resolução (LR) e com a referência de alta resolução (HR). Para tal, selecionou-se a subárea de Braço de Prata/Parque Ribeirinho Oriente (PRO), um setor de Lisboa que combina malha urbana densa, frentes ribeirinhas e espaços verdes estruturados, oferecendo um contexto exigente para a leitura de arestas e de padrões lineares (Figura 4.7). Nesta subárea, foram vetorizados manualmente

polígonos sobre a ortomagem HR degradada para 2,5 m, representando quatro classes cartograficamente estruturantes: edifícios, vias, espaços verdes e água. Estes polígonos foram mantidos invariantes e sobrepostos, sem qualquer edição, às imagens LR e às imagens SR inferidas pelo modelo, assegurando uma comparação à mesma escala e no mesmo enquadramento geográfico.

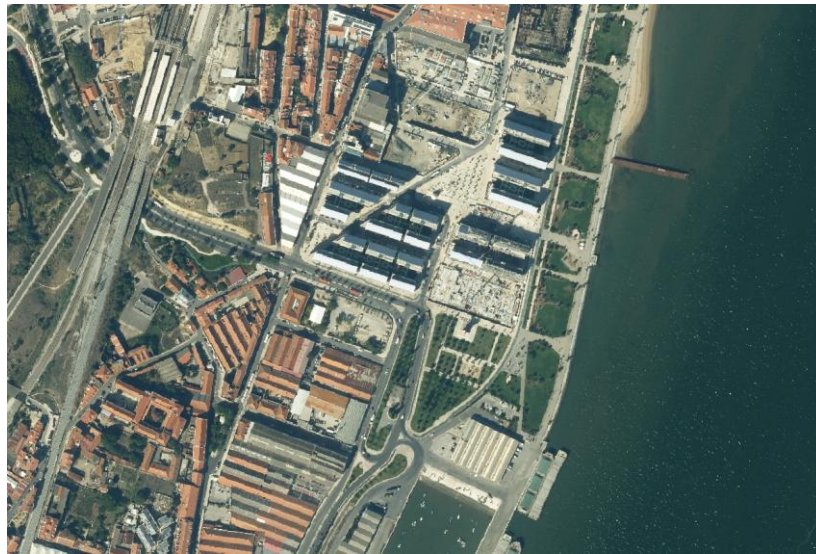






Figura 4.7 Enquadramento da subárea: Braço de Prata / PRO, 2023.

A inspeção visual foi novamente realizada em multiescala, combinando painéis LR → SR → HR de contexto com “zooms” em locais críticos. Em edifícios, observou-se uma aproximação consistente entre SR e HR na definição de edifícios e descontinuidades entre blocos, com redução do efeito de escada (bordos em escada, típico de LR) nas arestas oblíquas e nos cantos. As fronteiras dos polígonos de edificado apresentam, na SR, maior conformidade geométrica com os limites visíveis, o que facilita a interpretação e a vetorização subsequente. Em vias, particularmente nas ruas estreitas e nas interseções, a SR restabelece continuidade linear e reduz interrupções que surgem na LR por mistura espectral, aproximando o traçado observado ao padrão esperado em HR. Na água, as arestas água–terra mostram-se mais suaves e contínuas do que na LR e com menor contaminação para o edificado adjacente. Em espaços verdes, a SR melhora o contorno de jardins e taludes e clarifica interfaces vegetação–edificado, mantendo coerência cromática global (Figura 4.8).



Figura 4.8 Painel LR → HR → SR com polígonos sobrepostos (visão geral).

Tabela 4.1 Legenda das classes: edifícios (rosa), vias (laranja), espaços verdes (verde) e água (azul).

CLASSES	COR
EDIFICADO	 Rosa
VIAS	 Laranja
VEGETAÇÃO	 Verde
ÁGUA	 Azul

Esta leitura qualitativa foi replicada em todos os cenários experimentais, preservando a mesma geometria de referência para comparabilidade (Figura 4.9) e a mesma legenda de classes (Tabela 4.1). Na validação interna, onde treino e validação ocorrem no mesmo universo, o ganho visual é mais pronunciado e homogêneo, resultado expectável quando não há transferência espacial ou temporal. No ensaio de generalização Lisboa↔Cascais, o padrão mantém-se em edifícios e vias, embora com variação local em frentes costeiras e zonas periurbanas, onde a textura natural domina e o benefício da SR é mais subtil. No cenário 2018→2023, os ganhos persistem, mas o desfasamento temporal entre as datas das ortoimagens e as séries Sentinel-2 pode introduzir pequenas incoerências visuais nas imediações de obras recentes; nestes casos, diferenças aparentes entre SR e HR derivam sobretudo de mudança real e não do modelo.



Figura 4.9 Repetição controlada nas três abordagens. (a) Validação interna; (b) Cascais→Lisboa; (c) 2018→2023.

Para além da inspeção direta de contornos, recorreu-se à sobreposição dos polígonos para aferir estabilidade de fronteira. Na LR, as transições abruptas entre classes urbanas tendem a diluir-se em zonas de fronteira; na SR, o bordo dos polígonos coincide mais frequentemente com quebras radiométricas visíveis, reduzindo a ambiguidade espacial na delimitação de quarteirões e na leitura de eixos viários. Em termos cartográficos, isto traduz-se numa vetorização mais expedita e em menor necessidade de edição manual para corrigir descontínuos ou pequenos “saltos” de linha. Importa notar que esta melhoria visual não decorre de alterações de simbologia ou do pós-processamento, mas sim do próprio mecanismo adversarial + perceptual do gerador, concebido para reforçar texturas plausíveis e arestas salientes (Ledig et al., 2017; Wang et al., 2018).

Foram igualmente registados modos de falha pontuais: em sombras fortes de edifícios, a componente perceptual pode realçar texturas que aumentam a nitidez aparente sem representar necessariamente a geometria real; em arestas de alto contraste, surgiram ocasionalmente vestígios de “ringing” ou de padrões discretos associados ao *upsampling* aprendido (Odena, Dumoulin, & Olah, 2016). Estes efeitos revelaram-se localizados e não comprometeram a interpretação geral, mas recomendam prudência quando a decisão cartográfica depender exclusivamente de textura fina em condições radiométricas adversas.

No conjunto, a leitura com polígonos de referência demonstra que a SR acrescenta valor operacional nas classes que estruturam a cartografia urbana: os edifícios e as vias beneficiam de contornos mais fiéis

e de maior continuidade, a superfície de água apresenta margens mais estáveis e os espaços verdes revelam melhor separabilidade face ao edificado. Ao manter os polígonos invariantes, a avaliação isola o contributo do modelo, evidenciando que as melhorias observadas resultam do processo da super-resolução e não de artefactos de escala ou desalinhamento. Esta conclusão visual alinha-se com as expectativas teóricas para modelos SRGAN/ESRGAN e constitui um elo claro para a avaliação extensiva por classificação supervisionada apresentada na secção seguinte, onde se quantifica o impacto destes ganhos na discriminação temática (Congalton & Green, 2009; Foody, 2002).

4.3 Avaliação temática com Random Forest

A avaliação temática teve como objetivo quantificar em que medida a super-resolução acrescenta valor operativo à cartografia derivada do Sentinel-2, quando comparada com o produto original LR e com a referência HR. Para assegurar uma leitura justa do efeito “resolução”, a classificação supervisionada foi aplicada separadamente a cada domínio (LR, SR e HR), utilizando exatamente as mesmas amostras emparelhadas no espaço. Assim, cada píxel rotulado contribuiu com um triplo de observações (características RGB em LR, SR e HR) amostradas na mesma localização geográfica, o que isola o impacto da resolução espacial na capacidade discriminativa do classificador.





Evidenciando apenas os aspetos essenciais aos resultados, a avaliação temática baseou-se em 340 polígonos distribuídos por edifícios, vias, espaços verdes e água (Tabela 4.2), com amostragem ao nível do píxel empilhada espacialmente em LR, SR e HR; aplicou-se Random Forest com parametrização homogénea e partição 70/30 (treino/validação) em quatro cenários (validação interna, Lisboa↔Cascais, 2018→2023). Em cada cenário, correu-se uma experiência por domínio (LR, SR, HR) com partições sem interseção e zero partilha entre treino e validação, evitando *data leakage*.

Tabela 4.2 Inventário de amostras por classe.

CLASSES	Nº POLÍGONOS	% POLÍGONOS / CLASSE
ÁGUA	4	1
EDIFICADO	211	62
VEGETAÇÃO	36	11
VIAS	89	26
TOTAL	340	100

A legenda dos mapas de classificação resultantes do RF foi uniformizada para todos os cenários, de modo a garantir a comparabilidade entre resoluções. As quatro classes cartográficas de referência foram representadas com cores distintas: edifícios em vermelho, vias em cinzento, espaços verdes em verde e água em azul (Tabela 4.3). Esta convenção cromática facilita a interpretação visual e a análise temática, permitindo identificar rapidamente padrões de acerto e erro na classificação, bem como a influência da resolução espacial na separabilidade entre classes.

Tabela 4.3 Legenda das classes: edifícios (vermelho), vias (cinzento), espaços verdes (verde) e água (azul).

CLASSES	COR
EDIFICADO	 Vermelho
VIAS	 Cinzento
VEGETAÇÃO	 Verde
ÁGUA	 Azul

Perceciona-se, na leitura qualitativa dos mapas RF, um padrão coerente com o diagnóstico quantitativo: a super-resolução melhora a legibilidade temática sobretudo nas classes estruturais. Em edificado, a SR classifica arestas e volumetrias, reduzindo a aparência típica de 10 m e permitindo que o RF separe melhor telhados de espaços adjacentes. Nos “zooms” sobre quarteirões densos, observa-se maior continuidade dos limites dos edificios e delimitação mais limpa dos perímetros, com redução de mistura com vegetação contígua (Figura 4.10).

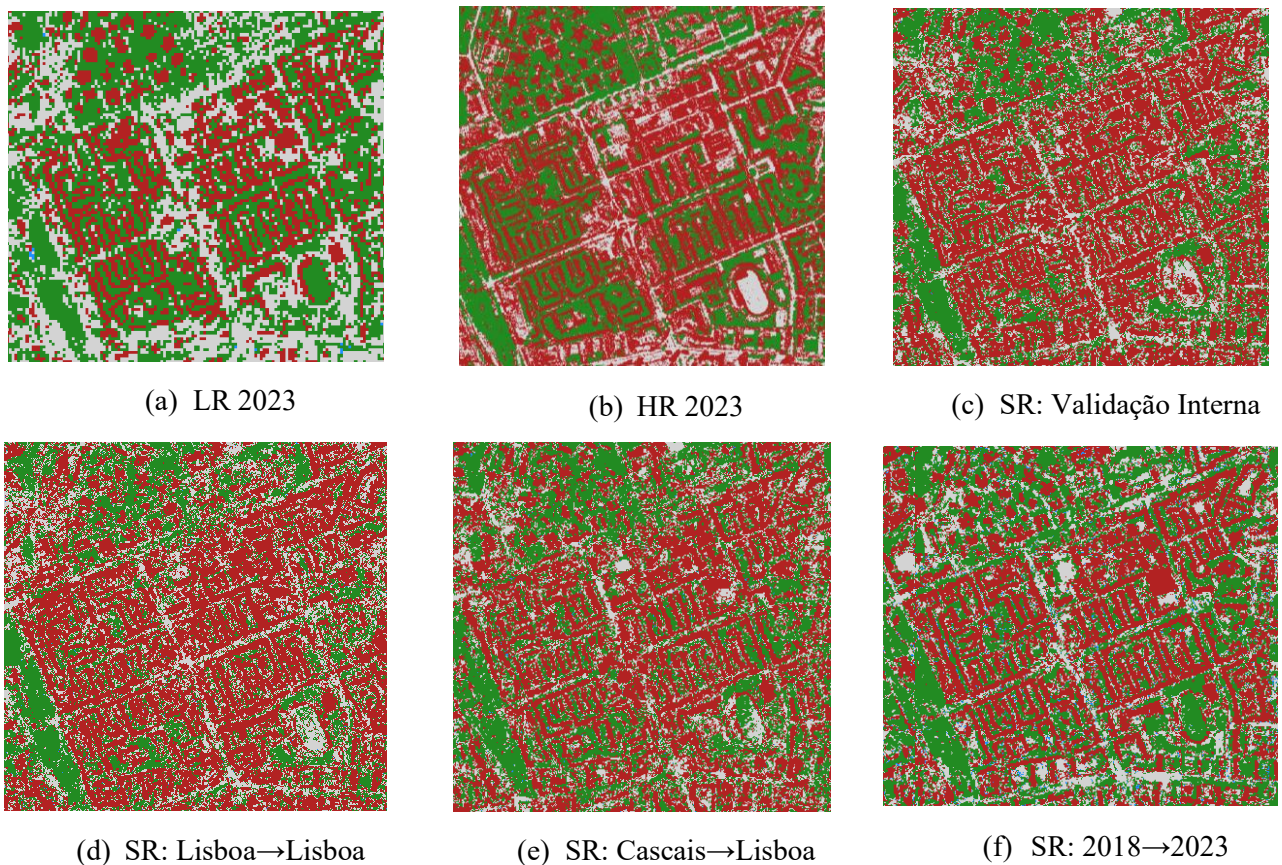


Figura 4.10 "Zooms" dos mapas RF nos diferentes cenários com ênfase no edificado.

Nas vias, a SR tem impacto particularmente visível: troços estreitos que em LR surgem fragmentados ou interrompidos passam a formar traçados contínuos, o que se traduz em menor omissão e melhor conectividade topológica. Em interseções e rotundas, os “zooms” mostram que a SR suaviza quebras de continuidade e ajuda o classificador a distinguir pavimento de taludes ou coberturas claras. Em

contrapartida, regista-se por vezes um ligeiro aumento de erros de classificação em superfícies muito claras (estacionamento, telhados metálicos ou betão exposto), sobretudo quando o brilho especular induz padrões de alto contraste (Figura 4.11).

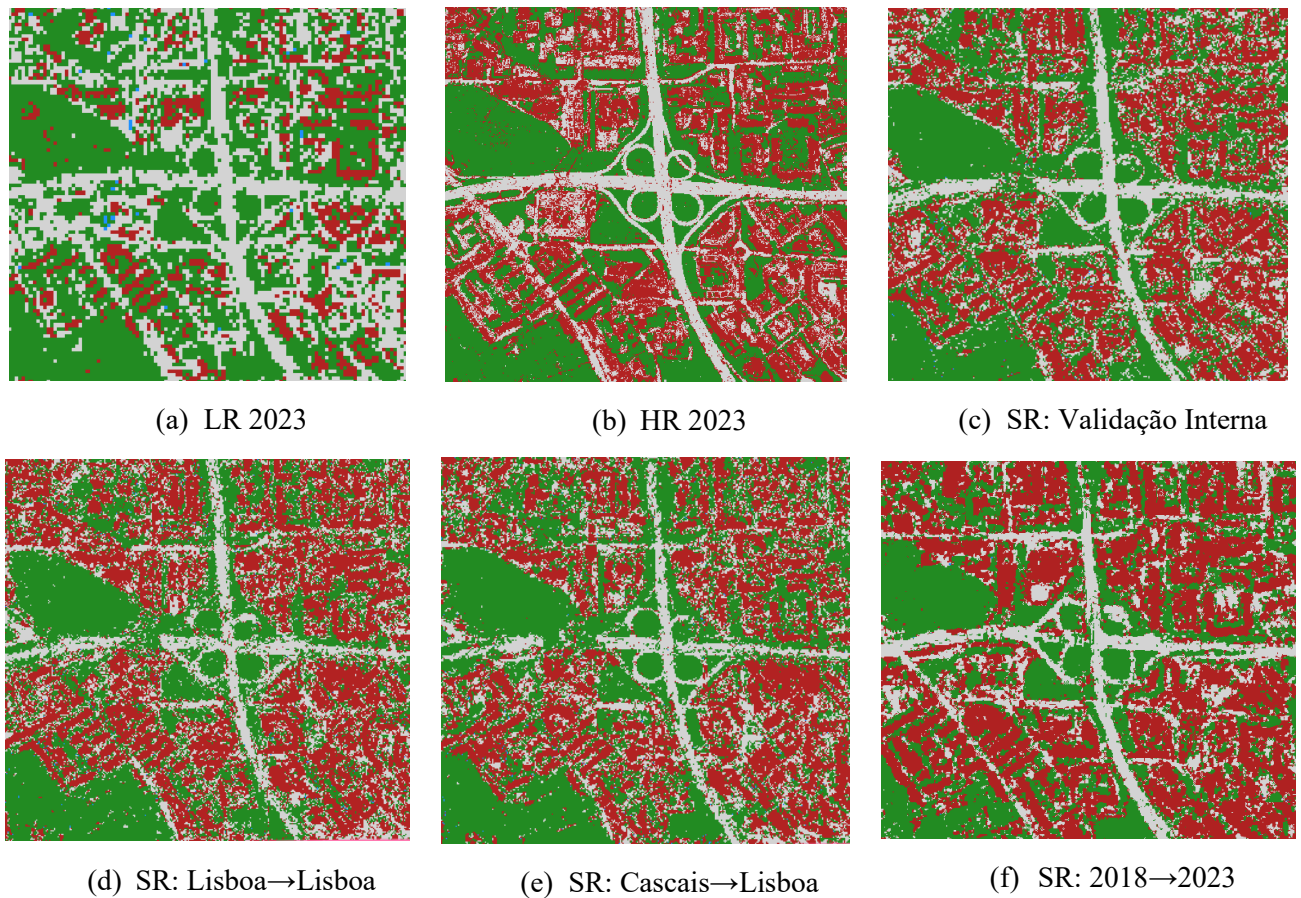


Figura 4.11 "Zooms" dos mapas RF nos diferentes cenários com ênfase nas vias.

A vegetação mantém um comportamento estável entre a LR e a SR. Nas manchas arbóreas compactas, a classificação é densa e coerente em ambos os produtos; em alinhamentos arbóreos ao longo de vias, a SR melhora a leitura das copas, mas o ganho temático é discreto. Os “zooms” em ruas arborizadas evidenciam que a fronteira árvore-pavimento é mais definida em SR, embora permaneça alguma confusão em sombras e em vegetação muito seca, onde a assinatura espectral se aproxima de solos nus (Figura 4.12).

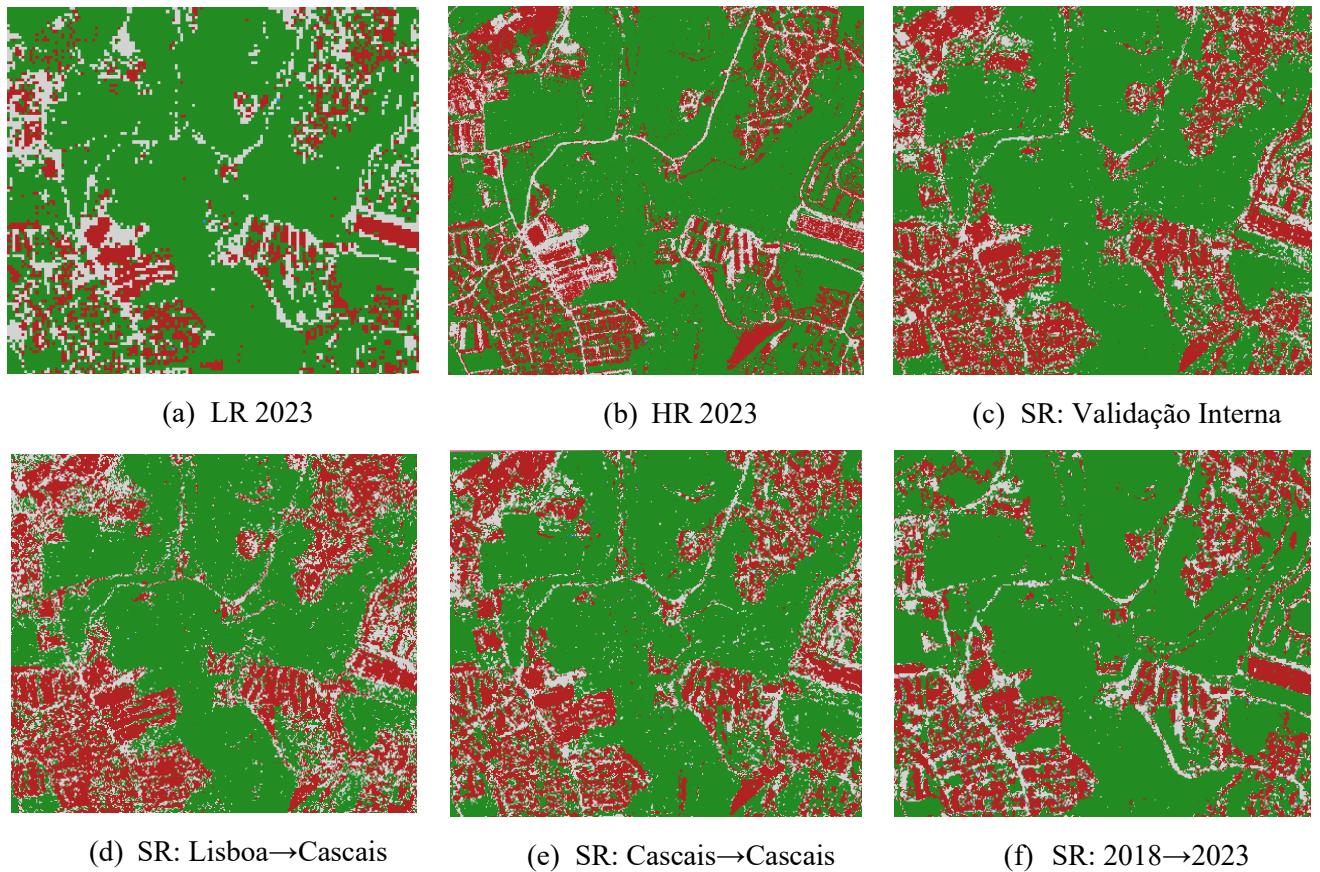


Figura 4.12 "Zooms" dos mapas RF nos diferentes cenários com ênfase na vegetação.

Na água, a leitura qualitativa explica o reduzido valor de F1-score (Tabela 4.4): as margens água-terra são zonas de alto contraste, com sombras e reflexos que a SR tende a enfatizar. Em zonas de docas e frente ribeirinha, notam-se pequenos trechos onde o espelhado da água adquire microtextura, levando o RF a sobreclassificar água em zonas de cais ou a confundir reflexos com superfícies pavimentadas muito escuras (Figura 4.13). Este efeito é localizado e não compromete a interpretação global, mas recomenda prudência na leitura de interfaces hidrológicas e, se necessário, a aplicação de um ligeiro pós-processamento nas margens.

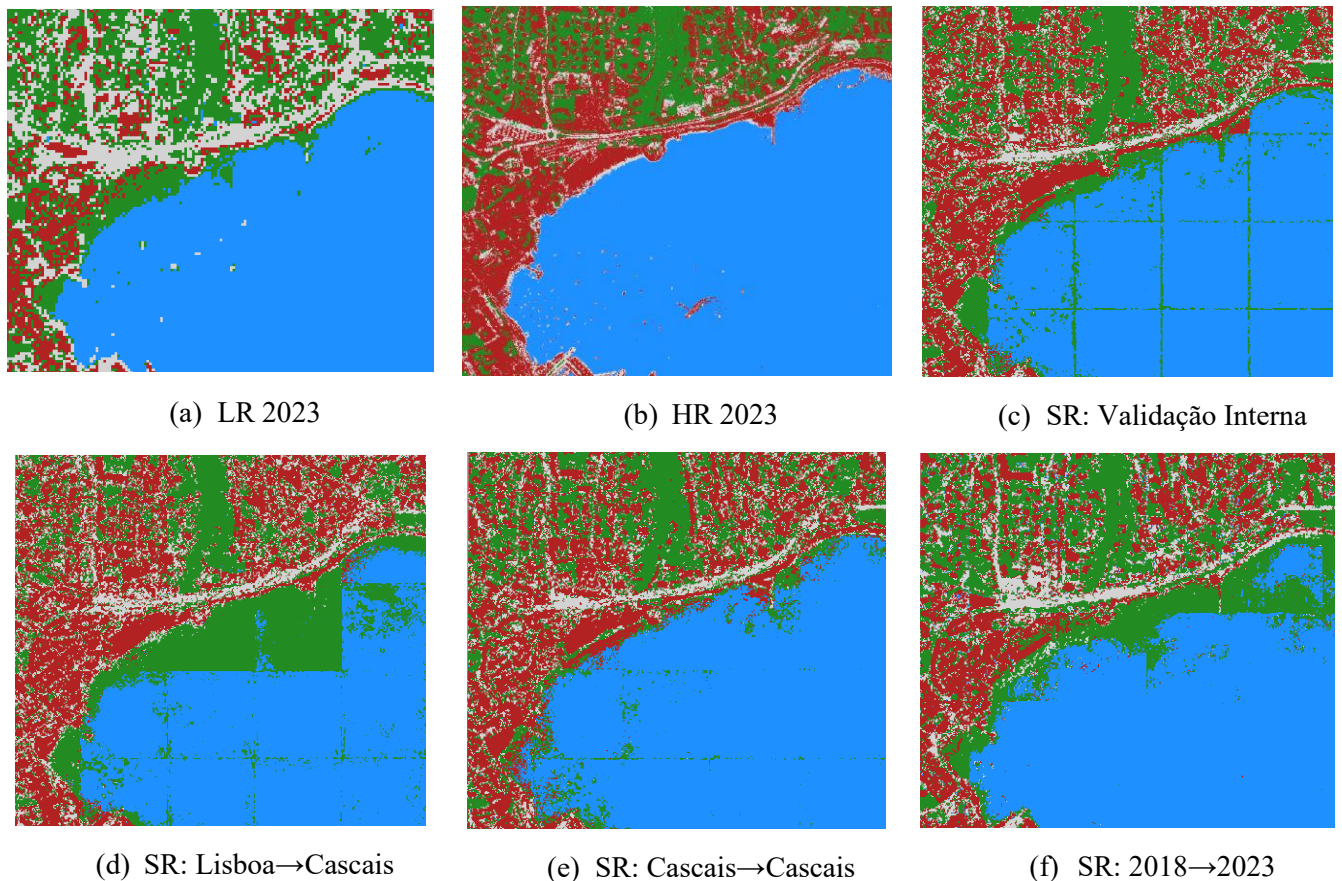


Figura 4.13 “Zooms” dos mapas RF nos diferentes cenários com ênfase na água.

Em transferência espacial (Lisboa→Cascais e inverso), a SR preserva os benefícios em edificado e vias, ainda que com intensidade um pouco menor, o que se percebe bem nos “zooms” comparáveis entre malha urbana densa e frentes periurbanas/costeiras. Em Cascais, onde a textura natural e a presença de superfícies arenosas são mais frequentes, a SR continua a ajudar na continuidade viária, mas pode introduzir microtextura em praias ou esplanadas, exigindo leitura atenta nesses contextos.

Finalmente, importa notar que alguns artefactos de SR são detetáveis nos “zooms” de bordos de altíssimo contraste: pequenas auréolas ao longo das arestas junto a cumeeiras muito claras e padrões discretos em telhados metálicos. Estes efeitos são pontuais e não dominam o mapa, mas explicam comissões residuais em vias/edificado e em margens de água. Do ponto de vista operacional, a combinação de SR com uma regra de pós-processamento leve (por exemplo, suavização morfológica nas classes) tende a conservar os ganhos de continuidade e aresta, mitigando os poucos casos de sobreclassificação.

Em síntese, a leitura qualitativa confirma a narrativa principal: a SR aproxima o comportamento temático ao de HR onde a geometria é crítica (edificado, vias), mantém-se neutra na vegetação e exige cautela localizada na água, sobretudo em frentes ribeirinhas com reflexos e sombras. Esta coerência entre “olho humano” e métricas por classe sustenta o valor cartográfico da SR em contexto urbano (Belgiu & Drăguț, 2016; Congalton & Green, 2009; Foody, 2002; Maxwell et al., 2018).

A análise foi ainda estratificada por concelho (Lisboa, Cascais) e por época (2018, 2023) para testar, respetivamente, a transferência espacial e a robustez temporal do modelo. No cruzamento Lisboa→Cascais (e inverso), onde as morfologias contrastam (malha densa vs. frente

costeira/periurbana), compara-se o F1-score por classe para verificar se os ganhos de SR persistem fora do domínio de treino. No par temporal 2018→2023, a repetição do protocolo sobre as mesmas geometrias de referência permite isolar diferenças devidas a alterações reais da paisagem (novas construções, requalificações) de flutuações de desempenho do classificador, reforçando a validade externa. A avaliação quantitativa é reportada a nível global e por classe com as métricas padrão, precisão global, precisão, recall e F1-score (Tabela 4.4), como indicadores de estabilidade; perante a assimetria de amostras (predomínio de edifícios), privilegia-se o F1 por classe e a leitura de confusões estruturais típicas em meio urbano (vias↔edifícios; água↔espaços verdes em margens), em linha com as boas práticas de avaliação temática (Congalton & Green, 2009; Foody, 2002).

Tabela 4.4 Exatidão de LR, de HR e de SR nas três abordagens. (a) LR; (b) HR; (c) SR: Validação interna; (d) SR: Lisboa→Cascais; (e) SR: Cascais→Lisboa; (f) SR: 2018→2023.

(a) Imagem Sentinel 2023

CLASSE	PRECISÃO	RECALL	F1-SCORE
ÁGUA	0,86	0,43	0,58
EDIFICADO	0,87	0,42	0,57
VEGETAÇÃO	0,60	0,99	0,75
VIAS	0,92	0,31	0,47
PRECISÃO GLOBAL		0,67	

(b) Ortoimagem 2023

CLASSE	PRECISÃO	RECALL	F1-SCORE
ÁGUA	1,00	0,67	0,80
EDIFICADO	0,83	0,97	0,72
VEGETAÇÃO	0,98	0,64	0,77
VIAS	0,92	0,65	0,62
PRECISÃO GLOBAL		0,70	

(c) Super-Resolução Lisboa + Cascais

CLASSE	PRECISÃO	RECALL	F1-SCORE
ÁGUA	0,81	0,56	0,56
EDIFICADO	0,80	0,82	0,68
VEGETAÇÃO	0,89	0,64	0,74
VIAS	0,70	0,77	0,54
PRECISÃO GLOBAL		0,62	

(d) Super-Resolução Lisboa → Cascais

CLASSE	PRECISÃO	RECALL	F1-SCORE
ÁGUA	0,77	0,38	0,53
EDIFICADO	0,80	0,51	0,69
VEGETAÇÃO	0,82	0,65	0,76
VIAS	0,65	0,63	0,51
PRECISÃO GLOBAL		0,61	

(d) Super-Resolução Cascais → Lisboa

CLASSE	PRECISÃO	RECALL	F1-SCORE
ÁGUA	0,87	0,39	0,54
EDIFICADO	0,78	0,70	0,66
VEGETAÇÃO	0,80	0,63	0,73
VIAS	0,66	0,86	0,53
PRECISÃO GLOBAL		0,59	

(e) Super-Resolução 2018 → 2023

CLASSE	PRECISÃO	RECALL	F1-SCORE
ÁGUA	0,94	0,64	0,76
EDIFICADO	0,55	0,93	0,69
VEGETAÇÃO	0,83	0,58	0,72
VIAS	0,75	0,49	0,51
PRECISÃO GLOBAL		0,64	

Além das métricas de exatidão, apresenta-se uma tabela de ganhos e um gráfico de ganhos construídos a partir dos resultados do RF, com a LR como referência. Estes dois elementos quantificam, por cenário e por classe, a diferença entre a SR e a LR (e entre HR e LR) na principal métrica, F1-score por classe, permitindo verificar de forma direta onde e quanto a super-resolução acrescenta valor face ao produto de 10 m. A tabela (Tabela 4.5) resume os ganhos por classe e cenário, enquanto o gráfico (Figura 4.14) visualiza esses ganhos, facilitando a leitura comparativa entre validação interna,

transferência espacial e robustez temporal, e evidenciando a aproximação de SR ao comportamento de HR nas classes mais críticas em contexto urbano.

Tabela 4.5 Ganhos F1-score por classe vs LR.

CENÁRIO	Δ F1-SCORE ÁGUA	Δ F1-SCORE EDIFICADO	Δ F1-SCORE VEGETAÇÃO	Δ F1-SCORE VIAS
HR	0,22	0,15	0,02	0,15
SR – VALIDAÇÃO INTERNA	-0,02	0,11	-0,01	0,07
SR - LISBOA→ CASCAIS	-0,05	0,12	0,01	0,04
SR - CASCAIS→ LISBOA	-0,04	0,09	-0,02	0,06
SR - 2018→2023	0,18	0,12	-0,03	0,04

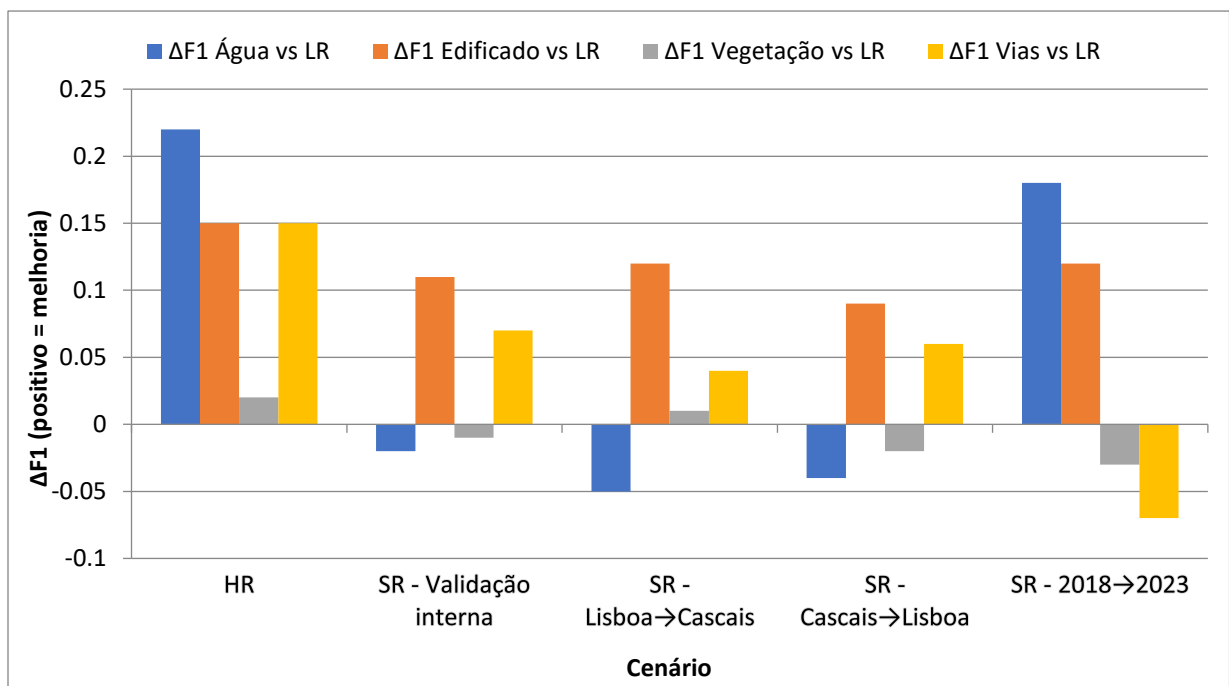


Figura 4.14 Gráfico de ganhos F1-score por classe vs LR.

Os resultados da classificação RF evidenciam padrões consistentes com a literatura em detecção remota urbana: classes estruturais e lineares beneficiam desproporcionalmente do aumento de detalhe espacial, ao passo que classes espectralmente homogêneas mostram menor variação (Breiman, 2001; Belgiu & Drăguț, 2016; Maxwell, Warner, & Fang, 2018). Na imagem de referência LR (10 m), a precisão global é 0,67, com desempenho díspar por classe (F1-score): Água = 0,58, Vegetação = 0,75, mas Edificado = 0,57 e Vias = 0,47. Este retrato é típico de resolução média em contexto urbano: massas de água e vegetação são relativamente fáceis de separar, enquanto mistura espectral e subamostragem penalizam arestas de edifícios e a continuidade de vias.

A referência HR (2,5 m) confirma o “teto” alcançável pelo RF quando a informação espacial é suficiente: a precisão global aumenta para 0,70 (sensível à distribuição de amostras), e observam-se ganhos substanciais em F1-score nas classes de maior interesse cartográfico: Água sobe para 0,80 (+0,22), Edificado sobe para 0,72 (+0,15) e Vias para 0,62 (+0,15), mantendo valores ligeiramente superiores para Vegetação (0,77, +0,02). Ou seja, a melhoria estruturante ocorre onde a resolução espacial é crítica: arestas mais nítidas e traçados viários mais contínuos traduzem-se em separabilidade temática superior (Congalton & Green, 2009; Foody, 2002).

Nos cenários SR, as tendências mantêm-se com nuances importantes. Na validação interna, Edificado atinge 0,68 (+0,11 vs. LR) e Vias 0,54 (+0,07), com Vegetação estável (0,77, 0,01) e Água a descer ligeiramente (0,56, -0,02). Na transferência espacial (Lisboa→Cascais e Cascais→Lisboa), os ganhos em Edificado (0,69 e 0,54; +0,12 e +0,09) e Vias (0,51 e 0,53; +0,04 e +0,06) persistem, refletindo a diferença morfológica entre malhas urbanas (densa vs. periurbana/costeira). Na robustez temporal (2018→2023), salienta-se o maior ganho observado na classe Água (0,76; +0,18), sinalizando que o reforço de arestas e texturas urbanas em SR é resiliente a mudança real da paisagem. Nesta mesma abordagem, Vias mantêm-se próximas da linha de base (0,51, +0,04); Água mantém-se elevada (0,76; +0,18) e Vegetação desce ligeiramente (0,72; -0,03). Em suma, Edificado e Vias apresentam ganhos consistentes face a LR nas SR e, especialmente, no par temporal para Edificado; Vegetação é neutra/estável; Água tende a perder por efeito de microtextura adicional junto a margens, o que pode produzir classificações indevidas em zonas pontuais, um resultado coerente com as observações visuais de interfaces água-terra. A perda de qualidade na identificação seria esperada, uma vez que na super-resolução foram utilizadas apenas as três bandas do RGB, não incluindo a banda do infravermelho.

A leitura destes resultados deve privilegiar métricas por classe (F1-score), dado que a precisão global é sensível ao não balanceamento entre classes e pode mascarar ganhos pontuais nas classes mais críticas para cartografia urbana (vias estreitas e arestas de edifícios). Em termos operacionais, a super-resolução aproxima o comportamento temático ao observado em HR nas classes estruturais, melhorando a legibilidade cartográfica (contornos, continuidade linear, separação de coberturas adjacentes) e, portanto, a utilidade dos produtos para vetorização e monitorização de mudança. As perdas observadas na classe Água recomendam, tal como no caso da vegetação, a utilização da banda do infravermelho que diferencia claramente a água e a vegetação de outras ocupações do solo.

No conjunto, a evidência empírica aponta para um benefício cartográfico tangível da SR face a LR: ganhos sistemáticos em Edificado e Vias em todos os cenários espaciais, robustez temporal especialmente marcada no edificado e neutralidade na vegetação. Estas conclusões são coerentes com o papel do RF enquanto classificador robusto e colinearidade, capaz de capitalizar informação espacial adicional quando esta é disponibilizada (Breiman, 2001; Belgiu & Drăguț, 2016; Maxwell et al., 2018), e com as boas práticas de avaliação temática, que recomendam análise por classe e matrizes de confusão sob partições sem fuga de informação (Congalton & Green, 2009; Foody, 2002).

É importante reconhecer limitações e ameaças à validade. Como referido, a utilização de apenas RGB (TCI) pode limitar a separabilidade em certas transições (p. ex., tipos de vegetação ou superfícies

artificiais com assinaturas espectrais semelhantes no visível). Além disso, a desproporção entre classes (mais edifícios) pode favorecer a precisão global; por isso, privilegiou-se F1-score por classe. Finalmente, a amostragem ao nível do píxel não elimina toda a autocorrelação espacial; no entanto, o emparelhamento rigoroso LR–SR–HR e a repetição por cenários espaciais/temporais mitigam vieses sistemáticos e sustentam a inferência sobre o efeito da resolução.

Em síntese, a avaliação temática com Random Forest fornece um quadro quantitativo coerente com a evidência visual: a super-resolução melhora a discriminação de vias e edifícios, estabiliza interfaces críticas e aproxima o desempenho do produto LR ao observado em HR, com ganhos consistentes nos diferentes cenários. Estes resultados demonstram que os benefícios perceptivos da SR se traduzem em utilidade cartográfica mensurável, respondendo ao terceiro objetivo desta dissertação e reforçando a viabilidade do uso de GAN como etapa de pré-processamento para mapeamento urbano com Sentinel-2.

4.4 Métricas de avaliação de reconstrução

As métricas de reconstrução foram usadas para quantificar, de forma objetiva, a proximidade entre os produtos de super-resolução (SR) e a referência de alta resolução (HR), complementando a leitura visual e a avaliação temática. As métricas MAE e MSE medem o erro radiométrico médio; SSIM e PSNR capturam, respetivamente, a semelhança estrutural e a visibilidade do erro numa escala logarítmica. Em cartografia urbana, onde arestas e padrões finos dominam (limites de edifícios, cruzamentos, interfaces água–terra), estas métricas são particularmente informativas, devendo, contudo, ser interpretadas à luz do compromisso percepção–distorção. As melhorias em PSNR/SSIM nem sempre se traduzem linearmente em ganho perceptivo e vice-versa (Wang, Bovik, Sheikh, & Simoncelli, 2004; Blau & Michaeli, 2018). Aqui, a implementação é mantida ao essencial: as métricas foram calculadas por mosaico, à mesma escala de saída (2,5 m), em RGB, como referência HR–LR e HR–SR como efeito do modelo, para todos os cenários do plano experimental (Tabela 4.6).

Tabela 4.6 Valores das métricas em todos os cenários.

	HR - LR	HR - SR: Validação Interna	HR - SR: Lisboa→Cascais	HR - SR: Cascais→Lisboa	HR - SR: 2018→2023
PSNR	10,71	13,73	13,71	13,42	12,64
SSIM	0,07	0,25	0,25	0,25	0,15
MSE	5525,43	2753,65	2766,97	2959,56	3540,75
MAE	56,57	35,24	35,49	36,54	44,93

A referência HR–LR estabelece o nível de partida: PSNR = 10,71 dB e SSIM = 0,072, com MSE = 5525,43 e MAE = 56,57. Estes valores refletem a discrepância esperada entre um produto de 10 m reamostrado e a referência a 2,5 m, sobretudo em zonas de alto contraste onde a mistura espectral em LR é mais penalizadora. Com a aplicação da super-resolução, observa-se um salto consistente em todas as métricas (Tabela 4.6; Tabela 4.7; Figura 4.15; Figura 4.16). Na validação interna, PSNR aumenta para 13,73 dB e SSIM para 0,254, enquanto MSE e MAE descem para 2753,65 e 35,24. Em termos diferenciais face à linha de base, isto traduz-se em +3,03 dB (PSNR) e +0,182 (SSIM), com reduções de –2771,8 (MSE) e –21,33 (MAE). A magnitude destes ganhos confirma que a SR aproxima radiometricamente e estruturalmente o mosaico reconstruído do alvo HR, sobretudo pela recuperação de arestas e texturas urbanas.

Em transferência espacial, os benefícios mantêm-se com grande estabilidade, ainda que ligeiramente atenuados pela mudança de morfologia entre regiões. No cenário Lisboa→Cascais, obtêm-se PSNR = 13,71 dB e SSIM = 0,251 (ganhos de +3,00 dB e +0,179 face a LR), com MSE = 2766,97 e MAE = 35,49 (reduções de -2758,5 e -21,08). No sentido Cascais→Lisboa, os valores são PSNR = 13,42 dB e SSIM = 0,253 (+2,71 dB; +0,181), com MSE = 2959,56 e MAE = 36,54 (-2565,9; -20,03). A leitura é coerente com a interpretação geoespacial: a SR continua a reforçar a definição de bordos e a continuidade linear em contextos urbanos distintos, apesar de diferenças de textura natural e padrões costeiros.

No eixo temporal (2018→2023), a melhoria mantém-se, mas com amplitude inferior, como é expectável perante mudanças reais da paisagem e possíveis desfasamentos temporais finos entre as ortoimagens de referência e as aquisições Sentinel-2. Os resultados agregados mostram PSNR = 12,64 dB e SSIM = 0,145 (+1,93 dB; +0,073), com MSE = 3540,75 e MAE = 44,93 (-1984,7; -11,64). Apesar de menores, estes ganhos continuam positivos e operacionais: a recuperação de arestas e a redução de mistura em bordos persistem, sobretudo no edificado e na rede viária, enquanto superfícies muito homogêneas (areais extensos) ou dinâmicas (reflexos especulares, sombra dura) geram, pontualmente, penalizações métricas que não anulam a tendência global.

Tabela 4.7 Ganhos Δ de SR face a LR por cenário e métrica.

Cenário	Δ PSNR (dB)	Δ SSIM	Δ MSE	Δ MAE
SR - Validação interna	3,025	0,1816	-2771,8	-21,332
SR - Lisboa→Cascais	3,004	0,1791	-2758,5	-21,076
SR - Cascais→Lisboa	2,711	0,1807	-2565,9	-20,033
SR - 2018→2023	1,933	0,0729	-1984,7	-11,639

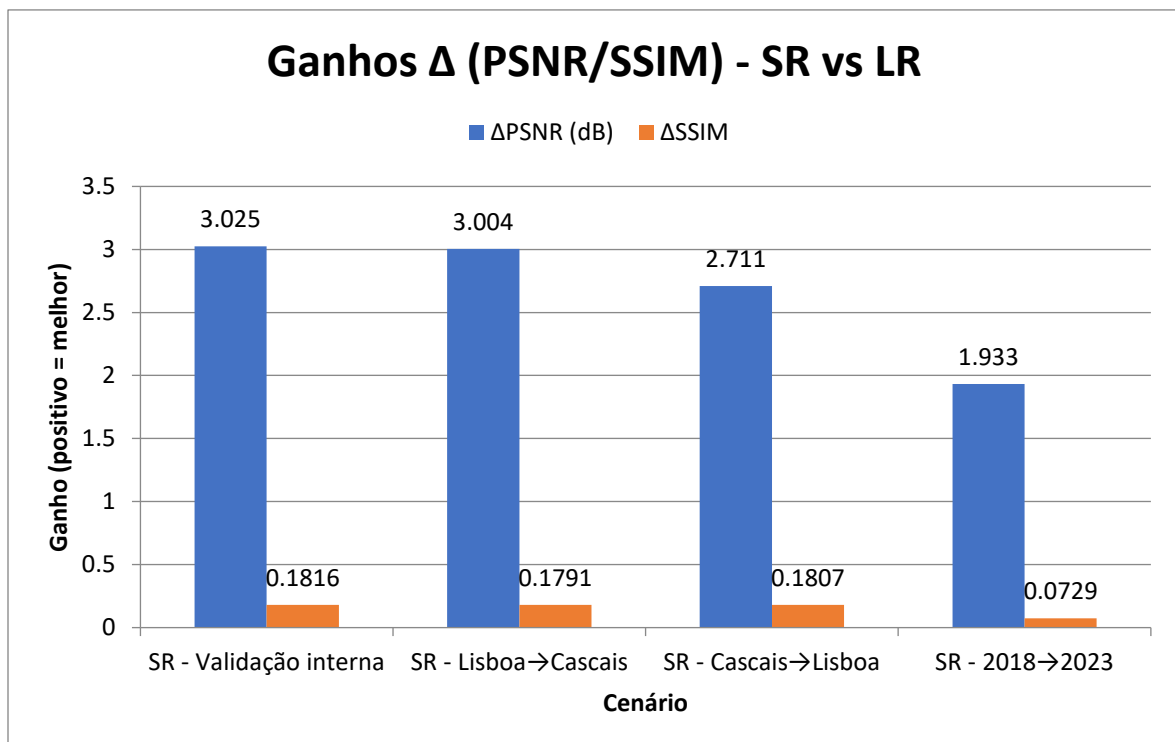


Figura 4.15 Ganhos Δ (PSNR/SSIM) - SR vs LR.

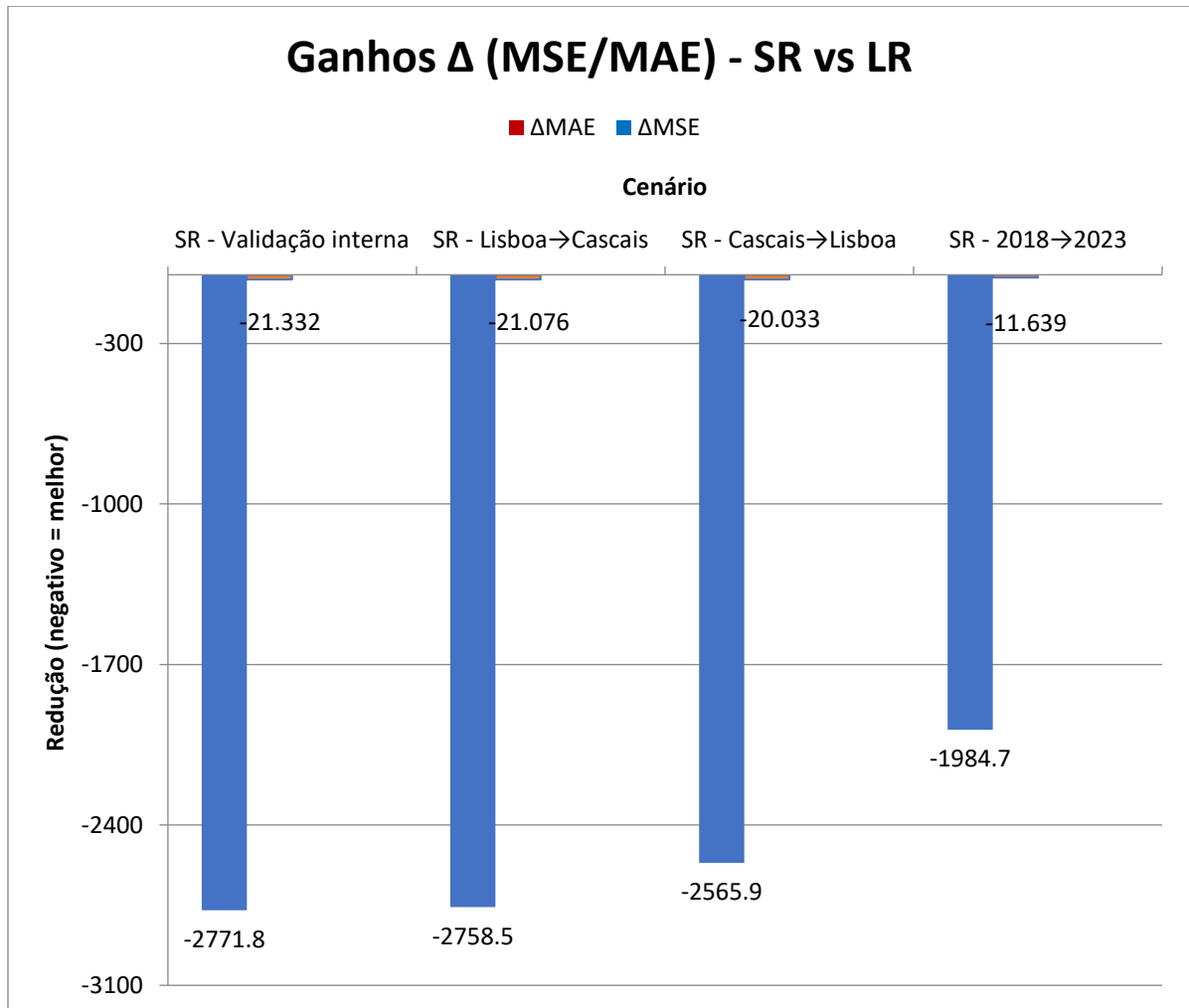


Figura 4.16 Ganhos Δ (MSE/MAE) - SR vs LR.

Em conjunto, os resultados evidenciam três pontos centrais. Primeiro, consistência: a SR melhora todos os indicadores em todos os cenários, com ganhos máximos em validação interna, robustos na transferência espacial e positivos no cenário temporal. Segundo, coerência com a leitura geoespacial: os maiores benefícios surgem onde a geometria conta: arestas de edifícios, continuidade de vias, interfaces abruptas, precisamente as zonas onde a utilidade cartográfica é mais sensível ao detalhe espacial. Terceiro, limitações mensuráveis: variações locais em PSNR/SSIM/MSE/MAE são explicáveis por fatores não modelares (desalinhamentos ao nível do píxel, desfasamento temporal, reflexos/ sombras), justificando a interpretação conjunta com a evidência visual e com a avaliação temática. Assim, as quatro métricas convergem na mesma conclusão: a super-resolução aproxima, de modo mensurável e consistente, o produto Sentinel-2 da referência HR nos domínios críticos para cartografia urbana, sustentando os ganhos de legibilidade e separabilidade reportados noutros subcapítulos (Wang et al., 2004; Blau & Michaeli, 2018).

4.5 Discussão integrada por objetivos

A discussão dos resultados é aqui organizada em torno dos três objetivos definidos, articulando a evidência visual, a avaliação temática e a análise dos registos de treino no quadro da engenharia geoespacial. No primeiro objetivo, desenvolver e avaliar um modelo de super-resolução baseado em GAN para aumentar a resolução espacial do Sentinel-2 a 2,5 m preservando a consistência espectral, os resultados confirmam a viabilidade técnica da abordagem ESRGAN adotada via Satlas-SR. As

comparações corregistadas LR→SR→HR mostram ganhos estáveis de nitidez e continuidade geométrica, sobretudo em malhas urbanas densas, sem evidência de deriva radiométrica sistemática quando a SR é observada à mesma escala da referência. Reconhecem-se, ainda assim, os compromissos clássicos entre percepção e distorção em GANs: em zonas de sombra ou superfícies com brilho especular ocorrem pontualmente microtexturas e pequenos halos ao longo das arestas (“ringing”); são efeitos localizados que não comprometem a leitura cartográfica global, mas que justificam uma seleção criteriosa da versão do modelo durante o treino, de modo a conciliar qualidade perceptual com fidelidade geométrica (Ledig et al., 2017; Wang et al., 2018; Blau & Michaeli, 2018).

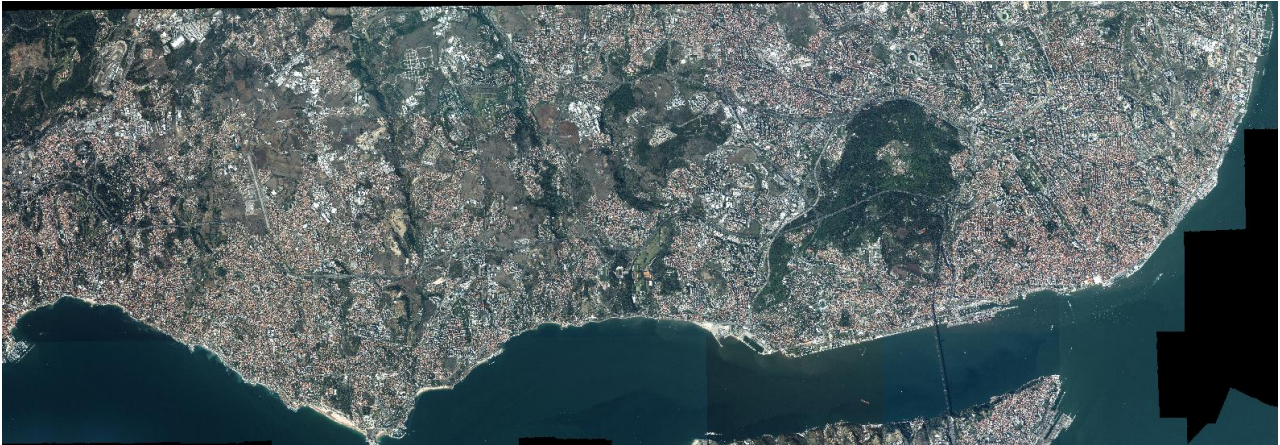
Relativamente ao segundo objetivo, validar a eficácia do modelo em cenários urbanos distintos comparando com a referência HR, o desenho experimental permitiu testar generalização espacial (Lisboa↔Cascais) e robustez temporal (2018→2023) sob controlo geoespacial rigoroso. A transferência entre concelhos evidencia que a SR mantém vantagens face ao LR quando a morfologia urbana muda, com benefício mais expressivo em edifícios e vias e mais moderado em frentes costeiras e zonas periurbanas, onde a textura natural domina. A robustez temporal confirma que, mesmo perante alterações reais da paisagem, a SR preserva o ganho relativo, ainda que com maior variabilidade nas imediações de obras recentes, onde a não coincidência temporal entre ortoimagens e imagens S2 podem introduzir diferenças que não decorrem do modelo. Em ambos os casos, o controlo de *data leakage* espacial/temporal, o corregistro à escala do píxel e o uso de HR degradada para 2,5 m como *ground-truth* sustentam a comparabilidade e reforçam a validade externa das conclusões (Goodfellow et al., 2014; Drusch et al., 2012; ESA, 2021).

No terceiro objetivo, analisar o impacto da melhoria de resolução na classificação da ocupação do solo, os resultados convergem entre a avaliação quantitativa com Random Forest e a leitura qualitativa com polígonos de referência delineados sobre HR e sobrepostos, invariantes, a LR e SR. O RF, aplicado separadamente a LR, SR e HR sobre as mesmas amostras emparelhadas no espaço, mostra que a SR se aproxima sistematicamente da referência e supera o LR nas classes estruturantes do contexto urbano; em paralelo, os painéis LR→SR→HR com polígonos confirmam, à mesma escala de 2,5 m, maior conformidade de arestas de edifícios, continuidade viária mais estável e interfaces água–terra e vegetação–edificado mais nítidas, traduzindo redução da mistura espectral nos bordos e menor necessidade de edição manual. Em edifícios e vias, onde a geometria fina é determinante, observam-se menos confusões (vias–edifícios) e melhor alinhamento entre fronteiras visuais e limites cartográficos; em água, os ganhos são sobretudo marginais, mas consistentes na definição de margem; em espaços verdes, a melhoria é dependente da textura e da sombra, mas beneficia da estabilização de contornos face ao edificado. A conjugação das métricas por classe com a evidência visual baseada em polígonos reforça a robustez dos resultados e demonstra que os ganhos perceptivos da SR têm utilidade operacional para a cartografia urbana, não se limitando a melhorias estéticas (Breiman, 2001; Belgiu & Drăguț, 2016; Congalton & Green, 2009; Foody, 2002; Ledig et al., 2017; Wang et al., 2018).

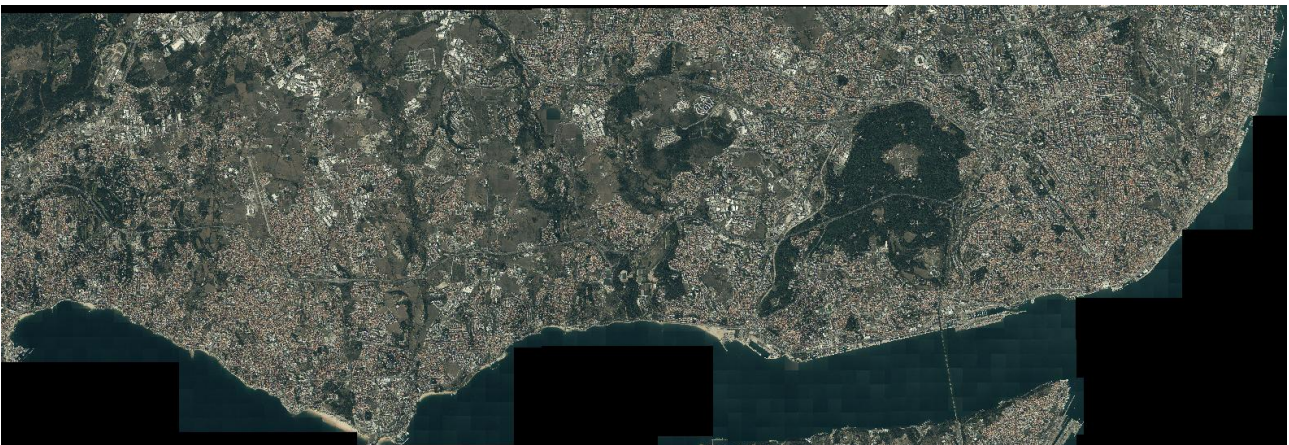
As limitações observadas enquadram-se nas escolhas metodológicas e nos constrangimentos operacionais. A utilização exclusiva da composição RGB (TCI) restringe a informação espectral; a inclusão futura de bandas NIR/SWIR ou de séries multi-temporais mais longas poderá aumentar a robustez temática, sobretudo em transições vegetação–edificado. O número de iterações foi reduzido face ao recomendado pelo repositório, impondo um teto prudente ao desempenho; ainda assim, a consistência dos ganhos ao longo dos cenários indica que o efeito principal se deve ao aumento de resolução e não à sobre otimização. (Wang et al., 2018; Wolters, Bastani, & Kembhavi, 2023; Zhu et al., 2018).

Em síntese, a dissertação cumpre os três objetivos propostos: demonstra a viabilidade de super-resolução Sentinel-2 para 2,5 m preservando coerência espectral e geoespacial (Figura 4.17); valida a

eficácia da abordagem em contextos urbanos distintos e sob mudança temporal (Figura 4.18); e comprova que os ganhos de detalhe se traduzem em melhoria mensurável na classificação temática. Estes resultados consolidam a super-resolução baseada em GAN como uma etapa credível de pré-processamento cartográfico para análise urbana com Sentinel-2, desde que enquadrada por boas práticas de registro e partição espacial/temporal.



(a) HR: orto 2023



(b) SR: validação interna

Figura 4.17 Comparação entre a imagem original em baixa resolução (LR) e a imagem de super-resolução (SR) no cenário de validação interna.

HR: orto 2023

SR: validação interna

(a)



(b)



Figura 4.18 Comparação visual entre a imagem HR e a imagem SR, apresentada em recortes localizados na zona do Estoril (a) e do Terreiro do Paço (b) (cenário de validação interna).

Conclusão

O presente trabalho de projeto centrou-se em explorar a super-resolução de imagens Sentinel-2 (aumento da resolução espacial de 10 m para ~2,5 m) através de redes neuronais generativas adversárias (GAN), com o intuito de melhorar a classificação da ocupação do solo em ambiente urbano. Foram estabelecidos três objetivos principais: (i) demonstrar a viabilidade de gerar imagens super-resolução coerentes com a referência de alta resolução (HR), preservando a consistência espectral e geométrica; (ii) validar a eficácia do modelo de super-resolução em contextos urbanos distintos e sob variação temporal, aferindo a capacidade de generalização espacial e temporal; e (iii) analisar o impacto do aumento de detalhe espacial na precisão da classificação temática de ocupação do solo. Todos estes objetivos foram cumpridos, conforme evidenciado pelos resultados obtidos. O estudo comprovou que é possível produzir imagens de super-resolução de Sentinel-2 com aproximadamente 2,5 m de GSD mantendo elevada fidelidade espectral/espacial, que o modelo GAN treinado consegue generalizar o ganho de resolução para áreas urbanas e épocas não utilizadas no treino, e que os incrementos de detalhe proporcionados pela super-resolução se traduzem em melhorias mensuráveis na classificação das classes de ocupação do solo, em comparação com a utilização direta das imagens de baixa resolução.

Do ponto de vista quantitativo, as imagens super-resolução (SR) geradas apresentaram melhorias significativas de qualidade face às imagens Sentinel-2 originais de baixa resolução (LR). Observou-se um aumento de aproximadamente 3 dB no PSNR (de ~10.7 para ~13.7) e um incremento expressivo no SSIM (de ~0,07 para ~0,25), indicando que as imagens SR se aproximam substancialmente mais das referências de alta resolução do que as imagens originais. Em paralelo, os erros médios quadrático e absoluto (MSE e MAE) reduziram para cerca de metade dos valores obtidos com LR, refletindo uma reconstrução muito mais fiel dos detalhes espaciais finos. Estas métricas convergem na mesma conclusão: a super-resolução aproximou de forma consistente o produto Sentinel-2 da imagem de referência HR nos aspetos críticos para a cartografia urbana, sustentando ganhos de legibilidade e separabilidade de alvos geográficos. Também a avaliação temática confirmou a vantagem quantitativa das imagens de super-resolução, por exemplo, nas classes urbanas estruturantes como edificado e vias, o classificador registou aumentos significativos de precisão e F1-score em comparação com os dados LR, aproximando os resultados daqueles obtidos com dados HR (alta resolução). Em suma, a adoção da super-resolução via GAN resultou em melhorias objetivas de qualidade de imagem e de desempenho na classificação quando comparada à resolução nativa do Sentinel-2.

Em termos qualitativos, os resultados reforçam a utilidade prática da super-resolução para leitura cartográfica urbana. Visualmente, as imagens SR exibem contornos e texturas mais nítidos e contínuos do que as imagens originais, sem introduzir distorções geométricas significativas. Constatou-se uma melhoria clara na definição dos contornos de edifícios, que na SR alinham melhor com os limites reais observados em ortoimagens de referência, facilitando a sua vectorização e reduzindo o efeito “escadeado” presente nas arestas oblíquas em LR. De igual modo, a rede viária urbana surge em SR com maior continuidade: ruas estreitas e interseções que no produto LR apareciam fragmentadas, devido à mistura espectral, são restabelecidas como segmentos lineares mais completos, aproximando-se do padrão esperado em HR. Nas interfaces entre classes, observam-se fronteiras mais estáveis e definidas, por exemplo, as margens água-terra tornam-se mais suaves e com menor contaminação de píxeis vizinhos, e as transições vegetação–edificado aparecem mais claras e coerentes. Estas melhorias indicam uma redução da ambiguidade e da mistura espectral nos bordos das classes em SR, o que se traduziu em menos confusões entre classes (e.g., entre vias e edifícios) e menor necessidade de correção manual dos mapas. Importa realçar que tais ganhos não são meramente estéticos, mas sim operacionais: a leitura geoespacial das imagens de super-resolução revela-se mais informativa para fins de cartografia urbana, permitindo extrair e discriminar elementos do território com maior confiança (Ledig et al., 2017; Wang

et al., 2018). Isto corrobora a noção de que os ganhos perceptivos da SR têm utilidade prática, alinhando-se com resultados de outros estudos que aplicaram técnicas GAN para melhorar a classificação de imagens Sentinel-2 (Breiman, 2001; Belgiu & Drăguț, 2016; Congalton & Green, 2009). Em síntese, a geração de detalhes adicionais via super-resolução traduziu-se numa melhoria real da legibilidade cartográfica e do desempenho temático, sobretudo nas classes onde a geometria fina é determinante (edifícios e vias).

Apesar dos sucessos alcançados, é importante reconhecer limitações neste estudo. Uma das restrições advém do uso exclusivo de imagens RGB (visível) do Sentinel-2 para treinar e aplicar o modelo; esta opção simplificou o processo, mas reduziu a informação espectral disponível, potencialmente limitando a discriminação de certas classes, em particular nas transições entre vegetação e área construída. A inclusão, em trabalhos futuros, de bandas adicionais (e.g., NIR) ou mesmo de séries multi-temporais mais extensas poderá aumentar a robustez da super-resolução e melhorar ainda mais a classificação temática, ao fornecer ao modelo pistas espectrais e temporais adicionais para distinguir materiais e coberturas do solo. Outra limitação relaciona-se com a capacidade computacional e de treino: devido a constrangimentos operacionais, o número de épocas/iterações de treino do GAN foi inferior ao ideal apontado na literatura, o que impôs um teto prudente ao desempenho atingido. Ainda assim, a consistência dos ganhos observados em todos os cenários sugere que o efeito principal (o aumento de detalhe espacial) foi concluído, não estando o modelo fortemente subtreinado nem sobreajustado. Adicionalmente, identificaram-se alguns artefactos típicos de redes GAN em casos particulares: por exemplo, ligeiros halos ou padrões de “píxelização” em torno de arestas muito contrastadas e texturas artificiais em regiões de sombra. Estes artefactos são localizados e não comprometem a interpretação global das imagens, mas salientam a necessidade de ajustes finos na modelação (por exemplo, na regularização ou escolha de perdas perceptuais vs. de fidelidade) para mitigar tais efeitos. Por fim, deve notar-se que os dados de referência HR usados (ortoimagens a 2,5 m) foram assumidos como *ground-truth* degradada; embora este procedimento seja comum em super-resolução supervisionada, existe sempre alguma discrepância entre a imagem sintética super-resolvida e uma imagem verdadeiramente adquirida a alta resolução. Esta abordagem, aliada ao foco numa região específica (Lisboa/Cascais) e a apenas duas épocas temporais, pode limitar a validade externa dos resultados, isto é, a garantia de que o modelo teria desempenho idêntico em outras regiões urbanas com características muito distintas ou em condições ambientais diferentes. Em suma, as conclusões obtidas devem ser interpretadas à luz destas limitações, que oferecem também oportunidades de melhoria em investigações futuras.

Com base nos resultados obtidos, podem ser elaboradas várias recomendações e linhas de trabalho futuro. Em primeiro lugar, sugere-se expandir a abordagem para novos conjuntos de dados, incluindo *datasets* multiespectrais e de maior resolução recentemente disponibilizados. Por exemplo, o uso de imagens de satélite de muito alta resolução (como WorldView ou PlanetScope) co-registradas com Sentinel-2, ou de novos *benchmarks* de super-resolução (e.g., SEN2NAIP ou MuS2) pode fornecer treino mais robusto e permitir avaliar até que ponto os ganhos observados se mantêm com dados de diferentes proveniências. Os modelos de próxima geração também merecem avaliação: as técnicas emergentes baseadas em *diffusion models* (modelos generativos difusivos) e as arquiteturas avançadas de *deep learning* (como Transformers voltados para super-resolução) podem superar algumas limitações dos GAN clássicos, oferecendo potencialmente imagens ainda mais fiéis e com menos artefactos. Adicionalmente, integrar informação de múltiplas imagens Sentinel-2 no tempo (super-resolução multi-imagem) é uma via promissora: agregando observações de várias datas, pode-se atenuar ruídos e obter detalhes mais estáveis, como demonstrado em iniciativas recentes de mapeamento global de alta resolução a partir de Sentinel-2. Outra direção consiste em aplicar a super-resolução a outras aplicações geoespaciais além da classificação estática de ocupação do solo: por exemplo, investigar o impacto da super-resolução em deteção de alterações temporais, monitorização de infraestruturas, ou mesmo na

melhoria de produtos derivados (como modelos digitais ou estimativas biofísicas) onde a resolução espacial seja um fator limitante. Por fim, recomenda-se avaliar a integração da etapa de super-resolução num fluxo de trabalho operacional completo de cartografia, por exemplo, incorporando as imagens SR num pipeline de classificação automática baseado em redes neuronais de segmentação semântica, para verificar se os ganhos de resolução se traduzem em melhorias equivalentes num sistema *end-to-end*. Estas perspetivas futuras, aliadas a uma contínua calibração e avaliação rigorosa, poderão consolidar e estender os contributos deste trabalho.

Em conclusão, este trabalho demonstrou que a super-resolução baseada em GAN pode ser uma ferramenta efetiva e fiável para melhorar a cartografia urbana a partir de dados Sentinel-2 de média resolução. Os resultados obtidos mostram que é possível aproximar a informação de 10 m do Sentinel-2 à qualidade de produtos de alta resolução de forma consistente e mensurável, potenciando a extração de informação temática mais precisa sem incorrer nos custos elevados de aquisição de imagem de muito alta resolução. Tal abordagem reveste-se de especial interesse em contextos com restrições orçamentais ou necessidades de atualização frequente, onde a disponibilidade gratuita e regular das imagens Sentinel-2, aliada a técnicas de super-resolução, pode fornecer uma solução intermediária de baixo custo para cartografia de pormenor. Em síntese, a dissertação cumpriu os objetivos a que se propôs e fornece evidências de que a aplicação de super-resolução por GAN constitui um valioso contributo metodológico para a área da Engenharia Geoespacial. Este contributo manifesta-se não só em ganhos objetivos de qualidade de imagem e de classificação, mas também na abertura de novas possibilidades operacionais, permitindo que dados de satélite moderadamente detalhados, como os do Sentinel-2, possam desempenhar um papel acrescido na cartografia urbana de precisão, de forma económica e eficaz.

Referência Bibliográfica

- Allen Institute for AI (AI2). (2023). Satlas Super-Resolution (training & inference code; dataset structure; configs). GitHub. <https://github.com/allenai/satlas-super-resolution>
- Anselin, L. (1995). Local Indicators of Spatial Association—LISA. *Geographical Analysis*, 27(2), 93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>
- Batool, A., & Byun, Y. (2022). An ensemble architecture based on deep learning model for click fraud detection in pay-per-click advertisement campaign. *IEEE Access*, 1–1. <https://doi.org/10.1109/ACCESS.2022.3211528>
- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Bermudez, J., et al. (2020). Super-resolution of Sentinel-2 imagery using deep learning techniques: Applications and challenges. *Remote Sensing of Environment*, 251, 112042.
- Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*. Springer.
- Blau, Y., & Michaeli, T. (2018). The perception–distortion tradeoff. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 6228–6237. <https://doi.org/10.1109/CVPR.2018.00652>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Chen, X., et al. (2017). High-resolution image synthesis with latent-variable models. *CVPR*, 1576–1585. <https://doi.org/10.1109/CVPR.2017.169>
- Choi, J., et al. (2023). Enhancing land-use classification using GAN-based super-resolution on Sentinel-2 data. *International Journal of Applied Earth Observation*.
- Congalton, R. G., & Green, K. (2009). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices* (2nd ed.). CRC Press.
- Creswell, A., White, T., Dumoulin, V., Arulkumaran, K., Sengupta, B., & Bharath, A. A. (2018). Generative adversarial networks: An overview. *IEEE Signal Processing Magazine*, 35(1), 53–65. <https://doi.org/10.1109/MSP.2017.2765202>
- Dong, C., Loy, C. C., He, K., & Tang, X. (2016). Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2), 295–307. <https://doi.org/10.1109/TPAMI.2015.2439281>

- Drusch, M., et al. (2012). Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 120, 25–36. <https://doi.org/10.1016/j.rse.2011.11.026>
- ESA. (2021). Sentinel-2 User Handbook. ESA. https://sentinels.copernicus.eu/documents/247904/685211/Sentinel-2_User_Handbook
- ESA. (2023). Sen2Cor – Level-2A Atmospheric Correction. <https://step.esa.int/main/snap-supported-plugins/sen2cor>
- Fakhfakh, R., Ben Ammar, A., & Ben Amar, C. (2017). Deep learning-based recommendation: Current issues and challenges. *International Journal of Advanced Computer Science and Applications*, 8. <https://doi.org/10.14569/IJACSA.2017.081209>
- Farris, C., Giaiotti, D., Miniussi, S., Sgubin, C., & Tudorov, N. (2022). An integrated approach for marine litter hot spots identification. <https://doi.org/10.36253/979-12-215-0030-1.20>
- Foody, G. M. (2002). Status of land cover classification accuracy assessment. *Remote Sensing of Environment*, 80(1), 185–201. [https://doi.org/10.1016/S0034-4257\(01\)00295-4](https://doi.org/10.1016/S0034-4257(01)00295-4)
- Gao, F., et al. (2022). Improving Sentinel-2 classification with GAN-based super-resolution. *Remote Sensing*.
- Gislason, P. O., Benediktsson, J. A., & Sveinsson, J. R. (2006). Random Forests for land cover classification. *Pattern Recognition Letters*, 27(4), 294–301. <https://doi.org/10.1016/j.patrec.2005.08.011>
- Goodfellow, I. (2016). NIPS 2016 Tutorial: Generative Adversarial Networks. arXiv:1701.00160.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.
- Goodfellow, I., Pouget-Abadie, J., Mirza, D., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *NeurIPS*, 27. <https://doi.org/10.48550/arXiv.1406.2661>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
- Huang, X., et al. (2023). MobileGAN: Skin lesion segmentation using a lightweight generative adversarial network.
- Huang, Y., et al. (2022). Challenges in GAN-based super-resolution for remote sensing. *IEEE Geoscience and Remote Sensing Letters*.

- Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. CVPR, 1125–1134. <https://doi.org/10.1109/CVPR.2017.132>
- Johnson, J., Alahi, A., & Fei-Fei, L. (2016). Perceptual losses for real-time style transfer and super-resolution. ECCV, 694–711. <https://arxiv.org/abs/1603.08155>
- Jolicœur-Martineau, A. (2019). The relativistic discriminator: A key element missing from standard GAN. ICLR Workshop. <https://arxiv.org/abs/1807.00734>
- Jouppi, N. P., Young, C., Patil, N., et al. (2017). In-datacenter performance analysis of a Tensor Processing Unit. ISCA, 1–12. <https://doi.org/10.1145/3079856.3080246>
- Kim, K. I., et al. (2008). Single-image super-resolution using sparse regression and natural image prior. IEEE TPAMI.
- Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv:1412.6980.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. NeurIPS, 25.
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. Neural Computation, 1(4), 541–551.
- Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi, W. (2017). Photo-realistic single image super-resolution using a generative adversarial network. CVPR, 4681–4690. <https://doi.org/10.1109/CVPR.2017.19>
- Li, H., et al. (2022). Impact of super-resolution on land use classification accuracy: A case study with Sentinel-2. Remote Sensing of Environment, 268, 112735.
- Li, J., & Roy, D. P. (2017). A global analysis of Sentinel-2 and Landsat-8 data revisits and implications for terrestrial monitoring. Remote Sensing of Environment, 193, 18–28. <https://doi.org/10.1016/j.rse.2017.02.021>
- Ma, X., et al. (2019). Super-resolution for remote sensing images: A systematic survey. IEEE Transactions on Geoscience and Remote Sensing, 57(8), 4440–4457.
- Maxwell, A. E., Warner, T. A., & Fang, F. (2018). Implementation of machine-learning classification in remote sensing: An applied review. International Journal of Remote Sensing, 39(9), 2784–2817. <https://doi.org/10.1080/01431161.2018.1433343>

- McCarthy, J., et al. (2006). The Dartmouth Conference: The Birth of Artificial Intelligence. *AI Magazine*.
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.
- Mittal, A., Moorthy, A. K., & Bovik, A. C. (2012). No-reference image quality assessment in the spatial domain. *IEEE Transactions on Image Processing*, 21(12), 4695–4708. <https://doi.org/10.1109/TIP.2012.2214050>
- Moreno-Torres, J. G., Raeder, T., Alaíz-Rodríguez, R., Chawla, N. V., & Herrera, F. (2012). A unifying view on dataset shift in classification. *Pattern Recognition*, 45(1), 521–530. <https://doi.org/10.1016/j.patcog.2011.06.019>
- Mou, L., et al. (2019). Super-resolution of remote sensing images using GANs. *International Journal of Remote Sensing*.
- Mountrakis, G., et al. (2011). Support vector machines in remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(3), 247–259.
- Novák, D., & Tóth, V. (2016). The role of orthophotos in the development of geographic information systems. *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, 64(1), 227–236.
- Odena, A., Dumoulin, V., & Olah, C. (2016). Deconvolution and checkerboard artifacts. *Distill*. <https://distill.pub/2016/deconv-checkerboard/>
- Pasqualotto, N., Delegido, J., Falanga Bolognesi, S., D’Urso, G., Belfiore, O. R., & Moreno, J. (2019). Canopy chlorophyll content and LAI estimation from Sentinel-2: Vegetation indices and Sentinel-2 Level-2A automatic products comparison. 2019 IEEE International Workshop on Metrology for Agriculture and Forestry (MetroAgriFor). <https://doi.org/10.1109/MetroAgriFor.2019.8909218>
- Polyak, B. T. (1964). Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, 4(5), 1–17.
- Rasmussen, C. E., & Williams, C. K. I. (2006). *Gaussian Processes for Machine Learning*. MIT Press.
- Roberts, D. R., Bahn, V., Ciuti, S., et al. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8), 913–929. <https://doi.org/10.1111/ecog.02881>
- Robson, C., et al. (2020). Improving land cover classification using extended multi-attribute profiles (EMAP). *Remote Sensing*, 12(9), 1392. <https://doi.org/10.3390/rs12091392>
- Rodríguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An assessment of the effectiveness of a Random Forest classifier for land-cover classification.

- ISPRS Journal of Photogrammetry and Remote Sensing, 67, 93–104.
<https://doi.org/10.1016/j.isprsjprs.2011.11.002>
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533–536.
- Russell, S., & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach*. Pearson.
- Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *ICLR*. <https://arxiv.org/abs/1409.1556>
- Smith, L. N. (2017). Cyclical learning rates for training neural networks. *WACV*, 464–472.
<https://doi.org/10.1109/WACV.2017.58>
- Sutskever, I., Martens, J., Dahl, G., & Hinton, G. (2013). On the importance of initialization and momentum in deep learning. *ICML*, 1139–1147.
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. MIT Press.
- Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46, 234–240.
- Toutin, T. (2004). Review article: Geometric processing of remote sensing images: Models, algorithms and methods. *International Journal of Remote Sensing*, 25(10), 1893–1924.
<https://doi.org/10.1080/0143116031000101611>
- Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Qiao, Y., & Change Loy, C. (2018). ESRGAN: Enhanced super-resolution generative adversarial networks. *ECCV Workshops*.
<https://arxiv.org/abs/1809.00219>
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 600–612.
<https://doi.org/10.1109/TIP.2003.819861>
- Wang, Z., Zhang, J., Li, Y., et al. (2020). Deep learning for image super-resolution: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Wang, Z., et al. (2021). Enhancing Landsat-8 satellite imagery using SRGAN. *IEEE Transactions on Geoscience and Remote Sensing*.
- Wang, Z., Simoncelli, E. P., & Bovik, A. C. (2003). Multiscale structural similarity for image quality assessment. *Asilomar Conference on Signals, Systems and Computers*, 9–13.
<https://doi.org/10.1109/ACSSC.2003.1292216>
- Wolters, P., Bastani, F., & Kembhavi, A. (2023). Zooming Out on Zooming In: Advancing Super-Resolution for Remote Sensing. *arXiv*. <https://arxiv.org/abs/2311.18082>

- Xu, J., et al. (2021). Training generative adversarial networks for satellite image enhancement: Computational and practical considerations. *IEEE Transactions on Geoscience and Remote Sensing*, 59(8), 6789–6802.
- Yan, Y., et al. (2019). Fine-grained attention and feature-sharing generative adversarial networks for single image super-resolution. arXiv:1911.10773. <https://arxiv.org/abs/1911.10773>
- Yang, W., Zhang, X., Tian, Y., Wang, W., & Xue, J.-H. (2019). Deep learning for single image super-resolution: A brief review. *IEEE Transactions on Multimedia*, 21(12), 3106–3121.
- Zhu, X. (2005). Semi-supervised learning literature survey. University of Wisconsin–Madison.
- Zhu, X., et al. (2018). Preserving spectral fidelity in super-resolution techniques for remote sensing. *IEEE Transactions on Geoscience and Remote Sensing*, 56(10), 5940–5952. <https://doi.org/10.1109/TGRS.2018.2829382>
- Zhu, X., et al. (2019). Multi-source data fusion for improved land use classification. *ISPRS International Journal of Geo-Information*.
- Zhu, X., et al. (2022). Challenges in super-resolution for remote sensing data using generative adversarial networks. *Remote Sensing*, 14(5), 1234–1248.