CLEI Electronic Journal, Volume 28, Issue 6, paper 3, October 2025

Deep learning and data integration for detecting trees entangled with utility lines

Artur André A. M. Oliveira
Universidade de São Paulo, Instituto de Matemática e Estatística
São Paulo, Brasil, 05508-090,
arturao@ime.usp.br

Advisor: **Roberto Hirata Jr.**
Universidade de São Paulo, Instituto de Matemática e Estatística
São Paulo, Brasil, 05508-090,
hirata@ime.usp.br

Abstract

Urban image classification is a challenging task due to factors such as occlusions, diverse appearances, and ambiguous visual contexts. This paper addresses these challenges by consolidating and discussing methodologies focused on tree-powerline interactions, a critical problem for urban safety and infrastructure planning. Using state-of-the-art Deep Learning Networks (DLNs), prior studies evaluated performance on a curated dataset of 11,000 labeled street-level images collected and annotated with the INvestigate and Analyze a CITY (INACITY) platform and the Street-Level Image Labeler (SLIL) tool. These evaluations revealed significant limitations in baseline models, with a maximum accuracy of 74.6% on this dataset. To address these limitations, earlier work introduced a “Challenging” class and applied semi-supervised learning techniques, including the Noisy Student protocol and Focal Loss, achieving recall rates of 83.7% for positive cases and 78.8% for negative cases. Dimensionality reduction using the Self-Supervised Neural Projection (SSNP) method was explored for clustering and visualization tasks, while the Supervised Decision Boundary Maps (SDBM) technique improved classifier interpretability by addressing scalability and clarity in decision boundary visualizations. This paper consolidates these contributions, presenting an integrated narrative that highlights tools, datasets, and methodologies for advancing urban image classification. These efforts provide scalable, interpretable, and accurate solutions for urban safety and infrastructure management challenges.

Keywords: Urban images. Computer Vision. Deep Learning. Instance Hardness.

1 Introduction

Urban image classification presents significant challenges due to the complexity of street-level imagery, where diverse elements such as buildings, vehicles, trees, and powerlines often overlap or occlude each other. Additional factors, such as varying environmental conditions and camera perspectives, exacerbate these difficulties, making automated analysis a demanding task. These challenges are particularly evident in detecting tree-powerline interactions, a critical issue for urban safety and infrastructure management. Mismanagement of these interactions can result in hazards such as electrical outages, fires, or accidents, necessitating accurate and efficient detection methods.

This paper synthesizes contributions from previous works, integrating advancements in dataset creation, learning methodologies, dimensionality reduction, and classifier interpretability. Specifically, we evaluate the performance of state-of-the-art Deep Learning Networks (DLNs) on a curated dataset of 11,000 labeled images, focusing on tree-powerline scenarios [1, 2]. To support this, we developed tools like the INvestigate and Analyze a CITY (INACITY) platform [3, 4] and the Street-Level Image Labeler (SLIL) [5], which streamline image collection and annotation processes, enabling the creation of high-quality datasets for urban analysis.

Baseline evaluations revealed that existing DLNs struggle with ambiguities and complexities in urban imagery, achieving a maximum accuracy of 74.6% [1]. To address these limitations, we introduced a "Challenging" class and applied advanced learning techniques, including the Noisy Student protocol and Focal Loss, achieving significant performance improvements [2]. Furthermore, we explored dimensionality reduction with the Self-Supervised Network Projection (SSNP) method for enhanced data visualization and clustering [6] and proposed the Supervised Decision Boundary Maps (SDBM) [7] technique to improve classifier interpretability.

By integrating these innovations and building on prior work, this study provides a robust framework for addressing the complexities of urban image classification, offering scalable and accurate methods with practical applications in urban safety and planning.

2 Review

This section examines existing research on the collection, classification, and analysis of urban street-level images, with an emphasis on applications related to greenery, trees, and powerlines. It explores methodologies and tools developed for these tasks, highlights their achievements, and identifies persistent challenges, such as handling complex urban scenes, interpreting ambiguous visual contexts, and addressing hard-to-classify instances.

Collecting and Analyzing Urban Images Urban studies [8, 9] have been facilitated since 2009 with Google Street View [10], which offers a collection of street-level urban images from locations across every continent. Additionally, Geo-located Imagery Databases (GIDs) such as KartaView [11] and Mappilary Vistas [12] provide similar services by turning urban image collection into a distributed, crowd-sourced effort by citizens. These initiatives, aligned with the principles of citizen science [13, 14], are both cost-effective and accessible, making the process of gathering data for urban studies more scalable.

While these GIDs enable large-scale urban studies, they often lack domain-specific annotations critical for detecting and analyzing specific elements, such as trees interacting with powerlines. Moreover, these GIDs do not fully address challenges inherent to street-level imagery, including occlusions, lighting conditions, and diverse camera perspectives.

Analysis of Urban Images The application of Deep Learning [15, 16] to urban imagery has driven advancements in Computer Vision tasks, for instance, object detection and scene classification. Wegner et al. [17] proposed a Convolutional Neural Network (CNN) [18] to detect and classify urban objects using both aerial and street-level imagery from Google Street View, showing that the fusion of these modalities yields better results than using either alone. Urban elements such as greenery [19], tree cover [20], shade provision [21], and even concepts like graffiti [22] can be quantified through a combination of Computer Vision heuristics and deep learning networks (DLNs) [19, 20, 21, 22].

Among the many elements captured in urban imagery, trees present unique challenges due to their different species, colors, shapes, etc. In the problem of interest here, trees pose even hard challenges because their interactions with power lines, which can cause electrical hazards, wildfires, and blackouts [23, 24, 25, 26, 27]. Managing vegetation to prevent such issues is crucial but must be done carefully to avoid negative consequences like spreading diseases or reducing shade, which can draw public criticism [28]. Monitoring these interactions often requires specialized indicators and substantial resources, which limits scalability [29, 30, 31, 32].

Using street-level imagery has significant advantages for analyzing tree-powerline interactions, particularly in urban settings, compared to aerial views for they are usually obstructed by dense canopies. Studies using Google Street View have demonstrated its potential for assessing urban vegetation [20], cataloging tree species [33], and estimating metrics like the Green View Index [20]. However, variability in tree appearances, coupled with ambiguities in visual contexts (e.g., distinguishing whether powerlines intersect or pass behind tree canopies), continues to challenge classification methods.

Gaps in Existing Approaches Despite advancements in urban image analysis, significant challenges remain. Existing tools and datasets lack the granularity and domain-specific annotations required for detecting complex interactions, such as trees entangled with powerlines. Most approaches struggle to address the variability and ambiguity inherent in urban scenes, particularly for hard-to-classify instances. Additionally, few studies propose frameworks to handle challenging images, such as defining a separate class for these cases or leveraging semi-supervised learning to enhance model performance. These gaps underscore the need for a robust methodology that can improve accuracy and interpretability in complex urban scenarios.

3 Methods and tools

In this section, we describe the tools and methods employed for urban image classification, with a focus on addressing tree-powerline interactions. Our contributions include the creation of datasets, enhancements to annotation and data collection tools, and techniques for improving model performance and interpretability. These efforts aim to tackle the unique challenges of urban imagery, including occlusions, varying perspectives, class imbalance, and ambiguous visual contexts.

3.1 Tools and methods for Urban Image Collection and Labeling

The INvestigate and Analyze a CITY (INACITY) platform [3] is a software tool for integrating urban data and analyzing together with street-level images. Extended during this research, the platform now supports a graph-oriented database to improve the performance and coverage of image collections from Google Street View. This enhancement enables more comprehensive analysis and facilitates the extraction of geospatial and image data from predefined regions of interest, contributing to a scalable, efficient workflow for urban studies. It is beyond the scope of this paper to enter into the details about the architecture of the platform. The details can be found in the original paper [3].

The Street-Level Image Labeler (SLIL) [5] was developed to facilitate the annotation process of urban images. Unlike generic tools, SLIL is tailored for the specific challenges of street-level imagery. It offers features such as direct visualization of image geolocation and simplified annotation interfaces, reducing the effort required for labeling large datasets. SLIL also supports other image domains, making it a versatile tool for various machine-learning tasks.

The curated datasets have overall 50,000 images collected with the INACITY platform, where we labeled 11,000 of them with the SLIL tool, focusing on tree-powerline interactions. These are from São Paulo and Porto Alegre cities in Brazil, and encompass diverse urban environments. The dataset balances scenarios with and without tree-wire interactions, and diverse urban infrastructure backgrounds [1, 2].

To address the ambiguity and complexity of certain images (e.g., distinguishing whether powerlines intersect or pass behind tree canopies), in the paper "Locating Urban Trees near Electric Wires" [2], we introduced a new "Challenging" class label. This class was incorporated into a semi-supervised learning setup using the Noisy Student protocol [34] and Focal Loss [35], which effectively handled the class imbalance and improved recall rates for both positive and negative cases. Together, these methods enhanced model performance on challenging data while maintaining generalization on simpler cases.

3.2 Advances in Dimensionality Reduction

Dimensionality reduction (DR) plays a critical role in understanding and addressing classification challenges in high-dimensional datasets. By projecting data into lower-dimensional spaces, DR techniques facilitate clustering, visualization, and exploration, offering valuable insights into underlying patterns and structures. For classification tasks, DR can improve model interpretability by highlighting separability between classes and uncovering relationships that might otherwise remain obscured in high-dimensional spaces.

The Self-Supervised Neural Projection (SSNP) [36] stands out among DR methods for its unique combination of desirable properties: scalability, stability, support for out-of-sample data, inverse mapping capabilities, and clustering effectiveness. Recognizing its potential for handling complex datasets, we conducted a detailed analysis to optimize SSNP's application and provide practical guidelines for its effective use [37].

Our study systematically examined the influence of various hyperparameters, including:

1. **Clustering Methods:** The impact of clustering algorithms such as K-means [38], DBSCAN [39], and Agglomerative Clustering [40] on pseudo-label generation.

2. **Neural Network Configurations:** Exploration of activation functions, regularization techniques, and weight initialization methods to optimize network performance.

3. **Projection Quality Metrics:** Metrics like trustworthiness, continuity, and neighborhood hit were used to evaluate and compare the output of SSNP with traditional DR methods like t-SNE, UMAP, and autoencoders.

Our analysis demonstrates that SSNP have competitive projection quality, particularly in separating clusters visually and preserving data structures in reduced dimensions. These findings advance the practical use of SSNP for tasks such as clustering, data exploration, and visualization in high-dimensional data spaces.

Supervised Decision Boundary Maps (SDBM) [41] improves the interpretability of classifiers. SDBM addresses the limitations of Decision Boundary Maps (DBM) [42], including scalability and multi-class complexity. Key features include:

1. **GPU-accelerated supervised projections:** Enables fast and fine-grained boundary visualization.
2. **Generic applicability:** Works with any single-output classifier.
3. **Improved interpretability:** Produces smoother, artifact-free maps.

Our follow-up study [6] demonstrates the stability of SDBM under dataset perturbations, such as noise or variations in training data. This robustness ensures the reliability of SDBM for analyzing real-world classifiers.

4 Experimental Results

In this section, we present the experimental setup and results for evaluating our proposed methods and tools. These experiments validate the effectiveness of our approaches in improving the classification of urban street-level images, particularly those involving challenging scenarios such as tree-powerline interactions.

4.1 Baseline DLN Performance on Urban Images

We evaluated the performance of state-of-the-art Deep Learning Networks (DLNs), including VGG16 [43], MobileNets [44, 45], ResNets [46], DenseNets [47], and EfficientNets [48], on a dataset of 11,000 labeled images collected via the INACITY platform and annotated with the SLIL tool. The dataset includes images from São Paulo and Porto Alegre, Brazil, with labels identifying tree-powerline intersections.

The highest accuracy achieved by plain DLNs was 74.6%, highlighting the difficulty of classifying urban scenes characterized by occlusions, varying environmental conditions, and ambiguous visual contexts. Figure 1 illustrates a GradCAM++ visualization [49], showing relevant regions for the DLN classification of an image of a tree canopy intersecting powerlines. These experiments emphasize the need for advanced techniques to handle challenging urban images. We refer the interested reader to our paper "Detecting tree and wire entanglements with deep learning" [1] for more details.

Figure 1: (Left) Image of a tree with branches and leaves in contact with overhead powerlines. (Right) The corresponding Grad-CAM++ image, showing (in red) that the whole extension of the wires passing through the tree canopy has been identified and relevant for classifying this image (as a true positive) by the DLN used.

4.2 Addressing Challenging Images with Semi-Supervised Learning

To improve performance on challenging images, in our paper "Locating Urban Trees near Electric Wires using Google Street View Photos: A New Dataset and A Semi-Supervised Learning Approach in the Wild" [2], we introduced a new "Challenging" class and applied the Noisy Student training protocol [34] with the Focal Loss cost function [35]. Using a semi-supervised learning setup, we combined 11,000 labeled images with 40,000 pseudo-labeled images generated during the training process. This approach significantly enhanced recall rates. Positive class recall improved from 66.5% to 83.7%. Negative class recall improved from 63.7% to 78.8%. Our findings revealed that excluding challenging images from training datasets improved generalization performance for all tested DLN architectures except VGG16, which achieved a significantly lower accuracy of 57.28% under this setup, compared to 73.63%. Before removing these images, DenseNet121 achieved an accuracy of 84.49%, which increased to 88.54% after their exclusion, highlighting the benefits of this approach. For a detailed analysis of these results, please refer to Appendix B of the thesis [50].

4.3 Dimensionality Reduction with SSNP

We evaluated the Self-Supervised Network Projection (SSNP) method [37] for dimensionality reduction (DR) across synthetic and real-world datasets to analyze its scalability, stability, and projection quality. These experiments systematically explored the effects of hyperparameter configurations and pseudo-labeling strategies, focusing on the quality of projections and the preservation of data structures in reduced spaces.

The evaluation involved datasets with varying complexity, using clustering algorithms such as K-means, DBSCAN, and Agglomerative Clustering to generate pseudo-labels. Metrics like trustworthiness, continuity, and neighborhood preservation were employed to assess how well SSNP retained meaningful relationships in the reduced data. Results demonstrated that SSNP maintained scalability with linear complexity, allowing efficient processing of large datasets while producing high-quality projections. The method exhibited stability across different clustering algorithms and hyperparameter configurations, showing minimal variations in projection quality under changes such as initialization or the number of clusters. Additionally, SSNP supported out-of-sample projections effectively, preserving neighborhood relationships when unseen data were incorporated without requiring retraining.

Our analysis suggests that the pseudo-label quality is very important for training SSNP. Among the clustering methods tested, K-means provided the most consistent and interpretable projections. Fine-tuning hyperparameters, including the number of clusters and network depth, further enhanced projection quality, as reflected in trustworthiness and continuity metrics. Compared to methods like PCA and autoencoders, SSNP consistently delivered better neighborhood preservation while maintaining computational efficiency.

4.4 Improving Classifier Interpretability with SDBM

The Supervised Decision Boundary Maps (SDBM) method [41] was developed to address the limitations of traditional Decision Boundary Maps (DBM), such as poor scalability and difficulty interpreting multi-class problems.

We conducted experiments using SDBM on diverse datasets, including MNIST [51], Fashion-MNIST [52], and UCI HAR [53], to evaluate its effectiveness. The results demonstrate that SDBM offers significant advantages:

- **Improved interpretability**: SDBM produces decision boundary maps that are smoother and easier to interpret than those generated by traditional DBM methods. This clarity is particularly important in understanding classification tasks with overlapping or complex decision zones.
- **Enhanced computational efficiency**: Leveraging GPU-accelerated supervised projections, SDBM can generate fine-grained maps in seconds, addressing the scalability issues of DBM.
- **General applicability**: SDBM is compatible with a variety of classifiers, making it a versatile tool for analyzing decision boundaries across different machine learning models.

Further analysis in "Stability Analysis of Supervised Decision Boundary Maps" [6] explored the robustness of SDBM under real-world conditions, including dataset perturbations such as noise and variations in training data. The results showed that SDBM maintains stable and consistent decision boundaries despite these changes, a property that is essential for practical applications in dynamic or noisy environments. This stability, combined with its interpretability and efficiency, makes SDBM particularly suited for analyzing machine learning classifiers in urban imagery and other complex scenarios.

The ability of SDBM to visualize decision boundaries not only aids in understanding the internal workings of classifiers but also provides insights into model performance and behavior, highlighting areas for improvement or refinement. Its utility in real-world applications underscores its value as a reliable tool for advancing classifier interpretability.

5 Conclusion

This study presents advances in urban image classification by addressing the complexities of urban scenes, such as occlusions, diverse perspectives, and ambiguous contexts. We enhanced the INACITY platform by integrating to it a graph-oriented database for scalable image collection, and developed the SLIL tool, which facilitates annotation processes and enables the creation of high-quality datasets. There tools make it easier to create a comprehensive dataset of 11,000 labeled images, and other 40,000 unlabeled ones, focusing on tree-powerline interactions.

Our evaluation of state-of-the-art Deep Learning Networks (DLNs) highlighted their limitations, with baseline models achieving a maximum accuracy of 74.6%. To overcome these challenges, we introduced a "Challenging" class and applied semi-supervised learning techniques, such as the Noisy Student protocol and Focal Loss, significantly improving recall rates for positive and negative classes. These methods proved effective in handling ambiguous and complex cases, demonstrating their robustness in real-world scenarios.

Additionally, we explored dimensionality reduction with the Self-Supervised Network Projection (SSNP) method, optimizing its application through hyperparameter tuning and clustering strategies. This improved the quality of projections while maintaining scalability and stability. To further enhance model interpretability, we proposed the Supervised Decision Boundary Maps (SDBM) technique, which addresses scalability and clarity issues in traditional methods. SDBM provides stable and efficient visualizations of decision boundaries, enabling deeper insights into classifier behavior.

By integrating these contributions, this work provides a framework for advancing urban image classification. The tools and methodologies developed are versatile and adaptable, offering solutions for a wide range of challenges in urban planning, infrastructure management, and environmental monitoring. As cities continue to grow, the approaches outlined in this study will support more accurate, scalable, and interpretable analyses of urban environments.

Acknowledgment

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, São Paulo Research Foundation (FAPESP) 22/15304-4, 18/10767-0 and 15/22308-2 and MCTI (law 8.248, PPI-Softex - TIC 13 - 01245.010222/2022-44).

References

[1] A. A. Oliveira, M. S. Buckeridge, and R. Hirata Jr, "Detecting tree and wire entanglements with deep learning," *Trees*, pp. 1–13, 2022.

[2] A. A. Oliveira, Z. Wang, and R. Hirata, "Locating urban trees near electric wires using google street view photos: A new dataset and a semi-supervised learning approach in the wild," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2022, pp. 4285–4293.

[3] A. A. Oliveira and R. Hirata, "Inacity - investigate and analyze a city," *SoftwareX*, vol. 15, p. 100777, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352711021000911

[4] A. A. Oliveira and R. Hirata Júnior, "Investigate and analyse a city - inacity," Master's thesis, Universidade de São Paulo, 2018. [Online]. Available: https://teses.usp.br/teses/disponiveis/45/45134/tde-04052018-170132/en.php

[5] A. A. Oliveira and R. Hirata Jr., "Street Level Image Labeler (SLIL)," Aug. 2020. [Online]. Available: https://github.com/arturandre/SLIL

[6] A. A. Oliveira, M. Espadoto, R. Hirata Jr, and A. C. Telea, "Stability analysis of supervised decision boundary maps," *SN Computer Science*, vol. 4, no. 3, p. 226, 2023.

[7] A. A. Oliveira, M. Espadoto, R. Hirata Jr., and A. C. Telea, "Supervised Decision Boundary Maps (SDBM)," 2022. [Online]. Available: https://github.com/mespadoto/sdbm

[8] A. G. Rundle, M. D. Bader, C. A. Richards, K. M. Neckerman, and J. O. Teitler, "Using google street view to audit neighborhood environments," *American journal of preventive medicine*, vol. 40, no. 1, pp. 94–100, 2011.

[9] J. S. Wilson, C. M. Kelly, M. Schootman, E. A. Baker, A. Banerjee, M. Clennin, and D. K. Miller, "Assessing the built environment using omnidirectional imagery," *American journal of preventive medicine*, vol. 42, no. 2, pp. 193–199, 2012.

[10] D. Anguelov, C. Dulong, D. Filip, C. Frueh, S. Lafon, R. Lyon, A. Ogale, L. Vincent, and J. Weaver, “Google street view: Capturing the world at street level,” *Computer*, vol. 43, no. 6, pp. 32–38, 2010.

[11] ®Grab and O. Contributors, “Hello, kartaview! - openstreetmap @ grab,” https://blog.improveosm.org/en/hello-kartaview/, 2020, access Ago 22nd, 2021.

[12] G. Neuhold, T. Ollmann, S. Rota Bulò, and P. Kontschieder, “The mapillary vistas dataset for semantic understanding of street scenes,” in *International Conference on Computer Vision (ICCV)*, 2017. [Online]. Available: https://www.mapillary.com/dataset/vistas

[13] M. V. e. a. Eitzel, “Citizen science terminology matters: Exploring key terms. citizen science: Theory and practice,” vol. 2(1):1, pp. 1–20, 2017.

[14] F. Biljecki and K. Ito, “Street view imagery in urban analytics and gis: A review,” *Landscape and Urban Planning*, vol. 215, p. 104217, 2021.

[15] F. Rosenblatt, “The perceptron: a probabilistic model for information storage and organization in the brain.” *Psychological review*, vol. 65, no. 6, p. 386, 1958.

[16] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, http://www.deeplearningbook.org.

[17] J. D. Wegner, S. Branson, D. Hall, K. Schindler, and P. Perona, “Cataloging public objects using aerial and street-level images-urban trees,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 6014–6023.

[18] Y. LeCun, B. Boser, J. Denker, D. Henderson, R. Howard, W. Hubbard, and L. Jackel, “Handwritten digit recognition with a back-propagation network,” *Advances in neural information processing systems*, vol. 2, 1989.

[19] T. Boehmer, C. Hoehner, A. Deshpande, L. B. Ramirez, and R. C. Brownson, “Perceived and observed neighborhood indicators of obesity among urban adults,” *International journal of obesity*, vol. 31, no. 6, p. 968, 2007.

[20] B. Y. Cai, X. Li, I. Seiferling, and C. Ratti, “Treepedia 2.0: applying deep learning for large-scale quantification of urban tree cover,” in *2018 IEEE International Congress on Big Data (BigData Congress)*. IEEE, 2018, pp. 49–56.

[21] X. Li and C. Ratti, “Mapping the spatial distribution of shade provision of street trees in boston using google street view panoramas,” *Urban Forestry & Urban Greening*, vol. 31, pp. 109–119, 2018.

[22] E. K. Tokuda, R. M. Cesar, and C. T. Silva, “Quantifying the presence of graffiti in urban environments,” in *2019 IEEE International Conference on Big Data and Smart Computing (BigComp)*, 2019, pp. 1–4.

[23] J. Ma, J. C. Cheng, F. Jiang, V. J. Gan, M. Wang, and C. Zhai, “Real-time detection of wildfire risk caused by powerline vegetation faults using advanced machine learning techniques,” *Advanced Engineering Informatics*, vol. 44, p. 101070, 2020.

[24] D. J. Clarke and J. G. White, “Towards ecological management of australian powerline corridor vegetation,” *Landscape and urban planning*, vol. 86, no. 3-4, pp. 257–266, 2008.

[25] J. Ahmad, A. S. Malik, L. Xia, and N. Ashikin, “Vegetation encroachment monitoring for transmission lines right-of-ways: A survey,” *Electric Power Systems Research*, vol. 95, pp. 339–352, 2013.

[26] L. Pucell, “Trees and electric lines,” https://extension.purdue.edu/extmedia/FNR/FNR-512-W.pdf, 2015, access Jan 22nd, 2021.

[27] P. Pakonen, “Characteristics of partial discharges caused by trees in contact with covered conductor lines,” *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 15, no. 6, pp. 1626–1633, 2008.

[28] J. Fakes, “Practical issues in line clearance and street trees,” *TREENET*, p. 39, 2000.

[29] W. A. Kenney, P. J. van Wassenaer, A. L. Satel *et al.*, “Criteria and indicators for strategic urban forest planning and management,” *Arboriculture & Urban Forestry*, vol. 37, no. 3, pp. 108–117, 2011.

[30] C. Ordóñez and P. N. Duinker, “An analysis of urban forest management plans in Canada: Implications for urban forest management,” *Landscape and Urban Planning*, vol. 116, pp. 36–47, 2013.

[31] W. Cheng and Z. Song, "Power pole detection based on graph cut," in *2008 Congress on Image and Signal Processing*, vol. 3. IEEE, 2008, pp. 720–724.

[32] Y. Jwa, G. Sohn, and H. Kim, "Automatic 3d powerline reconstruction using airborne lidar data," *Int. Arch. Photogramm. Remote Sens*, vol. 38, no. Part 3, p. W8, 2009.

[33] D. Liu, Y. Jiang, R. Wang, and Y. Lu, "Establishing a citywide street tree inventory with street view images and computer vision techniques," *Computers, Environment and Urban Systems*, vol. 100, p. 101924, 2023.

[34] Q. Xie, M.-T. Luong, E. Hovy, and Q. V. Le, "Self-training with noisy student improves imagenet classification," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 10 687–10 698.

[35] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal loss for dense object detection," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2980–2988.

[36] M. Espadoto, N. S. Hirata, and A. C. Telea, "Self-supervised dimensionality reduction with neural networks and pseudo-labeling." in *Proc. IVAPP*, 2021, pp. 27–37, SCITEPRESS.

[37] A. A. Oliveira, M. Espadoto, R. Hirata Jr, N. S. Hirata, and A. C. Telea, "Improving self-supervised dimensionality reduction: Exploring hyperparameters and pseudo-labeling strategies," in *Computer Vision, Imaging and Computer Graphics Theory and Applications: 16th International Joint Conference, VISIGRAPP 2021, Virtual Event, February 8–10, 2021, Revised Selected Papers.* Springer, 2023, pp. 135–161.

[38] S. Lloyd, "Least squares quantization in pcm," *IEEE transactions on information theory*, vol. 28, no. 2, pp. 129–137, 1982.

[39] M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise." in *Proc. KDD*, vol. 96, no. 34, 1996, pp. 226–231.

[40] L. Kaufman and P. J. Rousseeuw, *Finding groups in data: an introduction to cluster analysis.* John Wiley & Sons, 2009.

[41] A. A. Oliveira, M. Espadoto, R. Hirata Jr, N. S. Hirata, and A. C. Telea, "Sdbm: Supervised decision boundary maps for machine learning classifiers." in *VISIGRAPP (3: IVAPP)*, 2022, pp. 77–87.

[42] F. C. Rodrigues, M. Espadoto, R. Hirata Jr, and A. C. Telea, "Constructing and visualizing high-quality classifier decision boundary maps," *Information*, vol. 10, no. 9, p. 280, 2019.

[43] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[44] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," 2017.

[45] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan *et al.*, "Searching for mobilenetv3," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 1314–1324.

[46] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[47] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[48] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International Conference on Machine Learning.* PMLR, 2019, pp. 6105–6114.

[49] A. Chattopadhyay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, "Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks," *CoRR*, vol. abs/1710.11063, 2017. [Online]. Available: http://arxiv.org/abs/1710.11063

[50] A. A. Oliveira, “Overcoming challenging urban images - deep learning and data integration methods for detecting trees entangled with power lines,” PhD Thesis in Computer Science, University of São Paulo, São Paulo, 2023, advisor: Roberto Hirata Junior.

[51] L. Deng, “The mnist database of handwritten digit images for machine learning research,” *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 141–142, 2012.

[52] H. Xiao, K. Rasul, and R. Vollgraf, “Fashion-MNIST: A novel image dataset for benchmarking machine learning algorithms,” 2017, arXiv:1708.07747.

[53] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, “Human activity recognition on smartphones using a multiclass hardware-friendly support vector machine,” in *Proc. Intl. Workshop on Ambient Assisted Living*, 2012, pp. 216–223, springer.