
Fully Autonomous Neuromorphic Navigation and Dynamic Obstacle Avoidance

Anonymous Author(s)

Affiliation

Address

email

Abstract

1 Unmanned aerial vehicles could accurately accomplish complex navigation and ob-
2 stacle avoidance tasks under external control. However, enabling unmanned aerial
3 vehicles (UAVs) to rely solely on onboard computation and sensing for real-time
4 navigation and dynamic obstacle avoidance remains a significant challenge due to
5 stringent latency and energy constraints. Inspired by the efficiency of biological
6 systems, we propose a fully neuromorphic framework achieving end-to-end obsta-
7 cle avoidance during navigation with an overall latency of just 2.3 milliseconds.
8 Specifically, our bio-inspired approach enables accurate moving object detection
9 and avoidance without requiring target recognition or trajectory computation. Ad-
10 ditionally, we introduce the first monocular event-based pose correction dataset
11 with over 50,000 paired and labeled event streams. We validate our system on
12 an autonomous quadrotor using only onboard resources, demonstrating reliable
13 navigation and avoidance of diverse obstacles moving at speeds up to 10 m/s under
14 different light conditions, with energy consumption reduced to 21% compared to
15 traditional architecture.

16 1 Introduction

17 The utilization of UAVs across various applications expanded rapidly over the past decade [1].
18 Currently, most UAVs rely heavily on external aids such as positioning systems like Global Positioning
19 System (GPS) [2] for localization and ground stations [3] for navigation and dynamic obstacle
20 avoidance. However, such external aid is not feasible in all circumstances as it could be easily
21 jammed [4] or interfered in multiple scenarios, including dense urban areas [5], caves, or even war
22 zones [6]. Therefore, it is vital for UAVs to fully perform navigation and dynamic obstacle avoidance
23 tasks using only sensors and computing resources onboard, without any dependence on external
24 signals or infrastructure. Although applicable options are well researched, most solutions are designed
25 for larger platforms but not for tiny UAV systems [7]. Ranging sensors like the Li-DAR system could
26 provide accurate positioning information, but are too heavy and power-hungry to be deployed on tiny
27 autonomous systems [8]. Vision-based approach may be an appropriate way for tiny UAVs since,
28 firstly, visual sensors can be both lightweight and power-efficient [9, 10]. Secondly, visual algorithms
29 achieve state-of-the-art performance in multiple tasks. However, such high performance comes at
30 excessive computational and memory costs. Mainstream approaches like simultaneous localization
31 and mapping (SLAM) algorithms [11] and object recognition-based trajectory estimation methods
32 [12, 13] consume hundreds of megabytes to several gigabytes of memory and hundreds of gigaflops
33 [14]. Such high consumption makes tiny UAV autonomy challenging.

34 Neuromorphic hardwares provide a solution to this problem since the asynchronous and sparse nature
35 of their biomimetic data format could exceed the current standard of energy efficiency, computational
36 consumption, and task accuracy, and thus represents a paradigm shift compared to the traditional
37 computer vision approach [15]. Ideally, such data structure could lead to a data processing method
38 with higher processing speed and lower energy consumption, but contemporary methods treat event
39 data similar to a traditional image frame [16, 17] using "event frames" [18, 19], and hence fails to fully
Submitted to 39th Conference on Neural Information Processing Systems (NeurIPS 2025). Do not distribute.

40 leverage the inherent sparsity of event streams [20], resulting in a performance similar to traditional
41 methods. Inspired by the efficiency of biological systems, we observed that during prey capture,
42 frogs can accurately localize fast-moving insects while exhibiting significant disregard for stationary
43 objects. Further anatomical analysis of their visual neural pathways revealed a striking similarity
44 between the frog’s visual imaging mechanism and the working principles of event cameras; hence, by
45 leveraging this similarity, a purely neuromorphic dynamic obstacle avoidance approach—mimicking
46 frog visual neurons becomes feasible. This enables tiny UAVs to react to high-speed dynamic objects
47 with minimal energy consumption and latency, much like frogs do. The resulting reduction in power
48 and computational demands significantly enhances the performance of tiny UAVs in complex tasks,
49 paving the way for their full autonomy.

50 In this paper, we exhibit a fully neuromorphic pipeline. With only one monocular event camera and
51 an inertial measurement unit (IMU), the autonomous UAV could accomplish the navigation task and
52 the dynamic obstacle avoidance task simultaneously, purely with its onboard computing resources
53 without any external aid. In the navigation module, the quadrotor uses IMU data to navigate long
54 distances, and by coupling the visual-homing algorithm and event data, the quadrotor employs an
55 SCNN network to mitigate the effects of error drift. During the training process, despite the abundance
56 of existing datasets for pose correction, event-driven datasets remain conspicuously absent in current
57 research. To the best of our knowledge, we construct the first monocular event-based pose correction
58 dataset with 50,234 paired event streams, each labeled with its ground truth extrinsic obtained by a
59 motion capture system. This dataset can provide sufficient training data for event-based calibration,
60 thereby advancing the development of neuromorphic approaches in navigation and localization, and
61 facilitating a deeper understanding of the discrepancies between event-based modalities and other
62 sensing modalities (e.g., RGB), while laying the foundation for effective multi-modal information
63 fusion. In the obstacle avoidance module, by implementing our bio-inspired algorithm, the quadrotor
64 could suppress the events produced by static objects. With only events generated by dynamic
65 objects preserved, the algorithm bypasses target recognition and trajectory computation steps, directly
66 outputting evasion maneuvers, and reduces the latency of obstacle avoidance to only 2.3 milliseconds.
67 The significantly reduced latency provides UAVs with a longer time window for evasion maneuvers,
68 substantially enhancing their performance when encountering high-speed moving objects. The
69 comparative evaluation with other state-of-the-art dynamic obstacle approaches demonstrates the
70 superior performance of our neuromorphic architecture and bio-inspired algorithm. Additionally, we
71 have validated the effectiveness and robustness of our approach in real-world environments through
72 physical flight experiments under different light conditions, with energy consumption reduced to 21%
73 compared to the traditional structure.

74 In summary, our contributions to the community include:

- 75 • A fully neuromorphic framework enabling tiny UAVs to rely solely on onboard computation
76 and sensing for real-time navigation and dynamic obstacle avoidance.
77 • A bio-inspired approach enabling tiny UAVs to accurately avoid dynamic obstacles at speeds
78 up to 10 m/s with a latency of 2.3 milliseconds.
79 • An open-sourced monocular event-based pose correction dataset with over 50,234 paired
80 and labeled event streams.

81 2 Related Work

82 2.1 Neuromorphic Control of Quadrotors

83 While some studies discuss the topic of neuromorphic control on objects like larger robots and robot
84 arms [21, 22, 23], the neuromorphic control system on quadrotors remains an underexplored area in
85 research. A Viale et al. [24] proposed the first example of a neuromorphic vision-based controller
86 solving a high-speed UAV control task by using a spiking neuronal network with an Intel Loihi chip
87 [25]. Dupeyroux et al. [26] accomplished the task of UAV landing with a 3-layer spiking neuronal
88 network on Loihi, and recently, Paredes-Vallés et al. proposed the first fully neuromorphic vision
89 and control pipeline for controlling an autonomous quadrotor and made the quadrotor successfully
90 take off, fly along a given route, and then land [27]. The study of neuromorphic control of quadrotors
91 is highly restricted by the hardware performance of embedded neuromorphic processing platforms
92 [28, 29] in terms of the number of available neurons and synapses. The Intel Kapoho Bay with 2
93 Loihi chips [25] carries 262,100 neurons [24], and the SpiNNaker(SNN architecture) version [30]

94 has 768,000 neurons. Though higher-neuron neuromorphic platforms expand computational capacity,
95 they remain inadequate for tasks like optical flow estimation (requiring > 3.7M neurons [31]). In this
96 work, we use Speck [32], a neuromorphic SoC (System on Chip) with 327,000 neurons [33] that
97 could support at most 8 layers of SNNs.

98 **2.2 Visual-homing Algorithm**

99 Visual-homing comes from the idea that small insects such as ants and bees can navigate long
100 distances despite their tiny brains. The mechanism behind such behavior can be categorized into two
101 parts: path integration and drift error elimination. Cartwright and Collet [34] first proposed a snapshot
102 model that describes the homing behavior of bees, and researchers in the field of robotics use this
103 concept to develop efficient navigation algorithms for tiny robots [35, 36]. Subsequent researches
104 focus on reducing the memory required for visual-homing, and has been made in two directions. The
105 first is the reduction of the memory consumed by snapshots: Stürzl and Mallot [37] transformed the
106 snapshot into the frequency domain and remembered only the lowest-frequency component, and
107 reduced the size of the snapshot remarkably. The second direction is to increase the spacing between
108 snapshots. Denuelle and Srinivasan [38] proposed a study that uses the homing vector as a position
109 estimate relative to the snapshot, enabling the drone to navigate some distance toward the next
110 snapshot area. Van Dijk et al. [39] combined two directions and successfully deployed visual-homing
111 on a tiny 56-gram autonomous drone with one panoramic camera. For detailed biological concepts,
112 please take a look at the supplementary note 3.

113 **2.3 Frog-eye Receptive Field**

114 In nature, frogs' visual systems exhibit high-fidelity motion detection for fast-moving objects with
115 deliberate suppression of static background stimuli. The observed motion selectivity stems from
116 specialized receptive field organization in the anuran retinotectal system [40, 41]. During the
117 past decades, researchers conducted extensive research on such a mechanism and found that R3
118 ganglion cells respond to stimuli to ON-OFF brightness changes, create motion-sensitive detection
119 zones[42, 43, 44]. In the standard model of such detection zones, ERF (excitatory receptive field)
120 and IRF (inhibitory receptive field) generate symmetrical excitatory and inhibitory responses to
121 ON-OFF stimuli. Extending these findings, Hoshino et al.[45] identified a functional asymmetry in
122 the spatial organization of ERF and IRF. For detailed biological concepts, please take a look at the
123 supplementary note 3.

124 **2.4 Dynamic Obstacle Avoidance**

125 The dynamic obstacle avoidance problem for unmanned aerial vehicles has been widely researched in
126 recent years, but mainly in the aspect of quasi-static environment[46] and low-speed obstacles. Even
127 though existing literature that relies on monocular vision[47, 48, 49, 13], stereo vision[50, 51]and
128 depth camera [12, 52, 53] exhibits satisfactory performance on slow-moving objects like pedestrians
129 [54], their performance dealing with high-speed dynamic obstacles like thrown balls, birds[55] and
130 even other unmanned aerial vehicles cannot meet the requirement of real-time avoidance. However,
131 despite Falanga et al.[20] displaying the concept of using event stream directly, many researchers still
132 treat events in the form of "event frame" [19]. To the best of our knowledge, this is the first work
133 that implements low-latency (2.3 milliseconds) dynamic obstacle avoidance when the quadrotor is
134 executing a navigation task without the help of any external infrastructure.

135 **3 Methods**

136 In this section, we introduce our neuromorphic navigation and dynamic obstacle avoidance pipeline,
137 which includes a neuromorphic control framework that allocates computing resources to minimize
138 evasion latency and maximize navigation correctness, an event-visual-homing based end-to-end
139 method and a bio-inspired dynamic obstacle avoidance algorithm that reduces the latency of obstacle
140 detection to 2.3 milliseconds and applicable for dodging multiple high-speed obstacles when the
141 drone is navigating to its destination.

142 **3.1 Overview of Neuromorphic Control**

143 The neuromorphic control framework is implemented on the Speck Neuromorphic SoC [32] and
144 deployed on a small quadrotor for navigation and dynamic obstacle avoidance. The schematic of

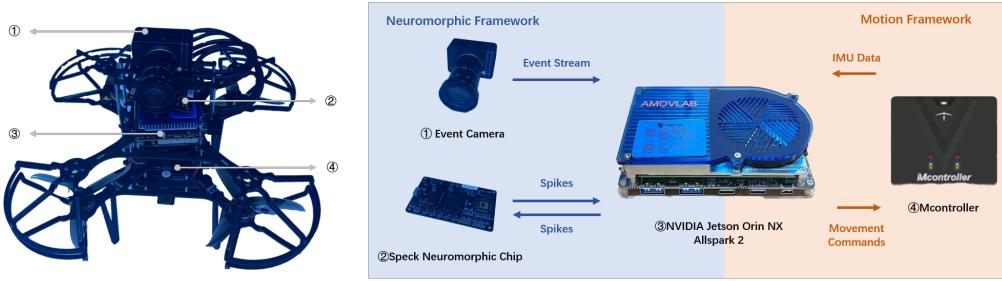


Figure 1: **Schematic of the neuromorphic quadrotor system** The left part is the quadrotor used in this work, total weight of 856 g; tip-to-tip diameter of 240 mm, with the numbers indicating the components in the right part. The right part is the hardware overview with the display of data flow, with components divided into two frameworks: the neuromorphic framework and the motion framework. One for processing neuromorphic data and the other one processing the movement control.

145 the quadrotor is illustrated in Fig. 1. In this framework, we assume the quadrotor first performs an
 146 outbound flight towards a designated target, which could be under any control law, including manual
 147 control, and then performs an inbound navigation and avoids multiple dynamic obstacles during
 148 this navigation in a fully autonomous fashion. Since our focus is on the navigation and dynamic
 149 obstacle avoidance during the inbound flight, we assume the outbound flight is performed without
 150 any collision, the environment is static (surroundings don't change), and no dynamic obstacles appear
 151 when the quadrotor is recalibrating the drift error of the IMU.

152 To minimize the obstacle avoidance latency under strict computational resource restriction [32] for
 153 longer response windows and higher success rates, we need to minimize the resources used for
 154 navigation. By introducing visual-homing, navigation during most of the flight is accomplished
 155 solely by odometry with negligible computational overhead, thereby reserving sufficient resources
 156 for obstacle avoidance. Moreover, both the calibration phase and obstacle avoidance module share
 157 the same monocular event camera, which not only reduces computational load but also significantly
 158 decreases the UAV's payload, ultimately enhancing its motion performance.

159 3.2 Event Visual-homing

160 **IMU** During the outbound flight, the quadrotor records all IMU information it produces without
 161 any correction. Generally, the IMU drift error stacks over time and will gradually become too large
 162 to provide applicable navigation information for the quadrotor[56]. A simplified version of the
 163 stack-over-time error can be defined as follows:

$$\delta r_N = \delta r_{N,0} + \delta v_{N,0,t} + \frac{1}{2}(g \cdot \delta \Theta_0 + b_{\alpha N})t^2 + \frac{1}{6}(g \cdot b_{gE})t^3 \quad (1)$$

164 where $\delta r_{N,0}$ is the initial position error, remains the same for all time; $\delta v_{N,0,t}$ is the initial velocity
 165 error with linear amplification; $g \cdot \delta \Theta_0$ denotes initial attitude angle error, $b_{\alpha N}$ denotes accelerometer
 166 error, and the third term exhibits quadratic divergence; $g \cdot b_{gE}$ denotes angular velocity error, and the
 167 last term exhibits cubic divergence. As shown in the formula, the IMU error exhibits approximately
 168 cubic drift over time. We let the quadrotor use a visual-homing algorithm to periodically return to
 169 snapshot positions, and recalibrate IMU drift error before the error aggregates too high. After the
 170 homing, the only error is the homing error during the recalibration process.

171 **Event-based Drift Error Recalibration Network** As we have shown in formula 1, we can
 172 periodically recalibrate the error caused by drift before it becomes too large, thereby keeping it at a
 173 relatively small level consistently. To fully utilize the advantage of event data and a neuromorphic
 174 framework, we use an SCNN (spiking convolutional neural network) with a Siamese structure [57].
 175 The feature extractor of the network extracts features from two continuous event streams with a
 176 temporal time window of 50 ms. The first event stream is filmed at snapshot position during the
 177 outbound flight, and the second event stream is filmed near snapshot position during the inbound
 178 flight. Two feature tensors are then concatenated and passed to the calibration module. Finally,

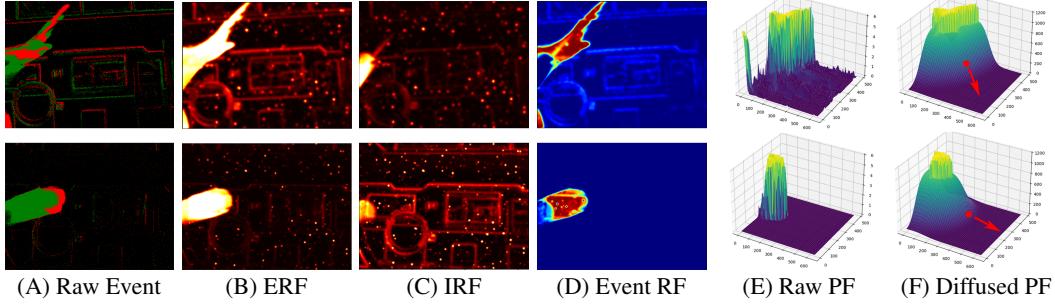


Figure 2: **Illustration of the workflow of the Event RF Model and the potential field movement command generator.** The UAV avoids dynamic obstacles' high-energy zones via the gradient descent method.

179 the network outputs a vector containing the relative x, y coordinate differences and the yaw angle
180 difference between two captured points.

181 To solve the scale issue which makes the network impossible to determine the absolute scale of object
182 brought by the monocular event camera, we train the network using data obtained in a similar-scale
183 environment and design a cyclic correction method where the UAV continues capturing event stream
184 from the corrected position and performs repeated correction until the position error output by the
185 network falls below a specific threshold. Training details are shown in supplementary note 6.

186 **Dataset** To address the absence of benchmark data for monocular event-based pose correction, we
187 constructed a novel dataset containing 50,234 event stream pairs, each precisely annotated with 4-DoF
188 relative pose ($\Delta x, \Delta y, \Delta z, \Delta \phi$) ground truth. There are four distinct indoor scenarios contained
189 in the dataset, and maximum object-camera proximity is constrained to a 10-meter range. The
190 camera was mounted on a DJI Ronin SC gimbal ($\pm 0.02^\circ$ stabilization accuracy) during shooting,
191 which eliminated the influence of pitch and roll angles while simulating the stabilized attitude of a
192 drone equipped with flight controllers. The ground truth of camera's shooting position is obtained
193 by a motion capture system with 12 Vicon Vero 2.2 motion capture camera, each featuring with a
194 resolution of 2048 x 1088 and a max frame rate of 330 Hz.

195 3.3 Bio-inspired Dynamic Obstacle Avoidance

196 **Event Receptive Field Model** The brightness-sensitive biological mechanism behind anuran
197 ganglion cells exhibits isomorphic correspondence with event-based vision sensing. By leveraging the
198 ERF-IRF spatial asymmetry, we proposed an Event RF (receptive field) model, used for suppressing
199 the event stream produced by static objects and background, and enhancing the event stream produced
200 by dynamic obstacles:

$$F(x, y, e_n, t) = \min(A_1 K(t, \tau_e) G(x, y, e_n), E_{th}) - \min(A_2 K(t - \Delta t, \tau_i) G(x, y, e_n), I_{th}) \quad (2)$$

201 where A_1 is ERF parameter, A_2 is IRF parameter, τ is energy decay parameter, Δt is IRF delay
202 parameter, and $K(t, \tau)$ is time kernel function:

$$K(t, \tau) = \begin{cases} e^{-t/\tau} & t \geq 0 \\ 0 & t < 0 \end{cases} \quad (3)$$

203 and the first term $\min(A_1 K(t, \tau_e) G(x, y, e_n), E_{th})$ is the ERF energy level, while E_{th} is ERF
204 energy threshold, the second term $\min(A_2 K(t - \Delta t, \tau_i) G(x, y, e_n), I_{th})$ is the IRF suppression
205 level, while I_{th} is the IRF suppression threshold.

206 e_n is the event passed to the model, with its coordinates and timestamp as (x_n, y_n, t_n) . $G(x, y, e_n)$
207 is the Gaussian kernel function:

$$G(x, y, e_n) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{(x - x_n)^2}{2\sigma_x^2} - \frac{(y - y_n)^2}{2\sigma_y^2}\right) \quad (4)$$

208 where σ_x and σ_y are standard deviations along the major and minor axes of the 2D elliptical Gaussian
209 function. In this model, the event stream from static objects is quickly suppressed by the IRF, which
210 drives the energy level close to zero. In contrast, the event stream produced by moving objects resists
211 suppression, allowing it to maintain a high energy level persistently, as shown in Fig. 2A–2D. A
212 complete mathematical proof is provided in supplementary note 2.

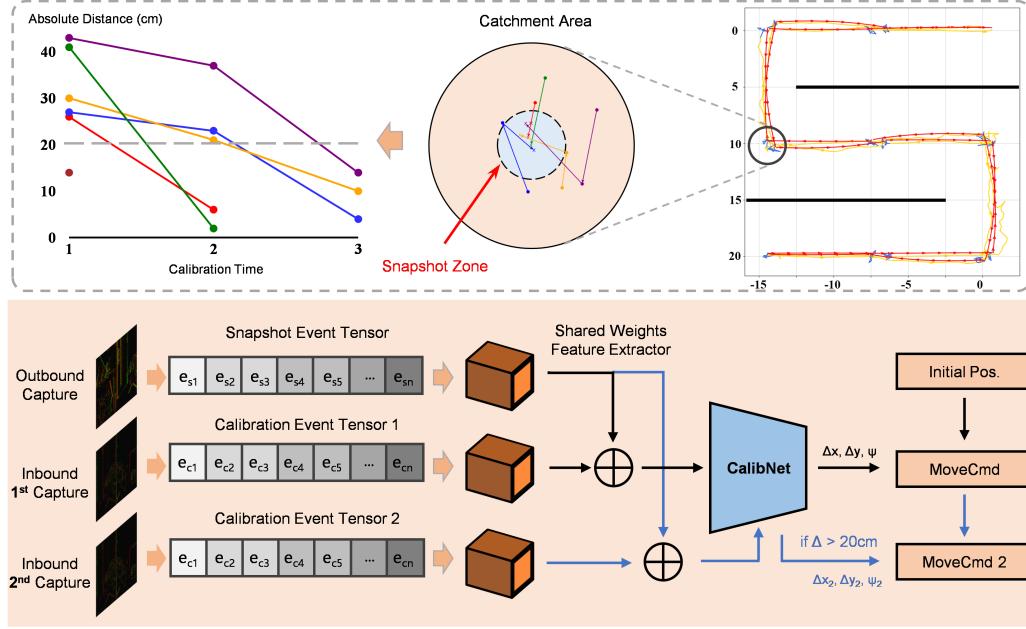


Figure 3: **Illustration of Visual-homing and CalibNet.** The quadrotor continuously calibrates itself in the catchment area until it reaches the snapshot zone.

213 **Potential Field Based Movement Command Generator** We proposed a potential field-based
 214 method to generate movement commands from the processed event stream obtained from the Event
 215 RF Model. By converting the energy map directly to the activation map, we can consider the event
 216 camera’s field of view as a 2-dimensional plane and construct potential on this plane based on the
 217 energy level of the event stream, as shown in Fig. 2E. After removing the points with excessively low
 218 potential in this potential field (here we set the threshold as half of the maximum potential), we can
 219 consider that the potential field fully represents the moving obstacles within the event camera’s field
 220 of view, as shown in Fig. 2F.

221 Since we are using a monocular event camera as the only sensor to capture the dynamic obstacle, the
 222 depth information of the dynamic obstacle cannot be obtained, which means obstacles far from the
 223 quadrotor can also be considered as dangerous objects that need to be avoided. To solve this, we use
 224 the Two-Pass Algorithm to make a connected component detection. Neglecting the potential level of
 225 points, we consider points with potential as 1 and points without potential as 0, and convert the map
 226 to a binary image I :

$$I = \{I(i, j) | I(i, j) \in \{0, 1\}, 1 \leq i \leq H, 1 \leq j \leq W\} \quad (5)$$

227 By making the connected component detection, we can assess the danger level of dynamic obstacles
 228 based on the proportion of their potential field regions occupying the entire field of view. For those
 229 dangerous potential field clusters, we define a dilation function:

$$(I \oplus B)(x, y) = \max I(x + i, y + j) \quad \text{for } (i, j) \in B \quad (6)$$

230 where \oplus is the symbol of dilation operation, (x, y) is the coordinate of the point in the plane surface,
 231 (i, j) is the offset in the structuring element B .

232 After the dilation process, since we can consider the position of the quadrotor in the center of the
 233 potential field map, we can now determine the motion direction and motion intensity of the quadrotor
 234 using gradient descent in the artificial potential field.

235 4 Experiments

236 4.1 Simulation Experiments

237 Before combining visual navigation and dynamic obstacle avoidance into a single neuromorphic
 238 system, we conducted separate experiments to verify the effectiveness of each part. The whole
 239 simulation experiment is in the Gazebo simulation environment.

240 **Visual-homing Navigation** We trained the Siamese Network using the dataset we constructed, and
 241 we test the navigation process in three different customized maps, with difficulty from low to high.
 242 The flight distance is 40 meters for the easiest map and 130 meters for the other two maps.

243 Fig. 3 shows the resulting trajectories for the proposed method. The quadrotor successfully and
 244 steadily followed the route of outbound flight and reached the starting point. Based on the calculated
 245 drift error propagation, we set the snapshot interval at 5-10 meters, with each snapshot occupying 240
 246 KB of storage space. Details on IMU error analysis can be found in supplementary note 8. We test
 247 the navigation procedure 10 times in each map, and in every single test, the quadrotor successfully
 248 reaches the destination.

249 **Dynamic Obstacle Avoidance** We use ESIM [45], an event camera simulator in Gazebo, to simulate
 250 event camera imaging effects for the quadrotor and conducted 300 dynamic obstacle avoidance tests.
 251 The dynamic obstacles were categorized into three groups based on size: coin-sized, tennis ball-sized,
 252 and basketball-sized. Each group was tested 25 times at four different distance ranges: within 0.2
 253 - 0.5m, 0.5-1m, 1-2m, and beyond 2m. We set the closest starting distance of obstacles at 0.2m
 254 since firstly, if the obstacle is too close to the quadrotor, the entire field of view will be occupied by
 255 the dynamic obstacles and the algorithm cannot make effective obstacle avoidance commands, and
 256 secondly, in real-world scenarios, it is generally impossible for dynamic to abruptly appear within the
 257 drone's immediate proximity.

258 For each obstacle detection, we also marked its centroid in the image frame and compared it with the
 259 centroid of the algorithm-processed event stream to validate the position error in dynamic obstacle
 260 detection, as shown in Table 1. Details about the calculation are provided in supplementary note 4.

Table 1: Centroid Difference Between RF Model and GT (m)

Obstacle Type	Distance	Mean	Median	Std. Dev.	M.A.D	SR
Coin-sized	0.2m - 0.5m	0.017872	0.016379	0.008949	0.001494	94%
	0.5m - 1m	0.017747	0.013403	0.009385	0.000794	92%
	1m - 2m	0.01375	0.012626	0.007721	0.020246	90%
	2m+	0.013197	0.012287	0.005258	0.00418	86%
Tennis-sized	0.2m - 0.5m	0.0213	0.022801	0.009291	0.005177	92%
	0.5m - 1m	0.017965	0.016874	0.007212	0.009258	98%
	1m - 2m	0.016847	0.014589	0.009541	0.000212	100%
	2m+	0.012912	0.010226	0.010277	0.005681	96%
Basketball-sized	0.2m - 0.5m	0.030569	0.029741	0.014125	0.00169	84%
	0.5m - 1m	0.027082	0.026069	0.013484	0.002642	96%
	1m - 2m	0.019617	0.019564	0.005942	0.00805	100%
	2m+	0.02219	0.016931	0.016632	0.010854	100%

261 To quantify the computational cost of the model, we recorded multiple event streams of dynamic
 262 obstacles and processed these event streams with our algorithm to calculate the processing time to
 263 evaluate the delay of our model. Since the model relies on generating IRF fields from prior processed
 264 events and applying decay on both ERF and IRF fields, biased results inevitably arise when processing
 265 arbitrarily cropped sections of the raw event stream. We use our algorithm to process the whole event
 266 stream and compute the ratio between the processing time and the total length of the event stream to
 267 obtain the unbiased average latency of 2.3 ms. Detailed data are shown in supplementary note 11.

268 **Combined Task Simulation** With the core algorithms proven, we then demonstrated the complete
 269 pipeline by combining the complex tasks together. Using the 3 maps we created in Gazebo (mentioned
 270 in 4.1.1), we made the quadrotor traverse the outbound route using odometry without any global
 271 position information and added randomly throwing dynamic obstacles when the quadrotor is flying
 272 during its inbound journey. Among 50 tests on each map, the success rate is 100% for the first map,
 273 98% for the second map, and 94% for the last map. Figure demonstrations and other details are
 274 shown in supplementary note 7.

275 4.2 Real-world Experiments

276 In this section, we conduct indoor experiments using our neuromorphic platform mentioned in section
 277 3.1, and also conduct extra indoor experiments under different extreme conditions (flicker condition
 278 and darkish condition) to validate the algorithm's robustness.

279 **Indoor Experiment** As previously mentioned, the main
 280 goal of the indoor experiment is to verify the effectiveness
 281 of our neuromorphic framework in a real-world setup and
 282 test our neuromorphic structure’s advantage in computa-
 283 tional resource consumption, energy consumption, and
 284 verify the performance of the framework with such low
 285 consumption in a tiny autonomous quadrotor. The exper-
 286 iment is conducted in a 10m * 10m flight arena. Three
 287 experimenters stationed at designated locations threw dy-
 288 namic obstacles at passing drones, and experimenters were
 289 instructed to remain stationary to prevent the quadrotor
 290 from misidentifying them as dynamic obstacles. During
 291 10 repeated trials, the quadrotor successfully avoided all dynamic obstacles and reached the destina-
 292 tion in every instance, as shown in Fig. 4. Details about obstacles are shown in supplementary note
 293 5.



Figure 4: Real-world Experiment The quadrotor avoids the tennis ball thrown by experimenters during navigation.

294 **Reduced Energy Consumption on Neuromorphic Hardware** We tested the energy consump-
 295 tion and run time between different setups, and a main observation is that the neuromorphic chip
 296 demonstrates a two orders of magnitude reduction in power consumption compared to conventional
 297 devices. Systems equipped with this neuromorphic chip achieve a 95% reduction in operational
 298 energy consumption (down to 5% of original levels) when executing identical tasks using the same
 299 algorithms. The total system energy consumption decreases to 21% of baseline values. Notably, in
 300 neuromorphic systems, the primary energy expenditure originates from three core processes: the
 301 onboard computer operations, data exchange between the neuromorphic chip and flight controller,
 302 and motion command execution. Details about the energy consumption of each architecture are
 303 displayed in supplementary note 10.

304 **Robustness Under Extreme Light Condition** Despite our main goal is to validate the advantage of
 305 our bio-inspired algorithm and fully neuromorphic framework, to better demonstrate the effectiveness
 306 of the unique event stream data modality, we need to test the framework under extreme light conditions
 307 to prove its robustness. We test the flight performance in the same arena under three different light
 308 conditions: light (10 - 100 lux), flicker (1 - 100 lux), dim (1 - 10 lux), and dark (0 - 1 lux). The
 309 result shows that the performance of the quadrotor is approximately the same under different light
 310 conditions, but it does not work in dark conditions. Experiment details are shown in supplementary
 311 note 9.

312 4.3 Comparison and Analysis

313 **Comparison with the State-of-
 314 the-Arts** Since this work is the
 315 first to implement a fully neu-
 316 romorphic pipeline on complex
 317 navigation and dynamic obsta-
 318 cle avoidance tasks, to provide a
 319 reference level, we compare our
 320 Event RF model to some related
 321 traditional approaches based on
 322 object recognition and trajectory
 323 estimation[58, 12, 20, 59], as
 324 shown in Table 2. These ap-
 325 proaches are tested under the

Table 2: Quantitative evaluation on Dynamic Obstacle Avoidance Task Only Method 4 utilized non-visual sensors, and experiments that did not employ onboard computational resources were specifically marked.

Method	Latency	Pos. Err.	SR
Method 1 [58]	19.12 ms	0.11 m	96.3
Method 2 [12]	39.49 ms	0.11 m	89.1
Method 3 [20]	3.56 ms	0.09 m	86.7
Method 4 (GTX 4090) [59]	14 ms	LiDAR	95.75
Method 4 (onboard) [59]	27 ms	LiDAR	86.5
Ours	2.34 ms	0.02 m	94.5

326 same simulation environment using their official codes. Since Li-DAR can get precise position
 327 information of obstacles, there is no position error for the Li-DAR method. Among all the works,
 328 our system achieves the lowest latency, significantly lower than other types of sensors. The only
 329 approach with comparable latency is the method 3 [20], which also employs an event camera but lacks
 330 a navigation module, thus allocating all computational resources to obstacle avoidance. Since our
 331 method does not perform object recognition or trajectory estimation, we cannot compare prediction

332 speed errors. However, in terms of positional error, our work also achieves the lowest. Regarding
333 obstacle avoidance success rate, our performance is very close to the best, with only a 2% gap.

334 **Analysis of Event RF Model** We delve into the parameter choosing for the Event RF model and
335 conduct simulation experiments on multiple dynamic obstacles of different sizes to evaluate the effect
336 of parameter selection on performance and the generalizability of the parameters. There are 3 pairs of
337 parameter in Event RF model: (A_1, A_2) , (τ_e, τ_i) , (E_{th}, I_{th}) , and 3 separate parameters: Δt , σ_x , σ_y ,
338 and the value of each parameter significantly affects the model's performance.

339 σ_x, σ_y , as the standard deviation along the major and minor axes of the 2D elliptical Gaussian
340 function, affects the size of the receptive field generated by each event. Under perfect motion
341 compensation, $\sigma_x = \sigma_y = 1$ makes the IRF just sufficient to suppress the stimulation caused
342 by the ERF. However, considering the limited computational resources, achieving perfect motion
343 compensation is challenging, along with the inherent noise introduced by the event camera itself.
344 Setting $\sigma_x = \sigma_y = 2$ could achieve a better result. Making σ_x and σ_y unequal could enable the
345 model to exhibit anisotropy, reducing sensitivity to motion in specific directions, especially when
346 setting different σ_x and σ_y for ERF and IRF individually.

347 Δt affects the delay of the IRF relative to the ERF. Higher Δt increases the size of high-energy
348 regions for dynamic objects in the model, thereby increasing the distance between the centroid of
349 the real obstacle and the centroid of the dynamic obstacle in the model, resulting in greater error.
350 However, if Δt is too small, the ERF will be rapidly overridden by the IRF, thereby reducing the
351 model's sensitivity to slow-moving objects. In experiments, we find $\Delta t = 5$ ms delivers optimal
352 performance, and this value is suitable for the vast majority of dynamic obstacle avoidance scenarios.

353 Through mathematical derivation, we found that the model achieves optimal performance when
354 $\frac{A_1}{A_2} = e^{-\Delta t/\tau_i}$, the model reaches optimal performance, and since $\tau_e < \tau_i$, we can determine the
355 values of A_1, A_2 and τ_e based on the value of τ_i . In biological systems, the value of τ_i typically
356 ranges from 25 to 50 ms. We conducted tests at 5-millisecond (ms) intervals and found the best value
357 as 25 ms. Therefore, $\frac{A_1}{A_2}$ should be 1.22, and when $\tau_e = 5$ ms we get the best result. Under ideal
358 conditions, setting $E_{th} = I_{th}$ would enable perfect static event cancellation. However, in practice,
359 sensor noise and firing threshold fluctuations in biological neurons necessitate permitting minor
360 deviations to prevent noise-induced false dynamic responses; here we choose $\frac{I_{th}}{E_{th}} = 1.2$ based on
361 our experimental testing. Details on the analysis process can be found in supplementary note 12.

362 5 Conclusion

363 This paper presents a fully autonomous neuromorphic navigation and dynamic obstacle avoidance
364 pipeline for tiny autonomous unmanned aerial vehicles. Its Event RF Model is the first bio-inspired
365 algorithm that could make the quadrotor bypass the object recognition and trajectory estimation
366 processes, thus avoiding dynamic obstacles in a real-time manner. By reducing the latency to 2.3 ms,
367 the model gives a much longer reaction time window for the quadrotor when facing dynamic obstacles
368 with speeds up to 10m/s. Comparative evaluations under identical experimental conditions prove
369 our neuromorphic approach outperforms current state-of-the-art solutions for autonomous UAVs,
370 delivering significantly lower latency in high-velocity dynamic obstacle avoidance while maintaining
371 comparable success rates under stringent onboard computational constraints. The dataset presented
372 in our work also establishes a solid foundation for further research on event-based pose calibration.
373 Moreover, with reduced energy consumption and robustness under various light conditions, this work
374 presents a substantial step toward neuromorphic sensing and controlling for UAVs, and exhibits the
375 great potential of neuromorphic architecture on tiny autonomous robots, revealing the possibility of
376 tiny autonomous robots to evolve to higher levels of operational capability and performance.

377 **Limitation and Future Work** The current work relies on IMU information for navigation, and
378 the monocular event camera could not obtain depth information of the dynamic obstacle. Future
379 work will further explore the Event RF Model's capabilities by leveraging its ability to distinguish
380 between dynamic and static objects. A stereo vision setup will be used to explore the possibility of
381 tiny autonomous neuromorphic quadrotors exploring and avoiding dynamic obstacles in completely
382 unfamiliar environments without relying on any prior information.

383 **References**

- 384 [1] Dan Cireşan, Ueli Meier, Jonathan Masci, and Jürgen Schmidhuber. A committee of neural
385 networks for traffic sign classification. In *The 2011 international joint conference on neural*
386 *networks*, pages 1918–1921. IEEE, 2011.
- 387 [2] Kimon P Valavanis and George J Vachtsevanos. *Handbook of unmanned aerial vehicles*.
388 Springer Publishing Company, Incorporated, 2014.
- 389 [3] Xiao Liang, Shirou Zhao, Guodong Chen, Guanglei Meng, and Yu Wang. Design and develop-
390 ment of ground station for uav/ugv heterogeneous collaborative system. *Ain Shams Engineering*
391 *Journal*, 12(4):3879–3889, 2021.
- 392 [4] Ali Krayani, Atm Shafiu Alam, Lucio Marcenaro, Arumugam Nallanathan, and Carlo Regaz-
393 zoni. Automatic jamming signal classification in cognitive uav radios. *IEEE Transactions on*
394 *Vehicular Technology*, 71(12):12972–12988, 2022.
- 395 [5] Nour El-Din Safwat, Roberto Sabatini, Alessandro Gardi, Ismail Mohamed Hafez, and Fatma
396 Newagy. Urban air mobility communication performance considering cochannel interference.
397 *IEEE Transactions on Aerospace and Electronic Systems*, 60(4):5089–5100, 2024.
- 398 [6] Roman Horbyk. “the war phone”: mobile communication on the frontline in eastern ukraine.
399 *Digital War*, 3(1):9–24, 2022.
- 400 [7] Sawyer B Fuller. Four wings: An insect-sized aerial robot with steering ability and payload
401 capacity for autonomy. *IEEE Robotics and Automation Letters*, 4(2):570–577, 2019.
- 402 [8] Thinal Raj, Fazida Hanim Hashim, Aqilah Baseri Huddin, Mohd Faisal Ibrahim, and Aini
403 Hussain. A survey on lidar scanning mechanisms. *Electronics*, 9(5):741, 2020.
- 404 [9] Sivakumar Balasubramanian, Yogesh M Chukewad, Johannes M James, Geoffrey L Barrows,
405 and Sawyer B Fuller. An insect-sized robot that uses a custom-built onboard camera and a neural
406 network to classify and respond to visual input. In *2018 7th IEEE International Conference on*
407 *Biomedical Robotics and Biomechatronics (Biorob)*, pages 1297–1302. IEEE, 2018.
- 408 [10] Stéphane Viollet, Stéphanie Godiot, Robert Leitel, Wolfgang Buss, Patrick Breugnon, Mohsine
409 Menouni, Raphaël Juston, Fabien Expert, Fabien Colonnier, Géraud L’Eplattenier, et al. Hard-
410 ware architecture and cutting-edge assembly process of a tiny curved compound eye. *Sensors*,
411 14(11):21702–21721, 2014.
- 412 [11] Hugh Durrant-Whyte and Tim Bailey. Simultaneous localization and mapping: part i. *IEEE*
413 *robotics & automation magazine*, 13(2):99–110, 2006.
- 414 [12] Zhefan Xu, Xiaoyang Zhan, Baihan Chen, Yumeng Xiu, Chenhao Yang, and Kenji Shimada.
415 A real-time dynamic obstacle tracking and mapping system for uav navigation and collision
416 avoidance with an rgbd camera. In *2023 IEEE International Conference on Robotics and*
417 *Automation (ICRA)*, pages 10645–10651. IEEE, 2023.
- 418 [13] Minghao Lu, Han Chen, and Peng Lu. Perception and avoidance of multiple small fast moving
419 objects for quadrotors with only low-cost rgbd camera. *IEEE Robotics and Automation Letters*,
420 7(4):11657–11664, 2022.
- 421 [14] Bruno Bodin, Harry Wagstaff, Sajad Saecdi, Luigi Nardi, Emanuele Vespa, John Mawer, Andy
422 Nisbet, Mikel Luján, Steve Furber, Andrew J Davison, et al. Slambench2: Multi-objective head-
423 to-head benchmarking for visual slam. In *2018 IEEE International Conference on Robotics and*
424 *Automation (ICRA)*, pages 3637–3644. IEEE, 2018.
- 425 [15] Guillermo Gallego, Tobi Delbrück, Garrick Orchard, Chiara Bartolozzi, Brian Taba, Andrea
426 Censi, Stefan Leutenegger, Andrew J. Davison, Jörg Conradt, Kostas Daniilidis, and Davide
427 Scaramuzza. Event-based vision: A survey. *IEEE Transactions on Pattern Analysis and*
428 *Machine Intelligence*, 44(1):154–180, 2022.
- 429 [16] Alex Zhu, Liangzhe Yuan, Kenneth Chaney, and Kostas Daniilidis. Ev-flownet: Self-supervised
430 optical flow estimation for event-based cameras. In *Robotics: Science and Systems XIV*.
431 Robotics: Science and Systems Foundation, June 2018.

- 432 [17] Alex Zihao Zhu, Liangzhe Yuan, Kenneth Chaney, and Kostas Daniilidis. Unsupervised
 433 event-based learning of optical flow, depth, and egomotion. In *Proceedings of the IEEE/CVF*
 434 *conference on computer vision and pattern recognition (CVPR)*, pages 989–997, 2019.
- 435 [18] Minggui Teng, Chu Zhou, Hanyue Lou, and Boxin Shi. Nest: Neural event stack for event-based
 436 image enhancement. In *European Conference on Computer Vision (ECCV)*, pages 660–676.
 437 Springer, 2022.
- 438 [19] Botao He, Haojia Li, Siyuan Wu, Dong Wang, Zhiwei Zhang, Qianli Dong, Chao Xu, and
 439 Fei Gao. Fast-dynamic-vision: Detection and tracking dynamic objects with event and depth
 440 sensing. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*,
 441 pages 3071–3078. IEEE, 2021.
- 442 [20] Davide Falanga, Kevin Kleber, and Davide Scaramuzza. Dynamic obstacle avoidance for
 443 quadrotors with event cameras. *Science Robotics*, 5(40):eaaz9712, 2020.
- 444 [21] Travis DeWolf, Kinjal Patel, Pawel Jaworski, Roxana Leontie, Joe Hays, and Chris Eliasmith.
 445 Neuromorphic control of a simulated 7-dof arm using loihi. *Neuromorphic Computing and*
 446 *Engineering*, 3(1):014007, 2023.
- 447 [22] Ioannis Polykretis, Lazar Supic, and Andreea Danilescu. Bioinspired smooth neuromorphic
 448 control for robotic arms. *Neuromorphic Computing and Engineering*, 3(1):014013, 2023.
- 449 [23] Abdulla Ayyad, Mohamad Halwani, Dewald Swart, Rajkumar Muthusamy, Fahad Almaskari,
 450 and Yahya Zweiri. Neuromorphic vision based control for the precise positioning of robotic
 451 drilling systems. *Robotics and Computer-Integrated Manufacturing*, 79:102419, 2023.
- 452 [24] Antonio Vitale, Alpha Renner, Celine Nauer, Davide Scaramuzza, and Yulia Sandamirskaya.
 453 Event-driven vision and control for uavs on a neuromorphic chip. In *2021 IEEE International*
 454 *Conference on Robotics and Automation (ICRA)*, pages 103–109. IEEE, 2021.
- 455 [25] Mike Davies, Narayan Srinivasa, Tsung-Han Lin, Gautham Chinya, Yongqiang Cao, Sri Harsha
 456 Choday, Georgios Dimou, Prasad Joshi, Nabil Imam, Shweta Jain, et al. Loihi: A neuromorphic
 457 manycore processor with on-chip learning. *IEEE Micro*, 38(1):82–99, 2018.
- 458 [26] Julien Dupeyroux, Jesse J Hagenaars, Federico Paredes-Vallés, and Guido CHE de Croon.
 459 Neuromorphic control for optic-flow-based landing of mavs using the loihi processor. In *2021*
 460 *IEEE International Conference on Robotics and Automation (ICRA)*, pages 96–102. IEEE,
 461 2021.
- 462 [27] Federico Paredes-Vallés, Jesse J Hagenaars, Julien Dupeyroux, Stein Stroobants, Yingfu Xu,
 463 and Guido CHE de Croon. Fully neuromorphic vision and control for autonomous drone flight.
 464 *Science Robotics*, 9(90):eadl0591, 2024.
- 465 [28] Le Zhu, Michael Mangan, and Barbara Webb. Neuromorphic sequence learning with an event
 466 camera on routes through vegetation. *Science Robotics*, 8(82):eadg3679, 2023.
- 467 [29] Ning Qiao, Hesham Mostafa, Federico Corradi, Marc Osswald, Fabio Stefanini, Dora Sum-
 468 islawska, and Giacomo Indiveri. A reconfigurable on-line learning spiking neuromorphic
 469 processor comprising 256 neurons and 128k synapses. *Frontiers in neuroscience*, 9:141, 2015.
- 470 [30] Francesco Galluppi, Christian Denk, Matthias C Meiner, Terrence C Stewart, Luis A Plana, Chris
 471 Eliasmith, Steve Furber, and Jörg Conradt. Event-based neural computing on an autonomous
 472 mobile platform. In *2014 IEEE international conference on robotics and automation (ICRA)*,
 473 pages 2862–2867. IEEE, 2014.
- 474 [31] Jesse Hagenaars, Federico Paredes-Vallés, and Guido De Croon. Self-supervised learning
 475 of event-based optical flow with spiking neural networks. *Advances in Neural Information*
 476 *Processing Systems*, 34:7167–7179, 2021.
- 477 [32] Man Yao, Ole Richter, Guangshe Zhao, Ning Qiao, Yannan Xing, Dingheng Wang, Tianxiang
 478 Hu, Wei Fang, Tugba Demirci, Michele De Marchi, et al. Spike-based dynamic computing with
 479 asynchronous sensing-computing neuromorphic chip. *Nature Communications*, 15(1):4464,
 480 2024.

- 481 [33] Ole Richter, Yannan Xing, Michele De Marchi, Carsten Nielsen, Merkourios Katsimpris,
 482 Roberto Cattaneo, Yudi Ren, Yalun Hu, Qian Liu, Sadique Sheik, et al. Speck: A smart event-
 483 based vision sensor with a low latency 327k neuron convolutional neuronal network processing
 484 pipeline. *arXiv preprint arXiv:2304.06793*, 2023.
- 485 [34] BA Cartwright and TS Collett. How honey bees use landmarks to guide their return to a food
 486 source. *Nature*, 295(5850):560–564, 1982.
- 487 [35] Aymeric Denuelle and Mandyam V Srinivasan. Bio-inspired visual guidance: From insect
 488 homing to uas navigation. In *2015 IEEE International Conference on Robotics and Biomimetics
 489 (ROBIO)*, pages 326–332. IEEE, 2015.
- 490 [36] Julien Dupeyroux, Julien R Serres, and Stéphane Viollet. Antbot: A six-legged walking robot
 491 able to home like desert ants in outdoor environments. *Science Robotics*, 4(27):eaau0307, 2019.
- 492 [37] Wolfgang Stürzl and Hanspeter A Mallot. Efficient visual homing based on fourier transformed
 493 panoramic images. *Robotics and Autonomous Systems*, 54(4):300–313, 2006.
- 494 [38] Aymeric Denuelle and Mandyam V Srinivasan. A sparse snapshot-based navigation strategy for
 495 uas guidance in natural environments. In *2016 IEEE International Conference on Robotics and
 496 Automation (ICRA)*, pages 3455–3462. IEEE, 2016.
- 497 [39] Tom van Dijk, Christophe De Wagter, and Guido CHE de Croon. Visual route following for
 498 tiny autonomous robots. *Science Robotics*, 9(92):eadk0310, 2024.
- 499 [40] Frédéric Gaillard, Michael A Arbib, Fernando J Corbacho, and Hyun Bong Lee. Modeling the
 500 physiological responses of anuran r3 ganglion cells. *Vision research*, 38(17):2551–2568, 1998.
- 501 [41] Younginha Jung, Sungmoo Lee, Jun Kyu Rhee, Chae-Eun Lee, Bradley J Baker, and Yoon-Kyu
 502 Song. Imaging electrical activity of retinal ganglion cells with fluorescent voltage and calcium
 503 indicator proteins in retinal degenerative rd1 blind mice. *bioRxiv*, pages 2023–12, 2023.
- 504 [42] Jerome Y Lettin, Humberto R Maturana, Warren S McCulloch, and Walter H Pitts. What the
 505 frog’s eye tells the frog’s brain. *Proceedings of the IRE*, 47(11):1940–1951, 2007.
- 506 [43] Horace B Barlow et al. Possible principles underlying the transformation of sensory messages.
 507 *Sensory communication*, 1(01):217–233, 1961.
- 508 [44] Kristian Donner and Carola AM Yovanovich. A frog’s eye view: Foundational revelations and
 509 future promises. In *Seminars in Cell & Developmental Biology*, volume 106, pages 72–85.
 510 Elsevier, 2020.
- 511 [45] Henri Rebecq, Daniel Gehrig, and Davide Scaramuzza. Esim: an open event camera simulator.
 512 In *Conference on robot learning (CoRL)*, pages 969–982. PMLR, 2018.
- 513 [46] Xin Zhou, Zhepei Wang, Hongkai Ye, Chao Xu, and Fei Gao. Ego-planner: An esdf-free
 514 gradient-based local planner for quadrotors. *IEEE Robotics and Automation Letters*, 6(2):478–
 515 485, 2020.
- 516 [47] Omid Esrafilian and Hamid D Taghirad. Autonomous flight and obstacle avoidance of a quadro-
 517 tor by monocular slam. In *2016 4th International Conference on Robotics and Mechatronics
 518 (ICROM)*, pages 240–245. IEEE, 2016.
- 519 [48] Yi Lin, Fei Gao, Tong Qin, Wenliang Gao, Tianbo Liu, William Wu, Zhenfei Yang, and Shaojie
 520 Shen. Autonomous aerial navigation using monocular visual-inertial fusion. *Journal of Field
 521 Robotics*, 35(1):23–51, 2018.
- 522 [49] Thomas Eppenberger, Gianluca Cesari, Marcin Dymczyk, Roland Siegwart, and Renaud Dubé.
 523 Leveraging stereo-camera data for real-time dynamic obstacle detection and tracking. In
 524 *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages
 525 10528–10535. IEEE, 2020.
- 526 [50] Helen Oleynikova, Dominik Honegger, and Marc Pollefeys. Reactive avoidance using embedded
 527 stereo vision for mav flight. In *2015 IEEE International Conference on Robotics and Automation
 528 (ICRA)*, pages 50–56. IEEE, 2015.

- 529 [51] Andrew J Barry, Peter R Florence, and Russ Tedrake. High-speed autonomous obstacle
530 avoidance with pushbroom stereo. *Journal of Field Robotics*, 35(1):52–68, 2018.
- 531 [52] Sikang Liu, Michael Watterson, Sarah Tang, and Vijay Kumar. High speed navigation for
532 quadrotors with limited onboard sensing. In *2016 IEEE international conference on robotics
533 and automation (ICRA)*, pages 1484–1491. IEEE, 2016.
- 534 [53] Albert S Huang, Abraham Bachrach, Peter Henry, Michael Krainin, Daniel Maturana, Dieter
535 Fox, and Nicholas Roy. Visual odometry and mapping for autonomous flight using an rgb-
536 d camera. In *Robotics Research: The 15th International Symposium ISRR*, pages 235–252.
537 Springer, 2017.
- 538 [54] Zhefan Xu, Yumeng Xiu, Xiaoyang Zhan, Baihan Chen, and Kenji Shimada. Vision-aided
539 uav navigation and dynamic obstacle avoidance using gradient-based b-spline trajectory opti-
540 mization. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages
541 1214–1220. IEEE, 2023.
- 542 [55] Junqing Ning, Haotian Zhang, and Quan Quan. Dynamic obstacle avoidance of quadcopters
543 with monocular camera based on image-based visual servo. In *2022 International Conference
544 on Unmanned Aircraft Systems (ICUAS)*, pages 150–156. IEEE, 2022.
- 545 [56] John H Wall, David M Bevly, et al. Characterization of various imu error sources and the effect
546 on navigation performance. In *Proceedings of the 18th international technical meeting of the
547 satellite division of the institute of navigation (ION GNSS 2005)*, pages 967–978, 2005.
- 548 [57] Jane Bromley, Isabelle Guyon, Yann LeCun, Eduard Säckinger, and Roopak Shah. Signature
549 verification using a "siamese" time delay neural network. *Advances in neural information
550 processing systems*, 6, 1993.
- 551 [58] Zhefan Xu, Xiaoyang Zhan, Yumeng Xiu, Christopher Suzuki, and Kenji Shimada. Onboard
552 dynamic-object detection and tracking for autonomous robot navigation with rgb-d camera.
553 *IEEE Robotics and Automation Letters*, 9(1):651–658, 2023.
- 554 [59] Zhefan Xu, Xinming Han, Haoyu Shen, Hanyu Jin, and Kenji Shimada. Navrl: Learning safe
555 flight in dynamic environments. *IEEE Robotics and Automation Letters*, 2025.

556 **NeurIPS Paper Checklist**

557 **1. Claims**

558 Question: Do the main claims made in the abstract and introduction accurately reflect the
559 paper's contributions and scope?

560 Answer: [Yes]

561 Justification: The paper's contributions are presented in the section 3 and 4.

562 Guidelines:

- 563 • The answer NA means that the abstract and introduction do not include the claims
564 made in the paper.
- 565 • The abstract and/or introduction should clearly state the claims made, including the
566 contributions made in the paper and important assumptions and limitations. A No or
567 NA answer to this question will not be perceived well by the reviewers.
- 568 • The claims made should match theoretical and experimental results, and reflect how
569 much the results can be expected to generalize to other settings.
- 570 • It is fine to include aspirational goals as motivation as long as it is clear that these goals
571 are not attained by the paper.

572 **2. Limitations**

573 Question: Does the paper discuss the limitations of the work performed by the authors?

574 Answer: [Yes]

575 Justification: The limitations are discussed at the end of the paper.

576 Guidelines:

- 577 • The answer NA means that the paper has no limitation while the answer No means that
578 the paper has limitations, but those are not discussed in the paper.
- 579 • The authors are encouraged to create a separate "Limitations" section in their paper.
- 580 • The paper should point out any strong assumptions and how robust the results are to
581 violations of these assumptions (e.g., independence assumptions, noiseless settings,
582 model well-specification, asymptotic approximations only holding locally). The authors
583 should reflect on how these assumptions might be violated in practice and what the
584 implications would be.
- 585 • The authors should reflect on the scope of the claims made, e.g., if the approach was
586 only tested on a few datasets or with a few runs. In general, empirical results often
587 depend on implicit assumptions, which should be articulated.
- 588 • The authors should reflect on the factors that influence the performance of the approach.
589 For example, a facial recognition algorithm may perform poorly when image resolution
590 is low or images are taken in low lighting. Or a speech-to-text system might not be
591 used reliably to provide closed captions for online lectures because it fails to handle
592 technical jargon.
- 593 • The authors should discuss the computational efficiency of the proposed algorithms
594 and how they scale with dataset size.
- 595 • If applicable, the authors should discuss possible limitations of their approach to
596 address problems of privacy and fairness.
- 597 • While the authors might fear that complete honesty about limitations might be used by
598 reviewers as grounds for rejection, a worse outcome might be that reviewers discover
599 limitations that aren't acknowledged in the paper. The authors should use their best
600 judgment and recognize that individual actions in favor of transparency play an impor-
601 tant role in developing norms that preserve the integrity of the community. Reviewers
602 will be specifically instructed to not penalize honesty concerning limitations.

603 **3. Theory assumptions and proofs**

604 Question: For each theoretical result, does the paper provide the full set of assumptions and
605 a complete (and correct) proof?

606 Answer: [Yes]

607 Justification: Proofs of all models and formulas used in this paper are provided in the
608 supplementary material.

609 Guidelines:

- 610 • The answer NA means that the paper does not include theoretical results.
611 • All the theorems, formulas, and proofs in the paper should be numbered and cross-
612 referenced.
613 • All assumptions should be clearly stated or referenced in the statement of any theorems.
614 • The proofs can either appear in the main paper or the supplemental material, but if
615 they appear in the supplemental material, the authors are encouraged to provide a short
616 proof sketch to provide intuition.
617 • Inversely, any informal proof provided in the core of the paper should be complemented
618 by formal proofs provided in appendix or supplemental material.
619 • Theorems and Lemmas that the proof relies upon should be properly referenced.

620 **4. Experimental result reproducibility**

621 Question: Does the paper fully disclose all the information needed to reproduce the main ex-
622 perimental results of the paper to the extent that it affects the main claims and/or conclusions
623 of the paper (regardless of whether the code and data are provided or not)?

624 Answer: [Yes]

625 Justification: All information needed to reproduced the result are provided in the section 3
626 of the paper.

627 Guidelines:

- 628 • The answer NA means that the paper does not include experiments.
629 • If the paper includes experiments, a No answer to this question will not be perceived
630 well by the reviewers: Making the paper reproducible is important, regardless of
631 whether the code and data are provided or not.
632 • If the contribution is a dataset and/or model, the authors should describe the steps taken
633 to make their results reproducible or verifiable.
634 • Depending on the contribution, reproducibility can be accomplished in various ways.
635 For example, if the contribution is a novel architecture, describing the architecture fully
636 might suffice, or if the contribution is a specific model and empirical evaluation, it may
637 be necessary to either make it possible for others to replicate the model with the same
638 dataset, or provide access to the model. In general, releasing code and data is often
639 one good way to accomplish this, but reproducibility can also be provided via detailed
640 instructions for how to replicate the results, access to a hosted model (e.g., in the case
641 of a large language model), releasing of a model checkpoint, or other means that are
642 appropriate to the research performed.
643 • While NeurIPS does not require releasing code, the conference does require all submis-
644 sions to provide some reasonable avenue for reproducibility, which may depend on the
645 nature of the contribution. For example
646 (a) If the contribution is primarily a new algorithm, the paper should make it clear how
647 to reproduce that algorithm.
648 (b) If the contribution is primarily a new model architecture, the paper should describe
649 the architecture clearly and fully.
650 (c) If the contribution is a new model (e.g., a large language model), then there should
651 either be a way to access this model for reproducing the results or a way to reproduce
652 the model (e.g., with an open-source dataset or instructions for how to construct
653 the dataset).
654 (d) We recognize that reproducibility may be tricky in some cases, in which case
655 authors are welcome to describe the particular way they provide for reproducibility.
656 In the case of closed-source models, it may be that access to the model is limited in
657 some way (e.g., to registered users), but it should be possible for other researchers
658 to have some path to reproducing or verifying the results.

659 **5. Open access to data and code**

660 Question: Does the paper provide open access to the data and code, with sufficient instruc-
661 tions to faithfully reproduce the main experimental results, as described in supplemental
662 material?

663 Answer: [No]

664 Justification: We will release the code upon acceptance.
665

666 Guidelines:

- 667 • The answer NA means that paper does not include experiments requiring code.
- 668 • Please see the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- 669 • While we encourage the release of code and data, we understand that this might not be
670 possible, so “No” is an acceptable answer. Papers cannot be rejected simply for not
671 including code, unless this is central to the contribution (e.g., for a new open-source
672 benchmark).
- 673 • The instructions should contain the exact command and environment needed to run to
674 reproduce the results. See the NeurIPS code and data submission guidelines (<https://nips.cc/public/guides/CodeSubmissionPolicy>) for more details.
- 675 • The authors should provide instructions on data access and preparation, including how
676 to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- 677 • The authors should provide scripts to reproduce all experimental results for the new
678 proposed method and baselines. If only a subset of experiments are reproducible, they
679 should state which ones are omitted from the script and why.
- 680 • At submission time, to preserve anonymity, the authors should release anonymized
681 versions (if applicable).
- 682 • Providing as much information as possible in supplemental material (appended to the
683 paper) is recommended, but including URLs to data and code is permitted.

684 **6. Experimental setting/details**

685 Question: Does the paper specify all the training and test details (e.g., data splits, hyper-
686 parameters, how they were chosen, type of optimizer, etc.) necessary to understand the
687 results?

688 Answer: [Yes]

689 Justification: Training and testing details are provided in the supplementary material.

690 Guidelines:

- 691 • The answer NA means that the paper does not include experiments.
- 692 • The experimental setting should be presented in the core of the paper to a level of detail
693 that is necessary to appreciate the results and make sense of them.
- 694 • The full details can be provided either with the code, in appendix, or as supplemental
695 material.

696 **7. Experiment statistical significance**

697 Question: Does the paper report error bars suitably and correctly defined or other appropriate
698 information about the statistical significance of the experiments?

699 Answer: [Yes]

700 Justification: Errors are provided in the table 1 and in the supplementary material.

701 Guidelines:

- 702 • The answer NA means that the paper does not include experiments.
- 703 • The authors should answer "Yes" if the results are accompanied by error bars, confi-
704 dence intervals, or statistical significance tests, at least for the experiments that support
705 the main claims of the paper.
- 706 • The factors of variability that the error bars are capturing should be clearly stated (for
707 example, train/test split, initialization, random drawing of some parameter, or overall
708 run with given experimental conditions).
- 709 • The method for calculating the error bars should be explained (closed form formula,
710 call to a library function, bootstrap, etc.)
- 711 • The assumptions made should be given (e.g., Normally distributed errors).
- 712 • It should be clear whether the error bar is the standard deviation or the standard error
713 of the mean.
- 714 • It is OK to report 1-sigma error bars, but one should state it. The authors should
715 preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis
716 of Normality of errors is not verified.

- 718 • For asymmetric distributions, the authors should be careful not to show in tables or
719 figures symmetric error bars that would yield results that are out of range (e.g. negative
720 error rates).
721 • If error bars are reported in tables or plots, The authors should explain in the text how
722 they were calculated and reference the corresponding figures or tables in the text.

723 **8. Experiments compute resources**

724 Question: For each experiment, does the paper provide sufficient information on the com-
725 puter resources (type of compute workers, memory, time of execution) needed to reproduce
726 the experiments?

727 Answer: [Yes]

728 Justification: The devices we used and the latency of the system are provided in section 3.
729 Training details are provided in the supplementary material.

730 Guidelines:

- 731 • The answer NA means that the paper does not include experiments.
732 • The paper should indicate the type of compute workers CPU or GPU, internal cluster,
733 or cloud provider, including relevant memory and storage.
734 • The paper should provide the amount of compute required for each of the individual
735 experimental runs as well as estimate the total compute.
736 • The paper should disclose whether the full research project required more compute
737 than the experiments reported in the paper (e.g., preliminary or failed experiments that
738 didn't make it into the paper).

739 **9. Code of ethics**

740 Question: Does the research conducted in the paper conform, in every respect, with the
741 NeurIPS Code of Ethics <https://neurips.cc/public/EthicsGuidelines>?

742 Answer: [Yes]

743 Justification: The research conducted in the paper conforms with the NeurIPS Code of
744 Ethics.

745 Guidelines:

- 746 • The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
747 • If the authors answer No, they should explain the special circumstances that require a
748 deviation from the Code of Ethics.
749 • The authors should make sure to preserve anonymity (e.g., if there is a special consid-
750 eration due to laws or regulations in their jurisdiction).

751 **10. Broader impacts**

752 Question: Does the paper discuss both potential positive societal impacts and negative
753 societal impacts of the work performed?

754 Answer: [NA]

755 Justification: There is no societal impact of the work performed.

756 Guidelines:

- 757 • The answer NA means that there is no societal impact of the work performed.
758 • If the authors answer NA or No, they should explain why their work has no societal
759 impact or why the paper does not address societal impact.
760 • Examples of negative societal impacts include potential malicious or unintended uses
761 (e.g., disinformation, generating fake profiles, surveillance), fairness considerations
762 (e.g., deployment of technologies that could make decisions that unfairly impact specific
763 groups), privacy considerations, and security considerations.
764 • The conference expects that many papers will be foundational research and not tied
765 to particular applications, let alone deployments. However, if there is a direct path to
766 any negative applications, the authors should point it out. For example, it is legitimate
767 to point out that an improvement in the quality of generative models could be used to
768 generate deepfakes for disinformation. On the other hand, it is not needed to point out
769 that a generic algorithm for optimizing neural networks could enable people to train
770 models that generate Deepfakes faster.

- 771 • The authors should consider possible harms that could arise when the technology is
772 being used as intended and functioning correctly, harms that could arise when the
773 technology is being used as intended but gives incorrect results, and harms following
774 from (intentional or unintentional) misuse of the technology.
775 • If there are negative societal impacts, the authors could also discuss possible mitigation
776 strategies (e.g., gated release of models, providing defenses in addition to attacks,
777 mechanisms for monitoring misuse, mechanisms to monitor how a system learns from
778 feedback over time, improving the efficiency and accessibility of ML).

779 **11. Safeguards**

780 Question: Does the paper describe safeguards that have been put in place for responsible
781 release of data or models that have a high risk for misuse (e.g., pretrained language models,
782 image generators, or scraped datasets)?

783 Answer: [NA]

784 Justification: The paper poses no such risks.

785 Guidelines:

- 786 • The answer NA means that the paper poses no such risks.
787 • Released models that have a high risk for misuse or dual-use should be released with
788 necessary safeguards to allow for controlled use of the model, for example by requiring
789 that users adhere to usage guidelines or restrictions to access the model or implementing
790 safety filters.
791 • Datasets that have been scraped from the Internet could pose safety risks. The authors
792 should describe how they avoided releasing unsafe images.
793 • We recognize that providing effective safeguards is challenging, and many papers do
794 not require this, but we encourage authors to take this into account and make a best
795 faith effort.

796 **12. Licenses for existing assets**

797 Question: Are the creators or original owners of assets (e.g., code, data, models), used in
798 the paper, properly credited and are the license and terms of use explicitly mentioned and
799 properly respected?

800 Answer: [Yes]

801 Justification: All assets used in this paper are properly credited, the license and terms of use
802 are explicitly mentioned and properly respected.

803 Guidelines:

- 804 • The answer NA means that the paper does not use existing assets.
805 • The authors should cite the original paper that produced the code package or dataset.
806 • The authors should state which version of the asset is used and, if possible, include a
807 URL.
808 • The name of the license (e.g., CC-BY 4.0) should be included for each asset.
809 • For scraped data from a particular source (e.g., website), the copyright and terms of
810 service of that source should be provided.
811 • If assets are released, the license, copyright information, and terms of use in the
812 package should be provided. For popular datasets, paperswithcode.com/datasets
813 has curated licenses for some datasets. Their licensing guide can help determine the
814 license of a dataset.
815 • For existing datasets that are re-packaged, both the original license and the license of
816 the derived asset (if it has changed) should be provided.
817 • If this information is not available online, the authors are encouraged to reach out to
818 the asset's creators.

819 **13. New assets**

820 Question: Are new assets introduced in the paper well documented and is the documentation
821 provided alongside the assets?

822 Answer: [Yes]

823 Justification: The model proposed in this paper is clearly documented.

824 Guidelines:

- 825 • The answer NA means that the paper does not release new assets.
826 • Researchers should communicate the details of the dataset/code/model as part of their
827 submissions via structured templates. This includes details about training, license,
828 limitations, etc.
829 • The paper should discuss whether and how consent was obtained from people whose
830 asset is used.
831 • At submission time, remember to anonymize your assets (if applicable). You can either
832 create an anonymized URL or include an anonymized zip file.

833 **14. Crowdsourcing and research with human subjects**

834 Question: For crowdsourcing experiments and research with human subjects, does the paper
835 include the full text of instructions given to participants and screenshots, if applicable, as
836 well as details about compensation (if any)?

837 Answer: [NA]

838 Justification: The paper does not involve crowdsourcing nor research with human subjects.

839 Guidelines:

- 840 • The answer NA means that the paper does not involve crowdsourcing nor research with
841 human subjects.
842 • Including this information in the supplemental material is fine, but if the main contribu-
843 tion of the paper involves human subjects, then as much detail as possible should be
844 included in the main paper.
845 • According to the NeurIPS Code of Ethics, workers involved in data collection, curation,
846 or other labor should be paid at least the minimum wage in the country of the data
847 collector.

848 **15. Institutional review board (IRB) approvals or equivalent for research with human
849 subjects**

850 Question: Does the paper describe potential risks incurred by study participants, whether
851 such risks were disclosed to the subjects, and whether Institutional Review Board (IRB)
852 approvals (or an equivalent approval/review based on the requirements of your country or
853 institution) were obtained?

854 Answer: [NA]

855 Justification: The paper does not involve crowdsourcing nor research with human subjects.

856 Guidelines:

- 857 • The answer NA means that the paper does not involve crowdsourcing nor research with
858 human subjects.
859 • Depending on the country in which research is conducted, IRB approval (or equivalent)
860 may be required for any human subjects research. If you obtained IRB approval, you
861 should clearly state this in the paper.
862 • We recognize that the procedures for this may vary significantly between institutions
863 and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the
864 guidelines for their institution.
865 • For initial submissions, do not include any information that would break anonymity (if
866 applicable), such as the institution conducting the review.

867 **16. Declaration of LLM usage**

868 Question: Does the paper describe the usage of LLMs if it is an important, original, or
869 non-standard component of the core methods in this research? Note that if the LLM is used
870 only for writing, editing, or formatting purposes and does not impact the core methodology,
871 scientific rigorosity, or originality of the research, declaration is not required.

872 Answer: [NA]

873 Justification: The core method development in this research does not involve LLMs as any
874 important, original, or non-standard components.

875 Guidelines:

- 876 • The answer NA means that the core method development in this research does not
877 involve LLMs as any important, original, or non-standard components.

878
879

- Please refer to our LLM policy (<https://neurips.cc/Conferences/2025/LLM>) for what should or should not be described.