Aligning Robot Navigation Behaviors with Human Intentions and Preferences

Haresh Karnan
Walker Department of Mechanical Engineering
The University of Texas at Austin
Austin, Texas, USA
haresh.miriyala@utexas.edu

***Abstract*—Advances in the fields of machine learning and artificial intelligence have enabled mobile robots to become increasingly intelligent, efficient, and autonomous. However, as these agents become more capable, there is a danger of *value misalignment*, i.e., that the agent's internal objectives may result in behaviors that do not align with the intentions and preferences of humans. Towards addressing this problem in robot navigation, this dissertation introduces novel algorithms and a large-scale dataset to align the internal objectives of mobile robots with human intentions and preferences, through the lens of imitation learning. The first contribution of this thesis is an imitation learning algorithm capable of learning navigation behaviors from physically different agents, such as humans, while effectively addressing the challenge of egocentric viewpoint mismatch, a prevalent issue in visual imitation learning. Next, to enable socially compliant robot navigation in human-occupied environments, a large-scale dataset of navigation demonstrations is introduced, which we show enables learning socially compliant navigation policies using imitation learning. Finally, in the context of off-road navigation, the dissertation introduces a self-supervised multi-modal representation learning algorithm, which, by utilizing easy-to-collect unconstrained robot experiences, learns relevant terrain representations that enable operator-preference-aligned off-road visual navigation, such as an exciting real-world deployment of the Spot robot semi-autonomously traversing a 3-mile trail successfully. In summary, the contributions of this dissertation take a step toward enabling mobile robots to navigate in alignment with human intentions and preferences.**

***Index Terms*—Imitation Learning, Reinforcement Learning, Self-Supervised Learning, Robot Navigation**

I. INTRODUCTION

Advancements in the field of machine learning and artificial intelligence have ushered in a new era where mobile robots are increasingly intelligent, efficient, and autonomous [1]–[5]. The significant growth in compute capabilities such as hardware accelerators, allows deep learning-based methods to learn from vast amounts of data, enabling real-time perception [6]–[8], planning [9], [10], and control [4], [11]–[15] in mobile robots. However, as these agents become more advanced, there's increasing concern over value misalignment—where their internal objectives may not align with human intentions and preferences. For instance, a vacuuming robot solely incentivized to gather more dust may potentially begin spewing dust to collect more rewards [16], an unintended behavior commonly known as reward hacking [16], [17] in the reinforcement learning community. Such discrepancies have been previously noted in robotics [17]–[21], emphasizing the pressing need to align the behavior of autonomous agents with human intentions and preferences.

This dissertation focuses on the value misalignment problem in the specific context of autonomous robot navigation in unstructured indoor and outdoor environments. The aim is to ensure that these robots not only navigate efficiently but also in a manner that aligns with human intentions and preferences. Classical heuristic-based methods [14], [22]–[24] excel in structured environments but face challenges in dynamic, real-world settings, such as navigating outdoor terrains and around humans. In particular, these methods struggle in unfamiliar, unmapped environments, unstructured off-road environments with diverse terrains, and scenarios demanding safe navigation around humans.

To address these limitations of classical navigation methods, the research community is increasingly employing learning-based techniques that can discern patterns from data [14], enabling adaptive robot behavior [4], [11], [25]–[29]. However, the efficacy of these methods critically depends on articulating a precise objective function that truly reflects human intentions and preferences for particular tasks. Misalignment, due to vague objectives, can produce unintended behaviors, like a robot trampling flowers or cutting impolitely through crowds, sometimes culminating in severe safety issues, especially involving vulnerable populations. Such public safety incidents have even resulted in bans on autonomous mobile robots in public spaces [30], [31].

In light of the value misalignment problem with autonomous robots, this dissertation adopts an imitation learning approach for robot navigation and presents three key contributions. First, we introduce VOILA [32], an Imitation from Observation (IFO) algorithm that learns a navigation policy by imitating video-only demonstrations from physically different experts, such as humans, overcoming egocentric viewpoint mismatch, a significant problem in visual imitation learning. Secondly, to enable socially compliant robot navigation, this dissertation introduces a large-scale dataset called SCAND [33] consisting of 25 miles of socially compliant teleoperated navigation demonstrations, which enables learning socially compliant navigation policies using imitation learning. Lastly, to enable off-road terrain awareness and navigation over terrains in an operator-aligned manner, this dissertation introduces STERLING, a self-

Fig. 1. Policy rollout trajectories of the VOILA agent (green) successfully imitating a demonstrated behavior (black) of patrolling a rectangular hallway clockwise. The demonstration consists of a video gathered by a human walking while using a handheld camera that is considerably higher than the robot's camera (introducing significant viewpoint mismatch). We see that the VOILA agent is able to successfully imitate the expert demonstration even in the presence of this egocentric viewpoint mismatch.

supervised terrain representation learning approach that learns terrain representations through multi-modal self-supervision from easy-to-collect unconstrained robot experience. In summary, this dissertation presents two novel algorithms and a large-scale dataset that enable mobile robots to navigate in unstructured indoor and outdoor environments in accordance with human intentions and preferences.

II. Contributions

In this section, we introduce the three contributions of this dissertation. First, we present VOILA, an Imitation from Observation (IfO) algorithm for autonomous navigation. We then introduce SCAND, a large-scale dataset of demonstrations for socially compliant robot navigation. Lastly, we present STERLING, a self-supervised representation learning approach to learn relevant terrain representations that enable operator-aligned visual terrain-aware navigation in outdoor environments.

A. *Visual Imitation Learning for Autonomous Navigation* [1]

The task of autonomous robot navigation is to enable a robot to navigate autonomously, with minimal or no human supervision during deployment from one location to another, involving sequential decision-making along the path. Reinforcement Learning is a branch of machine learning that addresses sequential decision-making problems such as the task of autonomous mobile robot navigation. Applying reinforcement learning-based approaches to physical robots requires a well-defined dense reward function that is informative of the task, aligns with the objective of the human operator, provides real-time feedback, and should not suffer from unintended behaviors when optimized. While defining such as reward function is hard, we seek to imitate navigation behaviors from demonstrator agents, such as humans. However, doing so introduces an egocentric viewpoint mismatch between the observation spaces of the demonstrator and imitator agents, as shown in Fig. 1, posing a major challenge for existing state-of-the-art imitation learning algorithms [34]–[37].

Towards addressing the egocentric viewpoint mismatch problem in visual imitation learning and to learn navigation policies by imitating physically different agents such as humans, we introduce Visual Observation-only Imitation Learning for Autonomous navigation (VOILA) [32]. VOILA introduces a novel reward function for IfO, overcoming egocentric viewpoint mismatch using existing viewpoint-invariant keypoint detectors such as SUPERPOINT [38]. We perform experiments in the physical world with the Clearpath Jackal mobile robot, and in a simulated photorealistic AirSim [39] simulator, and find that compared to the existing state-of-the-art adversarial IfO algorithm GAIfO [40], VOILA is sample-efficient while being successful at imitating video-only demonstrations even in the presence of significant egocentric viewpoint mismatch, and also generalizes to unseen environments, while solving the task, as intended by the human demonstrator. Fig. 1 shows the performance of VOILA in the physical robot experiments. The trajectories of different approaches are traced and superimposed on a map of the environment, used only for visualization purposes. We see that compared to a sub-optimal policy and a random policy, VOILA is able to successfully imitate the expert's video-only demonstration and thereby learn the task of clockwise hallway patrol.

[1]A recorded presentation of this work is available at: VOILA ICRA'22

Fig. 2. Five example scenarios from SCAND showing the RGB image and below it the accompanying pointcloud with the monocular image. From left to right, the scenarios have the tags "Street Crossing", "Narrow Doorway, "Navigating Large Crowds", "Vehicle Interaction", and "Crossing Stationary Queue."

B. Socially Compliant Robot Navigation [2]

Social navigation, or social compliance, is the capability of an autonomous agent, such as a robot, to navigate in a socially compliant manner in the presence of other autonomous agents such as humans. With the emergence of autonomously navigating mobile robots in human-populated environments (e.g., domestic service robots in homes and restaurants and food delivery robots on public sidewalks), incorporating social compliance becomes essential to ensuring safe and comfortable navigation in human presence.

The use of imitation learning algorithms, such as VOILA [32] described in the previous subsection, for social navigation is currently hindered by a lack of large-scale datasets that capture socially compliant robot navigation demonstrations in the wild. To fill this gap, we introduce Socially Compli-Ant Navigation Dataset (SCAND) [33]—a large-scale, first-person-view dataset of socially compliant robot navigation demonstrations[3]. The SCAND dataset contains 8.7 hours, 138 trajectories, and 25 miles of socially compliant, human-teleoperated driving demonstrations that comprise multi-modal data streams including 3D lidar, joystick commands, odometry, visual and inertial information, collected on two morphologically different mobile robots—a Boston Dynamics Spot and a Clearpath Jackal—by four different human demonstrators in both indoor and outdoor environments within the UT Austin campus. In addition to the multi-modal sensor data, SCAND also contains 12 coarse social interaction labels that occur along every trajectory. Fig 2 shows five example scenarios and their associated tags in SCAND.

We perform validation through physical robot experiments on two indoor scenarios, as shown in Fig. 3 and find that navigation policies learned by imitation learning on SCAND generate socially compliant behaviors. Specifically, we ask fourteen human participants to evaluate the behavior cloning agent trained on SCAND, and the `movebase` [22] agent (a classical, heuristic-based navigation stack available in ROS) on two questions related to safety and social compliance. On average, the participants felt the imitative policy trained using SCAND was more socially compliant and safe in comparison to the `movebase` agent in both scenarios shown in Fig. 3.

[2]A recorded presentation of this work is available at SCAND IROS'22
[3]SCAND is openly accessible via the Texas Robotics Dataverse

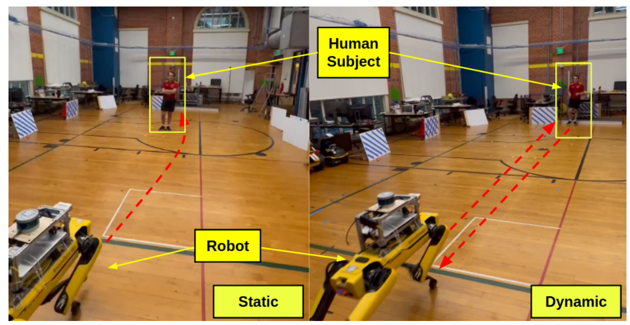

Fig. 3. Evaluating the local planner agent trained using Behavior Cloning on SCAND. Scenario on the left shows a stationary human in the robot's path and the scenario on the right shows a human walking to the location of the robot. The robot is evaluated on social compliance and safety as it navigates to its goal position.

C. Operator Preference-Aligned Off-Road Navigation [4]

Off-road navigation is emerging as a crucial capability for autonomous mobile robots envisioned for use in a growing number of outdoor applications such as agricultural operations [41], package delivery [42], and search and rescue [43]. A major challenge for existing mobile robots that are deployed in outdoor real-world environments is the capability to be terrain-aware, i.e., visually identify and distinguish different terrains. However, instilling robots with visual terrain awareness has been challenging, as prior works require large-labeled datasets [44]–[46], structured demonstrations [47], or learn task-specific behaviors [48]–[50] that may not generalize, or learn behaviors that may not align with operator's preferences.

In this work, we focus on self-supervised learning of terrain representations using multi-modal unconstrained robot experiences, eliminating the requirement of manually labeled large-scale datasets. We present a novel algorithm, Self-supervised TErrain Representation LearnING from unconstrained robot experience (STERLING) [51], [52], which harnesses onboard multi-modal data for effective terrain representation learning. Through STERLING, we can proactively query a human operator's terrain preferences, enabling operator-aligned visual navigation on autonomous mobile robots.

[4]This work is recently accepted for publication at CoRL'23 and an initial version was presented at the PT4R workshop at ICRA'23

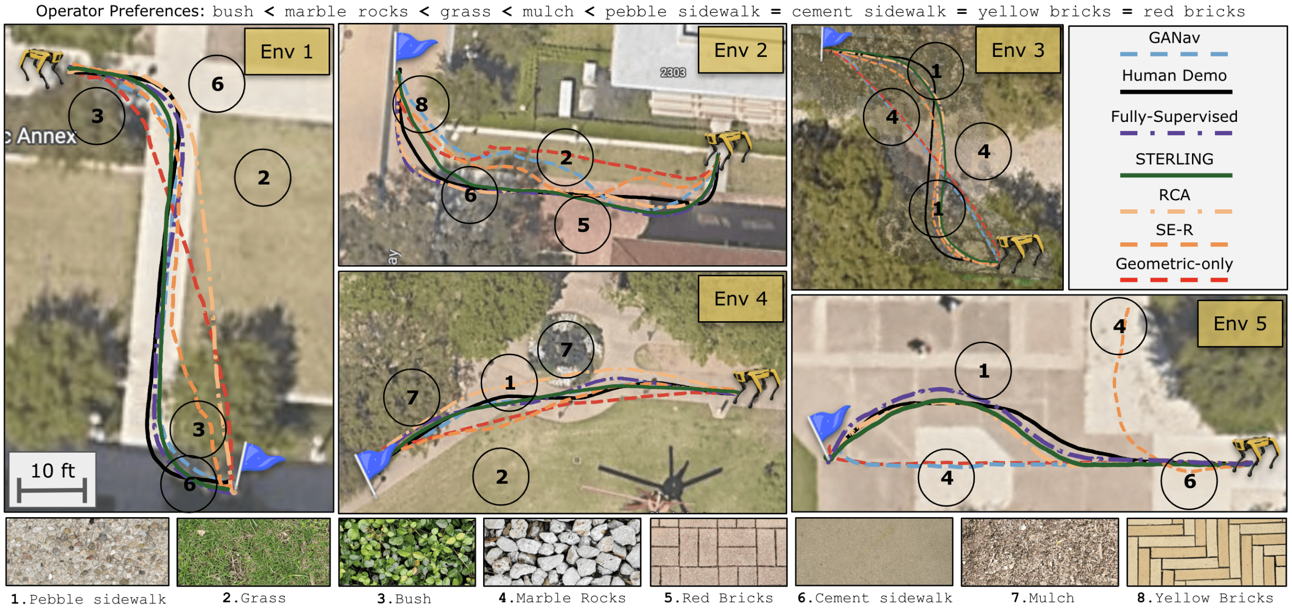

Fig. 4. Trajectories traced by different approaches in 5 environments containing 8 different terrains. The operator preferences are shown above. We see that STERLING navigates in an operator-preference aligned manner, by preferring `cement sidewalk, red bricks, pebble sidewalk, and yellow bricks` over `mulch, grass, marble rocks, and bush`, outperforming other baselines and performing on-par with the Fully-Supervised approach.

We evaluate STERLING against state-of-the-art visual off-road navigation algorithms such as GANav [46], RCA [48], SE-R [49], geometric-only planner [53] and a fully-supervised approach. We perform quantitative trials [5], comparing STERLING with the baseline approaches, within the UT Austin university campus. The trajectories traced by all algorithms, in five environments, along with the operator's preferences are shown in Fig. 4. We additionally perform a large-scale qualitative evaluation of STERLING by deploying it in a 3-mile trail [6] on a Spot robot [7]. The robot is tasked with semi-autonomously traversing the trail while adhering to operator-defined terrain preferences for the trail. Fig. 5 shows the 3-mile trajectory traced by the robot and the two failure cases that required manual intervention. This large-scale qualitative experiment demonstrates both the reliability of STERLING in diverse real-world conditions and its efficacy in operator-aligned off-road navigation.

Fig. 5. A large-scale qualitative evaluation of STERLING on a 3-mile outdoor trail. STERLING features successfully complete the trail with only two manual interventions (shown in red).

III. Conclusion

In this dissertation, we address the crucial challenge of aligning robot navigation objectives with human intentions and preferences in unstructured environments, by presenting three novel contributions. First, we introduced VOILA, a visual imitation learning algorithm that adeptly overcomes the common egocentric viewpoint mismatch problem. By leveraging a unique reward function, it successfully imitates, on a mobile robot, navigation demonstrations from physically different agents, such as humans. Secondly, we introduced SCAND, a large-scale dataset of teleoperated navigation demonstrations, which facilitates the learning of socially compliant navigation policies through imitation learning. Finally, we introduced STERLING, a self-supervised multi-modal representation learning algorithm. STERLING learns relevant terrain representations from easy-to-collect unconstrained robot experiences, making it feasible to discern operator terrain preferences and traverse in an operator-aligned manner. We demonstrated this on a 3-mile hike performed by the Spot robot, using STERLING features, which it completes successfully, demonstrating its robustness and efficacy in real-world conditions. Collectively, the contributions in this dissertation take a step towards addressing the value misalignment problem in robot navigation, enabling harmonious coexistence between robots and humans in diverse real-world environments.

[5] A video summary of quantitative experiments is available at: https://youtu.be/7WI41DfJQ2k

[6] Ann and Roy Butler Trail, Austin, TX, USA

[7] A video of the robot traversing the 3-mile trail is available at: https://youtu.be/dQb1XzocdtE

References

[1] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, X. Zhang, J. Zhao, and K. Zieba, "End to end learning for self-driving cars," 2016. [Online]. Available: https://arxiv.org/abs/1604.07316

[2] G. Kahn, A. Villaflor, B. Ding, P. Abbeel, and S. Levine, "Self-supervised deep reinforcement learning with generalized computation graphs for robot navigation," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 5129–5136.

[3] A. Kendall, J. Hawke, D. Janz, P. Mazur, D. Reda, J.-M. Allen, V.-D. Lam, A. Bewley, and A. Shah, "Learning to drive in a day," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 8248–8254.

[4] Y. Pan, C.-A. Cheng, K. Saigol, K. Lee, X. Yan, E. Theodorou, and B. Boots, "Agile autonomous driving using end-to-end deep imitation learning," in *Robotics: Science and Systems*, 2018.

[5] D. Silver, J. A. Bagnell, and A. Stentz, "Learning from demonstration for autonomous navigation in complex unstructured terrain," *The International Journal of Robotics Research*, vol. 29, no. 12, pp. 1565–1592, 2010.

[6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788.

[7] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.

[8] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark *et al.*, "Learning transferable visual models from natural language supervision," in *International Conference on Machine Learning*. PMLR, 2021, pp. 8748–8763.

[9] D. S. Chaplot, D. Pathak, and J. Malik, "Differentiable spatial planning using transformers," 2021. [Online]. Available: https://arxiv.org/abs/2112.01010

[10] A. Tamar, Y. Wu, G. Thomas, S. Levine, and P. Abbeel, "Value iteration networks," *Advances in neural information processing systems*, vol. 29, 2016.

[11] M. Bojarski, D. D. Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, X. Zhang, J. Zhao, and K. Zieba, "End to end learning for self-driving cars."

[12] H. Karnan, K. S. Sikand, P. Atreya, S. Rabiee, X. Xiao, G. Warnell, P. Stone, and J. Biswas, "Vi-ikd: High-speed accurate off-road navigation using learned visual-inertial inverse kinodynamics," *CoRR*, vol. abs/2203.15983, 2022. [Online]. Available: https://doi.org/10.48550/arXiv.2203.15983

[13] X. Xiao, B. Liu, G. Warnell, and P. Stone, "Motion control for mobile robot navigation using machine learning: a survey," 2020.

[14] ——, "Motion planning and control for mobile robot navigation using machine learning: a survey," 2020. [Online]. Available: https://arxiv.org/abs/2011.13112

[15] H. Kim, M. Jordan, S. Sastry, and A. Ng, "Autonomous helicopter flight via reinforcement learning," *Advances in neural information processing systems*, vol. 16, 2003.

[16] S. Russell, "P. norvig artificial intelligence: a modern approach," 2003.

[17] D. Hadfield-Menell, S. Milli, P. Abbeel, S. J. Russell, and A. Dragan, "Inverse reward design," *Advances in neural information processing systems*, vol. 30, 2017.

[18] D. Amodei, C. Olah, J. Steinhardt, P. Christiano, J. Schulman, and D. Mané, "Concrete problems in ai safety," 2016. [Online]. Available: https://arxiv.org/abs/1606.06565

[19] T. Everitt, "Towards safe artificial general intelligence," 2018.

[20] B. Ibarz, J. Leike, T. Pohlen, G. Irving, S. Legg, and D. Amodei, "Reward learning from human preferences and demonstrations in atari," 2018. [Online]. Available: https://arxiv.org/abs/1811.06521

[21] J. Lehman, J. Clune, D. Misevic, C. Adami, L. Altenberg, J. Beaulieu, P. J. Bentley, S. Bernard, G. Beslon, D. M. Bryson, N. Cheney, P. Chrabaszcz, A. Cully, S. Doncieux, F. C. Dyer, K. O. Ellefsen, R. Feldt, S. Fischer, S. Forrest, A. Fundefinedenoy, C. Gagńe, L. Le Goff, L. M. Grabowski, B. Hodjat, F. Hutter, L. Keller, C. Knibbe, P. Krcah, R. E. Lenski, H. Lipson, R. MacCurdy, C. Maestre, R. Miikkulainen, S. Mitri, D. E. Moriarty, J.-B. Mouret, A. Nguyen, C. Ofria, M. Parizeau, D. Parsons, R. T. Pennock, W. F. Punch, T. S. Ray, M. Schoenauer, E. Schulte, K. Sims, K. O. Stanley, F. Taddei, D. Tarapore, S. Thibault, R. Watson, W. Weimer, and J. Yosinski, "The surprising creativity of digital evolution: A collection of anecdotes from the evolutionary computation and artificial life research communities," *Artif. Life*, vol. 26, no. 2, p. 274–306, may 2020. [Online]. Available: https://doi.org/10.1162/artl_a_00319

[22] "ROS movebase navigation stack," http://wiki.ros.org/move_base, accessed: 2021-09-9.

[23] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numer. Math.*, vol. 1, no. 1, p. 269–271, dec 1959. [Online]. Available: https://doi.org/10.1007/BF01386390

[24] P. E. Hart, N. J. Nilsson, and B. Raphael, "A formal basis for the heuristic determination of minimum cost paths," *IEEE Transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.

[25] D. A. Pomerleau, *ALVINN: An Autonomous Land Vehicle in a Neural Network*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1989, p. 305–313.

[26] L. Tai, J. Zhang, M. Liu, and W. Burgard, "Socially compliant navigation through raw depth inputs with generative adversarial imitation learning," 2018.

[27] X. Xiao, B. Liu, G. Warnell, J. Fink, and P. Stone, "Appld: Adaptive planner parameter learning from demonstration," *IEEE Robotics and Automation Letters*, vol. 5, no. 3, pp. 4541–4547, 2020.

[28] H. Karnan, E. Yang, G. Warnell, J. Biswas, and P. Stone, "Wait, that feels familiar: Learning to extrapolate human preferences for preference aligned path planning," 2023.

[29] E. Yang, H. Karnan, G. Warnell, P. Stone, and J. Biswas, "Wait, that feels familiar: Learning to extrapolate human preferences for preference-aligned path planning," in *ICRA2023 Workshop on Pretraining for Robotics (PT4R)*, 2023.

[30] "'our streets are made for people': San francisco mulls ban on delivery robots," https://www.theguardian.com/sustainable-business/2017/may/31/delivery-robots-drones-san-francisco-public-safety-job-loss-fears-marble, accessed: 2022-08-31.

[31] "Toronto bans automated robots from sidewalk operation until further study," https://smartcitiesconnect.org/toronto-bans-automated-robots-from-sidewalk-operation-until-further-study/, accessed: 2022-08-31.

[32] H. Karnan, G. Warnell, X. Xiao, and P. Stone, "Voila: Visual-observation-only imitation learning for autonomous navigation," 2021. [Online]. Available: https://arxiv.org/abs/2105.09371

[33] H. Karnan, A. Nair, X. Xiao, G. Warnell, S. Pirk, A. Toshev, J. Hart, J. Biswas, and P. Stone, "Socially compliant navigation dataset (scand): A large-scale dataset of demonstrations for social navigation," 2022. [Online]. Available: https://arxiv.org/abs/2203.15041

[34] M. Bain and C. Sammut, "A framework for behavioural cloning," in *Machine Intelligence 15, Intelligent Agents [St. Catherine's College, Oxford, July 1995]*. GBR: Oxford University, 1999, p. 103–129.

[35] S. Ross, G. J. Gordon, and J. A. Bagnell, "A reduction of imitation learning and structured prediction to no-regret online learning," 2010. [Online]. Available: https://arxiv.org/abs/1011.0686

[36] J. Ho and S. Ermon, "Generative adversarial imitation learning," *Advances in neural information processing systems*, vol. 29, 2016.

[37] C. Finn, P. Christiano, P. Abbeel, and S. Levine, "A connection between generative adversarial networks, inverse reinforcement learning, and energy-based models," *arXiv preprint arXiv:1611.03852*, 2016.

[38] D. DeTone, T. Malisiewicz, and A. Rabinovich, "Superpoint: Self-supervised interest point detection and description," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2018, pp. 224–236.

[39] 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.

[40] F. Torabi, G. Warnell, and P. Stone, "Generative adversarial imitation from observation," *arXiv preprint arXiv:1807.06158*, 2018.

[41] T. Wang, B. Chen, Z. Zhang, H. Li, and M. Zhang, "Applications of machine vision in agricultural robot navigation: A review," *Computers and Electronics in Agriculture*, vol. 198, p. 107085, 2022.

[42] "Agility robotics launches next generation of digit: World's first human-centric, multi-purpose robot made for logistics work," 2023. [Online]. Available: https://agilityrobotics.com/news/2022/future-robotics-l3mjh

[43] X. Xiao, J. Dufek, and R. R. Murphy, "Autonomous visual assistance for robot operations using a tethered uav," 2019.

[44] M. Wigness, S. Eum, J. G. Rogers, D. Han, and H. Kwon, "A rugd dataset for autonomous navigation and visual perception in unstructured outdoor environments," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 5000–5007.

[45] P. Jiang, P. Osteen, M. Wigness, and S. Saripalli, "Rellis-3d dataset: Data, benchmarks and analysis," 2020.

[46] T. Guan, D. Kothandaraman, R. Chandra, A. J. Sathyamoorthy, K. Weerakoon, and D. Manocha, "Ganav: Efficient terrain segmentation for robot navigation in unstructured outdoor environments," 2021. [Online]. Available: https://arxiv.org/abs/2103.04233

[47] K. S. Sikand, S. Rabiee, A. Uccello, X. Xiao, G. Warnell, and J. Biswas, "Visual representation learning for preference-aware path planning," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 11 303–11 309.

[48] X. Yao, J. Zhang, and J. Oh, "Rca: Ride comfort-aware visual navigation via self-supervised learning," *arXiv preprint arXiv:2207.14460*, 2022.

[49] J. Zürn, W. Burgard, and A. Valada, "Self-supervised visual terrain classification from unsupervised acoustic feature learning," *IEEE Transactions on Robotics*, vol. 37, no. 2, pp. 466–481, 2020.

[50] T. Miki, J. Lee, J. Hwangbo, L. Wellhausen, V. Koltun, and M. Hutter, "Learning robust perceptive locomotion for quadrupedal robots in the wild," *Science Robotics*, vol. 7, no. 62, p. eabk2822, 2022.

[51] H. Karnan, E. Yang, D. Farkash, G. Warnell, J. Biswas, and P. Stone, "Self-supervised terrain representation learning from unconstrained robot experience," 2023. [Online]. Available: https://openreview.net/pdf?id=ShKj-4wLWyv

[52] ——, "Self-supervised terrain representation learning from unconstrained robot experience," in *ICRA2023 Workshop on Pretraining for Robotics (PT4R)*, 2023.

[53] J. Biswas, "Amrl autonomy stack," https://github.com/ut-amrl/graph_navigation, 2013.