photonics

MDPI

Article

Deep Learning-Based Robust Visible Light Positioning for High-Speed Vehicles

Danjie Li [1,2,3,4,†], Zhanhang Wei [1,2,3,4,†], Ganhong Yang [1,2,3,4], Yi Yang [1,2,3,4], Jingwen Li [1,2,3,4], Mingyang Yu [1,2,3,4], Puxi Lin [1,2,3,4], Jiajun Lin [1,2,3,4], Shuyu Chen [1,2,3,4], Mingli Lu [5], Zhe Chen [1,2,3,4], Zoe Lin Jiang [6] and Junbin Fang [1,2,3,4,*]

1 Guangdong Provincial Key Laboratory of Optical Fiber Sensing and Communications, Guangzhou 510632, China
2 Guangdong Provincial Engineering Technology Research Center on Visible Light Communication, Guangzhou 510632, China
3 Guangzhou Municipal Key Laboratory of Engineering Technology on Visible Light Communication, Guangzhou 510632, China
4 Department of Optoelectronic Engineering, Jinan University, Guangzhou 510632, China
5 Academic Affairs Office of Beijing Vocational College of Agriculture, Beijing 102442, China
6 School of Computer Science and Technology, Harbin Insitute of Technology, Shenzhen 518055, China
* Correspondence: tjunbinfang@jnu.edu.cn
† These authors contributed equally to this work.

Abstract: Robustness is a key factor for real-time positioning and navigation, especially for high-speed vehicles. While visible light positioning (VLP) based on LED illumination and image sensors is widely studied, most of the VLP systems still suffer from the high positioning latency and the image blurs caused by high-speed movements. In this paper, a robust VLP system for high-speed vehicles is proposed based on a deep learning and data-driven approach. The proposed system can significantly increase the success rate of decoding VLP-LED user identifications (UID) from blurred images and reduce the computational latency for detecting and extracting VLP-LED stripe image regions from captured images. Experimental results show that the success rate of UID decoding using the proposed BN-CNN model could be higher than 98% when that of the traditional Zbar-based decoder falls to 0, while the computational time for positioning is decreased to 9.19 ms and the supported moving speed of our scheme can achieve 38.5 km/h. Therefore, the proposed VLP system can enhance the robustness against high-speed movement and guarantee the real-time response for positioning and navigation for high-speed vehicles.

Keywords: visible light positioning; high speed; deep learning; motion blur; diffusion blur

Citation: Li, D.; Wei, Z.; Yang, G.; Yang, Y.; Li, J.; Yu, M.; Lin, P.; Lin, J.; Chen, S.; Lu, M.; et al. Deep Learning-Based Robust Visible Light Positioning for High-Speed Vehicles. *Photonics* **2022**, *9*, 632. https://doi.org/10.3390/photonics9090632

Received: 12 July 2022
Accepted: 27 August 2022
Published: 2 September 2022

1. Introduction

Visible light positioning (VLP) is considered a promising technology for indoor positioning since it can simultaneously provide the functionalities of high-accuracy positioning and illumination. In recent years, with the increasing applications of intelligent mobile machines, such as mobile robots and automated guided vehicles, the potential application of VLP is extended from human indoor navigation to machine navigation in indoor environment, underground space or tunnels [1–3]. According to the devices for receiving light signal, current VLP technologies can be divided into two categories: image sensor (IS)-based VLP or photodiode (PD)-based VLP. Compared with PD-based VLP systems [4–10], IS-based VLP is easier to implement or integrate with current mobile terminals and mobile equipment, which are usually equipped with high-resolution complementary metal–oxide–semiconductor (CMOS) sensor cameras. Taking advantage of high resolution, complementary metal–oxide–semiconductor (CMOS) sensor cameras are utilized in most of the VLP systems to receive the light signals transmitted from VLP-LED lamps and

to obtain angle-of-arrival (AOA) information for position calculation with triangulation functions [1,11,12]. Furthermore, with the wide use of CMOS sensor cameras, IS-based VLP is easier to integrate with current mobile terminals and mobile machines.

Utilizing the rolling shutter effect of CMOS sensor cameras, image sensor-based (IS-based) VLP systems usually receive ON/OFF light signals in time sequence as bright/dark stripes in a captured image. Successful decoding the VLP-LED stripes can identify the unique ID (UID) sent from VLP-LED lamps and then the precise position can be calculated, while decoding failures will cause positioning failures. Positioning failures are harmful to real-time positioning and navigation for moving objects, especially for high-speed applications. Once a position failure occurs, an IS-based VLP system needs to reacquire a new image and retry decoding. As a consequence, the VLP system will introduce additional positioning latency and cannot support real-time positioning for high-speed moving objects. Therefore, the real-time performance of IS-based VLP depends not only on the positioning time, but also on the positioning robustness (i.e., the success rate of positioning).

Most of the IS-based VLP systems focus on increasing the positioning accuracy or reducing the positioning time to improve the real-time response for positioning [13–18]. Few researchers have begun to study the robustness problem of VLP systems, although it is more critical for the practicality and availability of VLP systems, especially for high-speed moving objects. Xie et al. proposed a proximity-based visible light localization LED-ID detection and recalibration method to improve the robustness of their system by using a machine learning approach (Fisher classifier or linear support vector machine) to increase the success rate of identifying the UID of VLP-LED lamps from the captured images [19]. However, their work only considers the low-speed scenarios where the captured images are not deteriorated by high-speed motion effects. Additionally, their system can only support low-speed moving objects due to the high computational time. In fact, for high-speed vehicles (e.g., autonomous vehicles in tunnels), IS-based VLP systems will encounter some practical issues including motion blur, diffusion blur, etc. [20–22]. Motion blur or diffusion blur caused by high-speed movement will deteriorate the quality of a captured image, as well as the success rate of identifying or decoding the VLP-LED UID and then the accuracy of the VLP system.

Since the traditional decoding algorithm is not effective for these blurred images, in this work, a robust VLP system for high-speed vehicles is proposed based on a deep learning (DL) and data-driven approach. First, a DL framework with an eight-layer batch-normalized convolutional neural network (BN-CNN) is proposed to increase the success rate of identifying VLP-LED UIDs from the captured image with motion blur caused by high-speed movement. The batch normalization layer can accelerate the training of the neural network and improve the generalization ability of it, as well as the recognition accuracy of LED striped images [23,24]. Second, the DL framework is further trained to identify VLP-LED UIDs from images with diffusion blur, which may be introduced by out-of-focus blur due to rapid movement or smog in tunnels, etc. Third, a lightweight fast region of interest (ROI) detection algorithm is proposed to decrease the computational time of extracting the VLP-LED stripe image region from a captured image. Therefore, the proposed VLP system is robust and fast for real-time positioning and navigation for high-speed vehicles.

Experimental results show that the success rate of UID decoding using the proposed BN-CNN model could reach 98.4% when the moving speed is larger than 6.9 km/h, respectively, while that of the traditional Zbar-based decoder falls to 0. For diffusion-blurred images, the success rate of UID decoding using the proposed BN-CNN model could reach 98.5% when the diffusion radius is larger than 4 pixels, while that of the traditional Zbar-based decoder falls to 0. Moreover, the computational time for positioning is decreased to 9.19 ms and the supported moving speed of our scheme can achieve 38.5 km/h.

The rest of this paper is organized as follows. Section 2 introduces the system architecture, the proposed DL-based decoder with BN-CNN model, and the lightweight ROI

detection algorithm of the proposed DL-based robust VLP system. The experimental setup and results are shown and discussed in Section 3. Section 4 concludes the work.

2. The Proposed DL-Based Robust VLP Systems

2.1. System Architecture

As shown in Figure 1, the proposed DL-based robust VLP system mainly includes two parts: the VLP-LED lamp as the visible light signal transmitter and the mobile terminal with a VLP module equipped with a CMOS sensor camera as the visible light signal receiver and position estimator. At the transmitter side, the unique ID (UID) of a VLP-LED lamp is first encoded into an interleaved-two-five (ITF) codeword, which is further used to modulate the VLP-LED lamp with on–off keying (OOK) modulation to transmit ON/OFF visible light signals. The UID of each VLP-LED lamp is coded as a 4-digit decimal number, and the number on each bit is independently coded as a 5-digit binary number. One frame includes 20 data bits and 1 stop bit, and we use different modulation times to transmit each bit of data. Since the modulation frequency of the ON/OFF signals is high enough, it will not cause flickers perceivable to human eyes. In addition, the VLP-LED lamp with this OOK modulation always shows horizontal stripes in the image because the exposure sequence of CMOS image sensors is always top-down exposure line by line due to the rolling shutter mechanism.

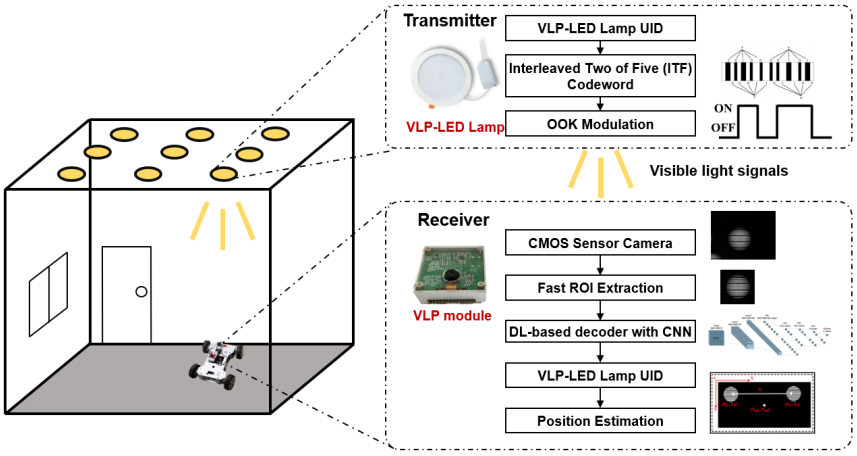

Figure 1. The proposed DL-based robust VLP system.

The visible light signal is transmitted through an air channel. At the receiver side, the transmitted visible light signals in time sequence are received with a rolling shutter CMOS sensor camera as bright or dark stripes within the VLP-LED lamp region in a captured image. A lightweight image processing algorithm is proposed to detect and extract the ROI, i.e., the VLP-LED stripe image, from the captured image quickly. The lightweight fast ROI extraction can decrease the computational time of image processing as well as the positioning latency. After that, the proposed DL-based decoder with BN-CNN is utilized to identify the UID from the extracted VLP-LED stripe image region. The proposed DL-based decoder is composed of BN-CNN with 8 layers and it can significantly increase the success rate of identifying UIDs from captured images with high-speed motion blurs and diffusion blurs. Therefore, the robustness of the proposed VLP system for high-speed vehicles can be greatly improved. The proposed VLP system can also alleviate the problem of stripes width variation caused by interference between the bright or dark stripes. Assuming that the image captured contains the fringe images of two or more VLP LEDs, the image preprocessing algorithm can be used to identify and cut out the fringe images of each VLP LED, and then put them into BN-CNN successively to obtain the UID corresponding to the VLP LEDs. Through the accurate decoding result of BN-CNN, we can use the triangulation algorithm to accurately obtain the positioning coordinate of the vehicle in the system.

2.2. The Proposed DL-Based Decoder with BN-CNN

For motion-blurred or diffusion-blurred VLP-LED strip images, conventional image processing algorithms decode by screening the information of each pixel, which is severely disturbed by blurred pixels, and the decoding success rate of conventional image processing algorithms will be greatly reduced once clear imaging is not possible.

Compared with other models of deep learning, CNN can directly use the image as the input of the network to establish the mapping relationship between the binary data and the classification results. Therefore, we use CNN and add the Batch Normalization layers which can effectively solve the vanishing gradient problem of CNN to build the VLP decoder model. The convolutional neural network structure proposed in this paper uses multilayer convolution to expand the perceptual field of feature data, so as to regionally extract image features, overcome the interference of blurred pixels to information pixels in motion-blurred and diffusion-blurred VLP images, and correctly classify the LED-UID corresponding to the image with the training model to improve the recognition accuracy of the VLP system for blurred images and solve the VLP system's robustness problem for application on high-speed moving vehicles. The advantage of the proposed BN-CNN decoding in this work is the use of a batch normalization layer and a dropout layer to prevent model overfitting. The batch normalization layer and dropout layer enable the model to accurately identify LED images with different blur forms and different blur levels, and improve the robustness of the VLP system in case of high-speed movement and image sensor focus adjustment time delay.

As shown in Figures 2 and 3, the proposed BN-CNN model mainly includes 8 layers.

1. The first layer is the Input layer, which is read in VLP-LED stripe images with the size of 800 × 800 pixels.
2. In Layer Conv1, the convolutional layer uses 32 convolutional kernels of 3 × 3 size with the step of (1, 1) to extract features and generate 32 feature maps. Let the input array be X_i and the output array be Y_i, then, the convolutional layer performs the extraction of feature maps according to the following equation.

$$Y_i = b_i + \sum_j W_{ji} * X_i \tag{1}$$

where b_i is the bias of the neuron, $w_j i$ is the weight of the neuron, and $*$ denotes the convolution operations.

After the feature mapping in the convolutional layer, we use zero to fill the edge pixels of the 798 × 798 feature maps and obtain 32 feature maps with the size of 800 × 800. A batch normalization layer is utilized to normalize the feature maps output from the convolution layer and then the average value and variance of the feature maps are limited to the range of [0, 1]. It helps accelerate the convergence speed of the proposed BN-CNN model and improve the generalization ability of the model.

Here, the batch normalization layer introduces the mean and variance of each batch into the convolutional neural network, while the mean and variance of different batches are generally different. Therefore, the batch normalization layer is equivalent to adding random noise to the process of convolutional neural network training, which can play a role in preventing model overfitting. On the other hand, the batch normalization layer adjusts the input of any neuron of the neural network to a standard normal distribution with mean 0 and variance 1, so that the input value of the activation layer falls in the region where the nonlinear function is more sensitive to the input, i.e., a smaller input change leads to a larger gradient change, which avoids the vanishing gradient problem and speeds up the convergence of the neural network. Further, with the excitation function of the RELU activating layer, the output of the batch normalization layer is mapped nonlinearly into the max pooling layer in Layer Conv1. The max pooling layer uses a max pooling with the step of (2, 2) to calculate and output the maximum value of the data corresponding to the sliding process of

the polling window. The max pooling layer implements a separate dimensionality reduction of each feature map to reduce the connection between layers, as well as the amount of data to be computed and stored. It also reduces the risk of model overfitting. After the max pooling layer, the size of the output feature map is reduced to 400 × 400.

3. Layer Conv2 has a similar structure to that of Layer Conv1, while it employs 64 convolutional kernels of 3 × 3 size with the step of (1, 1) to further extract higher-dimension features from the 32 feature maps generated by Layer Conv1 and to generate 64 feature maps. After the batch normalization layer, RELU activation layer and max pooling layer in Layer Conv2, 64 feature maps with the size of 200 × 200 will be generated for next processing.
4. In Layer M1, the dropout layer is introduced to randomly discardsome hidden neurons to avoid overfitting of the training model. Through randomly discarding 50% of hidden neurons during forward propagation, the dropout layer can reduce the complex co-adaptive relationships between neurons and avoid relying on the linkage relationships between neurons. As a result, the dropout layer can alleviate the occurrence of overfitting and improve the robustness of CNN training. Additionally, randomly discarding some neurons in the network is equivalent to averaging over many different neural networks and it can significantly improve the generalization ability of the CNN model.
   After the dropout layer, the flatten layer transforms the 64 × 200 × 200 feature maps into a 64 × 200 × 200 one-dimensional feature array, which will be fed into the fully connected layer.
5. In Layer FC1, the fully connected layer uses 512 connected nodes to convert the input feature array into 512 scored values, which are propagated forward to the batch normalization layer. The batch normalization layer is used to adjust the input of the RELU activation layer to a standard normal distribution, which can speed up network training and prevent the vanishing gradient problem. The normalized scored values are input into the RELU activation layer to add nonlinear factors into the neural network, which makes the neural network can adapt to more complex problems.
6. Layer M2 only contains a dropout layer, which is used to further reduce the co-adaptive relationship between the neurons of the neural network. There is no need for adding a flatten layer between layer FC1 and layer FC2.
7. Layer FC2 uses a fully connected layer and a batch normalization layer, as in the layer FC1. After the last fully connected layer with the number of connected nodes of N, the feature values are finally output as N classification results corresponding to N LED-UIDs.
8. Finally, we input the N classification outputs from layer FC2 to the Output layer. In order to obtain the accurate results of the 9 classification outputs for backpropagation and simplify the calculation of the loss function, we use softmax to calculate in the output layer. The output layer uses a softmax classifier to convert the N classification outputs into a classification percentage that sums to 1. The softmax classifier equation is:

$$f(V_i) = \frac{e^{(}V_i)}{\sum_i e^{(}V_i)} \tag{2}$$

V_i is the ith input signal in the output layer, the denominator indicates that there are j output signals (neurons) in the output layer, and the exponential sum of all the input signals in the output layer is calculated. $f(V_i)$ is the output of the ith neuron, and the formula is used to calculate the probability distribution of the original input image data after feature extraction of the convolutional neural network to finally obtain the closest to each identifier. The probability distribution is then used in the classifier to obtain the loss value L of the current model calculated according to Equation (3) to

back-propagate the convolution kernel (weight matrix) of the optimized convolution layer so that the loss value L keeps decreasing.

$$L = -\sum_i Y_i * log\hat{Y}_i \quad (3)$$

After the training process of continuous gradient descent, we obtain the BN-CNN model that can classify the training set accurately, and the data of the test set are used to obtain the classification probability of the test set images for the N identifiers through the feature extraction of this model, and the string of identifiers is converted into the byte code output, i.e., LED-UIDs, so as to achieve accurate recognition of the images of the VLP system.

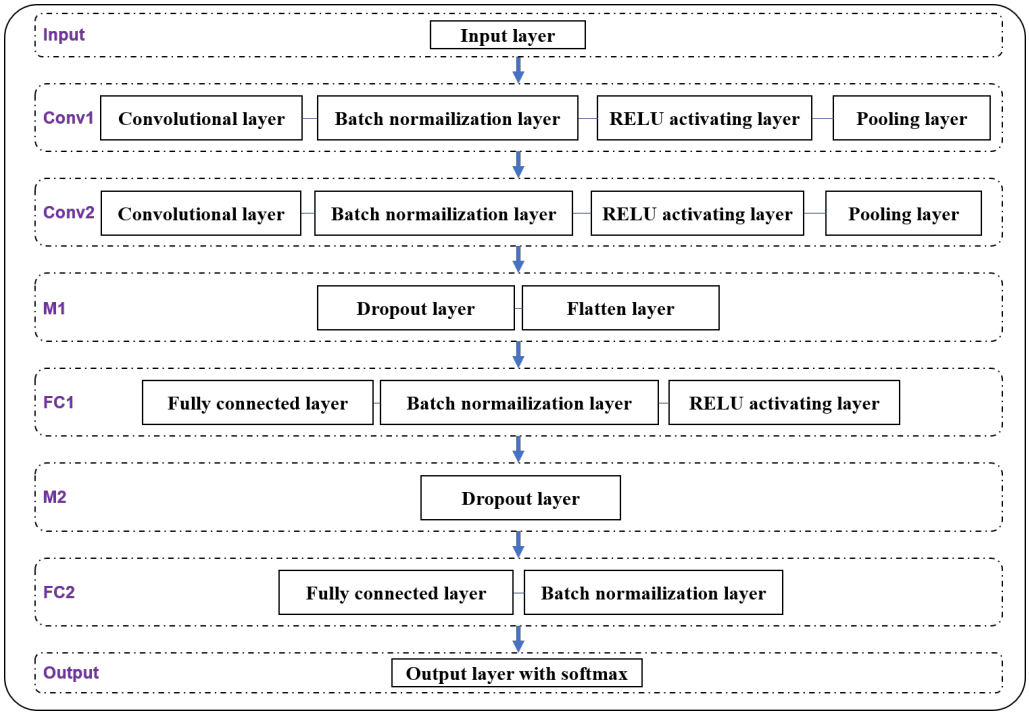

Figure 2. Layers of the proposed BN-CNN model.

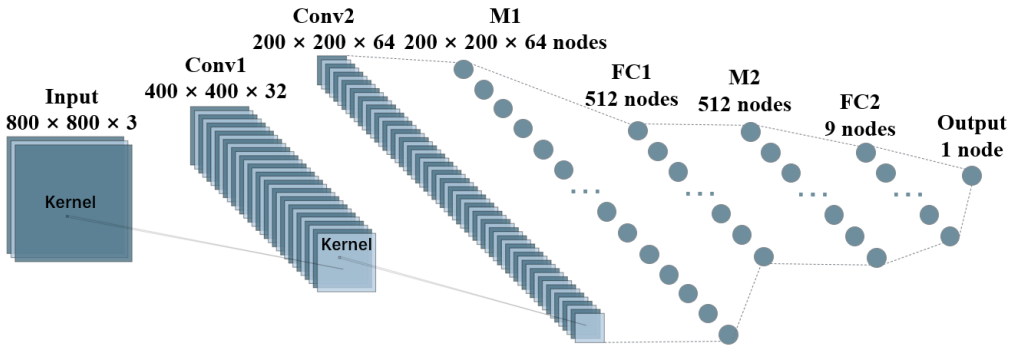

Figure 3. Data dimension of the proposed BN-CNN model.

2.3. Lightweight Fast ROI Detection Algorithm for Extracting VLP-LED Stripe Images

In this paper, we propose a lightweight fast ROI detection algorithm to extract VLP-LED stripe images. The proposed ROI detection algorithm searches the edges of VLP-LED stripe image regions using a variable step size in a back-and-forth searching mode, instead of column-by-column and pixel-by-pixel. The improved ROI detection algorithm can reduce the time for detecting VLP-LED stripe image regions and speed up the VLP-LED

stripe image extraction. Therefore, the positioning speed of the proposed VLP system can be significantly increased to provide real-time positioning for high-speed vehicles. As shown in Figure 4, the details of the proposed algorithm are as follows.

1. First, the captured image is scanned from left to right with an initial step length (denoted as la, which is 9 pixels in this paper) to detect whether the column under scanning contains enough number of bright pixels. If so, the column is inside a VLP-LED stripe image region since the low exposure setting of the CMOS camera makes sure that only the VLP-LED stripe image region contains bright pixels;
2. Then, from the "bright" column which is detected, we use a smaller step length, denoted as lb, which is 1 pixel in this paper)to scan the image back and forth to precisely determine the left edge and the right edge of the VLP-LED stripe image region;
3. Similarly, the upper edge and the lower edge of the VLP-LED stripe image region can be determined. The VLP-LED stripe image can also be extracted, as shown in Figure 5.

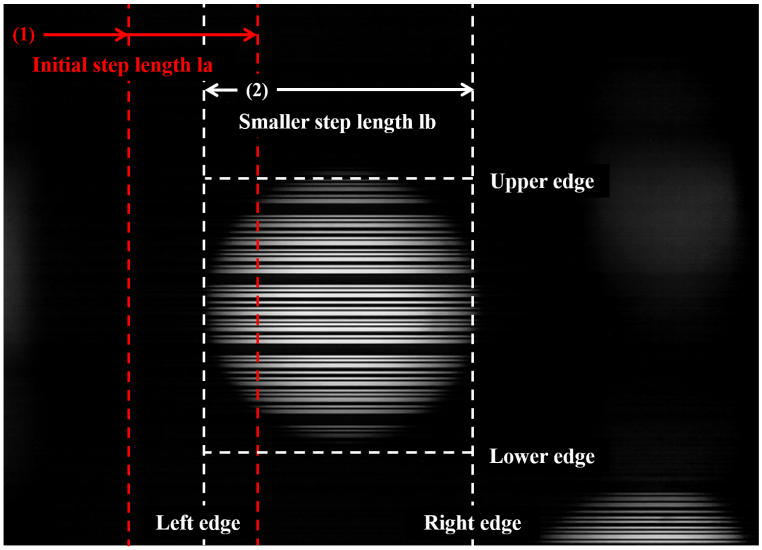

Figure 4. Lightweight ROI detection algorithm for extracting VLP-LED stripe image region.

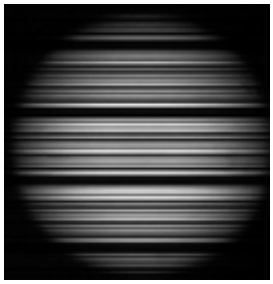

Figure 5. The extracted VLP-LED stripe image.

Therefore, the computational complexity of detecting the VLP-LED stripe image region can be reduced to a several-times-lower level depending on the value of the initial step length (la). In theory, the lightweight fast ROI detection algorithm can increase the detection speed by more than 9 times compared with the original pixel-by-pixel detection algorithm by reducing the scanning of non-informative pixels.

3. Experimental Results

3.1. Experimental Setup

A series of experiments were conducted to evaluate the performance of the proposed robust DL-based VLP system for blurred images in terms of the success rate of UID decoding and the computational time of decoding.

As shown in Figures 6 and 7, 9 VLP-LED lamps are installed on the roof of a shelf with a height of 175 cm and a size of 181 cm × 181 cm. The VLP-LED lamps are 17.5 cm diameter commercial LED lamps with our self-designed VLP modulator and the distance between two nearby VLP-LED lamps is 62 cm. A Raspberry Pi 3B+ development kit (Raspberry Pi Foundation, Cambridge, The United Kingdom of Great Britain and Northern Ireland) with a 1600 × 1200 pixels CMOS camera is used as the VLP module to capture images containing VLP-LED stripe images. The processor of the Raspberry Pi 3B+ development kit is a Quad Core 1.2GHz Broadcom BCM2837 64bit CPU. The CMOS camera isOmniVision's OV5647 CMOS image sensor chip equipped with a lens with a diagonal field of view (FOV) of 60.6 degrees (Waveshare, Shenzhen, China).

Figure 6. The VLP-LED lamp.

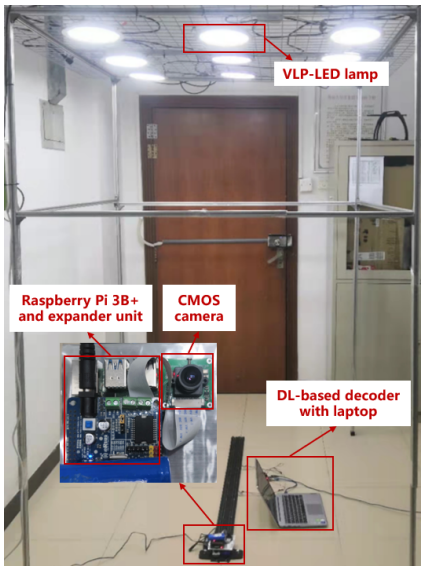

Figure 7. Experimental environment andhardware.

For each VLP-LED lamp, the VLP module was placed at random positions around and under the lamp in four orthogonal directions, and 50 images were collected for each direction. We cropped the image to change the size of the input spots to improve the robustness of the CNN model, which is heavily influenced by the training set, so as to consider the light spot captured in more situations. Therefore, the image dataset used in our experiments contains 50 × 4 × 9 = 1800 images. Note that all the 1800 images are clear, without any motion blur or diffusion blur.

In our experiments, Tensorflow gpu version 1.4.0 (Google Brain, Santa Clara Country, CA, USA), Keras version 2.2.5 (Google, Santa Clara Country, CA, USA), CUDA version 10.0 (NVIDIA, Santa Clara Country, CA, USA) and cudnn version 7.6.4.38 (NVIDIA, Santa Clara Country, CA, USA) were used for the training and testing of the proposed BN-CNN models.

3.2. Success Rate of UID Decoding for Normal Clear Images

We used 1440 VLP-LED stripe images (160 images per VLP-LED) as the train set to train the BN-CNN model described in Section 2.2 and took the remaining 360 images as the test set to evaluate the success rate of UID decoding of the proposed DL approach, comparing with that of traditional decoding algorithms.

As shown in Figure 8, the loss value in the training set (denoted as train loss) and the loss value in the test set (denoted as test loss) drop sharply after 10 epochs and converge to 0.235 and 0.340 after 100 epochs. Meanwhile, the success rate of UID decoding in the training set (denoted as train acc) and the success rate of UID decoding in the test set (denoted as test acc) also reach 0.963 and 0.986 after 80 epochs and reach 0.999 and 0.992 after 100 epochs. The data in Figure 8 show that the proposed BN-CNN model has excellent performances of high accuracy and fast convergence.

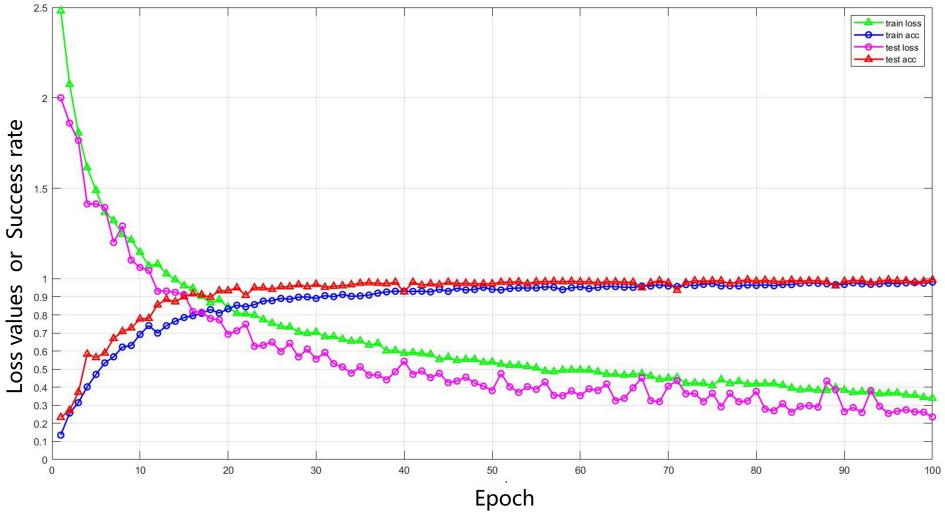

Figure 8. Loss values and success rate of the proposed BN-CNN model.

The 1800 images in the dataset were also tested using a traditional UID decoder based on Zbar [25], which is a popular open source software suite for reading barcodes from various sources. The success rates of UID decoding using the proposed BN-CNN versus that using the traditional algorithm are shown in Table 1. Since Zbar is a robust barcode reader, the Zbar-based decoder can successfully identify the UID in the stripe images with a rate of 99.9%, i.e., only two images cannot be correctly decoded. As for the proposed BN-CNN, the success rate for training dataset is 99.9%, while the success rate for test dataset is 99.2%. The average success rate of using the proposed BN-CNN is 99.7%, which is close to that of the the Zbar-based decoder.

Table 1. The success rate of UID decoding using the proposed BN-CNN versus the Zbar-based decoder.

UID Decoding Method	Success Rate
The proposed BN-CNN for 1440 images in train set	99.9%
The proposed BN-CNN for 360 images in test set	99.2%
Zbar-based decoder for 1800 images in data set	99.9%

3.3. Success Rate of UID Decoding for Motion-Blurred Images

Motion blur is the blur seen in moving objects in a photograph or a single frame of film or video. It happens because objects move and the image being recorded changes during the recording of a single exposure, due to rapid movement or long exposure. Since the VLP-LED lamps are usually fixed, while the VLP module is installed on a vehicle moving fast, the high-speed movement will introduce motion blur into the images captured by the VLP module. As shown in Figure 9, an example with moving speed of 5.0 km/h and the length of movement in the image of 12 pixels is constructed to verify this phenomenon. To assess the motion blur caused by high-speed movements with the system proposed in this paper and make sure the moving speed of the vehicle is up to 38.5 km/h, we used Matlab's 2-D filter function to simulate the motion blur with different motion length in pixels and used the vehicle speed detection approaches to evaluate the speed we need corresponding to the length of movement in the motion blur images due to physical experimental equipment limitations [26–28]. Using Equation (4), the speed of the vehicle was calculated with the following parameters: focal length f = 3.37 mm, CMOS pixel size s_x = 1.4 µm, shutter speed T = 3.4 ms, distance between vehicle and LED z = 175 cm. The estimation vehicle speeds of the motion blur images were shown in Table 2.

$$v = \frac{zKs_x}{Tf} \tag{4}$$

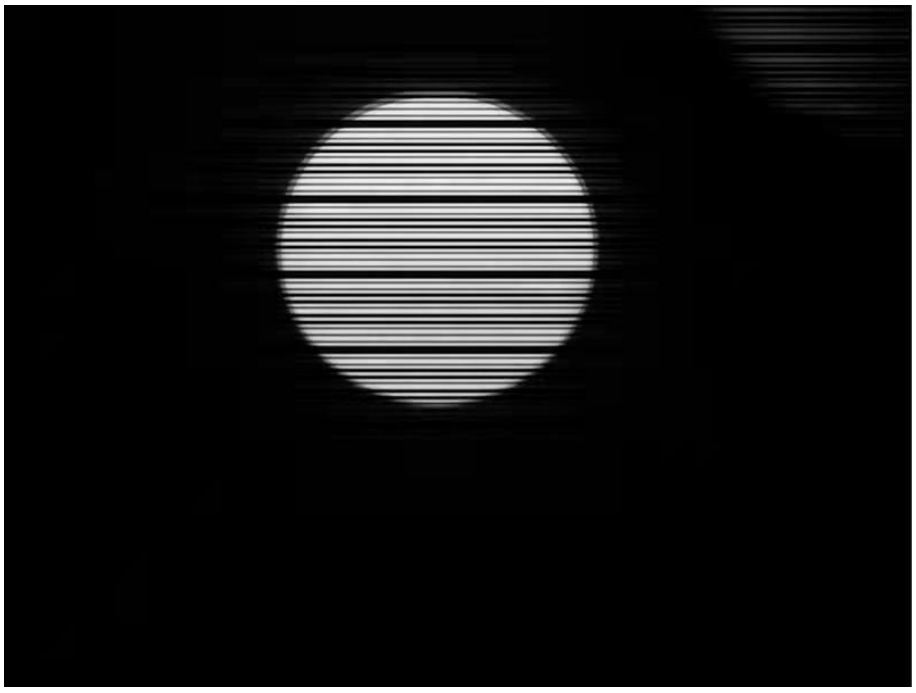

Figure 9. The VLP-LED stripe images with a motion blur length of 12 in pixels.

As shown in Table 2 and Figure 10, motion blur in two directions (diagonal direction and vertical direction to the stripes) with 21 kinds of motion blur lengths was introduced into each image in the original dataset. Therefore, we have a total of 2 × 21 × 1800 = 75,600 motion-blurred images to test the robustness of the proposed DL-based VLP system. Note that the motion blur length varies from 2 pixels to 15 pixels with 1-pixel increment and from 15 pixels to 50 pixels with 5-pixel increment.

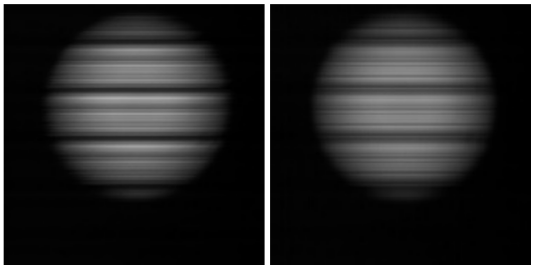

Figure 10. The VLP-LED stripe images with motion blur in diagonal direction and vertical direction to the stripes and with a motion blur length of 40 in pixels.

Table 2. The vehicle speed of the motion blur images using the vehicle speed detection approach.

Diagonal Motion Blur Length K (Pixel)	v_y (km/h)
2	1.5
3	2.3
4	3.0
5	3.8
6	4.6
7	5.4
8	6.2
9	6.9
10	7.7
11	8.5
12	9.2
13	10.0
14	10.7
15	11.5
20	15.4
25	19.2
30	23.1
35	26.9
40	30.8
45	34.6
50	38.5

As shown in Figures 11 and 12, for diagonal motion-blurred images, when the motion blur length reaches 9 pixels, the success rate of UID decoding using the traditional Zbar-based decoder falls to 0, while that of using the proposed BN-CNN model remains 98.4%. Furthermore, the success rate of UID decoding using the proposed BN-CNN model remains higher than 90.9% until the motion blur length reaches 40 pixels and it is still higher than 60.9% even when the motion blur length reaches 100 pixels. For vertical motion-blurred images, when the motion blur length reaches 7 pixels, the success rate of UID decoding using the Zbar-based decoder falls to 0, while that of using the proposed BN-CNN model remains 98.4%. Additionally, the success rate of UID decoding using the proposed BN-CNN model remains higher than 87.8% until the motion blur length reaches 40 pixels, and it is still higher than 64.4% even when the motion blur length reaches 100 pixels.

To conclude, the proposed BN-CNN model is robust for VLP-LED stripe images with motion blur introduced by high-speed movement, compared with the traditional Zbar-based decoder. Note that the parameters of the proposed BN-CNN model used in this experiment are the same as those used in the experiment described in Section 3.2, i.e., the proposed BN-CNN model trained by clear images still has a high success rate for images with heavy motion blur. It means that the proposed BN-CNN model has high robustness and adaptivity for motion-blurred images, and then the robustness of IS-based VLP systems is improved for high-speed vehicles.

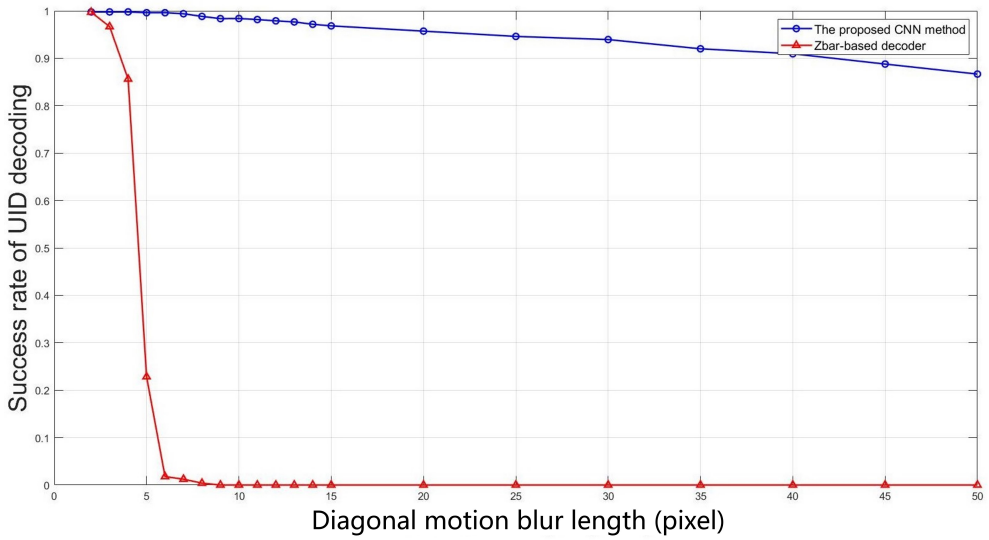

Figure 11. Success rate of UID decoding for diagonal motion-blurred images using the proposed BN-CNN model versus that using the Zbar-based decoder.

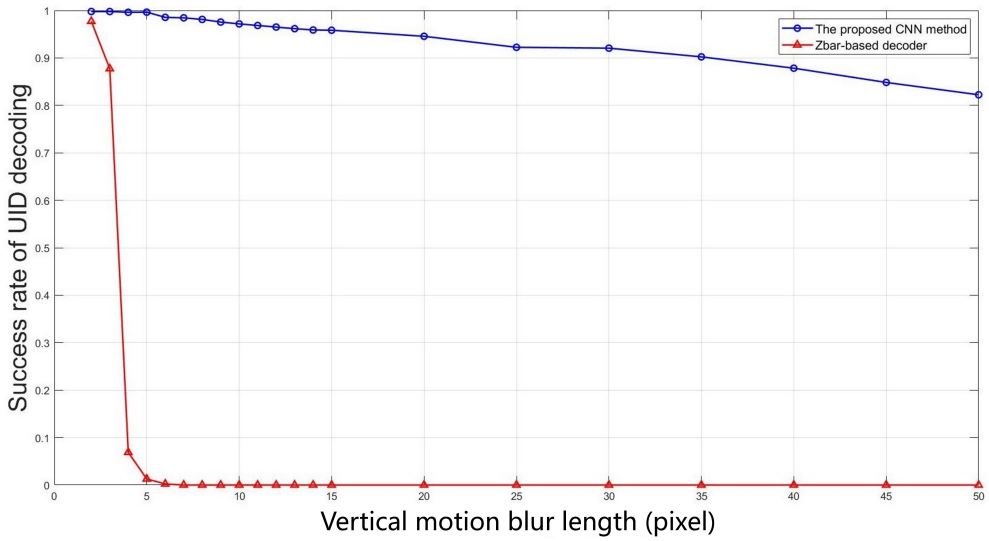

Figure 12. Success rate of UID decoding for vertical motion-blurred images using the proposed BN-CNN model versus that using the Zbar-based decoder.

3.4. Success Rate of UID Decoding for Diffusion-Blurred Images

Diffusion blur is the blur caused by the camera being out of focus in a photograph or a single frame of film or video. The image being recorded changes when the camera is too late to focus because of the high-speed motion of the camera on the vehicle and the focus delay of the camera. In this experiment, Matlab's 2-D filter function was also used to simulate the diffusion blur with different diffusion radius in pixels. As shown in Figure 13, diffusion blur with 23 kinds of diffusion radius was introduced into each image in the original dataset. Therefore, we have a total of 23 $\times$ 1800 = 41,400 diffusion-blurred images to test the robustness of the proposed DL-based VLP system. Note that the diffusion blur radius varies from 2 pixels to 20 pixels with a 1-pixel increment and 20 pixels to 40 pixels with a 5-pixel increment.

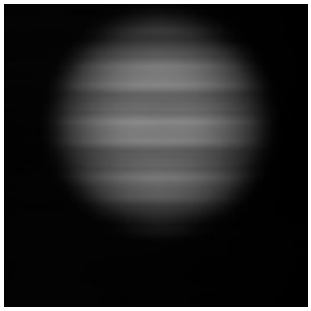

Figure 13. The VLP-LED stripe images with diffusion blur with a diffusion blur radius of 40 in pixels.

As shown in Figure 14, for diffusion-blurred images, when the diffusion radius reaches 4 pixels, the success rate of UID decoding using the traditional Zbar-based decoder falls to 0, while that of using the proposed BN-CNN model remains 98.5%. Furthermore, the success rate of UID decoding using the proposed BN-CNN model remains higher than 90.5% until the diffusion radius reaches 16 pixels and it is still higher than 58.5% even when the diffusion radius reaches 40 pixels.

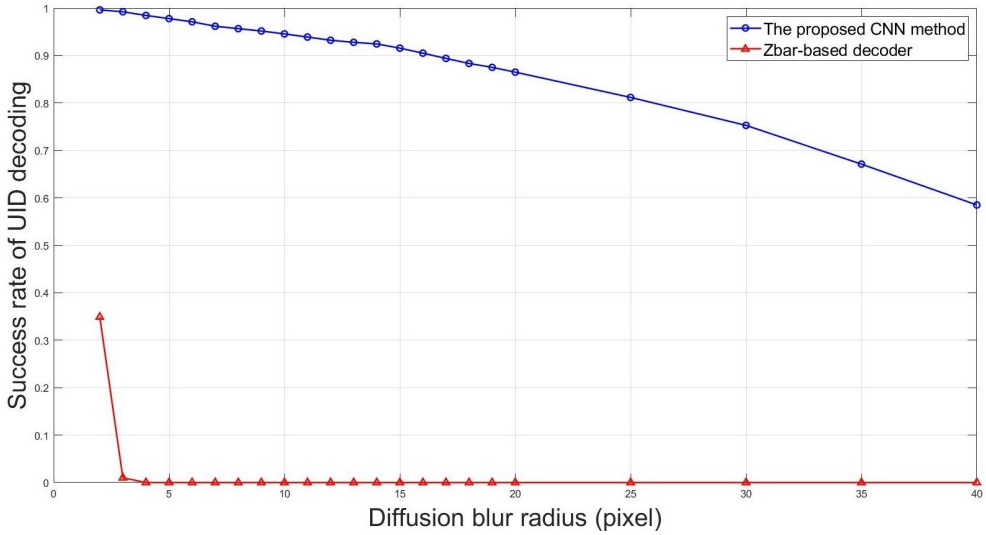

Figure 14. Success rate of UID decoding for diffusion blurred images using the proposed BN-CNN model versus that using the Zbar-based decoder.

To conclude, the proposed BN-CNN model is robust for VLP-LED stripe images with diffusion blur introduced by focus delay of the camera in high-speed movement or smog in tunnels, compared with the traditional Zbar-based decoder. Note that the parameters of the proposed BN-CNN model used in this experiment are the same as before, i.e., the proposed BN-CNN model trained by clear images still has a high success rate for not only images with heavy motion blur, but also images with heavy diffusion blur. It means that the proposed BN-CNN model has high robustness and adaptivity for both motion-blurred images and diffusion-blurred images, and then improves the robustness of IS-based VLP systems for high-speed vehicles.

3.5. Positioning Time

The positioning time of the VLP algorithm includes the computational time of ROI detection and the computational time of UID decoding. In our experiments, we used Windows 10 (Microsoft, Redmond, WA, USA) with Intel(R) Core(TM) i7-8550U CPU @ 1.80 GHz and NVIDIA GeForce MX150 GPU (Lenovo, Beijing, China). As shown in Table 3, the overall positioning time of the VLP algorithm based on Wiener image restoration [29] and Zbar [25] is 99.54 ms, while the overall positioning time of the proposed DL-based VLP

algorithm is reduced to 9.19 ms. The computational time of extracting a VLP-LED stripe image region from a captured image is about 1 ms, while the computational time for each UID decoding using the proposed BN-CNN model is about 8.18 ms.

Table 3. The time of the proposed DL-based robust VLP system.

VLP Algorithm	Used Time of ROI Detection (ms)	Used Time of UID Decoding (ms)	Overall Positioning Time (ms)
DL-based VLP	1.01	8.18	9.19
Zbar-based VLP	1.01	98.53	99.54

4. Conclusions

This work proposes a robust VLP system for high-speed vehicles based on a deep learning and data-driven approach. The proposed DL-based decoder with a BN-CNN model can increase the success rate of identifying (or decoding) VLP-LED UIDs from the captured images with motion blur and diffusion blur caused by high-speed movement. Furthermore, a lightweight fast ROI detection algorithm is also proposed to reduce the computational latency for detecting and extracting VLP-LED stripe image regions from the captured images. Experimental results show that the success rate of UID decoding using the proposed BN-CNN model could be higher than 98% when that of the traditional Zbar-based decoder falls to 0, while the computational time for positioning is decreased to 9.19 ms and the speed of vehicles is 6.73 km/h. Therefore, the proposed VLP system can significantly enhance the robustness against high-speed movement and guarantee the real-time response for positioning and navigation for high-speed vehicles.

Author Contributions: Conceptualization, D.L. and Z.W.; methodology, D.L.; software, Z.W. and G.Y.; validation, Z.W.; formal analysis, D.L. and Z.W.; investigation, G.Y., Y.Y., J.L. (Jingwen Li), M.Y. and P.L.; resources, M.L., Z.C., Z.L.J. and J.F.; data curation, J.L. (Jiajun Lin) and S.C.; writing—original draft preparation, D.L. and Z.W.; writing—review and editing, D.L., Z.W., G.Y., Y.Y., J.L. (Jingwen Li), M.Y., P.L., J.L. (Jiajun Lin), S.C., M.L., Z.C., Z.L.J. and J.F.; visualization, G.Y. and Y.Y.; supervision, M.L., Z.C., J.F. and Z.L.J.; project administration, J.F.; funding acquisition, J.F. All authors contributed to the critical reading and writing of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the National Key Research and Development Program of China (No. 2018YFB1801900, No. 2019YFE0123600), National Natural Science Foundation of China (No. 62171202, No. 62075088), Guangdong Provincial Postgraduate Education Innovation Project (No. 2019SFKC08), Project of Guangzhou Industry Leading Talents (CXLJTD-201607), European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 872172 (TESTBED2 project: www.testbed2.org (accessed on 1 September 2022)).

Data Availability Statement: The data that supports the findings of this study are available within the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. He, J.; Zhou, B. Vehicle positioning scheme based on visible light communication using a CMOS camera. *Opt. Express* **2021**, *29*, 27278–27290. [CrossRef] [PubMed]
2. Guan, W.; Chen, S.; Wen, S.; Tan, Z.; Song, H.; Hou, W. High-accuracy robot indoor localization scheme based on robot operating system using visible light positioning. *IEEE Photonics J.* **2020**, *12*, 7901716. [CrossRef]
3. Hussain, B.; Wang, Y.; Chen, R.; Cheng, H.; Yue, C. LiDR: Visible Light Communication-Assisted Dead Reckoning for Accurate Indoor Localization. *IEEE Internet Things J.* **2022**, *9*, 15742–15755. [CrossRef]
4. Armstrong, J.; Sekercioglu, Y.A.; Neild, A. Visible light positioning: A roadmap for international standardization. *IEEE Commun. Mag.* **2013**, *51*, 68–73. [CrossRef]
5. Yasir, M.; Ho, S.W.; Vellambi, B.N. Indoor position tracking using multiple optical receivers. *J. Lightwave Technol.* **2015**, *34*, 1166–1176. [CrossRef]

6. Steendam, H. A 3-D positioning algorithm for AOA-based VLP with an aperture-based receiver. *IEEE J. Sel. Areas Commun.* **2017**, *36*, 23–33. [CrossRef]
7. Wu, Y.C.; Chow, C.W.; Liu, Y.; Lin, Y.S.; Hong, C.Y.; Lin, D.C.; Song, S.H.; Yeh, C.H. Received-signal-strength (RSS) based 3D visible-light-positioning (VLP) system using kernel ridge regression machine learning algorithm with sigmoid function data preprocessing method. *IEEE Access* **2020**, *8*, 214269–214281. [CrossRef]
8. Meng, X.; Jia, C.; Cai, C.; He, F.; Wang, Q. Indoor High-Precision 3D Positioning System Based on Visible-Light Communication Using Improved Whale Optimization Algorithm. *Photonics* **2022**, *9*, 93. [CrossRef]
9. Martínez-Ciro, R.A.; López-Giraldo, F.E.; Luna-Rivera, J.M.; Ramírez-Aguilera, A.M. An Indoor Visible Light Positioning System for Multi-Cell Networks. *Photonics* **2022**, *9*, 146. [CrossRef]
10. You, X.; Yang, X.; Jiang, Z.; Zhao, S. A Two-LED Based Indoor Three-Dimensional Visible Light Positioning and Orienteering Scheme for a Tilted Receiver. *Photonics* **2022**, *9*, 159. [CrossRef]
11. Zhu, B.; Zhu, Z.; Wang, Y.; Cheng, J. Optimal optical omnidirectional angle-of-arrival estimator with complementary photodiodes. *J. Lightwave Technol.* **2019**, *37*, 2932–2945. [CrossRef]
12. Do, T.H.; Yoo, M. An in-depth survey of visible light communication based positioning systems. *Sensors* **2019**, *16*, 678. [CrossRef] [PubMed]
13. Hsu, C.W.; Liu, S.; Lu, F.; Chow, C.W.; Yeh, C.H.; Chang, G.K. Accurate indoor visible light positioning system utilizing machine learning technique with height tolerance. In Proceedings of the 2018 Optical Fiber Communications Conference and Exposition (OFC), San Diego, CA, USA, 11–15 March 2018; Volume 3, pp. 1–3.
14. Chuang, Y.C.; Li, Z.Q.; Hsu, C.W.; Liu, Y.; Chow, C.W. Visible light communication and positioning using positioning cells and machine learning algorithms. *Opt. Express* **2019**, *27*, 16377–16383. [CrossRef] [PubMed]
15. Lin, P.; Hu, X.; Ruan, Y.; Li, H.; Fang, J.; Zhong, Y.; Zheng, H.; Fang, J.; Jiang, Z.L.; Chen, Z. Real-time visible light positioning supporting fast moving speed. *Opt. Express* **2020**, *28*, 14503–14510. [CrossRef]
16. Li, H.; Huang, H.; Xu, Y.; Wei, Z.; Yuan, S.; Lin, P.; Wu, H.; Lei, W.; Fang, J.; Chen, Z. A fast and high-accuracy real-time visible light positioning system based on single LED lamp with a beacon. *IEEE Photonics J.* **2020**, *12*, 7906512. [CrossRef]
17. Guan, W.; Huang, L.; Wen, S.; Yan, Z.; Liang, W.; Yang, C.; Liu, Z. Robot Localization and Navigation Using Visible Light Positioning and SLAM Fusion. *J. Lightwave Technol.* **2021**, *39*, 7040–7051. [CrossRef]
18. Lin, D.C.; Chow, C.W.; Peng, C.W.; Hung, T.Y.; Chang, Y.H.; Song, S.H.; Lin, Y.S.; Lin, Y.; Lin, K.H. Positioning unit cell model duplication with residual concatenation neural network (RCNN) and transfer learning for visible light positioning (VLP). *J. Lightwave Technol.* **2021**, *39*, 6366–6372. [CrossRef]
19. Xie, C.; Guan, W.; Wu, Y.; Fang, L.; Cai, Y. The LED-ID detection and recognition method based on visible light positioning using proximity method. *IEEE Photonics J.* **2018**, *10*, 7902116. [CrossRef]
20. Guan, W.; Chen, X.; Huang, M.; Liu, Z.; Wu, Y.; Chen, Y. High-speed robust dynamic positioning and tracking method based on visual visible light communication using optical flow detection and Bayesian forecast. *IEEE Photonics J.* **2018**, *10*, 7904722. [CrossRef]
21. Su, S.; Heidrich, W. Rolling shutter motion deblurring. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1529–1537.
22. Schöberl, M.; Fößel, S.; Bloss, H.; Kaup, A. Modeling of image shutters and motion blur in analog and digital camera systems. In Proceedings of the 2009 16th IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, 7–10 November 2009.
23. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In Proceedings of the International Conference on Machine Learning (ICML), Lile, France, 6–11 July 2015; pp. 448–456.
24. Santurkar, S.; Tsipras, D.; Ilyas, A.; Mądry, A. How does batch normalization help optimization? In Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS'18), Montréal, QC, Canada, 3–8 December 2018; pp. 2488–2498.
25. ZBar Bar Code Reader. Available online: http://zbar.sourceforge.net/ (accessed on 29 June 2022).
26. Lin, H.Y.; Li, K.J.; Chang, C.H. Vehicle speed detection from a single motion blurred image. *Image Vis. Comput.* **2008**, *26*, 21327–21337. [CrossRef]
27. Xu, T.-F.; Zho, P. Object's translational speed measurement using motion blur information. *Measurement* **2010**, *43*, 1173–1179.
28. Mohammadi, J.; Akbari, R. Vehicle speed estimation based on the image motion blur using radon transform. In Proceedings of the 2010 2nd International Conference on Signal Processing Systems (ICSPS), Dalian, China, 5–7 July 2010; V1-243.
29. Orieux, F., Giovannelli, J. F., Rodet, T. Bayesian estimation of regularization and point spread function parameters for Wiener–Hunt deconvolution. *J. Opt. Soc. Am. A* **2010** *27*, 1593–1607. [CrossRef] [PubMed]