

# Article A Novel Bio-Inspired Motion Direction Detection Mechanism in Binary and Grayscale Background

Yuxiao Hua <sup>1</sup>, Yuki Todo <sup>2,\*</sup>, Zheng Tang <sup>1,\*</sup>, Sichen Tao <sup>1</sup>, Bin Li <sup>2</sup> and Riku Inoue <sup>2</sup>

- <sup>1</sup> Faculty of Engineering, University of Toyama, Toyama-shi 930-8555, Japan
- <sup>2</sup> Faculty of Electrical and Computer Engineering, Kanazawa University, Kanazawa-shi 920-1192, Japan
- \* Correspondence: yktodo@ec.t.kanazawa-u.ac.jp (Y.T.); ztang@eng.u-toyama.ac.jp (Z.T.)

Abstract: The visual system plays a vital role in the daily life of humans, as more than 90 percent of the external information received by the human brain throughout the day comes from the visual system. However, how the human brain processes the received visual information remains a mystery. The information received from the external through the visual system can be divided into three main categories, namely, shape features, color features, and motion features. Of these, motion features are considered the key to deciphering the secrets of the visual system due to their independence and importance. In this paper, we propose a novel bio-inspired motion direction detection mechanism using direction-selective ganglion cells to explore the mystery of motion information extraction and analysis. The mechanism proposed in this paper is divided into two parts: local motion direction detection neurons and global motion direction detection neurons; the former is used to extract motion direction information from the local area, while the latter infers global motion direction from the local motion direction information. This mechanism is more consistent with the biological perception of the human natural visual system than the previously proposed model and has a higher biological plausibility and greater versatility. It is worth mentioning that we have overcome the problem in which the previous motion direction detection model could only be applied in the binary background by introducing the horizontal cells. Through the association formed by horizontal cells and bipolar cells, this model can be applied to recognizing problems of motion direction detection on a grayscale background. To further validate the effectiveness of the proposed model, a series of experiments with objects of different sizes, shapes, and positions are conducted by computer simulation. According to the simulation results, this model has been proven to have high accuracy rates regardless of objects' sizes, shapes, and positions in all experiments. Furthermore, the proposed model is verified to own more stable accuracy rates and stronger noise immunity by comparing it with the recognized superior classical convolutional neural network in a background of different percentage noise.

**Keywords:** artificial visual system; motion direction detection; direction-selective ganglion cells; horizontal cells; convolutional neural network

MSC: 68T07

## 1. Introduction

Over the past few decades, the artificial visual system has become one of the most popular research areas in artificial intelligence [1], and, at the same time, it has been widely applied in daily life and the cutting-edge technology of human beings such as intelligent robots, autonomous driving, video surveillance, and human–computer interaction [2]. However, current research on artificial visual systems mainly focuses on the application of deep learning models such as Convolutional Neural Networks (CNN), which perform well in terms of detection accuracy [3], but still lack robustness in complex environments and can not fully explain how the human visual system achieves some of the physiological functions. Thus, in an attempt to build a more powerful artificial visual system, making it



Citation: Hua, Y.; Todo, Y.; Tang, Z.; Tao, S.; Li, B.; Inoue, R. A Novel Bio-Inspired Motion Direction Detection Mechanism in Binary and Grayscale Background. *Mathematics* 2022, *10*, 3767. https://doi.org/ 10.3390/math10203767

Academic Editor: Jianping Gou

Received: 22 September 2022 Accepted: 9 October 2022 Published: 13 October 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).



closer to the features of the visual system of organisms is considered the key [4]. Through previous research, the visual system of organisms is considered as having three main abilities, namely, the extraction of color, shape, and movement features from external objects [5]. Of these, due to the independence and uniqueness of the motion feature, its analysis and extraction is considered as key to deciphering the secrets of the visual system [6].

Actually, research into the mechanism of motion detection has been underway for over half a century. In 1952, H.B. Barlow identified the presence of a neural mechanism in the retina of the frog by using pulse discharges in the optic nerve as an indicator of activity [7]. In the same year, Kuffler demonstrated that discharge in the retina was not only for cold-blooded animals but also mammals by flashing small dots of light on the light-adapted cat retina [8]. In addition, in 1956, the first bio-inspired visual model was proposed by Hassenstein and Reichardt based on the analysis of the optomotor behaviour of Chlorophanus viridis beetles [9]. This Hassenstein–Reichardt detector consisted of two mirror-symmetrical subunits with neurons responding more vigorously to specific directions, which could be widely used to explain the mechanism of motion direction detection in the retina of various organisms [10]. In 1959, D.H. Huble identified the existence of a "directional selection mechanism" in the retina of cats. Based on this mechanism, neurons in the retina would give more vigorous response when a stimulus unit moved in a particular direction (called the preferred direction), whereas motion in the opposite direction would not provoke an obvious response [11]. In 1965, H.B. Barlow designed experiments on the organization and its corresponding mechanism of directional selectivity within the receptive field of a rabbit retina. According to these experiments, Barlow proposed a new conclusion: ganglion cells respond to motion in a particular direction and belong to a subset of bipolar cells; these bipolar cells could respond to the corresponding stimuli in their adjacent receptive field. In addition, the presence of a laterally connecting inhibitory element from one of these regions was experimentally confirmed, which could lead to the hypothesis that horizontal cells also played a role in the process of motion direction detection [12]. Through these experimental results as well as hypotheses, a prototype of bio-inspired model for object motion direction detection could be derived. However, even so, the research into the motion direction detection mechanism in organisms was still only at the cellular level. Although we could speculate on which cells played a role in motion direction detection, how the neurons in them combined together to achieve their functions remains unknown. In 1989, three basic requirements for motion direction detection models were proposed as the first common principles in neural network computations by Alexander Borst and Martin Egelhaaf, which were 'two inputs', 'nonlinear interaction', and 'asymmetry' [13]. These three requirements constituted the guidelines for future design models and laid the foundation for the emergence of more reliable models in the future. In 2000, Clifford and Ibbotson linked the concept of theoretical models to the neural circuits of biological systems and proposed a basic mechanism of visual motion detection [14]. However, this mechanism did not take into account the fact that the nonlinear mechanism of the dendrites played an important role in motion direction detection [15].

In our previous studies, we proposed a novel Dendritic Neural Model (DNM) that mimicked the nonlinear interaction among inputs to the dendrites [16]. We could modify the location and type of synapses in the dendritic branches of this model to make it appropriate for a particular task such as classification problems [17], speed forecasting [18], and directional selectivity problems. Furthermore, we had proposed a model based on DNM to simulate the Direction Selective Ganglion Cells (DSGCs) for detecting the object motion direction [19]. However, this previously proposed model is only applicable to object motion direction detection on binary images. Thus, we innovated on the previously proposed model by introducing horizontal cells to make the model capable of discriminating between objects and background pixels during the motion direction detection of moving objects. The newly proposed model consists of two main components: Local Direction Detection Neuron (LDDN) and Global Direction Detection Neuron (GDDN), which can be applied

to the problem of detecting the direction of object motion on a grayscale background compared to the previous model. This paper will introduce the proposed mechanism in two main subsections. The first part will be divided into three parts, where we will first give a brief introduction to the mechanism and structure of DNM to set the scene for the rest of the presentation, and then describe in detail how LDDN and GDDN can work together to achieve object motion direction detection. We assume that the LDDN extracts elementary local motion direction information by detecting the grayscale difference between the central pixel and the surrounding pixels at different moments, and GDDN will take a simple summation and infer the global motion direction. During this process, the horizontal cells can serve to detect the grayscale difference of the surrounding pixels. In the next section, the performance of this model will be evaluated by computer simulation on several datasets. These datasets consist of images of different sizes, shapes, motion directions, and backgrounds with different proportions of noise. The experimental results show that these models have extremely high accuracy in detecting the moving direction of the objects regardless of their size and shape. In order to further demonstrate the superiority of the model, we compare their performance with the highly recognized method CNN [20] and find that, compared with CNN, these models have higher accuracy and noise resistance. In summary, this paper proposes an object motion direction detection model with not only better bio-interpretation, but also higher accuracy and noise immunity.

#### 2. Model and Method

## 2.1. Dendritic Neuron Model

Neurons are made up of three main parts, namely the cell body, axon, and dendrites. Dendrites receive information from other neurons through the parts called synapses, and transmit the information to the cell body via axons, which are the main physiological activities of neurons. In recent years, dendrites have been proved to play an important role in a variety of neuronal computations in humans, including motion direction detection operations [21]. In 1983, Poggio, Torre, and Koch proposed that there was a nonlinear logical relationship between different kinds of synapses in the dendrites of retinal nerve cells. According to the experiments, they find that the excitatory synapse will be intercepted if an activated inhibitory synapse is located on its way to the cell body; this interaction can be described as AND logic as shown in Figure 1 [22], and this relationship can be represented as the following equation:

$$Output = Output_a \cdot \overline{Output_b} \tag{1}$$

Among this equation, 'a' denotes an excitatory input, and 'b' denotes an inhibitory input. Each input is logically 0 or 1, and the inhibitory input has a logical NOT effect. Thus, the cell body (soma) will only output the result 1 when a = 0 and b = 1.



**Figure 1.** A basic model used for describing the nonlinear relationship between excitatory dendrites (a) and inhibitory dendrites (b).

Furthermore, Figure 2 shows the case of a dendrite with multiple branches. According to the research by Koch, Poggio, and Torre, when a plurality of synapses is present on a dendrite, a given excitatory signal will be intercepted by all the inhibitory signals on its direct path to the soma [23], and this process can be represented by the following equation:



$$Output = \overline{Output_a} \cdot Output_b + \overline{Output_a} \cdot Output_c \cdot \overline{Output_d}$$
(2)

**Figure 2.** An advanced model used for describing the nonlinear relationship in the dendrite with multiple branches. The a, b, c, d here stand for the different dendrites.

Based on the above research, a novel model called the Dendritic Neuron Model was proposed in 2014, which aimed to compensate for the nonlinear relationship between dendrites that was not taken into account in traditional neural models [24]. In this model, synapses on each branch will receive information from the external world (logical 0 or logical 1) and process it through a one-input one-output sigmoid function; then, output results from synapses on the same branch will be processed with a multiplication function, which is equivalent to logical AND; outputs from the multiple branches will be integrated by the summation function, which corresponds to the logical OR relationship between the different branches in the neuron; finally, the soma will receive outputs from all branches, process it with another sigmoid function, and transmit the obtained result to the other neurons. During this process, if the output signal exceeds the soma threshold, the neuron will be activated and provide facilitation effects; otherwise, it will provide inhibitory effects. The whole process can be represented by the example shown in Figure 3, and the operation implemented in Figure 3 could be represented as the following equation:

$$Output = Output_{a1} \cdot Output_{b1} \cdot Output_{c1} + Output_{a2} \cdot \overline{Output_{b2}} \cdot \overline{Output_{c2}} + Output_{a3} \cdot Output_{b3} \cdot \overline{Output_{c3}}$$
(3)

The experimental examples show that this model can be used to simulate the directionally selective ganglion cells in the retina by unsupervised learning and demonstrates a high accuracy rate in detecting the motion direction of objects on a binary background. Based on the dendritic neuron model, we hypothesize that the nonlinear relationship also exists between the cells in the human retina and proposed new models which can be applied in the motion direction in grayscale backgrounds.



**Figure 3.** The structure of the Dendritic Neuron Model with several branches and a soma. The input signals are a, b, and c, respectively, which are received by the synapses on branches numbered 1, 2, and 3, and the received signals are labeled as a1, b1, c1, etc.

#### 2.2. Local Motion Direction Detection Neuron

The specific structure and operating principle of the proposed local motion direction neuron are described in this section. According to the DNM proposed above, we assume that there are several Local Direction Detection Neurons (LDDN) that are only responsible for the detection of specific movement directions (upward, upper-rightward, rightward, lower-rightward, downward, lower-leftward, leftward, and upper-leftward) existing in the local visual area. For each type of LDDN, the basic construction is the same. The main components of LDDN, including the Photoreceptor Cells (PCs), ON-OFF Bipolar Cells (ON-OFF BCs), ON-OFF Horizontal Cells (ON-OFF HCs), and Ganglion Cells (GCs), form the whole structure as shown in Figure 4. PCs are the only cells that can perform direct interaction with optical signals in the retina [25], so they are responsible for collecting the light signals from the corresponding receptive field and transmitting them to other cells after converting them into an electrical signal through photoelectric conversion. After receiving the electrical signal transmitted by the PCs, the OF-OFF BCs generate a facilitating or inhibiting signal depending on whether the state of its corresponding receptive field has changed. ON-OFF HCs have synaptic connections with two adjacent PCs, receive the electric signals sent from both parties [26], and modulate their output through incurring inhibition effect. The output electrical signals from the ON-OFF BCs and ON-OFF HCs are gathered in their connected GCs, and they will be transmitted to the approximately 50 brain regions through their axons after being managed [27].

We assume that the LDDN acquires local information by scanning each part of the global receptive field with a local receptive field of  $3 \times 3$  areas. To introduce the operating principle of LDDN in detail, we define that, at a certain time point, the PCs receive optical signals from a local two-dimensional receptive field of size M × N. The location of the pixel in this receptive field at this moment can be defined with the coordinate (i, j, t); thus, the positions of eight adjacent pixels around the (i, j, t) can be represented as Figure 5.



Figure 4. The structure of local motion direction detection neurons.



**Figure 5.** Representation of the central pixel and its neighboring pixels at the time points before and after the movement.

Thus, eight types of LDDN responsible for eight different directions that can receive inputs from the local receptive fields are defined as shown in Figure 6.



**Figure 6.** Eight types of LDNN responsible for eight different directions (**a**–**h**). The orange area represents the center pixel before the movement, and the blue area represents its position after the movement; only movements in the preferred direction will trigger the corresponding LDDN, and the triggered LDDN is shown in red in the image.

When a light spot falls within the receptive field, we can call the pixel where the light spot is located the central pixel. The PCs receive the optical signal from the central pixel, convert it into an electrical signal, and transmit it to its connected ON-OFF BCs and ON-OFF HCs. The value of the output electrical signal is the grayscale value of the corresponding receptive field, which can be represented by X(i, j, t). When the light spot moves to any other adjacent pixel after a time interval  $\Delta t$ , the output of the PCs will also change to  $X(i, j, t + \Delta t)$  as shown in Figure 7. At this moment, the ON-OFF BCs connected to these PCs, which is called  $BC_1$ , will perform a subtraction operation on the electrical signals transmitted by the PCs at both time points.  $BC_1$  only needs to determine whether the state of the receptive field has changed, not whether it is a light-dark transition or a dark-light transition, so the result of the subtraction needs to be taken as an absolute value. The exact procedure for this operation can be expressed as the following formula:

$$BC_1 = \begin{cases} Excitation & |X(i,j,t) - X(i,j,t + \Delta t)| > L\\ Inhibition & |X(i,j,t) - X(i,j,t + \Delta t)| <= L \end{cases}$$
(4)

The 'L' here is the threshold for the minimum grayscale value difference, and only grayscale value differences greater than L will be recognized as having a change of state. At the same time, the eight types of ON-OFF BCs corresponding to the directions of motion needed to be detected, which can be called  $BC_2$ , will also detect the state of the corresponding pixels and make an excitation response if the state of the pixel has changed. The specific process is demonstrated in the following equation:

$$BC_{2} = \begin{cases} Excitation & |X(i+\alpha, j+\beta, t) - X(i+\alpha, j+\beta, t+\Delta t)| > L\\ Inhibition & |X(i+\alpha, j+\beta, t) - X(i+\alpha, j+\beta, t+\Delta t)| <= L\\ & \alpha \in \{-1, 0, 1\}, \beta \in \{-1, 0, 1\} \end{cases}$$
(5)

After the steps above have been performed, the movement direction of the light spot can be initially determined, but the results are still subject to significant error, as the distinction between the object and the background before and after the motion cannot be specifically discerned by the BCs only. Therefore, ON-OFF HCs are introduced to regulate the preliminary results by interfering with the output signals of  $BC_1$  as well as  $BC_2$  and eliminate some of the background pixels that are mistaken as moving objects. In general, ON-OFF HCs determine whether generating a suppression signal to modulate the output results of the  $BC_1$  and  $BC_2$  based on the grayscale difference of their corresponding region at time t and  $t + \Delta t$ . ON-OFF HCs will perform this process according to the following formula:

$$HCs = \begin{cases} Inhibition & |X(i,j,t) - X(i+\alpha,j+\beta,t+\Delta t)| > L\\ Excitation & |X(i,j,t) - X(i+\alpha,j+\beta,t+\Delta t)| <= L\\ & \alpha \in \{-1,0,1\}, \beta \in \{-1,0,1\} \end{cases}$$
(6)

The output of  $BC_1$ ,  $BC_2$  and ON-OFF HCs will be transferred together to the GCs for integrated processing, and the GCs will only be fired when both ON-OFF BCs in the preferred direction are activated, and the ON-OFF HCs do not produce an inhibitory effect. The output of the GCs can be expressed as the following formula, the variables in the formula represent the output signals from  $BC_1$ ,  $BC_2$ , and HCs, respectively:

$$GCs = Out_{BC_1} \cdot Out_{BC_2} \cdot Out_{HCs} \tag{7}$$

As an example, in Figure 7, we use all eight kinds of LDDN to scan the receptive filed of  $5 \times 5$  region of 25 pixels at time *t* as well as  $t + \Delta t$ . During the scanning process, the position of the center pixel of the LDDN local receptive field is denoted as (i, j) (i = 1, 2, ..., M; j = 1, 2, ..., N). For the sake of simplicity, only the first five scanned images have been represented in the figure.



**Figure 7.** Example of scanning a  $5 \times 5$  receptive field with eight types of LDDN. The subfigure (**a**–**e**) here represents the scanning process of LDDNs with (1,1) to (1,5) as the center point. The red area represents the object before the movement, and the blue area represents the object after the movement. The LDDN will extract the information of the central pixel and the surrounding pixels from the  $3 \times 3$  area and infer whether to activate the neurons in the corresponding direction. As shown in (**a**), none of the LDDN have been activated when the location of central pixel is (1,1); however, the LDDN corresponding to the rightward and the downward are activated when the location of central pixel is (1,2), as shown in (**b**).

#### 2.3. Global Direction Detection Neuron

After each LDDN has scanned for the  $M \times N$  sized receptive field, eight scanned images of the corresponding activation positions of the LDDNs will be obtained, which can be described as local motion characteristics maps. Then, the GDDN finds the LDDN with the highest times of excited positions based on the local motion characteristics maps, and the direction corresponding to this LDDN is then inferred as the global motion direction of the object.

To illustrate the function of the GDDN more clearly, how the GDDN accomplishes the motion direction detection of a U-shaped object in the background of  $6 \times 6$  grayscale is shown in Figure 8. In this case, the grayscale values of all pixels in the background are randomly selected, and the grayscale values of each pixel of the object at time *t* and  $t + \Delta t$ are the same. The activated local locations in the scanned images are indicated in red, and it is easy to observe that the LDDN responsible for detecting the lower-rightward is activated most often. It is activated at coordinates (2,2), (2,4), (4,2), (4,3), and (4,4), respectively, five times in total. After making statistics for the eight scanning results of the LDDN, the motion direction corresponding to the LDDN with the highest number of activations (lower-rightward) will be identified as the global motion direction of the object, which is consistent with reality.



**Figure 8.** (a) Demonstrates how GDDN enables motion detection of a U-shaped object moving down-rightward in a  $6 \times 6$  grayscale background. The two images (left) represent the position of the object at time *t* and time  $t + \Delta t$ , respectively, and the red areas in the eight images (right) represent the location and number of activated LDDN. (b) The red areas in this diagram represent the locations where LDDN has been activated in this receptive field; (c) radar plot of the GDDN detection results, from which it can be seen that the LDDN corresponding to the down-rightward was activated most frequently.

#### 3. Experiments and Analysis

Several experiments are conducted to further validate the accuracy and reliability of this model; the design of the experiments and the analysis of the results are described in detail in this section.

## 3.1. Experiments with Objects of Different Shapes and Sizes on a Noiseless Background

First, we plan to apply the model to the direction recognition problem performed by the computer simulation. A data set of grayscale images, each containing  $32 \times 32$  pixels, is generated as the background of the experiments. The background is filled with randomly generated grayscale pixels and does not change over time within the same set of experiments. Then, several experimental objects will be set at random positions in the background. These objects consist of several contiguous pixel points with the same randomly set grayscale values, and will perform motion in eight random directions (rightward, lower-rightward, downward, lower-leftward, leftward, upper-leftward, upward and uprightward). Two consecutive frames during this motion will be extracted and detected by the LDDN and GDDN proposed in this paper. Finally, we will compare the direction of motion of the object, and analyze the effectiveness of this model. In addition, considering that some variables such as size and motion direction of the objects may affect the experiment performance, the motion of the objects of eight different sizes (pixel scale of 1, 2, 4, 16, 32, 64, 128, 256) in eight directions are implemented as experimental examples.

The first experimental object is a dotted object of 1-pixel size. As shown in Figure 9, a dotted object moving in the direction of leftward is used as an example here, and we can observe the two consecutive frames of the experimental object moving in leftward at time *t* and  $t + \Delta t$  through Figure 9a. The scanned images provided by the eight kinds of LDDN are presented in Figure 9b, and we can easily find that only the LDDN which is responsible

10 of 16

for the leftward is excited, while all the other LDDN are resting. Thus, the GDDN infers the leftward as the global motion direction, which is consistent with the facts. The data set for the first experiment has a total of 7200, 900 in each direction, and the specific accuracy rates are shown in Table 1.



**Figure 9.** (a) In the first accuracy experiment, the position of the object in its initial state and the position of the object after moving down one pixel. (b) In the first accuracy experiment, the scanned image obtained by GDDN, which shows that only the LDDN corresponding to the downward direction is activated.

Then, several experiments are implemented for the parallelogram objects, rectangular objects, and completely irregular objects of different sizes; parts of the experimental images are shown in Figure 10. Through these experiments, it is found that the excited LDDN is generally located at the edge of the objects, which corresponds to the edge-sensitivity and the local motion direction selectivity in the human visual system, which demonstrates that the model has high biological plausibility. According to the experiment results shown in Table 1, we can find that the accuracy rate is slightly lower when the object is one pixel, as the object is too small to be confused with the background. This situation is acceptable because even human naked eye vision has difficulty in accurately determining the direction of motion of very small objects on a greyscale background. In addition, in all other cases, the model consisting of LDDN and GDDN shows a high accuracy rate in the detection of motion direction of the object with different sizes and shapes on the greyscale background. Based on the above experimental results, we can find that the model has a high accuracy rate in detecting the motion direction of different objects on a greyscale background. Moreover, as the size of the object becomes larger and the shape becomes more complex, the accuracy rate will not decrease. Thus, we are able to conclude that this model can detect the object motion direction regardless of their shapes and sizes, which demonstrates high reliability and versatility.



**Figure 10.** Experiments of a 32 pixel sized rectangular object on a greyscale background with random directional motion. (a) here represent the state of the object before motion movement, and the state of the object after motion movement. (b) here represents the location as well as number of times the eight kinds of LDDNs were activated.

**Table 1.** Experimental results of LDDN and GDDN on motion direction detection with different scales of objects. The table mainly consists of the data on object size, motion direction and accuracy, which the arrows representing the eight different motion directions.

Object Size	Motion Direction	1	$\nearrow$	$\rightarrow$	$\searrow$	$\downarrow$	$\checkmark$	$\leftarrow$	ĸ	Total
1-pixel	No. of samples	900	900	900	900	900	900	900	900	7200
	Correct No.	900	888	882	878	863	846	848	838	6943
	Accuracy	100%	98.6%	98%	97.5%	95.8%	94%	94.2%	93.1%	96.4%
2-pixel	No. of samples	3000	3000	3000	3000	3000	3000	3000	3000	24,000
	Correct No.	3000	3000	3000	2999	2999	2999	2997	2998	23,992
	Accuracy	100%	100%	100%	99.9%	99.9%	99.9%	99.9%	99.9%	99.9%
4-pixel	No. of samples	3000	3000	3000	3000	3000	3000	3000	3000	24,000
	Correct No.	3000	3000	3000	3000	3000	3000	3000	3000	24,000
	Accuracy	100%	100%	100%	100%	100%	100%	100%	100%	100%
16-pixel	No. of samples	3000	3000	3000	3000	3000	3000	3000	3000	24,000
	Correct No.	3000	3000	3000	3000	3000	3000	3000	3000	24,000
	Accuracy	100%	100%	100%	100%	100%	100%	100%	100%	100%
32-pixel	No. of samples	3000	3000	3000	3000	3000	3000	3000	3000	24,000
	Correct No.	3000	3000	3000	3000	3000	3000	3000	3000	24,000
	Accuracy	100%	100%	100%	100%	100%	100%	100%	100%	100%
64-pixel	No. of samples	3000	3000	3000	3000	3000	3000	3000	3000	24,000
	Correct No.	3000	3000	3000	3000	3000	3000	3000	3000	24,000
	Accuracy	100%	100%	100%	100%	100%	100%	100%	100%	100%
128-pixel	No. of samples	3000	3000	3000	3000	3000	3000	3000	3000	24,000
	Correct No.	3000	3000	3000	3000	3000	3000	3000	3000	24,000
	Accuracy	100%	100%	100%	100%	100%	100%	100%	100%	100%
256-pixel	No. of samples	3000	3000	3000	3000	3000	3000	3000	3000	24,000
	Correct No.	3000	3000	3000	3000	3000	3000	3000	3000	24,000
	Accuracy	100%	100%	100%	100%	100%	100%	100%	100%	100%

## 3.2. Comparison Experiments with CNN under Adding Different Types of Noise

Although LDDN and GDDN can achieve high motion direction detection accuracy on noise-free grayscale backgrounds, interference information may exist in the background in some non-ideal situations. In order to further validate the reliability of the model, several experiments are adopted to verify the noise immunity of this model. To make the results more convincing, we conduct a series of comparison experiments using the Convolutional Neuron Network (CNN) as a reference. CNN is considered to be a reasonable and convincing reference because, in recent years, it has shown impressive performance on many computer vision problems, and it was also verified in 2017 that it can be applied to motion direction detection of pedestrians, which has a high accuracy rate. A classical and representative CNN method called LeNet-5 is used as the reference in our experiments, and its construction is shown in Figure 11. A convolutional layer with  $5 \times 5$  kernel and

zero padding is used as the first layer, and the number of input and output neurons is 2 and 16, respectively; the second layer is a pooling layer with  $5 \times 5$  kernel; the third layer is still a convolutional layer with a  $5 \times 5$  kernel, but the input and output become 16 and 32; the fourth layer is also a pooling layer with a  $5 \times 5$  kernel; the fifth layer is a linear layer with  $8 \times 8 \times 32$  inputs and outputs. We generate a database of images of 10 different sizes of objects moving in random directions, 1000 of each size, for a total of 10,000 images. In the learning phase, 800 images of each size are randomly selected from the database as the learning data set. During the testing phase, the residual 2000 images will be used as the testing data set.



Figure 11. Structure of the CNN used in the comparison experiment.

For the comparison experiments, we use the same grayscale image data set as in the accuracy experiments as the background image, and different percentages of static background noise are added on them. Static background noise is added by randomly selecting any number of pixels on the grayscale background which are not in contact with the objects, and resetting their grayscale value at random. During the motion of objects, this kind of noise will not change over time. Several parts of the experimental procedure in the comparison experiments are recorded as Figure 12.

The results of the comparison experiments are summarized in Table 2, and the corresponding linear chart of accuracy rate change is also represented in Figure 13a–d. Based on the experimental results, we can easily find that both the CNN and LDDN and GDDN perform well at low noise levels. However, as we continue to increase the noise percentage, the performance of the CNN becomes poorer. When the noise percentage reaches up to 10%, the detection accuracy rate of CNN with small objects (4~128 pixels) will decrease by 50–60%, while the detection accuracy rate with large objects (256 pixels) will also decrease by 30%. When the noise percentage reaches 30%, the detection accuracy rate of CNN for the small objects has dropped to below 30%, and, for large objects, it is also below 40%, which is not very ideal. However, for the LDDN with GDDN proposed in this paper, it has a slightly higher detection accuracy due to the increase in noise percentage, its detection accuracy only has a small drop and remains above 90%, which proves that the proposed model has good noise immunity.



**Figure 12.** (a) Experimental example of an object size of four pixels used for comparison experiments; (b) experimental example of an object size of 16 pixels used for comparison experiments.

**Table 2.** Data statistics of the comparison experiments of LDDN and GDDN, and CNN, under the static background noise.

Noise Type	0%		10%		20%		30%	
Proportion	LDDN & GDDN	CNN						
1-pixel	99.18%	61.50%	91.51%	16.50%	83.44%	15.50%	77.11%	15.50%
2-pixel	99.45%	60.50%	94.39%	16.00%	88.77%	15.50%	83.63%	17.00%
4-pixel	99.85%	71.50%	98.14%	21.00%	96.37%	18.00%	94.50%	18.00%
16-pixel	100.00%	85.50%	99.96%	20.00%	99.87%	19.00%	99.72%	17.50%
32-pixel	100.00%	89.50%	99.99%	24.00%	99.91%	20.00%	99.86%	21.00%
64-pixel	100.00%	97.50%	100%	31.00%	100%	25.50%	100%	29.00%
128-pixel	100.00%	97.50%	100%	37.00%	100%	35.50%	100%	31.00%
256-pixel	100.00%	98.50%	100%	62.00%	100%	47.00%	100%	39.00%



Figure 13. Cont.



**Figure 13.** Line graph which compares the accuracy rates of LDDN and GDDN as well as CNN in comparison experiments with various percentages of noise added to the grayscale background. Among these four images, (**a**) depicts an experiment with the object of 4 pixels, (**b**) depicts an experiment with the object of 16 pixels, (**c**) depicts an experiment with the object of 128 pixels, and (**d**) depicts an experiment with the object of 256 pixels.

#### 4. Conclusions and Perspectives

An efficient mechanism for detecting the object motion direction in the grayscale background is proposed in this paper. The inspiration of this mechanism comes from the research on DSGCs, which assumes that organisms can process information about object motion direction in the retinal layer. Though the specific biological mechanism for motion direction detection of human beings still remains a black box, we refer to the physiological role of individual neurons and propose the LDDN and GDDN as a possibility. Eight kinds of LDDN preferred in different directions have been used to scan over a two-dimensional image in this mechanism; then, the GDDN will analyze the statistics of the scanned results and infer the global motion direction of the object.

A series of experiments are conducted to verify the validity and reliability of this mechanism. Experimental results demonstrate that the mechanism can accurately detect the motion direction of an object on grayscale background regardless of its shapes, sizes, and conditions. In addition, we also conduct comparison experiments with CNN to further demonstrate the superiority of this mechanism. The experimental results show that the proposed mechanism performs better in detecting accuracy and noise resistance than CNN on the motion direction detection of objects on a grayscale background. Furthermore, this mechanism is designed according to the real retinal configuration of the living creature, which is more credible and acceptable than CNN. We believe that this mechanism presents an acceptable and explainable possibility for the black box of motion direction detection mechanism in humans, and can provide further inspiration for the study of the human visual and cerebral nervous systems.

In future work, this research will be extended in various ways. In terms of the application scope, we will extend the mechanism to motion direction detection on colored backgrounds. Since colored images can be considered as a special superposition of greyscale images, detection of object motion on a greyscale background can be the basis for future motion detection on a colored background. In terms of functionality, this mechanism will be extended to orientation detection and motion velocity detection of the objects. We believe

that this mechanism of local detection paired with global detection can be applied to other functions as well. Furthermore, the application of this mechanism in 3D scenes is also in the future research program, and the practical application value of this mechanism is also worth studying.

**Author Contributions:** Conceptualization, Y.H., Z.T. and Y.T.; methodology,Y.H., Z.T. and Y.T.; software, S.T., Y.H., B.L. and R.I.; validation, Y.H., Z.T. and Y.T.; writing—original draft preparation, Y.H.; writing—review and editing, Y.H., Z.T. and Y.T.; visualization, Y.H. and S.T.; supervision, Z.T. and Y.T.; project administration, Z.T. and Y.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

**Conflicts of Interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Abbreviations

The following abbreviations are used in this manuscript:

- MDPI Multidisciplinary Digital Publishing Institute
- DOAJ Directory of Open Access Journals
- TLA Three Letter Acronym
- LD Linear Dichroism

## References

- 1. Fasel, B. An introduction to bio-inspired artificial neural network architectures. *Acta Neurol. Belg.* 2003, *103*, 6–12. [PubMed]
- Kim, K.; Lee, S.; Kim, J.Y.; Kim, M.; Yoo, H.J. A 125 GOPS 583 mW network-on-chip based parallel processor with bio-inspired visual attention engine. *IEEE J. Solid-State Circuits* 2008, 44, 136–147. [CrossRef]
- 3. Lindsay, G.W. Convolutional neural networks as a model of the visual system: Past, present, and future. *J. Cogn. Neurosci.* 2021, 33, 2017–2031. [CrossRef]
- 4. Medina, J. Brain Rules: 12 Principles for Surviving and Thriving at Work, Home, and School; Pear Press: Seattle, WA, USA 2011.
- 5. Viviani, P.; Aymoz, C. Colour, form, and movement are not perceived simultaneously. Vis. Res. 2001, 41, 2909–2918. [CrossRef]
- 6. Mauss, A.S.; Vlasits, A.; Borst, A.; Feller, M. Visual circuits for direction selectivity. Annu. Rev. Neurosci. 2017, 40, 211. [CrossRef]
- 7. Barlow, H.B. Summation and inhibition in the frog's retina. J. Physiol. 1953, 119, 69. [CrossRef] [PubMed]
- 8. Kuffler, S.W. Discharge patterns and functional organization of mammalian retina. J. Neurophysiol. 1953, 16, 37–68. [CrossRef]
- 9. Hassenstein, B.; Reichardt, W. Systemtheoretische analyse der zeit-, reihenfolgen-und vorzeichenauswertung bei der bewegungsperzeption des rüsselkäfers chlorophanus. Z. Naturforschung B **1956**, *11*, 513–524. [CrossRef]
- 10. Yan, C.; Todo, Y.; Kobayashi, Y.; Tang, Z.; Li, B. An Artificial Visual System for Motion Direction Detection Based on the Hassenstein–Reichardt Correlator Model. *Electronics* **2022**, *11*, 1423. [CrossRef]
- 11. Hubel, D.H.; Wiesel, T.N. Receptive fields of single neurones in the cat's striate cortex. J. Physiol. 1959, 148, 574. [CrossRef]
- 12. Barlow, H.; Levick, W.R. The mechanism of directionally selective units in rabbit's retina. *J. Physiol.* **1965**, *178*, 477. [CrossRef] [PubMed]
- 13. Borst, A.; Egelhaaf, M. Principles of visual motion detection. Trends Neurosci. 1989, 12, 297-306. [CrossRef]
- Clifford, C.W.; Ibbotson, M. Fundamental mechanisms of visual motion detection: Models, cells and functions. *Prog. Neurobiol.* 2002, 68, 409–437. [CrossRef]
- 15. Taylor, W.R.; He, S.; Levick, W.R.; Vaney, D.I. Dendritic computation of direction selectivity by retinal ganglion cells. *Science* 2000, 289, 2347–2350. [CrossRef] [PubMed]
- Todo, Y.; Tang, Z.; Todo, H.; Ji, J.; Yamashita, K. Neurons with multiplicative interactions of nonlinear synapses. *Int. J. Neural Syst.* 2019, 29, 1950012. [CrossRef]
- 17. Tang, Y.; Ji, J.; Gao, S.; Dai, H.; Yu, Y.; Todo, Y. A pruning neural network model in credit classification analysis. *Comput. Intell. Neurosci.* **2018**, 2018, 9390410. [CrossRef]
- 18. Song, Z.; Tang, Y.; Ji, J.; Todo, Y. Evaluating a dendritic neuron model for wind speed forecasting. *Knowl.-Based Syst.* **2020**, 201, 106052. [CrossRef]
- 19. Han, M.; Todo, Y.; Tang, Z. Mechanism of Motion Direction Detection Based on Barlow's Retina Inhibitory Scheme in Direction-Selective Ganglion Cells. *Electronics* **2021**, *10*, 1663. [CrossRef]
- 20. Dominguez-Sanchez, A.; Cazorla, M.; Orts-Escolano, S. Pedestrian movement direction recognition using convolutional neural networks. *IEEE Trans. Intell. Transp. Syst.* 2017, *18*, 3540–3548. [CrossRef]

- 21. Euler, T.; Detwiler, P.B.; Denk, W. Directionally selective calcium signals in dendrites of starburst amacrine cells. *Nature* 2002, 418, 845–852. [CrossRef]
- 22. Koch, C.; Poggio, T.; Torre, V. Retinal ganglion cells: A functional interpretation of dendritic morphology. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **1982**, *298*, 227–263. [PubMed]
- 23. Poggio, T.; Torre, V.; Koch, C. Computational vision and regularization theory. In *Readings in Computer Vision*; Elsevier: Amsterdam, The Netherlands, 1987; pp. 638–643.
- 24. Todo, Y.; Tamura, H.; Yamashita, K.; Tang, Z. Unsupervised learnable neuron model with nonlinear interaction on dendrites. *Neural Netw.* **2014**, *60*, 96–103. [CrossRef]
- 25. Schnapf, J.L.; Baylor, D.A. How photoreceptor cells respond to light. Sci. Am. 1987, 256, 40–47. [CrossRef] [PubMed]
- Chapot, C.A.; Euler, T.; Schubert, T. How do horizontal cells 'talk' to cone photoreceptors? Different levels of complexity at the cone–horizontal cell synapse. J. Physiol. 2017, 595, 5495–5506. [CrossRef] [PubMed]
- 27. Aranda, M.L.; Schmidt, T.M. Diversity of intrinsically photosensitive retinal ganglion cells: Circuits and functions. *Cell. Mol. Life Sci.* **2021**, *78*, 889–907. [CrossRef] [PubMed]