# GUESs: Generative modeling of Unknown Environments and Spatial Abstraction for Robots

Extended Abstract

Francesco Riccio
DIAG, Sapienza University of Rome
Rome, Italy
riccio@diag.uniroma1.it

Roberto Capobianco
DIAG, Sapienza University of Rome
Rome, Italy
capobianco@diag.uniroma1.it

Daniele Nardi
DIAG, Sapienza University of Rome
Rome, Italy
nardi@diag.uniroma1.it

## ABSTRACT

Representing unknown and missing knowledge about the environment is fundamental to leverage robot behavior and improve its performance in completing a task. However, reconstructing spatial knowledge beyond the sensory horizon of the robot is an extremely challenging task. Existing approaches assume that the environment static and features repetitive patterns (e.g. rectangular rooms) or that it can be all generalized with pre-trained models. Our goal is to remove such assumptions and to introduce a novel methodology that allows the robot to represent unknown spatial knowledge in dynamic and unstructured environments. To this end, we exploit generative learning to (1) learn a distribution of spatial landmarks observed during the robot mission and to (2) generate missing information in real-time. The proposed approach aims at supporting planning and decision-making processes needed for robot behaviors. In this paper, we describe architecture modeling the proposed approach and a first validation on a mobile platform.

## KEYWORDS

Robot learning; Knowledge representation; Generative learning

**ACM Reference Format:**
Francesco Riccio, Roberto Capobianco, and Daniele Nardi. 2020. GUESs: Generative modeling of Unknown Environments and Spatial Abstraction for Robots . In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 3 pages.

## 1 INTRODUCTION

The representation of the environment and its structural characteristics, is fundamental to enable effective and autonomous robots behaviors [16, 20]. Several approaches assume that the map of the environment is given, or it can be generated before the robot deployment [4, 8]. However, as we want robots to fulfill complex tasks in different and varying environments, such an assumption does not always hold; consequently, robots have to explore their world while completing their tasks. Examples of such environments are search and rescue [12], door-to-door delivery [13], planetary exploration [17], visual inspection [1] and mining [15]; in these applications a map is not typically available to the robot before deployment. In fact, the robot has to perform SLAM and accomplish its mission simultaneously. Hence, the robot has to efficiently explore the environment and understand its structure, appearances and the key landmarks that characterize (and define) it. Such a skill is particularly well-suited for cognitive robots operating in unknown environments. We attack the problem of enabling an autonomous robot to explore an unknown environment, and represent portion of it not yet perceived through sensors. Several approaches consider the problem of estimating unknown parts of environment. These methods usually rely on frontier-based [10, 16] or gain-based [14, 21] techniques. Only few approaches attempt to explicitly reconstruct the portion of the environment not (yet) observable by the robot, either with pre-trained models [2, 18], spoken instructions [6], or structure prediction, based on geometric features [3]. These methods require a pre-trained model of the world that tells the robot how to classify the environment and the expected structure. However, such approaches are often inaccurate and hardly generalize to dynamic environment.

In this paper we introduce a novel approach that – in contrast to methods proposed in the literature – enables a robot to perform online map prediction and *guess* the structure of the environment based upon previous observations collected during operation. To this end, we introduce GUESs, a generative modeling approach of unknown environments and spatial abstraction. GUESs is designed to refine and improve map prediction during the robot operation, by relying upon a Variational AutoEncoder (VAE) [11] paired with a Generative Adversarial Network (GAN) [7]. The autoencoding is used to learn a latent representation of the structure of the environment, while the GAN is used to generate expected future observations by the robot sensors.

Our ultimate goal is to provide a robot with the ability to represent portions of the world beyond its sensory horizon, in order to support reasoning in partially known environments. This contribution aims at introducing our methodology, and at confirming the insights upon which we built GUESs. Hence, we describe our approach, and we show how it can be used to predict laser scans of a mobile robot. The key contribution is a novel generative approach to address partial observability and missing knowledge in robotic applications. GUESs does not assume prior knowledge nor a pre-trained model of the environment. We perform a validation of our approach on a mobile robot visiting a simulated environment.

## 2 GUESS

GUESs is a deep iterative algorithm based on generative learning [7] and variational autoencoding [11]. Its underlying architecture is completely agnostic with respect to the type of data and the application of deployment. Here, we focus on validating its conceptual

*Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.* 

building blocks and insights on spatial knowledge modeled through 2D laser scans. Hence, we formalize GUESs to tackle the problem of map prediction by exploiting laser sensor readings collected during the robot mission. We configure GUESs to accumulate lasers scans, while roaming the environment and to predict laser scans, expected to be perceived at a given point $t'$ in the future. The algorithm assumes no prior knowledge about the environment and the weights of both VAE and GAN are randomly initialized. Then, at each execution GUESs iteratively refines its estimations as soon as new observations are available. In fact, it continuously performs online aggregation [19] of new data samples to allow the robot to quickly adapt to unpredictable events and to generalize to new environments.

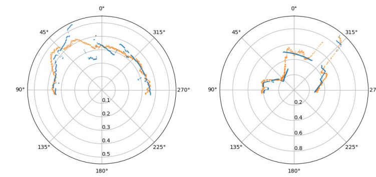
To infer missing knowledge not (yet) perceivable through robot sensors, we exploit generative adversarial learning. It has been demonstrated that GANs achieve remarkable results in learning data distributions to generate new samples and/or completing missing knowledge [5, 9]. However, since they need to be fed with large datasets and configured with complex networks with high dimensionality inputs and outputs, GANs cannot be easily deployed in dynamic and real-time applications. In robotic settings, for example, large datasets and computational costs are assumptions that cannot always be satisfied. Thus, in order to exploit the potential of GANs in robotics, we need to alleviate dimensionality constraints and network complexity – yet guaranteeing robust performance. To meet such a compelling requirement, we reduce the dimensionality of the generative network input by learning a lower-dimensionality latent representation of it. We exploit a autoencoder [11] to learn the latent – and more compact – representation of the input data. Specifically, a variational autoencoder is used to learn the distribution of input data over the latent space, which allows us to perform sampling directly in the latent space and to generate batches of inputs of lower dimensionality to be fed to the generative network.

In fact, to enable GUESs to evaluate sequences $T$ of time-correlated inputs $\boldsymbol{x}_{t:t+T}$ (e.g. 2D laser scans) and to predict a future observation $x'$, we substitute raw sensor data $\boldsymbol{x}_{t:t+T}$ with sequences of their latent representation $z_{t:t+T}$ that are iteratively learned through variational autoencoding. This allows for a significant reduction in the dimensionality of the input, yet preserving an informative representation used to predict missing knowledge.

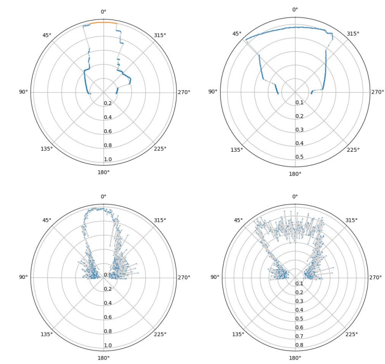
## 3 EXPERIMENTAL EVALUATION

The goal of the experimental evaluation conducted in this section is to validate the key insights of GUESs, and to demonstrate that our approach is feasible and practical in robotics. We validate GUESs in an indoor office environment where a robot is tasked to navigate while collecting data in real-time and refining its predictions iteratively.

We simulated a training session where the system is randomly initialized and ran for 600 iterations. Figure 1 shows the inference of the VAE and GAN over samples collected by the robot during its mission at the 600th iteration. In Figure 1(a), from left to right, samples represent a corner of a large room and the end of a corridor. The blue line represents the original laser scan $x$, while the orange line represents the VAE reconstructed scan $x'$. It is worth remarking



(a) VAE encodings



(b) GAN generations

**Figure 1: GUESs inference samples. 1(a) and 1(b) show the performance and generalization capabilities of the two networks after 600 iterations. In 1(a) the original scans (blue) and VAE reconstructed (orange) are reported, while 1(b) pairs of reference scans (top-row) and generated scans (bottom-row).**

that the latent encodings used in this experiment reduce the dimensionality of the inputted laser scans to L=16 and – as shown in the figure – yet maintain an informative representation. As for the VAE encodings, Figure 1(b) reports two laser scan samples. The top-row shows the original laser scan, while the bottom-row the generator predictions – samples are associated columnwise. From left to right, scans represent a corridor and a large room. As shown in the figure, even though predictions are cluttered, the generator successfully learns the target distributions and it is able to generalize to different areas of the environment. Most importantly, Figure 1(b) confirms that a more compact – but equally informative – representation of the input data, can be used as a surrogate to significantly reduce dimensionality of data and dimension of the networks of the GAN.

## 4 CONCLUSION

In this paper we introduced a novel architecture to achieve online training and inference of missing spatial knowledge for a mobile robot. We validated the ability of GUESs to (1) exploit a learned latent representation of environment and to (2) extend the robot sensory horizon at real-time. However, the performance of the approach is still limited, and it has to be further improved in order to support decision-making. Our ultimate goal is to extend the robot horizon to predict complete spatial entities and to generalize the architecture to objects in the environment.

## REFERENCES


[1] Alireza Ahrary, Amir A.F. Nassiraei, and Masumi Ishikawa. 2007. A study of an autonomous mobile robot for a sewer inspection system. *Artificial Life and Robotics* 11, 1 (01 Jan 2007), 23–27. https://doi.org/10.1007/s10015-006-0392-x

[2] S. Bai, J. Wang, F. Chen, and B. Englot. 2016. Information-theoretic exploration with Bayesian optimization. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 1816–1822. https://doi.org/10.1109/IROS.2016.7759289

[3] H. J. Chang, C. S. G. Lee, Y. Lu, and Y. C. Hu. 2007. P-SLAM: Simultaneous Localization and Mapping With Environmental-Structure Prediction. *IEEE Transactions on Robotics* 23, 2 (April 2007), 281–293. https://doi.org/10.1109/TRO.2007.892230

[4] Michael Jae-Yoon Chung*, Andrzej Pronobis*, Maya Cakmak, Dieter Fox, and Rajesh P. N. Rao. 2016. Autonomous Question Answering with Mobile Robots in Human-Populated Environments. In *Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Daejeon, Korea. https://doi.org/10.1109/IROS.2016.7759146

[5] Jiankang Deng, Shiyang Cheng, Niannan Xue, Yuxiang Zhou, and Stefanos Zafeiriou. 2018. UV-GAN: Adversarial Facial UV Map Completion for Pose-Invariant Face Recognition. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[6] Felix Duvallet, Matthew R. Walter, Thomas Howard, Sachithra Hemachandra, Jean Oh, Seth Teller, Nicholas Roy, and Anthony Stentz. 2016. *Inferring Maps and Behaviors from Natural Language Instructions.* Springer International Publishing, 373–388. https://doi.org/10.1007/978-3-319-23778-7_25

[7] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative Adversarial Nets. In *Advances in Neural Information Processing Systems 27*, Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, and K. Q. Weinberger (Eds.). Curran Associates, Inc., 2672–2680. http://papers.nips.cc/paper/5423-generative-adversarial-nets.pdf

[8] G. Grisetti, C. Stachniss, and W. Burgard. 2007. Improved Techniques for Grid Mapping With Rao-Blackwellized Particle Filters. *IEEE Transactions on Robotics* 23, 1 (Feb 2007), 34–46. https://doi.org/10.1109/TRO.2006.889486

[9] Kapil D. Katyal, Katie M. Popek, Chris Paxton, Joseph L. Moore, Kevin C. Wolfe, Philippe Burlina, and Gregory D. Hager. 2018. Occupancy Map Prediction Using Generative and Fully Convolutional Networks for Vehicle Navigation. *CoRR* abs/1803.02007 (2018). arXiv:1803.02007 http://arxiv.org/abs/1803.02007

[10] Mohammad Al Khawaldah and Andreas Nüchter. 2015. Enhanced frontier-based exploration for indoor environment with multiple robots. *Advanced Robotics* 29, 10 (2015), 657–669. https://doi.org/10.1080/01691864.2015.1015443 arXiv:https://doi.org/10.1080/01691864.2015.1015443

[11] Diederik P. Kingma and Max Welling. 2013. Auto-Encoding Variational Bayes. *CoRR* abs/1312.6114 (2013).

[12] Stefan Kohlbrecher, Johannes Meyer, Thorsten Graber, Karen Petersen, Uwe Klingauf, and Oskar von Stryk. 2014. Hector Open Source Modules for Autonomous Mapping and Navigation with Rescue Robots. In *RoboCup 2013: Robot World Cup XVII*, Sven Behnke, Manuela Veloso, Arnoud Visser, and Rong Xiong (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 624–631.

[13] Raffaele Limosani, Raffele Esposito, Alessandro Manzi, Giancarlo Teti, Filippo Cavallo, and Paolo Dario. 2018. Robotic delivery service in combined outdoorâĂŞindoor environments: technical analysis and user evaluation. *Robotics and Autonomous Systems* 103 (2018), 56 – 67. https://doi.org/10.1016/j.robot.2018.02.001

[14] E. Nelson and N. Michael. 2015. Information-theoretic occupancy grid compression for high-speed information-based exploration. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 4976–4982. https://doi.org/10.1109/IROS.2015.7354077

[15] Tobias Neumann, Alexander Ferrein, Stephan Kallweit, and Ingrid Scholl. 2014. Towards a mobile mapping robot for underground mines. In *Proc. of the 2014 PRASA, RobMech and AfLaI Int. Joint Symposium, Cape Town, South Africa.*

[16] S. Oßwald, M. Bennewitz, W. Burgard, and C. Stachniss. 2016. Speeding-Up Robot Exploration by Exploiting Background Information. *IEEE Robotics and Automation Letters* 1, 2 (July 2016), 716–723. https://doi.org/10.1109/LRA.2016.2520560

[17] K. Otsu, A. Agha-Mohammadi, and M. Paton. 2018. Where to Look? Predictive Perception With Applications to Planetary Exploration. *IEEE Robotics and Automation Letters* 3, 2 (April 2018), 635–642. https://doi.org/10.1109/LRA.2017.2777526

[18] Andrzej Pronobis and Rajesh P. N. Rao. 2017. Learning Deep Generative Spatial Models for Mobile Robots. In *Proceedings of the 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Vancouver, BC, Canada. https://doi.org/10.1109/IROS.2017.8202235 arXiv:cs.RO/1610.02627

[19] Stéphane Ross, Geoffrey J Gordon, and Drew Bagnell. 2011. A Reduction of Imitation Learning and Structured Prediction to No-Regret Online Learning. In *International Conference on Artificial Intelligence and Statistics*. 627–635.

[20] Cyrill Stachniss and Wolfram Burgard. 2003. Exploring Unknown Environments with Mobile Robots Using Coverage Maps. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence (IJCAI'03)*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1127–1132. http://dl.acm.org/citation.cfm?id=1630659.1630821

[21] Joan Vallvé and Juan Andrade-Cetto. 2015. Potential information fields for mobile robot exploration. *Robotics and Autonomous Systems* 69 (2015), 68 – 79. https://doi.org/10.1016/j.robot.2014.08.009 Selected papers from 6th European Conference on Mobile Robots.