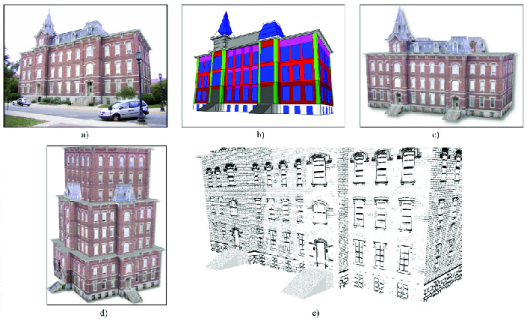
Ph.D. thesis – deep generative models for direct and inverse procedural generation

Application deadline: March 15th 2025
Online application

Location	Champs-sur-Marne (near Paris, France)
Salary	≈1800€ net/month
Contract	Fixed-term (36 months)
Remote work	Partial
Start date	Before May 1st 2025
Contacts	Nicolas Audebert (nicolas.audebert@ign.fr)
	Julien Perret (julien.perret@ign.fr)

Keywords: deep learning, procedural generation, geospatial data, generative models.

(a) City layout generation with GANs (source).



(b) Inverse procedural generation of buildings [20].

1 Research project

Training large-scale deep learning models for remote sensing image analysis requires vast amounts of labeled imagery. While Earth observation data is available in large quantities thanks to imaging programs such as the European Sentinel-2 constellation or the French SPOT and BDORTHO initiatives, this mass of data is unlabeled. Indeed, few remote sensing images are actually labeled with semantic information that can be used to train supervised deep neural networks.

To alleviate this obstacles, machine learning research has turned towards the generation of *synthetic* datasets. The democratization of "generative" artificial intelligence has made it possible to produce large-scale labeled datasets by generating diverse images in known configurations. Procedural generation is another technique well-known in the video game industry, that has been use to quickly and efficiently synthesize large-scale 3D virtual worlds.

2 Scientific goals

The main goal of this Ph.D. thesis is to combine the strenghts of **deep learning** with **procedural generation** based on grammars, with applications to geospatial modeling.

Procedural generation consists in algorithms for content creation, especially tailored to video games [7]. These techniques produce coherent virtual worlds [1] that can be used to model or simulate the real world. Procedural generation has regained traction in recent years as it can be used to synthesize large amounts of labeled synthetic data, on which deep neural networks can then be trained [11, 10]. Previously, procedural generation has mostly leveraged four broad types of approaches: exploration [17], constraint satisfaction [18], grammars [15], and statistical learning [5, 14].

Grammars (or assimilated, such as the L-system [21]) are particularly interesting. Indeed, they use a formal language that defines which instances are acceptable objects that can be generated. A grammar allows users to include specialized knowledge and is *interpretable*. However, manual definition of grammars require expert knowledge and often entails an iterative "try-and-retry" workflow. In comparison, learning-based procedural generation trains models to learn how to generate objects based on a existing collection. Yet, these models are not constrained and might generate objects that are not acceptable, e.g. houses without any doors. In addition, recent approaches based on deep neural networks need large-scale datasets to excel, which is often not an option when working with specialized data. Geographic entities in urban areas, typically roads and buildings, have to follow strict geometrical priors that are known to be difficult to satisfy [6].

This Ph.D. therefore looks into **hybrid approaches at the intersection of symbolic procedural generation**, especially grammars, **and deep learning**. Hybrid approaches could be able to learn from fewer examples, while being better at enforcing the expert constraints defined in a user-based grammar [9]. We have two main goals:

1. Designing generative neural architectures in which the outputs are constrained by a user-defined grammar. Doing so guarantees that the model can only generate acceptable objects to the user [8]. More specifically, we will investigate:
   - how to constrain segmentation maps from a supervised model so that they satisfy a grammar on the spatial relationships between objects,
   - or generative models that can only generate instances that are the results of successive rule applications from a given grammar, e.g. for buildings.
2. Develop models that can infer part of or an entire grammar based on a training dataset (inverse procedural generation [12] and grammar inferencea). In particular, we will look for:
   - methods that can automatically learn the terminal symbols of the formal language, for example using prototype learning [4, 13],
   - then for methods that can infer production rules of a grammar that can produce the training instances [2].

These generation algorithms will be applied to 2D and 3D geospatial data for city generation [16, 19, 3], such as cadaster maps (building and parcel footprints), buildings 3D models (either manually produced or automatically extracted from Lidar point clouds) and land cover/land use maps.

3 Applicant profile

The ideal applicant has a master level education (M.Sc. or M.Eng.) specialized in one of the following fields: data science, video games, geoinformation. He or she has previous experience in programming, especially with the Python language ecosystem. Previous knowledge of project management tools, such as Git, is a plus. English proficiency is mandatory. French is not required, although it can help in everyday life. Although not required, a first experience with procedural generation, generative models or geospatial data is welcome.

4 Workplace

The National institute for geographic and forest information (IGN) is a French governmental agency under the ministry of Ecology and Forests. Its main role is to produce and disseminate reference data and representations (paper and online maps, geovisualizations) relevant to the understanding of the French national territory, its forests, and their evolution. Thanks to its engineering school, ENSG-Géomatique, and its pluridisciplinary research laboratories, the institute fosters a strong and high-level innovation culture in several fields (geodesy, forest management, photogrammetry, artificial intelligence, spatial analysis, visualization…).

The LASTIC[1] is a mixed research unit under the umbrella of IGN-ENSG and University Gustave Eiffel. The laboratory works fundamental and applied researchers in geographical information sciences and technologies. The STRUDEL team focuses on spatio-temporal structures to analyze territories. In particular, its main researches are extracting and structuring knowledge about territories, its characteristics, its evolution, and how to reuse this knowledge for simulation. The team is involved in multiple research projects leveraging synthetic data for simulation, digital twins and unsupervised learning.

Management: This is a fixed-term 36 months contract, full-time. It is tied to a student registration to the MSTIC graduate school. Salary is approximately 1800€/month.

Location: The lab is located in the ENSG engineering school at Champs-sur-Marne (77), near Paris by train (RER A).

Advantages:

- Flexible remote work after a starting period
- Sports facilities available on-site
- Sports and cultural associations available in the institute
- Access to the campus cafeteria
- Travel card 75% paid up by the employer, biking subsidies

Process: resume, on-site or remote interview including a technical test. Application online:
https://www.ign.fr/nous-rejoindre/offres-emploi/doctorant-e-modeles-generatifs-profonds-pour-la-gen

[1] LAboratory in Sciences and Technologies for Geographical Information https://www.umr-lastig.fr

References

[1] Abdul Latif et al. "A Critical Evaluation of Procedural Content Generation Approaches for Digital Twins". In: *Journal of Sensors* 2022 (July 29, 2022), e5629645. URL: https://www.hindawi.com/journals/js/2022/5629645/.

[2] Peter Belcak, David Hofer, and Roger Wattenhofer. "A Neural Model for Regular Grammar Induction". In: *2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA)*. 2022 21st IEEE International Conference on Machine Learning and Applications (ICMLA). Dec. 2022, pp. 401–406. URL: https://ieeexplore.ieee.org/abstract/document/10069747.

[3] Tom Kelly. "CityEngine: An Introduction to Rule-Based Modeling". In: *Urban Informatics*. Ed. by Wenzhong Shi et al. Singapore: Springer, 2021, pp. 637–662. URL: https://doi.org/10.1007/978-981-15-8983-6_35.

[4] Oscar Li et al. "Deep Learning for Case-Based Reasoning Through Prototypes: A Neural Network That Explains Its Predictions". In: *Proceedings of the AAAI Conference on Artificial Intelligence* 32.1 (1 Apr. 29, 2018). URL: https://ojs.aaai.org/index.php/AAAI/article/view/11771.

[5] Jialin Liu et al. "Deep Learning for Procedural Content Generation". In: *Neural Computing and Applications* 33.1 (Jan. 1, 2021), pp. 19–37. URL: https://doi.org/10.1007/s00521-020-05383-8.

[6] Bruno Vallet and Franck Taillandier. "Fitting Constrained 3D Models in Multiple Aerial Images." In: *BMVC*. 2005. URL: https://www.robots.ox.ac.uk/~phst/BMVC2005/papers/176/Vallet_Taillandier_BMVC2005.pdf.

[7] Georgios N. Yannakakis and Julian Togelius. "Generating Content". In: *Artificial Intelligence and Games*. Ed. by Georgios N. Yannakakis and Julian Togelius. Cham: Springer International Publishing, 2018, pp. 151–202. URL: https://doi.org/10.1007/978-3-319-63519-4_4.

[8] Matt J. Kusner, Brooks Paige, and José Miguel Hernández-Lobato. "Grammar Variational Autoencoder". In: *Proceedings of the 34th International Conference on Machine Learning - Volume 70*. ICML'17. Sydney, NSW, Australia: JMLR.org, Aug. 6, 2017, pp. 1945–1954.

[9] Léo Géré, Nicolas Audebert, and Philippe Rigaux. "Improved Symbolic Drum Style Classification with Grammar-Based Hierarchical Representations". In: *International Society for Music Information Retrieval Conference 2024*. San Francisco, United States, Nov. 2024. URL: https://hal.science/hal-04660056.

[10] Alexander Raistrick et al. "Infinigen Indoors: Photorealistic Indoor Scenes Using Procedural Generation". In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024, pp. 21783–21794. URL: https://openaccess.thecvf.com/content/CVPR2024/html/Raistrick_Infinigen_Indoors_Photorealistic_Indoor_Scenes_using_Procedural_Generation_CVPR_2024_paper.html.

[11] Alexander Raistrick et al. "Infinite Photorealistic Worlds Using Procedural Generation". In: *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). June 2023, pp. 12630–12641. URL: https://ieeexplore.ieee.org/document/10204989.

[12] O. Šťava et al. "Inverse Procedural Modeling by Automatic Generation of L-systems". In: *Computer Graphics Forum* 29.2 (2010), pp. 665–674. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-8659.2009.01636.x.

[13] Junxian He, Taylor Berg-Kirkpatrick, and Graham Neubig. "Learning Sparse Prototypes for Text Generation". In: *Proceedings of the 34th International Conference on Neural Information Processing Systems*. NIPS '20. Red Hook, NY, USA: Curran Associates Inc., Dec. 6, 2020, pp. 14724–14735.

[14] Matthew Guzdial, Sam Snodgrass, and Adam J. Summerville. *Procedural Content Generation via Machine Learning: An Overview*. Synthesis Lectures on Games and Computational Intelligence. Cham: Springer International Publishing, 2022. URL: https://link.springer.com/10.1007/978-3-031-16719-5.

[15] Roland van der Linden, Ricardo Lopes, and Rafael Bidarra. "Procedural Generation of Dungeons". In: *IEEE Transactions on Computational Intelligence and AI in Games* 6.1 (Mar. 2014), pp. 78–89. URL: https://ieeexplore.ieee.org/abstract/document/6661386.

[16] Yoav I. H. Parish and Pascal Müller. "Procedural Modeling of Cities". In: *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH '01. New York, NY, USA: Association for Computing Machinery, Aug. 1, 2001, pp. 301–308. URL: https://doi.org/10.1145/383259.383292.

[17] Julian Togelius et al. "Search-Based Procedural Content Generation: A Taxonomy and Survey". In: *IEEE Transactions on Computational Intelligence and AI in Games* 3.3 (Sept. 2011), pp. 172–186.

[18] Ruben Smelik et al. "Semantic Constraints for Procedural Generation of Virtual Worlds". In: *Proceedings of the 2nd International Workshop on Procedural Content Generation in Games*. PCGames '11. New York, NY, USA: Association for Computing Machinery, June 28, 2011, pp. 1–4. URL: https://doi.org/10.1145/2000919.2000928.

[19] R. Cura, J. Perret, and N. Paparoditis. "STREETGEN: IN-BASE PROCEDURAL-BASED ROAD GENERATION". In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* II-3-W5 (Aug. 20, 2015), pp. 409–416. URL: https://isprs-annals.copernicus.org/articles/II-3-W5/409/2015/.

[20] Daniel G. Aliaga, Paul A. Rosen, and Daniel R. Bekins. "Style Grammars for Interactive Visualization of Architecture". In: *IEEE Transactions on Visualization and Computer Graphics* 13.4 (2007), pp. 786–797.

[21] Jean-Eudes Marvie, Julien Perret, and Kadi Bouatouch. "The FL-system: A Functional L-system for Procedural Geometric Modeling". In: *The Visual Computer* 21.5 (June 1, 2005), pp. 329–339. URL: https://doi.org/10.1007/s00371-005-0289-z.