Proceeding Paper

Cyber–Physical System for Traffic Sign Detection and Recognition †

Nikolay Gospodinov and Georgi Krastev *

Department of Computer Systems and Technologies, Angel Kanchev University of Ruse, 7004 Ruse, Bulgaria; ngospodinov@uni-ruse.bg

* Correspondence: geork@uni-ruse.bg

† Presented at the 4th International Conference on Communications, Information, Electronic and Energy Systems (CIEES 2023), Plovdiv, Bulgaria, 23–25 November 2023.

Abstract: The objective of this paper is to introduce a module that works with smart glasses that to detect and recognize traffic signs on the road. The developed module encompasses the detection and classification of traffic signs, thereby enhancing safety and convenience for users during their travels. Employing innovative recognition algorithms, the software distinguishes between various types of signs, including speed limits, hazard warnings, and directional indicators. Furthermore, the system prioritizes user convenience by providing an intuitive and easily com-prehensible interface. This facilitates quick and precise access to information about traffic signs for both drivers and pedestrians, irrespective of their technological experience. Our endeavor not only illustrates the functionality of the system but also underscores its significance in augmenting road user safety. This innovative approach reflects our commitment to advancing technologies that not only offer intelligent solutions but also simplify everyday tasks for individuals. The paper establishes the sustainability and efficacy of this prototype, thereby laying the groundwork for future development and potential large-scale implementation of the system. This presents an exciting opportunity to enhance the mobility and safety for road users through innovative technologies and intelligent solutions.

Keywords: intelligent system; traffic sign recognition; artificial intelligence; smart glasses

1. Introduction

In recent years, with the development of computer technologies in the automotive industry, road sign recognition systems have been a big part of it, and in particular in autonomous navigation [1]. The reason for this is the effort of many researchers who have researched and analyzed different road sign recognition methods [2]. Part of this research is based on artificial intelligence and deep learning [3]. Most road sign recognition algorithms are based on computer vision and the use of neural networks that use images to train and test algorithms [4]. In a study by Kaoutar Sefrioui Boujemaa (2017) [5], two approaches are used: (1) the C-CNN approach, based on color segmentation techniques and convolutional neural networks; (2) the fast R-CNN approach, based on region-based convolutional neural networks.

Some researchers use a three-stage real-time traffic sign recognition system with the help of the SVM method (support vector machine) for color enhancement and the HOG method (histogram of oriented gradients) to extract red regions in the image [6]. The content of the traffic signs found is identified by analyzing the color information and classified according to the two classifiers.

Another study, by Claus Bahlmann (2005) [7], used color, shape, and motion information for creating a system for traffic sign recognition. The proposed system uses a detection and tracking framework based on AdaBoost, color sensitive Haar wavelet features, temporal information propagation, and a Bayesian classification with temporal hypothesis fusion.

The study builds on the standard computer vision-based traffic sign detection systems with LSTM blocks for recognizing objects resembling road signs and by presenting a prototype working on devices with minimal hardware requirements, with the possibility of connection between the driving student and the driving teacher.

Nowadays, people are increasingly frequently victims of traffic accidents. In Bulgaria, there were 6609 traffic accidents on the nation's roads between 1 January and 31 December 2022, resulting in 531 fatalities and 8422 injuries to drivers. In total, 6656 of the 8422 accident victims who sustained injuries were only mildly hurt, while 1766 were critically hurt. Table 1 shows a comparative description of the types of road casualties for 2021 and 2022 [8].

Table 1. Road victims who have died or been injured in Bulgaria.

Types of Participants in Traffic Accidents	Dead	Wounded	Year
Drive	291	3975	2022
	295	3599	2021
	−4	+376	Diffs. 2022–2021
Passenger	146	2855	2022
	170	2551	2021
	−24	+304	Diffs. 2022–2021
Pedestrian	94	1579	2022
	94	1449	2021
	0	+130	Diffs. 2022–2021
Road worker	0	13	2022
	2	10	2021
	−2	+3	Diffs. 2022–2021
Total	531	8422	2022
	561	7609	2021
	−30	+813	Diffs. 2022–2021

These dark statistics show the dire situation Bulgaria is in when it comes to dealing with this type of incident. There are different causes of road accidents, and they are divided into the following types:

- Traffic accidents, deaths, and injuries due to the fault of the drivers of road vehicles according to offenses committed (first offense only).
- Traffic accidents, deaths, and injuries due to improper actions by the passengers.
- Traffic accidents, deaths, and injuries due to improper actions by pedestrians.

Many traffic accidents are the result of not knowing the road signs and the basic rules of the road.

Traffic sign detection and correct recognition is a potential remedy for the issue of image classification with AR smart spectacles. Finding the class to which an input image of a road sign belongs is the aim. The process of doing this involves building and training an artificial neural network on images. The neural network's task is to identify the class to which the image of a road sign belongs, considering the benefits of a particular strategy that leverages deep learning techniques in cyber–physical systems.

The development includes the detection and recognition of road signs. Information about a recognized sign is transmitted to the user by means of audible display of the name of the road sign.

The third generation of augmented reality glasses, the Moverio BT-300 (Figure 1), are the foundation of the current research. Their benefits include audio support, video support, and excellent image contrast.

The developed model is part of a larger cyber–physical system, which includes other diverse modules: a module for object detection, a module for hazardous object detection, a module for fire detection, and more.

Figure 1. Moverio BT-300 augmented reality glasses.

2. Methods

A prototype of a cyber–physical system for the detection and recognition of traffic signs was developed using the BT-300 augmented smart glasses. The glasses work on an Android like operating system called Moverio OS. They include a camera, a 9-axis motion sensor, and they support MP3 audio formats and have also other functions (Table 2).

Table 2. BT-300 Smart glasses—technical characteristics.

Name		Characteristics
Model Number		BT-300
Material		OLED
Supported movie formats		MP4 (MPEG4/H.264+AAC), MPEG2 (H.264+AAC), VP8
Supported still image formats		JPEG, PNG, BMP, GIF
Supported audio formats		WAV, MP3, AAC
Wi-Fi standards		802.11a, 802.11b, 802.11g, Wi-Fi 4 (802.11n), Wi-Fi 5
Internal memory	Main memory	2GB
	User memory	16GB Processor manufacturer
Processor manufacturer		Intel
Driving method		Monocrystalline silicon active matrix
Maximum refresh rate		30 Hz
Filetype		EN-67 Manual (PDF), EN-2 Manual (PDF), EN-2 (PDF)

With the Moverio BT-300 glasses, you may display 3D material on both sides of the screen while simultaneously projecting a picture for both eyes. As an illustration, if the recipient is a person learning the rules of the road, they can simultaneously receive and transmit real-time information about the different types of traffic signs by making use of the integrated mic and camcorder [9].

3. Implementation

The traffic sign detection and identification module serve as the foundation of the software prototype.

A schematic of the software implementation may be seen in Figure 2. There are six steps to the implementation process:

- Step 1: Creating/modifying a dataset of traffic sign image samples for training and testing the traffic sign detection and recognition model based on neural networks;
- Step 2: Developing a model for detection and recognition of traffic signs based on YOLOv8 and long short-term memory (LSTM) networks;
- Step 3: Training the developed model with the provided dataset;
- Step 4: Implementing the trained model on Android-based software with TensorFlow;
- Step 5: Testing the module for traffic sign detection and recognition;
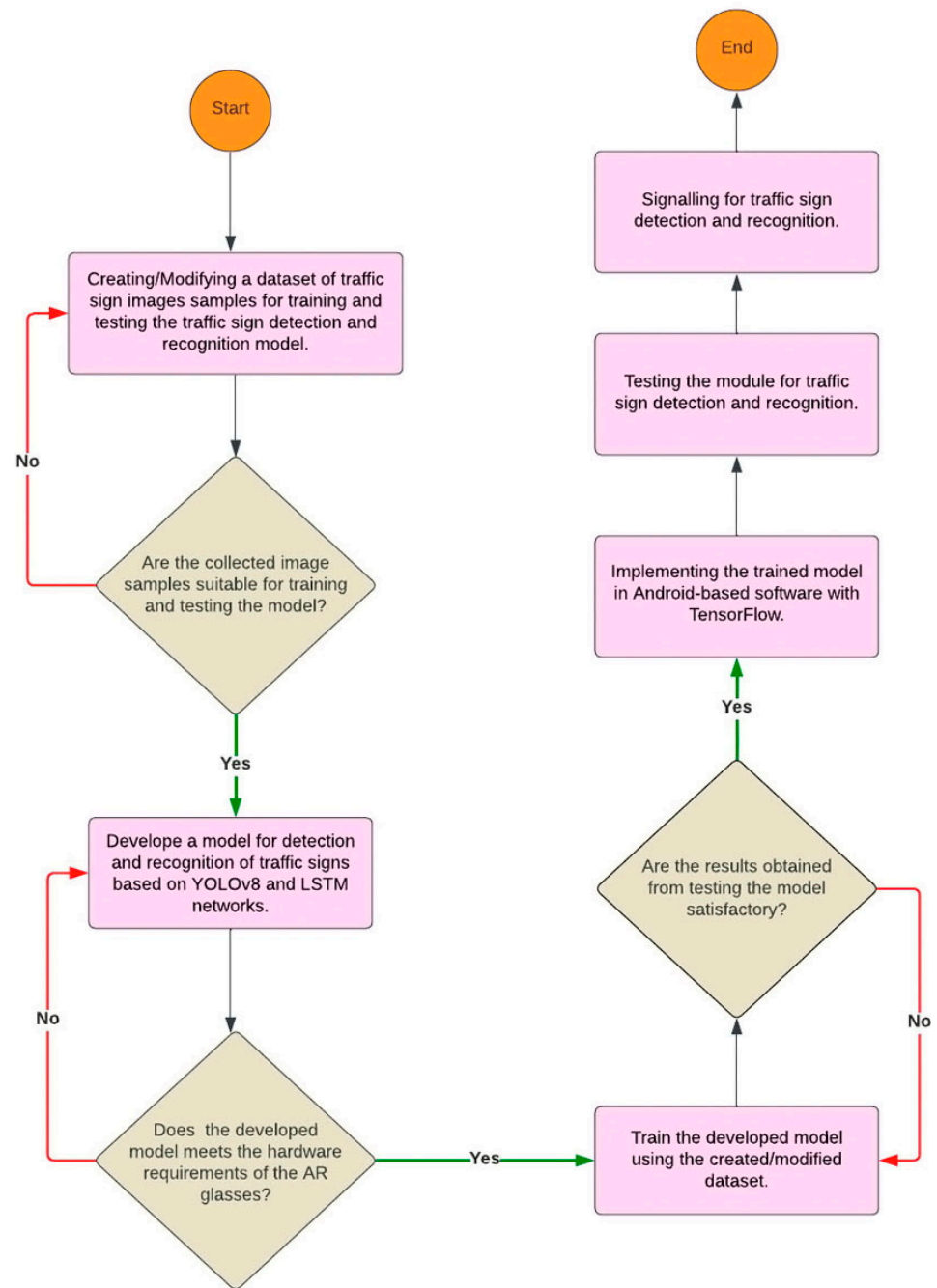- Step 6: Signaling for traffic sign detection and recognition.

Figure 2. Software module implementation architecture.

3.1. Creating/Modifying a Dataset of Traffic Sign Image Samples for Training and Testing the Traffic Sign Detection and Recognition Model Based on Neural Networks

The first stage is essential for the established model to operate well. The dataset of image samples must be exact and adhere to any standards that may be applied. The training and testing phases of the programmed model depend greatly on image selection. The likelihood of the neural network being faked at this point is fairly significant, necessitating careful monitoring of the image samples utilized. The following requirements were set and complied with in the compilation of the dataset:

- The greatest possible number of image samples for training and testing. For the purpose of this research, over 100,000 images were used;
- The image samples are of a high resolution. The image sample resolution was between 1280×720 and 1920×1080;

- The image samples are from different seasons (spring, autumn, winter);
- The image samples are from different times of day (morning, afternoon, evening);
- The image samples are from different weather conditions (rain, snow, bright sun).

3.2. Developing a Model for Detection and Recognition of Traffic Signs Based on YOLOv8 and Long Short-Term Memory Networks

At the second stage, a neural network based on artificial intelligence was created. The neural network has to fulfill the following requirements:

- The neural network model ought to function on hardware-light devices (such as mobile phones and smart glasses);
- The model has to be executable in real-time;
- Road signs of different shapes must be recognized by the neural network;
- The model must be able to separate other items from traffic signs that have the same color and form as the traffic signs, such as billboards and safety equipment.

In Figure 3, the principal scheme of the traffic sign detection and recognition model is shown.
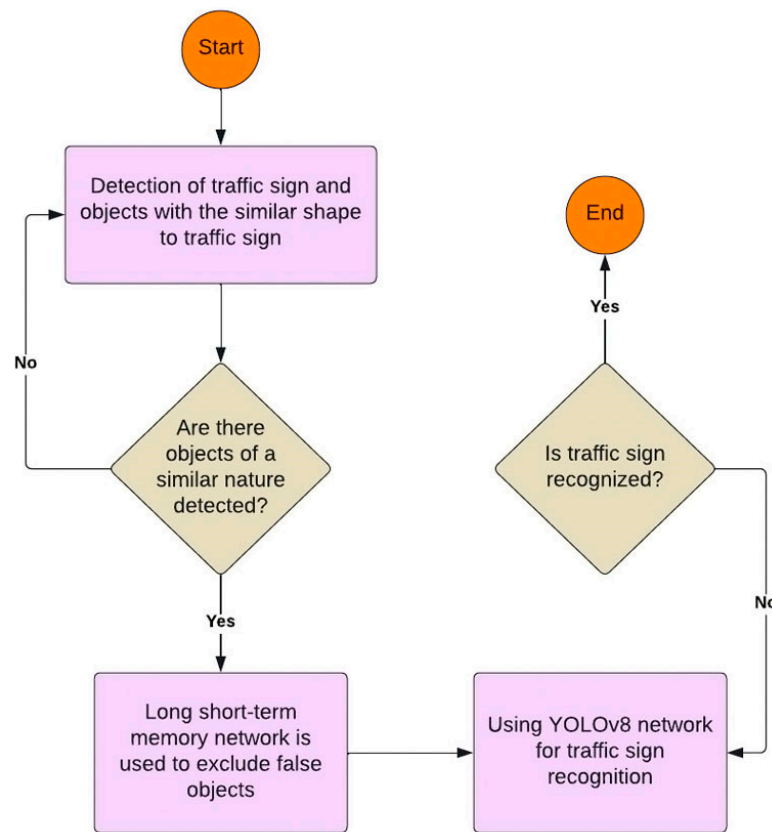

Figure 3. Scheme of the traffic sign detection and recognition model.

3.2.1. The Long Short-Term Memory (LSTM) Network Is Used to Exclude False Objects from Real Traffic Signs

A recurrent neural network (RNN) architecture specifically created to solve the issue of capturing long-term dependencies in sequential data, which is referred to as long short-term memory [10]. The moniker underscores the LSTM's capacity to choose to discard unimportant information in the short term while simultaneously retaining significant information over time by approaching it as a binary classification issue, where the objective is to categorize a given object as either phony or legitimate [11]. LSTM is used for distinguishing real road signs from objects that resemble them.

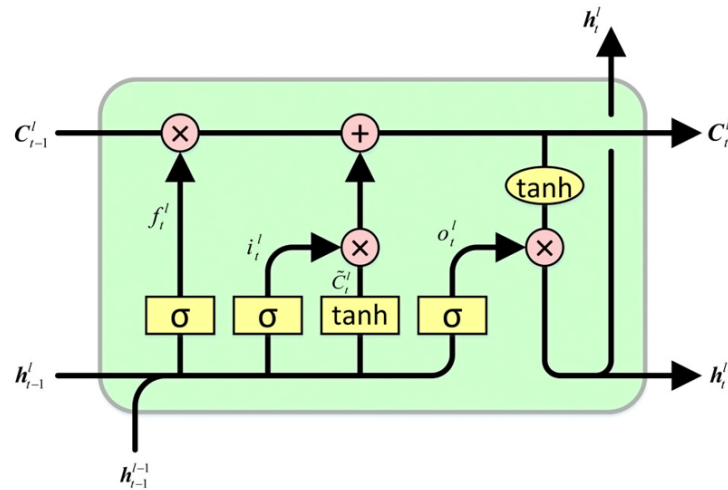The LSTM network is made up of LSTM cells (Figure 4).

Figure 4. LSTM cell.

Each LSTM cell has an LSTM state—C_t. This state is used by the next LSTM cell. The next LSTM cell can read information, write information, or reset information. Every LSTM cell has three gates and a new memory network.

The forget gate controls whether the memory cell is reset to 0. Based on the prior hidden state and the fresh input data, the LSTM neural network will now decide which components of the cell state (long-term memory) are pertinent. In order to do this, the following equation is used:

$$F_t = \sigma\left(W_f.[h_{t-1}, X_t] + b_f\right). \tag{1}$$

The input gate controls whether the memory cell is updated. Based on the prior hidden state and the current input data, the goal of this stage is to determine what new information needs to be added to the network's long-term memory (cell state). In order to do this, the following equation is used:

$$i_t = \sigma(W_i.[h_{t-1}, X_t] + b_i). \tag{2}$$

The new memory network is a neural network that employs the tanh activation function. It is trained to produce a "new memory update vector" by fusing the prior hidden state with the present input data [12]. This vector contains data from the input and accounts for the context that the preceding concealed state gave. The new memory update vector defines how much the long-term memory's individual components (cell state) should be modified in light of the most recent information. All this is represented by the following equation:

$$\hat{C}_t = tanh(W_c.[h_{t-1}, X_c] + b_c). \tag{3}$$

The LSTM network's cell state, which serves as its long-term memory, is updated using the outcome of the interaction between the input gate filter and the new memory update. Only the pertinent parts of the new memory update are added to the cell state because the input gate filter controls the output of the new memory update through pointwise multiplication:

$$C_t = i_t.\hat{C}_t + f_t.C_{t-1}. \tag{4}$$

The output gate controls whether the information of the current cell is visible. The new hidden state is established in the last step of an LSTM utilizing the freshly updated cell state, the prior hidden state, and the fresh input data. It is the output gate that makes this choice. The LSTM network's ultimate hidden state is ascertained using this gate. The prior concealed state, the new input data, and the updated cell state are all inputs used in this stage. The output gate serves as a filter since simply releasing the modified cell state would

reveal too much information. The output gate, which determines which components of the updated cell state are significant and ought to be output as the new hidden state, is a sigmoid-activated network that serves as a filter:

$$O_t = \sigma(W_o.[h_{t-1}, X_t] + b_o). \tag{5}$$

3.2.2. YOLOv8 Is Used for Traffic Sign Detection and Recognition

YOLOv8 stands for You Only Look Once version 8. The architecture was released in January 2023.

In YOLOv8, its backbone is the same as that of YOLOv5, and the C3 (three convolutional layers) module is replaced by a C2f (two faster convolutional layers) module based on the cross stage partial (CSP) idea [7]. All YOLO versions are supported by YOLOv8, which may also switch between them at will. Additionally, it may function on a variety of hardware platforms (CPU-GPU), showing its high versatility. The YOLOv8 network architecture diagram is shown in Figure 5.
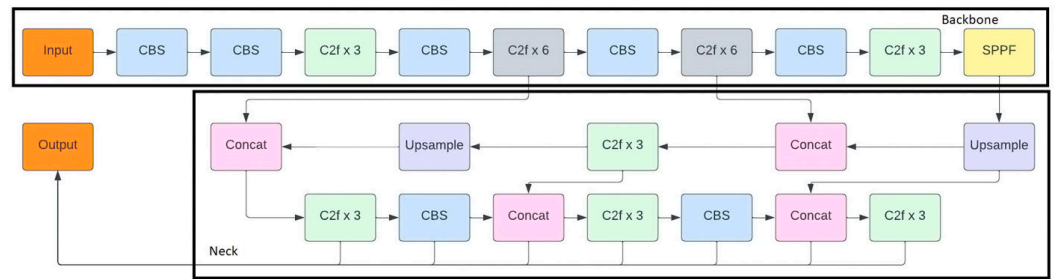
Figure 5. YOLOv8 architecture diagram.

The main difference between YOLOv8 and YOLOv5 is:

- Anchor-free detection;
- Mosaic augmentation.

3.3. *Training the Developed Model with the Provided Dataset*

The training procedure begins after the model is constructed and finishes when the testing is successful. The method is crucial, and the outcomes demonstrate how well the model functions.

The model that will be put to the test is trained using a predetermined collection of image samples. It includes image samples of traffic signs, image samples of things that seem like traffic signs, and image samples without traffic signs. Once the testing results are successful, the model is prepared for use in the Android-based smart glasses application. In Figure 6, a result from the training and testing of the model is shown.

Figure 6. Result from the training and testing of the model.

3.4. Implementing the Trained Model on Android-Based Software with TensorFlow

The development of a prototype program for smart glasses with augmented reality requires the use of several software tools and platforms.

The development and training of a model for the detection and recognition of traffic signs using software. The two sorts of software platforms used are those for putting the trained model into use and those for designing the user interface of the prototype software. PyCharm (Version 2023.1) and Jupyter Notebook (Version 6.4.8) are the first software platforms used, followed by Android Studio. In order for the model to work on an Android-based application, it must be converted to a suitable format. The used format is the TensorFlow format.

3.5. Testing the Module for Traffic Sign Detection and Recognition

This step is performed entirely on the device. At this stage, it was checked that the software was working efficiently and that it was performing its main task, namely the recognition of road signs. For a successfully completed task, recognition of all road signs within the frame and their correct labeling is considered necessary. Figure 7 shows the result of the software running in real time.

Figure 7. Result from testing the model on the Android application.

3.6. Signaling for Traffic Sign Detection and Recognition

In this step, the names of the labeled traffic signs are transmitted in audio format to the user sequentially and one at a time. This is carried out so that even users with visual problems can use the application.

4. Conclusions

In conclusion, the development and implementation of a cyber–physical system for detecting and recognizing traffic signs represent a step forward in road safety and literacy. This system utilizes technologies like computer vision, deep learning, and real time data processing to improve the accuracy and efficiency of identifying traffic signs on our roads.

The potential advantages of this system are significant. It can greatly decrease the risk of accidents by providing precise information to both drivers and pedestrians enabling them to make decisions and adapt to changing road conditions.

However, there are also challenges that need attention, including concerns regarding privacy, and cybersecurity as a necessity for reliable hardware and software use. Moreover, ensuring that this system remains accessible and affordable for a range of users in various regions is crucial for its widespread acceptance and success.

As technology continues to advance and as society moves forward, the role played by cyber–physical systems in detecting and recognizing traffic signs will only grow more critical.

To ensure the safety and efficiency of our roads it is crucial that researchers and engineers collaborate. By refining and implementing these systems collectively, we can pave the way for a future in transportation that maximizes road safety and optimizes traffic management.

References

1. Diakaki, C.; Papageorgiou, M.; Papamichail, I.; Nikolos, I. Overview and analysis of Vehicle Automation and Communication Systems from a motorway traffic management perspective. Transportation. *Res. Part A Policy Pract.* **2015**, *75*, 147–165. [CrossRef]
2. Saadna, Y.; Behloul, A. An overview of traffic sign detection and classification methods. *Int. J. Multimed. Inf. Retr.* **2017**, *6*, 193–210. [CrossRef]
3. Aghdam, H.H.; Heravi, E.J.; Puig, D. Recognizing traffic signs using a practical deep neural network. In *Robot 2015: Second Iberian Robotics Conference*; Springer International Publishing: Cham, Switzerland, 2015; pp. 399–410. [CrossRef]
4. Aghdam, H.H.; Heravi, E.J.; Puig, D. A practical approach for detection and classification of traffic signs using Convolutional Neural Networks. *Robot. Auton. Syst.* **2016**, *84*, 97–112. [CrossRef]
5. Boujemaa, K.S.; Berrada, I.; Bouhoute, A.; Boubouh, K. Traffic sign recognition using convolutional neural networks. In Proceedings of the 2017 International Conference on Wireless Networks and Mobile Communications (WINCOM), Rabat, Morocco, 1–4 November 2017; pp. 1–6. [CrossRef]
6. Fatin Zaklouta, F.Z.; Bogdan Stanciulescu, B.S. Real-time traffic sign recognition in three stages. *Robot. Auton. Syst.* **2014**, *62*, 16–24. [CrossRef]
7. Bahlmann, C.; Zhu, Y.; Ramesh, V.; Pellkofer, M.; Koehler, T. A system for traffic sign detection, tracking, and recognition using color, shape, and motion information. In Proceedings of the IEEE Proceedings. Intelligent Vehicles Symposium, Las Vegas, NV, USA, 6–8 June 2005; pp. 255–260. [CrossRef]
8. Ministry of Interior Affairs. *Road Transport Injuries in 2022*; General Directorate "National Police": Sofia, Bulgaria, 2022.
9. Epson Moverio. Available online: https://moverio.epson.com (accessed on 20 November 2023).
10. Van Houdt, G.; Mosquera, C.; Nápoles, G. A Review on the Long Short-Term Memory Model. *Artif. Intell. Rev.* **2020**, *53*, 5929–5955. [CrossRef]
11. Loue, H.; Duan, X.; Guo, J. Dc-yolov8: Small-size object detection algorithm based on camerasensor. *Electronics* **2023**, *12*, 2323. [CrossRef]
12. Sepp, H.; Jürgen, S. Long Short-Term Memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef]