WILEY | Hindawi

Research Article

Panoramic Motion Perception Inspired Fly Visual Brain Joint Neural Network on Omnidirectional Collision Detection

Lun Li, Zhuhong Zhang [ID], Wensheng Jia [ID], and Jiaxuan Lu

College of Big Data and Information Engineering, Guizhou University, Guizhou Provincial Characteristic Key Laboratory of System Optimization and Scientific Computing, Guiyang, Guizhou 550025, China

Correspondence should be addressed to Zhuhong Zhang; zhzhang@gzu.edu.cn

Biological systems have a great number of visual motion detection neurons, some of which can preferentially react to specific visual regions. Nevertheless, little work has been performed about how they can be used to develop neural network models for omnidirectional collision detection. Hereby, an artificial fly visual brain neural network with presynaptic and postsynaptic subnetworks, for the first time, is developed to detect the changes of visual motion in panoramic scenes. Herein, the presynaptic subnetwork, which originates from the preferential response characteristics of five fly visual neurons, responds to all the moving objects in the panoramic field; the postsynaptic network, which is based on the properties of the angle and height detection neurons in the fly's brain system, collects the excitatory intensities of the visual neurons, and outputs the real-time activities of the main object closest to the panoramic camera. Hereafter, a computational model is constructed to implement omnidirectional collision detection, relying upon the artificial visual brain neural network and three functional neurons. The theoretical analysis has verified that the collision detection model's computational complexity depends mainly on the image input resolution. Three experimental conclusions can be clearly drawn: (i) the motion characteristics of the main object in the panoramic environment can be clearly exhibited in the neural network; (ii) the collision detection model can not only outperform the compared models but also successfully perform omnidirectional collision detection; and (iii) it spends 0.24 s or so to execute visual information processing per frame with the resolution of 120×80.

1. Introduction

As a classic and popular topic, collision detection often touches upon many of engineering problems, e.g., vehicle collision [1], navigation [2], and mobile robot [3]. Even though radars and ultrasonic sensors can monitor whether a moving object encounters possible dangers, their collision detection accuracy depends strongly on the quality of the hardware measurement. Thereby, they necessarily cause economically expensive costs in engineering applications, due to the requirement of high-precision collision detection. This motivates researchers to probe into bioinspired, low-cost, and high-precision collision detection models from the perspective of artificial intelligence.

Biological systems cannot only capture visual information, but also rapidly discover moving objects and accurately judge if a danger occurs. Particularly, flies, which possess such merits as high-speed motion, prompt response, and collision decision-making, can detect spatiotemporal intensities and avoid imminent collision. They may be viewed as an optical instrument to synthesize visual information and can perform visual information processing and decision-making by means of visual brain neurons [4–9]. This is a nice source for constructing artificial visual brain models to solve specific engineering problems. However, can the intrinsic characteristics of the visual neurons be employed to construct neurocomputational models for detecting the behavior changes of the moving objects in the horizontal, vertical, and equatorial directions? Can the response characteristics of the visual brain neurons be jointly employed to construct artificial visual brain systems for panoramic collision detection? These motivate us to

explore a new neurocomputational model for omnidirectional collision detection, by means of the visual response mechanisms of the horizontal, vertical, and equatorial motion detection neurons in the fly visual system and those of the central complex and mushroom body in the fly's brain system.

The current work concentrates on exploring bioinspired computational models which are required to not only detect changes in panoramic motion but also transmit real-time collision warning signals. The main contributions are summarized as follows:

(i) Based on the preferential response properties of the five direction-selective neurons in the fly visual system, a feedforward presynaptic neural network with five subnetworks is originally constructed to check visual changes in the panoramic scene. Therein, the response characteristics of the backward and equatorial motion detection neurons are, for the first time, borrowed to develop two visual neural networks, in order to detect visual changes in the lower half field of the panoramic scene and on the horizontal equator. Two new symmetric inhibition models are designed to execute the function of lateral inhibition between the neural nodes. Also, a motion direction detection region is skillfully divided into multiple response subregions in terms of two symmetric parabolic curves.

(ii) Originated from the characteristics of the angle and height detection neurons and the response mechanism of processing visual neural signals in the fly's brain system, a postsynaptic network is originally constructed to characterize the morphological changes of the main object closest to the panoramic camera.

(iii) Inspired by the characteristics of the joint perception and response in the fly visual brain system, the abovementioned presynaptic and postsynaptic neural networks are combined to construct a fly visual brain joint neural network (FVBJNN). Thereafter, inspired by spike-transmitting mechanisms in the brain system, a FVBJNN-based fly visual brain joint omnidirectional collision detection (FVBJOCD) model, for the first time, is constructed to implement omnidirectional collision early warning in the panoramic scene.

To our best knowledge, no work has been conducted to discuss the omnidirectional collision detection in terms of bio-inspired collision detection models. FVBJNN as an artificial visual brain neural network differs from any reported computational model and especially our models, i.e., AFVNN and AFVJPNN [8, 9], due to their biological inspirations and different problems to be solved. The two reported models can only catch the change of visual movement in the case where one or more objects move on the upper half of the plane of the panoramic field, but so does FVBJNN in the panoramic field, and particularly can detect the change of an object's equatorial

motion. On the other hand, FVBJNN differs differs from our previous model (RMPNN) [10] and those (LGMDs) developed by Yue's group [11, 12], since the two models originate from locust visual neurophysiologic discoveries.

The rest of the paper is organized as follows. Section 2 gives some reviews on related work. Section 3 presents some fly neurophysiologic findings, by which an artificial visual brain model is drawn to display the fly's visual response characteristics. FVBJNN is designed in Section 4, which involves its architecture and neural layer design. Section 5 develops the model of the FVBJOCD mentioned above and formulates the related algorithm. The experiments are exhaustively performed in Section 6. Some conclusions are drawn in Section 7.

2. Survey on Related Work

2.1. Fly's Visual Neuron. Since two neurophysiologists, Bishop and Keehn [13], discovered motion detection neurons in the fly visual system, electro- and neurophysiologists [14–18] have comprehensively probed into the relation between visual neurons and acquired multiple directional detection neurons capable of detecting the direction of visual movement. Particularly, two specific neurons of $T4$ and $T5$ have been verified to be able to detect the excitatory or inhibitory activities of direction-selective neurons in the four cardinal directions [5]. They can respond to visual movements by the ON and OFF channels as in the lamina layer when an object moves with strong or weak brightness in the field of view. Besides, there exists a wide-field neuron which can send feedback signals to the lamina layer and dynamically adjust visual information processing [19]. Keles and Frye [20]claimed that the neuron LC11 could specifically react to the dynamic small objects in a dynamic and panoramic scene. In the same year, Dickson's group [21] revealed that the fly's lobula columnar neurons could preferentially react to the visual stimuli of the imminent looming object. Such two findings indicated that the fly's brain neural system could receive and process the changes of visual nerve signals by means of visual neurons' motion detection activities [5, 22–24]. Furthermore, Takemura et al. [25] constructed a connectome circuit model and proved that the connection of motion-sensitive neurons was consistent with the neurons' preference direction. Subsequently, Shinomiya et al. [26, 27] studied and simulated the properties of edge motion detection of several functional neurons in terms of six elementary motion detectors. Recently, a computational model, discovered by Tanaka and Clark [28] was developed to quickly decide whether the fly encountered a small object. This can provide us key insights into exploring neurocomputational models.

2.2. Fly's Brain Behavior. The fly brain system can provide neuroscientists with abundant insights to discuss its behavior characteristics, due to the abilities of learning, memory, and prompt decision-making [29–31]. Therein, the central complex (CX) and mushroom body (MB) in the

central nervous system include crucial units such as the protocerebral bridge (PB) and fan-shaped body (FB). [32]. After Kohl et al. [33] pointed out that the fly brain's connectomic discipline should aim to discover the complete wiring diagram of the brain, several research groups [34–36] exhaustively carried out the related studies of PB and FB. For example, Pfeiffer and Homberg [36] asserted that CX as the command center of flight, learning, memory, and autonomous attitude control [22] was a functional organization with the structural array of sixteen slices on the left and right axes.

To simulate the behavioral attributes of the fly's brain, Shih et al. [37] presented a neurocomputational model that involved the response mechanisms of the premotor center of the central brain and four sensory modality centers, by which the brain neural structure was revealed. Lin et al. [38] reconstructed the circuit of PB and transformed it into a common three-dimensional framework. Subsequently, some works [39, 40] validated that MB, as the long-term memory center of the fly's olfactory sense, was a crucial tissue in visual learning and memory. Additionally, Hu et al. [35] found the nociceptive mechanism in the central brain. Their experiments verified that the central nervous system played a decisive role in processing nociceptive information and direct avoidance behavior, particularly the elliptic body could control its spatial navigation. Although these neurophysiologic achievements are valuable in identifying the functions of the brain, it remains unclear how the neural circuits of motor stimulation and decision-making can be characterized by computational models.

2.3. Artificial Neurocomputational Model. The fly's visual brain system dominates its motion behavior. By an analog to the functional mechanisms of its visual information processing and decision-making behaviors, some computational models have been developed to solve engineering examples such as motion detection, object navigation, object tracking, as well as collision avoidance.

2.3.1. Motion Detection. Motion detection involves how to estimate and track the object's location, gesture, motion speed, etc. Since Hassenstein and Reichardt [41] constructed an elementary motion detector (EMD) to check the change of horizontal visual movement; a series of valuable computational models suitable for navigation and object tracking have been reported [42–50], inspired by the biovisual information processing mechanisms of different living organisms such as flies and locust. Especially, the fly visual system as a valuable organism is a natural inspiration for detecting the state of moving objects. Öğmen and Gagné [42] designed two fly visual neural models to perceive the motion cues of objects. Afterwards, Eichner et al. [43] proposed multiple kinds of motion detectors by virtue of the fly's optic nerve structure. Shiozaki and Kazama [44] acquired two independent neural processing circuits in the brain of the fly that could encode multiple clues in the process of navigation and decision-making. Therein, some reported neurocomputational models can be used to estimate the angular and velocity of a moving object [45] and control the object's navigation [46]. Missler et al. [47] initially developed a fly visual neural network which is able to evaluate the position and speed of a small moving object in the black or white background. Thereafter, some object tracking models were reported one after another [48–50]. For example, Sun et al. [48] proposed a computational model to perform face recognition and object tracking, and later, claimed that the model could track any type of object in the spatiotemporal environment. Also, Aptekar et al. [50] constructed a figure-based fly visual tracking model to predict an object's motion behavior. Additionally, Yue's group [51, 52] designed a visual tracking model which comprised of a hybrid visual neural network and a tracking detection scheme in terms of the visual nerve structures of flies and locusts, and later, an autonomous micro biological robot was adopted to detect the model's object tracking performance.

2.3.2. Local-Regional Collision Detection. As another challenging topic in computer vision, collision detection involves how to probe into bio-visual models which is able to perceive imminent objects and transmit collision early warning signals in real time. Therein, visual neurophysiologic findings have validated that the fly will be in an excitatory state and output high excitatory activities if an object approaches it. This can provide us with enough inspirations for exploring collision detection-related neurocomputational models. Herein, some groups have constructed several bio-inspired models in terms of visual information processing mechanisms of living organisms. Yue's group [11, 12, 53–56] proposed multiple kinds of locust visual neural networks, based on the properties of direction-sensitive neurons in the locust vision system. Particularly, Fu et al. [12] suggested two artificial locust visual neural networks in terms of the locust visual nerve mechanism and small object detectors. Additionally, inspired by visual information processing mechanisms of other insects such as hippocamp, pigeon, and frog, several collision detection models were reported [57, 58] as well. For example, the looming-sensitive neurons in the pigeon's brain could be used to construct collision detection models. Moreover, our group [8, 9] intensively discussed the topic of collision detection and acquired two computational models suitable for detecting the motion characteristics of the moving objects in the wide or narrow visual scene, relying upon the typical response mechanisms of the fly visual system.

2.3.3. Omnidirectional Collision Detection. Usually, achieved by panoramic cameras, mobile robots, or multi-camera systems [59], panoramic collision detection touches upon how to capture the visual clues in panoramic scenes and detect possible hazards around fixed or moving objects. Some initial works have involved omnidirectional position estimation and object tracking [60–63]. For example, Shimizuhira et al. [60] constructed an artificial vision system to estimate the objects' location by an autonomous mobile robot with three cameras; Hirabayashi et al. [61] suggested

an omnidirectional detection system, relying upon the combination of image division and multiple object detection methods; and Taheri et al. [62] reviewed the progress of omnidirectional wheel detection and also put forward some challenging topics on robotic navigation. To our best knowledge, there has been no bio-inspired omnidirectional collision detection model or system. Therefore, related to the factors of cost, effectiveness, and efficiency, it is desired to construct omnidirectional collision detection models which can detect the behavior changes of moving objects and avoid possible collision.

3. Simplified Fly Visual Brain Model

The fly visual system comprises two striking compound eyes, each with ommatidia arranged hexagonally. Each ommatidium, which includes eight photoreceptors (R1-R8) [32, 37, 64], consists of photoreceptors. The main role of the visual system consists in capturing visual signals and hierarchically performing visual information processing through five neural layers: photoreceptor, retina, lamina, medulla, and lobula plate [36, 65].

The photoreceptor layer receives external visual signals and perceives the change of visual movement. Also, it transforms visual optical flow into neural signals and transmits them to the retina layer.

The retina layer consists of ommatidia, each of which as an optic apparatus not only continually receives neural signals at the corresponding photoreceptor in the photoreceptor layer but also checks if there exist visual movements. Also, it needs to implement figure-ground discrimination and depolarization. This implicates that retina nodes can carry out noise elimination and object extraction. However, each ommatidium is required to deliver its membrane potential or activity to the counterpart in the next layer after depolarization processing [4, 66].

The lamina layer, a crucial optic flow processing layer with lamina units, not only receives the ommatidia's neural signals in the retina layer and synthesizes them into excitatory or inhibitory activities by projection, but also executes lateral inhibition between the neural nodes. Precisely, it is mainly composed of cartridge (cart) and on-off (oo) nodes [47]. Each cart node takes on neural signal synthesis and outputs an excitatory or inhibitory activity at each moment by signal projection. Nevertheless, each oo node performs lateral inhibition based on its adjacent nodes, while extracting visual signals by means of ON and OFF pathways.

The medulla layer, which consists of m nodes, is a local motion direction detection layer. It generates the local motion-directional activities of the nodes through a classical correlative detector. Subsequently, these nodes produce local directional activities and output them to the lobula plate.

The lobula neurons in the lobula platelayer preferentially collect the local activities produced by the medulla layer, and later, output its own activities to measure the global direction changes of one or more objects. Herein, it is pointed out that the lobula platelayer (LP) comprises about 60 neurons, among which the lobula plate tangential cells (LPTCs), e.g., H1, H3, V1, V2, and EH, are used to extract visual information from the optic flow [17, 67]. H1 preferentially react to ipsilateral forward motion in the horizontal direction, but inhibits reverse motion. H3 can only respond to ipsilateral backward motion. Besides, V1 is a spike detection neuron capable of selectively responding to ipsilaterally downward movements in the vertical direction. Conversely, V2 as a spike-transmitting neuron can only react to ipsilateral upward motion specifically. Additionally, EH, an equatorial motion detection neuron, makes a binocular response to equatorial motion. Subsequently, each motion-sensitive neuron, together with the abovementioned four visual neural layers, forms a visual information processing circuit.

However, the fly's brain, which includes three main functional parts: forebrain, central brain, and tritocerebrum, can exhibit a series of complex motion behaviors, e.g., collision detection and learning. Especially, based on the introduction on the fly's brain behavior in Section 2.2, the forebrain as the largest component of the fly's brain includes MB and CX. Meanwhile, PB and FB as high-level activity areas control the whole behavior of the fly and regulate its information integration, learning, memory, movement control, etc. [35, 68, 69]. Therein, PB includes sixteen neural nodes evenly distributed on a sine-like curve, based on the neurophysiologic findings. Once PB and FB receive excitatory activities produced by visual neurons, they will evaluate whether there is a danger. If yes, they will make a decision to avoid possible collision jointly. The angle and height detection neurons F1 and F5 in the first and fifth layers of FB are responsible for learning and memorizing the morphological angle and vertical height of some moving object, respectively.

Summarily, the process of the fly's visual brain information processing includes two stages. One is to generate the excitatory activities of visual neurons based on the visual information processing mechanism, the other is to decide whether to avoid the current obstacles after integrating the acquired activities into the brain's nerve system. The simplified visual brain information processing mechanism can be formulated by Figure 1.

4. Fly Visual Brain Joint Neural Network

This section firstly elaborates the medulla layer's regional division scheme. Secondly, related to one such scheme, the architecture of the neural network (FVBJNN) is formulated in Section 4.2. Finally, its neural layers are designed in Sections 4.3 and 4.4, based on an analogy to the functional properties of the above neural layers.

4.1. Medulla Region Division. As mentioned in Section 3, H1 and H3 preferentially respond to local motion activities produced by medulla nodes in the left-to-right (LR) and right-to-left (RL) subregions as in Figure 2, respectively, while V1 and V2 do so orderly in the up-to-down (UD) and back-to-forward (BF) subregions. Also, EH only prefers to respond to local motion-directional activities generated by the medulla nodes on the equatorial trajectory. These show that H1, H3, V1, and V2, orderly named LRN, RLN, UDN, and BFN, can only react to visual movements in their
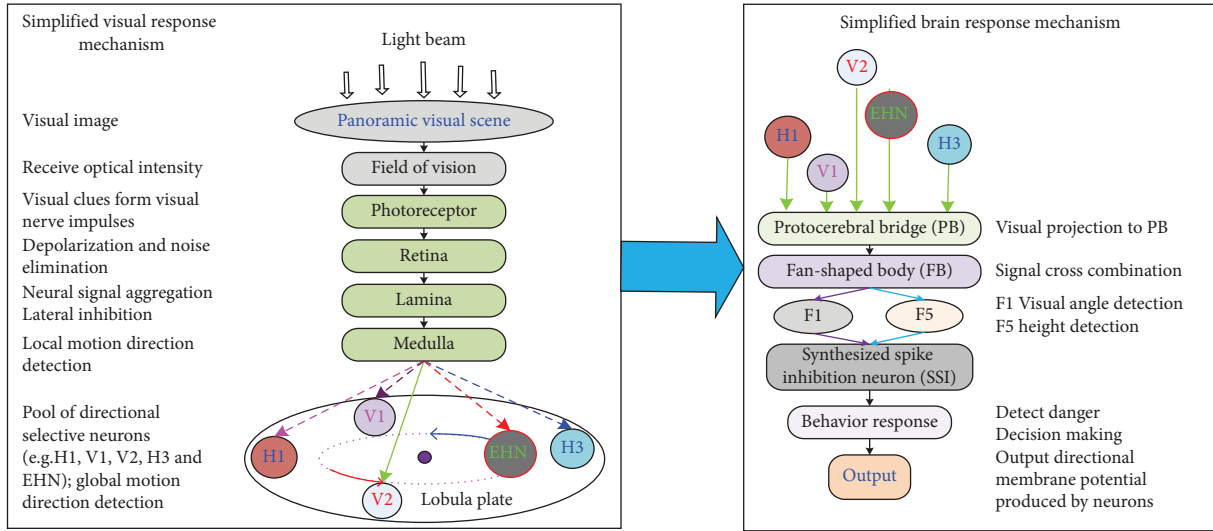
FIGURE 1: Illustrative diagram on the fly's information transmission. Herein, once a visual scene is captured in the panoramic field, the fly visual system executes information processing through five visual neural layers, in which each neuron (e.g., H1, H3, V1, V2, or EH) only reacts to its preferential region in the medulla layer, while extracting visual characteristics in its preferential sub region. Subsequently, the extracted visual characteristics are transmitted to the central complex in the brain neural system and then sent to the functional neurons after symmetric cross-transmission processing, e.g., neurons F1 and F5 in the sectoral body. After that, some brain neurons interact with each other and detect possible dangers.
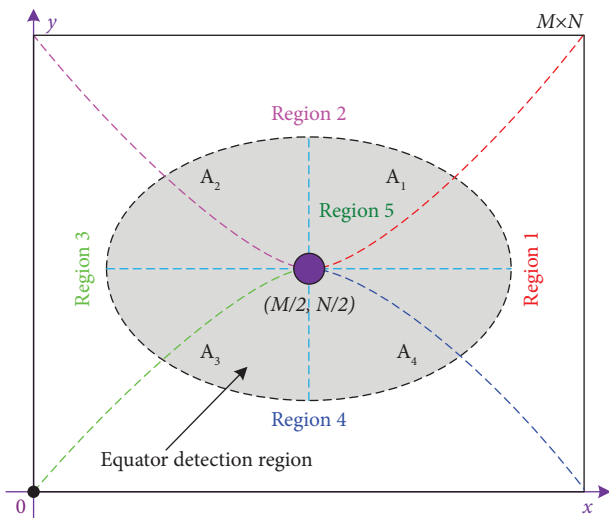

FIGURE 2: Illustrative diagram on the medulla layer's regional division. Wherein, the whole medulla layer expressed by a rectangle region is divided into four subregions (regions 1–4) by two parabolic curves. Furthermore, an elliptic curve is approximately regarded as an equatorial one, after which the elliptic region is segmented into four subregions (A_1, A_2, A_3, and A_4) by the two cardinal lines.

preferential regions, and meanwhile EH called EHN only pays close attention to the changes of visual motion on the equatorial trajectory. Based on such an inspiration, the medulla layer with $M \times N$ nodes, represented by a rectangle in Figure 2 is divided into four subregions (regions k, $1 \leq k \leq 4$) in terms of two symmetrical parabolic curves. The upward parabolic curve passes through three coordinate vertices of $(M/2, N/2)$, $(M, 0)$, and (M, N), while the

downward one goes through $(M/2, N/2)$, $(0, 0)$, and $(0, N)$. LRN and RLN respond to the nodes' activities in Regions 3 and 1, respectively, whereas so do UDN and BFN in Regions 2 and 4 in turn. In addition, a horizontal elliptic curve with the center $(M/2, N/2)$ as well as the long axis a and the short axis b are chosen to approximately formulate the equatorial curve. The equatorial region, i.e., Region 5, is divided into four subregions, since EHN is required to detect visual motion changes on the equator.

4.2. The Topological Architecture on FVBJNN. Related to Figures 1 and 2, FVBJNN, schematically illustrated by Figure 3 below consists of presynaptic and postsynaptic networks. The presynaptic neural network, which is based on the mechanism of visual response and the preferential response characteristics of the aforementioned visual neurons (LRN, RLN, UDN, and BFN), comprises five fly visual neural networks (FVNNs) with equal importance, i.e., LR-, RL-, UD-, BF-, and EH-FVNNs. The former four networks, LR-, RL-, UD-, and BF-FVNNs, are orderly employed to detect the changes of LR, RL, UD, and BF visual motion in the panoramic field, but so does the latter one (BF-FVNN) on the horizontal equatorial trajectory. They share three sequentially connected visual information processing layers (photoreceptor (P), retina (R), and lamina (L)), but include different neurons and response subregions as in the medulla layer (M-layer), namely, LR-, RL-, UD-, BF-, and EH-FVNNs orderly include subregions 3, 1, 2, and 4 and the equatorial curve that are consistent with those in Figure 2 and also involve in the neurons of LRN, RLN, UDN, BFN, and EHN, respectively. LR- and RL-FVNNs are designed to orderly measure visual motion changes in the left and right visual regions as in the panoramic scene. UD- and BF-FVNNs detect visual motion changes in the upper and backward visual

regions, respectively. EH-FVNN is specifically responsible for detecting the change of visual movement on the equatorial curve, relying upon the related equatorial region. On the other hand, the postsynaptic neural network consists of two neural layers PB and FB as well as two neurons F1 and F5. The PB-layer comprises sixteen nodes (PB$_i$, $1 \leq i \leq 16$) which receive the excitatory activities of the visual neurons in the presynaptic neural network and transform them into spike signals. The FB-layer includes eight nodes (FB$_i$, $1 \leq i \leq 8$), each of which collects the outputs of the two symmetrically arranged PB nodes and changes them into pulse signals through a spike-transmitting mechanism. Finally, the neurons of F1 and F5 gather the outputs of the nodes in the FB-layer and output the morphological angle and vertical height of the main moving object closest to the panoramic camera.

4.3. Presynaptic Neural Network.

Recall that the presynaptic neural network consists of LR-, RL-, UD-, BF-, and EH-FVNNs in Figure 3(a). Herein, we take BF- and EH-FVNNs for example to explicitly formulate their design inspirations. Their schematics are given in Figure 4.

4.3.1. BF-FVNN.

The subnetwork incorporates an output neuron (BFN), three neural layers (P-, R- and L-layers), and the medulla subregion in Figure 3(a) that corresponds to the region 4 in Figure 2, among which the P-layer is the same as that in our previous work [5]. The detailed designs are summarized below.

(1) P-Layer. This layer captures the illuminance intensity or grayscale at each pixel point in the current grayscale frame, while filtering out additional noise signals, in order to extract one or more moving object (s) in the panoramic scene. Precisely, let the layer include $M \times N$ photoreceptors displayed in a matrix form; $J_f(i, j)$ stands for the luminance intensity at the pixel (i, j) in the frame f with size $M \times N$. Each photoreceptor coincides with a matched pixel point in the frame. The ViBe method [69] is used to create a foreground image of G_f with size $M \times N$ in terms of J_f, by which the output of the photoreceptor (i, j), $H_f(i, j)$, is computed by the following equation:

$$H_f(i, j) = \begin{cases} G_f(i, j), & \text{if } G_f(i, j) = J_f(i, j), \\ 0, & \text{else,} \end{cases} \quad (1)$$

with $1 \leq i \leq M$ and $1 \leq j \leq N$. Equation (1) shows that those valuable visual signals in J_f are extracted to reflect the behavior changes of some moving objects.

(2) R-Layer. The layer comprises $M \times N$ retina nodes with the same structure as that in the P-layer. Each node (i, j) captures the neural signal $H_f(i, j)$ at frame f and processes it by a smoothening signal processing mechanism, namely,

$$P_f(i, j) = H_f(i, j) + \sum_{l=1}^{n_{st}} v_l H_{f-l}(i, j), 1 \leq i \leq M, 1 \leq j \leq N. \quad (2)$$

with maximal time step n_{st}, where v_l as a persistence coefficient is decided by the standard sigmoid function, i.e., $v_l = \text{Sig}(l)$. Equation (2) emerges that, once one or more objects are present in the panoramic scene, some retina nodes will stay excited for a short time, based on the historical visual signals.

(3) L-Layer. The layer includes two sequentially connected visual neural layers, i.e., the cartridge (cart) and on-off (oo) ones, each of which has $M \times N$ nodes with the same structure as that in the R-layer and performs signal projection, filtering, or shunting inhibition. In the cart layer, each cart node (i, j) receives neural signals at the retina nodes around the counterpart in the R-layer and synthesizes them to output an activity of $R_f(i, j)$ in terms of signal projection. Based on one such information processing mechanism, the conventional Gaussian filter is used to produce the activity of the node (i, j), namely,

$$R_f(i, j) = \sum_{0 \leq k, l \leq 2} P_f(i + k - 1, j + l - 1) w_{kl}, \quad (3)$$

with $1 \leq i \leq M, 1 \leq j \leq N$, where w_{kl} denotes the element at the position (k, l) in the convolution kernel matrix $\mathbf{W}$ given by the following equation:

$$\mathbf{W} = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}. \quad (4)$$

Thereafter, since the node (i, j) needs to depolarize the low activity outputted by its counterpart in the R-layer, a first-order low-pass filter is adopted to output its activity, namely,

$$\frac{dC_t(i, j)}{dt} = \frac{1}{\tau} [R_t(i, j) - C_t(i, j)], \quad (5)$$

with positive speed attenuation factor τ.

Besides, each oo node (i, j) in the oo layer transforms the activity of the counterpart in the cart layer into an oo signal of $O_t(i, j)$ by means of the shunting inhibition mechanism [6], as follows:

$$\frac{dO_t(i, j)}{dt} = -AO_t(i, j) + (B - O_t(i, j)) \\ \times (C_t(i, j) - g(C_t(i, j))), \quad (6)$$

with positive decay factor A and excitation amplitude B, where $g(.)$ is the one-frame time-delay function. Equation (5) shows that, if $O_t(i, j)$ becomes large with time t, the oo node (i, j) will stay excited and encounter inhibition conversely, which can characterize the node's response behaviors.

(4) M-Layer. The layer, which comprises of $M \times N$ medulla nodes with the same structure as that in the oo layer, preferentially pays close attention to the changes of local motion of the medulla nodes in the backward subregion (Region 4) as in Figure 2. Each node in the region receives two activities of excitation and inhibition outputted by the
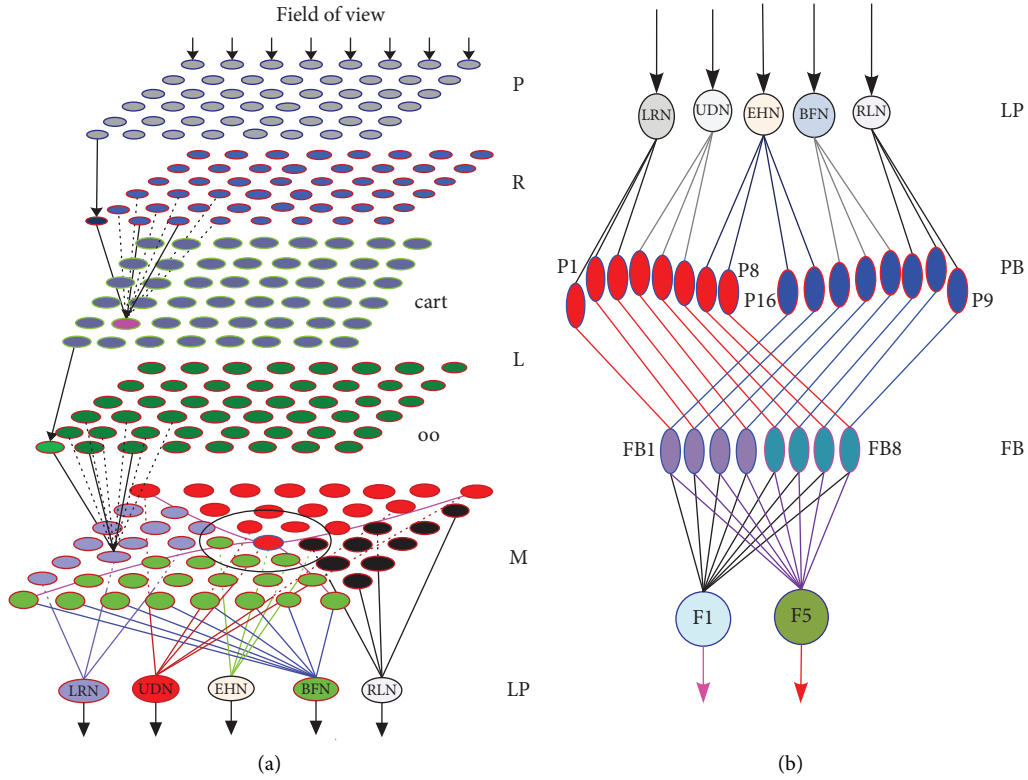
FIGURE 3: Schematic illustration of the FVBJNN's architecture. FVBJNN comprises of the presynaptic and postsynaptic neural networks. (a) Presynaptic neural network is acquired by simply simulating the neural mechanism of visual response in Figure 1, but (b) postsynaptic neural network originates from the fly's brain response mechanism. In the schematic diagram (a), the symbols of P, R, L, M, and LP are the abbreviations of photoreceptor, retina, lamina, medulla, and lobula plate, respectively. The L-layer is formed of two sublayers, i.e., cartridge (cart) node and on-off (oo) node sublayers. It is highlighted that the M-layer is completely consistent with that in Figure 2. Once a grayscale frame is inputted in to the P-layer, it is processed by the five neural layers hierarchically. Afterwards, LRN, RLN, UDN, BFN, and EHN independently generate motion-directional activities and transmit such activities to the PB-layer. The FB-layer receives the excitatory activities generated by the PB-layer and later implements a symmetric cross-transfer mechanism. Finally, the excitatory activities produced by the FB-layer are simultaneously transmitted to F1 and F5, by which some important neural clues are extracted to detect omnidirectional collision. (a) Presynaptic neural network. (b) Postsynaptic neural network.

counterpart in the oo layer and later generates an activity to evaluate its state by the elementary motion detector (EMD) in Figure 5(a) and the lateral inhibition mechanism.

Precisely, each node (i, j) inhibits the five lower nodes in its 3×3 neighborhood as in Figure 5(b), and outputs an excitatory activity by the following equation:

$$E_f^{BF}(i, j) = o_f(i, j) - w \frac{1}{5} \sum_{0 \leq k, l \leq 2} o_f(i + k, j + l) g(o_f(i + k, j + l)), \tag{7}$$

with $k \neq 2$, $(k, l) \neq (1, 1)$, $1 \leq i \leq M/2$, and $n_l^{BF}(i) \leq j \leq n_r^{BF}(i)$, where $n_l^{BF}(i)$ and $n_r^{BF}(i)$ are the column numbers of the left most and right most medulla nodes in the ith row as in region 4, respectively; w is a local inhibition weight of the

node's signal strength in the nonpreferred direction. On the other hand, the node (i, j) is inhibited by the five lower nodes in its 3×3 neighborhood as in Figure 5(b) and outputs an inhibitory activity by the following equation:

$$I_f^{BF}(i, j) = w \left(o_f(i, j) - \frac{1}{5} \sum_{0 \leq k, l \leq 2} o_f(i - k, j + l) g(o_f(i - k, j + l)) \right), \tag{8}$$
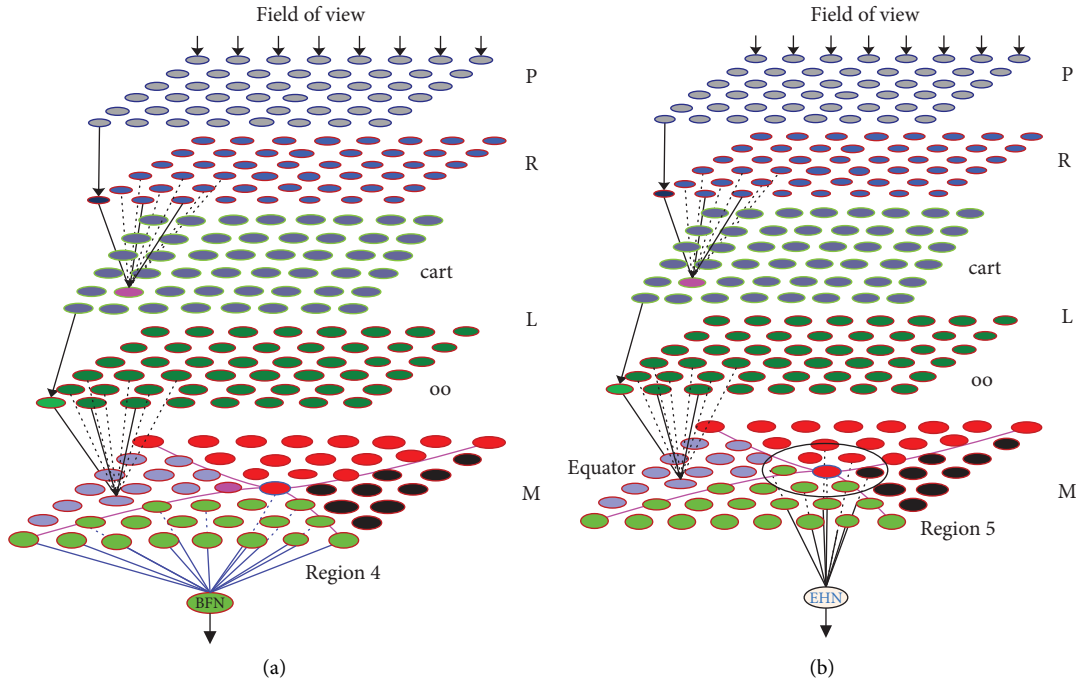
FIGURE 4: The schematic diagrams of the illustrative neural networks. (a) The schematic of BF-FVNN. (b) The schematic of EH-FVNN.
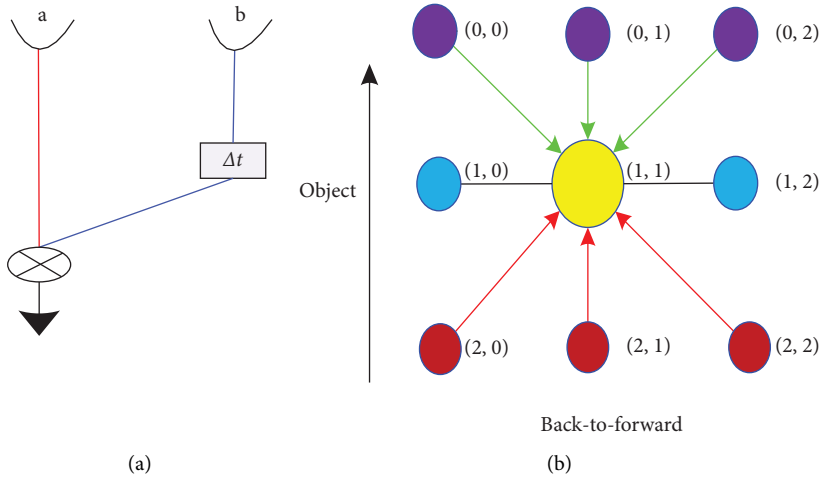

FIGURE 5: Schematic illustration on EMD and lateral inhibition. (a) A unidirectional motion detector which detects the direction of a single moving object. (b) The schematic of the lateral inhibition mechanism; when an object moves along the direction from bottom to top, the central node (1, 1) is activated by the five nodes in the lower half-plane but inhibited by the five nodes in the upper half-planer. (a)EMD [41]. (b) Lateral inhibition between nodes.

with $k \neq 0$, $(k, l) \neq (1, 1)$. Accordingly, the node produces an activity by the following equation:

$$M_f^{\mathrm{BF}}(i, j) = \mathrm{abs}\big((E_f^{BF}(i, j) - I_f^{BF}(i, j) \big). \qquad (9)$$

If $M_f^{\mathrm{BF}}(i, j) > 0$, the node is in an excitatory state, and especially if $M_f^{\mathrm{BF}}(i, j)$ is large, there exists at least one object to move toward the center point (camera or robot) along the back-to-up direction.

(5) Neuron BFN. The neuron collects the activities of medulla nodes only in Region 4, and subsequently produces a converged activity by the following equation:

$$S_f^{\mathrm{BF}} = 2 \times \mathrm{Sig}\left(\frac{\mathrm{Sum}_f^{\mathrm{BF}}}{m^{\mathrm{BF}}} \right) - 1, \qquad (10)$$

where

$$\mathrm{Sum}_f^{\mathrm{BF}} = \sum_{i=1}^{M/2} \sum_{j=n_l^{\mathrm{BF}}(i)}^{n_r^{\mathrm{BF}}(i)} \mathrm{abs}\big(M_f^{\mathrm{BF}}(i,j)\big), \qquad (11)$$

and m^{BF} denotes the number of the medulla nodes in region 4 as in Figure 2. If S_f^{BF} is large, an object will move toward the center point ($M/2$, $N/2$) along the back-to-up direction. Conversely, an object will deviate from the point gradually even when there is no moving object in the region.

4.3.2. EH-FVNN. The network detects the changes of visual motion on the equatorial curve around a fixed point, which involves three points: (i) judging whether there exists a moving object on the curve; (ii) if yes, it will detect whether the object moves clockwise (cw) or counterclockwise (ccw);

and (iii) it computes the output activity if an object makes an either cw or ccw movement.

As related to Figure 2, the equatorial region, surrounded by the elliptic curve is uniformly divided into four sub-regions A_i with $1 \le i \le 4$ by means of the center point ($M/2$, $N/2$) and both horizontal and vertical axes. A_1 and A_4 are the right-upward and right-downward subregions with sector angles $90°$ in turn, whereas A_2 and A_3 denote the left-upward and left-downward ones, respectively. Similar to BF-FVNN, EH-FVNN also sequentially executes the P-, R-, and L-layers. Hereafter, each medulla node (i, j) within the equatorial region captures the activities of oo nodes around the counterpart in the oo layer, and then outputs an excitatory or inhibitory activity. Precisely, each node (i, j) in A_1 produces an excitatory activity by the following equation:

$$E_f^{A_1}(i,j) = o_f(i,j) - \frac{1}{3}w \sum_{0 \le k,l \le 2} o_f(i-k,j-l)g\big(o_f(i-k,j-l)\big), \qquad (12)$$

with $k, l \neq 2$ and $(k, l) \neq (1.1)$. Also, it generates an inhibitory activity by the following equation:

$$I_f^{A_1}(i,j) = w\left(o_f(i,j) - \frac{1}{3} \sum_{0 \le k,l \le 2} o_f(i+k,j+l)g\big(o_f(i+k,j+l)\big) \right), \qquad (13)$$

with $k, l \neq 0$, $(k, l) \neq (1, 1)$, $M/2 \le i \le M$, and $N/2 \le j \le n_r^{A_1}(i)$. Herein, $n_r^{A_1}(i)$ denotes the column number of the rightmost medulla node in the ith row as in A_1. Accordingly, the node (i, j) elicits an excitatory or inhibition activity by the following equation:

$$M_f^{A_1}(i,j) = E_f^{A_1}(i,j) - I_f^{A_1}(i,j). \qquad (14)$$

Furthermore, all the nodes in the region A_1 generate a gathered activity by the following equation:

$$\mathrm{EH}_f^{A_1} = 2\left(\mathrm{Sig}\left(\frac{\mathrm{Sum}_f^{A_1}}{m^{A_1}} \right) - 0.5 \right), \qquad (15)$$

where m^{A_1} denotes the number of nodes in A_1 and

$$\mathrm{Sum}_f^{A_1} = \sum_{i=M/2}^{M} \sum_{j=N/2}^{n_r^{A_1}(i)} \mathrm{abs}\big(M_f^{A_1}(i,j)\big). \qquad (16)$$

Similarly, the nodes in the ith region A_i can also create a converged activity of $\mathrm{EH}_f^{A_i}$ with $2 \le i \le 4$. This way, when an object moves on the fixed equatorial curve, the maximum of the four activities $\mathrm{EH}_f^{A_i}$ with $1 \le i \le 4$, i.e., c_f, is taken to measure the motion changes of the object on the curve, which is given by the following equation:

$$c_f = \max\left\{ \mathrm{EH}_f^{A_i}, 1 \le i \le 4 \right\}. \qquad (17)$$

It is highlighted that, in the case where there is no moving object in the equatorial region, if an object makes an equatorial movement on the equatorial curve, the synthesized excitatory intensity at each medulla node on the curve will keep almost identical at each frame. Therefore, an index I_f can be used to identify whether an object makes an equatorial movement, defined by the following equation:

$$I_f = \begin{cases} 1, & \text{if } \sigma_f < \max_s, \\ 0, & \text{else,} \end{cases} \qquad (18)$$

where $\max_s$ is the submaximum in the set of $\left\{ \mathrm{EH}_f^{A_i}, 1 \le i \le 4 \right\}$ and σ_f is the standard deviation of $\{c_k\}_{k=f-n_{\mathrm{st}}}^{f}$. Again, an equatorial motion trajectory can be formulated approximately by a sine or cosine-like curve. Thereby, the movement trajectory of an object near the equatorial curve can be represented approximatively by either $d_f = c_f \sin f$ or $e_f = c_f \cos f$. Hence, a ccw identification index is designed to judge whether an object makes a ccw movement, which is given by the following equation:

$$I_f^{\mathrm{ccw}} = \begin{cases} 1, & \text{if } \big(I_f = 1\big) \wedge \big(\sigma_f^{n_{\mathrm{st}}} < \sigma_{th}\big), \\ 0, & \text{else,} \end{cases} \qquad (19)$$

where $\sigma_f^{n_{st}}$ and $\wedge$ are the standard deviation of $\{d_k\}_{k=f-n_{st}}^{f}$ and the AND logical operation, respectively. Similarly, a cw identification index can be used to check if an object makes a cw movement, designed by the following equation:

$$I_f^{cw} = \begin{cases} 1, & \text{if } \left(I_f = 1\right) \wedge \left(\sigma_f^e < \sigma_{th}\right), \\ 0, & \text{else,} \end{cases} \quad (20)$$

where σ_f^e is the standard deviation of $\{e_k\}_{k=f-n_{st}}^{f}$.

In terms of the abovementioned model designs in (11)–(18), EHN can not only perceive the excitatory intensity of the moving object on the equatorial trajectory at each frame, but also identify whether the object moves counterclockwise or clockwise. Accordingly, the output of EHN at frame f can be defined by the following equation:

$$S_f^{EH} = \begin{cases} d_f, & \text{if } I_f^{ccw} = 1, \\ e_f, & \text{elseif } I_f^{cw} = 1, \\ 0, & \text{else.} \end{cases} \quad (21)$$

4.3.3. LR-, RL-, and UD-FVNNs.
Each of the three networks also includes three shared neural layers (P, R, and L), one specific medulla subregion, and one output neuron. Their neural layers are completely the same as those in BF-FVNN. Besides, their neurons only respond to respective preferential sub regions in the medulla layer.

(1) UD-FVNN: The medulla subregion only includes the nodes in region 2 as in Figure 2. The activity of each node (i, j) at frame f, i.e., $M_f^{UD}(i, j)$, is computed by replacing "BF" with "UD" in equations (6)–(8) as in the BF-FVNN. Therein, the notations of i and j range from $M/2$ to M and from $n_l^{UD}(i)$ to $n_r^{UD}(i)$ respectively, in which $n_l^{UD}(i)$ and $n_r^{UD}(i)$ denote individually the column numbers of the leftmost and rightmost nodes in the ith row as in region 2. Additionally, the output of the UDN can be acquired by equation (9) after "BF" is displaced by "UD."

(2) LR- and RL-FVNNs: The LR-FVNN's medulla subregion only includes the nodes in region 3 as in Figure 2, in which the activity of the node (i, j), $M_f^{LR}(i, j)$, can be computed by replacing "BF" with "LR" in equations (6)–(8); i and j range from 1 to $N/2$ and from $n_l^{LR}(j)$ to $n_r^{LR}(j)$, respectively. The RL-network's medulla subregion only involves in the nodes in region 1 as in Figure 2. Herein, the activity of each node (i, j), $M_f^{RL}(i, j)$, can be acquired similarly, where j changes from $N/2$ to N and so does i from $n_l^{RL}(j)$ to $n_r^{RL}(j)$. Notice that $n_l^{LR}(j)$ and $n_r^{LR}(j)$ denote the row numbers of the upmost and downmost medulla nodes in the jth column as in region 3, and so do $n_l^{RL}(j)$ and $n_r^{RL}(j)$ in region 1. Additionally, the outputs of LRN and RLN are similarly computed by equation (9) after replacing "BF" with "LR" and "RL," respectively.

4.4. Postsynaptic Neural Network.
Related to Figure 3(b), the network consists of the PB- and FB-layers as well as the neurons of F1 and F5. The former layer, i.e., PB-layer, comprises sixteen neural nodes which receive and process the excitatory activities outputted by the neurons in the presynaptic neural network. The latter one, i.e., FB-layer, is composed of eight nodes, each of which collects the outputs of the two symmetrically arranged PB nodes and process them through a spike-transmitting mechanism. F1 and F5 collect the outputs of all the nodes in the FB-layer simultaneously and later generate their activities to measure the changes of the visual angle and height of the main object.

4.4.1. PB-Layer.
As mentioned in Section 3, the sixteen nodes (PB_i, $1 \le i \le 16$) in the PB-layer are uniformly arranged on a sine-like curve segment neurophysiologically. Herein, the outputs of LRN, UDN, EHN, BFN, and RLN are received by all the neural nodes in the sets of $\{PB_1, PB_2, PB_3\}$, $\{PB_4, PB_5, PB_6\}$, $\{PB_7, PB_8, PB_{15}, PB_{16}\}$, $\{PB_{12}, PB_{13}, PB_{14}\}$, and $\{PB_9, PB_{10}, PB_{11}\}$, respectively. For example, the output of LRN is collected by each node in set $\{PB_1, PB_2, PB_3\}$ simultaneously. Thereafter, a constraint information transmission mechanism is used to ensure that each node PB_i outputs an excitatory activity or remains unresponsive by the following equation:

$$PB_f(i) = \begin{cases} S_f(i), & \text{if } S_f(i) > 0, \\ 0, & \text{else,} \end{cases} \quad (22)$$

with $1 \le i \le 16$, where $S_f(i)$ represents the input of PB_i at frame f.

4.4.2. FB-Layer.
The nodes (FB_i, $1 \le i \le 8$) in the layer are distributed on the horizontal line. A symmetrical cross-connection transmission mechanism is employed to ensure that each node acquires the outputs of the two symmetrical nodes in the PB-layer, namely, the ith node receives the outputs of PB_i and PB_{17-i} with $1 \le i \le 8$. Subsequently, a specific weighted linear combination transmission mechanism is used to generate the node's activity by the following equation:

$$FB_f(i) = \mu PB_f(i) + (1 - \mu)PB_f(17 - i), \quad (23)$$

where μ is a weight coefficient taking 0.1 if $i = 1, 2, 5$ or 6, and 0.9 otherwise.

4.4.3. Neuron F1.
As an angle detector, F1, which can make a strong response to the main object, detects the change of the visual angle (θ) of the object. It collects the activities outputted by the eight nodes in the FB-layer, and then executes a spike mechanism to evaluate whether it is in an excitatory state. More precisely, let f_0 stand for the first frame number in which the moving object is captured by the camera, and f_{max} represents the maximal time step which the danger occurs. Usually, f_0 and f_{max} are two fixed numbers defined by the user, in which their difference denotes the

version of visibility. Subsequently, F1 firstly filters out the low activities by the following equation:

$$G_f = \max\{FB_f(i), 1 \le i \le 8\} \times \sin(\theta_f), \qquad (24)$$

where the visual angle at frame f is estimated by the following equation:

$$\theta_f = \begin{cases} \dfrac{\pi}{6} \times \dfrac{k_f}{f_0}, & \text{if } k_f < f_o, \\[2ex] \dfrac{\pi}{2} \times \dfrac{k_f - f_o}{f_{\max}}, & \text{else,} \end{cases} \qquad (25)$$

with $k_f = f \bmod (f_0 + f_{\max})$. Herein, equation (25) shows that if k_f approaches either f_0 or $f_0 + f_{\max}$, θ_f will be equal to 30° or 90° approximately. This indicates that, once some moving object appears in the visual field or enters a given danger zone, F1 will become excitatory. Thus, equations (24) and (25) are exploited to characterize part of the neuron's properties appropriately. Then, a spike mechanism is chosen to detect whether there exists a spike, which is given by the following equation:

$$S_f^{F1\,\text{spike}} = \begin{cases} 1, & \text{if } G_f > 0, \\ 0, & \text{else.} \end{cases} \qquad (26)$$

Hence, F1 outputs an excitatory activity by the following equation:

$$V_f^{F1} = \begin{cases} P_f^{F1}, & \text{if } \sum\limits_{i=0}^{n_{st}-1} S_{f-i}^{F1\,\text{spike}} \ge n_{st}, \\[2ex] 0, & \text{else,} \end{cases} \qquad (27)$$

where P_f^{F1} is defined by the following equation:

$$P_f^{F1} = \begin{cases} \dfrac{1}{n_{st}} \sum\limits_{i=0}^{n_{st}-1} S_{f-i}^{F1\,spike}, & \text{if } S_f^{F1\,spike} = 0, \\[2ex] G_f, & \text{else.} \end{cases} \qquad (28)$$

4.4.4. Neuron F5. In the panoramic scene, the height from the camera to the ground and the minimum danger distance between the camera and the main object directly determine whether a collision alarm will be transmitted. Therefore, F5 as a height detection neuron is adopted to detect the visual height of the object. Herein, it collects the activities generated by the eight nodes in the FB-layer, and then generates an excitatory intensity by the following equation:

$$g_f = \max\{FB_f(i), 1 \le i \le 8\} \times \frac{h}{l}, \qquad (29)$$

where h and l, usually given by user, denote the visually maximal vertical height and horizontal distance (i.e., the distance between frames f_0 and $f_0 + f_{\max}$), respectively. Accordingly, a spike mechanism is designed to identify whether

F5 generates a spike at frame f, which is given by the following equation:

$$S_f^{F5\,\text{spike}} = \begin{cases} 1, & \text{if } g_f > 0, \\ 0, & \text{else.} \end{cases} \qquad (30)$$

Such a mechanism guides F5 to output an activity by the following equation:

$$V_f^{F5} = \begin{cases} P_f^{F5}, & \text{if } \sum\limits_{i=0}^{n_{st}-1} S_{f-i}^{F5\,\text{spike}} \ge n_{st}, \\[2ex] 0, & \text{else,} \end{cases} \qquad (31)$$

where P_f^{F5} is given by the following equation:

$$P_f^{F5} = \begin{cases} n_{st} \prod\limits_{i=0}^{n_{st}-1} g_{f-i}, & \text{if } g_f = 0, \\[2ex] g_f, & \text{else.} \end{cases} \qquad (32)$$

5. Fly Visual Brain Omnidirectional Collision Detection

This section first displays a so-called FVBJNN-based omnidirectional collision detection model.

FVBJOCD is given in Section 5.1. Second, the model's algorithm formulation and computational complexity are given in Section 5.2.

FVBJOCD as a collision detection model, schematically illustrated by Figure 6 which is exploited to measure changes in the panoramic scene and transmit collision alarms. Once multiple objects move in the scene (e.g., Figure 6(a)), the model is required to promptly take one of the objects as the main object closest to the panoramic camera and decide whether an alarm is transmitted through an early warning scheme. Specifically, related to Figure 6(b), FVBJOCD is formed of FVBJNN and three functional neurons, i.e., feedforward inhibition (FFI), synthetic spike inhibition (SSI), and collision detection (CD). Therein, FVBJNN captures successive visual signals and outputs the activities of F1 and F5. FFI prevents FVBJOCD from transmitting false collision signals in the case where an object suddenly oscillates in the panoramic field. SSI, along with FFI, generates possible spike signals, after which CD decides whether a danger happens.

5.1. Omnidirectional Collision Detection

5.1.1. Neuron FFI. Recall that $J_f(i,j)$ represents the luminance intensity of photoreceptor (i,j) at frame f. Herein, a conventional high-pass filter with time constant τ is used to eliminate visual noise signals at each photoreceptor in the P-layer as in FVBJNN by the following equation:

$$\frac{dF_t(i,j)}{dt} = J_t(i,j) - \frac{1}{\tau} F_t(i,j), 1 \le i \le M, 1 \le j \le N, \qquad (33)$$

where $F_t(i,j)$ is employed to smoothen the luminance intensity at time t captured by photoreceptor (i,j) in the P-layer. Afterwards, all the nodes in the P-layer generated a gathered activity by the following equation:
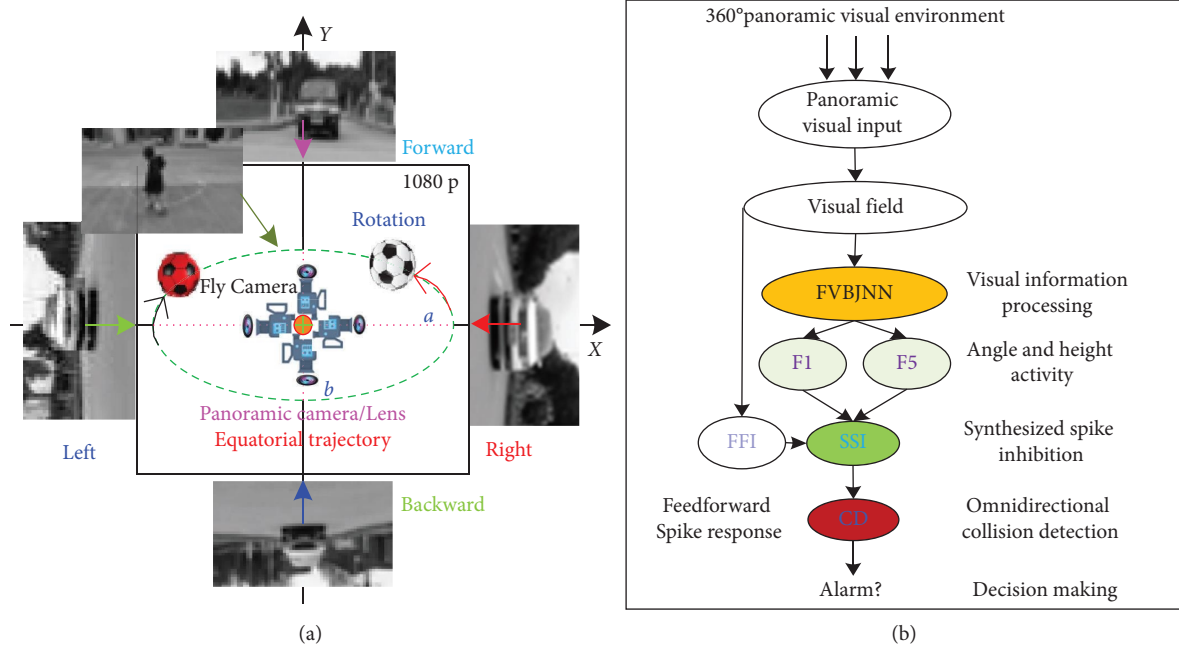
Figure 6: FVBJNN's collision contour and decision-making model. (a) Omnidirectional contour structure. (b) Schematic diagram on collision detection.

$$S_f^{\text{FFI}} = 2\left(\text{Sig}\left(\text{Sum}_f^{\text{FFI}}\right) - 0.5\right), \tag{34}$$

where

$$\text{Sum}_f^{\text{FFI}} = \text{FFI}_f + \sum_{l=0}^{n_{\text{st}}-1} \nu_l \times \text{FFI}_{f-l}, \tag{35}$$

and

$$\text{FFI}_f = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} \text{abs}\left(F_f(i,j)\right). \tag{36}$$

Finally, a spike inhibition model is developed to check if FFI delivers a spiking signal by the following equation:

$$S_f^{\text{FFI spike}} = \begin{cases} 1, & \text{if } S_f^{\text{FFI}} \leq T_{\text{PF}}, \\ 0, & \text{else,} \end{cases} \tag{37}$$

where T_{PF} is a dynamic threshold [53], i.e., $T_{\text{PF}} = T_0 + a_F \times F_f$ with positive T_0 and k_F.

5.1.2. Neuron SSI. To characterize the changes of motion in the panoramic scene, SSI synthesizes the activities of F1 and F5 as in FVBJNN. When an object moves near the equatorial trajectory in Figure 2, SSI produces a synthesized activity by the following equation:

$$S_f^{Eq} = \begin{cases} \rho\left(V_f^{F1} + V_f^{F5}\right)\sin(f), & \text{if } I_f^{ccw} = 1, \\ \rho\left(V_f^{F1} + V_f^{F5}\right)\cos(f), & \text{elseif } I_f^{cw} = 1, \\ 0, & \text{else,} \end{cases} \tag{38}$$

with adjustable intensity parameter ρ. Conversely, when an object makes an on-equatorial movement, a converged activity, generated by the SSI is computed by the following equation:

$$S_f^{Rad} = \begin{cases} \rho\left(V_f^{F1} + V_f^{F5}\right), & \text{if}\left(V_f^{F1} > 0\right) \wedge \left(V_f^{F5} > 0\right), \\ \rho\left(V_f^{F1} \times V_f^{F5}\right), & \text{else.} \end{cases} \tag{39}$$

Summarily, in terms of the index of I_f as in equation (18), SSI generates an activity by the following equation:

$$V_f^{SSI} = \begin{cases} S_f^{\text{Eq}}, & \text{if } I_f = 1, \\ S_f^{\text{Rad}}, & \text{else.} \end{cases} \tag{40}$$

Hence, a spike detection model is used to judge whether a spike inside SSI will be delivered by the following equation:

$$S_f^{SSI\,\text{spike}} = \begin{cases} 1, & \text{if } V_f^{SSI} \geq T_f, \\ 0, & \text{else,} \end{cases} \tag{41}$$

where T_f is decided by one dynamic threshold scheme [8].

5.1.3. Neuron CD. To judge whether the panoramic field includes a possible danger in terms of the activities and spike signals produced by FFI and SSI, CD as an alarm-transmitting neuron determines whether to transmit a collision alarm. Precisely, it first creates a responsive activity by the following equation:

$$V_f^{\text{CD}} = \begin{cases} V_f^{\text{FFI}}, & \text{if } \left(\sum_{l=0}^{n_{\text{st}-1}} S_{f-l}^{\text{FFI spike}} \geq n_{\text{st}} \right) \vee \left(\sum_{l=0}^{n_{\text{st}-1}} S_{f-l}^{\text{SSI spike}} < n_{\text{st}} \right), \\ V_f^{\text{SSI}}, & \text{else,} \end{cases} \tag{42}$$

where $\vee$ denotes the OR logical operation. Then, a spike detection scheme is constructed to check if CD disseminates an alarm, defined by the following equation:

$$\text{Alarm}_f^{\text{CD}} = \begin{cases} 1, & \text{if } \left(\sum_{l=0}^{n_{\text{st}-1}} S_{f-l}^{\text{SSI}} \geq n_{\text{st}} \right) \vee \left(\left| S_f^{\text{Eq}} \right| \geq T_f \right). \\ 0, & \text{else.} \end{cases} \tag{43}$$

5.2. Algorithm Formulation and Computational Complexity.
Based on the model design of FVBJOCD, the algorithm formulation is given in Algorithm 1 below.

The FVBJOCD's computational cost is mainly determined by steps 1-5. Within a run period, the ViBe method in step 1 is used to extract the current foreground image G_f in the P-layer, which needs to operate M_1 times with $M_1 = 2\,MN$. Furthermore, since each of the five subnetworks in Section 4.3 is required to enforce M_2 arithmetic operations with $M_2 = (43 + 2n_c)\,MN + 7$, step 2 computes 5 M_2 times; step 3 needs M_3 operations with $M_3 = 6n_c + 8$; step 4 executes M_4 operations to check if a spike happens with $M_4 = 2\,MN + 4$. However in order that FVBJOCD can transmit early warning signals, step 5 computes the value of T_f in equation (40) with at most $2n_c^2$ times, and also implements M_5 times with $M_5 = 2n_c^2 + 4$. Summarily, FVBJOCD runs at most M_6 times for a given input frame. Herein, M_6 is computed by the following equation:

$$M_6 = (217 + 10n_c)\text{MN} + 4n_c^2 + 6n_c + 51. \tag{44}$$

Since n_c takes a small value, The FVBJOCD's complexity in the worst case is decided by $O\,(217\,MN)$.

6. Experimental Study

To check if FVBJOCD can implement omnidirectional collision detection, the FVBJOCD's parameters setting is given in Section 6.1. Section 6.2 discusses whether the five subneural networks in the presynaptic neural network as in Section 4.3 can emerge their preferential response performances in terms of seven kinds of real or virtual video sequences. Subsequently, the performance test of the postsynaptic neural network in Section 4.4 is examined in Section 6.3 based on three real video sequences. Furthermore, FVBJOCD is compared with five neurocomputational models (AFVNN (2015) [8], RMPNN (2017) [10], LGMD1 (2018) [55], FVSCDM (2021) [9], and MLG1 (2022) [70]) with the aspect of collision detection performance in Section 6.4. The experimental analyses on the FVBJOCD's efficiency and sensitivity are executed in Section 6.5. Finally, an experimental summary is displayed in Section 6.6.

6.1. Environmental Parameters Setting.
The whole experiments are executed on the Windows 10 system with i7CPU/ 3.60 GHz and RAM/16G. Source codes are written in Visual Studio 2013 with Open CV3.4.0. The videos are shot by a Sony HDR-CX700 and Canon HF-R86/806 or generated by computer. Each video sequence is recorded at a frame rate of 50 fps or 25 fps and separated into 8 bit grayscale images. Related to the works [5–7, 44, 50, 52], the parameters setting is defined in Table 1.

6.2. Performance Test on the Presynaptic Neural Network

6.2.1. Panoramic Radial and Equatorial Motion Detection.
Three video sequences in Figure 7 are taken to examine if the five subnetworks (LR-, RL-, UD-, BF-, and EH-FVNNs) in FVBJNN can respond to visual changes in their preferred regions.

After the three series of frames in Figure 7 are inputted to the presynaptic neural network in FVBJNN, the subnetworks output own response curves given in Figure 8. Figures 7(a) and 8(a) indicate that LR-, RL-, UD-, and BF-FVNNs present unique performance characteristics. They are all activated at frame 50 or so, and generate different or similar response curves able to characterize their response performances. Herein, LR- and RL- FVNNs emerge two similar response curves after the video 1 is inputted. More precisely, they output own increasing excitatory intensities within frames 50 and 376 or so, but reduce their activities after frame 376, for which the two footballs synchronously approach the visual center in regions 3 and 1 before 363 and later stay at the visual center. On the other hand, UD- and BF-FVNNs are more and more excitatory from frame 50 to frame 545 or so, but maintain an inhibited state after frame 545 or so. Their output curves implicate that the upper and lower footballs synchronously move toward the visual center in regions 2 and 4 before frame 545 and keep static after frame 545, which coincides with the real movement status of the footballs. It is highlighted that, since the left and right footballs approach the visual center faster than the upper and lower ones, the two response curves, created by LR- and RL-FVNNs ascend the own peaks faster than those generated by UD- and BF-FVNNs. These correspond with the behavioral changes of radial movements of the four footballs in Figure 7(a), and thus the four subnetworks can be exploited to formulate the intrinsically preferential response characteristics of the related neurons in radial movement scenes.

As associated to Figure 7(b), Figures 7(c) and 8(a) clearly emphasize that, when a football makes a ccw or cw equatorial movement by Figures 7(b) or 7(c), EH-FVNN can always identify the movement pattern and direction of

Input: Grayscale frame I_f;
Output: $S_f^{LR}, S_f^{RL}, S_f^{UD}, S_f^{BF}, S_f^{EH}, T_f$, and V_f^{CD};
Step 1: Create foreground image G_f at frame f by equation (2);
Step2: Generate BF-FVNN's excitatory activity and output S_f^{BF} by equation (10); likewise, compute $S_f^{LR}, S_f^{RL}, S_f^{UD}$, and S_f^{EH} for LR-, RL-, UD- and EH-FVNNs, respectively;
Step 3: Compute SSI's excitatory intensity S_f^{Eq}, S_f^{Rad}, and V_f^{SSI} by equations (36)–(38);
Step 4: Utilize equation (37) to detect if a spike occurs inside FFI;
Step 5: Check if a collision warning signal occurs by equation (43);
Step 6: Stop collision detection if no requirement is satisfied, and return to step 2 otherwise.

ALGORITHM 1: Omnidirectional collision detection algorithm on FVBJOCD

TABLE 1: Parameters setting on FVNN and FVBJOCD.

Parameter name	Value	Parameter name	Value
Decay factor A	5	Excitation amplitude B	10
Speed attenuation τ	0.02	Elliptic half axis a (b)	2(1)
Spike threshold n_{st}	6	Intensity parameter ρ	0.1~0.9
Resolution $M \times N$	120×80	Local inhibition weight w	10^{-3}~0.1
The rate of h/l	0.4~0.6	—	—

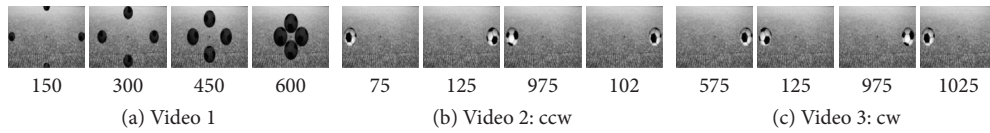| 150	300	450	600	75	125	975	102	575	125	975	1025
	(a) Video 1				(b) Video 2: ccw				(c) Video 3: cw		

FIGURE 7: Illustrative panoramic frames. Each video is expressed only by four frames; the frame number is mentioned under each image. (a) Involves in 600 frames used to examine whether LR-, RL-, UD-, and BF-FVNNs can react to visual movements in their preferred regions. Based on the regional division of the panoramic visual region similar to that in Figure 2, the up-to-down and down-to-up black footballs in regions 2 and 4 approach the camera within frames 50 and frame 360 individually, while so do the left-to-right and right-to-left footballs in regions 3 and 1 within frames 50 and 540, respectively; (b) and (c) include 1025 frames, where two identical footballs make a ccw movement in video 2 and a cw movement in video 3 along the same equatorial curve with the center of origin, respectively; the two videos are chosen to detect if EH-FVNN can recognize the patterns of ccw and cw motion of the identical footballs and detect the related changes of visual motion. (a) Video 1. (b) Video 2: ccw. (c) Video 3: cw.
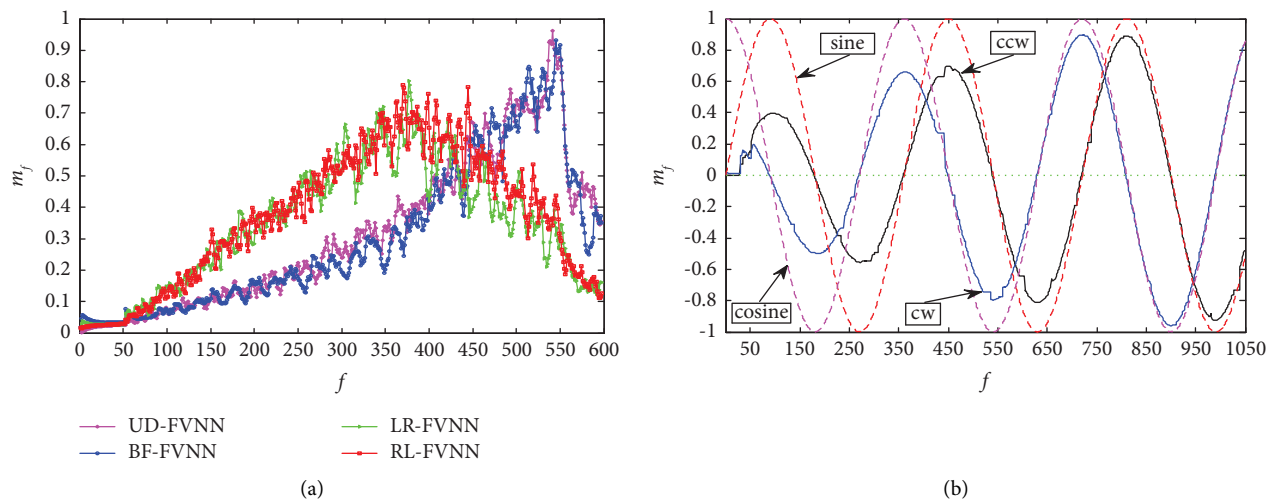

FIGURE 8: Subnetworks' response curves based on the matched visual scenes in Figure 7; m_f stands for the activity of each subnetwork at frame (f). (a) Response curves on UD-, BF-, LR-, and RL-FVNNs. (b) Response curves on EH-FVNN. (a) Video 1. (b) Videos 2 and 3.

a moving object on the equatorial trajectory and accordingly generate a sine- or cosine-like response curve. In the case where the football makes cw movements, the cosine-like response curve acquired by EH-FVNN can almost periodically reach the peaks at frames 56, 364 and 722, and 1050, and falls down the feet at frames 187, 543, and 902. On the other hand, when it scrolls along the equatorial curve counterclockwise, the acquired sine-like response curve can almost periodically ascend the peaks at frames 98, 452, and 810, and get into the valley bottoms at frames 275, 632, and 988. Therefore, EH-FVNN can not only recognize the pattern of cw or ccw movement, but also rationally describe the change of equatorial movement. Herein, it is emphasized that EH-FVNN does not occur the phenomenon of time-delayed response for the cw equatorial movement by comparison against the cw movement, as the patterns of the cw and ccw movements are defined by the cosine and sine curves as in Section 4.3.2, respectively, i.e., $e_f = c_f \cos f$ and $d_f = c_f \sin f$. This indicates that EH-FVNN can preferentially respond to the pattern of equatorial movement.

6.2.2. Panoramic Compound Movement Detection. Related to Figure 9(a), Figure 10(a) illustrates that, once the football moves along the S-shaped curve in the panoramic scene, the subnetworks generate excitatory curve segments in their preferred visual regions. We take UD-FVNN for example to analyze the intrinsic properties of LR-, RL-, UD-, and BF-FVNNs. It is activated at frame 48, after which its excitatory intensity enlarges increasingly and reach the maximum at frame 316 in that the football moves along the left-to-rightS-shaped curve to pass through the upper visual subregion (region 2) within frames 48 and 316. Subsequently, since the football leaves region 2 and moves along the S-shaped curve in the right and lower subregions (regions 1 and 4) within frames 316 and 963, UD-FVNN does not make any response. Finally, it generates an increasing excitation curve segment within frames 963 and 1050, as the football to appear in the left visual region (Region 2) once again. These verifies that the model design in Section 4.3 is rational, namely, UD-FVNN is required to only respond to the changes of visual motion in region 2. Herein, it is highlighted that, both LR- and UD-FVNNs increasingly excite within frames 49 and 130, since the football rolls simultaneously in the common part of the upper and left regions 2 and 3.

Furthermore, it can be seen from Figures 9(b1) and 10(b) that, if a football makes irregular movements and goes through regions 1–4 and the equator-like curve, the five subnetworks can output period-like activity curve segments. For example, within frames 1 and 141, UD-, LR-, and EH-FVNN keep excitatory or inhibitory, but BF-FVNN has no any response to the movements of the football. Once the football makes periodic-like circle movements, the activity curves, outputted by the neural networks change almost periodically within 142 and 956. Especially, each of them can only yield excitatory or inhibitory activities in their preferred regions in the panoramic scene.

6.2.3. Real Panoramic Movement Detection. Based on the scene setting in Figure 6(a), four real video sequences are taken synchronously by four identical cameras on the same bracket at an intersection of our campus, of which each sequence only pays close attention to the visual movements on the upper, lower, left, or right visual scene. Hereafter, they are synthesized to form a panoramic scene in Figure 11(a) with a total of 750 frames. In the synthesized panoramic scene, one truck and three cars slowly move simultaneously toward the respective cameras at the intersection from top to bottom, from bottom to top, from left to right, and from right to left, respectively, which causes a collision hazard.

Figure 11(b) indicates that EH-FVNN has no response to the panoramic scene due to no equatorial movement. When the vehicles in regions 1–4 approach the camera gradually, all the subnetworks except EH-FVNN become excitatory within frames 42 and 445 or so, and later, their excitation curves rise with time gradually, which coincides with the changes of visual motion in the panoramic scene. Herein, compared with UD- or BF-FVNN, each of LR-and RL-FVNNs can produce relatively high activities within frames 450 and 750, since the two cars in the left and right directions approach the cameras at a faster speed. Thereby, FVBJNN can detect the changes of motion in the panoramic scene.

Summarily, Figures 7, 10, and 11(b) can conclude that each of the five subnetworks only reacts to its preferred subregion in the panoramic scene and also produce a response curve to rationally evaluate the movement changes of one or more moving objects. This shows that FVBJNN can detect visual movement changes in any panoramic scene by means of the subnetworks.

6.3. Performance Test on the Postsynaptic Neural Network. Recall that the neurons F1 and F5 are used to detect the changes of the visual height and angle of the main object between the panoramic camera and the object, respectively. When a car approaches the camera gradually, the postsynaptic neural network is desired to produce a series of large visual angles and heights, in order to measure the degree of which the main object approaches the camera. As related to Figures 12 and 13, when $h = 0$, the two neurons keep weakly excitatory, while their response curves descend after frame 1430. If h is large, e.g., $h = 140$, their excitations cannot rationally describe the behavioral changes of the car. Nevertheless, if $h = 50$, the neurons can output proper excitatory intensity curves within frames 1050 and 1500. It is highlighted that their response behaviors are highly consistent with time and can be therefore used to formulate the change of the contour of visual movement. Totally, the experiment implies that a rational camera height can help FVBJNN enhance the performance of motion detection.

6.4. Performance Comparison on Collision Detection. To evaluate the collision detection performance of FVBJOCD, three real panoramic scenes, which involve horizontal, vertical, equatorial and omnidirectional collision detection, are taken to check if the model, together with the six

(a1) Video1: S-type movement.

(a2) Movement trajectory based on (a1).

(b1) Video2: Compound movement.

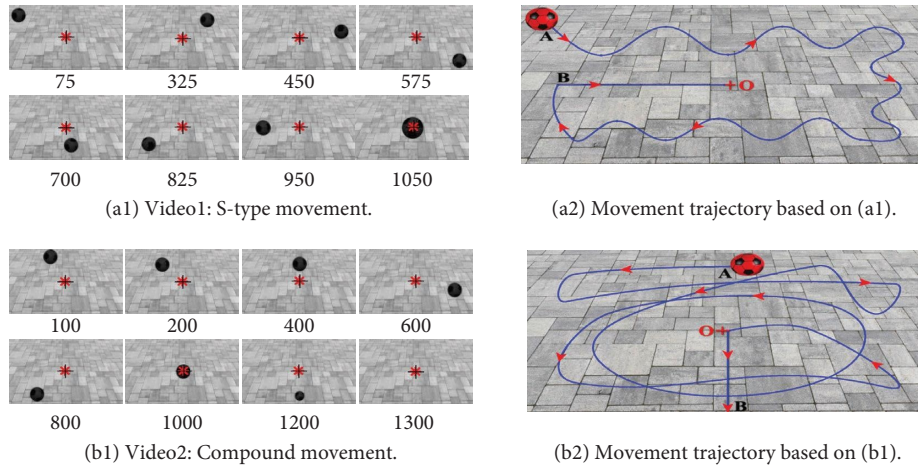(b2) Movement trajectory based on (b1).

FIGURE 9: Two sequence frames (a1) and (b1) as well as the related movement trajectories are picked up to check if the five sub networks can react to nonradial and nonequatorial motion. In (a2), a football starts at the upper left point near the ground, and then rolls along a S-shaped curve. Related to the regional division of a panoramic visual scene such as that in Figure 2, the football passes through four subregions (regions 1–4) and one equator-like trajectory. Finally, it moves toward the central point (O). However, in (b2), the football starts from the up-front point A and then goes to the point of reely. Subsequently, it radially vanishes from the field of view along the vertical-down line. (a1) Video 1: S-type movement. (a2) Movement trajectory based on (a1). (b1) Video2: Compound movement. (b2) Movement trajectory based on (b1).
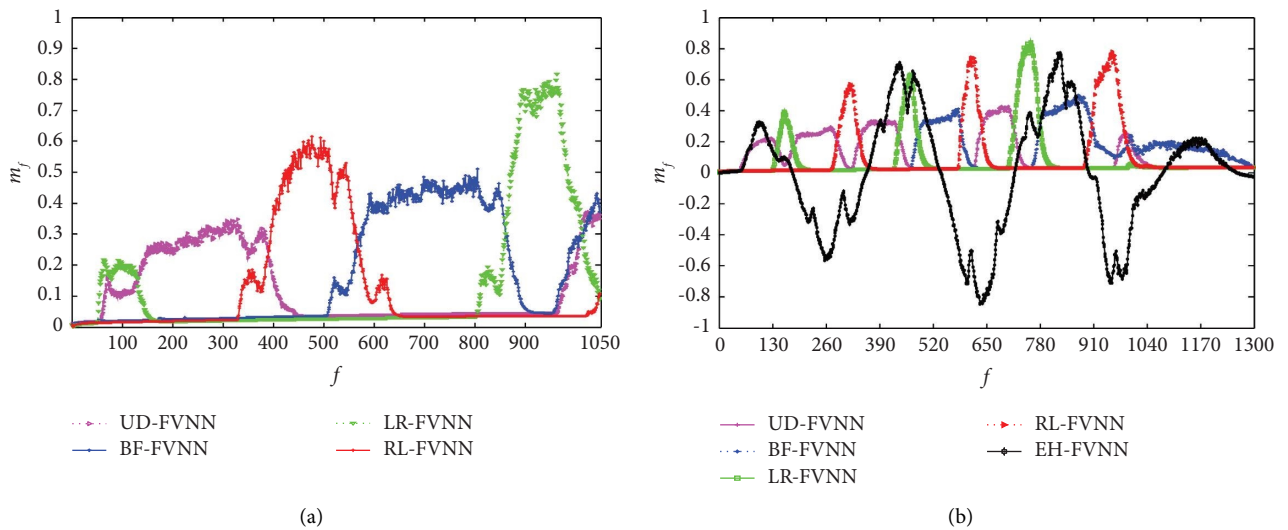

(a)

(b)

FIGURE 10: The subneural networks' response curves related to the matched visual scenes in Figure 9. (a) Video 1. (b) Video 2.

compared models (AFVNN, LGMD1, RMPNN, FVSCDM, MLG1, and FVBJOCD), can successfully execute panoramic collision detection in terms of one dynamical threshold scheme (i.e., T_f) mentioned in equation (39). It is illustrated that among the six compared models, AFVNN can only receive the image frames with size 128×128 per frame, due to the inherent structure.

6.4.1. Horizontal or Vertical Collision Detection.
Figure 14 displays two panoramic scenes, each of which formulates the state of two cars approaching the camera in the horizontal or vertical directions. In Figure 14(a), the left and right white cars approach the camera at their own speeds, in which the right car approaches the camera at a faster speed. In such a case, the collision region is the pixel interval of [690, 710]. In Figure 14(b), the up-to-down blue car approaches the camera at a constant speed. The down-to-up white car, however, does so at a fast speed, and hence, leads to a collision region of [695, 705].

Figures 14(a) and 15(a) verify that, when the two cars gradually approach the camera at respective speeds in the horizontal direction, the compared models (AFVNN, LGMD1, RMPNN, MLG1, and FVSCDM) exhibit different collision detection performances. Based on their design principles, AFVNN and LGMD1 can only acquire relatively weak excitatory intensities with time, since they can capture those visual movements in the visual regions with a visual
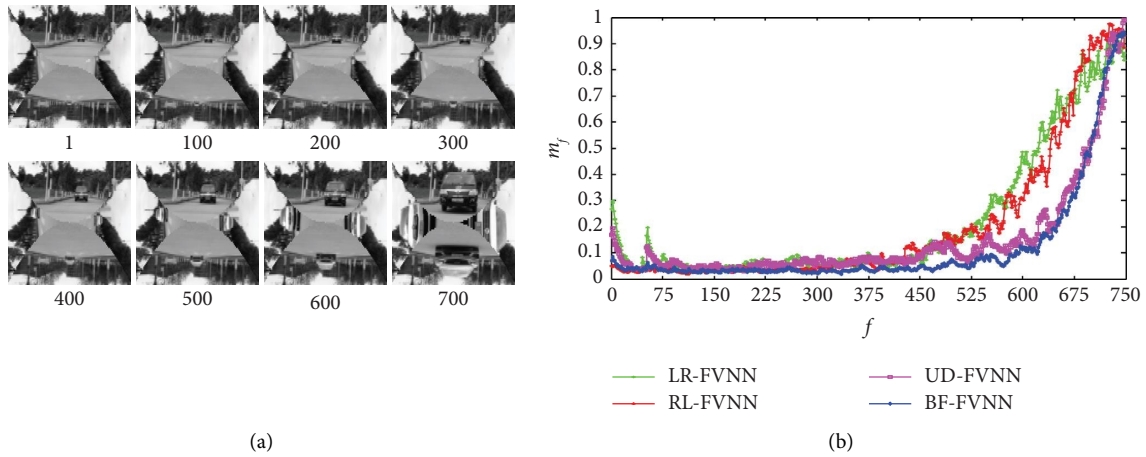
(a)

(b)

FIGURE 11: Panoramic video scene and excitation intensity response curves: (a) eight illustrative example frames of a synthesized panoramic video sequence and (b) the neural response curves acquired by the four subnetworks.

(a) Video2:Vertical height 0 cm.   (b) Video3:Vertical height 50 cm.   (c) Video4:Vertical height 140 cm
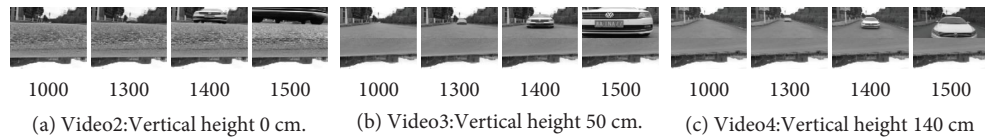
FIGURE 12: Three illustrative video frame sequences on radial movement. The version of vertical height denotes the distance between the camera and the ground. The frame sequences are selected to detect how the camera's height influences the outputs of F1 and F5. Herein, when the camera's height is 50 or 140 cm, a car drives toward the camera at a slow speed from frame 1 to frame 1400 and later approaches the camera at a fast speed from frame 1401 to frame 1500. Also, the car approaches the camera at a slow speed from frame 1360 to frame 1500 when the camera's height is 0 cm. It is emphasized that despite the same visual scene, the three image sequences indicate that the car's contour changes with the camera's height. In other words, the higher the height of the camera, the smaller the car's contour. (a) Video 2: vertical height 0 cm. (b) Video 3: vertical height 50 cm. (c) Video 4: vertical height 140 cm.
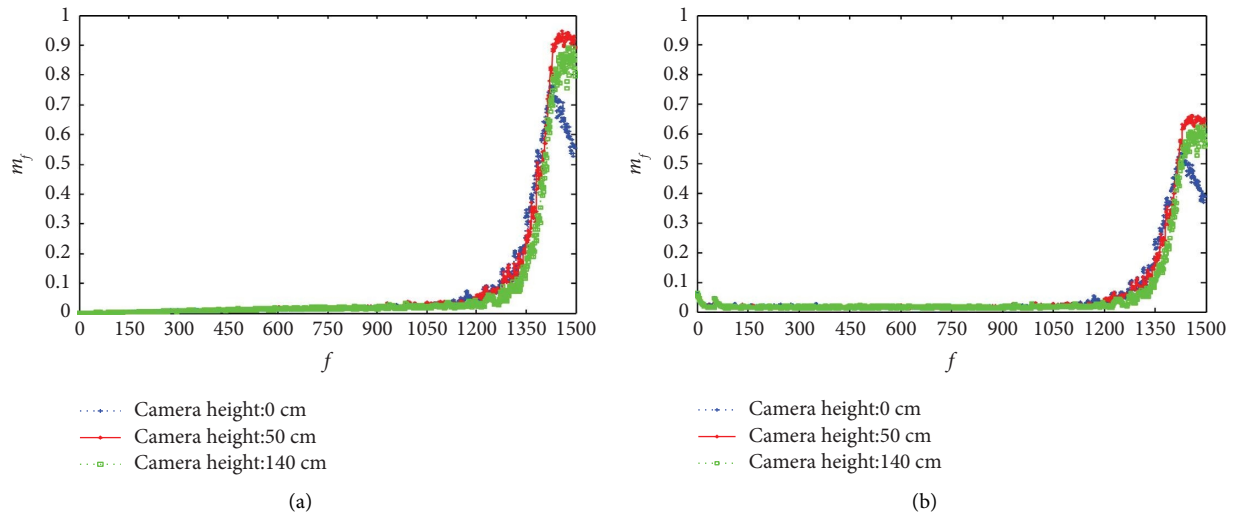

(a)

(b)

FIGURE 13: FVBJNN's response curves related to the matched visual scenes in Figure 12. (a) Response curves of neuron F1. (b) Response curves of neuron F5.

angle of 60° or 120° on the upper half plane. RMPNN cannot obtain any excitatory intensity at each frame because the panoramic scene does not involve in rotational motion. MLG1, as a looming spatial localization neural network, can detect the spatial position changes of moving objects.

Compared with AFVNN and LGMD1, it can only output a weak excitatory intensity at each moment, since its design inspiration requires its excitatory intensity at each moment to be decided by visual stimuli within a narrow subregion with a visual angle of 22.5°. FVSCDM behaves best among

95 190   380   570   665 750 95   190   380   570   665   750

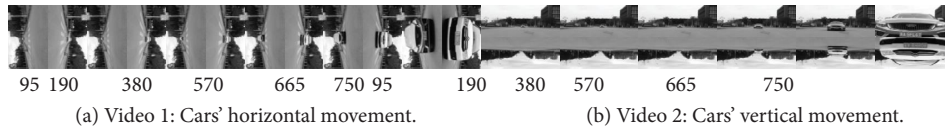(a) Video 1: Cars' horizontal movement.   (b) Video 2: Cars' vertical movement.

FIGURE 14: Illustrative frames of two videos: (a) two cars approach the camera at respective speeds along the left and right directions and (b) two cars move toward the camera at different speeds along the upward and backward directions.
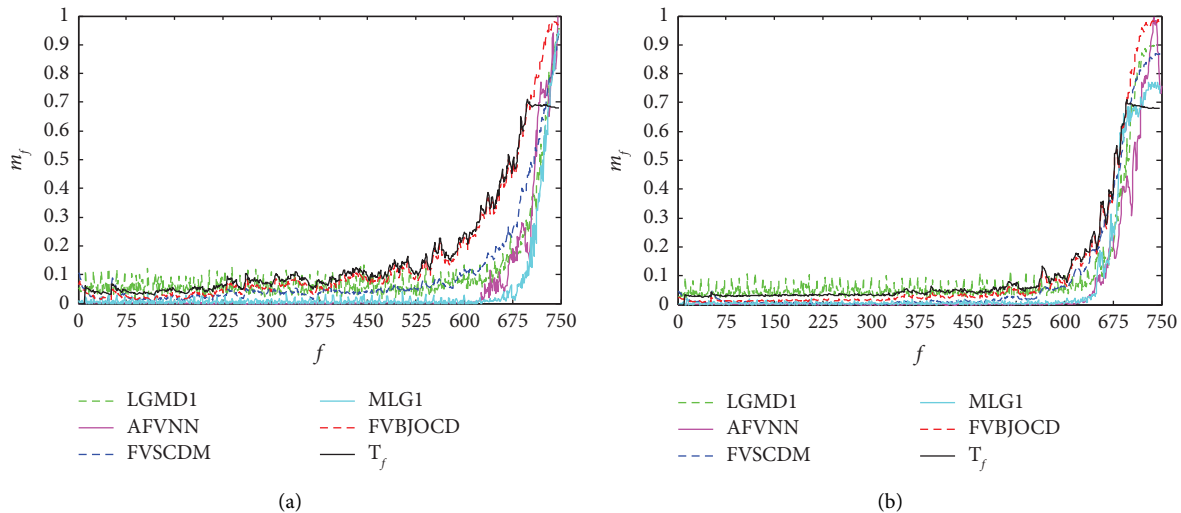

(a)   (b)

FIGURE 15: Collision detection and the comparison of excitatory intensity curves related to Figure 14. T_f is the dynamic threshold curve as in equation (39). (a) Left and right collision response. (b) Upward and backward collision response.

the compared models, as it can perceive changes in visual motion as in the whole upper half-plane and can execute edge detection as well. On the other hand, compared with each of the above five chosen models, FVBJOCD can produce a relatively high excitatory intensity at each frame after frame 500, since it can capture the complete visual movement information of the cars. Furthermore, it outperforms the compared models regarding the effect of collision detection. The threshold curve hints that it can accurately transmit collision early warning signals after frame 697. AFVNN, LGMD1, MLG1, and FVSCDM can only generate narrow and time-delayed collision regions [717, 737], [730, 750], [733, 750], and [727, 747], respectively.

Figure 15(b) demonstrates that, when the two cars move toward the camera along the direction from top to bottom and from bottom to top, respectively, the models can produce their excitatory curves based on Figure 14(b). In such a case, RMPNN has no response due to no rotational movement. LGMD1 and FVSCDM, which can only respond to the motion information of the up-to-bottom car on the upper-half plane, generate similar and increasing excitatory intensity curves and transmit collision early warning signals after frame 710 or so. Additionally, the excitatory curve of FVBJOCD is decided by the motion changes of the down-to-up car and exceeds the curves of AFVNN, LGMD1, MLG1, and FVSCDM, since the car approaches the camera faster than the up-to-down one. As related to the threshold curve of T_f, FVBJOCD can rationally touch upon collision

early warning after frame 695, whereas so do LGMD1 and FVSCDM after at least frame 705, and hence, result in time-delayed early warning. MLG1 and AFVNN can only trigger collision warning after frame 720 or so; particularly, AFVNN can only derive a relatively low excitatory intensity curve and broadcast late warning signals, since it can only focus on the motion changes of the up-to-down car in front of the upper half field of view with 60°.

6.4.2. Equatorial Collision Detection. The video sequence in Figure 16(a) is taken to examine whether FVBJOCD can trigger equatorial collision. Therein, the girl moves 64.8 seconds clockwise along the elliptic curve with long axis 2 m and short axis 1 m, and rotates 9 laps along the curve. It is emphasized that, since AFVNN, LGMD1, and FVSCDM cannot react to the pattern of equatorial movement, they do not participate in comparison. Based on the visual scene, Figure 16(b) presents the threshold curve and the excitatory curves of FVBJOCD, MLG1, and RMPNN. Herein, MLG1 cannot effectively implement equatorial collision detection, since its output activities are not consistent with the change of visual movement on the equatorial curve. Especially, even though RMPNN can produce a response curve, it cannot recognize the movement pattern of the girl for which it is designed to recognize the pattern of circular movement. Also, it can only transmit a few of early warning signals, and thus it is difficult to perform collision detection. FVBJOCD, however, can identify the girl's equatorial movement in
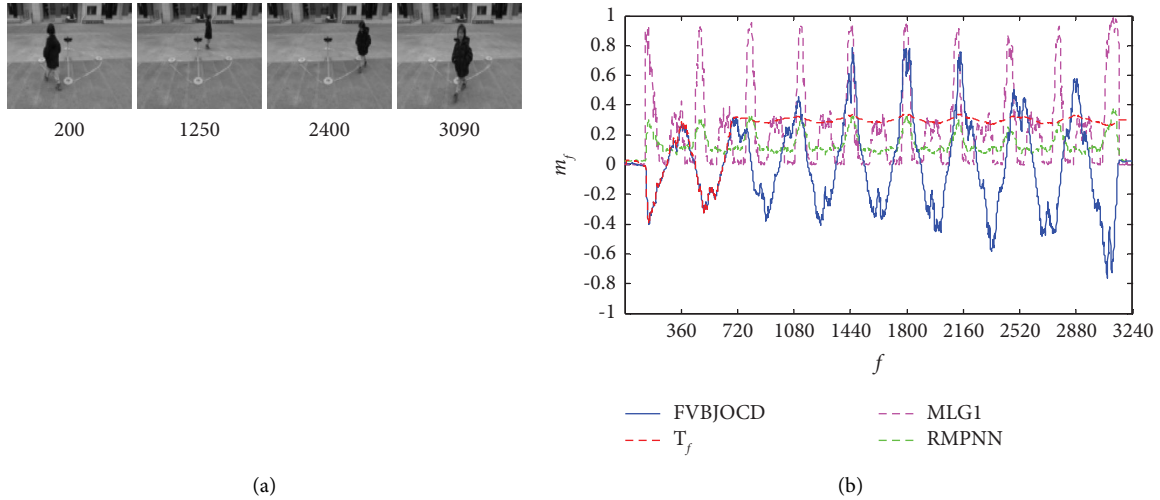
(a)

(b)

FIGURE 16: Equatorial collision detection. (a) cw equatorial movement and (b) cw response and threshold curve.

terms of EH-FVNN and generates excitatory or inhibitory curve segments periodically. Also, it triggers collision warning at frame 156 or so and engenders a collision interval of [156, 3240].

6.4.3. Omnidirectional Collision Detection.

Take Figure 11(a) for example to examine whether FVBJOCD can execute omnidirectional collision detection in a panoramic scene. It is highlighted that, since there does not exist the pattern of equatorial motion in the scene, RMPNN does not participate in the comparison. Figure 17 confirms that FVBJOCD can generate a high excitatory intensity curve segment which exceeds the threshold curve, since the four cars approach the camera synchronously at different speeds, especially the left-to-right car approaches the camera at a fast speed in comparison with the other cars. The compared models, however, are activated to produce high activities at a slow speed and transmit late collision early warning signals. Herein, FVSCDM can output a relatively high intensity curve by comparison with AFVNN, MLG1, and LGMD1, as the left and right edges can be detected. FVBJOCD can make a strong excitatory response and acquire a rational collision region of [672, 750] by the left-to-right car. Among the compared four models, AFVNN, MLG1, and LGMD1 can send collision early warning signals by their collision regions [721, 741], [711, 731], and [712, 732], respectively; FVSCDM performs relatively well in the collision region [705, 725], since it can output a high excitatory intensity than each of the other three models at each frame f after frame 610 or so. Additionally, FVBJOCD can accurately transmit early warning signals, whereas the four compared models can only send late collision early warning signals in that they cannot detect any visual motion change on the visual lower half-plane in the panoramic environment.

Summarily, based on the experiments on collision detection, the performance differences between FVBJOCD and the compared models can be found in Table 2. The table verifies that RMPNN can only recognize the equatorial movement of the girl and send collision early warning signals after frame 156. Thereby, after solving the four panoramic visual scenes, RMPNN can get the lowest success rate of 25% regarding collision detection. MLG1, LGMD1, AFVNN, and FVSCDM can transmit their collision alarms by the left and right moving cars in Car I as well as the upper and lower cars in Car II and Car III, but fail to recognize the pattern of equatorial movement in Girl. It is pointed out that, although the four models can perform collision detection with a success rate of 75%, they easily cause time-delayed early warning, which can be known by comparing their collision regions with the real CR intervals in Table 2. Relatively, FVSCDM behaves well by comparison against LGMD1, MLG1, and AFVNN, since the collision regions, acquired by the latter three models more prominently deviate from the corresponding real collision regions. Besides, not only FVBJOCD can accurately recognize the movement pattern appearing in each of the four panoramic scenes, but also its collision regions acquired in the scenes are close to the true collision regions. Therefore, it can promptly and accurately transmit early warning signals in the panoramic scenes, and acquire the highest success rate by comparison with the compared models.

6.5. Sensitivity and Efficiency Analysis.

Although FVBJNN includes eight parameters (i.e., M, N, A, B, τ, a, b, and w), the former seven parameters are usually fixed or defined by the user. Again, w is used to control the intensity that each node is inhibited at any moment by its neighboring nodes in the M-layer, in which its value is within 0 and 0.01. Therefore, only the input image resolution $M \times N$ dominates the performance of each subneural network in FVBJNN. Furthermore, the omnidirectional collision detection model in Section 5 includes six parameters (f_0, f_{max}, n_{st}, h, l, and ρ). Herein, the former five parameters are defined by the user, and ρ as an amplitude parameter is exploited to adjust the excitatory degree of neuron SSI. If it is large, the phenomenon of false early warning will occur, and conversely,
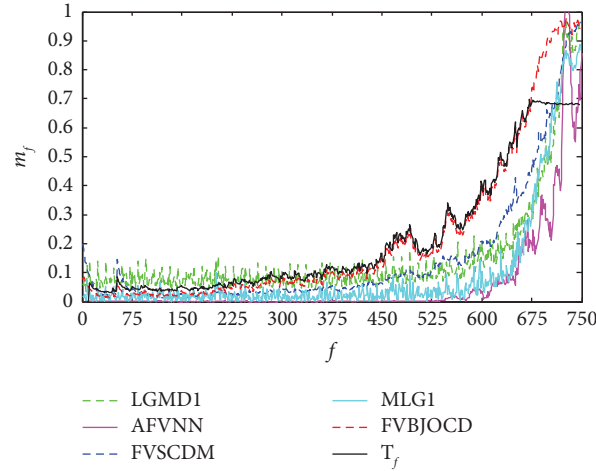
FIGURE 17: Comparison of collision detection performances of the four models based on the panoramic scene in Figure 11(a) and the threshold curve of T_f.

TABLE 2: Comparison of experimental results of the six collision detection models.

Fundamental materials							Experimental result				
Type	ToF	MP	TF	CR	Model		Car I	Car II	Girl	Car III	SR (%)
—	—	—	—	—	LGMD1	CR	[730, 750]	[707, 727]	—	[712, 732]	75
Car I	750	HM	702	[697, 750]	AFVNN	CR	[717, 737]	[719, 739]	—	[721, 741]	75
Car II	750	VM	700	[695, 750]	FVSCDM	CR	[727, 747]	[702, 722]	—	[705, 725]	75
Girl	3240	EM	691	[686, 3220]	RMPNN	CR	—	—	[156, 3240]	—	25
Car III	750	OM	677	[672, 750]	MLG1	CR	[733, 750]	[717, 737]	[138, 3240]	[711, 731]	75
—	—	—	—	—	FVBJOCD	CR	[697, 717]	[695, 715]	[686, 3240]	[672, 692]	100

Note: Cars I, II, III, and the girl denote the image frame sequences in Figures 11(a), 14(a), 14(b), and 16(a), respectively. SR stands for the success rate of collision detection or early collision warning. Some versions are abbreviated, i.e., ToF, total of frames; MP, motion pattern; TF, trigger frame; CR, collision region; HM, horizontal movement; VM, vertical movement; EM, equatorial movement; OM, omnidirectional movement.

FVBJOCD will delay the early warning. Consequently, it is suggested that ρ take a value within 0.6 and 0.9. Herein, we take the visual scene in Figure 10(a) for example to examine how the efficiencies of the abovementioned six models depend on the image input resolution.

Table 3 illustrates that, once the resolution becomes large increasingly, each model takes much more runtime to implement collision detection. Therefore, the frame's resolution directly influences the models' efficiency. Additionally, a high-resolution video prompts the models to generate early collision warnings, while conversely, delayed warnings may occur. Thereby, to make a rational trade-off between effect and efficiency, it is suggested that M and N take values within 110 and 130 and within 75 and 85, respectively.

By comparison against LGMD1, MLG1, and FVSCDM, FVBJOCD can acquire the highest efficiency in each resolution setting, and meanwhile our previous model, i.e., FVSCDM, can also implement collision detection with high efficiency. However, LGMD1 and MLG1 need much more runtime to execute their functional modules, while their excitatory curves rise slowly and easily which cause late collision early warning. Additionally, whereas AFVNN and RMPNN only require less runtime to handle each image frame, they can only detect changes in visual motion as in the narrow or on the circle curve, and cannot adapt to the scene

of omnidirectional collision. Thereby, such two models are clearly inferior to the other aforementioned models.

6.6. Summary. Based on the above twelve virtual or real video sequences, FVBJNN and FVBJOCD have been examined with the aspects of response preference, collision detection performance, and execution efficiency, in order to exhibit their intrinsic characteristics, advantages, disadvantages, and applicable scopes. Meanwhile, five computational models, i.e., AFVNN, LGMD1, FVSCDM, MLG1, and RMPNN, are chosen to compare with FVBJOCD. Related to the above experimental results, some distinctive differences between them can be highlighted below:

(i) FVBJNN, which consists of five subnetworks (LR-, RL-, UD-, BF-, and EH-FVNNs), is used to detect the change of omnidirectional movement in the panoramic scene. Each subnetwork has been verified to be able to respond to visual movements in the preferential region of the panoramic scene, even if one or more moving objects makes complex and irregular movements. Particularly, EH-FVNN can well perceive the changes of visual motion on the equatorial curve. Generally, the vertical height from the panoramic camera to the ground usually

TABLE 3: Comparison of average runtime at each frame.

$M \times N$ Res	LGMD1 CR/RT	MLG1 CR/RT	FVSCDM CR/RT	AFVNN CR/RT	RMPNN CR/RT	FVBJOCD CR/RT
100×60	[723, 743]/0.19	[684, 704]/0.13	[663, 683]/0.11	—	—/0.15	[692, 712]//0.11
120×80	[712, 732]/0.48	[711, 731]/0.32	[697, 717]/0.24	[721, 741]/0.02	—/0.18	[672, 692]/0.24
128×128	[695, 715]/1.43	[721, 741]/0.98	[680, 700]/0.61	[732, 750]/0.17	—/0.32	[656, 676]/0.63
240×160	[683, 703]/7.60	[741, 750]/5.07	[640, 660]//2.91	—	—/1.20	[574, 594]//2.73

Note: Res, CR, and RT denote the abbreviations of resolution ratio, collision region, and runtime (second), respectively.

influences the performance of movement detection, and thus a proper camera height is desired to satisfy the requirement of visual movement detection.

(ii) FVBJOCD, which comprises of FVBJNN and an omnidirectional collision detection model, is designed to transmit collision alarms when a hazard occurs imminently. The experimental results validate that it can disseminate collision alarms when one or more objects radially approach the panoramic camera, especially when an object moves on the equatorial curve. The results in Table 2 have clearly verified that FVBJOCD can successfully enforce omnidirectional collision early warning in panoramic environments, regardless of the pattern of visual motion. However, LGMD1, AFVNN, and FVSCDM can only achieve early warning with the success rate of 75%, since they are specially designed to satisfy the requirement of forward but not omnidirectional collision warning. Even though MLG1 can achieve a 75% success rate in collision early warning and adapt to the region localization of spatial motion, it cannot perform omnidirectional collision early warning well. RMPNN can only acquire the success rate of 25% for collision detection, as it can only detect the change of rotational movement.

(iii) The analyses of sensitivity and efficiency show that ρ and $M \times N$ are the crucial factors that influence the FVBJOCD's efficiency and effect on collision detection. It is recommended that ρ, M, and N take values in the intervals [0.6, 0.9], [110, 130], and [75, 85], respectively. The compared models need to make great improvements in terms of efficiency and effectiveness when confronted with implementing omnidirectional collision detection.

7. Conclusion

Omnidirectional collision detection has been rarely cultivated over recent decades. However, such a topic will be popular in computer vision, as the visual brain system is a natural bio-inspiration for developing omnidirectional collision detection models. Hereby, FVBJNN is constructed to perceive changes in visual motion as in the panoramic scene, while a collision detection model (FVBJOCD) is developed to perform omnidirectional collision detection. Three conclusions can be drawn below:

(i) FVBJNN can effectively detect the whole changes of objects in the panoramic scene, relying upon the five subnetworks (LR-, RL-, UD-, BF-, and EH-FVNNs) in the presynaptic neural network. The subnetworks can preferentially respond to specific motion clues. Also, in terms of the postsynaptic neural network, FVBJOCD can output the excitatory intensities of angle and height detection neurons to rationally evaluate motion changes in the panoramic scene. Summarily, FVBJOCD can not only detect the change of visual movement omnidirectionally, but also transmit collision alarms with a high success rate.

(ii) As an omnidirectional collision detection model with the computational complexity of O ($217\,MN$), FVBJOCD can decide which object is approaching the camera at the fastest speed. Also, it can successfully transmit collision alarms if a possible danger appears in the panoramic field of view, and only take about 0.24 s to process each image frame with size 120×80.

(iii) AFVNN, LGMD1, and FVSCDM can only perform collision detection in the *upper half* field of view and cannot detect the changes of motion on the posterior half-plane and the equatorial trajectory. RMPNN and MLG1 can partially recognize the pattern of equatorial movement, but cannot preferentially execute collision detection in the panoramic scene. FVBJOCD, however, can execute *omnidirectional* collision detection in the panoramic scenes, and can also implement collision early warning successfully.

Finally, even though FVBJOCD has exhibited some prominent advantages over the compared models with the aspect of panoramic collision detection, some issues are kept open. For example, it is necessary to further study the setting scheme of the intrinsic parameters so that the model can automatically adapt to different visual scenes. Besides, whereas the current work has initially touched upon omnidirectional collision detection, it keeps open how FVBJOCD can adapt to panoramic collision avoidance. In the future, we will investigate a FVBJOCD-based automatic collision avoidance system able to be transplanted into FPGA/ARM.

Data Availability

No underlying data were collected or produced in this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] B. Brik and A. Ksentini, "Toward optimal MEC resource dimensioning for a vehicle collision avoidance system: a deep learning approach," *IEEE Network*, vol. 35, no. 3, pp. 74–80, 2021.

[2] J. Y. Gai, L. R. Xiang, and L. Tang, "Using a depth camera for crop row detection and mapping for under-canopy navigation of agricultural robotic vehicle," *Computers and Electronics in Agriculture*, vol. 188, Article ID 106301, 2021.

[3] Z. G. Zhou, L. H. Li, A. Fürsterling, H. J. Durocher, J. Mouridsen, and X. P. Zhang, "Learning-based object detection and localization for a mobile robot manipulator in SME production," *Robotics and Computer-Integrated Manufacturing*, vol. 73, pp. 102229–102312, 2022.

[4] A. Borst, M. Drews, and M. Meier, "The neural network behind the eyes of a fly," *Current Opinion in Physiology*, vol. 16, pp. 33–42, 2020.

[5] T. Schilling, A. H. Ali, A. Leonhardt, A. Borst, and J. Pujol-Martí, "Transcriptional control of morphological properties of direction-selective T4/T5 neurons in Drosophila," *Development*, vol. 146, no. 2, Article ID dev169763, 2019.

[6] H. Wei, H. Y. Kyung, C. Desplan, P. J. Kim, and C. Desplan, "The diversity of lobula plate tangential cells (LPTCs) in the Drosophila motion vision system," *Journal of Comparative Physiology*, vol. 206, no. 2, pp. 139–148, 2020.

[7] H. S. Cheong, I. Siwanowicz, and G. M. Card, "Multi-regional circuits underlying visually guided decision-making in Drosophila," *Current Opinion in Neurobiology*, vol. 65, pp. 77–87, 2020.

[8] Z. Zhang, S. G. Yue, and G. P. Zhang, "Fly visual system inspired artificial neural network for collision detection," *Neurocomputing*, vol. 153, pp. 221–234, 2015.

[9] 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.

[10] B. Hu, S. Yue, and Z. Zhang, "A rotational motion perception neural network based on asymmetric spatiotemporal visual information processing," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 11, pp. 2803–2821, 2017.

[11] H. Wang, Q. Fu, H. Wang, P. Baxter, J. Peng, and S. Yue, "A bioinspired angular velocity decoding neural network model for visually guided flights," *Neural Networks*, vol. 136, pp. 180–193, 2021.

[12] Q. Fu, C. Hu, J. Peng, F. C. Rind, and S. Yue, "A robust collision perception visual neural network with specific selectivity to darker objects," *IEEE Transactions on Cybernetics*, vol. 50, no. 12, pp. 5074–5088, 2020.

[13] L. G. Bishop and D. G. Keehn, "Two types of neurones sensitive to motion in the optic lobe of the fly," *Nature*, vol. 212, no. 5068, pp. 1374–1376, 1966.

[14] G. Ammer, R. M. Vieira, S. Fendl, A. Borst, and S. Footnotes, "Anatomical distribution and functional roles of electrical synapses in Drosophila," *Current Biology*, vol. 32, no. 9, pp. 2022–2036.e4, 2022.

[15] K. M. Boergens, C. Kapfer, M. Helmstaedter, W. Denk, and A. Borst, "Full reconstruction of large lobula plate tangential cells in Drosophila from a 3D EM dataset," *PLoS One*, vol. 13, no. 11, pp. 02078288–e207915, 2018.

[16] M. Meier and A. Borst, "Extreme compart mentalization in a Drosophila amacrine cell," *Current Biology*, vol. 29, no. 9, pp. 1545–1550.e2, 2019.

[17] K. Hausen, "Functional characterization and anatomical identification of motion sensitive neurons in the lobula plate of the blowfly calliphoraerythrocephala," *ZeitschriftFürNaturforschung C*, vol. 31, no. 9-10, pp. 629–634, 1976.

[18] H. Eckert, "Functional properties of the H1-neurone in the third optic ganglion of the Blowfly,Phaenicia," *Journal of Comparative Physiology*, vol. 135, no. 1, pp. 29–39, 1980.

[19] J. C. Tuthill, A. Nern, G. M. Rubin, and M. B. Reiser, "Widefield feedback neurons dynamically tune early visual processing," *Neuron*, vol. 82, no. 4, pp. 887–895, 2014.

[20] M. F. Keles and M. A. Frye, "Object-detecting neurons in Drosophila," *Current Biology*, vol. 27, no. 5, pp. 680–687, 2017.

[21] R. Sen, M. Wu, K. Branson, A. Robie, G. M. Rubin, and B. J. Dickson, "Moonwalker descending neurons mediate visually evoked retreat in Drosophila," *Current Biology*, vol. 27, no. 5, pp. 766–771, 2017.

[22] J. J. Omoto, M. F. Keleş, B. C. Nguyen et al., "Visual input to the Drosophila central complex by developmentally and functionally distinct neuronal populations," *Current Biology*, vol. 27, no. 8, pp. 1098–1110, 2017.

[23] J. P. Kumar, "The fly eye: through the looking glass," *Developmental Dynamics*, vol. 247, no. 1, pp. 111–123, 2017.

[24] F. G. Richter, S. Fendl, J. Haag, M. S. Drews, and A. Borst, "Glutamate signaling in the fly visual system," *iScience*, vol. 7, pp. 85–95, 2018.

[25] S. Y. Takemura, A. Bharioke, Z. Y. Lu et al., "A visual motion detection circuit suggested by Drosophila connectomics," *Nature*, vol. 500, no. 7461, pp. 175–181, 2013.

[26] K. Shinomiya, T. Karuppudurai, T. Y. Lin, Z. Y. Lu, C. H. Lee, and I. A. Meinertzhagen, "Candidate neural substrates for off-edge motion detection in Drosophila," *Current Biology*, vol. 24, no. 10, pp. 1062–1070, 2014.

[27] E. Serbe, M. Meier, A. Leonhardt, and A. Borst, "Comprehensive characterization of the major presynaptic elements to the Drosophila OFF motion detector," *Neuron*, vol. 89, no. 4, pp. 829–841, 2016.

[28] R. Tanaka and D. A. Clark, "Object-displacement-sensitive visual neurons drive freezing in Drosophila," *Current Biology*, vol. 30, no. 13, pp. 2532–2550.e8, 2020.

[29] U. Hanesch, K. F. Fischbach, and M. Heisenberg, "Neuronal architecture of the central complex in Drosophila melanogaster," *Cell and Tissue Research*, vol. 257, no. 2, pp. 343–366, 1989.

[30] G. Liu, H. Seiler, A. Wen et al., "Distinct memory traces for two visual features in the Drosophila brain," *Nature*, vol. 439, no. 7076, pp. 551–556, 2006.

[31] F. Toepfer, R. Wolf, and M. Heisenberg, "Multi-stability with ambiguous visual stimuli in Drosophila orientation behavior," *PLoS Biology*, vol. 16, no. 2, pp. e2003113–e2003124, 2018.

[32] K. Ito, K. Shinomiya, M. Ito et al., "A systematic nomenclature for the insect brain," *Neuron*, vol. 81, no. 4, pp. 755–765, 2014.

[33] J. Kohl and G. Jefferis, "Neuroanatomy: decoding the fly brain," *Current Biology*, vol. 21, no. 1, pp. 19-20, 2011.

[34] G. S. Boyan and H. Reichert, "Mechanisms for complexity in the brain: generating the insect central complex," *Trends in Neurosciences*, vol. 34, no. 5, pp. 247–257, 2011.

[35] W. Hu, Y. Peng, J. Sun et al., "Fan-shaped body neurons in the Drosophila brain regulate both innate and conditioned

nociceptive avoidance," *Cell Reports*, vol. 24, no. 6, pp. 1573–1584, 2018.

[36] K. Pfeiffer and U. Homberg, "Organization and functional roles of the central complex in the insect brain," *Annual Review of Entomology*, vol. 59, no. 1, pp. 165–184, 2014.

[37] C. T. Shih, O. Sporns, S. Yuan et al., "Connectomics-based analysis of information flow in the Drosophila brain," *Current Biology*, vol. 25, no. 10, pp. 1249–1258, 2015.

[38] C. Y. Lin, C. C. Chuang, T. E. Hua et al., "A comprehensive wiring diagram of the proto-cerebral bridge for visual information processing in the Drosophila brain," *Cell Reports*, vol. 3, no. 5, pp. 1739–1753, 2013.

[39] Q. Liu, Y. Xing, J. Tian et al., "Gap junction networks in mushroom bodies participate in visual learning and memory in Drosophila," *Elife*, vol. 5, pp. 1–18, 2016.

[40] A. Pavlowsky, J. Schor, P. Y. Plaçais, and T. Preat, "A GABAergic feedback shapes dopaminergic input on the Drosophila mushroom body to promote appetitive long-term memory," *Current Biology*, vol. 28, no. 11, pp. 1783–1793.e4, 2018.

[41] B. Hassenstein and W. Reichardt, "System theoretischeanalyse der zeit-reihenfolgen-und vorzeichenauswertungbei der bewegungsperzeption des rüsselkäferschlorophanus," *ZeitschriftfürNaturforschung B*, vol. 11, no. 9-10, pp. 513–524, 1956.

[42] H. Öǧmen and S. Gagné, "Neural network architectures for motion perception and elementary motion detection in the fly visual system," *Neural Networks*, vol. 3, no. 5, pp. 487–505, 1990.

[43] H. Eichner, M. Joesch, B. Schnell, D. Reiff, and A. Borst, "Internal structure of the fly elementary motion detector," *Neuron*, vol. 70, no. 6, pp. 1155–1164, 2011.

[44] H. M. Shiozaki and H. Kazama, "Parallel encoding of recent visual experience and self-motion during navigation in Drosophila," *Nature Neuroscience*, vol. 20, no. 10, pp. 1395–1403, 2017.

[45] H. Wang, J. Peng, P. Baxter, C. Zhang, Z. Wang, and S. Yue, "A model for detection of angular velocity of image motion based on the temporal tuning of the Drosophila," *Artificial Neural Networks and Machine Learning – ICANN 2018*, pp. 37–46, 2018.

[46] D. Goldschmidt, P. Manoonpong, and S. Dasgupta, "A neurocomputational model of goal-directed navigation in insect-inspired artificial agents," *Frontiers in Neurorobotics*, vol. 11, p. 20, 2017.

[47] J. M. Missler and F. A. Kamangar, "A neural network for pursuit tracking inspired by the fly visual system," *Neural Networks*, vol. 8, no. 3, pp. 463–480, 1995.

[48] H. Sun, L. Liu, and A. Guo, "A neurocomputational model of figure-ground discrimination and target tracking," *IEEE Transactions on Neural Networks*, vol. 10, no. 4, pp. 860–884, 1999.

[49] S. Günel, H. Rhodin, D. Morales, J. H. Campagnolo, P. Ramdya, and P. Fua, "DeepFly3D, a deep learning-based approach for 3D limb and appendage tracking in tethered, adult Drosophila," *Elife*, vol. 8, Article ID e48571, 2019.

[50] J. W. Aptekar, P. A. Shoemaker, and M. A. Frye, "Figure tracking by flies is supported by parallel visual streams," *Current Biology*, vol. 22, no. 6, pp. 482–487, 2012.

[51] Q. Fu and S. Yue, "Mimicking fly motion tracking and fixation behaviors with a hybrid visual neural network," in *Proceedings of the 2017 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 1636–1641, Macau, Macao, December 2017.

[52] Q. Fu, N. Bellotto, C. Hu, and S. Yue, "Performance of a visual fixation model in an autonomous micro robot inspired by Drosophila physiology," in *Proceedings of the 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pp. 1802–1808, Kuala Lumpur, Malaysia, June 2018.

[53] S. G. Yue and F. C. Rind, "Collision detection in complex dynamic scenes using an LGMD-based visual neural network with feature enhancement," *IEEE Transactions on Neural Networks*, vol. 17, no. 3, pp. 705–716, 2006.

[54] H. Meng, K. Appiah, S. Yue et al., "A modified model for the lobula giant movement detector and its FPGA implementation," *Computer Vision and Image Understanding*, vol. 114, no. 11, pp. 1238–1247, 2010.

[55] Q. Fu, C. Hu, J. Peng, and S. Yue, "Shaping the collision selectivity in a looming sensitive neuron model with parallel ON and OFF pathways and spike frequency adaptation," *Neural Networks*, vol. 106, pp. 127–143, 2018.

[56] 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. 4-5, pp. 443–460, 2020.

[57] A. H. Gaede, V. B. Baliga, G. Smyth, C. Gutiérrez-Ibáñez, D. L. Altshuler, and D. R. Wylie, "Response properties of optic flow neurons in the accessory optic system of hummingbirds versus zebra finches and pigeons," *Journal of Neurophysiology*, vol. 127, no. 1, pp. 130–144, 2022.

[58] H. Liang and T. Morie, "A motion detection model inspired by hippocampal function and its applications to obstacle detection," *Neurocomputing*, vol. 129, pp. 59–66, 2014.

[59] E. Menegatti and T. Pajdla, "Omnidirectional robot vision," *Robotics and Autonomous Systems*, vol. 58, no. 6, pp. 745-746, 2010.

[60] W. Shimizuhira, K. Fujii, and Y. Maeda, "Fuzzy behavior control for autonomous mobile robot in dynamic environment with multiple omnidirectional vision system," in *Proceedings of the 2004 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, pp. 3412–3417, Sendai, Japan, September 2004.

[61] M. Hirabayashi, K. Kurosawa, R. Yokota et al., "Flying object detection system using an omnidirectional camera," *Forensic Science International: Digital Investigation*, vol. 35, pp. 301027–301029, 2020.

[62] H. Taheri and C. X. Zhao, "Omnidirectional mobile robots, mechanisms and navigation approaches," *Mechanism and Machine Theory*, vol. 153, pp. 103958–104028, 2020.

[63] Y. Tang, Z. Gao, F. Lin, Y. Li, and F. Wen, "Visual adaptive tracking for monocular omnidirectional camera," *Journal of Visual Communication and Image Representation*, vol. 55, pp. 253–262, 2018.

[64] S. Wang, I. Segev, A. Borst, and S. Palmer, "Maximally efficient prediction in the early fly visual system may support evasive flight maneuvers," *PLoS Computational Biology*, vol. 17, no. 5, Article ID e1008965, 2021.

[65] K. Davie, J. Janssens, D. Koldere et al., "A single-cell transcriptome atlas of the aging Drosophila brain," *Cell*, vol. 174, no. 4, pp. 982–998.e20, 2018.

[66] S. Sen, D. Cao, R. Choudhary et al., "Genetic transformation of structural and functional circuitry rewires the Drosophila brain," *Elife*, vol. 3, Article ID e04407, 2014.

[67] A. N. King and A. Sehgal, "Molecular and circuit mechanisms mediating circadian clock output in the Drosophila brain," *European Journal of Neuroscience*, vol. 51, no. 1, pp. 268–281, 2020.

[68] J. I. Raji and C. J. Potter, "The number of neurons in Drosophila and mosquito brains," *PLoS One*, vol. 16, no. 5, Article ID e0250381, 2021.

[69] Z. Zheng, J. S. Lauritzen, E. Perlman et al., "A complete electron microscopy volume of the brain of adult *Drosophila melanogaster*," *Cell*, vol. 174, no. 3, pp. 730–743.e22, 2018.

[70] H. Luan, Q. Fu, Y. Zhang, M. Hua, S. Chen, and S. Yue, "A looming spatial localization neural network inspired by mlg1 neurons in the crab neohelice," *Frontiers in Neuroscience*, vol. 15, Article ID 787256, 2021.