

High-frequency LED-headlight for faster pseudo image generation and object detection with event-based cameras

L.Hänsel¹, T. Bertram²

1: L-LAB, Germany

2: University Dortmund, Germany

1. Abstract

Event-based cameras, also known as Dynamic Vision Sensors (DVSs), are biologically-inspired sensors based on a fundamentally different operating principle than conventional cameras. Instead of continuous image acquisition, they asynchronously capture changes in brightness at the pixel level and output them as events characterized by their position, timestamp, and the sign of the change. Due to their high temporal resolution and dynamic range, these sensors are increasingly important for object detection in machine vision, particularly in autonomous driving. Nevertheless, efficiently using fast and asynchronous event streams for precise and robust object detection with low latency remains a key challenge. Most previous approaches are based on the integration of asynchronously recorded events, to generate so-called pseudo-frames, which are then used for object detection. However, these methods are associated with relatively high latency times, as the accumulation of the events typically covers periods of at least 50 milliseconds. This paper presents an approach to minimize accumulation time using a high-frequency Pulse-Width Modulated (PWM) Light Emitting Diode (LED) headlight system. Modern automotive headlight systems feature high pulse-width modulation frequencies, allowing the projection of light distributions at high frame rates. Through these artificially generated brightness changes invisible to human drivers, the number of registered events within a time interval increases significantly. The contribution at hand describes a prototype combining a PWM LED headlight system with an event-based camera. In a first step data was recorded in static settings for various objects. The analyses show a reduction in the integration time required for the generation of pseudo-frames resulting in an improved perception of the surroundings. This approach opens up new possibilities for faster and more efficient object detection in the context of autonomous driving at nighttime.

Keywords: autonomous driving, event-based camera, object detection, pseudo-image generation, accumulation time reduction, low-latency processing

2. Introduction

Given the initial successes of automated driving vehicles in terms of safety and comfort, research in this field continues at full speed [1]. Efficient and accurate environment detection is one of the key tasks in the development of an automated driving vehicle. Studies have shown that the introduction of automated vehicles can effectively reduce and actively prevent traffic accidents [2]. The performance indicators of environment detection, such as latency, robustness, and efficiency, influence the detection and interpretation capabilities of the automated vehicle. These indicators are primarily determined by the sensors used [3]. Event-based cameras are a promising alternative to the sensor systems already established in this field. These bio-inspired sensors are modeled on the human visual system. [4]. Unlike conventional camera sensors, which operate on a synchronous principle, event cameras work asynchronously. The pixels operate independently of each other and capture the observed scene in terms of changes in brightness over time.

The change in brightness L at time t_k is defined by the logarithm of the photocurrent $L = \log(I_{\text{photo}})$ in the individual pixel k at position x_k, y_k . Thus, Δt_k represents the time that has elapsed since the corresponding pixel last showed a reaction. The mathematical definition can be found in equation (1) [5], [6], [7].

$$\Delta L(x_k, y_k, t_k) = L(x_k, y_k, t_k) - L(x_k, y_k, t_k - \Delta t_k) \quad (1)$$

$$\Delta L(x_k, y_k, t_k) = p_k C$$

The change in brightness can be defined as the temporal contrast threshold C [6]. To prevent the sensor from becoming oversaturated, C must reach a certain threshold value, which represents the reaction limit at which the pixel shows a reaction. If the change does not reach the required threshold, no output is generated for this pixel. The variable p_k stands for the so-called polarity of the threshold value. The pixels of the event camera thus represent the polarity of the change between the time intervals Δt_k . The polarity indicates the sign of the change and can therefore be negative or positive. An increase in brightness in the respective pixel leads to a positive polarity and a decrease to a negative polarity. They are also referred to as ON or OFF events and have the definition range $\{+1, -1\}$. The corresponding threshold values C^+ and C^- can be adjusted individually. Consequently, the output does not provide individual images, but rather a sequence of quasi-continuously recorded brightness change events. These events can be described mathematically by a so-called quadruple and follow the scheme: $e[x_k, y_k, p_k, t_k]$. In addition to the polarity already mentioned, the event quadruple contains the position in x_k, y_k coordinates of the sensor pixel matrix and the time t_k , which indicates when the change took place. The mathematical definition is shown in equation (2).

$$e = e[x_k, y_k, p_k, t_k], \text{ with } p_k = \begin{cases} +1, & \text{if } C \geq C^+ \\ -1, & \text{if } C \leq C^- \end{cases} \quad (2)$$

Due to the lack of output when the reaction threshold is not reached, the event camera has a significant disadvantage in terms of the cognition of static or lowdynamic scenes. Due to the lack of dynamics, few or no changes in brightness are registered. As a result, the sensor is no longer able to fully capture the scene in its entirety [5], [6], [7]. In contrast, a conventional camera always captures the static scene in its entirety. However, this advantage is partially offset by the large amount of redundant image data. When there is a very high level of dynamics within the scene, the conventional sensor is limited by its exposure time and the resulting motion blur [8]. In this respect the event camera has a clear advantage due to its biologically-inspired concept and can capture the scene in its entirety in all its complexity. To clearly demonstrate the output of the different sensor concepts between conventional cameras and event cameras, a comparison is illustrated in Figure 1. This intuitive example illustrates the incoherent outputs of the different sensor concepts. While the conventional camera sensor captures the dot at different angles at specific points in time, the event camera only outputs the brightness changes in the scene. In this regard the data stream consists of the sequence of brightness changes or the event quadruples caused within the pixels by the rotating black dot.

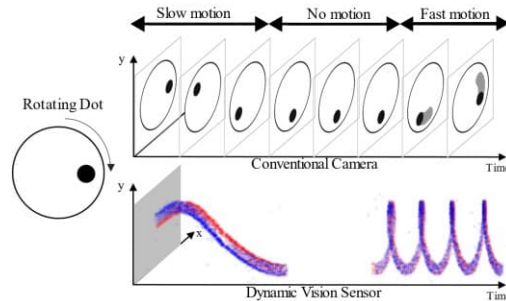


Figure 1: Comparison of the output of a conventional camera sensor and an event-based camera (red represents positive polarity, blue represents negative polarity) [9].

In addition to the advantages already mentioned, the benefits of the biologically-inspired sensor are summarized below. The concept of asynchronous operation of each individual pixel allows dynamic values of up to 140 dB to be achieved. This far exceeds the 60-80 dB achieved by conventional camera sensors. A high dynamic range is particularly advantageous in scenarios with high brightness differences, such as when driving through a tunnel. Conventional camera sensors are often unable to adequately display such contrasts, resulting in overexposed or underexposed areas in the image [8]. The event-based camera on the other hand can capture both very bright and very dark areas at the same time, resulting

in a more accurate cognition of the environment. The reduced amount of data from the event camera compared to conventional image data, combined with the asynchronous transmission method, results in a temporal resolution in the microsecond range. This makes it possible to achieve a measurement rate of up to 10,000 samples per second. This high temporal resolution makes it possible to precisely capture and analyze fast movements and react appropriately in time-critical situations. Due to the high sampling rate and the absence of a conventional exposure process, there is no blurring even at high dynamics. This leads to a clearer perception of the environment especially in agile situations with little ambient light. In addition, motion blur is a common problem with conventional cameras, especially during fast movements or in poor lighting conditions. The event-based camera eliminates this problem by continuously detecting changes in light intensity and transmitting them immediately, rather than relying on periodic exposure. As a result of the event-based camera's asynchronous operating concept redundant data transfer is reduced to a minimum. At the same time this, in combination with the reduced data volume, ensures significantly reduced energy consumption of just a few milliwatts [5].

Nevertheless, the asynchronous operating principle also poses a major challenge in terms of evaluation using conventional, image-based features. These features serve as the basis for object detection in neural networks. In order to generate these from the asynchronous event data stream, there are a number of approaches and different forms of representation, which are then fed into the detection network [10]. With regard to the generation of so-called pseudo-frames, the asynchronous brightness changes are accumulated in the form of events over a certain period of time. The resulting pseudo-frames are also referred to as event frames and consist of a so-called two-dimensional event histogram with two channels each for OFF and ON events. The procedure for creating such a two-dimensional event frame for different accumulation times Δt is shown in the Figure 2 below. This shows that longer accumulation time, the resulting event frame contains more information. This results in a conflict of interest: on the one hand, the event histogram should contain as much information as possible, but on the other hand, detection latency is extremely important in the context of automated driving. In addition, the high temporal resolution of the event-based camera is almost completely discarded. An accumulation time of 50 ms has become established in the literature and has since been considered state of the art [10], [11], [12], [13].

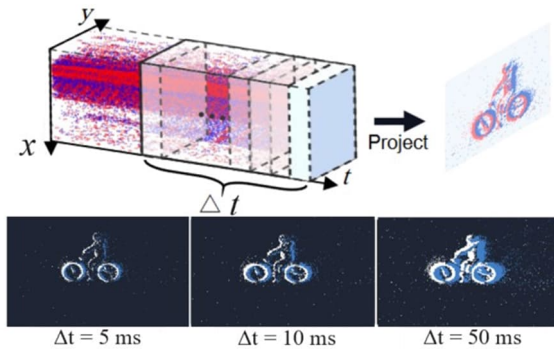


Figure 2: Procedure for pseudo-frame generation for different accumulation times. The color scheme for pseudo-frames has been changed from redblue to white-blue to improve contrast.

To compensate for this disadvantage and reduce the accumulation time, this paper uses a PWM-LED headlight system. The high-frequency intensity changes on the object lead to a corresponding increase in event density. This makes it possible to reduce the accumulation time for generating the pseudo-frames and thus maintain the speed advantage of the event-based camera.

3. Methods

A partially static setup is used to investigate the reduction in the accumulation time of event frames in connection with a PWM-LED headlight system. The measurement project is carried out at the light tunnel parking lot of FORVIA HELLA in Lippstadt. In order to ensure the comparability of the measurements, an advanced driving assistance dummy system from 4D Active is also used. Here, the pedestrian or cyclist is moved along a fixed trajectory with the aid of an electric motor and a belt system [14]. The object is located 30 m away from the vehicle and crosses the road at a speed of 5 km/h. In addition, the distance traveled between the starting point and the end point is 11 m. For a better understanding, the measurement setup is shown in Figure 3 below.

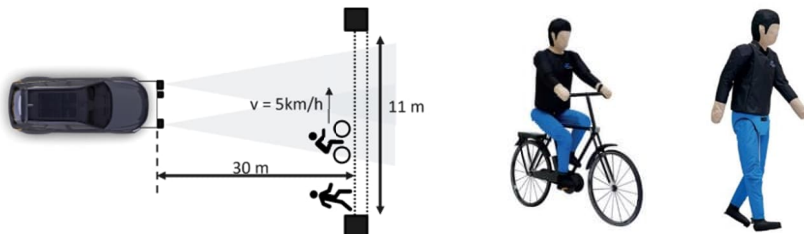


Figure 3: Measurement setup.

The vehicle is a test carrier equipped with a modified PWM- LED headlight system. The selected light distributions are Economic Commission for Europe (ECE)certified low beam and high beam distributions. The PWM frequency is fixed at a value of 500 Hz. A housing has also been designed for the sensor technology, which contains a conventional Red-Green-Blue (RGB) camera from Baumer “VCXU-23C” and a Prophesee “EVK4” event-based camera. The RGB camera is used to annotate the event data. In addition, a “TI 1843 Boost” automotive radar is installed, which functions as a reference sensor for distance and speed. The sensor system is located near the left headlight housing in order to also examine the resulting viewing angle. Figure 4 shows the test vehicle and the housing with the sensor system used.

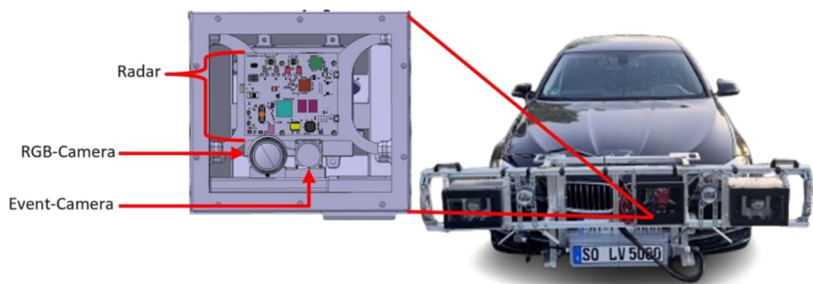


Figure 4: Test vehicle and sensor housing.

An integrated microcontroller with a trigger frequency of 50 Hz is used to synchronize the various sensor modalities with each other on a temporal level. In addition, the sensors are intrinsically and extrinsically calibrated to each other. This ensures that event data can be annotated based on the RGB camera. In this regard, the ground truth values in the form of bounding boxes based on the RGB camera are transferred to the event-based camera using a reprojection approach. To do this, the pixel positions of the bounding box on the RGB images are projected onto the corresponding pixel position of the event-based camera using the calculated reprojection matrix. When reducing the accumulation time to values below 20 ms, the bounding boxes are calculated using linear interpolation based on the work of Lu et al. [12]. Due to the small baseline of 4.5 cm between the two cameras, infinity depth alignment is assumed, based on the work of Gehrig et al. [15], [16]. Data acquisition and control of the microcontroller are carried out using a laptop with a Intel Core i5 13500H CPU, 16 GB of RAM, 8 TB Corsair MP600 PRO XT hard drive and RTX 4060 GPU. In addition, the electric motor for moving the dummies can be operated by an external device. Basically, a measurement consists of a total of one run, in which the dummy moves from the start to the end positions. With regard to the different measurements, the object (pedestrian, cyclist) and the selected ECE-compliant light distribution (low beam, high beam) are varied. For the final evaluation, the recorded event data is transferred to the

ISAL 2025 – Proceedings 6

previously twodimensional event histograms. Different accumulation times of 5 ms, 10 ms, and 50 ms are selected in order to perform a comparison based on the average precision (AP). This metric is a widely used performance indicator for information retrieval and object recognition in machine learning. It quantifies the balance between precision (proportion of correct detections) and recall (proportion of detected objects) across the corresponding class. For a deeper understanding of the metric, please refer to Solawetz [17]. A modified YOLOv8 detection network from the work of Verma et al. is used for object detection [18]. The network was previously trained on the dynamic *event camera dataset for driving scenarios* by Gehrig et al. to create a common starting point [15]. The accumulation time is constant 50 ms. The annotations were limited to the pedestrian and cyclist classes for the training process. Furthermore, the recording vehicle uses a conventional halogen headlight without modulation frequency.

4. Results

The results can be divided into a total of four different scenarios, which are presented in Table 1.

Table 1: Average precision (AP) of the four different scenarios with different accumulation times.

Accumulation Time	Pedestrian (Low Beam)	Pedestrian (High Beam)	Cyclist (Low Beam)	Cyclist (High Beam)
5 ms	0.68	0.74	0.72	0.77
10 ms	0.89	0.90	0.91	0.93
50 ms	0.88	0.89	0.91	0.92

Reducing the accumulation time from 50 ms to 10 ms results in an improvement in AP of 0.075 on average across all scenarios. This can be explained by the higher event rate of the objects due to the PWM signal of the headlight. This generates more events in a shorter period of time. This also leads to a more pronounced tail effect, which occurs primarily in darker scenes. An example of the tail effect is shown in Figure 5 below. With long accumulation times, this manifests itself as a kind of veil of negative events and can be interpreted as a kind of motion blur effect from event-based cameras.

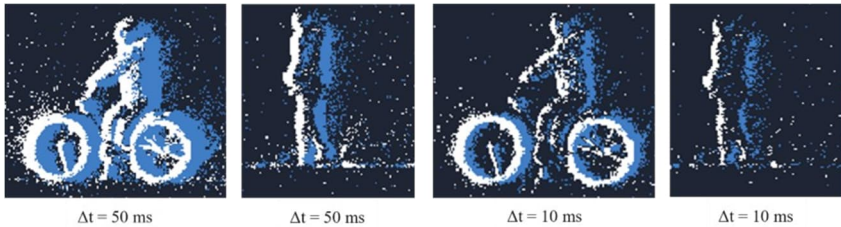


Figure 5: The tail effect, visible as the blue veil behind the object, can be reduced by reducing the accumulation time Δt . Reducing Δt also results in a sparser representation of the two-dimensional event histogram respectively pseudo-frame.

This behavior can be attributed to the integration time of the photodiode in the event-based camera pixel, i.e., the time required to detect a change in contrast. In bright lighting conditions, the integration time is shorter, allowing events to be detected more quickly, while in darker lighting conditions it is longer, resulting in slower responsiveness. This delayed response time occurs exclusively with OFF events and is related to the pixel design at the hardware level [19]. It is also noticeable that the AP is generally lower for pedestrians than for cyclists. This can be explained by the more distinctive features of cyclist, which are mainly caused by the retro-reflectors on the bicycle. Here, the PWM signal from the headlight system is optimally reflected, leading to a significantly higher number of events in this area. The number of events is on average 19.8 % higher under the same boundary conditions of the scene and accumulation times. This is also evident when viewing the scene with high beam and low beam. When illuminating the object with high beam, significantly more events are caused by the object than with low beam. Here, the difference in the respective scenery is 9.9 % on average in the measurement with the cyclist and 5.8 % on average in relation to the pedestrian. This inevitably leads to a denser two-dimensional histogram and thus to more features within the pseudo-image in a shorter accumulation time. The higher AP can also be derived from this. When the accumulation time is reduced to 5 ms, the AP drops. This can be attributed to the significantly sparser representation, which does not contain enough features to enable comparable robust detection in regard to the training data.

In addition, performance continues to decline in scenes where low beam distribution is used. This can in turn be explained by the drop in the event rate for low beam distribution, which further supports the sparseness.

5. Conclusion and Outlook

The contribution at hand presents a system consisting of an event camera and a pair of high-frequency PWM-LED headlights. It investigates the extent to which the high-frequency intensity changes of the headlight can be used to reduce the pseudoimage generation time for object detection by event-based cameras. The focus here is on reducing the so-called

accumulation time. This refers to the accumulation of detected asynchronous brightness changes into pseudo-frames. In order to achieve an adequate comparison from a practical point of view, a dummy in form of a pedestrian and a cyclist is used, which is moved along a fixed trajectory with the aid of an electric motor. The vehicle is located 30 m away in a static state. The scene is varied in terms of the object and the light distribution of high beam and low beam. The AP is used to compare detection performance, and the number of events is also related to detection performance. In principle, a higher AP can be achieved with an accumulation time reduced by a factor of five. This can be attributed to the so-called tail effect, the influence of which can be realized primarily by reducing the accumulation time. In addition, it can be seen that the retroreflectors on the bicycle lead to a higher event density by reflecting the PWM-LED headlight system. These distinctive features also explain differences in the detection performance of the objects. With a further reduction of the accumulation time by a factor of ten, the AP drops significantly, which can be explained by the low information content and the associated sparseness of the pseudo frames. The same applies to the selected low beam or high beam measurement. In this regard, the event rate increases by 9.9 % or 5.8 % depending on the object when illuminated with high beams. This in turn leads to higher detection performance.

Basically, the system in the configuration used shows increased detection performance while reducing the accumulation time by a factor of five.

Nevertheless, further investigations are needed to determine the extent to which the system functions dynamically and whether PWM offers any advantages at all in a full dynamic environment. At the same time the noise behavior in a dynamic environment should be analyzed. In addition, the parameterization of the event camera with regard to the contrast threshold for positive and negative events is limited to the default values. Optimized parameterization could remedy the tail effect. Prior filtering would also be conceivable in this regard. Another point would be the choice of an optimized representation form of the pseudo-frame, which takes into account the temporal component in an additional channel of the histogram. Furthermore, a concrete analysis of the PWM frequency of the headlight system would be desirable in order to be able to optimize it depending on the detection performance.

6. References

- [1] S. Zoting, "Autonomous Vehicle Market Size to Hit USD 2,752.80 BN by 2033." Accessed: Aug. 21, 2024. [Online]. Available: <https://www.precedenceresearch.com/autonomous-vehicle-market>
- [2] M. Kühn and L. Hannawald, "Driver Assistance and Road Safety," in *Handbook of Driver Assistance Systems: Basic Information*,

- Components and Systems for Active Safety and Comfort*, H. Winner, S. Hakuli, F. Lotz, and C. Singer, Eds., Cham: Springer International Publishing, 2016, pp. 69–90. doi: 10.1007/978-3-319-12352-3_4.
- [3] D. J. Yeong, G. Velasco-Hernandez, J. Barry, and J. Walsh, “Sensor and Sensor Fusion Technology in Autonomous Vehicles: A Review,” *Sensors*, vol. 21, no. 6, p. 2140, Mar. 2021, doi: 10.3390/s21062140.
- [4] G. Chen, H. Cao, J. Conradt, H. Tang, F. Rohrbein, and A. Knoll, “Event-Based Neuromorphic Vision for Autonomous Driving: A Paradigm Shift for Bio-Inspired Visual Sensing and Perception,” *IEEE Signal Process. Mag.*, vol. 37, no. 4, pp. 34–49, Jul. 2020, doi: 10.1109/MSP.2020.2985815.
- [5] G. Gallego *et al.*, “Event-Based Vision: A Survey,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 1, pp. 154–180, Jan. 2022, doi: 10.1109/TPAMI.2020.3008413.
- [6] P. Lichtsteiner, C. Posch, and T. Delbruck, “A 128×128 120 dB 15 μs Latency Asynchronous Temporal Contrast Vision Sensor,” *IEEE J. Solid-State Circuits*, vol. 43, no. 2, pp. 566–576, Feb. 2008, doi: 10.1109/JSSC.2007.914337.
- [7] G. Gallego, J. E. A. Lund, E. Mueggler, H. Rebecq, T. Delbruck, and D. Scaramuzza, “Event-Based, 6-DOF Camera Tracking from Photometric Depth Maps,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 40, no. 10, pp. 2402–2412, Oct. 2018, doi: 10.1109/TPAMI.2017.2769655.
- [8] G. C. Holst and T. S. Lomheim, *CMOS/CCD sensors and camera systems*, Second edition. Winter Park, FL, Bellingham, Washington, USA: JCD Publishing ; SPIE Press, 2011.
- [9] C. Wang, C. Fang, Y. Zou, J. Yang, and M. Sawan, “SpikeSEE: An Energy-Efficient Dynamic Scenes Processing Framework for Retinal Prostheses,” Sep. 16, 2022, *arXiv*: arXiv:2209.07898. doi: 10.48550/arXiv.2209.07898.
- [10] N. Zubić, D. Gehrig, M. Gehrig, and D. Scaramuzza, “From Chaos Comes Order: Ordering Event Representations for Object Recognition and Detection,” Aug. 30, 2023, *arXiv*: arXiv:2304.13455. doi: 10.48550/arXiv.2304.13455.
- [11] E. Perot, P. de Tournemire, D. Nitti, J. Masci, and A. Sironi, “Learning to Detect Objects with a 1 Megapixel Event Camera,” in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2020, pp. 16639–16652. Accessed: Jul. 05, 2023. [Online]. Available: <https://papers.nips.cc/paper/2020/hash/c213877427b46fa96cff6c39e837ccee-Abstract.html>
- [12] D. Lu, L. Kong, G. H. Lee, C. S. Chane, and W. T. Ooi, “FlexEvent: Event Camera Object Detection at Arbitrary Frequencies,” Dec. 09, 2024, *arXiv*:

- arXiv:2412.06708. Accessed: Feb. 06, 2025. [Online]. Available: <http://arxiv.org/abs/2412.06708>
- [13] B. Zhou and J. Jiang, "Deep Event-based Object Detection in Autonomous Driving: A Survey," May 07, 2024, *arXiv*: arXiv:2405.03995. Accessed: May 15, 2024. [Online]. Available: <http://arxiv.org/abs/2405.03995>
- [14] 4activeSystems, "4activeSB - system to perform AEB/AES-VRU tests," 4activeSystems - Solutions for ADAS and autonomous driving. Accessed: Jun. 07, 2025. [Online]. Available: <https://www.4activesystems.at/4activesb/>
- [15] M. Gehrig, W. Aarents, D. Gehrig, and D. Scaramuzza, "DSEC: A Stereo Event Camera Dataset for Driving Scenarios," Mar. 10, 2021, *arXiv*: arXiv:2103.06011. Accessed: May 31, 2023. [Online]. Available: <http://arxiv.org/abs/2103.06011>
- [16] D. Gehrig and D. Scaramuzza, "Low-latency automotive vision with event cameras," *Nature*, vol. 629, no. 8014, pp. 1034–1040, May 2024, doi: 10.1038/s41586-024-07409-w.
- [17] J. Solawetz, "mAP in Object Detection: Mean Average Precision Explained," Roboflow Blog. Accessed: Jun. 09, 2025. [Online]. Available: <https://blog.roboflow.com/mean-average-precision/>
- [18] A. A. Verma, B. Chakravarthi, A. Vaghela, H. Wei, and Y. Yang, "eTraM: Event-based Traffic Monitoring Dataset," Apr. 02, 2024, *arXiv*: arXiv:2403.19976. Accessed: Apr. 04, 2024. [Online]. Available: <http://arxiv.org/abs/2403.19976>
- [19] Prophesee, "FAQ — Metavision SDK Docs 5.1.1 documentation." Accessed: Jun. 10, 2025. [Online]. Available: <https://docs.prophesee.ai/stable/faq.html?highlight=tail#what-is-the-root-cause-of-tail-effect-or-the-delayed-response-observed-for-negative-polarity-events-in-low-light-conditions-with-imx636-sensor>