Thermal-to-Depth Driven Safe Navigation for UAVs in Degraded Environments

Hürkan Şahin[1], Van Huyen Dang[1], Erdal Kayacan [1]

***Abstract*— We propose an end-to-end thermal navigation framework for UAVs operating in GPS-denied and visually degraded environments. By coupling monocular thermal-to-depth estimation with depth-based safe navigation, the framework allows predicted depth from non-radiometric thermal images to be used directly for policy learning and control. In addition, a depth-driven reward formulation is designed to encourage safer navigation behavior. Results on a custom dataset demonstrate substantially improved depth estimation accuracy, supporting the promise of the proposed approach for thermal UAV navigation in challenging environments.**

***Index Terms*— Thermal imaging, Thermal-to-depth estimation, Recurrent neural networks, UAV, SLAM, Reinforcement learning, Reward engineering**

I. Introduction

Maritime transport carries around 90–95% of global trade due to its cost efficiency, making it a critical backbone of international logistics [1]. The global merchant fleet consists of about 90,000 ships, with roughly 1–2% reaching end-of-life each year [2]. Shipbreaking involves dismantling these vessels and separating materials into waste and reusable components. However, ships contain hazardous materials such as asbestos, PCBs, oil residues, and heavy metals, which, if not properly managed, pose significant risks to both human health and the environment [3].

Unmanned aerial vehicle (UAV) provide an effective solution for operating in such hazardous and confined environments, reducing risks for human operators while leveraging onboard sensors for monitoring and localization, thereby improving situational awareness and mission success [4].

Thermal-infrared cameras offer important advantages in degraded environments [5], as they detect infrared radiation without requiring visible light, enabling operation through smoke, dust, or haze. Despite these benefits, thermal cameras introduce several challenges for reliable simultaneous localization and mapping (SLAM) integration [6]. Their 14/16-bit high dynamic range is incompatible with standard 8-bit vision algorithms, while automatic gain control (AGC) leads to temporal inconsistencies, non-uniformity correction (NUC) interrupts data streams, and low texture reduces feature detectability. These limitations necessitate tailored adaptations of thermal imagery for robust SLAM.

*This work was partially supported by the Horizon Europe Grant Agreement No. 101136056 and No. 101070405

[1]Hürkan Şahin, Van Huyen Dang, and Erdal Kayacan are with the Automatic Control Group (RAT), Paderborn University, 33098 Paderborn, Germany {huerkan.sahin, van.huyen.dang, erdal.kayacan}@upb.de

To address these challenges, prior work has explored a range of thermal-based SLAM and visual odometry approaches to improve robustness in degraded environments. A feature-based monocular SLAM framework is introduced in [7], integrating thermal image denoising, semantic segmentation, and hybrid point-line tracking to enhance performance in dynamic and visually degraded scenarios. A fully self-supervised approach for estimating depth and ego-motion from monocular thermal video is presented in [8], incorporating temporally consistent mapping to improve contrast and structural information.

Furthermore, a semi-direct visual odometry system combining raw thermal and depth data is proposed in [9], including a dedicated NUC handling module to mitigate sensor interruptions. Learning-based navigation is also explored in [10], where long-wave infrared (LWIR) imagery is used together with the TrajNet model within a model predictive control framework to enable reliable nighttime navigation. Finally, a monocular thermal visual odometry approach for outdoor environments is presented in [11], addressing scale ambiguity through road segmentation and reducing pose estimation failures caused by NUC via a predictive triggering mechanism.

Building on these developments, recent work aims to unify navigation modules into a single end-to-end framework to reduce latency. A common approach separates perception and control, where perception maps visual inputs to waypoints or latent representations [12], which are then used for trajectory tracking [13] or policy learning via reinforcement learning and related methods [14]–[16]. While this modular design facilitates sim-to-real transfer, it may introduce information loss due to representation compression.

Alternatively, end-to-end approaches directly map images to motion commands using deep reinforcement learning (DRL), enabling the generation of motion primitives, collision-free trajectories, or velocity commands [17]–[20]. However, such approaches still struggle in degraded visual conditions due to the limitations of raw image representations.

To address these limitations, this work integrates thermal-to-depth estimation [21] with the VDS-Nav framework [22] and jointly optimizes the perception and control pipeline, with a particular focus on challenging environments such as ship recycling yards, where lighting conditions are highly variable, spaces are confined, and hazardous materials are present. The key contributions of this paper are as follows:

- We propose a unified framework that combines monocular thermal-to-depth estimation with volumetric depth-

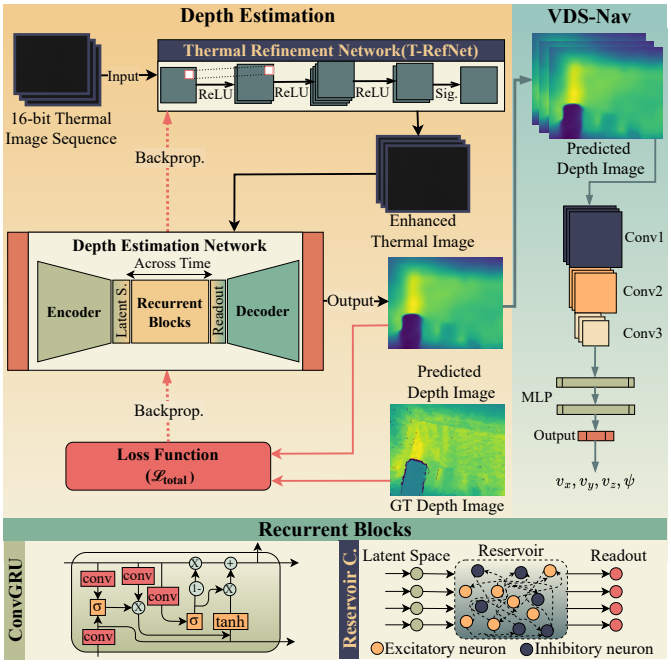


Fig. 1. Overview of the proposed thermal-to-depth navigation pipeline. A sequence of raw 16-bit long-wave infrared (LWIR) images is first processed by the T-RefNet module to enhance structural information and normalize the input. The refined thermal sequence is then fed into a depth estimation network composed of an encoder, recurrent blocks, and a decoder, producing temporally consistent dense depth maps. The predicted depth images are subsequently used as input to the policy network, which consists of a lightweight convolutional feature extractor followed by an MLP to generate control commands.

based safe navigation, aiming to enable direct policy learning from predicted depth sequences without relying on latent space representations, and to facilitate obstacle avoidance in complex and hazardous environments.

- We introduce a depth-driven reward formulation that leverages explicit sensor information and volumetric inputs to strengthen the correlation between observations and actions, improving generalization and enabling effective sim-to-real transfer within the VDS-Nav framework.
- We develop a lightweight thermal-to-depth network tailored for non-radiometric thermal imagery and validate, through simulation and real-world UAV experiments, the effectiveness of both thermal depth estimation and depth-based navigation independently, demonstrating improved depth accuracy and robust navigation performance under low-light and degraded conditions.

The remainder of this paper is organized as follows. Section II provides a brief overview of our methodology. Section III details the experimental setup, dataset, and real-time results. Finally, conclusions are presented in Section IV.

II. Methodology

A unified perception and control pipeline is proposed to enable depth-based navigation from thermal imagery in a mapless setting using model-free DRL. The overall framework is composed of two main stages: thermal-to-depth estimation and depth-based policy learning.

Non-radiometric thermal imagery is known to present significant challenges for depth estimation and SLAM, including low contrast, high dynamic range, and weak structural cues that hinder reliable feature extraction. To mitigate these limitations, a lightweight thermal refinement network, T-RefNet, is introduced to enhance structural visibility and normalize raw 16-bit thermal inputs. The refined thermal sequence is subsequently processed by a depth estimation network consisting of an encoder, recurrent blocks, and a supervised decoder. Temporal consistency is enforced through the recurrent module, allowing stable and geometrically coherent depth predictions to be obtained from thermal-only inputs. In this way, meaningful spatial structure can be recovered despite the inherent limitations of thermal data, and dense depth maps are generated as an intermediate representation for navigation.

Building upon the predicted depth representation, a depth-based navigation policy is learned following the VDS-Nav paradigm. A model-free proximal policy optimization (PPO) [23] algorithm is adopted due to its stability and practical deployment advantages. Instead of compressing observations into a latent space, the policy operates directly on a sequence of predicted depth images, preserving spatial information and reducing representation loss. Each depth image has a resolution of 224×224 and is updated at $50\,\text{Hz}$, while a memory queue is maintained to provide temporal context across consecutive observations.

The policy network is implemented using a lightweight convolutional feature extractor followed by a compact MLP. The feature extractor consists of three convolutional layers with increasing channel dimensions, each followed by ReLU activation, while the MLP comprises two fully connected layers with ELU activations and a linear output layer that predicts control commands.

For training, a depth-driven reward design is employed that relies directly on observation space information rather than privileged states. In contrast to conventional approaches that depend on ground-truth positions or predefined waypoints, the proposed reward formulation utilizes raw depth inputs to strengthen the correlation between state-action pairs and rewards. This formulation can be interpreted from a constrained optimization perspective, where navigation objectives are encoded as reward terms and safety requirements are imposed as constraints. As a result, the agent is guided to move toward open space while avoiding obstacles using only depth perception, improving learning efficiency and reducing the sim-to-real gap.

Overall, the proposed framework tightly integrates thermal perception and control by mapping thermal image sequences to control actions through an intermediate depth representation, eliminating the need for latent space encoding and enabling robust navigation in degraded environments.

[1]FLIR Boson: https://oem.flir.com/products/boson

III. Experiments

In this section, we present experimental results for thermal-to-depth estimation and real-time VDS-Nav performance using depth inputs from a Intel RealSense D455, highlighting robust UAV navigation. The integration of these two components is left for future work.

A. Thermal-to-depth estimation results

To comprehensively evaluate the proposed approach, we conducted experiments on two different thermal–depth datasets: (i) the indoor-dark subset of VIVID++ [25], which is recorded with a radiometric thermal camera, and (ii) a custom dataset collected with a non-radiometric thermal sensor and a depth camera. This dual evaluation setup enables us to assess performance under both radiometric and non-radiometric conditions. The thermal data were captured with a Flir Boson+[1] non-radiometric shuttered camera (640×512). The dataset comprises approximately 65,000 samples, covering diverse lighting conditions—bright, dark, and semi-lit—and including scenes with both hot and cold objects to improve robustness across thermal distributions.

For comparison, we include both RGB-trained depth estimation networks (ZoeDepth [26], DepthAnything-V2 [24]) and thermal-specific approach from the literature [8]. Since ZoeDepth and DepthAnything-V2 were trained on RGB images, we pre-processed our thermal inputs by mapping them to RGB format before inference. Regarding [8], we retrained and evaluated the model on our non-radiometric dataset using the sequences we collected. We use key metrics to quantitatively evaluate our methods against the baselines, such as mean absolute relative error (AbsRel), root mean squared error (RMSE), and accuracy under thresholds $1.25, 1.25^2, 1.25^3$ (a1, a2, a3).

TABLE I

Evaluation results of thermal-to-depth networks on a custom indoor dataset acquired with nonradiometric thermal and depth sensors. Best values are shown in bold.

Model	AbsRel	RMSE	a1	a2	a3
Shin (Max.) [8]	0.262	1.273	0.589	0.890	0.960
ZoeDepth [26]	0.243	1.110	0.605	0.885	0.954
DepthAnything-V2 [24]	0.267	1.043	0.571	0.863	0.931
ResNet8+GRU (ours)	0.109	0.516	0.886	0.943	0.969
MobileNet+GRU (ours)	0.085	0.453	0.911	0.951	0.971
Eff-B0+GRU (ours)	0.079	**0.424**	0.920	0.955	0.971
Eff-B0+RC (ours)	**0.076**	0.439	**0.929**	**0.965**	**0.981**

Eff-B0: EfficientNet-B0 backbone; GRU: ConvGRU;

Table I presents the evaluation on the proposed non-radiometric dataset, which is considerably more challenging due to fluctuating pixel intensities caused by auto-scaling and internal heating. The models trained purely on RGB inputs (ZoeDepth, DepthAnything-V2) perform poorly, with higher AbsRel and lower accuracies. Regarding [8], as illustrated in Fig. 2d, the method employs radiometric-specific preprocessing and performs reasonably on the VIVID++ dataset, but on non-radiometric data its consistency degrades when hot or cold objects enter or leave the scene. In contrast, our models remain robust, with RC combined with EfficientNet-B0 giving slightly better results on the non-radiometric dataset (AbsRel $= 0.076$, $a1 = 0.929$), as also shown in Fig. 2. Although the difference compared to the GRU-based model is not large, it is notable that the RC variant, with

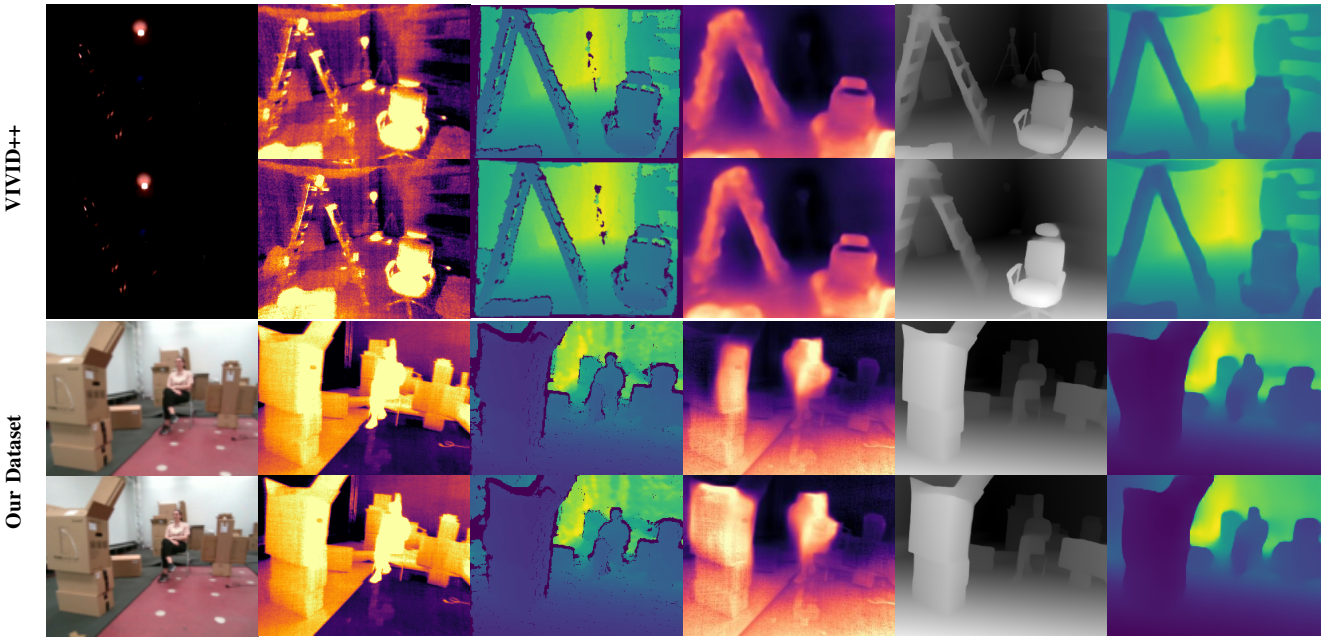
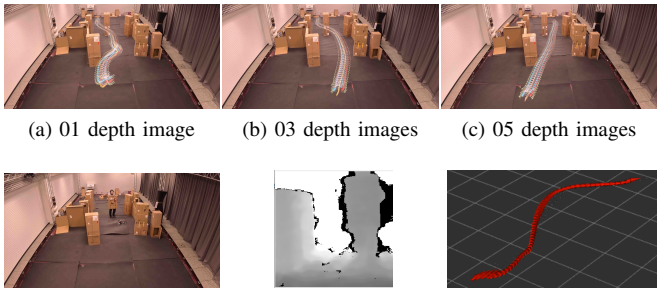


Fig. 2. Qualitative comparison across two datasets. Top: VIVID++; bottom: our dataset. Each row shows two temporally adjacent frames. Columns: (a) RGB, (b) thermal, (c) thermal-aligned ground-truth depth, (d) Shin et al. [8], (e) DepthAnything-V2 [24] (RGB-only), (f) Our representative proposed model with RC.



(a) 01 depth image (b) 03 depth images (c) 05 depth images

(d) Dynamic scene (e) Depth (f) Trajectory

Fig. 3. We established static and dynamic scenes, illustrated in 3a–3c, to evaluate VDS-Nav with 1, 3, and 5 depth images. We challenge the drone by involving a human (see Fig. 3d) where the depth image is shown in Fig. 3e and the complete trajectory is depicted in Fig. 3f.

a lighter architecture, performs better in the more variable non-radiometric conditions. These findings underscore the importance of radiometric invariance and demonstrate that our approach generalizes well across both radiometric and non-radiometric settings.

B. VDS-Nav Sim-to-Real Transfer

Policies are trained in simulation using NVIDIA IsaacLab [27] and the RL-Games framework [28]. A corridor-like environment with randomized obstacle placements is used to promote robustness and generalization. Training is performed with PPO [23] over 2000 episodes using parallel environments, resulting in approximately 2.0×10^6 training steps.

To assess the effectiveness of VDS-Nav, comparisons are made against a state-of-the-art baseline based on depth collision encoder (DCE) [14], where depth observations are compressed into a latent space. This enables a direct comparison between policies learned from raw depth sequences and those relying on a perception-control modularization.

For real-world validation, the trained policies are deployed on a drone equipped with an Intel RealSense D455 and an onboard NVIDIA Orin NX. Control commands predicted by the policy are executed via the PX4 flight stack. Evaluation is conducted in both static and dynamic environments, as illustrated in Fig. 3, by varying the number of depth images (1, 3, and 5) used in the observation space.

The results in Fig. 3 and Table II demonstrate that leveraging a sequence of depth images significantly improves navigation performance. Policies using multiple depth frames achieve higher success rates, shorter traversal times, and smoother trajectories compared to single-frame policies. In particular, the policy with five depth images achieves a 100% success rate in both static and dynamic scenarios, while the single-image policy shows reduced robustness, especially in dynamic environments.

These findings highlight the importance of temporal depth information and confirm the effectiveness of VDS-Nav for sim-to-real transfer in challenging navigation tasks.

TABLE II

EVALUATION OF POLICIES TRAINED WITH OUR PROPOSED VDS-NAV IN REAL TIME WITH VARYING NUMBERS OF DEPTH IMAGES.

# images	Static obstacles				Dynamic obstacles			
	SR [%] ↑	ALT [s] ↓	ARMSJ ↓	MaxJ ↓	SR [%]	ALT [s]	ARMSJ	MaxJ
01	80	10.1	3.75	14.14	60	12.23	3.87	14.68
03	100	8.5	3.67	13.0	-	-	-	-
05	100	7.3	3.3	12.18	100	8.2	3.98	15.83

SR: success rate; ALT: average lap time; MaxJ: maximum jerk.
ARMSJ: average root mean square jerk;

IV. CONCLUSIONS AND FUTURE WORK

This paper presents a unified perspective on thermal-based perception and depth-driven navigation for UAVs operating in GPS-denied and visually degraded environments. We first demonstrate that reliable depth estimation can be achieved from non-radiometric thermal imagery using a lightweight recurrent architecture, maintaining robustness under low texture, noise, and fluctuating thermal conditions. In parallel, we show that the proposed VDS-Nav framework enables end-to-end navigation by directly inferring control commands from depth observations, where volumetric depth inputs improve performance and depth-based reward design enhances safety and sim-to-real transfer.

Although both components are validated independently through simulation and real-world experiments, their integration has not yet been realized. The presented results indicate that combining thermal-to-depth estimation with depth-based policy learning has strong potential for enabling robust navigation in challenging environments such as low-light, confined, and hazardous industrial settings.

Future work will focus on tightly integrating the thermal-to-depth pipeline with VDS-Nav to enable fully end-to-end thermal-driven navigation. This includes evaluating the system in complex real-world scenarios, such as ship recycling environments, where lighting variability, confined spaces, and environmental hazards pose significant challenges. Furthermore, we aim to extend the approach to outdoor applications, including forest exploration and infrastructure inspection, to further assess scalability, robustness, and generalization capabilities.

REFERENCES

[1] Y.-C. Chang, N. Wang, and O. S. Durak, "Ship recycling and marine pollution," *Marine Pollution Bulletin*, vol. 60, no. 9, pp. 1390–1396, 2010. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0025326X10002225

[2] T. Solakivi, T. Kiiski, T. Kuusinen, and L. Ojala, "The european ship recycling regulation and its market implications: Ship-recycling capacity and market potential," *Journal of Cleaner Production*, vol. 294, p. 126235, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652621004558

[3] NGO Shipbreaking Platform, "Impact Report 2018/2019," 2019, [Online]. Available: https://web.archive.org/web/20210303095411/https://shipbreakingplatform.org/wp-content/uploads/2020/06/NGOSBP-Bi-Annual-Report-18-19.pdf. [Accessed: 29-Mar-2026].

[4] M. Lyu, Y. Zhao, C. Huang, and H. Huang, "Unmanned aerial vehicles for search and rescue: A survey," *Remote Sensing*, vol. 15, no. 13, 2023.

[5] Y.-S. Shin and A. Kim, "Sparse depth enhanced direct thermal-infrared SLAM beyond the visible spectrum," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2918–2925, 2019.

[6] J. Jiang, X. Chen, W. Dai, Z. Gao, and Y. Zhang, "Thermal-inertial SLAM for the environments with challenging illumination," *IEEE Robotics and Automation Letters*, vol. 7, no. 4, pp. 8767–8774, 2022.

[7] Y. Wu, L. Wang, L. Zhang, Y. Bai, Y. Cai, S. Wang, and Y. Li, "Improving autonomous detection in dynamic environments with robust monocular thermal SLAM system," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 203, pp. 265–284, 2023.

[8] U. Shin, K. Lee, B.-U. Lee, and I. S. Kweon, "Maximizing self-supervision from thermal image for effective self-supervised learning of depth and ego-motion," *IEEE Robotics and Automation Letters*, vol. 7, no. 3, pp. 7771–7778, 2022.

[9] X. Chen, W. Dai, J. Jiang, B. He, and Y. Zhang, "Thermal-depth odometry in challenging illumination conditions," *IEEE Robotics and Automation Letters*, vol. 8, no. 7, pp. 3988–3995, 2023.

[10] A. NG, D. PB, J. Shalabi, S. Jape, X. Wang, and Z. Jacob, "Thermal Voyager: a comparative study of RGB and thermal cameras for nighttime autonomous navigation," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 14 116–14 122.

[11] P. V. K. Borges and S. Vidas, "Practical infrared visual odometry," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 8, pp. 2205–2213, 2016.

[12] M. Müller, A. Dosovitskiy, B. Ghanem, and V. Koltun, "Driving policy transfer via modularity and abstraction," *arXiv preprint arXiv:1804.09364*, 2018.

[13] A. Loquercio, E. Kaufmann, R. Ranftl, A. Dosovitskiy, V. Koltun, and D. Scaramuzza, "Deep drone racing: From simulation to reality with domain randomization," *IEEE Transactions on Robotics*, vol. 36, no. 1, pp. 1–14, 2019.

[14] M. Kulkarni and K. Alexis, "Reinforcement learning for collision-free flight exploiting deep collision encoding," *arXiv preprint arXiv:2402.03947*, 2024.

[15] J. Fu, Y. Song, Y. Wu, F. Yu, and D. Scaramuzza, "Learning deep sensorimotor policies for vision-based autonomous drone racing," in *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2023, pp. 5243–5250.

[16] J. Xing, L. Bauersfeld, Y. Song, C. Xing, and D. Scaramuzza, "Contrastive learning for enhancing robust scene transfer in vision-based agile flight," in *2024 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2024, pp. 5330–5337.

[17] E. Camci, D. Campolo, and E. Kayacan, "Deep reinforcement learning for motion planning of quadrotors using raw depth images," in *2020 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2020, pp. 1–7.

[18] H. I. Ugurlu, X. H. Pham, and E. Kayacan, "Sim-to-real deep reinforcement learning for safe end-to-end planning of aerial robots," *Robotics*, vol. 11, no. 5, p. 109, 2022.

[19] F. Sadeghi and S. Levine, "Cad2rl: Real single-image flight without a single real image," *arXiv preprint arXiv:1611.04201*, 2016.

[20] I. Geles, L. Bauersfeld, A. Romero, J. Xing, and D. Scaramuzza, "Demonstrating agile flight from pixels without state estimation," *arXiv preprint arXiv:2406.12505*, 2024.

[21] H. Sahin, H. X. Pham, V. H. Dang, A. Yegenoglu, and E. Kayacan, "Thermal image refinement with depth estimation using recurrent networks for monocular orb-slam3," 2026. [Online]. Available: https://arxiv.org/abs/2603.14998

[22] V. H. Dang, A. Redder, H. X. Pham, A. Sarabakha, and E. Kayacan, "Vds-nav: Volumetric depth-based safe navigation for aerial robots–bridging the sim-to-real gap," *IEEE Robotics and Automation Letters*, vol. 10, no. 10, pp. 11 038–11 045, 2025.

[23] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal policy optimization algorithms," *arXiv preprint arXiv:1707.06347*, 2017.

[24] L. Yang, B. Kang, Z. Huang, Z. Zhao, X. Xu, J. Feng, and H. Zhao, "Depth Anything V2," 2024. [Online]. Available: https://arxiv.org/abs/2406.09414

[25] A. J. Lee, Y. Cho, Y. sik Shin, A. Kim, and H. Myung, "ViViD++: Vision for visibility dataset," 2022. [Online]. Available: https://arxiv.org/abs/2204.06183

[26] S. F. Bhat, R. Birkl, D. Wofk, P. Wonka, and M. Müller, "ZoeDepth: zero-shot transfer by combining relative and metric depth," 2023. [Online]. Available: https://arxiv.org/abs/2302.12288

[27] M. Mittal, C. Yu, Q. Yu, J. Liu, N. Rudin, D. Hoeller, J. L. Yuan, R. Singh, Y. Guo, H. Mazhar, A. Mandlekar, B. Babich, G. State, M. Hutter, and A. Garg, "Orbit: A unified simulation framework for interactive robot learning environments," *IEEE Robotics and Automation Letters*, vol. 8, no. 6, pp. 3740–3747, 2023.

[28] D. Makoviichuk and V. Makoviychuk, "rl-games: A high-performance framework for reinforcement learning," https://github.com/Denys88/rl_games, May 2021.