

EREBUS: End-to-end Robust Event Based Underwater Simulation

Hitesh Kyatham^{1*}, Arjun Suresh^{1*}, Aadi Palnitkar¹, Yiannis Aloimonos¹, Xiaomin Lin^{1,2†}

Abstract—The underwater domain presents a vast array of challenges for roboticists and computer vision researchers alike, such as poor lighting conditions and high dynamic range scenes. In these adverse conditions, traditional vision techniques struggle to adapt and lead to suboptimal performance.

Event-based cameras present an attractive solution to this problem, mitigating the issues of traditional cameras by tracking changes in the footage on a frame-by-frame basis. In this paper, we introduce a pipeline which can be used to generate realistic synthetic data of an event-based camera mounted to an AUV (Autonomous Underwater Vehicle) in an underwater environment for training vision models. We demonstrate the effectiveness of our pipeline using the task of rock detection with poor visibility and suspended particulate matter, but the approach can be generalized to other underwater tasks.

Keywords: DVS, Simulation, Underwater Robotics, computer vision, perception pipeline

I. INTRODUCTION

Underwater environments present unique challenges for robotic perception, including low-light conditions, high dynamic range scenes, and turbid water with poor visibility. Traditional frame-based imaging systems often struggle in these conditions due to motion blur, limited temporal resolution, and sensitivity to lighting fluctuations. Event cameras, with their high temporal resolution, low latency, high dynamic range and comparatively lower power consumption [1], offer a promising alternative for robust visual sensing in such scenarios.

Event-based vision is a bio-inspired paradigm where sensors asynchronously record changes in brightness at each pixel, rather than capturing frames at fixed intervals. This results in a stream of events that encode motion and scene dynamics with exceptional temporal fidelity. While event cameras have demonstrated success on terrestrial and aerial robotics platforms [2]–[4], their application in underwater domains remains relatively underexplored. Recent efforts have begun to highlight their potential for underwater perception, particularly in tasks involving high-speed motion or variable lighting, such as marine life monitoring or navigating autonomous underwater vehicles (AUVs) in complex environments [5] [6]. While event cameras offer distinct advantages, we do not propose them as a wholesale replacement for conventional frame-based systems. Instead, we advocate for their complementary

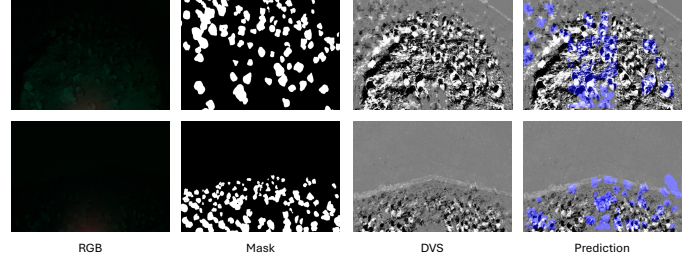


Fig. 1. Illustration of EREBUS frame-by-frame image processing

integration, especially in scenarios with rapid motion or low-light conditions, where event data can enhance perception capabilities and enable faster, more robust inference. While there is extensive literature and datasets available for vision sensors in the underwater domain, the event based paradigm remains under served in this regard. To address both this and the previously mentioned points, we make the following novel contributions.

- **Simulation pipeline for underwater event vision:** We present a fully functional pipeline that utilizes the physics based underwater simulation capabilities of Blender [7] along with event stream generation and downstream object detection.
- **Demonstration of event cameras potential for underwater environments:** We demonstrate, through simulation and detection performance, that event-based sensing is highly promising for challenging underwater conditions.
- **Few-shot segmentation from simulated event data:** We show that realistic, physics-based simulations can be used to train a YOLO model with only a few annotated samples highlighting the potential of synthetic event data for bootstrapping robust perception models with minimal labeling effort.
- **Extensible simulation framework for marine robotics:** Our Blender-based simulation setup is easily modifiable and scene-agnostic it can easily be adapted to simulate coral reefs, shipwrecks, marine life, or environmental monitoring scenarios. This enables broad applicability for developing and testing event-based models without requiring expensive real-world underwater data collection.

* Equal Contributors

† Corresponding Author

¹Maryland Robotics Center, University of Maryland, College Park, MD 20742, USA {hkyatham, apalnitk, arjsur, jyaloimono}@umd.edu

²Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD 21218, USA {xlin52}@jhu.edu

- **Foundational groundwork for a public dataset and benchmark:** We aim to extend this pipeline into an open-source tool and potentially release a benchmark dataset of underwater event streams supporting future research in event-based underwater robotics.

II. RELATED WORK

A. Underwater Simulations

Owing to the logistical difficulties encountered in obtaining data in the underwater domain, simulations have proven to be a reliable method to generate convincing training data for vision models and a testing ground for underwater robotics algorithms [8]–[10]. Similar simulations have also been built around specific underwater tasks, such as oyster detection and counting on the seabed [11]–[13]. These simulations are engineered to address certain challenges inherent to the underwater domain, which are expanded on below.

B. Challenges of the Underwater Domain

The underwater domain is a common avenue for robotics applications, with AUVs being used for structural health monitoring, mine detection and oceanic exploration [14] [15]. However, deployed AUVs commonly face the issue of poor visibility and high occlusion. Several efforts have been undertaken to circumvent these common roadblocks. [16] and [17] utilize light attenuation models to estimate the visibility properties of the occluded image, while [18] leverages the light polarization information contained within the Stokes vector as the foundation to create a polarimetric imaging model. However, underwater lighting has been shown to adversely affect marine life by several studies [19]–[21], thus to ensure the ecological soundness of AUV applications underwater, alternative strategies must be considered.

C. Event Based Cameras

While these works present viable avenues for mitigating the issues faced underwater, event based cameras also present a promising alternative. Unlike conventional frame-based cameras that record entire images at fixed intervals, event-based cameras operate asynchronously at the pixel level recording events only when a change in brightness is detected. This affords them several advantages; namely their superior temporal resolution, high dynamic range and low power consumption [1]. Extensive prior work has been done in utilizing their properties for robotics, such as for live obstacle detection and avoidance for quadrotors [22] and for live reconstruction of 3-D scenery during high speed egomotion [23]. Additionally, their applications have been studied in self-driving cars [24], and to quadrupedal robots [25]. Despite the volume of literature pertaining to their applications, event based cameras remain underutilized in the underwater domain.

D. Applications of Event Cameras to Underwater Domain

Recent research has explored event-based vision for various underwater applications. For AUV docking, event cameras paired with spiking neural networks have improved target

detection and docking precision in low-visibility settings [5], [26]. Event data has also been fused with RGB imagery to enhance underwater images by reducing haze, uneven lighting, and color distortion [27]. Additionally, event cameras have enabled high-resolution flow visualization by tracking laser-illuminated particles, aiding fluid dynamics analysis in environmental and medical studies [28].

In addition, collision avoidance systems [29] [30] for Autonomous Underwater Vehicles (UUVs) are beginning to incorporate event-based sensors. By integrating visual event data with auditory cues, these systems can perform real-time obstacle detection and improve the autonomy and safety of underwater operations in complex or obstacle-rich settings.

This does not address the inherent difficulties in acquiring real underwater data; simulations can mitigate this issue by providing an avenue for the generation of convincing event based imagery.

E. Event Based Simulations

Simulation pipelines dedicated towards event-based imagery are varied in approach. Earlier works have primarily focused on the logarithmic intensity differences between successive frames [31] [32], but this approach struggles with fast moving scenes. More recent techniques like EventGAN [33] leverage a generative adversarial network to generate convincing synthetic event data. Additionally, full 3D simulations of visual scenes are used, such as ESIM [34], an OpenGL based simulation that takes in a scene with a moving camera, rendering the changes in brightness as the camera moves along its trajectory. [35] leverages the DAVIS (Dynamic and Active-pixel Vision Sensor) in a wide domain of synthetic and real environments for visual odometry and SLAM.

These simulators have laid the groundwork for event data generation but they are not specific to the underwater domain. In contrast, our work introduces a dedicated simulation and perception pipeline tailored for event-based vision in marine environments.

III. METHODOLOGY

In this work, we present a pipeline for simulating event-based data in underwater environments and using it for downstream perception tasks, such as object segmentation. Our methodology is designed to replicate realistic underwater scenes using Blender, generate synthetic event streams via an event camera simulator, and train lightweight neural models for effective object detection with limited supervision. The following subsections detail each stage of our pipeline.

A. Underwater Scene Simulation in Blender

We first construct a synthetic underwater environment using Blender. The scene consisted of a random distribution of rocks on the seabed, with thousands of suspended particles, to simulate a relatively occluded shallow-water domain. These particles introduce natural visual noise, mimicking conditions in which suspended solids or plankton may affect visibility and contrast in marine settings. To simulate realistic motion, a

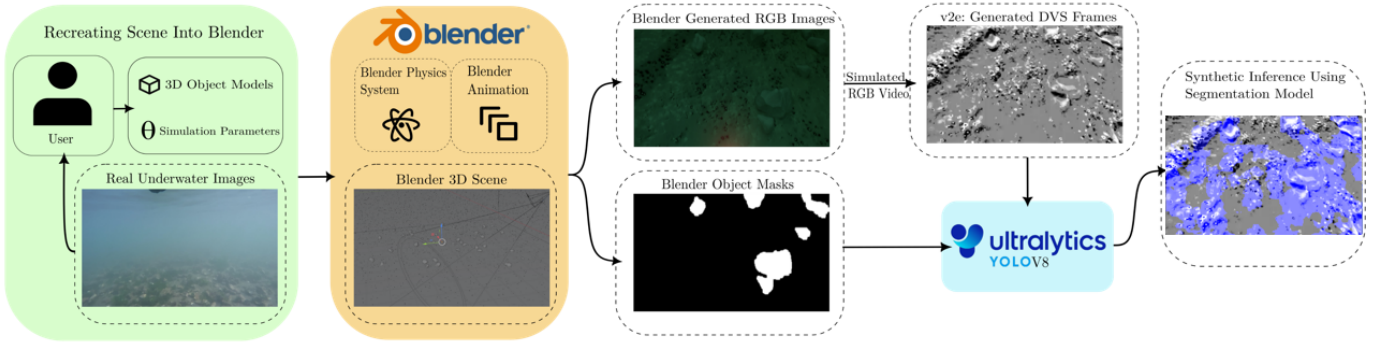


Fig. 2. A depiction of the pipeline of EREBUS. The Blender generated RGB image was brightened for demonstration purposes.

virtual camera is animated to traverse the scene, with the intent of emulating the movement of an autonomous underwater vehicle (AUV) navigating through the environment. The scene features limited visibility, with a spotlight mounted to the camera being the only source of illumination in the scene. The Blender engine is used to render a sequence of RGB images in 1920x1080 resolution, along with the corresponding segmentation masks of the rocks on the seafloor, capturing the temporal evolution of the scene as observed by the moving camera.

B. Event Stream Generation

The rendered image sequence is then processed using an event camera simulator. This simulator models the behavior of a Dynamic Vision Sensor (DVS), converting intensity images into asynchronous event streams. Each event is triggered based on a change in pixel-level brightness, producing a time-ordered sequence of (x, y, polarity, timestamp) events. This stream is visualized as a DVS video, representing the kind of output a real event camera would generate under similar conditions [36].

C. Object Segmentation using YOLO with Few-Shot Learning

To demonstrate the utility of the simulated event data, we train a custom object detection and segmentation model using the YOLO (You Only Look Once) architecture [37]. Specifically we utilize the YOLOv8-n, the smallest available model by weight count, as it is the most suited to be mounted on an edge device, where limited storage capacity is a concern. The DVS video frames are used as input to the model, and a small annotated dataset is sufficient to train YOLOv8 to reliably detect rock formations in the scene. This few-shot learning approach leverages the spatiotemporal richness of event data and highlights its efficiency in low-data regimes as shown in Fig. 1. We obtained a maximum mAP (Mean Average Precision) of 0.83 on a training set of 10 images over a similarly sized validation set for 250 epochs. Despite the small size of the training set, we obtained impressive results for the live inference over the synthetic DVS footage, sporting a large number of detections with high confidence intervals.

D. Enhancing Robustness through Simulated Particle Noise and Adaptive Training

Building upon the object detection capabilities demonstrated in the previous subsection, we further explore the robustness of our YOLO-based segmentation model under conditions that more closely resemble real-world underwater environments. A major challenge in such settings is particle noise caused by suspended sediments, bubbles, or other particulate matter, which can significantly distort visual inputs and reduce detection reliability.

To emulate these conditions, we introduce varying levels of particle noise in our simulation by systematically increasing both the size and the number of particles. Our observations show that when the particle size exceeds twice the baseline configuration, the resulting DVS output becomes heavily affected. This degrades object detection performance due to the increase in background clutter and occlusion effects in the event stream as shown in Fig. 3

Crucially, because the simulation environment allows fine-grained control over such disturbances, we are able to incorporate particle noise into the training process itself. By exposing the model to noisy conditions during training, we enable it to learn noise-invariant features, improving its ability to detect target objects (e.g., rock formations) despite visual interference. This customized training strategy enhances the model's robustness and further demonstrates the advantage of combining synthetic DVS data with few-shot learning.

E. Proposed Pipeline Summary

We propose this simulation-to-perception pipeline as a foundational framework for developing and evaluating event-based vision algorithms in underwater robotics. While we focused on rock counting for this iteration of the simulation, the approach is highly flexible and could be capable of supporting a wide range of marine tasks. Our preliminary results indicate that even with minimal supervision, the pipeline enables effective segmentation from event-based inputs validating its potential for real-world use.

IV. CONCLUSIONS AND FUTURE WORK

In this work, we propose a simulation framework to generate synthetic event data for marine scenes. The simulation is

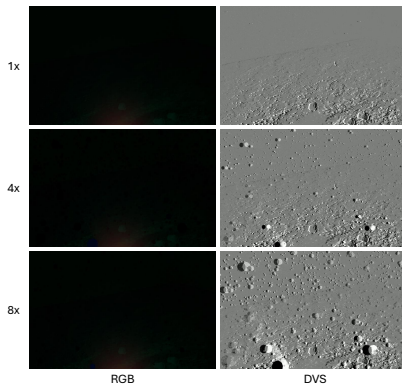


Fig. 3. Illustration showing particle sizes ranging from the default to 8 larger. Best viewed at 400% zoom. Note: the background is not completely black.

intended to replicate the scenario in which an AUV navigates a realistic underwater scene. We present preliminary results from our simulations, showcasing event stream outputs in representative marine scenarios. These results demonstrate the viability of event-based vision for underwater robotics, even without extensive noise modeling. Our goal is to develop high-fidelity event stream simulations tailored for underwater environments, serving as a foundation for training and benchmarking event-based perception algorithms in an underwater domain. By contributing to the emerging intersection of event-based vision and marine robotics, we aim to enable more robust, efficient, and adaptable underwater perception systems. We also envision this work paving the way for real-world deployments, closing the gap between simulation and field-ready marine sensing technologies.

A promising future direction lies in fusing event-based and traditional frame-based data streams. Such hybrid systems could capitalize on the strengths of both modalities, thus leveraging the structural richness of RGB frames and the temporal precision of event streams to enhance detection reliability and computational efficiency under challenging underwater conditions.

Looking ahead, we aim to broaden the applicability of the sim, by simulating a more diverse assortment of environments which are realistically encountered by underwater vehicles. Future iterations of our pipeline can incorporate physics-based noise modeling and more varied sources of water occlusion to further improve realism. Other tasks such as species monitoring could also potentially be addressed with a modified version of our simulation. Our future work consists of a comprehensive synthetic event based dataset which can be open-sourced and made available to all, as a means to improve the performance of mounted vision models on autonomous underwater vehicles in a wide range of commonly encountered domains. Extensive testing of the trained vision models will follow, on event-based data in diverse underwater domains.

REFERENCES

- [1] G. Gallego, T. Delbrück, G. Orchard, C. Bartolozzi, B. Taba, A. Censi, S. Leutenegger, A. J. Davison, J. Conradt, K. Daniilidis, and D. Scaramuzza, "Event-based vision: A survey," *CoRR*, vol. abs/1904.08405, 2019. [Online]. Available: <http://arxiv.org/abs/1904.08405>
- [2] S. Zhu, Z. Tang, M. Yang, E. Learned-Miller, and D. Kim, "Event camera-based visual odometry for dynamic motion tracking of a legged robot using adaptive time surface," 2023. [Online]. Available: <https://arxiv.org/abs/2305.08962>
- [3] A. Bhattacharya, M. Cannici, N. Rao, Y. Tao, V. Kumar, N. Matni, and D. Scaramuzza, "Monocular event-based vision for obstacle avoidance with a quadrotor," 2024. [Online]. Available: <https://arxiv.org/abs/2411.03303>
- [4] J. Martinez-de Dios, A. Gmez Eguluz, J. Rodriguez-Gmez, R. Tapia, and A. Ollero, "Towards uas surveillance using event cameras," in *2020 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, 2020, pp. 71–76.
- [5] F. Zhang, Y. Zhong, L. Chen, and Z. Wang, "Event-based circular detection for auv docking based on spiking neural network," *Frontiers in neurorobotics*, vol. 15, p. 815144, 01 2022.
- [6] S. Takatsuka, N. Miyamoto, H. Sato, Y. Morino, Y. Kurita, A. Yabuki, C. Chen, and S. Kawagucci, "Millisecond-scale behaviours of plankton quantified in situ and in vitro using the event-based vision sensor (evs)," *bioRxiv*, 01 2023.
- [7] B. O. Community, *Blender - a 3D modelling and rendering package*, Blender Foundation, Stichting Blender Foundation, Amsterdam, 2018. [Online]. Available: <http://www.blender.org>
- [8] A. Amer, O. Ivarez Tun, H. . Uurlu, J. Le Fevre Sejersen, Y. Brodskiy, and E. Kayacan, "Unav-sim: A visually realistic underwater robotics simulator and synthetic data-generation framework," in *2023 21st International Conference on Advanced Robotics (ICAR)*, 2023, pp. 570–576.
- [9] E. Potokar, S. Ashford, M. Kaess, and J. G. Mangelson, "Holocean: An underwater robotics simulator," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022, pp. 3040–3046.
- [10] J. Song, H. Ma, O. Bagoren, A. V. Sethuraman, Y. Zhang, and K. A. Skinner, "Oceansim: A gpu-accelerated underwater robot perception simulation framework," 2025. [Online]. Available: <https://arxiv.org/abs/2503.01074>
- [11] X. Lin, N. Jha, M. Joshi, N. Karapetyan, Y. Aloimonos, and M. Yu, "Oystersim: Underwater simulation for enhancing oyster reef monitoring," in *OCEANS 2022, Hampton Roads*. IEEE, Oct. 2022, p. 16. [Online]. Available: <http://dx.doi.org/10.1109/OCEANS47191.2022.9977233>
- [12] X. Lin, V. Mange, A. Suresh, B. Neuberger, A. Palnitkar, B. Campbell, A. Williams, K. Baxevasi, J. Mallette, A. Vera, M. Vincze, I. Rekleitis, H. G. Tanner, and Y. Aloimonos, "Odyssey: Oyster detection yielded by sensor systems on edge electronics," 2025. [Online]. Available: <https://arxiv.org/abs/2409.07003>
- [13] C. Wei, Y. Bai, C. Liu, Y. Zhu, C. Wang, and X. Li, "Unsupervised underwater shipwreck detection in side-scan sonar images based on domain-adaptive techniques," *Scientific Reports*, vol. 14, no. 1, p. 12687, 2024.
- [14] A. Wibisono, M. Piran, H.-K. Song, and B. M. Lee, "A survey on unmanned underwater vehicles: Challenges, enabling technologies, and future research directions," *Sensors*, vol. 23, p. 7321, 08 2023.
- [15] J. F. Waldner and A. Sadhu, "A systematic literature review of unmanned underwater vehicle-based structural health monitoring technologies," *Journal of Infrastructure Intelligence and Resilience*, vol. 3, no. 4, p. 100112, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2772991524000318>
- [16] M. Roser, M. Dunbabin, and A. Geiger, "Simultaneous underwater visibility assessment, enhancement and improved stereo," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 3840–3847.
- [17] J. Zhou, D. Liu, D. Zhang, and W. Zhang, "Light attenuation and color fluctuation for underwater image restoration," in *Asian Conference on Computer Vision*. Springer, 2022, pp. 374–389.
- [18] Y. Wei, P. Han, F. Liu, and X. Shao, "Enhancement of underwater vision by fully exploiting the polarization information from the stokes vector," *Optics Express*, vol. 29, pp. 22 275–22 287, 06 2021.

- [19] D. L. Stanton and J. R. Cowart, "The effects of artificial light at night (alan) on the circadian biology of marine animals," *Frontiers in Marine Science*, vol. 11, p. 1372889, 2024.
- [20] M. Geoffroy, T. Langbehn, P. Priou, Ø. Varpe, G. Johnsen, A. Le Bris, J. A. Fisher, M. Daase, D. McKee, J. Cohen *et al.*, "Pelagic organisms avoid white, blue, and red artificial light from scientific instruments," *Scientific Reports*, vol. 11, no. 1, p. 14941, 2021.
- [21] R. Kochevar, "Effects of artificial light on deep sea organisms: recommendations for ongoing use of artificial lights on deep sea submersibles," Technical Report to the Monterey Bay National Marine Sanctuary Research, Tech. Rep., 1998.
- [22] N. J. Sanket, C. M. Parameshwara, C. D. Singh, A. V. Kuruttukulam, C. Fermüller, D. Scaramuzza, and Y. Aloimonos, "Evdodgenet: Deep dynamic obstacle dodging with event cameras," 2020. [Online]. Available: <https://arxiv.org/abs/1906.02919>
- [23] T. Xiong, J. Wu, B. He, C. Fermüller, Y. Aloimonos, H. Huang, and C. A. Metzler, "Event3dgs: Event-based 3d gaussian splatting for high-speed robot egomotion," 2024. [Online]. Available: <https://arxiv.org/abs/2406.02972>
- [24] A. I. Maqueda, A. Loquercio, G. Gallego, N. Garca, and D. Scaramuzza, "Event-based vision meets deep learning on steering prediction for self-driving cars," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 5419–5427.
- [25] B. Forrai, T. Miki, D. Gehrig, M. Hutter, and D. Scaramuzza, "Event-based agile object catching with a quadrupedal robot," in *2023 IEEE International Conference on Robotics and Automation (ICRA)*, 2023, pp. 12 177–12 183.
- [26] W. Maass, "Networks of spiking neurons: the third generation of neural network models," *Neural networks*, vol. 10, no. 9, pp. 1659–1671, 1997.
- [27] X. Bi, P. Wang, W. Guo, F. Zha, and L. Sun, "Rgb/event signal fusion framework for multi-degraded underwater image enhancement," *Frontiers in Marine Science*, vol. 11, p. 1366815, 2024.
- [28] C. E. Willert and J. Klinner, "Event-based imaging velocimetry: an assessment of event-based cameras for the measurement of fluid flows," *Experiments in Fluids*, vol. 63, no. 6, p. 101, 2022.
- [29] E. Dalmas, F. Danneville, F. Elbahhar, M. Bocquet, and C. Loyez, "A review of neuromorphic sound source localization and echolocation-based navigation systems," *Electronics*, vol. 13, no. 24, p. 4858, 2024.
- [30] N. K. N. Dadson and C. Barbalata, "Marine event vision: Harnessing event cameras for robust object detection in marine scenarios," in *Proceedings of the Winter Conference on Applications of Computer Vision (WACV) Workshops*, February 2025, pp. 1520–1529.
- [31] Y. Bi and Y. Andreopoulos, "Pix2nvs: Parameterized conversion of pixel-domain video frames to neuromorphic vision streams," in *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 1990–1994.
- [32] J. Kaiser, J. C. Vasquez Tieck, C. Hubschneider, P. Wolf, M. Weber, M. Hoff, A. Friedrich, K. Wojtasik, A. Roennau, R. Kohlhaas, R. Dillmann, and J. M. Zillner, "Towards a framework for end-to-end control of a simulated vehicle with spiking neural networks," in *2016 IEEE International Conference on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAN)*, 2016, pp. 127–134.
- [33] A. Z. Zhu, Z. Wang, K. Khant, and K. Daniilidis, "Eventgan: Leveraging large scale image datasets for event cameras," 2019. [Online]. Available: <https://arxiv.org/abs/1912.01584>
- [34] H. Rebecq, D. Gehrig, and D. Scaramuzza, "Esim: an open event camera simulator," 10 2018.
- [35] E. Mueggler, H. Rebecq, G. Gallego, T. Delbruck, and D. Scaramuzza, "The event-camera dataset and simulator: Event-based data for pose estimation, visual odometry, and slam," *The International Journal of Robotics Research*, vol. 36, no. 2, p. 142149, Feb. 2017. [Online]. Available: <http://dx.doi.org/10.1177/0278364917691115>
- [36] Y. Hu, S.-C. Liu, and T. Delbruck, "v2e: From video frames to realistic dvs events," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 1312–1321.
- [37] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," 2016. [Online]. Available: <https://arxiv.org/abs/1506.02640>