Learning to Control Drones in Natural Environments: A Survey

Antonio Loquercio and Davide Scaramuzza

***Abstract*— Lightweight, autonomous drones are soon expected to be used in a wide variety of tasks such as aerial surveillance, delivery, or monitoring of existing architectures. A large body of literature in robotic perception and control exists. Existing methods are mature, but not robust, therefore hindering drones' deployment in natural environments, like a city or a forest. Indeed, in unstructured and dynamic scenarios, drones face numerous challenges to navigate autonomously in a feasible and safe way. A recent line of research has exploited the "perception-awareness" of deep learning techniques to unlock autonomous flight in uncontrolled environments. This is motivated by the insight that traditional methods relying on global state estimates in the form of robot poses are doomed to fail because of the inherent difficulties of pose estimation at high speed along with their inability to adequately cope with dynamic environments. In this paper, we survey existing learning-based methods for drone navigation and identify open areas of research for future work.**

I. INTRODUCTION

Safe and reliable navigation of autonomous systems, e.g. unmanned aerial vehicles (UAVs), is a challenging open problem in robotics. Being able to successfully navigate while avoiding obstacles is indeed crucial to unlock many applications of robotics, e.g. surveillance, construction monitoring, delivery, and emergency response [1], [2], [3]. A robotic system facing the aforementioned tasks should simultaneously solve many challenges in perception, control, and localization. These become particularly difficult when working in uncontrolled environments, *e.g.* forests or streets of cities, as the one illustrated in Fig. 1. In those cases, the autonomous agent is not only expected to navigate while avoiding collisions but also to safely interact with other agents present in the environment, such as pedestrians or cars.

The traditional approach to tackle this problem is a two step interleaved process consisting of (i) automatic localization in a given map (using GPS, visual and/or range sensors), and (ii) computation of control commands to allow the agent to avoid obstacles while achieving its goal [1], [4]. Even though advanced SLAM algorithms enable localization under a wide range of conditions [5], visual aliasing, dynamic scenes, and strong appearance changes can drive the perception system to unrecoverable errors. Moreover, keeping the perception and control blocks separated not only hinders any possibility of positive feedback between them, but also introduces the challenging problem of inferring control commands from 3D maps. Recently, new approaches based on deep learning have offered a way to learn end-to-end flying policies, tightly coupling perception and control [6], [7], [8], [9]. The most successful of those methods are based on supervised-learning, since they offer an effective and sample efficient way to learn flying policies, but leave open the issue of the domain shift between the teacher and the learner. Even though systems based on deep learning achieve remarkable results, we believe that those systems are still not ready to completely replace traditional "map-localize-plan" approaches for drone navigation. Indeed, we speculate that learning-based and traditional approaches are going to complement each other and enable drones to accomplish the most challenging tasks.

The authors are with the Robotics and Perception Group, Dep. of Informatics University of Zurich and Dep. of Neuroinformatics of the University of Zurich and ETH Zurich, Switzerland `http://rpg.ifi.uzh.ch`.

Fig. 1. Enable a drone to autonomously fly through highly unstructured environments such as the streets of a city, represents a challenging task for robotics. Current methods based on deep learning have made a first step in this direction, but still many technical and theoretical questions remain open.

II. OVERVIEW OF DEEP LEARNING FOR DRONE NAVIGATION

A wide variety of techniques for drone navigation and obstacle avoidance can be found in the literature. At high level, these methods differ for the kind of sensory input and processing employed to control the flying platform. We divide those approaches in three different categories:

- Classical approaches based on mapping, localization, and planning.
- Imitation Learning methods.
- Reinforcement Learning approaches.

In the following, we survey existing works from these cathegories and identify open areas of research for future work.

A. *Map-Localize-Plan*

An Unmanned Aerial Vehicle is usually provided with GPS, range, and visual sensors to estimate the system state, infer the presence of obstacles, and perform path planning [1], [4]. Nevertheless, those systems are still prone to fail in environments where the GPS signal is weak as in *e.g.* the streets of a city or a forest trail. In addition, it is still not clear how to detect and avoid the static and dynamic obstacles that are present in those environment. The prevalent approach in GPS-denied scenarios is SLAM, where the robot simultaneously builds a map of the environment and self-localizes in it [5]. These approaches may fail, however, when localizing in a map that was created in significantly different conditions or during periods of high acceleration (because of motion blur and loss of feature tracking). Additionally, enforcing global consistency leads to a larger computational complexity and a significant difficulty in coping with dynamic environments. Indeed, SLAM methods enable navigation only in a "predominantly-static world" where waypoints and (optionally) collision-free trajectories can be statically defined.

B. Imitation Learning Methods

Recently, there has been an increasing research effort in directly learning control policies from raw sensory data using imitation learning. Given its relatively low sample complexity (i.e. not much sample data is required to generalize) and its implementation simplicity, supervised learning has become the predominant tool used to learn visual-motor policies [6], [7], [10], [8], [11], [12]. The supervisory signal may come from a human expert [6], hard-coded trajectories [11], or model predictive control [12]. However, when working in the streets of a city, it can be both tedious and dangerous to collect a large set of expert trajectories, or evaluate partially trained policies [6]. A possible approach is to collect data in simulation and then transfer the learned policy to the real world. To generate very basic navigation policies, however, it is required either a lot of photorealistic data [9], or some real world examples [13]. Therefore, in order to safely and efficiently acquire data, the authors of [8] proposed to use cameras mounted on cars and bicycles. Tightly coupling perception and control, the resulting visual motor policy unlocks good generalization performance on a set of environments unseen during training. Similar works in [7], [10] trained a deep neural network from video collected by a mountain hiker to detect forest trails. The results of the detection were then used to make a UAV fly through forests. Clearly, the main disadvantage of these approaches consists in the domain-shift between the expert providing supervision and the learning agent. Indeed, in most of the aforementioned methods the learned flying policies are limited to planar motion, not fully exploiting the agile dynamics of quadrotors. An interesting direction for future work should be unifying supervised learning with model-based approaches for drone control. By taking the best of the two worlds, the resulting approach will provide not only generalizing and sample efficient solutions, but will also ease the domain-shift problem and fully exploit drones' dynamics.

C. Reinforcement Learning (RL) Methods

In the last few years, RL-based algorithms have gained a lot of popularity in the research community [9], [14], [15]. With respect to supervised methods, RL offers the advantage of not requiring an expert to imitate since the learning signal comes from direct interaction with the environment. This automatically solves the domain-shift problem. The cost to be paid for such a feature is that RL methods have extremely high sample complexity and therefore require a large amount of robot experience to learn generalizing policies. This has mainly hindered RL methods to be widely adopted in the quadrotors literature. Indeed, robotic experience is both costly and dangerous to acquire with safety critical platforms such as drones. A promising approach has been to use simulations in order to get cheap training data for reinforcement learning tasks while testing the learned policy in the real world [9], [13]. Clearly, this method suffers from the domain shift between simulation and reality and might require some real-world data to be able to generalize [13]. On the one hand available simulators can already model the quadrotors dynamics very well [16]. On the other they still fail to effectively model drag effects [17] which are potentially crucial to learn how to behave when flying close to structure. In addition, while there has been a lot of effort in making simulators more photorealistic [18], [19], the quality of rendered images still does not enable learned policies to fully transfer to real world. Nonetheless, some pioneering work has been done in [9], where the authors learned a simple flying policy by using only simulated robot experience.

RL methods have raised more success on grounded platforms [20], [21], [22]. When the robot can afford to experience collision, good driving behaviour can only be learned from real-world data. Another approach consists in making most of the experiences in simulation, and then fine-tune the learned policies with real-world data [9]. Eventually, if the robotic platform can be precisely modeled, model-based RL techniques [20], [21] can be used to learn advanced driving policies. We believe that future work will take advantage of known quadrotors models to efficiently learn generalizing flying policies with model-based reinforcement learning and leverage recent computer vision techniques [23] to make policies learned in simulation better generalize to the real world.

III. CONCLUSIONS

This paper presents a survey on existing deep learning methods for drone navigation. Even though they have a stronger perception awareness with respect to classical approaches, learning based solutions are still failing to exploit the dynamics and agility of drones. Indeed, we believe that learning-based and traditional approaches are going to complement each other and allow drones to be used in a wide variety of navigation related applications such as aerial delivery or search and rescue.

References

[1] S. Scherer, J. Rehder, S. Achar, H. Cover, A. Chambers, S. Nuske, and S. Singh, "River mapping from a flying robot: state estimation, river detection, and obstacle mapping," *Autonomous Robots*, vol. 33, no. 1-2, pp. 189–214, 2012.

[2] N. Michael, S. Shen, K. Mohta, Y. Mulgaonkar, V. Kumar, K. Nagatani, Y. Okada, S. Kiribayashi, K. Otake, K. Yoshida *et al.*, "Collaborative mapping of an earthquake-damaged building via ground and aerial robots," *Journal of Field Robotics*, vol. 29, no. 5, pp. 832–841, 2012.

[3] M. Faessler, F. Fontana, C. Forster, E. Mueggler, M. Pizzoli, and D. Scaramuzza, "Autonomous, vision-based flight and live dense 3D mapping with a quadrotor MAV," *J. Field Robot.*, vol. 33, no. 4, pp. 431–450, 2016.

[4] S. Shen, Y. Mulgaonkar, N. Michael, and V. Kumar, "Multi-sensor fusion for robust autonomous flight in indoor and outdoor environments with a rotorcraft MAV," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 4974–4981.

[5] S. Lynen, T. Sattler, M. Bosse, J. Hesch, M. Pollefeys, and R. Siegwart, "Get out of my lab: Large-scale, real-time visual-inertial localization," in *Robotics: Science and Systems XI*, jul 2015.

[6] S. Ross, N. Melik-Barkhudarov, K. S. Shankar, A. Wendel, D. Dey, J. A. Bagnell, and M. Hebert, "Learning monocular reactive UAV control in cluttered natural environments," in *IEEE Int. Conf. Robot. Autom. (ICRA)*, 2013, pp. 1765–1772.

[7] A. Giusti, J. Guzzi, D. C. Cirean, F. L. He, J. P. Rodrguez, F. Fontana, M. Faessler, C. Forster, J. Schmidhuber, G. D. Caro, D. Scaramuzza, and L. M. Gambardella, "A machine learning approach to visual perception of forest trails for mobile robots," *IEEE Robotics and Automation Letters*, 2016.

[8] A. Loquercio, A. I. Maqueda, C. R. Del Blanco, and D. Scaramuzza, "Dronet: Learning to fly by driving," *IEEE Robotics and Automation Letters*, 2018.

[9] F. Sadeghi and S. Levine, "CAD2rl: Real single-image flight without a single real image," in *Robotics: Science and Systems XIII*, jul 2017.

[10] N. Smolyanskiy, A. Kamenev, J. Smith, and S. Birchfield, "Toward low-flying autonomous mav trail navigation using deep neural networks for environmental awareness," *IEEE/RSJ Int. Conf. Intell. Robot. Syst. (IROS)*, 2017.

[11] D. Gandhi, L. Pinto, and A. Gupta, "Learning to fly by crashing," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, sep 2017.

[12] G. Kahn, T. Zhang, S. Levine, and P. Abbeel, "PLATO: Policy learning using adaptive trajectory optimization," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, may 2017.

[13] Y. Zhu, R. Mottaghi, E. Kolve, J. J. Lim, A. Gupta, L. Fei-Fei, and A. Farhadi, "Target-driven visual navigation in indoor scenes using deep reinforcement learning," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*, may 2017.

[14] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," *arXiv preprint arXiv:1509.02971*, 2015.

[15] J. Schulman, S. Levine, P. Abbeel, M. Jordan, and P. Moritz, "Trust region policy optimization," in *Proceedings of the 32nd International Conference on Machine Learning (ICML-15)*, 2015, pp. 1889–1897.

[16] F. Furrer, M. Burri, M. Achtelik, and R. Siegwart, "Robot operating system (ros)," *Studies Comp.Intelligence Volume Number:625*, 2016.

[17] M. Faessler, A. Franchi, and D. Scaramuzza, "Differential flatness of quadrotor dynamics subject to rotor drag for accurate tracking of high-speed trajectories," *IEEE Robotics and Automation Letters*, pp. 620–626, 2018.

[18] S. Shah, D. Dey, C. Lovett, and A. Kapoor, "Airsim: High-fidelity visual and physical simulation for autonomous vehicles," in *Field and Service Robotics*, 2017.

[19] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," in *Proceedings of the 1st Annual Conference on Robot Learning*, 2017, pp. 1–16.

[20] P. Drews, G. Williams, B. Goldfain, E. A. Theodorou, and J. M. Rehg, "Aggressive deep driving: Combining convolutional neural networks and model predictive control," in *Proceedings of the 1st Annual Conference on Robot Learning*, 2017.

[21] G. Williams, N. Wagener, B. Goldfain, P. Drews, J. M. Rehg, B. Boots, and E. A. Theodorou, "Information theoretic mpc for model-based reinforcement learning," in *Robotics and Automation (ICRA), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1714–1721.

[22] G. Kahn, A. Villaflor, B. Ding, P. Abbeel, and S. Levine, "Self-supervised deep reinforcement learning with generalized computation graphs for robot navigation," *arXiv preprint arXiv:1709.10489*, 2017.

[23] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *The IEEE International Conference on Computer Vision (ICCV)*, Oct 2017.