

RESSALVA

Atendendo solicitação do(a) autor(a), o texto completo desta dissertação será disponibilizado somente a partir de 10/07/2026.

GUSTAVO ROTA COLLEGIO

Road detection via deep learning: unsupervised domain
adaptation for urban and suburban environment



Presidente Prudente - SP

2025

Gustavo Rota Collegio

Road detection via deep learning: unsupervised domain adaptation for urban and suburban environment

Thesis presented to the Graduate Program in Cartographic Sciences of São Paulo State University – Presidente Prudente Campus, as a partial requirement for obtaining the degree of Master in Cartographic Science.

Advisor: Dr. Aluir Porfírio Dal Poz

Co-advisor: Dr. Antonio Gaudencio
Guimarães Filho

Presidente Prudente - SP

2025

C697r

Collegio, Gustavo Rota

Road detection via deep learning: unsupervised domain adaptation for urban and suburban environment / Gustavo Rota Collegio. -- Presidente Prudente, 2025

101 p.

Dissertação (mestrado) - Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Tecnologia, Presidente Prudente

Orientador: Aluir Porfirio Dal Poz

Coorientador: Antonio Gaudencio Guimarães Filho

1. Road detection. 2. Unsupervised Domain Adaptation. 3. Domain Adversarial Neural Network. 4. U-Net. 5. Semantic Segmentation. I. Título.

Potential Impact of This Research

The road detection task is essential for several purposes, as it provides direct applications to real-world activities. Although algorithms based on Deep Learning (DL) are the predominant approach for handling the challenges of the road network, their performance is still limited, facing difficulties such as the need for large labeled datasets and sensitivity to domain shift effects. Overcoming these challenges represents a key expected impact of this research.

By integrating the technique called Unsupervised Domain Adaptation (UDA) and achieving generalization across urban and suburban domains with different spatial resolutions in aerial imagery, the approach strengthens its potential for large-scale road network mapping. This capability allows the research to directly support the Sustainable Development Goals (SDGs) of the the United Nations (UN). In particular, it contributes to sustainable infrastructure (SDG 9) and to sustainable cities and communities through improved transportation systems (SDG 11), reinforcing the societal relevance of the study.

Impacto Potencial desta Pesquisa

A tarefa de detecção de rodovias é essencial para diversos propósitos, pois fornece aplicações diretas em atividades do mundo real. Embora os algoritmos baseados em Aprendizado Profundo (Deep Learning - DL) sejam a abordagem predominante para lidar com os desafios da malha viária, tal desempenho ainda é limitado, havendo dificuldades como a necessidade de grandes conjuntos de dados rotulados e a sensibilidade aos efeitos de mudança de domínio. Mitigar esses desafios representa um impacto esperado central desta pesquisa.

Ao integrar a técnica denominada Adaptação de Domínio Não Supervisionada (Unsupervised Domain Adaptation – UDA) e alcançar a generalização entre domínios urbanos e suburbanos com diferentes resoluções espaciais em imagens aéreas, a abordagem potencializa seu potencial para o mapeamento da rede viária em larga escala. Essa capacidade permite que a pesquisa apoie diretamente os Objetivos de Desenvolvimento Sustentável (ODS) da Organização das Nações Unidas (ONU). Em particular, contribui para a infraestrutura sustentável (ODS 9) e para cidades e comunidades sustentáveis por meio da melhoria dos sistemas de transporte (ODS 11), reforçando a relevância social do estudo.

CERTIFICADO DE APROVAÇÃO

TÍTULO DA DISSERTAÇÃO: ROAD DETECTION VIA DEEP LEARNING: UNSUPERVISED DOMAIN ADAPTATION FOR URBAN AND SUBURBAN ENVIRONMENT

AUTOR: GUSTAVO ROTA COLLEGIO

ORIENTADOR: ALUIR PORFIRIO DAL POZ

COORDENADOR: ANTONIO GAUDENCIO GUIMARÃES FILHO

Aprovado como parte das exigências para obtenção do Título de Mestre em Ciências Cartográficas, área: Aquisição, Análise e Representação de Informações Espaciais pela Comissão Examinadora:



Prof. Dr. ALUIR PORFIRIO DAL POZ (Participação Presencial)

Departamento de Cartografia / Faculdade de Ciências e Tecnologia de Presidente Prudente - FCT/Unesp

CONFERÊNCIA

Prof. Dr. DANIEL RODRIGUES DOS SANTOS (Participação Virtual)

Seção de Engenharia Cartográfica / Instituto Militar de Engenharia



Prof. Dr. MAURÍCIO GALO (Participação Presencial)

Departamento de Cartografia / Faculdade de Ciências e Tecnologia de Presidente Prudente - FCT/Unesp

Presidente Prudente, 10 de julho de 2025

To my grandmother (in memoriam)

*Whose passing left a void, but whose teachings, love, and example remain alive
in my heart, guiding me at every step of the journey of my life.*

To my parents, my sister, and my girlfriend

Who always believed in me, even in times of uncertainty, offering love, encouragement, and unconditional support.

Acknowledgements

I thank God for granting me the opportunity, the strength, and the wisdom to carry out this research, and for fostering in me a sincere interest and dedication to what I do.

I thank my parents for their constant support and love throughout my life, which have shaped and strengthened my character.

I thank my sister for all the conversations we have shared, for her trust, and for her unwavering support.

I thank my loving girlfriend for her patience, support, and encouragement, especially during moments when I needed to stay focused on this work.

I thank my advisor, Prof. Aluir P. Dal Poz, for the trust he has had in me since my time as an undergraduate student with early research experience, and for consistently motivating me to aim higher and bring out my best.

I thank my co-advisor, Dr. Antonio G. G. Filho, who, even from afar in the Brazilian Army, trusted that this challenging master's thesis could be accomplished. His dedication and high-level of expertise have always been inspiring since I was an undergraduate student.

I thank Prof. Ayman Habib for welcoming me to Purdue University with his research group.

I thank all my friends from the PPGCC for the conversations, discussions, and enjoyable moments we have shared.

I thank FCT/UNESP and all the professors from the PPGCC for the knowledge and experiences I gained throughout the course.

This study was financed, in part, by the São Paulo Research Foundation (FAPESP), Brasil. Process Number 2024/02040-4 and Process Number 2021/03586-2.

This work was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 (Grants 88887.821783/2023-00 and 88887.936740/2024-00).

Commit to the Lord whatever you do, and he will establish your plans.

Proverbs 16:3 NIV

Abstract

The road network is one of the main ways of transport in the world, providing various applications for human activities, such as urban planning, economy (logistics), autonomous navigation, and others. Currently, methods based on Deep Learning (DL) are the main approaches for obtaining road network information. However, this task is still challenging due to the domain shift effect. Since the characteristics and properties that define a domain for training are distinct from a domain designed for testing (source domain and target domain), the inference of pre-trained models to new areas is hindered, leading to a drop in classification accuracy. To address this challenge, this work explores an Unsupervised Domain Adaptation (UDA) approach, which leverages the strategy of the Domain Adversarial Neural Network (DANN) adapted to the U-Net architecture. Without the need for labeled target data to mitigate domain shift, our adaptation deals with differences in geographical contexts, from urban to suburban areas, and different spatial resolutions of aerial imagery. The experimental results demonstrate that the UDA approach outperforms the source-trained model (i.e., baseline) in all experiments, which reaches the main purpose of UDA. Particularly, in the Distrito Federal state (DF), our study area, UDA improves the road extraction performance by 3.78% in $F1_{\text{Score}}$ and 3.0% in IoU when compared to the baseline. The qualitative analysis further illustrates that the UDA approach relies on geometric structures and contextual cues to produce domain-invariant features, which contributes to better generalization across domains, although it may suppress road extraction-relevant features.

Keywords: Road detection. Unsupervised Domain Adaptation. Domain Adversarial Neural Network. U-Net. Semantic Segmentation.

Resumo

A malha viária é um dos principais meios de transporte no mundo, sendo fundamental para diversas aplicações, como planejamento urbano, economia (logística), navegação autônoma, entre outras. Atualmente, métodos baseados em aprendizado profundo (*Deep Learning* – DL) são as principais abordagens para obtenção de informação da rede viária. No entanto, essa tarefa ainda é desafiadora devido ao efeito chamado *domain shift*. Uma vez que as características e propriedades que definem um domínio para treino são distintas de um domínio designado para teste (domínio fonte e domínio alvo), a inferência de modelos pré-treinados é limitada, acarretando em uma queda na acurácia da classificação. Para lidar com esse desafio, este trabalho explora uma abordagem chamada *Unsupervised Domain Adaptation* (UDA), que utiliza a estratégia da *Domain Adversarial Neural Network* (DANN) adaptada à arquitetura U-Net. Sem a necessidade de rótulos no domínio alvo para mitigar o efeito da mudança de domínio, a adaptação proposta lida com diferenças em contextos geográficos, de áreas urbanas para suburbanas, assim como para diferentes resoluções espaciais de imageamento aéreo. Os resultados experimentais demonstram que UDA supera o modelo treinado apenas com o domínio fonte (*baseline*) nos três experimentos realizados, atingindo o principal objetivo da UDA. Especificamente no estado do Distrito Federal (DF), área de estudo, UDA melhora o desempenho da extração de vias em 3,78% em $F1_{Score}$ e 3,0% em IoU quando comparado com a *baseline*. A análise qualitativa ainda demonstra que a extração de características em UDA prioriza a princípio propriedades geométricas ou informações contextuais para produzir características invariantes ao domínio de modo que desconsidera características relevantes à tarefa de extração de rodovias, a fim de alcançar melhor generalização.

Palavras-chave: Detecção de rodovias. Adaptação de Domínio não Supervisionada. Domain Adversarial Neural Network. U-Net. Segmentação Semântica.

List of Figures

Figure 1 – Road model.	19
Figure 2 – Comparing patch classification and semantic labeling.	22
Figure 3 – Unsupervised Domain Adaptation in the RS&P area.	24
Figure 4 – DL as a subset of ML and AI.	27
Figure 5 – Artificial neuron structure.	29
Figure 6 – CNN representation by units: X represents the input units; A and B represent the feature maps generated; max represents the pooling operation; and F represents the fully connected layer.	31
Figure 7 – Inside a CNN.	32
Figure 8 – Activation functions: ReLU, $Tanh$, and Sigmoid. In blue color: activation functions; in red color: derivatives.	33
Figure 9 – Pooling operation representation. The top part shows the transformation of a feature map from $28 \times 28 \times 64$ to $14 \times 14 \times 64$. The bottom part highlights the same operation on a single channel to emphasize spatial downsampling.	34
Figure 10 – VGG-16 illustration. The @ symbol is the feature map dimension.	35
Figure 11 – Training error (left) and test error (right) with 20-layer and 56-layer networks (CIFAR-10 dataset).	36
Figure 12 – Residual learning block.	36
Figure 13 – Derivates of a function to follow the function downhill to a minimum.	38
Figure 14 – Backpropagation illustration: forward and backward pass.	42
Figure 15 – Example of tasks in the DL context.	43
Figure 16 – Fully convolutional network. Segmentation g.t. stands for segmentation ground truth.	44
Figure 17 – The U-Net architecture.	45
Figure 18 – Semantic segmentation strategies incorporated in DeepLabv3+: (a) a combination of encoder-decoder structure and a spatial pyramid pooling; (b) an example of a 3×3 atrous depthwise convolution with a dilation rate of 2.	46
Figure 19 – The DeepLabv3+ model.	47
Figure 20 – Approaches in ML/DL. a) traditional; and (b) transfer learning.	50
Figure 21 – An overview of different settings of domain adaptation.	53
Figure 22 – Examples of domain pairs used in the experiments to UDA approach.	55
Figure 23 – DANN architecture for UDA.	56
Figure 24 – Urban areas in Brasília city (DF).	58
Figure 25 – Cheng et al. (2017)’s dataset to support UDA analysis: grayscale-RGB case.	60

Figure 26 – Domain pairs from UAV imagery. Panels (a) and (b) correspond to the source and target domains, respectively. MM stands for I-65 Mile Markers.	61
Figure 27 – Source and target domains of DF dataset. Panels (a) and (b) correspond to the source and target domains, respectively.	61
Figure 28 – Flowchart of UDA data preprocessing pipeline.	62
Figure 29 – Overall architecture of the proposed approach: U-Net is adapted to the DANN framework. The encoder acts as the feature extractor, the decoder as the label predictor, and the additional domain discriminator completes the model as a full UDA approach.	64
Figure 30 – DANN-based adaptation for DeepLabv3+. ASPP stands for Atrous Spatial Pyramid Pooling and GRL for Gradient Reversal Layer.	67
Figure 31 – Attention mechanism for U-Net skip connections.	68
Figure 32 – Prediction based on the influence of the domain regularization parameter.	71
Figure 33 – UDA grayscale-to-RGB predictions on the target domain. The 3rd and 4th rows correspond to the highlighted close-ups in a) and b). The last column corresponds to the Attention U-Net baseline.	73
Figure 34 – UDA predictions on the target domain. Rows: (a) Target image; (b) Reference data; (c)-(e) Baseline, UDA, and Supervised for DeepLabv3+, (f)-(h) for U-Net, and (i)-(k) for Attention U-Net. FPs: red color; FNs: cyan color.	75
Figure 35 – Encoder feature maps' visualization by Score-CAM. Weighted class activation maps on image overlap from the coarse to the fine level of the encoder. Columns: (a) Target image; (b)-(d) 2nd conv. block, 4th conv. block, and ASPP for DeepLabv3+; (e)-(g) 2nd, 4th, and 5th conv. block for ResU-Net; (h)-(j) 2nd, 4th, and 5th conv. block for Att-ResU-Net.	76
Figure 36 – DF predictions on the target domain. False positives: red color; false negatives: cyan color.	79
Figure 37 – First case of encoder maps visualized by Score-CAM. Top two rows: weighted class activation maps on image overlap. Bottom row: emphasized map.	80
Figure 38 – Second case of encoder maps visualized by Score-CAM. Top two rows: weighted class activation maps on image overlap. Bottom row: emphasized map.	81
Figure 39 – Third case of encoder maps visualized by Score-CAM. Top two rows: weighted class activation maps on image overlap. Bottom row: emphasized map.	82
Figure 40 – First case of decoder maps visualized by Score-CAM. Top two rows: weighted class activation maps on image overlap. Bottom row: emphasized map.	83

Figure 41 – Second case of decoder maps visualized by Score-CAM. Top two rows: weighted class activation maps on image overlap. Bottom row: emphasized map.	84
Figure 42 – Third case of decoder maps visualized by Score-CAM. Top two rows: weighted class activation maps on image overlap. Bottom row: emphasized map.	84
Figure 43 – Comparison of the best-performing UDA-based models in each experiment. $F1_{\text{score}}$ and IoU are shown for Experiment 1 (Grayscale-to-RGB), Experiment 2 (Colombia UAV-to-I65 UAV), and Experiment 3 (DF Area 10-to-DF Area 1).	86

List of Tables

Table 1	– ResNet architectures proposed for ImageNet.	37
Table 2	– Specifications of the aerial photogrammetry survey.	59
Table 3	– Datasets’ specifications.	59
Table 4	– UAV dataset’s characteristics for source and target domains.	60
Table 5	– Library and software specifications.	63
Table 6	– Quantitative analysis of the domain regularize term in the target domain prediction. The first λ term corresponds to the encoder updating and the second one to the domain discriminator module.	71
Table 7	– Quantitative performance of predictions on the target domain (%). The best results are in bold, and the second-best results are underlined. *U-Net modified by Cheng et al. (2017).	72
Table 8	– Target domain’s predictions Metrics, where values in bold are the best and the values underlined are the second best.	74
Table 9	– Target domain performance when both domains share the same GSD of 20 cm.	77
Table 10	– Target domain performance: Area 10 with 20 cm of GSD and Area 1 with 25 cm of GSD. The best results are in bold and the second ones are underlined.	78
Table 11	– Influence of the domain lambda regularizer in terms of $F1_{\text{Score}}$ and IoU. The best results are in bold.	78

List of abbreviations and acronyms

Adam	Adaptive Moment Estimation
AdaGrad	Adaptive Gradient Algorithm
AI	Artificial Intelligence
ANN	Artificial Neural Network
API	Application Programming Interface
ASPP	Atrous Spatial Pyramid Pooling
BCE	Binary Cross-Entropy
BN	Batch Normalization
CNN	Convolutional Neural Network
DA	Domain Adaptation
DANN	Domain Adversarial Neural Network
DF	<i>Distrito Federal</i> (Federal District)
DIP	Digital Image Processing
DL	Deep Learning
DSM	Digital Surface Model
DTM	Digital Terrain Model
FN	False Negatives
FP	False Positives
FPN	Feature Pyramid Network
GC	Gradient Centralization
GD	Gradient Descent
GIS	Geographic Information System
GSD	Ground Sample Distance

I-65	Interstate 65
I2I	Image-to-Image
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
IoU	Intersection over Union
ISPRS	International Society for Photogrammetry and Remote Sensing
LiDAR	Light Detection and Ranging
ML	Machine Learning
PNG	Portable Network Graphics
ReLU	Rectified Linear Unit
ResNet	Residual Network
RGB	Red, Green, and Blue bands
RMSProp	Root Mean Square Propagation
RS&P	Remote Sensing and Photogrammetry
SGD	Stochastic Gradient Descent
SIRGAS	<i>Sistema de Referência Geodésico para as Américas</i> (Geocentric Reference System for the Americas)
TL	Transfer Learning
TN	True Negatives
TP	True Positives
UAV	Unmanned Aerial Vehicle
UDA	Unsupervised Domain Adaptation
UTM	Universal Transverse Mercator
VC	Vapnik–Chervonenkis
VGG	Visual Geometry Group

Contents

1	INTRODUCTION	18
1.1	Initial considerations	18
1.2	State-of-the-art	21
1.3	Objectives	25
1.3.1	General	25
1.3.2	Specifics	25
1.4	Contributions	25
1.5	Justification and relevance	26
1.6	Master's thesis overview	26
2	THEORETICAL FOUNDATION	27
2.1	Deep Learning	27
2.1.1	Machine Learning	28
2.1.1.1	Neural Networks history	28
2.1.1.2	Artificial Neural Network	29
2.1.1.3	Supervised and unsupervised learning	30
2.1.2	Convolutional Neural Networks	31
2.1.2.1	VGG architecture	34
2.1.2.2	ResNet architecture	35
2.1.3	Training and optimization	37
2.1.3.1	Gradient-based optimization	37
2.1.3.2	Optimization for deep models	39
2.1.3.2.1	Stochastic Gradient Descent	40
2.1.3.2.2	Adam	40
2.1.3.3	Backpropagation	41
2.1.4	Semantic segmentation	42
2.1.4.1	U-Net	44
2.1.4.2	DeepLabv3+	45
2.1.4.3	Loss functions	47
2.2	Domain Adaptation in Deep Learning	50
2.2.1	Transfer Learning	50
2.2.2	Key Principles of Domain Adaptation	52
2.2.3	Unsupervised Domain Adaptation	53
2.2.3.1	Domain Adversarial Neural Network	55
3	MATERIAL AND METHODS	58

3.1	Study Area	58
3.2	Data collection	59
3.2.1	Cheng et al. (2017)'s dataset	59
3.2.2	UAV dataset	60
3.2.3	DF dataset	61
3.3	UDA Data Preprocessing Pipeline	62
3.4	Available computational resources	63
3.5	Method	64
3.5.1	Semantic Segmentation Models Adaptation	66
3.5.1.1	DANN-based DeepLabv3+	67
3.5.1.2	DANN-based Attention U-Net	67
3.5.2	Evaluation of the results	68
4	RESULTS AND ANALYSIS	70
4.1	Experiment I	70
4.2	Experiment II	73
4.3	Experiment III	77
4.4	Synthesis of Experimental Findings	85
5	CONCLUSION AND FUTURE WORKS	88
	REFERENCES	90

1 Introduction

1.1 Initial considerations

The road network is one of the main modes of transport in the world, which provides various applications for human activities. Expansion or restriction among regions, urban planning (public policy), economy (e.g., logistics), autonomous navigation systems, smart city development, etc. are just a few examples of the importance of the road network (Wang et al., 2016; Abdollahi et al., 2020). However, obtaining accurate information for road network mapping is still a challenge. Particularly, this is due to the constant changes in roads and their irregular distribution across a territory.

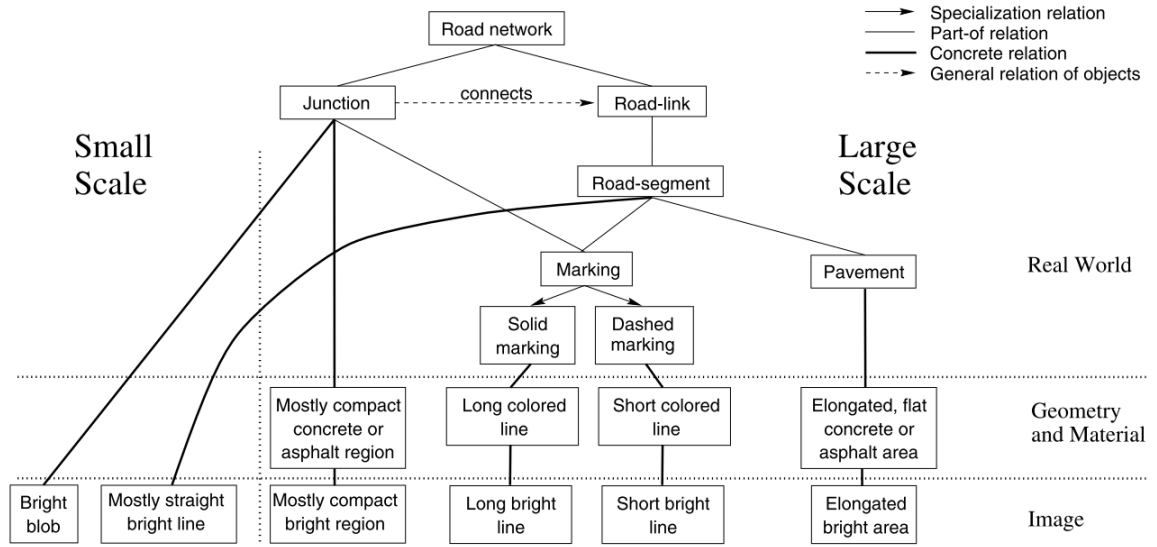
Road features such as width, length, type, jurisdiction, and others are essential for obtaining the road network and, especially, for ensuring consistency and full usage. Images taken from sensors boarded in aerial or orbital platforms are a quick and economical source of getting new spatial information over the roads or updating them in cartographic databases. Vosselman et al. (1995) state that using this data type, road network mapping is based on information about geometrical, radiometrical, topological, functional, and contextual characteristics.

For instance, roads do not stop without a reason, they intersect and build a network (topology) that connects cities (functionality). In geometric and radiometric terms, roads are elongated and have maximum curvature (geometric information), while according to the radiometric information, their surface is usually homogeneous with contrast to adjacent areas. On the other hand, these road characteristics have relations with objects in their surroundings (contextual aspect), like fly-overs, occlusion caused by trees and vehicles, different targets with similar spectral and radiometric information (buildings, rivers, parking lots), etc.

It is evident that, despite the eminently practical value of getting information about the roads in aerial images, a complex and dynamic context is inserted. Baumgartner and Hinz (2000), for this end, proposed a road model (Figure 1). This road model, which comprises multiple scales and describes the road network at different levels, aims to analyze how the road network in the object-space (*real world* level) is related to the image-space (*image* level). Essentially, this relation depends, mainly, on the surrounds objects, topology, geometry, acquisition methods, and scale. For example, at large scale, i.e., in higher-resolution images, a better geometry of the roads is obtained, just like the substructures on the roads. Otherwise, at a small scale, roads are presented only a few pixels wide, and small objects (e.g., trees and cars) and their surroundings have little

influence, compared to a large scale. Moreover, the contextual model named by Vosselman et al. (1995) is split by Baumgartner and Hinz (2000) in local and global context: the first one describes relations between objects of different types, while the other one describes the context like in main regions (e.g., urban, rural, forest).

Figure 1 – Road model.



Source: Baumgartner and Hinz (2000).

However, that road model, as pointed out by Filho (2022), is not unique and does not cover all types of roads but is worth it as a theoretical basis. Consequently, by assuming all these characteristics and relations about the road network, the road detection task can be considered one of the most challenging and recurrent research topics in the Remote Sensing and Photogrammetry (RS&P) community (Wang et al., 2025). The development of road network mapping is, therefore, relevant for the scientific community, as well as public policies and private organizations.

Owing to aspects like cost, low accuracy, time-consuming, and the possibility of generating data already obsolete to obtaining or updating the information of the road network manually, different approaches for road detection have been developed since the 70s. As reported by Dal Poz, Zanin and Do Vale (2006), with the use of aerial images, and according to the degree of automation involved, those approaches are categorized in semiautomatic or automatic methods, with Bajcsy and Tavakoli (1976) being pioneers in this context.

The use of automatic methods, since the last decade, contextualizes the specific subarea of Machine Learning (ML), i.e., Deep Learning (DL), as the main approach for the specific topic of road detection, proposed initially by Mnih and Hinton (2010). Specifically, this is a result of the advances in Computer Vision for many areas of knowledge since the

2000s (Lecun; Bengio; Hinton, 2015), making an impact on the RS&P area (Ma et al., 2019; Zhang; Zhang, 2022).

Convolutional Neural Networks (CNNs) are the algorithms most utilized for road detection in this DL context. According to Filho (2022), CNNs in the road detection scenario obtain high accuracy due to the automatic feature extraction of the objects (deep levels of abstraction), flexibility (e.g., possibility to use multisource or multimodal data), scalability to work with a large amount of data and focus on the study over data of multiple arrays, such as images. Liu et al. (2022a), from another viewpoint, suggest that, despite the advantages, the high accuracy is related to the use of benchmark sets, with the aim of comparison of models. This implies that if there is a set of training under specific conditions aiming at a test inference (unknown set by the network), an accuracy decrease can be obtained, or the model cannot generalize for the testing set, i.e., an overfitting (Xu et al., 2022; Peng et al., 2022).

Usually, a dataset is split to train, validate, and test a model. In principle, it is expected that the training data have enough variability for generalization at the inference (test). However, in practice, this may not hold for further inferences. For example, the road network can appear under different scenarios: paved, unpaved, or vicinal roads; urban or rural areas; differences in scale or spatial resolution (scale-space factor); and other aspects (e.g., acquisition method or conditions). These scenarios, emphasized mainly by the local context aspect, establish that there is a set for training and another for testing with different characteristics, which define, respectively, a source domain and a target domain. This divergence between domains is caused, in general, by different probability distributions of the data that generate a set. From the statistic concept, this is known as domain shift (Ben-David et al., 2006; Daume; Marcu, 2011).

A strategy to mitigate the domain shift effect would be to use the transfer learning (TL) technique (Pan; Yang, 2010; Yosinski et al., 2014), which involves pre-training a CNN on the source domain and fine-tuning it on the target domain. However, for this, it would be necessary to label some data of the target domain, which elucidates another limitation: in the RS&P area, the habitual is the availability of raw data, without labels of objects. It becomes fundamental, hence, to find viable solutions to minimize the domain shift problem, with the possibility that learned characteristics of a labeled set lead to the segmentation of the road automatically without the necessity of manual labeling.

Domain Adaptation (DA), which is a particularity of TL, is a manner to approach these aforementioned challenges (Patel et al., 2015; Tuia; Persello; Bruzzone, 2016; Wang; Deng, 2018). More specifically, to the unsupervised case (Unsupervised Domain Adaptation - UDA), where the adaptation needs labeled data just in the source domain, in order to infer in the target domain without the necessity of labeling (Ganin et al., 2016; Toldo et al., 2020).

According to Patel et al. (2015), DA is a fundamental problem in ML, but only in 2015 was explored in the Computer Vision area, mainly for the classification of objects via UDA (Ganin et al., 2016) and, later, for semantic segmentation (Hoffman et al., 2016). In the RS&P area, as a consequence, the approach is recent, with most advances since 2020 (Xu et al., 2022).

As for the specific topic of road detection, few studies have been made by UDA according to the literature (Iqbal et al., 2023). While CNNs have shown potential to handle many inherent variations in road characteristics, challenges regarding the divergence between domains (e.g., location, contextual model of roads, spatial resolution, etc.), and reducing labeling efforts, exploring the use of CNNs under the UDA paradigm still suggests several challenges. Addressing this gap presents a fundamental opportunity for research to achieve higher accuracy in the detection of the road network through DL.

Motivated by this context, this work presents a tailored adaptation of the U-Net architecture to the UDA setting by leveraging the Domain Adversarial Neural Network (DANN) framework. Our UDA-based road detection approach particularly targets domain shift challenges that arise from variations in geographical context (e.g., urban vs. suburban areas) and spatial resolution, aiming to enhance road detection performance while minimizing the need for manual labeling.

The subsequent section reviews the current state-of-the-art in DL for road detection and UDA approaches relevant to these challenges.

1.2 State-of-the-art

Several approaches for road extraction and road detection¹ have emerged since the pioneering work by Bajcsy and Tavakoli (1976). In this context, the road extraction task is related to two subtasks: delineation and recognition. By convention, these subtasks are categorized as automatic or semiautomatic approaches. According to Dal Poz, Zanin and Do Vale (2006), semiautomatic methods address only the geometric delineation of the roads (algorithm), leaving the high-level decisions (i.e., recognition) to a human operator to set the meaning "road" to the object, while the automatic methods, in turn, address both subtasks.

Despite the traditional methods that are densified in the literature, such as road tracking, active contour models (snakes), filtering and grouping, morphological profile analysis, and dynamic programming (Dal Poz et al., 2012; Wang et al., 2016), since the 2010s, the classification methods have been predominant. According to Abdollahi et al. (2020), classification methods, supervised or unsupervised, are automatic and based on

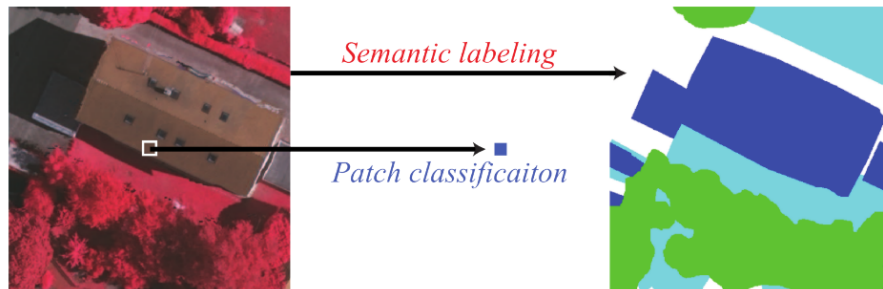
¹ We use the terms road extraction and road detection interchangeably, as is common in the literature.

ML/DL techniques. These DL-based methods involve, mainly, the road detection (i.e., surface) or the road centerline extraction (e.g., skeleton recognition).

Road detection belongs to the task of semantic segmentation or classification at the pixel level² in DL. It is worth noticing that based on the traditional Digital Image Processing (DIP), image segmentation is one of the most difficult tasks in this field (Gonzalez; Woods, 2008). For this purpose, DL-based algorithms were initially approached by patch CNNs and later extended to deep CNNs.

With patch CNNs, the road detection task suffered from the problem of little context in the patch, or on the other hand, by computational resources (larger patch), as stated by Li et al. (2016). However, due to the advances in DL, fully convolutional architectures (Fully Convolutional Networks - FCNs) revolutionized the context of semantic segmentation in the Computer Vision field, learning densely predict semantic labels for each location instead of a single label per patch (Figure 2). A landmark for this is the work of Long, Shelhamer and Darrell (2015), and consequently, for the road detection topic (Zhong et al., 2016; Henry; Azimi; Merkle, 2018).

Figure 2 – Comparing patch classification and semantic labeling.



Source: Volpi and Tuia (2016).

Inspired by the FCN, the U-Net architecture, from the work by Ronneberger, Fischer and Brox (2015), provided several developments in the process of semantic segmentation in general, which contextualizes the term of deep CNNs. Essentially, due to the strategy of skip connections, which can recover spatial information from the lower level up to the higher level of abstraction to reconstruct the masks more accurately. These architectures are named as encoder-decoder type.

In this context, Cheng et al. (2017) under the multitask aspect, proposed an encoder-decoder CNN to detect the road and its centerline simultaneously; Lu et al. (2022), in the same task, added the edge of the roads. Zhang, Liu and Wang (2018), Bandara, Valanarasu and Patel (2022) and Wang et al. (2022) explored the residual learning in deeper CNNs to mitigate the vanishing gradient, using, exclusively, the U-Net architecture. In parallel, Filho (2022) innovated the road extraction topic by fusing multimodal data in

² Explained in Subsection 2.1.4.

DL, using orthoimages and LiDAR (Light Detection and Ranging) data with the U-Net network.

Recently, to make the performance of DL-based models more robust, these deep CNNs have been adopting the strategy of scene context. Those approaches are mainly based on the attention mechanism to get more information about the structure of the roads (Xu et al., 2018; Shao et al., 2022; Dai; Zhang; Zhang, 2023; Qiu et al., 2023). Such approaches can be extended to Transformer-based models, which might be the most recent approach for this end (Xiong et al., 2024; Zhu et al., 2024). In a few words, Transformer-based models can aggregate rich spatial contextual (global and local) information on the roads, with the ability to capture long-range relationships, contrasting with CNNs which are focused on localized information (Gao et al., 2024).

Although current approaches, mainly based on encoder-decoder architectures, adopt different strategies to achieve higher detection accuracy, they still require a large amount of labeled data for high performance when inferring. Additionally, when in the presence of a domain shift effect, the accuracy can significantly drop. The UDA technique, to overcome these challenges, has recently attracted the attention of the community, enabling the segmentation of objects without the need for labeling on the model inference.

To contextualize, DA is a traditional problem for ML (Ben-David et al., 2006; Daume; Marcu, 2011; Glorot; Bordes; Bengio, 2011b; Chopra; Balakrishnan; Gopalan, 2013), but only in the last decade have gained traction in the Computer Vision field (Patel et al., 2015; Xu et al., 2022). The breakthrough for DA in DL was marked by the possibility of extracting deep learning features and minimizing the domain shift into one process, and also the possibility of adapting the domain without labels. This concerns the most difficult case of DA, the UDA approach.

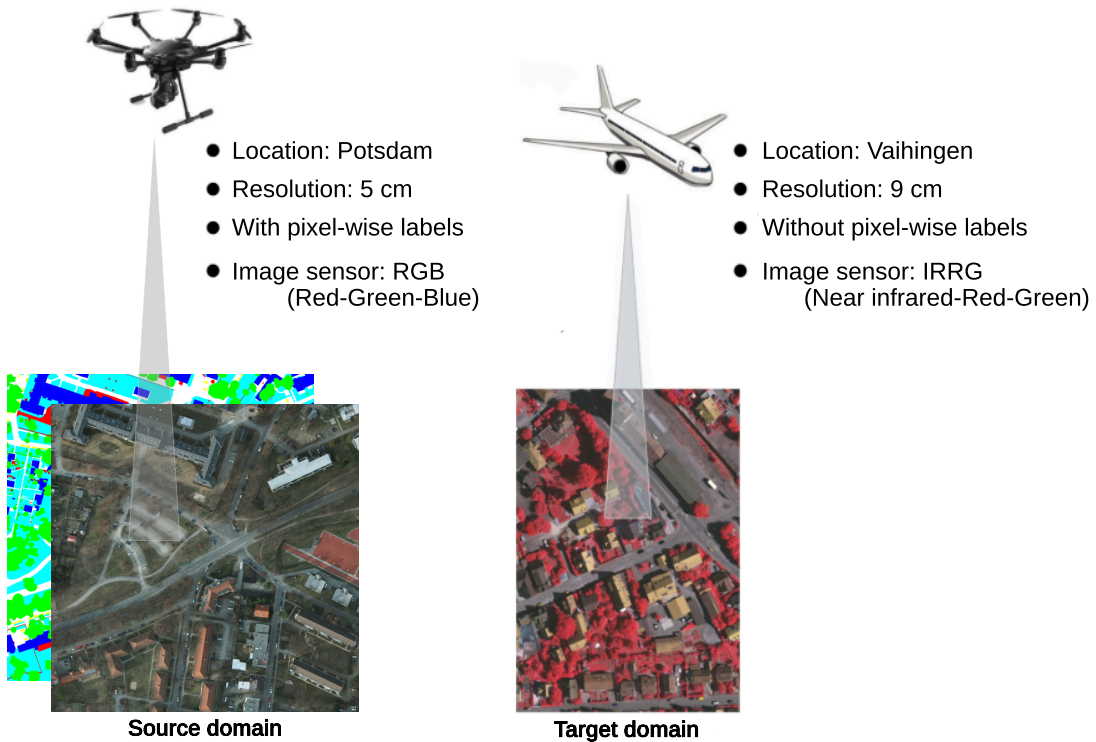
The study by Ganin et al. (2016) was a landmark for UDA, and later by Hoffman et al. (2016) and Tsai et al. (2018) for semantic segmentation. Since then, several algorithms for UDA have emerged. Essentially, UDA can be categorized as representation matching and appearance adaptation networks. The first category aims to align the extracted features between both domains (source and target) in a domain-invariant latent space (Ganin et al., 2016; Wittich; Rottensteiner, 2019). Appearance adaptation, the second one, is associated with methods that transform images from a target domain so that the generated images have a style (appearance) that is similar to the images from a source domain; the *image-to-image* (I2I) term can also be found (Wittich; Rottensteiner, 2021; Vega et al., 2021).

In both approaches, the adversarial training is the most used (Goodfellow et al., 2014). It includes a classifier to predict the samples of a domain, and a discriminator to distinguish from each domain the sample belongs to. The domain shift effect is, then, adversarially mitigated, like a "game" of min-max, between the generator/classifier and

the discriminator of the network (Chen; Pan; Chong, 2022).

In the RS&P community, the domain gap includes, mainly, the location (e.g. scene context), the spatial resolution, the system sensor, the condition of the acquisition, and other factors, as exemplified in Figure 3. According to the literature, the main applications have been in the land use and land cover classification (Tuia; Persello; Bruzzone, 2016; Elshamli et al., 2017; Benjdira et al., 2019), with methodological focuses in a more recent period (Xu et al., 2022).

Figure 3 – Unsupervised Domain Adaptation in the RS&P area.



Source: Adapted from Chen et al. (2022).

Based on several aspects, different approaches have been proposed for the multiclass problem in high-resolution images, such as: in terms of multimodal fusion (RGB images and Digital Models, e.g., Digital Surface Model) (Zhou et al., 2023), attention blocks to extract features in different levels between the domains (Chen; Pan; Chong, 2022; Chen et al., 2022), the use of transformers (Liu et al., 2022b; Ma et al., 2023), and also, for different epochs, like the change detection case (Vega et al., 2021; Vega et al., 2022).

In the specific topic of road detection, few studies have been made by optic data through UDA. Shamsolmoali et al. (2020) addressed UDA in adversarial training with the use of a Feature Pyramid Network (FPN) to minimize the difference between the source and target domain, via a multiscale strategy. Later, Zhang et al. (2021) utilized this FPN with a self-training stage to mitigate the intra-domain discrepancy in the target domain (generation of pseudo labels with evaluation of quality). Lu et al. (2021) introduced an

alignment at a local and global level, in order to adaptively adjust the weights of each domain, in a cross-domain case. Iqbal et al. (2023) proposed a simultaneous task: road and centerline detection, with the aid of constraints on the topology of the road network via loss function for domain adaptation.

Finally, in UDA, the work by Ganin et al. (2016), namely DANN, has encouraged future research, due to the advantages of adaption to DL frameworks, and specifically, in the backpropagation calculus. In RS&P area, Elshamli et al. (2017) were pioneers in this sense and, later, Vega et al. (2022). However, the authors do not address the specific cases of roads, nor the deep CNNs usage for each particular purpose. This gap highlights an opportunity to advance the state-of-the-art in semantic segmentation in general, and in road detection in particular.

1.3 Objectives

1.3.1 General

Developing a method for automatic road detection using UDA, leveraging the DANN framework in urban and suburban environment.

1.3.2 Specifics

- Adapting the U-Net architecture to UDA for road detection in urban and suburban environment;
- Evaluating other semantic segmentation models adapted with the DANN framework;
- Evaluating the performance of road detection with and without the application of UDA using the DANN framework;
- Evaluating qualitatively the segmentation results based on the UDA approach.

1.4 Contributions

Addressing the critical gap in UDA for road detection in aerial imagery, this work makes the following key contributions:

- We extend the DANN framework, originally designed for object classification, to the task of high-resolution aerial road segmentation according to U-Net, addressing domain shift challenges caused by variations in geographical context and spatial resolution;

- We demonstrate the flexibility of the proposed approach by adapting two additional architectures (DeepLabV3+ and Attention U-Net), and also handling very high-resolution images;
- We conduct an in-depth analysis of hyperparameters tuning, with particular focus on the critical DANN lambda value. Our analysis provides insights into optimizing adversarial training stability for road detection;
- We perform an interpretability analysis of the UDA process through the Score-CAM method, offering insights into how the model behaves when inferring on unlabeled target domain data.

1.5 Justification and relevance

The first relevance of the dissertation is evidenced by the International Society for Photogrammetry and Remote Sensing (ISPRS) on current and relevant topics. As the Commission II (Photogrammetry), from the Working Group II/IV, the following terms of reference are promising: 1) machine learning, and deep learning; 2) scene classification, semantic (instance) segmentation, and object detection; and 3) unsupervised, supervised, weakly supervised, transfer, and self-supervised learning.

The second relevance is highlighted by three aspects: 1) road detection is fundamental to public policies, private activities, and the scientific community, once the semantic segmentation is one of the most challenging topics in DL; 2) this work mitigates the necessity of large labeled datasets, which in turn, it is a time-consuming process and, therefore, addresses one of the main problems in DL, bringing innovation for the specific topic of road detection and segmentation in general; and 3) the solution is attractive in practical terms, as DL algorithms need high computational complexity, to get models more capable for generalization, as the UDA case, reduces the necessity of other training like transfer learning.

1.6 Master's thesis overview

This master's thesis is structured in five chapters. Chapter 2 presents the theoretical foundation, in order to provide support to DL, and DA methods. Chapter 3 describes the materials and the proposed method in this research. Chapter 4 presents the results as well as the analyses. Chapter 5 concludes the research with directions for future works. Lastly, the selected references are presented.

5 Conclusion and Future Works

This work aimed to develop a method for automatic road detection by adapting the U-Net architecture to UDA in an urban environment. The approach, which is based on the DANN framework, supposes that available labeled road data in a domain can achieve promising road network detection on unlabeled target images belonging to another domain even under the domain shift effect, such as geographical environments and image spatial resolution.

According to the obtained results, the proposed method based on the U-Net architecture was able to accomplish the objectives of the research. The developed framework, initially aimed at U-Net, was later extended to other encoder-decoder models, which indicates adaptability of the DANN framework to other models while maintaining the overall training setting to UDA.

The proposed approach could outperform the source-trained model (i.e., baseline) in all three experiments, which reaches the main purpose of UDA. Although the DF dataset is the main study area of this MSc research, the other experiments provided meaningful insights into how UDA works in different settings, such as prototyping steps to dive deeper into DL for UDA, and its performance depending on the architecture choice.

The approach demonstrates the possibility of dealing with very high-resolution images, as in Experiment II in Section 4.2 handling UAV imagery. Although particularized for this datatype, the experiment leaves as a takeaway that the UDA gains vary depending on architecture. For instance, ResU-Net and Att-ResU-Net benefited more from UDA due to the skip connections and attention mechanisms strategies than multi-scale context from DeepLabv3+. Furthermore, performance remains limited by intrinsic the data involved; UAV imagery challenges remain even in supervised settings, indicating a fundamental difficulty in modeling such data, rather than a failure of the UDA approach alone.

The UDA method also reveals that although the domains might have different probability distributions, the larger the domain shift, the more effective the adaptation, once they are still related. Experiment III in Section 4.3 suggested that when domains share similar characteristics (e.g., the same GSD), the UDA brings minor improvements. However, when domain shift enlarges, the baseline model degrades, while the UDA model improves performance, as evaluated by an increase of 3–4% in $F1_{\text{Score}}$ and IoU metrics. In addition, the λ term tuning is critical. The original DANN implementation suggests a schedule but our study indicated that an empirical evaluation is necessary, as a small and fixed λ achieved the optimal value.

The UDA feature extraction process highlights that the model relies primarily on

geometric elements or context information to produce domain-invariant features. The reconstruction process, likewise, remains focused on these domain-agnostic features and suppresses road extraction-relevant characteristics to achieve better generalization. On the other hand, this pattern may lead to a potential limitation. We would expect that the model prioritizes road-related features, given the road extraction task. However, this reconstruction strategy along with the absence of task-specific guidance may hinder road-focused segmentation to enhance its performance.

Moreover, while UDA exempts the need for target domain labels during training, this advantage might become another limitation. Since UDA relies on adversarial learning, which is inherently challenging to optimize, the absence of target labels may complicate real-time performance assessments and hyperparameter tuning. Although all labels were available for evaluation in this study, in real-world scenarios where such labels are unavailable, fine-tuning might be more difficult.

According to these considerations, we suggest a few directions for future work:

- This study adopts a fully unsupervised domain adaptation approach, relying solely on unlabeled target data. However, when a small set of target labels is available, semi-supervised strategies may improve performance. Alternatively, self-training with pseudo-labels can serve as a practical solution in low-label scenarios;
- Our approach, which is based on the original DANN implementation, employs a single discriminator at the output of the encoder. As the literature suggests, future work could address multi-level discriminators or place the discriminator at the decoder output, analyzing the segmentation road maps in the original input dimension;
- While UDA encourages the learning of domain-invariant features, this may come at the cost of task specificity. Incorporating attention mechanisms or road auxiliary tasks (e.g., road centerline, road boundaries) may help the decoder toward more road-focused representations, potentially improving segmentation accuracy.

References

- Abdollahi, A. et al. Deep Learning Approaches Applied to Remote Sensing Datasets for Road Extraction: A State-Of-The-Art Review. *Remote Sensing 2020, Vol. 12, Page 1444*, Multidisciplinary Digital Publishing Institute, v. 12, n. 9, p. 1444, may 2020. ISSN 2072-4292. Available on: <<https://www.mdpi.com/2072-4292/12/9/1444/htm>>. Cited 2 times on pages 18 and 21.
- Abraham, N.; Khan, N. M. A Novel Focal Tversky loss function with improved Attention U-Net for lesion segmentation. *Proceedings - International Symposium on Biomedical Imaging*, IEEE Computer Society, v. 2019-April, p. 683–687, oct 2018. ISSN 19458452. Available on: <<https://arxiv.org/abs/1810.07842v1>>. Cited on page 48.
- Bajcsy, R.; Tavakoli, M. Computer Recognition of Roads from Satellite Pictures. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6, n. 9, p. 623–637, 1976. ISSN 0018-9472, 2168-2909. Available on: <<http://ieeexplore.ieee.org/document/4309568/files/35/BajcsyTavakoli-1976-ComputerRecognitionofRoadsfromSatellitePictu.pdf>>. Cited 2 times on pages 19 and 21.
- Ballesteros, J. R.; Sanchez-Torres, G.; Branch-Bedoya, J. W. A gis pipeline to produce geoai datasets from drone overhead imagery. *ISPRS International Journal of Geo-Information 2022, Vol. 11, Page 508*, Multidisciplinary Digital Publishing Institute, v. 11, p. 508, 9 2022. ISSN 2220-9964. Available on: <<https://www.mdpi.com/2220-9964/11/10/508/htm>>. Cited on page 60.
- Bandara, W. G. C.; Valanarasu, J. M. J.; Patel, V. M. Spin road mapper: Extracting roads from aerial images via spatial and interaction space graph reasoning for autonomous driving. In: IEEE. *2022 International Conference on Robotics and Automation (ICRA)*. 2022. p. 343–350. Cited on page 22.
- Baumgartner, A.; Hinz, S. Multi-Scale Road Extraction Using Local and Global Grouping Criteria. In: *International Archives of Photogrammetry and Remote Sensing*. Amsterdam: , 2000. Cited 2 times on pages 18 and 19.
- Ben-David, S. et al. Analysis of Representations for Domain Adaptation. *Advances in Neural Information Processing Systems*, v. 19, 2006. Cited 4 times on pages 20, 23, 54, and 55.
- Ben-David, S. et al. A theory of learning from different domains. *Machine Learning*, Kluwer Academic Publishers, v. 79, n. 1-2, p. 151–175, oct 2010. ISSN 15730565. Available on: <<https://link.springer.com/article/10.1007/s10994-009-5152-4>>. Cited on page 54.
- Bengio, Y. Deep Learning of Representations for Unsupervised and Transfer Learning. v. 27, p. 17–37, 2012. Available on: <<https://dl.acm.org/doi/10.5555/3045796.3045800>>. Cited on page 52.
- Benjdira, B. et al. Unsupervised Domain Adaptation Using Generative Adversarial Networks for Semantic Segmentation of Aerial Images. *Remote Sensing 2019, Vol. 11, Page 1369*, Multidisciplinary Digital Publishing Institute, v. 11, n. 11, p. 1369, jun 2019.

ISSN 2072-4292. Available on: <<https://www.mdpi.com/2072-4292/11/11/1369/htm>>. Cited on page 24.

Boser, B. E.; Guyon, I. M.; Vapnik, V. N. A training algorithm for optimal margin classifiers. In: *Proceedings of the fifth annual workshop on Computational learning theory*. 1992. p. 144–152. Cited on page 29.

Breiman, L. Random forests. *Machine learning*, Springer, v. 45, p. 5–32, 2001. Cited on page 29.

Caruana, R. Multitask Learning. *Machine Learning*, Springer Netherlands, v. 28, n. 1, p. 41–75, 1997. ISSN 08856125. Available on: <<https://link.springer.com/article/10.1023/A:1007379606734>>. Cited on page 51.

Chen, J. et al. Unsupervised Domain Adaptation for Semantic Segmentation of High-Resolution Remote Sensing Imagery Driven by Category-Certainty Attention. *IEEE Transactions on Geoscience and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 60, 2022. ISSN 15580644. Cited on page 24.

Chen, L.-C. et al. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: *Proceedings of the European conference on computer vision (ECCV)*. 2018. p. 801–818. Cited 3 times on pages 45, 46, and 47.

Chen, X.; Pan, S.; Chong, Y. Unsupervised Domain Adaptation for Remote Sensing Image Semantic Segmentation Using Region and Category Adaptive Domain Discriminator. *IEEE Transactions on Geoscience and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 60, 2022. ISSN 15580644. Cited on page 24.

Cheng, G. et al. Automatic Road Detection and Centerline Extraction via Cascaded End-to-End Convolutional Neural Network. *IEEE Transactions on Geoscience and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 55, n. 6, p. 3322–3337, jun 2017. ISSN 01962892. Cited 7 times on pages 10, 13, 17, 22, 59, 60, and 72.

Chopra, S.; Balakrishnan, S.; Gopalan, R. Dlid: Deep learning for domain adaptation by interpolating between domains. In: Citeseer. *ICML workshop on challenges in representation learning*. 2013. v. 2, n. 6. Cited 2 times on pages 23 and 55.

Dai, L.; Zhang, G.; Zhang, R. Radanet: Road augmented deformable attention network for road extraction from complex high-resolution remote-sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 61, 2023. ISSN 15580644. Cited on page 23.

Dal Poz, A. P. et al. Object-space road extraction in rural areas using stereoscopic aerial images. *IEEE Geoscience and Remote Sensing Letters*, v. 9, n. 4, p. 654–658, 2012. ISSN 1545598X. Cited on page 21.

Dal Poz, A. P.; Zanin, R. B.; Do Vale, G. M. Automated extraction of road network from medium-and high-resolution images. *Pattern Recognition and Image Analysis*, Springer, v. 16, n. 2, p. 239–248, apr 2006. ISSN 10546618. Available on: <<https://link.springer.com/article/10.1134/S1054661806020118>>. Cited 2 times on pages 19 and 21.

- Daume, H.; Marcu, D. Domain Adaptation for Statistical Classifiers. *Journal Of Artificial Intelligence Research*, American Association for Artificial Intelligence, v. 26, p. 101–126, sep 2011. Available on: <<http://arxiv.org/abs/1109.6341>>. Cited 3 times on pages 20, 23, and 52.
- Duchi, J.; Hazan, E.; Singer, Y. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of machine learning research*, v. 12, n. 7, 2011. Cited on page 40.
- Elshamli, A. et al. Domain Adaptation Using Representation Learning for the Classification of Remote Sensing Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Institute of Electrical and Electronics Engineers, v. 10, n. 9, p. 4198–4209, sep 2017. ISSN 21511535. Cited 2 times on pages 24 and 25.
- Farabet, C. et al. Learning hierarchical features for scene labeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 35, n. 8, p. 1915–1929, 2013. ISSN 01628828. Cited on page 42.
- Filho, A. G. G. *Deep Convolutional Neural Network (DCNN) para extração automática de rede viária a partir da FUSÃO de dados varredura LASER aerotransportado e imagens de altíssima resolução em ambiente urbano*. Tese (Doutorado) — FCT/UNESP, aug 2022. Available on: <<https://repositorio.unesp.br/handle/11449/237021>>. Cited 14 times on pages 19, 20, 22, 27, 33, 35, 36, 48, 58, 59, 61, 62, 77, and 78.
- Fukushima, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics*, Springer, v. 36, n. 4, p. 193–202, 1980. Cited on page 28.
- Ganin, Y.; Lempitsky, V. Unsupervised Domain Adaptation by Backpropagation. *32nd International Conference on Machine Learning, ICML 2015*, International Machine Learning Society (IMLS), v. 2, p. 1180–1189, sep 2014. Available on: <<https://arxiv.org/abs/1409.7495v2>>. Cited on page 55.
- Ganin, Y. et al. Domain-Adversarial Training of Neural Networks. *Advances in Computer Vision and Pattern Recognition*, Springer London, v. 17, n. 9783319583464, p. 189–209, may 2016. ISSN 21916594. Available on: <<https://arxiv.org/abs/1505.07818v4>>. Cited 11 times on pages 20, 21, 23, 25, 52, 53, 54, 55, 56, 66, and 68.
- Gao, L. et al. Ddctnet: A deformable and dynamic cross-transformer network for road extraction from high-resolution remote sensing images. *IEEE Transactions on Geoscience and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 62, p. 1–19, 2024. ISSN 15580644. Cited on page 23.
- Garcia-Garcia, A. et al. A survey on deep learning techniques for image and video semantic segmentation. *Applied Soft Computing*, Elsevier, v. 70, p. 41–65, sep 2018. ISSN 1568-4946. Cited on page 43.
- Glorot, X.; Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. In: *Proceedings of the Thirteenth International Confer; PMLR*. 2010. p. 9–246–256. Cited 2 times on pages 35 and 66.
- Glorot, X.; Bordes, A.; Bengio, Y. Deep Sparse Rectifier Neural Networks. *undefined*, 2011. Cited on page 33.

- Glorot, X.; Bordes, A.; Bengio, Y. Domain adaptation for large-scale sentiment classification: A deep learning approach. In: *Proceedings of the 28th international conference on machine learning (ICML-11)*. 2011. p. 513–520. Cited 2 times on pages 23 and 55.
- Gonzalez, R. C.; Woods, R. E. *Digital image processing*. Upper Saddle River, N.J.: Prentice Hall, 2008. ISBN 9780131687288 013168728X 9780135052679 013505267X. Cited 2 times on pages 22 and 42.
- Goodfellow, I. et al. Generative Adversarial Networks. *Communications of the ACM*, Association for Computing Machinery, v. 63, n. 11, p. 139–144, jun 2014. ISSN 15577317. Available on: <<https://arxiv.org/abs/1406.2661v1>>. Cited on page 23.
- Goodfellow, I. J.; Bengio, Y.; Courville, A. *Deep Learning*. : MIT Press, 2016. Cited 11 times on pages 27, 28, 29, 30, 31, 32, 33, 37, 38, 39, and 40.
- Haykin, S. *Neural Networks and Learning Machines*. : Pearson, 2009. (Pearson International Edition). ISBN 9780131293762. Cited 3 times on pages 28, 29, and 30.
- He, K. et al. Deep Residual Learning for Image Recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, v. 2016-Decem, p. 770–778, dec 2015. ISSN 10636919. Available on: <<https://arxiv.org/abs/1512.03385v1>>. Cited 5 times on pages 34, 35, 36, 37, and 67.
- Henry, C.; Azimi, S. M.; Merkle, N. Road segmentation in sar satellite images with deep fully convolutional neural networks. *IEEE Geoscience and Remote Sensing Letters*, Institute of Electrical and Electronics Engineers (IEEE), v. 15, n. 12, p. 1867–1871, dez. 2018. ISSN 1558-0571. Available on: <<http://dx.doi.org/10.1109/LGRS.2018.2864342>>. Cited on page 22.
- Hinton, G. E.; Osindero, S.; Teh, Y. W. A fast learning algorithm for deep belief nets. *Neural computation*, Neural Comput, v. 18, n. 7, p. 1527–1554, 2006. ISSN 0899-7667. Available on: <<https://pubmed.ncbi.nlm.nih.gov/16764513/>>. Cited on page 29.
- Hoeser, T.; Kuenzer, C. Object Detection and Image Segmentation with Deep Learning on Earth Observation Data: A Review-Part I: Evolution and Recent Trends. *Remote Sensing 2020, Vol. 12, Page 1667*, Multidisciplinary Digital Publishing Institute, v. 12, n. 10, p. 1667, may 2020. ISSN 2072-4292. Available on: <<https://www.mdpi.com/2072-4292/12/10/1667/htm>>. Cited 2 times on pages 35 and 43.
- Hoffman, J. et al. FCNs in the Wild: Pixel-level Adversarial and Constraint-based Adaptation. dec 2016. Available on: <<https://arxiv.org/abs/1612.02649v1>>. Cited 2 times on pages 21 and 23.
- Hornik, K. Approximation capabilities of multilayer feedforward networks. *Neural Networks*, Pergamon, v. 4, n. 2, p. 251–257, jan 1991. ISSN 0893-6080. Cited on page 30.
- Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. *32nd International Conference on Machine Learning, ICML 2015*, International Machine Learning Society (IMLS), v. 1, p. 448–456, feb 2015. Available on: <<https://arxiv.org/abs/1502.03167v3>>. Cited 2 times on pages 35 and 65.

- Iqbal, J. et al. Leveraging topology for domain adaptive road segmentation in satellite and aerial imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier, v. 206, p. 106–117, dec 2023. ISSN 0924-2716. Cited 2 times on pages 21 and 25.
- Jadon, S. A survey of loss functions for semantic segmentation. *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB 2020*, Institute of Electrical and Electronics Engineers Inc., oct 2020. Cited 2 times on pages 47 and 49.
- Jordan, M. I. et al. An introduction to variational methods for graphical models. *Learning in graphical models*, Springer, p. 105–161, 1998. Cited on page 29.
- Khan, R. *Where did the Binary Cross-Entropy Loss Function come from? Towards Data Science*. 2019. Available on: <<https://towardsdatascience.com/where-did-the-binary-cross-entropy-loss-function-come-from-ac3de349a715>>. Cited on page 43.
- Kingma, D. P.; Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. Cited 2 times on pages 40 and 41.
- Krizhevsky, A.; Sutskever, I.; Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, v. 25, 2012. Cited 2 times on pages 29 and 34.
- Lecun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* 2015 521:7553, Nature Publishing Group, v. 521, n. 7553, p. 436–444, may 2015. ISSN 1476-4687. Available on: <<https://www.nature.com/articles/nature14539>>. Cited 4 times on pages 20, 31, 32, and 33.
- LeCun, Y. et al. Backpropagation Applied to Handwritten Zip Code Recognition. *Neural Computation*, MIT Press - Journals, v. 1, n. 4, p. 541–551, dec 1989. ISSN 0899-7667. Cited on page 29.
- LeCun, Y. et al. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, v. 86, n. 11, p. 2278–2323, 1998. ISSN 00189219. Cited on page 29.
- Li, P. et al. Road network extraction via deep learning and line integral convolution. In: IEEE. *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. 2016. p. 1599–1602. Cited on page 22.
- Lin, T.-Y. et al. Focal loss for dense object detection. In: *Proceedings of the IEEE international conference on computer vision*. 2017. p. 2980–2988. Cited on page 48.
- Liu, P. et al. Survey of Road Extraction Methods in Remote Sensing Images Based on Deep Learning. *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science* 2022 90:2, Springer, v. 90, n. 2, p. 135–159, feb 2022. ISSN 2512-2819. Available on: <<https://link.springer.com/article/10.1007/s41064-022-00194-z>>. Cited on page 20.
- Liu, W. et al. Bispase Domain Adaptation Network for Remotely Sensed Semantic Segmentation. *IEEE Transactions on Geoscience and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 60, 2022. ISSN 15580644. Cited on page 24.

- Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, v. 07-12-June, p. 431–440, oct 2015. ISSN 10636919. Cited 3 times on pages 22, 43, and 44.
- Lu, X. et al. Cross-domain road detection based on global-local adversarial learning framework from very high resolution satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier, v. 180, p. 296–312, oct 2021. ISSN 0924-2716. Cited on page 24.
- Lu, X. et al. Cascaded Multi-Task Road Extraction Network for Road Surface, Centerline, and Edge Extraction. *IEEE Transactions on Geoscience and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 60, p. 1–14, 2022. ISSN 0196-2892. Available on: <<https://ieeexplore.ieee.org/document/9756441/>>. Cited on page 22.
- Ma, J. Segmentation Loss Odyssey. may 2020. Available on: <<https://arxiv.org/abs/2005.13449v1>>. Cited 3 times on pages 47, 48, and 49.
- Ma, L. et al. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier, v. 152, p. 166–177, jun 2019. ISSN 0924-2716. Cited on page 20.
- Ma, X. et al. Unsupervised Domain Adaptation Augmented by Mutually Boosted Attention for Semantic Segmentation of VHR Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 61, 2023. ISSN 15580644. Cited on page 24.
- McCulloch, W. S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics*, Springer, v. 5, p. 115–133, 1943. Cited on page 28.
- Merkel, D. et al. Docker: lightweight linux containers for consistent development and deployment. *Linux j*, v. 239, n. 2, p. 2, 2014. Cited on page 63.
- Milletari, F.; Navab, N.; Ahmadi, S. A. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. *Proceedings - 2016 4th International Conference on 3D Vision, 3DV 2016*, Institute of Electrical and Electronics Engineers Inc., p. 565–571, jun 2016. Available on: <<https://arxiv.org/abs/1606.04797v1>>. Cited on page 48.
- Minsky, M.; Papert, S. An introduction to computational geometry. *Cambridge tiass., HIT*, v. 479, n. 480, p. 104, 1969. Cited on page 28.
- Mnih, V.; Hinton, G. E. Learning to detect roads in high-resolution aerial images. In: Springer. *Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part VI 11*. 2010. p. 210–223. Cited on page 19.
- Oktay, O. et al. Attention u-net: Learning where to look for the pancreas. 4 2018. Available on: <<https://arxiv.org/abs/1804.03999v3>>. Cited on page 68.
- Pan, S. J.; Yang, Q. A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*, v. 22, n. 10, p. 1345–1359, 2010. ISSN 10414347. Cited 4 times on pages 20, 50, 51, and 52.

- Patel, V. M. et al. Visual Domain Adaptation: A survey of recent advances. *IEEE Signal Processing Magazine*, Institute of Electrical and Electronics Engineers Inc., v. 32, n. 3, p. 53–69, may 2015. ISSN 10535888. Cited 3 times on pages 20, 21, and 23.
- Patterson, J.; Gibson, A. *Deep learning: A practitioner's approach*. : " O'Reilly Media, Inc.", 2017. Cited on page 27.
- Peng, J. et al. Domain Adaptation in Remote Sensing Image Classification: A Survey. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 15, p. 9842–9859, 2022. ISSN 21511535. Cited on page 20.
- Qiu, L. et al. A semantics-geometry framework for road extraction from remote sensing images. *IEEE Geoscience and Remote Sensing Letters*, Institute of Electrical and Electronics Engineers Inc., v. 20, 2023. ISSN 15580571. Cited on page 23.
- Quinlan, J. R. Induction of decision trees. *Machine learning*, Springer, v. 1, p. 81–106, 1986. Cited on page 29.
- Raina, R. et al. Self-taught learning: Transfer learning from unlabeled data. *ACM International Conference Proceeding Series*, v. 227, p. 759–766, 2007. Available on: <<https://dl.acm.org/doi/10.1145/1273496.1273592>>. Cited on page 51.
- Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv:1505.04597 [cs]*, sep 2015. Available on: <<http://arxiv.org/abs/1505.04597>>. Cited 3 times on pages 22, 44, and 45.
- Rosenblatt, F. The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological review*, American Psychological Association, v. 65, n. 6, p. 386, 1958. Cited on page 28.
- Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning representations by back-propagating errors. *nature*, Nature Publishing Group UK London, v. 323, n. 6088, p. 533–536, 1986. Cited 2 times on pages 28 and 41.
- Salehi, S. S. M.; Erdogmus, D.; Gholipour, A. Tversky loss function for image segmentation using 3d fully convolutional deep networks. In: Springer. *International workshop on machine learning in medical imaging*. 2017. p. 379–387. Cited on page 48.
- Shamsolmoali, P. et al. Road Segmentation for Remote Sensing Images using Adversarial Spatial Pyramid Networks. *IEEE Transactions on Geoscience and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 59, n. 6, p. 4673–4688, aug 2020. Available on: <<http://arxiv.org/abs/2008.04021>>. Cited on page 24.
- Shao, L.; Zhu, F.; Li, X. Transfer learning for visual categorization: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, Institute of Electrical and Electronics Engineers Inc., v. 26, n. 5, p. 1019–1034, may 2015. ISSN 21622388. Cited 2 times on pages 50 and 51.
- Shao, S. et al. Road extraction convolutional neural network with embedded attention mechanism for remote sensing imagery. *Remote Sensing 2022, Vol. 14, Page 2061*, Multidisciplinary Digital Publishing Institute, v. 14, p. 2061, 4 2022. ISSN 2072-4292. Available on: <<https://www.mdpi.com/2072-4292/14/9/2061>>. Cited on page 23.

- Shimodaira, H. Improving predictive inference under covariate shift by weighting the log-likelihood function. *Journal of Statistical Planning and Inference*, North-Holland, v. 90, n. 2, p. 227–244, oct 2000. ISSN 0378-3758. Cited on page 51.
- Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, International Conference on Learning Representations, ICLR, sep 2014. Available on: <<https://arxiv.org/abs/1409.1556v6>>. Cited 3 times on pages 34, 35, and 45.
- Srivastava, N. et al. Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, JMLR. org, v. 15, n. 1, p. 1929–1958, 2014. Cited on page 65.
- Szegedy, C. et al. Going Deeper with Convolutions. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, v. 07-12-June-2015, p. 1–9, sep 2014. ISSN 10636919. Available on: <<https://arxiv.org/abs/1409.4842v1>>. Cited 2 times on pages 34 and 35.
- Tieleman, T.; Hinton, G. Lecture 6.5-rmsprop, coursera: Neural networks for machine learning. *University of Toronto, Technical Report*, v. 6, 2012. Cited on page 40.
- Toldo, M. et al. Unsupervised Domain Adaptation in Semantic Segmentation: a Review. *Technologies*, MDPI, v. 8, n. 2, may 2020. ISSN 22277080. Available on: <<https://arxiv.org/abs/2005.10876v1>>. Cited on page 20.
- Tsai, Y. H. et al. Learning to Adapt Structured Output Space for Semantic Segmentation. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, p. 7472–7481, feb 2018. ISSN 10636919. Available on: <<https://arxiv.org/abs/1802.10349v3>>. Cited on page 23.
- Tuia, D.; Persello, C.; Bruzzone, L. Domain adaptation for the classification of remote sensing data: An overview of recent advances. *IEEE Geoscience and Remote Sensing Magazine*, Institute of Electrical and Electronics Engineers Inc., v. 4, n. 2, p. 41–57, jun 2016. ISSN 21686831. Cited 2 times on pages 20 and 24.
- Vapnik, V.; Levin, E.; Lecun, Y. Measuring the vc-dimension of a learning machine. *Neural Computation*, v. 6, n. 5, p. 851–876, 1994. Cited on page 54.
- Vega, P. J. S. et al. Domain-Adversarial Neural Networks for Deforestation Detection in Tropical Forests. *IEEE Geoscience and Remote Sensing Letters*, Institute of Electrical and Electronics Engineers Inc., v. 19, 2022. ISSN 15580571. Cited 2 times on pages 24 and 25.
- Vega, P. J. S. et al. An unsupervised domain adaptation approach for change detection and its application to deforestation mapping in tropical biomes. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier, v. 181, p. 113–128, nov 2021. ISSN 09242716. Available on: <<https://linkinghub.elsevier.com/retrieve/pii/S092427162100232X>>. Cited 3 times on pages 23, 24, and 51.
- Volpi, M.; Tuia, D. Dense semantic labeling of sub-decimeter resolution images with convolutional neural networks. *IEEE Transactions on Geoscience and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 55, n. 2, p. 881–893, aug 2016. Available on: <<http://dx.doi.org/10.1109/TGRS.2016.2616585>>. Cited on page 22.

- Vosselman, G. et al. Road tracing by profile matching and Kaiman filtering. *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, Birkhäuser Basel, p. 265–274, 1995. Available on: <https://link.springer.com/chapter/10.1007/978-3-0348-9242-1_25>. Cited 2 times on pages 18 and 19.
- Wang, H. et al. Score-cam: Score-weighted visual explanations for convolutional neural networks. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, IEEE Computer Society, v. 2020-June, p. 111–119, 10 2019. ISSN 21607516. Available on: <<https://arxiv.org/abs/1910.01279v2>>. Cited on page 69.
- Wang, M.; Deng, W. Deep visual domain adaptation: A survey. *Neurocomputing*, Elsevier, v. 312, p. 135–153, oct 2018. ISSN 0925-2312. Cited 3 times on pages 20, 52, and 53.
- Wang, W. et al. A review of road extraction from remote sensing images. *Journal of Traffic and Transportation Engineering (English Edition)*, Elsevier, v. 3, n. 3, p. 271–282, jun 2016. ISSN 2095-7564. Cited 2 times on pages 18 and 21.
- Wang, X. et al. Deep learning-based methods for road extraction from remote sensing images: A vision, survey, and future directions. *IEEE Geoscience and Remote Sensing Magazine*, IEEE, 2025. Cited on page 19.
- Wang, Y. et al. DDU-Net: Dual-Decoder-U-Net for Road Extraction Using High-Resolution Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 60, jan 2022. Available on: <<http://dx.doi.org/10.1109/TGRS.2022.3197546>>. Cited on page 22.
- Wittich, D.; Rottensteiner, F. Adversarial Domain Adaptation for the Classification of Aerial Images and Height Data Using CONvolutional Neural Networks. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Copernicus GmbH, IV-2-W7, n. 2/W7, p. 197–204, sep 2019. ISSN 2194-9042. Available on: <<https://doi.org/10.5194/isprs-annals-IV-2-W7-197-2019>>. Cited on page 23.
- Wittich, D.; Rottensteiner, F. Appearance based deep domain adaptation for the classification of aerial images. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier B.V., v. 180, p. 82–102, oct 2021. ISSN 09242716. Cited on page 23.
- Xiong, Y. et al. Pvt-unet: Road extraction in remote sensing imagery based on u-shaped pyramid vision transformer neural network. *ACM International Conference Proceeding Series*, Association for Computing Machinery, p. 199–204, 1 2024. Available on: <<https://dl.acm.org/doi/10.1145/3647649.3647682>>. Cited on page 23.
- Xu, M. et al. The Eyes of the Gods: A Survey of Unsupervised Domain Adaptation Methods Based on Remote Sensing Data. *Remote Sensing 2022, Vol. 14, Page 4380*, Multidisciplinary Digital Publishing Institute, v. 14, n. 17, p. 4380, sep 2022. ISSN 2072-4292. Available on: <<https://www.mdpi.com/2072-4292/14/17/4380>>. Cited 4 times on pages 20, 21, 23, and 24.
- Xu, Y. et al. Building extraction in very high resolution remote sensing imagery using deep learning and guided filters. *Remote Sensing*, MDPI AG, v. 10, n. 1, jan 2018. ISSN 20724292. Cited on page 23.

- Yong, H. et al. Gradient centralization: A new optimization technique for deep neural networks. In: Springer. *Computer Vision—ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I* 16. 2020. p. 635–652. Cited on page 77.
- Yosinski, J. et al. How transferable are features in deep neural networks? *Advances in Neural Information Processing Systems*, Neural information processing systems foundation, v. 4, n. January, p. 3320–3328, nov 2014. ISSN 10495258. Available on: <<https://arxiv.org/abs/1411.1792v1>>. Cited 2 times on pages 20 and 51.
- Zadrozny, B. Learning and evaluating classifiers under sample selection bias. *Proceedings, Twenty-First International Conference on Machine Learning, ICML 2004*, p. 903–910, 2004. Available on: <<https://dl.acm.org/doi/10.1145/1015330.1015425>>. Cited on page 51.
- Zhang, A. et al. *Dive into deep learning*. : Cambridge University Press, 2023. Cited on page 41.
- Zhang, L. et al. Stagewise Unsupervised Domain Adaptation with Adversarial Self-Training for Road Segmentation of Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 60, aug 2021. Available on: <<http://dx.doi.org/10.1109/TGRS.2021.3104032>>. Cited on page 24.
- Zhang, L.; Zhang, L. Artificial Intelligence for Remote Sensing Data Analysis: A review of challenges and opportunities. *IEEE Geoscience and Remote Sensing Magazine*, Institute of Electrical and Electronics Engineers Inc., v. 10, n. 2, p. 270–294, jun 2022. ISSN 21686831. Cited on page 20.
- Zhang, Z.; Liu, Q.; Wang, Y. Road Extraction by Deep Residual U-Net. *IEEE Geoscience and Remote Sensing Letters*, Institute of Electrical and Electronics Engineers Inc., v. 15, n. 5, p. 749–753, may 2018. ISSN 15580571. Cited on page 22.
- Zhong, Z. et al. Fully convolutional networks for building and road extraction: Preliminary results. In: IEEE. *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. 2016. p. 1591–1594. Cited on page 22.
- Zhou, S. et al. DSM-Assisted Unsupervised Domain Adaptive Network for Semantic Segmentation of Remote Sensing Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, Institute of Electrical and Electronics Engineers Inc., v. 61, 2023. ISSN 15580644. Cited on page 24.
- Zhu, X. et al. Road extraction from remote sensing imagery with spatial attention based on swin transformer. *Remote Sensing*, v. 16, n. 7, 2024. ISSN 2072-4292. Available on: <<https://www.mdpi.com/2072-4292/16/7/1183>>. Cited on page 23.

Curricular Data

IDENTIFICATION	
Date of Birth	September 19, 2000
Nationality	Brazilian
Name in Bibliographic Citations	COLLEGIO, G. R.; COLLEGIO, GUSTAVO ROTA
Lattes Curriculum	< http://lattes.cnpq.br/3180803944754871 >
ORCID	< https://orcid.org/0000-0002-8633-8293 >
ACADEMIC FORMATION	
2018-2023	Bachelor of Science: Cartographic and Surveying Engineering Faculty of Science and Technology - FCT, São Paulo State University - UNESP
2023-2025	Master of Science in Cartographic Sciences Graduate Program in Cartographic Sciences - PPGCC Faculty of Science and Technology - FCT, São Paulo State University - UNESP