

# Attention-Driven LPLC2 Neural Ensemble Model for Multi-Target Looming Detection and Localization

Anonymous Authors

**Abstract**—Lobula plate/lobula columnar, type 2 (LPLC2) visual projection neurons in the fly’s visual system possess highly looming-selective properties, making them ideal for developing artificial collision detection systems. The four dendritic branches of individual LPLC2 neurons, each tuned to specific directional motion, enhance the robustness of looming detection by utilizing radial motion opponency. Existing models of LPLC2 neurons either concentrate on individual cells to detect centroid-focused expansion or utilize population-level strategies to obtain global collision information. However, their potential for addressing multi-target collision scenarios remains largely untapped. In this study, we propose a numerical model for LPLC2 populations, leveraging a bottom-up attention mechanism driven by motion-sensitive neural pathways to generate attention fields (AFs). This integration of AFs with highly nonlinear LPLC2 responses enables precise and continuous detection of multiple looming objects emanating from any region of the visual field. We begin by conducting comparative experiments to evaluate the proposed model against two related models, highlighting its unique characteristics. Next, we test its ability to detect multiple targets in dynamic natural scenarios. Finally, we validate the model using real-world video data collected by unmanned aerial vehicles (UAVs). Experimental results demonstrate that the proposed model excels in detecting, distinguishing, and tracking multiple looming targets with remarkable speed and accuracy. This advanced ability to detect and localize looming objects, especially in complex and dynamic environments, holds great promise for overcoming collision-detection challenges in UAVs.

**Index Terms**—*Drosophila* LPLC2, neural model, attention, collision detection, multi-target motion, UAV

## I. INTRODUCTION

The ability to quickly and accurately identify approaching objects is critical for the survival of animals [1]. Various species accomplish this by distinguishing between different types of motion based on the unique movement characteristics of objects [2]–[4]. These efficient collision detection mechanisms also hold significant promise for applications in autonomous driving, drone technology, and robotic navigation. However, machine learning-based collision detection methods often require large training datasets [5], [6], and sensor-based approaches, such as LiDAR, are typically expensive and limited in application contexts [7].

Insects exhibit simple yet highly effective collision avoidance systems [4]. Inspired by the lobula giant movement detector (LGMD) neurons in locusts, researchers have developed low-power, fast, and real-time collision detection systems [8], [9]. However, traditional LGMD-based bio-inspired neural networks face limitations, particularly in their selectivity. These networks struggle to precisely differentiate looming motion from other movement types, often responding to receding

or translating movements, which are irrelevant to collision detection [10].

Collision detection systems based on directionally sensitive neurons in fruit flies *Drosophila* have shown promising results [11], [12]. Recent biological studies suggest that the lobula plate/lobula columnar, type 2 (LPLC2) visual projection neurons in fruit flies are better suited for looming perception due to their ultra-selective properties [13], [14]. These neurons, with cross-shaped dendrites spanning specific layers, respond strongly to outward motion from the center of their receptive field while being inhibited by inward motion [13].

Recent studies have investigated implementing LPLC2’s extreme selectivity through nonlinear integration [15], [16] and machine learning techniques to predict their properties and tuning [5]. While these approaches highlight the effectiveness of bio-inspired LPLC2 models for collision detection, they overlook a critical aspect: the LPLC2 neural ensemble consists of numerous neurons, each covering a portion of the visual field. This raises two key questions: (1) How can LPLC2 neurons collectively recognize expansion originating from any location within the visual field? (2) How can the responses of the LPLC2 population be harnessed to achieve multi-target collision detection?

In this study, we introduce a multi-attention mechanism to the LPLC2 neural network model (mLPLC2) to address these challenges. Inspired by the LPLC2 population’s ability to cover the entire visual field of the fruit fly (Fig. 1), we develop attention fields (AFs) that focus on suspicious local motion signals. Using a nonlinear integration approach, we simulate the response of LPLC2 neurons and continuously extract local motion information from regions of interests. This integration is more energy-efficient and effective than independently calculating the response of each LPLC2 neuron.

To evaluate the proposed mLPLC2 model, we used computer-generated stimuli with various motion features to test its functional characteristics. We then assessed its performance in dynamic natural scenarios with multiple looming objects, as well as in real-world UAV scenes facing diverse collision threats. The mLPLC2 model showed notable improvements over traditional LGMD-based systems, particularly in its ability to adaptively focus on fast-approaching objects emanating from any region of the visual field.

The structure of this paper is as follows: Section II reviews typical bio-inspired models for looming detection. Section III describes the proposed model in detail. Section IV presents experimental evaluations of the model. Finally, Section V summarizes the findings and discusses future directions.

## II. RELATED WORK

This section provides a brief review on the most relevant works regarding (1) development of LGMDs for detecting looming objects, and (2) functional characteristics of individual LPLC2 neurons and their population dynamics.

### A. Bio-inspired LGMD Looming Detector

The looming-sensitive LGMD1-based neural network, inspired by the locust visual system, was introduced by Rind and Bramwell in 1996 [17]. They demonstrated that visual information is processed through excitatory inputs and lateral inhibition within the presynaptic structure, which shapes LGMD1's unique selectivity. This numerical model responds most strongly to looming objects, exhibits brief excitation to receding objects, and generates weak, short-term responses to translating stimuli. LGMD2, a neighboring counterpart to LGMD1, also functions as a looming detector. Its presynaptic excitation is thought to be mediated by ON and OFF cells [18]. In 2015, Fu and Yue developed the first LGMD2-based visual neural network, modeling the ON and OFF mechanisms [19]. This model separates luminance changes into parallel channels, encoding excitatory and inhibitory signals through spatiotemporal computations similar to the LGMD1 model. It successfully reproduces LGMD2's specific selectivity, responding exclusively to light-to-dark luminance changes, such as those caused by dark objects with centrifugal expanding edges [20].

However, both LGMD1 and LGMD2 computational models demonstrate limitation on well distinguishing radial motion from other patterns, such as translation, compromising their selectivity for collision detection.

### B. LPLC2 Physiology and Model

In 2017, Klapoetke et al. identified an ultra-selective ensemble of looming-detecting neurons, LPLC2, in *Drosophila* [13]. Individual LPLC2 neurons possess dendritic arbors that form four distinct fields, each occupying a separate layer of the lobula plate. Within each lobula plate layer, the directionally selective signals from T4 and T5 neurons are gathered. More specifically, these presynaptic, directionally selective neurons are tuned to one of the cardinal directions: front-to-back, back-to-front, upwards, and downwards, serves as the local optic flow sensors, directly providing inputs to LPLC2 ensemble [21], [22] (Fig. 1a).

The four dendritic branches of an LPLC2 neuron form a radial, outward-oriented structure that aligns with the edges of a looming object expanding from the center of its receptive field (Fig. 1b). The direction of maximum outward dendritic spread is aligned with the preferred direction of T4 and T5 cells terminating in specific layers [13]. This dendritic architecture is highly specialized for detecting centrifugal motion from a small looming object whose visual expansion matches the receptive field center of the LPLC2 neuron.

By contrast, common motion stimuli, such as translation, only partially align with this structure and are suppressed by strong inhibitory responses to centripetal motion [13].

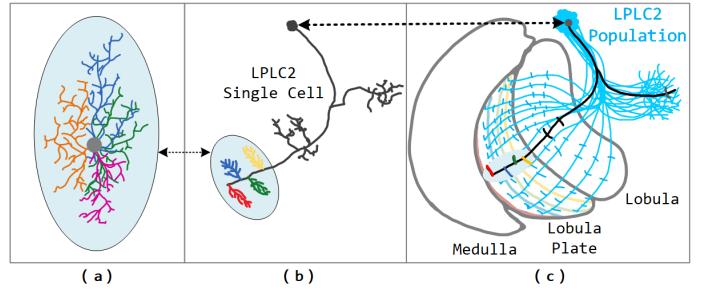


Fig. 1. The schematic anatomical structure of LPLC2 neurons is as follows: (a) Each arm of LPLC2's cross-shaped primary dendrites ramifies in one of the lobula plate layers and extends along that layer's preferred motion direction. (b) A single LPLC2 cell integrates directional signals from four sub-layers of lobula plate. (c) LPLC2 cells are a population of  $\sim 80$  visual projection neurons, with their dendrites collectively covering the lobula plate.

In 2022, Hua et al. further refined the understanding of LPLC2's ultra-selective characteristics via highly nonlinear neural computation, demonstrating its strong responsiveness to dark, centroid-emitted centrifugal motion patterns while remaining nearly unresponsive to motion originating outside its receptive field center [15].

However, a single LPLC2 neuron can only process visual information within its receptive field, a limited area approximately  $60^\circ$  in diameter [14]. In contrast, LGMD neurons receive stimuli through a single giant dendritic arbor that spans the entire visual field of one compound eye [17]. While this broad coverage enables LGMD neurons to detect looming objects, it limits their ability to localize or distinguish multiple objects in different parts of the visual field. Accurately identifying colliding objects from various locations within the global visual field remains a significant challenge.

The LPLC2 population, consisting of approximately  $\sim 80$  visual projection neurons, offers a potential solution. With their dendrites collectively covering the lobula plate, the population enables regional detection of motion across the entire visual field [13], [22]. This study explores how LPLC2 neurons distributed across different regions respond to motion using bottom-up salience generation, thus provides insights into their potential to separate global visual information into regional processing.

## III. FORMULATION OF THE NEURAL NETWORK

This section begins with an introduction to the computational presynaptic neuropils of LPLC2. Following this, we describe the multiple-attention mechanism employed in the model. The algorithm for the proposed mLPLC2 model is outlined in pseudocode in Algorithm 1, with the parameter settings detailed in Table I.

### A. Modeling of Presynaptic Neuropils

The retinal layer is derived from the compound eye structure of *Drosophila*, arranged as a 2D-matrix of photoreceptors( $P$ ), with the amount corresponding to the resolution of the input visual images. Each photoreceptor captures the gray-scale

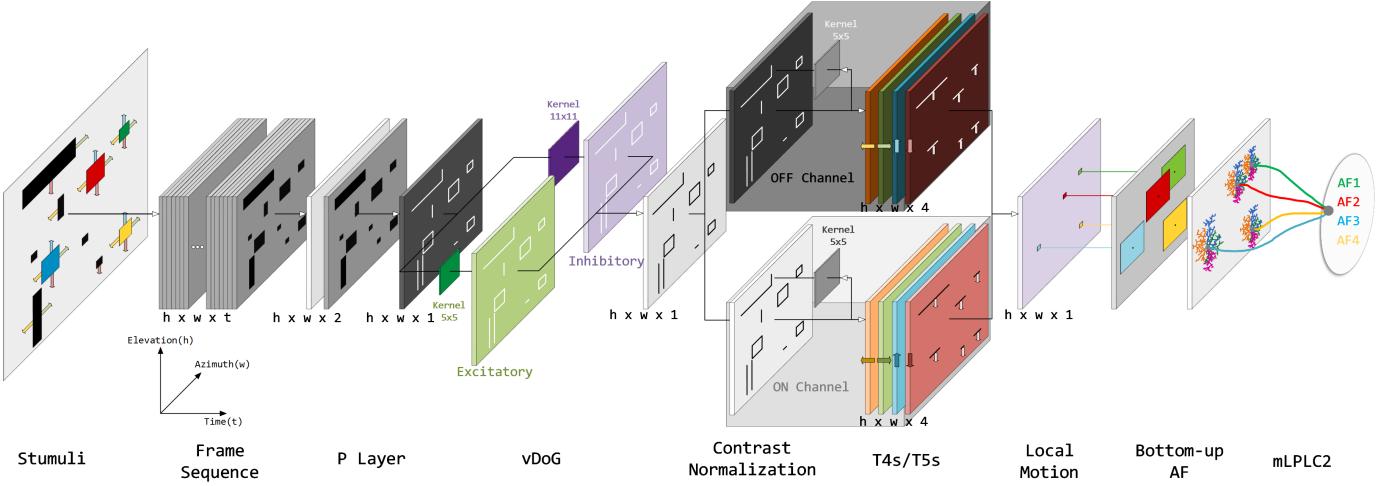


Fig. 2. The flowchart of the proposed mLPLC2 neural network outlines the following steps. The stimulation is input into the neural network, where the difference between consecutive frames is calculated to extract edge expansion information. Excitation and inhibition signals from the lamina monopolar cells (LMCs) are integrated using the vDoG (variant Difference of Gaussian) simulation with polarity selectivity. Motion signals representing brightness increases and decreases are sent into the ON and OFF channels, respectively, where they undergo contrast normalization. A triple-correlation model simulates T4 and T5 cells to extract motion in four cardinal directions. Directional signals are integrated to generate the local motion signal. The most suspicious local motion signal is selected to drive a bottom-up process, identifying a new attention field (AF). The mLPLC2 model non-linearly integrates motion signals from multiple AFs, enabling the detection and localization of multiple looming objects.

brightness and computes the difference with the previous frame to extract the brightness changes, which is defined as

$$L(x, y, t) = P(x, y, t) - P(x, y, t - 1), \quad (1)$$

where  $P(x, y, t)$  and  $P(x, y, t - 1)$  represent the gray-scale brightness of two successive frames.

In the lamina layer, a spatial mechanism named variant of Difference of Gaussians (vDoG) with ON/OFF-contrast selectivity [23] is used to simulate excitation and inhibition signal produced by large monocular cells (LMCs) [24]. The vDoG removes redundant background motion and enhances the selectivity to moving edges. The excitation can be calculated as

$$E(x, y, t) = \iint_{D_e} L(x+u, y+v, t) \cdot K_e(u+r_{D_e}, v+r_{D_e}) du dv, \quad (2)$$

where  $r_{D_e}$  stands for the radius of excitatory zone around LMC located at  $(x, y)$ . The inhibition  $I(x, y, t)$  conforms to the same computation that is omitted here. The Gaussian kernel  $K_e$  and  $K_i$  can be calculated as

$$K_{e,i}(u, v) = \frac{1}{2\pi\sigma^2} e^{-\frac{u^2+v^2}{2\sigma^2}}, \quad (3)$$

where  $\sigma$  represents the standard deviation  $\sigma_e$  of inner Gaussian, and  $\sigma_i$  of outer Gaussian. Subsequently, the broader inhibitory Gaussian is subtracted from the narrower excitatory one with polarity selectivity in order to match the following separation of signals into parallel ON/OFF channels, and the operation is consistent with the previous research [23].

After that, the LMCs split signals into ON and OFF pathways, an essential phenomenon observed in biological visual systems [22], [25], [26]. The ON pathway carries information about luminance increments, whereas the OFF pathway signals

luminance decrements. This mechanism can be realized by ‘half-wave’ rectifier which is defined as

$$\begin{aligned} L_{ON}(x, y, t) &= \frac{L_{vDoG}(x, y, t) + |L_{vDoG}(x, y, t)|}{2}, \\ L_{OFF}(x, y, t) &= \frac{|L_{vDoG}(x, y, t)| - L_{vDoG}(x, y, t)}{2}. \end{aligned} \quad (4)$$

In the medulla layer, trans-medullary, medulla-intrinsic, centrifugal, and complex tangential cells, together, transmit and process temporally filtered signals [22]. Within this process, contrast normalization is applied to both ON and OFF pathways through spatial integration of feedback neural signals [27], [28]. Let  $N$  stands for both ON and OFF interneurons, the entire process is defined as

$$N(x, y, t) = \tanh\left(\frac{N(x, y, t)}{\varepsilon + \hat{N}(x, y, t)}\right), \quad (5)$$

$$\hat{N}(x, y, t) = \iint_{D_c} N(x+u, y+v, t) \cdot K_c(u+r_{D_c}, v+r_{D_c}) du dv, \quad (6)$$

where  $D_c$  stands for the impact zone around a medulla cell located at  $(x, y)$  with radius at  $r_{D_c}$ . The Gaussian kernel  $K_c$  can be calculated similarly to Eq. 3.

Subsequently, the visual signal is transmitted to the direction-selective T4 cells in the medulla and T5 cells in the lobula [21], [22]. We use a triple-correlation model [29] based on Hassenstein-Reichardt Correlators (HRC) [30] to simulate every single type of T4(T5) cell perceives motion from one out of four cardinal directions in ON(OFF) channel [31], [32]. Each T4/T5 samples from two locations: the position itself, and a certain distance in one of the four cardinal directions.

Here we apply a first-order low-pass filter to obtain the delayed signal  $D(x, y, t)$  as

$$D(x, y, t) = \alpha_1 \cdot N(x, y, t) + (1 - \alpha_1) \cdot N(x, y, t - 1), \quad (7)$$

$$\alpha_1 = \tau_i / (\tau_i + \tau_1),$$

where  $\tau_1$  is a delay and  $\tau_i$  is the time interval between two consecutive frames in milliseconds.

Taking the rightward motion detector  $M_{r_i}(x, y, t)$  as an example here, it correlates triple input signals from two positions, at different time as

$$M_{r_i}(x, y, t) = N(x, y, t) \cdot D(x, y, t) \cdot D(x + sd, y, t), \quad (8)$$

$$- \beta \cdot N(x + sd, y, t) \cdot D(x, y, t) \cdot D(x + sd, y, t),$$

where  $\beta$  is a bias coefficient at null-preferred motion.  $i$  stands for current step in the dynamic delay mechanism and  $N(x, y, t)$  denotes the contrast-normalized information from ON/OFF channels. For each local T4(T5) cell, we sum the outputs of HRCs at various distances from the cell itself as the local motion perception result for T4(T5) cells. Next, the outputs of a local cell correlating neighboring cells at various distances are accumulated, and a first-order low-pass filter is then applied to introduce a delay, yielding the directional signals for T4 (T5) cells, as the following:

$$\hat{M}(x, y, t) = \sum_{i=1}^{N_c} M_{r_i}(x, y, t), \quad \alpha_2 = \tau_i / (\tau_i + \tau_{N_c}), \quad (9)$$

$$T4_r(x, y, t) = \alpha_2 \hat{M}(x, y, t) + (1 - \alpha_2) \hat{M}(x, y, t - 1),$$

where  $N_c$  is the number of connected neurons and  $\tau_{N_c}$  is the corresponding dynamic delay for T4(T5) cells in milliseconds.

In four distinct layers of the lobula plate, LPLC2 neuron receive inputs from four sub-populations of both T4 and T5 cells [13], [15]. The local directional motion signals from T4 and T5 cells at each position  $(x, y)$  at time  $t$  are integrated into  $LM(x, y, t)$  to facilitate subsequent calculations. Here, we take the rightward motion computation as the example:

$$LM_r(x, y, t) = f(r(x, y, t) - l(x, y, t)),$$

$$r(x, y, t) = T4_r(x, y, t)^{\gamma_1} + T5_r(x, y, t)^{\gamma_2}, \quad (10)$$

$$l(x, y, t) = T4_l(x, y, t)^{\gamma_1} + T5_l(x, y, t)^{\gamma_2},$$

where  $f(\cdot)$  denotes the Leaky-ReLU activation function.  $r(x, y, t)$  and  $l(x, y, t)$  represent the rightward and leftward motion information of T4/T5 cells integrated from the ON and OFF channels, respectively.  $\gamma_1$  and  $\gamma_2$  are exponential weights for the ON and OFF channels.

Subsequently, two directional systems integrate the magnitude of local motion  $LM(x, y, t)$  at time  $t$  via horizontal motion  $H(x, y, t)$  and vertical motion  $V(x, y, t)$  as

$$H(x, y, t) = \max(LM_r(x, y, t), LM_l(x, y, t)),$$

$$V(x, y, t) = \max(LM_d(x, y, t), LM_u(x, y, t)), \quad (11)$$

$$LM(x, y, t) = \|H^2(x, y, t) + V^2(x, y, t)\|_2.$$

## B. Bottom-up Multi-Attention Based Nonlinear Integration

Next, we introduce the multiple-attention mechanism driven by bottom-up motion salience that equips the LPLC2 model with the capability for multi-target looming detection, forming the mLPLC2 framework. When multiple objects approach from various regions of the visual field, the proposed mechanism enables mLPLC2 to promptly detect and focus on these objects from the very first moment.

At each time  $t$ , we search for suspicious motion cues from the whole visual field. When a target is on a direct approaching course, its edges will appear to expand outward from a single point in the observer's view. The expanding edges recorded in the  $LM$  2D-matrix start from a single point and gradually increase in size. We select  $(\hat{x}, \hat{y})$  that maximizes motion magnitude  $LM(x, y, t)$  as the centroid of new attention field, which is retinotopically aligned with the initial position of the expanding motion. Then the new attention field  $AF^{n+1}(x^{n+1}, y^{n+1})$  for suspicious local motion  $LM(\hat{x}, \hat{y}, t)$  will be incorporated into the set of all existing attention fields  $AF^N$  when the following conditions are satisfied: (1) the location does not belong to any existing AFs; (2) it has the highest local motion magnitude outside the existing AFs, and this magnitude exceeds the threshold  $T_a$ . Accordingly, the whole process is defined as the following:

$$(\hat{x}, \hat{y}) = \operatorname{argmax}_{x,y}(LM(x, y, t)),$$

$$(x, y) \in \mathbb{C}_\Omega(AF^0(x_0, y_0) \cup \dots \cup AF^n(x_n, y_n)),$$

$$AF^N = AF^N \cup AF^{n+1}(x^{n+1}, y^{n+1}), \text{ if } LM(\hat{x}, \hat{y}, t) > T_a. \quad (12)$$

$\Omega$  denotes for the entire visual field and  $AF^0(x_0, y_0) \sim AF^n(x_n, y_n)$  represent current existing attention fields.

Next, directional signals covered by each existing attention fields  $AF^N$  is integrated in a highly nonlinear manner. Specifically for each LPLC2 cell corresponding to a single  $AF^n$ , each arm of its cross-shaped primary dendrites ramifies in one of the lobula plate layers and extends along that layer's preferred motion direction [13]. Within  $AF^n$ , the motions toward four cardinal directions at time  $t$  are mapped into four quadrant:

$$Q_1^n(t) = \iint_{\Gamma_1^n} (LM_r(u, v, t) + LM_u(u, v, t)) du dv;$$

$$Q_2^n(t) = \iint_{\Gamma_2^n} (LM_l(u, v, t) + LM_u(u, v, t)) du dv;$$

$$Q_3^n(t) = \iint_{\Gamma_3^n} (LM_l(u, v, t) + LM_d(u, v, t)) du dv;$$

$$Q_4^n(t) = \iint_{\Gamma_4^n} (LM_r(u, v, t) + LM_d(u, v, t)) du dv, \quad (13)$$

where  $\Gamma_1^n \sim \Gamma_4^n$  denotes the four quadrants in the  $AF^n$  divided by the centroid of  $AF^n$  located at  $(\hat{x}, \hat{y})$ .

Only when all four dendritic arbors have received their preferred motion signal from  $AF^n$ , its postsynaptic  $LPLC2^n(t)$

will be activated. The attention-driven regional information is integrated as

$$\begin{aligned} LPLC2^n(t) = & \text{bool}\left(\prod_{i=1}^4 Q_i^n(t)\right) \\ & \times (Q_1^n(t) + Q_2^n(t) + Q_3^n(t) + Q_4^n(t)), \end{aligned} \quad (14)$$

where  $\text{bool}(\cdot)$  function outputs 1 once none of  $Q_1 \sim Q_4$  equals 0. Importantly, if the LPLC2 cell in current  $AF^k$  remains inactive or shows minimal activation after a certain period, the  $AF^k$  will disappear, mathematically expressed as follows:

$$AF^N = AF^N \setminus AF_{(x_k, y_k)}^k, \text{ if } \sum_{i=t-d}^t LPLC2^k(i) < T_d, \quad (15)$$

where  $T_d$  is a threshold and  $d$  denotes the time duration over which responses are accumulated within the current AF. However, the mLPLC2 model ensures that at least one AF is always retained.

In summary, the proposed model continuously monitors potential collision areas, generating response information that provides detailed and precise data about looming objects in the given scenario. The parameters of the mLPLC2 neural network are empirically derived, inspired by biological functions. With its feed-forward structure, this model operates without the need for training that is free of data.

TABLE I  
PARAMETER CONFIGURATION

Parameter	Description	Value
$\{r_{D_e}, r_{D_i}, r_{D_c}\}$	the impact radius of LMC/medulla cell	{5,11,5}
$\{\sigma_e, \sigma_i, \sigma_c\}$	standard deviation in $K_e, K_i, K_c$	{10,20,20}
$\varepsilon$	coefficient in contrast normalization	0.2
$N_c$	the number of connected neurons	5
$\{\tau_1, \tau_{N_c}\}$	time delay in milliseconds	{80,80 40}
$\beta$	bias coefficient for HRC	1.5
$\{\gamma_1, \gamma_2\}$	exponent on the ON and OFF signals	{0.9,0.5}
$R_{AF}$	the radius of a single attention field	40
$T_a$	threshold for establishing attention field	10
$T_d$	threshold for conserving attention field	5000
$d$	time length for accumulating response	$10 \cdot \tau_i$

#### IV. EXPERIMENTAL EVALUATION

In this section, we first explore the specific selectivity of mLPLC2 by setting up comparative experiments in Subsection IV-A. Then, we challenge mLPLC2 using multiple imminently appearing objects embedded in dynamic natural scenes in Subsection IV-B. Finally, in Subsection IV-C, we evaluate the practicality of mLPLC2 using real-world footage captured from UAVs.

##### A. Model Characteristic Study

To investigate the basic functionality of mLPLC2, we tested its responses to dark objects approaching, receding, bright objects approaching, receding, translating motion, and large-scale grating motion (input sampling resolution at  $100 \times 100$ , length at 50 frames, and frequency at 33 Hz). The LGMD2 model [19] with specific selectivity and the single-LPLC2 model [15] without the attention mechanism were used as comparisons. Notably, we also limit the AF of mLPLC2 to

---

##### Algorithm 1: LPLC2 with Multi-Attention Mechanism

---

```

for  $t = 2 : T$  do
  Input:  $P(x, y, t)$ 
  // Computation of Pre-synaptic Neuropils
  // Retina layer
  Compute  $L(x, y, t)$  via Eq. (1);
  // Lamina layer
  Compute  $L_{vDoG}(x, y, t)$  via Eq. (2)~(3);
  Split the visual signal into ON/OFF via Eq. (4);
  // Medulla layer
  Compute HRCs via Eq. (5)~(8);
  Compute directional-selective neurons via Eq. (9);
  // Lobula plate layer
  Integrate local motion via Eq. (10), (11);
  // Operation of Attention Fields
  // search for suspicious motion cue
  bool NewAttention = 1; double Maxpoint = 0;
  for  $y = 1 : Height \text{ of visual field}$  do
    for  $x = 1 : Width \text{ of visual field}$  do
      // Judge whether to set a new AF
      if  $(x, y) \in \text{any existing AF}$  then
         $\text{NewAttention} = 0;$ 
      // localize the new AF centroid
      if  $MaxPoint^n(\hat{x}, \hat{y}, t) < LM(x, y, t)$  then
         $MaxPoint^n(\hat{x}, \hat{y}, t) = LM(x, y, t);$ 
    end for
  end for
  // Update AFs
  if NewAttention then
    Generate a new  $AF^n$  at  $MaxPoint^n(\hat{x}, \hat{y})$ ;
  // Integrate response for all existing AFs
  for  $AF^n = 1 : n$  do
    Integrate response of  $AF^n$  as  $LPLC2^n(t)$ 
    if  $\sum_{i=t-d}^t LPLC2^n(i) < T_d$  then
      Abandon  $AF^n$ 
  end for
  Output:  $LPLC2_{(x_1, y_1)}^1(t) \sim LPLC2_{(x_n, y_n)}^n(t)$ 
end for

```

---

a maximum of 1, allowing aLPLC2 to participate in the experiment. The experimental results are presented in Figure 3, with snapshots of different motion patterns displayed in the first row.

The experimental results clearly demonstrate the similar selectivity of LPLC2, aLPLC2, and mLPLC2. Specifically, they all exhibit the strongest response to black objects approaching from the center of the visual field and show weaker responses to bright approaching objects. They are not responsive to receding objects with a contracting trend, nor do they react to translating or large-scale grating motion. In contrast, the LGMD2 model exhibited a markedly different response, reacting to bright, retreating objects. Additionally, it was disturbed by motion patterns with translational characteristics, such as

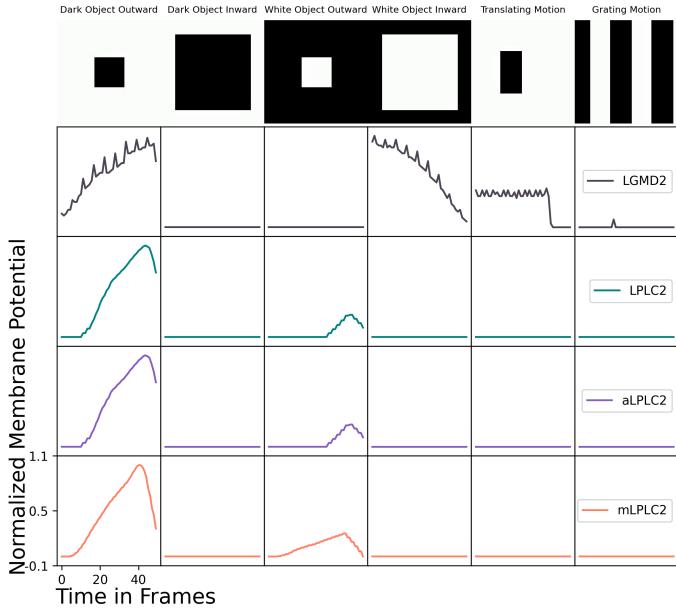


Fig. 3. Different models' responses to various motion characteristics (aLPLC2 refers to the scenario where the attention area of mLPLC2 is limited to 1). In the figure, each square represents a complete experiment, with the x-axis indicating time in frames and the y-axis representing the corresponding membrane potential of the models. The LPLC2-based neural network model exhibits selectivity to only approaching stimuli.

translating motion. Evidently, the LPLC2-based model demonstrates superior performance in detecting only approaching objects. As the edge of an object exceeds the receptive field of a single LPLC2 cell, the cell can no longer detect the edge expansion, causing the membrane potential to decrease. This phenomenon is evident from frame 40 onward in columns 1 and 3 of Figure 3. Overall, the performance of aLPLC2 and mLPLC2 is very similar to the LPLC2 model, showing an extremely selective response to expanding objects. The black object that elicited the strongest response was selected to conduct subsequent experiments.

#### B. Performance Against Shifting Natural Scenes

Hua et al.'s LPLC2 model only responds to expanding objects located at the center of the visual field [15]. In contrast, we explore mLPLC2's response to approaching objects with expanding centers located at various positions. Stimuli consisting of four phases were created to simulate the motion of objects confined to different regions of the visual field (input sampling resolution at  $320 \times 240$ , length at 400 frames, and frequency at 33 Hz). Each phase is characterized by specific moving objects: (1) a dark object located at the center of the visual field; (2) a dark object in the first quadrant; (3) a dark object in the third quadrant; and (4) two dark objects located in the second and fourth quadrants, respectively. Each phase contains 100 frames, with the first 40 frames showing the objects approaching, frames 40 to 60 featuring the objects stationary, and the final 40 frames depicting the objects receding. The background features a real-world scene shifting left at a constant speed. Notably, in the fourth phase, the movements of the two objects occur synchronously. The

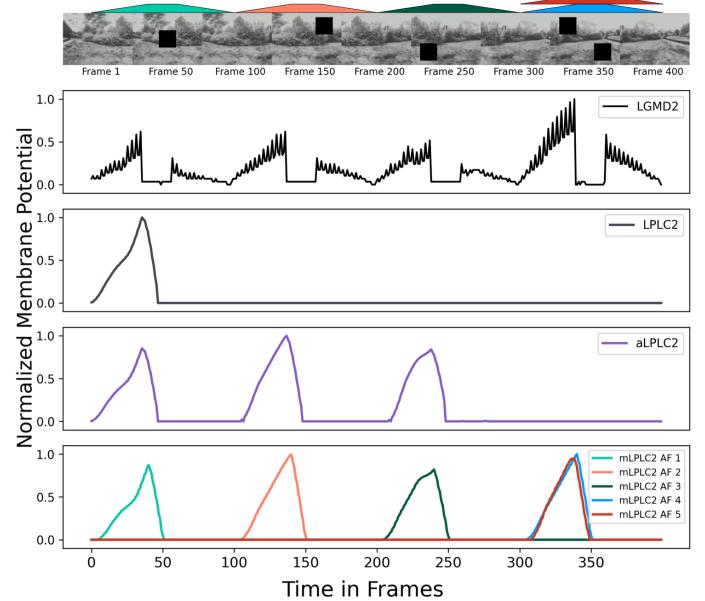


Fig. 4. Different models' responses to objects from various regions of the visual field across four phases. The horizontal and vertical axes represent time in frames and membrane potential, respectively. The widths and heights of the differently colored shapes above the graph indicate the appearance period and size of the respective objects. In the face of multi-target motion, only mLPLC2 can accurately distinguish and identify the looming objects.

results are presented in Figure 4, with key frames and the duration of object presence with image size change over time displayed above.

Based on the comparative experiments, it was observed that under cluttered background interference, the LGMD2 model exhibited a slight response to retreating dark objects. LGMD2 can respond to dark objects in any region of the visual field, although its membrane potential fluctuations are more pronounced compared to the LPLC2-inspired model. However, LGMD2 cannot effectively distinguish the multi-object looming motion in the fourth stage. The ultra-selective LPLC2-inspired model remain silent to the receding inward motion of the black objects in the later stages of each phase. However, LPLC2 model can only fire against the centroid-emitted centrifugal motion, which is insufficient to address the looming objects in stages 2, 3, and 4. The aLPLC2 reactions timely and accurately to outward motion emanated from different visual regions in stages 2 and 3. However, it still struggles to effectively identify and localize multiple simultaneously approaching collision objects in stage 4. In contrast, the mLPLC2 model with multiple-attention mechanism accurately responds to objects emanating from different locations, including two objects approaching simultaneously across various phases. Throughout the entire experimental process, mLPLC2 correctly generated 5 distinct AFs localizing the different looming objects over time.

To further investigate the multi-target detection capability of mLPLC2, we generated more complex simulated stimuli. Six dark objects appeared at frames 5, 15, 25, 45, 55, and 65 in different positions within the global visual field of mLPLC2 then move centrifugally. The background featured a chaotic

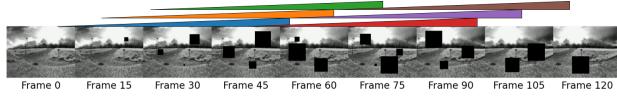


Fig. 5. mLPLC2’s responses to objects originating from different regions of the visual field and at various times, amidst a shifting natural background. The presence period and size of each object are represented by the widths and heights of differently colored shapes at the top of the graph. The center positions and presence periods of AFs are shown in the legend, corresponding closely to the 6 looming objects. In response to these 6 objects, mLPLC2 continuously detects and locks onto multiple looming targets, demonstrating effective responsiveness.

real outdoor scene shifting to the left at a constant speed. The experimental results are presented in Figure 5 with key frames and the duration of object presence displayed above.

When encountering collision objects that exhibit a centripetal outward expansion pattern at various positions and times, mLPLC2 can lock onto these objects within a maximum of 7 frames (23 ms) to establish an AF for each of looming targets. Interestingly, this AF persists for approximately 20 frames (66 ms) after the object disappears. This persistence is attributed to the dynamic background at the object’s location, which induces a luminance change in the Retina layer upon the object’s vanishing, ultimately leading to the excitation of the corresponding AF. This observation is illustrated in the post-peak regions of the curves shown in Figure 5. Due to the influence of high-contrast scenes, the edges of certain collision objects may merge with the background during their expansion, and the edges of multiple simultaneously appearing objects may overlap at certain moments. This complexity results in varying peak responses across different Attention Fields. However, thanks to the ultra-selectivity of LPLC2 neurons and the well-defined attention rules (as outlined in Eq. 15), mLPLC2 remains capable of recognizing these challenging collisions and establishing fatal AFs.

Experimental evidence demonstrates that, in complex dynamic real-world scenes, mLPLC2 can accurately localize and respond to every single looming object.

#### C. Real World Challenges in UAV Scenarios

Finally, we evaluated the practical performance of mLPLC2 using 4 common UAV footage examples available online. These videos include various scenarios where drones en-

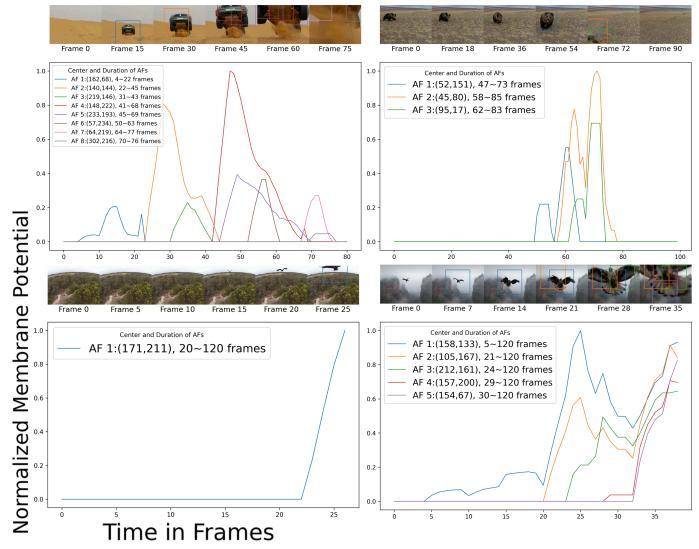


Fig. 6. This figure illustrates mLPLC2’s responses to dangerous colliding scenes captured by UAVs. In each subplot, the x-axis represents time in frames, and the y-axis denotes the normalized membrane potential. The first subplot depicts a drone filming a leaping car. The second subplot shows a drone facing an approaching bear. The third and fourth subplots feature drones confronted by raptors. This figure demonstrates that mLPLC2 effectively detects and localizes real-world hazards encountered by UAVs.

counter hazards, such as approaching predator attacks (see Figure 6).

mLPLC2 demonstrated timely responses to various real-world threats faced by drones. However, two notable observations emerged from the experiments: First, mLPLC2 generates multiple AFs when facing large objects in its field of view. This occurs because the expanding edges of the approaching object surpass the AF of an individual LPLC2 cells receptive field. In such cases, a single AF is insufficient to capture the full expansion motion, prompting the formation of additional AFs along the expanding edges. This effect is evident in the first subplot with the leaping car and in the fourth subplot with the approaching raptor.

Second, mLPLC2 also generates multiple AFs when the expansion center of a moving object shifts. This behavior results from the unique motion integration pattern of LPLC2 cells, which only process motion expanding radially from the AF center in four primary directions. When the expansion center shifts, the original AF can no longer track the motion effectively, leading to the formation of new AFs. This phenomenon is observed in the first and second subplots, where the expansion centers of both the upward-leaping car shift and the running bear, triggering multiple AFs in response.

## V. CONCLUSION AND DISCUSSIONS

This paper introduces an attention-driven mLPLC2 neural ensemble model for continuously detecting and localizing multiple looming targets. Based on the experimental results, the following conclusions can be drawn: (1) The mLPLC2 neural network retains the ultra-selectivity of the LPLC2 model for radial motion expansion. (2) With its multiple-attention mechanism, mLPLC2 accurately and promptly re-

sponds to multiple radial motion objects originating from different focused regions across the visual field.

The model's ultra-selectivity for expanding objects makes it exceptionally well-suited for collision detection tasks. The attention mechanism addresses a key limitation of the original single-LPLC2 model, which was restricted to responding only to objects emanating from the center of the whole visual field. This advancement enables precise, large-scale collision detection and localization across multiple regions, potentially providing vehicles or robots equipped with this system with more detailed and comprehensive collision scenario information. As a result, it facilitates more intelligent and precise avoidance maneuvers during navigation.

The mLPLC2 model's capabilities make it particularly advantageous for UAV applications in three-dimensional motion spaces, where wide visual coverage and precise multi-dimensional collision avoidance actions are essential. However, the mLPLC2 model inherits certain limitations from the biological LPLC2 neurons in *Drosophila*. Since individual LPLC2 neurons are confined to processing motion information within their receptive fields, they cannot track objects moving beyond these local regions. Furthermore, existing nonlinear integration methods heavily rely on the biological structure of LPLC2's dendrite. To address these challenges, future work would explore combining the global motion cue detection of the LGMD model with the mLPLC2 framework. This integration sheds light upon enhancing the system's ability to handle moving and expanding objects across the entire visual field in complex and dynamic scenes.

## REFERENCES

- [1] F. Gabbiani, H. G. Krapp, and G. Laurent, "Computation of object approach by a wide-field, motion-sensitive neuron," *Journal of Neuroscience*, vol. 19, no. 3, pp. 1122–1141, 1999.
- [2] W. Schiff, J. A. Caviness, and J. J. Gibson, "Persistent fear responses in rhesus monkeys to the optical stimulus of 'looming,'" *Science*, vol. 136, no. 3520, pp. 982–983, 1962.
- [3] J. Szarker and D. Tomsic, "Neuronal correlates of the visually elicited escape response of the crab *chasmagnathus* upon seasonal variations, stimuli changes and perceptual alterations," *Journal of Comparative Physiology A*, vol. 194, pp. 587–596, 2008.
- [4] F. T. Muijres, M. J. Elzinga, J. M. Melis, and M. H. Dickinson, "Flies evade looming targets by executing rapid visually directed banked turns," *Science*, vol. 344, no. 6180, pp. 172–177, 2014.
- [5] B. Zhou, Z. Li, S. Kim, J. Lafferty, and D. A. Clark, "Shallow neural networks trained to detect collisions recover features of visual looming-selective neurons," *Elife*, vol. 11, p. e72067, 2022.
- [6] R.-A. Rill and K. B. Faragó, "Collision avoidance using deep learning-based monocular vision," *SN Computer Science*, vol. 2, no. 5, p. 375, 2021.
- [7] L. Lu, G. Fasano, A. Carrio, M. Lei, H. Bavle, and P. Campoy, "A comprehensive survey on non-cooperative collision avoidance for micro aerial vehicles: Sensing and obstacle detection," *Journal of Field Robotics*, vol. 40, no. 6, pp. 1697–1720, 2023.
- [8] S. Yue and F. C. Rind, "A collision detection system for a mobile robot inspired by the locust visual system," in *Proceedings of the 2005 IEEE international conference on robotics and automation*. IEEE, 2005, pp. 3832–3837.
- [9] C. Hu, F. Arvin, C. Xiong, and S. Yue, "Bio-inspired embedded vision system for autonomous micro-robots: The lgmd case," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 9, no. 3, pp. 241–254, 2016.
- [10] Q. Fu, H. Wang, C. Hu, and S. Yue, "Towards computational models and applications of insect visual systems for motion perception: A review," *Artificial life*, vol. 25, no. 3, pp. 263–311, 2019.
- [11] L. Li, Z. Zhang, and J. Lu, "Artificial fly visual joint perception neural network inspired by multiple-regional collision detection," *Neural Networks*, vol. 135, pp. 13–28, 2021.
- [12] B. Gu, J. Feng, and Z. Song, "Looming detection in complex dynamic visual scenes by interneuronal coordination of motion and feature pathways," *Advanced Intelligent Systems*, vol. 6, no. 9, p. 2400198, 2024.
- [13] N. C. Klapoetke, A. Nern, M. Y. Peek, E. M. Rogers, P. Breads, G. M. Rubin, M. B. Reiser, and G. M. Card, "Ultra-selective looming detection from radial motion opponency," *Nature*, vol. 551, no. 7679, pp. 237–241, 2017.
- [14] N. C. Klapoetke, A. Nern, E. M. Rogers, G. M. Rubin, M. B. Reiser, and G. M. Card, "A functionally ordered visual feature map in the drosophila brain," *Neuron*, vol. 110, no. 10, pp. 1700–1711, 2022.
- [15] M. Hua, Q. Fu, J. Peng, S. Yue, and H. Luan, "Shaping the ultra-selectivity of a looming detection neural network from non-linear correlation of radial motion," in *2022 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2022, pp. 1–8.
- [16] J. Zhao, S. Xi, Y. Li, A. Guo, and Z. Wu, "A fly inspired solution to looming detection for collision avoidance," *IScience*, vol. 26, no. 4, 2023.
- [17] F. C. Rind and D. Bramwell, "Neural network based on the input organization of an identified neuron signaling impending collision," *Journal of neurophysiology*, vol. 75, no. 3, pp. 967–985, 1996.
- [18] M. Oshea and C. F. Rowell, "The neuronal basis of a sensory analyser, the acridid movement detector system. ii. response decrement, convergence, and the nature of the excitatory afferents to the fan-like dendrites of the lgmd," *Journal of Experimental Biology*, vol. 65, no. 2, pp. 289–308, 1976.
- [19] Q. Fu and S. Yue, "Modelling LGMD2 visual neuron system," in *2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP)*. IEEE, 2015, pp. 1–6.
- [20] Q. Fu, "Motion perception based on ON/OFF channels: A survey," *Neural Networks*, vol. 165, pp. 1–18, 2023.
- [21] M. S. Maisak, J. Haag, G. Ammer, E. Serbe, M. Meier, A. Leonhardt, T. Schilling, A. Bahl, G. M. Rubin, A. Nern *et al.*, "A directional tuning map of drosophila elementary motion detectors," *Nature*, vol. 500, no. 7461, pp. 212–216, 2013.
- [22] A. Borst and L. N. Groschner, "How flies see motion," *Annual review of neuroscience*, vol. 46, no. 1, pp. 17–37, 2023.
- [23] Q. Fu and S. Yue, "Modelling Drosophila motion vision pathways for decoding the direction of translating objects against cluttered moving backgrounds," *Biological Cybernetics*, vol. 114, no. 4, pp. 443–460, 2020.
- [24] I. A. Meinertzhagen and S. O'neil, "Synaptic organization of columnar elements in the lamina of the wild type in drosophila melanogaster," *Journal of comparative neurology*, vol. 305, no. 2, pp. 232–263, 1991.
- [25] M. Joesch, B. Schnell, S. V. Raghu, D. F. Reiff, and A. Borst, "On and off pathways in drosophila motion vision," *Nature*, vol. 468, no. 7321, pp. 300–304, 2010.
- [26] A. Borst, J. Haag, and A. S. Mauss, "How fly neurons compute the direction of visual motion," *Journal of Comparative Physiology A*, vol. 206, no. 2, pp. 109–124, 2020.
- [27] Q. Fu, J. Peng, and S. Yue, "Bioinspired contrast vision computation for robust motion estimation against natural signals," in *2021 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2021, pp. 1–8.
- [28] Q. Fu, Z. Li, and J. Peng, "Harmonizing motion and contrast vision for robust looming detection," *Array*, vol. 17, p. 100272, 2023.
- [29] D. A. Clark, J. E. Fitzgerald, J. M. Ales, D. M. Gohl, M. A. Silies, A. M. Norcia, and T. R. Clandinin, "Flies and humans share a motion estimation strategy that exploits natural scene statistics," *Nature Neuroscience*, vol. 17, no. 2, pp. 296–303, 2014.
- [30] B. Hassenstein and W. Reichardt, "Systemtheoretische analyse der zeit-, reihenfolgen-und vorzeichenauswertung bei der bewegungsperzeption des rüsselkäfers chlorophanus," *Zeitschrift für Naturforschung B*, vol. 11, no. 9–10, pp. 513–524, 1956.
- [31] A. Borst and M. Egelhaaf, "Principles of visual motion detection," *Trends in neurosciences*, vol. 12, no. 8, pp. 297–306, 1989.
- [32] H. Eichner, M. Joesch, B. Schnell, D. F. Reiff, and A. Borst, "Internal structure of the fly elementary motion detector," *Neuron*, vol. 70, no. 6, pp. 1155–1164, 2011.