applied sciences

MDPI

Article

Research on Road Pattern Recognition of a Vision-Guided Robot Based on Improved-YOLOv8

Xiangyu Zhang [1] and Yang Yang [2,*]

[1] Swinburne College, Shandong University of Science and Technology, Jinan 250031, China; skdxiangyuzhang@126.com
[2] College of Mechanical and Electrical Engineering, Shandong University of Science and Technology, Qingdao 266590, China
* Correspondence: yang.yang@sdust.edu.cn

Abstract: In order to promote the accurate recognition and application of visual navigation robots to the environment, this paper carried out research on the road pattern recognition of a vision-guided robot based on improved YOLOv8 on the basis of road pattern calibration and experimental sampling. First, an experimental system for road image shooting was built independently, and 21 different kinds of road pattern image data were obtained by sampling roads with different weather conditions, road materials, and degrees of damage. Second, the road pattern recognition model based on the classical neural network Resnet 18 was constructed for model training and testing, and the initial recognition of road pattern was realized. Third, the YOLOv8 target detection model was introduced to build the road pattern recognition model based on YOLOv8n, and the model was trained and tested, improving road pattern recognition accuracy and recognition response speed by 3.1% and 200%, respectively. Finally, to further improve the accuracy of road pattern recognition, improvement research was carried out on the YOLOv8n road pattern recognition model based on the C2f-ODConv module, the AWD adaptive weight downsampling module, the EMA attention mechanism, and the collaboration of the three modules. Three network architectures, classical CNN (Resnet 18), YOLOv8n, and improved YOLOv8n, were compared. The results show that four different optimization models can further improve the accuracy of road pattern recognition, among which the accuracy of the improved YOLO v8 road pattern recognition model based on multimodule cooperation is the highest, reaching more than 93%.

Keywords: road pattern; improved YOLOv8; recognition accuracy; road image

check for updates

Citation: Zhang, X.; Yang, Y. Research on Road Pattern Recognition of a Vision-Guided Robot Based on Improved-YOLOv8. *Appl. Sci.* **2024**, *14*, 4424. https://doi.org/10.3390/app14114424

Academic Editor: Nicola Pio Belfiore

Received: 4 April 2024
Revised: 19 May 2024
Accepted: 20 May 2024
Published: 23 May 2024

Copyright: © 2024 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/).

1. Introduction

Road type and pavement information are prior information for various tasks, such as vehicle control and path planning [1]. The wheeled robot's recognition results for the road state directly affect its decision-making results. Therefore, the autonomous recognition capability of the wheeled robot for the road state determines the running performance and control capability of the wheeled robot. Accurate recognition of the road state is the premise and necessary condition for operational decision making and real-time path planning adjustment. For a vision-guided wheeled robot, it is particularly important to use visual information for road previews and accurate recognition of the road state.

Yamaguchi et al. [2] increased background classification and used convolutional neural networks and a self-organizing map to detect road cracks, which improved the accuracy of crack detection and reduced false detection. Ouma et al. [3] proposed a triple transform method based on wavelet morphology that used an RGB camera for the detection of initial linear cracks in asphalt pavements, providing a reliable method for the formal identification of early linear structural damage to asphalt pavements. Tedeschi et al. [4] developed an OpenC-based automatic pavement damage recognition system

that achieved the recognition of three common pavement distresses: potholes, longitudinal or transversal cracks, and fatigue cracks. Wu et al. [5,6] proposed a lidar-based method of geometric parameter identification and model reconstruction for uneven roads and abnormal pavements. The semantic segmentation method of uneven features of non-paving roads was studied to achieve the classification of uneven features with different sizes and elevation differences. Maeda et al. [7] used a smartphone to take a large number of road damage images, and used a convolutional neural network to construct a damage detection model to achieve accurate identification of eight types of damaged pavements. Zhao Jian et al. [8] measured the driving data of an experimental vehicle under four kinds of road conditions: compacted soil road, sandy land, good asphalt road, and icy or snowy road. A high-dimensional random forest surface recognition model was established, and the SHAP interpretation method was used to extract the correlation between each feature and road type. Based on this, a reduced-dimensional random forest pavement classification was designed to accurately identify the real vehicle driving pavement type. Yang et al. [9] proposed a feature pyramid, hierarchical boosting network, and average intersection over union to achieve pavement crack detection. Feng et al. [10] proposed a small-sample bridge pavement crack segmentation model based on a multi-scale feature fusion network, which improved the accuracy and performance of bridge pavement crack segmentation. Wang et al. [11] proposed a vehicle road terrain classification method based on acceleration, enabling the road contour to be calculated by vehicle speed and a one-quarter-dynamic model. Liu et al. [12] proposed a multidimensional feature fusion and recognition technology for off-road pavements that integrates acceleration features and image-depth features to detect passable areas of off-road pavements. Shi et al. [13] used semi-supervised learning to carry out research on robot terrain classification based on vibration, which significantly improved classification accuracy. Chen et al. [14] used two-dimensional and three-dimensional images of the road surface as network input, and used deep learning based on a multi-branch framework to realize the segmentation of road strip repair, potholes, looseness, bridge joints, and the recognition of cracks and block repair. Dimastrogiovanni et al. [15] studied a planetary terrain detection method based on a support vector machine by estimating the motion state of the probe vehicle and the physical variables related to the interaction between the vehicle and the environment. By improving the AlexNet model, Wang et al. [16] constructed a common road-type image perception recognition model with higher network training speed and road image recognition accuracy. Qin et al. [17] developed a road classification output model through signal time–frequency processing and PNN with the object of steep mass acceleration, unsprung mass acceleration, and rattle space. Wang et al. [18] constructed a pavement classification model based on structural reparameterization and adaptive attention to realize the rapid and accurate identification of complex pavements, such as asphalt, cement, ice and snow, sand, flower brick, slate, wet, and slippery. Wang et al. [19] combined convolutional neural networks with support vector machines to achieve RGB image recognition and classification of six common terrains: grass, mud, sand, asphalt, gravel, and hydrops. Xu et al. [20] proposed a road surface apparent damage image recognition method based on historical information, introduced a mechanism of using historical information to create initial constraints for damage identification, trained the VGG-16 network to extract damage features, and finally used the improved genetic algorithm to achieve a significant increase in recognition speed under the condition of ensuring recognition accuracy. Chen et al. [21] proposed a thermal–RGB fusion image-based pavement damage detection model to achieve accurate detection of pavement damage. Kou et al. [22] established the BAS-BP road recognition model and constructed an adaptive fuzzy control of electromagnetic hybrid suspension based on road recognition. Dewangan et al. [23] constructed a road classification network model based on a convolutional neural network and realized the recognition of five main types of roads: curvy, dry, ice, rough, and wet. Zhang et al. [24] proposed a forward road adhesion coefficient prediction method based on image recognition. Road segmentation and road type identification are realized, and the forward road adhesion coefficient is ob-

tained using DeeplabV3+, a semantic segmentation network, and a MobilNetV2 lightweight convolutional neural network. Yousefzadeh et al. [25] measured surface roughness using accelerometers coupled with high-speed distance sensors and other methods, combined with an artificial neural network, to achieve road profile estimation. Du et al. [26] proposed a road profile elevation inversion and roughness estimation method based on vehicle vibration response signals. The international roughness index is solved by using the vehicle body vibration response signal as the measured value, and the road roughness is accurately evaluated by combining multi-vehicle collaborative estimation. Bai et al. [27] proposed a deep neural network based on vibration multilayer perception to realize terrain classification and recognition, and completed the terrain classification and recognition of planetary detectors. Yang et al. [28] proposed a method based on semantic segmentation to realize intelligent detection of asphalt pavement cracks in highways and other scenarios. Conducting experiments on two different road types, dry and wet, Šabanovič et al. [29] studied road type, condition recognition, and friction coefficient estimation based on a deep neural network (DNN) and a video image sensor. Li et al. [30] collected laser radar and image data for spatiotemporal matching, and proposed a three-dimensional shape and size extraction method for vehicles based on road space division, road segmentation, and laser point cloud gathering, as well as a vehicle labeling method covering steps such as target filtering and classification, identification difficulty division, three-dimensional bounding box calibration, and label information supplement. Cheng et al. [31] proposed a new DyVTC learning framework for robot terrain classification based on vibration. Chen et al. [32] proposed a pavement damage image classification method based on a VGG-based shallow deep convolutional neural network model that realized the classification and recognition of five kinds of damage images of small-sample asphalt pavement transverse cracks, longitudinal cracks, looseness, cracks, and pits. Andrades et al. [33] analyzed the vibration generated by tire-rolling movement, used machine learning techniques to analyze the signal, and used the self-organizing map (SOM) algorithm to classify and estimate the road surface. Xiao et al. [34] proposed a pavement crack recognition method based on an improved Mask R-CNN model, which can completely identify, locate, and extract cracks with high precision. Yousaf et al. [35] proposed a top-down scheme for the detection and localization of potholes in pavement images to achieve accurate recognition of pothole images. Zhao et al. [36] divided the road state into five categories: dry, wet, snow, ice, and water; the road state feature database was constructed to realize road state recognition based on a support vector machine according to the color and texture feature vectors. Bonfitto et al. [37] proposed a vehicle sideslip angle estimation algorithm based on combined regression and classification artificial neural networks to estimate the sideslip angle and identify the road conditions of dry, wet, and icy road conditions. Ouma et al. [38], casting pavement image segmentation for pothole detection as a problem of clustering multivariate features within mixed pixels, proposed a low-cost urban asphalt pavement pit detection method based on two-dimensional visual images. Wang et al. [39] extracted the time domain and the combined features of the time, frequency, and time-frequency domains of the original vibration signal and constructed an online classification model using a random forest algorithm to realize adaptive recognition of four different terrains. Liang et al. [40] proposed a real-time method for identifying road unevenness with serial acceleration signals and an unevenness-correlated adaptive suspension damping control; the long short-term memory network is used to identify the time domain features of the signal to realize the recognition and classification of road roughness. Yiğit et al. [41] used SVM, MLP, SGD, GNB, and extremely randomized trees techniques to estimate road types based on the brake pressure pulses of ABS.

It can be seen from the previous that there are many studies on road recognition in the existing literature reports, most of which are on road diseases, and there are few studies on road pavement types. Even if there are studies involving pavement types, most of them are aimed at one variable, such as the dry and wet states of the same pavement, or the identification of pavements of 3–5 different materials (asphalt, cement, sand, tiles, slate) in

the same state, and the research difficulty is low. In fact, the existing roads not only have changes in pavement materials or dry and wet conditions, but also have changes in weather conditions, pavement materials, damage degree, and other conditions. In particular, the coupling of various changing conditions leads to many kinds of real forms of pavement, and the difficulty of recognition is also greatly increased, which brings challenges to the outdoor application of wheeled robots. However, quickly and accurately identifying the type and comprehensive condition of the pavement ahead in a complex outdoor environment is a key prerequisite for the wheeled robot to predict whether it can pass safely. In order to promote the accurate recognition of environmental roads and the good application of indoor and outdoor environments by visual navigation robots, this paper will sample 21 kinds of road pavement types covering different weather conditions, different pavement materials, and different degrees of damage on the basis of building a pavement image capture system. By constructing a pavement pattern recognition model based on a classical neural network and the YOLO v8 target detection model, the problem of road pattern recognition is studied. On this basis, by studying the improvement and optimization method of the YOLOv8n pavement pattern recognition model, the best recognition model for improving the accuracy of pavement pattern recognition is explored.

Our innovation and contributions in this paper are threefold:

(1) Through the coupling of weather conditions, pavement materials, damage degree, and other pavement conditions, 21 kinds of complex pavement patterns are constructed, which provides effective and reliable experimental sample data for the effective recognition of pavement patterns under difficult conditions.

(2) Research on pavement recognition of three deep learning algorithms and two Yolo target detection algorithms is carried out. Through comparative analysis, Yolov8n is determined to be the basic framework model for pavement pattern recognition, which provides a reliable foundation model for accurate pavement detection and model improvement research.

(3) The improvement of the YOLOv8n road pattern recognition model was carried out based on the C2f-ODConv module, the AWD adaptive weight downsampling module, the EMA attention mechanism, and three-module collaboration, respectively, and the best improved YOLOv8 recognition model to realize accurate pavement detection was determined.

The research route of this paper is as follows. The first section summarizes the current progress and shortcomings of the road surface recognition problem, and points out the research content and innovation of this paper. The second section introduces the experimental platform for pavement image capture, describes the specific selection and sampling process of the pavement environment, and introduces the classification of the pavement image. In Section 3, the road pattern recognition model based on Resnet 18 is constructed, and the recognition results of the model for the road pattern are studied. In Section 4, a road pattern recognition model based on YOLOv8 is constructed, and the recognition results are compared with the road pattern recognition model based on Resnet 18. In Section 5, four kinds of improved YOLO v8 road recognition models are constructed, and the results of road recognition are studied. Section 6 introduces the research conclusions and related work.

2. Experiment and Data Acquisition Processing

2.1. Experimental Design

2.1.1. Experimental Platform Construction

A pavement image shooting experiment system integrated with a PC, USB camera, and driving image acquisition device was built, as shown in Figure 1. The camera parameters are shown in Table 1. When sampling pavement images, the camera was connected by USB to the computer, the OpenCV library was called in Anaconda, and the camera was driven by the Python 3.11 program to complete the shooting and save the dataset.



Figure 1. Pavement image capture system.

Table 1. Camera parameters.

Resolution Ratio	Frame Rate	Whether There Is Distortion	Driving Mode	Focusing Mode
640 × 480	30 frames per second	Distortionless	USB interface	Manual focus possible
Power supply mode	**Type**	**Lens size**	**Wide angle**	
USB power supply	DF200-1080p	2.8 mm	100°	

2.1.2. Control of Experimental Variable Condition

In order to construct an accurate AGV pavement pattern recognition model, comprehensive pavement recognition data are required, so pavement image data with different conditions, such as different weather conditions, pavement materials, and degrees of damage, need to be collected in the sampling and jointly form the pavement pattern classification. Among these are the following:

(1) Different weather conditions: consider dry pavement, wet pavement, and waterlogged pavement (three categories).

(2) Different pavement materials: consider asphalt road, concrete road, soil road, sand road, and masonry road (five categories).

(3) Pavement damage degree: consider good pavement, slightly damaged pavement, and heavily damaged pavement (three categories). Among these, good pavement refers to intact pavement with no cracks, aging, or pits; in slightly damaged pavement, there is a certain amount of pavement damage, but the mobile robot can pass slowly; with heavy damage, the pavement shows large cracks or pits, etc., and the mobile robot would be at high risk of damage.

2.1.3. Specific Selection and Sampling of the Pavement Environment

The road pattern image shooting process was completed by two people, one of whom was responsible for camera scene selection and adjustment, as well as control of height and angle. In order to improve the randomness of the recognition data, some appropriate adjustments were made to height and angle during sampling. The other person was responsible for using a computer program to complete the shooting and saving. Finally, the photos were summarized and screened.

In the process of selecting different pavements, asphalt road pavements are more common, most of which are intact and have high similarity. In order to diversify the collected image data, different roads, different angles, different traffic signs, and different environments on both sides of the pavement were photographed during the sampling process. At the same time, some fallen leaves and branches were added to make the images more distinguishable.

During the masonry road sampling process, different masonry shapes (diamond and rectangle) and color (gray, red, and yellow) were selected, while some fallen leaves were added to create different scenarios to achieve a variety of masonry road conditions for sampling and to improve the randomness of the data samples.

During the sampling of concrete roads, due to the great difference in the degree of damage to concrete roads, concrete roads were defined as good roads, slightly damaged pavements, and heavily damaged pavements. Then, according to the pavement conditions in different weather conditions, combined with the internal and surrounding environment of the school, the concrete roads were divided into multiple categories for shooting. Good concrete pavement also has smooth and rough points; some deciduous flowers were added to distinguish them; according to the different roughnesses of the surfaces, slightly damaged pavements were also collected for diversity. Due to the serious surface damage of the heavily damaged pavements, there were many obvious depressions, so waterlogged pavement after rain was selected for shooting.

The road condition of sand pavements is more homogenized, so during the shooting process, selective sampling was carried out in different locations and under different road conditions around the gravel roads, combining the vehicles parked on the road and the flowers and plants growing nearby to realize the diversity of sand pavement sampling.

In the process of soil road shooting, the color, texture, and hardness of soil and the presence or absence of plant growth were used as the basis for classification. Some soil roads are seriously damaged, and there will be water waterlogged after rain, which is suitable for shooting waterlogged, heavily damaged pavement. By sampling soil roads in many places, combined with the change of pavement state in different weather conditions, the image sampling of various types of soil roads was completed.

After completion of each shooting, summarize, classify, and filter, create conditions to increase multi-scene shooting according to actual conditions, and finally unify the folders with English names, unify the categories, delete the pictures with serious homogenization, and retain the actual effective pictures as the basic experimental data for the recognition model.

2.2. Pavement Image Data and Its Classification

As shown in Figure 2, within a period of time, road surface photos were taken according to the above experimental methods around roads under different conditions, and 3178 groups of road surface photos of 21 different modes, such as asphalt + good + dry or asphalt + good + wet, were collected. The sampling quantity of each category is shown in Table 2. Folders were established according to classification, and each category was placed separately.



(**1**) asphalt+ good + dry



(**2**) asphalt+ good + wet



(**3**) asphalt+slight + dry

Figure 2. *Cont.*

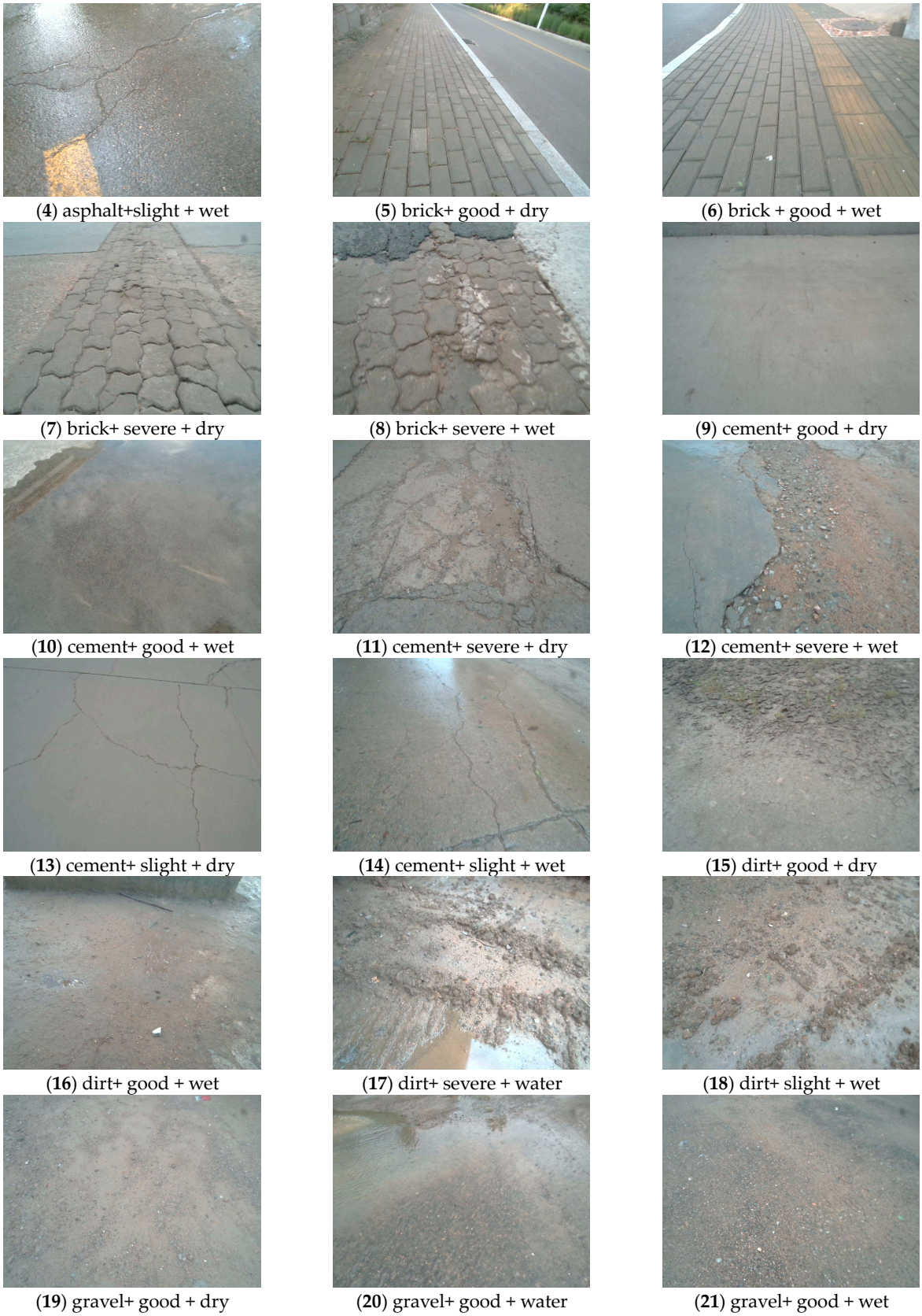


Figure 2. Pavements with different conditions.

Table 2. Sampling quantity of pavements in different modes.

asphalt	**Classification**	**good + dry**	**good + wet**	**slight + dry**	**slight + wet**	-	-
	Quantity	318	216	182	92	-	-
brick	**Classification**	**good + dry**	**good + wet**	**severe + dry**	**severe + wet**	-	-
	Quantity	481	301	32	57	-	-
cement	**Classification**	**good + dry**	**good + wet**	**severe + dry**	**severe + wet**	**slight + dry**	**slight + wet**
	Quantity	231	205	62	84	117	147
dirt	**Classification**	**good + dry**	**good + wet**	**severe + water**	**slight + wet**	-	-
	Quantity	108	118	26	52	-	-
gravel	**Classification**	**good + dry**	**good + water**	**good + wet**	-	-	-
	Quantity	95	170	84	-	-	-

3. Research on Pavement Pattern Recognition Based on Resnet 18

3.1. Construction of a Pavement Pattern Recognition Model Based on Resnet 18

ResNet (Residual Networks) is a deep residual neural network proposed to solve the gradient disappearance and network degradation problems caused by the increase in neural network depth. Its main idea is to avoid direct identity mapping deep in the network by introducing a residual block structure that enables input and output shortcut connections. In view of the real-time requirements of AGV road adaptability recognition, a ResNet 18 network structure with a simple network structure and a small parameter amount is selected in this study, as shown in Figure 3, which can quickly realize feature extraction and recognition of different road surfaces.

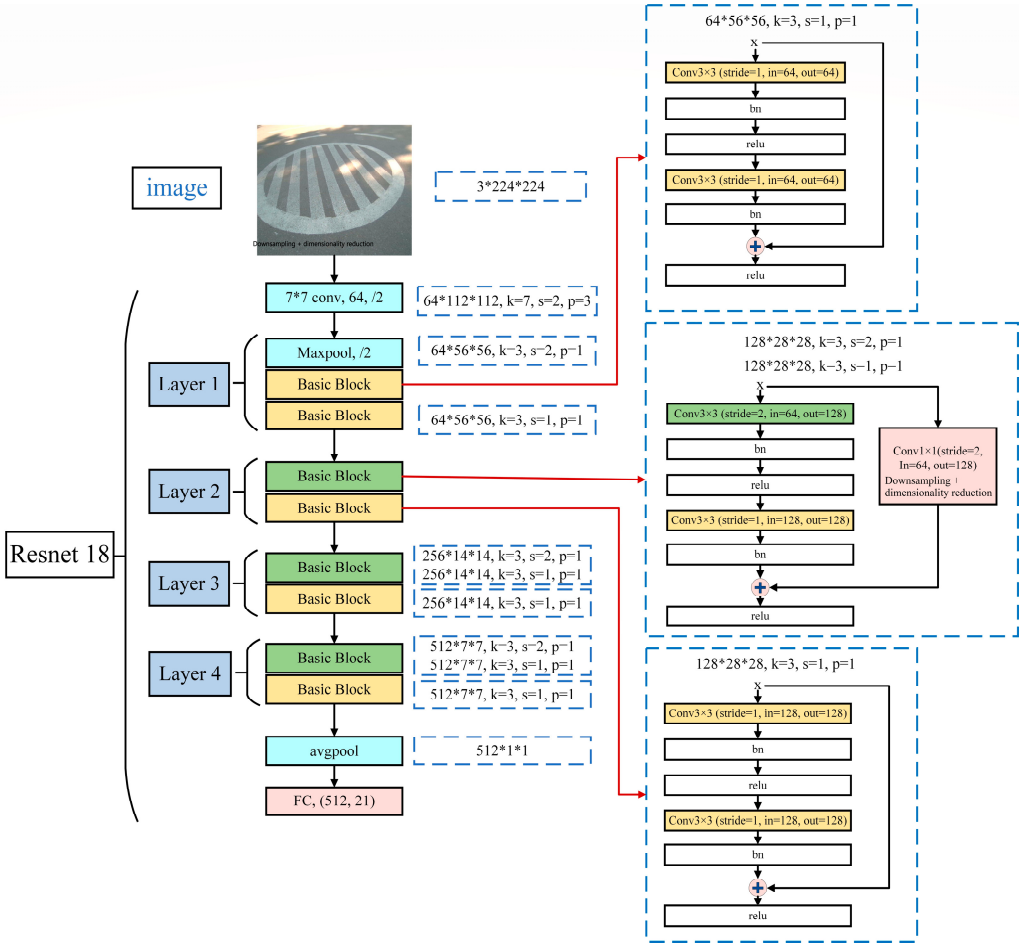


Figure 3. Network structure of Resnet 18.

The 18 convolutional layers contained in the representative network structure of ResNet 18, as shown in Figure 1 above, consist of 1 convolution layer of size 7 × 7, 1 maximum pooling layer, 8 residual structures, 1 global average pooling layer, 1 full connection layer, and 1 softmax output layer. The concrete structure is as follows:

(1) Input layer: input the pavement images with a size of 3 × 224 × 224.

(2) Convolutional layer: 7 × 7 convolutional layers and maximum pooling layers are used to reduce the dimensionality of input images.

(3) Residual block: There are 8 residual blocks in total, each composed of two convolutional layers and one shortcut connection, which is used to solve the gradient disappearance and gradient explosion problems in deep convolutional neural networks. The residual block consists of two 3 × 3 convolutional layers with the same number of output channels, each of which is followed by a batch normalization layer and a ReLU activation function. Then, we skip these two convolution operations through the cross-layer data path and add the input directly before the final ReLU activation function. Such a design requires that the outputs of the two convolutional layers be identical in shape to the inputs so that they can be added. If you want to change the number of channels, you need to introduce an extra 1 × 1 convolutional layer to transform the input into the desired shape before adding it. A common representation of a staggered network is

$$y_l = h(x_l) + F(x_l, W_l)$$

$$x_{l+1} = f(y_l)$$

If the situations of dimension upgrading or dimension reduction are not considered, $h(.)$ is a direct mapping, and $f(.)$ is the activation function, generally using ReLU, then the residual block can be expressed as

$$x_{l+1} = x_l + F(x_l, W_l)$$

For a deeper layer L, its relationship to layer l can be expressed as

$$x_L = x_l + \sum_{i=l}^{L-1} F(x_i, W_i)$$

(4) Global average pooling layer: Perform global average pooling on the feature map and convert the feature map into a one-dimensional vector.

(5) Fully connected layer: The size of the fully connected layer is 21, which is used for the classification output.

(6) Output layer: Using the softmax activation function to generate the probability distribution of 21 categories.

ResNet 18 is usually trained based on a backpropagation algorithm with stochastic gradient descent and optimizes the model parameters by minimizing the cross-entropy loss function, which improves the generalization ability of the model.

3.2. Analysis of Pavement Pattern Recognition Results Based on Resnet 18

According to the 3178 photos of 21 different modes of pavement taken by the experiment, the pavement pattern recognition model of Resnet 18 is used for pattern recognition research. Among these, 2550 photos are randomly selected from 3178 photos as the training set, and the remaining 628 photos are used as the test set. The recognition results obtained are shown in Table 3.

Table 3. Pavement pattern recognition results based on Resnet 18.

Model	Imagesize	Parameters	GFLOPS	Top1acc-val	Infer-Time	FPS
Resnet 18	224	11,187,285	1.82	0.884	2.7	370

It can be seen from the table that the accuracy of pavement pattern recognition is 88.4% when the pavement pattern recognition model based on Resnet 18 recognizes the small data samples of 21 different modes of pavement. Based on the convolutional neural network, pavement pattern recognition can be realized. The recognition of 628 pavement image data in the test set takes 2.7 s, and the frame rate of the pavement pattern recognition model based on Resnet 18 is 370.

On this basis, the AlexNet and MobileNet deep learning algorithms, respectively, are used to recognize pavement patterns under the same training set and test set data sample conditions. The research results show that the pavement pattern classification recognition accuracies of MobileNet and AlexNet are only 0.718 and 0.714, respectively. They are much lower than the 0.884 of Resnet 18 under the same data sample.

4. Research on Road Pattern Recognition Based on YOLOv8

4.1. Construction of Pavement Pattern Recognition Model Based on YOLOv8n

YOLO (You Only Look Once) can get the detection box and category of all targets in the picture through one inference, so it has the advantages of fast recognition speed and strong generalization ability. In order to achieve faster and more accurate image classification, target monitoring, and instance segmentation, Ultralytics launched YOLO v8, a new target detection model, in 2023. The model realizes real-time adjustment of the number of channels of different scale models by designing a C2f structure, and by referring to the design of the ELAN module, the YOLO v8 network has the advantages of lightweight and rich gradient information.

The C2f module consists of a Conv module, a Split, and several Bottlenecks. The beginning of C2f is a Conv module. The input feature image of C2f first passes through Conv and outputs a feature image with a shape of h × w × c. The Split divides the feature image into two groups of h × w × 0.5c feature images in the channel dimension to increase the richness of feature information. The two divided sets of feature images are used as independent branches for cross-layer concat stitching, and one set of feature images is used as the input of n Bottlenecks. Bottleneck is a residual structure. In C2f, n Bottlenecks are stacked continuously to form a main branch, and the output of each Bottleneck is used as an independent branch for cross-layer concat stitching. The feature images output by the main branch of Bottleneck, the two sets of feature images output by Split and the feature images output by n Bottlenecks are stitched in dimensions. Finally, the feature image with a channel number of 0.5 × (n + 2)c is obtained, and the height and width of the feature image remain unchanged. The h × w × 0.5 × (n + 2)c feature image passes through a terminal Conv, and the final obtained feature image is used as the output of C2f.

YOLOv8 has a faster speed, higher accuracy, and fewer parameters. The model is divided into different versions, such as n, s, m, l, and x, according to the depth and width. As the size of the model increases from n to x, the time and computation required for training gradually increase. Although the smaller model has lower accuracy, it has fewer parameters and a faster speed, making it suitable for the application of small computing equipment or models. Our dataset contains more than 3000 images, which cannot be counted as a large model, so YOLO v8n will be used as the main model for relevant research.

In this section, the AGV pavement classification model is based on the YOLO v8n backbone network [42,43]. As shown in Figure 4, different pavement images with an input size of 640*640*3 are processed for feature extraction. The calculation formula for the feature map is as follows:

Feature_new = (Feature_old − kernel + 2* padding)/stride + 1

i.e.:

$$H_n = \frac{H - f + pad * 2}{stride} + 1$$

$$W_n = \frac{W - f + pad * 2}{stride} + 1$$

Further, through global average pooling and linearization processing, it is converted into a one-dimensional vector input to a linear classifier to accurately and efficiently classify 21 categories of images of different pavements.

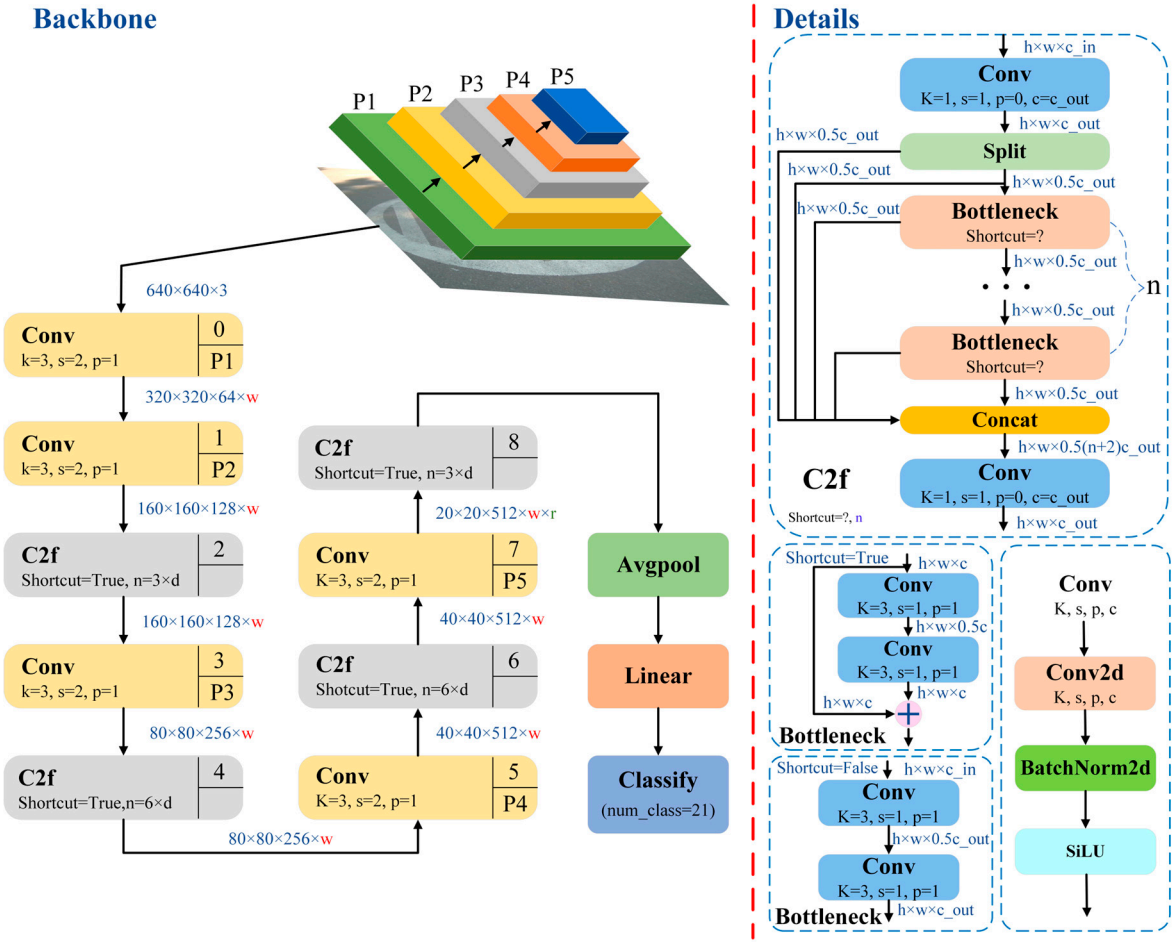


Figure 4. Classification principle of the YOLO v8n algorithm.

4.2. Analysis of Pavement Pattern Recognition Results Based on YOLO v8n

According to the 3178 photos of 21 different modes of pavement taken by the experiment, the pavement pattern recognition model based on YOLO v8n is used for pattern recognition research. Consistent with the pavement pattern recognition based on Resnet 18, 2550 photos were randomly selected from 3178 photos as the training set, and the remaining 628 photos were used as the test set. The recognition results are shown in Table 4.

Table 4. Results of road pattern recognition based on YOLO v8n.

Model	Imagesize	Parameters	GFLOPS	Top1acc-val	Infer-Time	FPS
YOLO v8	224	1,111,317	0.19	0.915	0.8	1250

It can be seen from the table that the accuracy of pavement pattern recognition can reach more than 90% when the pavement pattern recognition model based on YOLO v8n identifies the small data samples of 21 different modes of pavement. Compared with the pavement pattern recognition model based on Resnet 18, the recognition accuracy is improved by 3.1%, and the recognition accuracy is further improved. Therefore, the pavement pattern recognition accuracy based on YOLO v8 is higher. At the same time, after the training of the recognition model is completed, the recognition time of the 628 pavement

image data in the test set is only 0.8 s. Compared with the road pattern recognition model based on Resnet 18, the test recognition time is shortened by more than 200%.

In order to verify the effectiveness of YOLO v8n, the yolov8l model with moderate computational load and comprehensive performance was selected from the other four series models of YOLO v8, including yolov8s, yolov8m, yolov8l, and yolov8x, for pavement pattern recognition. The results were compared with those of YOLO v8n.

According to the comparative analysis of Tables 4 and 5, the recognition accuracy of Yolov8*l* is 0.825. Compared with the road pattern recognition model based on YOLO v8n, the recognition accuracy is reduced by 9%, the test recognition time is extended by 150%, and the response speed and recognition accuracy are far lower than those of Yolov8n.

Table 5. Results of road pattern recognition based on YOLO v8*l*.

Model	Parameters	GFLOPS	Top1acc-val	Infer-Time	FPS
YOLO v8	36,226,645	99.1	0.825	2	500

Therefore, the pavement pattern recognition model based on YOLO v8n has the advantages of better recognition accuracy, response speed, and recognition effect than Resnet 18. However, although the recognition accuracy of the pavement pattern recognition model based on YOLO v8n reaches 91.5%, the recognition accuracy still needs to be further improved in order to improve its applicability.

5. Research on Pavement Recognition Based on Improved YOLO v8

The recognition accuracy of the model depends on the structure and performance of the recognition model. In order to further improve recognition accuracy, this section will use YOLOv8n as the benchmark model to find a pavement pattern recognition model with higher recognition accuracy by exploring the improved method of YOLO v8n.

5.1. Research on Improvement of YOLO v8 Pavement Recognition Model Based on the C2f-ODConv Module

Dynamic convolution means that in the convolution process, the weight of the convolution kernel is not fixed but can be dynamically adjusted according to the different input data. The advantage of dynamic convolution is that it can make the convolution kernel better adapt to the characteristics of the input data and improve the performance of the convolution network.

ODConv [44] uses a multidimensional attention mechanism to learn complementary attention along the four dimensions of the kernel space through a parallel strategy, which extends the dynamic properties of one dimension in CondConv and considers the dynamics of the null domain, input channel, output channel, and other dimensions at the same time; thus, it is called a full-dimensional dynamic convolution. In this paper, ODConv is used to replace the convolution inside the Bottleneck in the C2f module of the YOLO v8 road recognition model to form a new C2f-ODConv module, as shown in Figure 5.

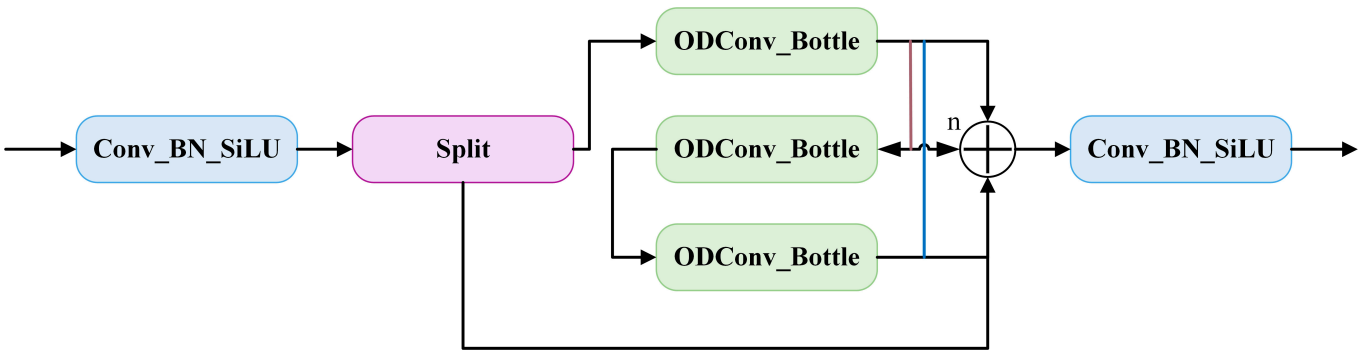


Figure 5. C2f-ODConv module structure.

On this basis, the improved YOLO v8 pavement recognition model based on the C2f-ODConv module is defined as Yolo v8n-C, as shown in Figure 6.

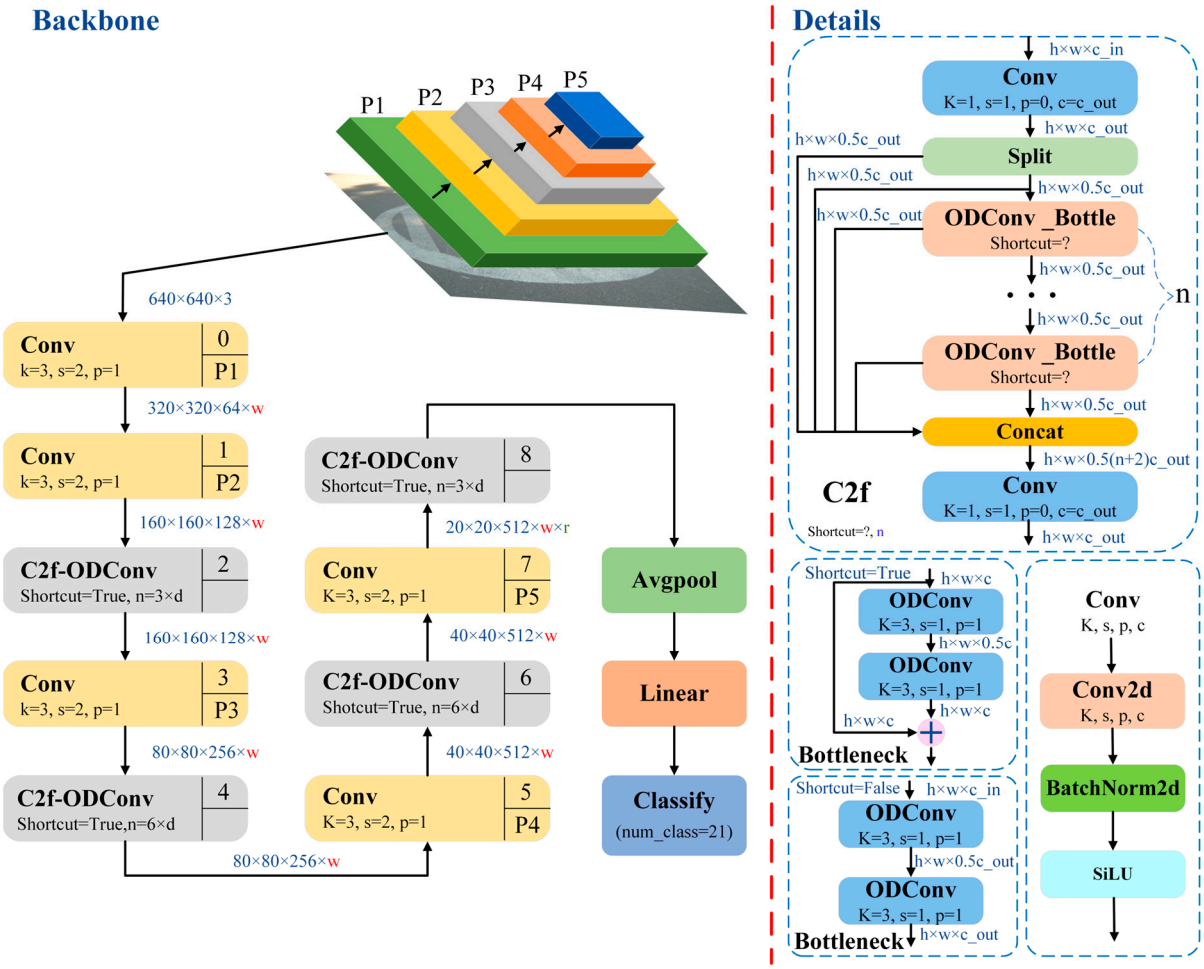


Figure 6. Improved YOLO v8 pavement recognition model based on the C2f-ODConv module.

5.2. Research on Improvement of YOLO v8 Pavement Recognition Model Based on the AWD Adaptive Weight Downsampling Module

Traditional downsampling methods, such as convolution with MaxPooling and a step size of 2, lose the relative importance of elements. Influenced by the idea of RFAConv [45], the AWD (adaptive weight downsampling) downsampling module was designed. As shown in Figure 7, AWD fully considers the relationship between elements when downsampling. The upper branch fuses the attention mechanism, uses the average pooling method to obtain the importance of the element, and finally obtains the weight of the element through the softmax operation. Meanwhile, the following branches use group convolution [46] to reduce the number of parameters and calculations. Finally, multiplication and summation operations are performed to obtain the final downsampling results.

On this basis, the AWD downsampling module is used to replace some Conv modules in the YOLO v8 pavement recognition model constructed above, and the improved YOLO v8 pavement recognition model based on the AWD adaptive weight downsampling module is obtained, which is defined as Yolo v8n-A, as shown in Figure 8.

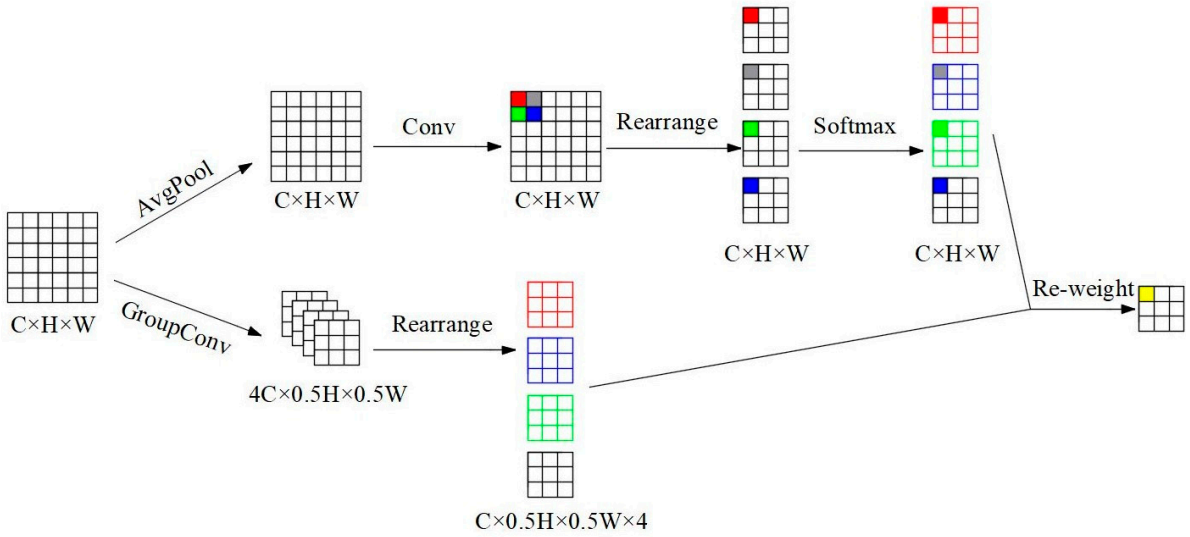


Figure 7. AWD structure.

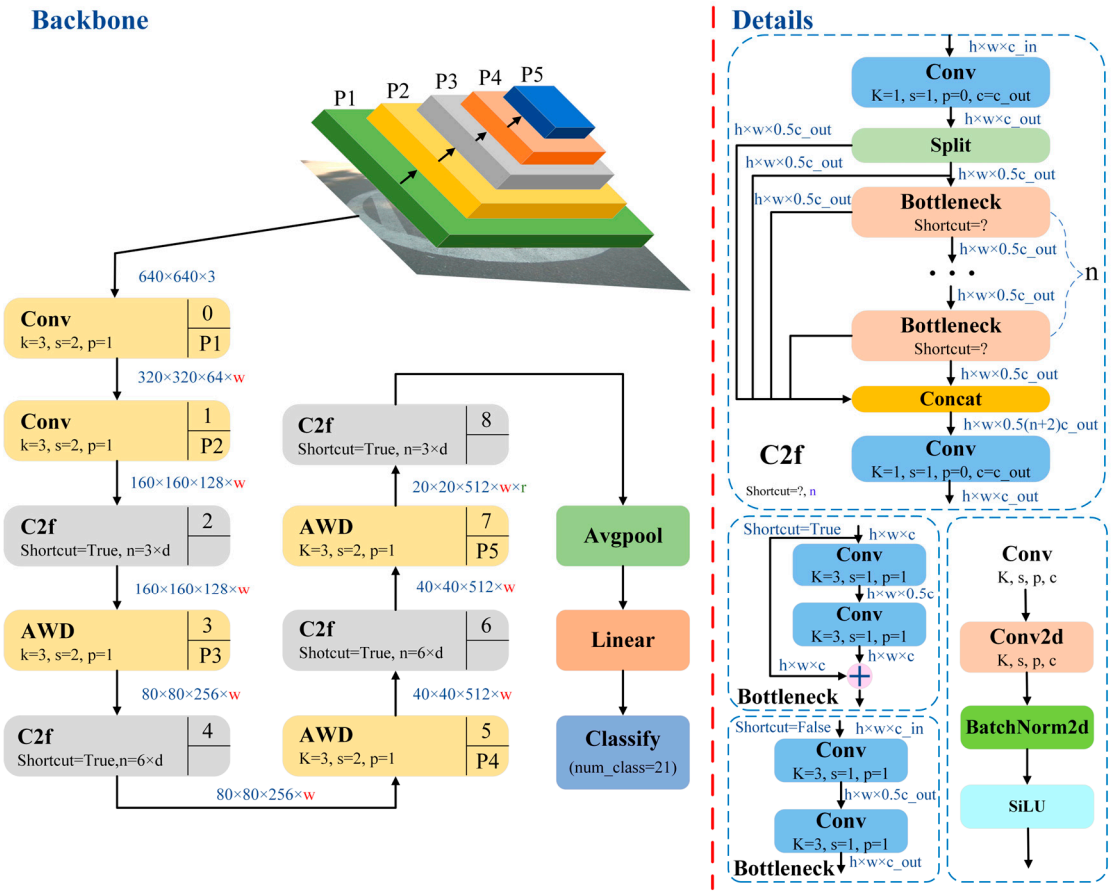


Figure 8. Improved YOLO v8 pavement recognition model based on the AWD adaptive weight downsampling module.

5.3. Research on Improvement of YOLO v8 Pavement Recognition Model Based on the EMA Attention Mechanism

Attention mechanism, a deep learning optimization strategy that mimics human attention, constructs dynamic weights by selecting relevant and irrelevant information features to help the network record positional relationships and estimate the importance

of different information, accordingly weakening useless information and strengthening important information to improve network efficiency.

There are three main types of attention mechanisms, namely channel attention, spatial attention, and hybrid attention mechanism. Among these, channel attention extracts attention between channels by modeling cross-dimensional interactions. For example, the SE module is a representative channel-attention module. Spatial attention establishes cross-space and cross-channel information interactions by semantic dependencies between spatial and channel dimensions in the feature map. For example, the convolution block attention module (CBAM) is a typical spatial-attention module. The hybrid attention mechanism combines channel attention and spatial attention to learn target features in multiple dimensions.

In order to improve the accuracy of pavement pattern recognition, this paper introduces a new hybrid attention mechanism: efficient multiscale attention without dimensionality reduction (EMA) [47]. Its structure is shown in Figure 9. EMA can achieve efficient learning of multiscale attention through a sequential processing method based on a grouping structure.

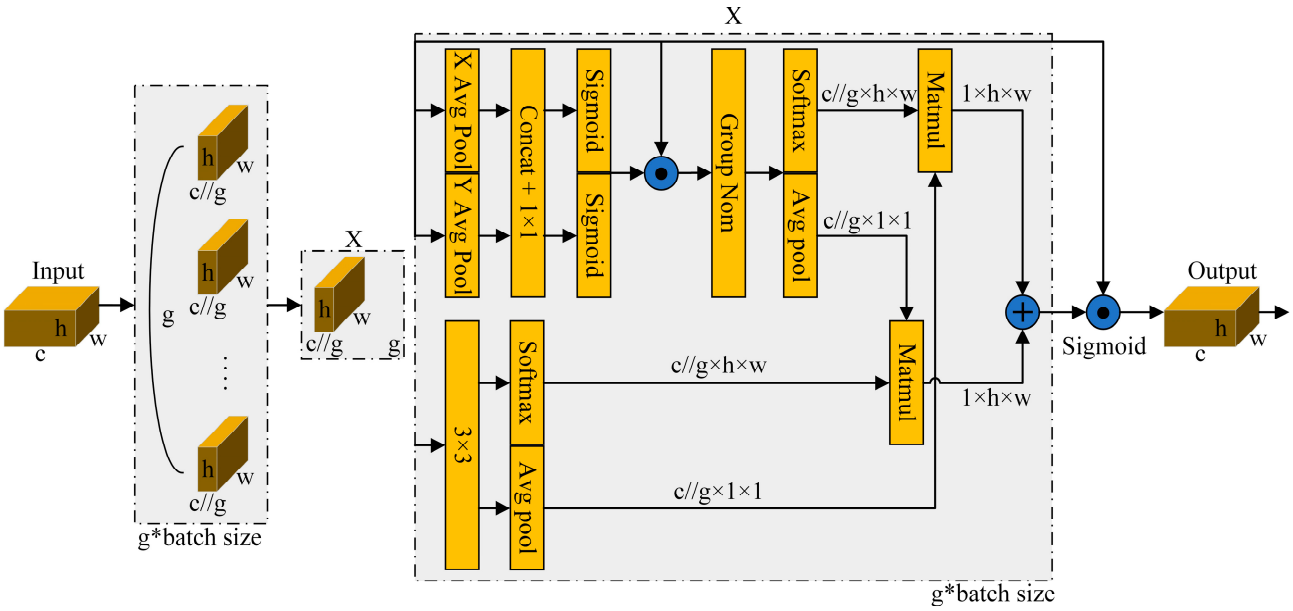


Figure 9. EMA structure.

The EMA module is added before the Avgpool module of the YOLO v8 pavement recognition model constructed above, and the rest of the model remains unchanged. The improved YOLO v8 pavement recognition model based on the EMA attention mechanism module is defined as Yolo v8n-E, as shown in Figure 10.

5.4. Research on Comprehensive Improvement of the YOLO v8 Pavement Recognition Model Based on Multimodule Collaboration

In the process of improving the YOLO v8 pavement recognition model, the three separate improvement points introduced or improved modules mentioned in the paper have their own advantages and characteristics. Based on the advantages of the aforementioned C2f-ODConv module, the AWD adaptive weight downsampling module, and the EMA attention mechanism module, the YOLO v8 pavement recognition model constructed above is further improved by the multimodule collaboration of the C2f-ODConv module, the AWD adaptive weight downsampling module, and the EMA attention mechanism. The multimodule collaborative YOLO v8 road recognition model is thus obtained, which is defined as Yolo v8n-CAE, as shown in Figure 11.

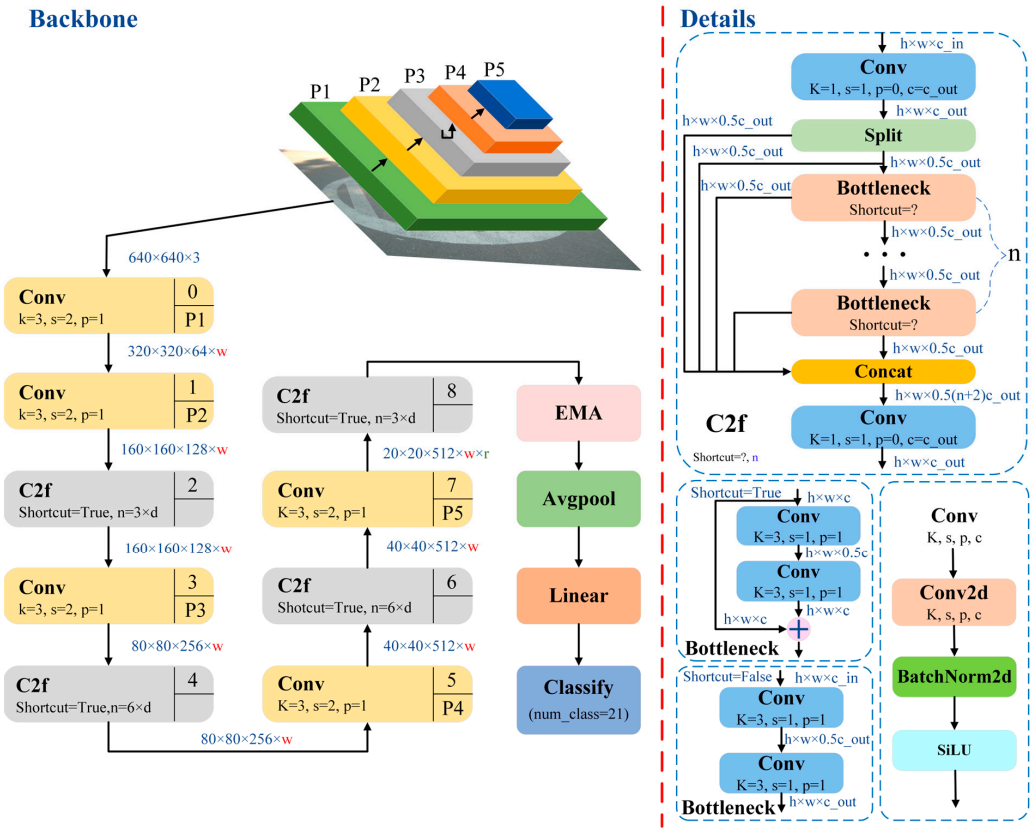


Figure 10. The improved YOLO v8 pavement recognition model based on the EMA attention mechanism module.

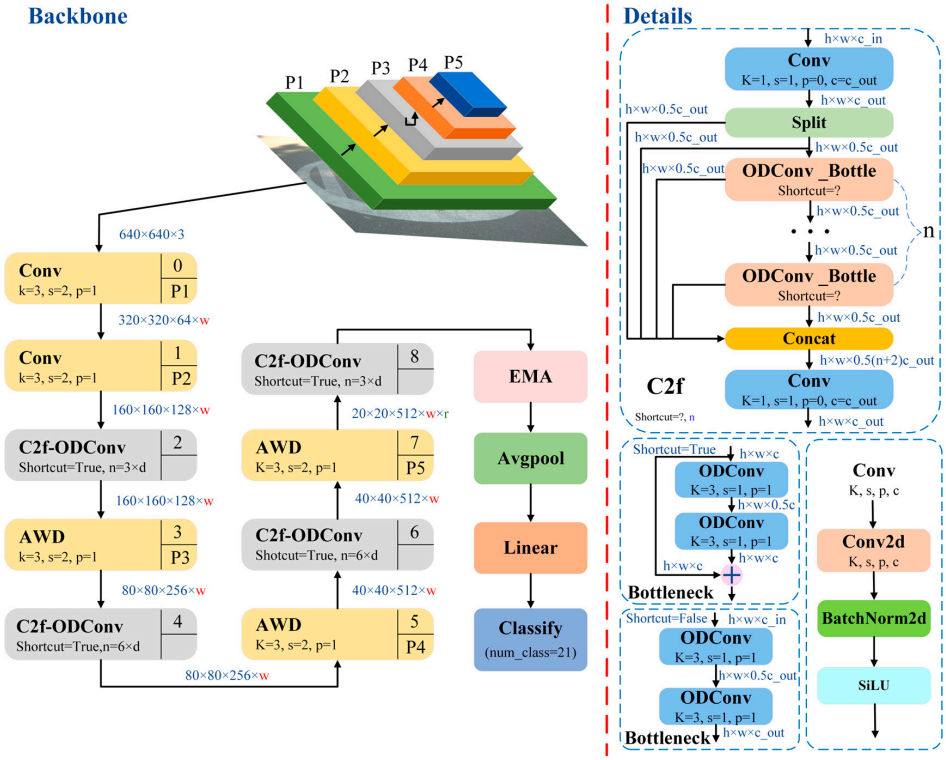


Figure 11. Improved YOLO v8 road recognition model based on multimodule collaboration.

5.5. Analysis of Pavement Pattern Recognition Results Based on Improved YOLO v8

According to the 3178 pavement photos of 21 different modes taken in the experiment, the improved YOLO v8 road recognition model based on the C2f-ODConv module, the improved YOLO v8 pavement recognition model based on the AWD adaptive weight downsampling module, the improved YOLO v8 road recognition model based on the EMA attention mechanism, and the improved YOLO v8 pavement recognition model based on multimodule collaboration are used for pattern recognition research. Similarly, 2550 photos are randomly selected from 3178 photos as training sets, and the remaining 628 photos are used as test sets. The recognition results are shown in Table 6.

Table 6. The road pattern recognition results based on improved YOLO v8.

Model	Improved Point	Imagesize	Parameters	GFLOPS	Top1acc-val
Yolo v8n-C	Yolo v8n + C2f-ODConv	224	1,146,273	0.10	0.927
Yolo v8n-A	Yolo v8n + AWD	224	832,437	0.17	0.921
Yolo v8n-E	Yolov8n + EMA	224	1,121,685	0.19	0.921
Yolo v8n-CAE	Yolo v8n + C2f-ODConv + AWD + EMA	224	877,761	0.09	0.932

From the table, it can be seen that the pavement pattern recognition accuracy of the four pavement pattern recognition models based on the improved YOLO v8 to identify the small data samples of 21 different modes of roads is more than 92%. Compared with the pavement pattern recognition model based on Resnet 18, the recognition accuracy is increased by 3.7% or more. Compared with the pavement pattern recognition model based on YOLO v8n, the recognition accuracy is increased by 0.6% or more, and the highest is increased by 1.7%. Recognition accuracy is significantly improved. Therefore, the improved YOLO v8 is effective for pavement pattern recognition accuracy.

Among the three improved YOLO v8 pavement pattern recognition models based on single module improvement, the improved YOLO v8 pavement pattern recognition model based on the C2f-ODConv module has the highest accuracy of 92.7%. Among the pavement pattern recognition models based on improved YOLO v8 obtained by four different improvement methods, the improved YOLO v8 pavement pattern recognition model based on multimodule collaboration has the highest recognition accuracy, and the recognition accuracy is more than 93%. Therefore, the improved YOLO v8 pavement recognition model based on multimodule collaboration has the best recognition effect. When the recognition model is selected by accuracy-orienting, the improved YOLO v8 pavement recognition model based on multimodule collaboration is the best model.

6. Conclusions

In order to realize the visual navigation robot's accurate recognition of and navigation control in the environment, this paper carries out research on the image recognition of pavement patterns. By building a data acquisition test system, 21 different patterns of pavement images, such as different weather conditions, pavement materials, and degrees of pavement damage, are collected and the pavement environment dataset is formed. The classical neural network algorithm Resnet 18 and the emerging image target detection algorithm YOLOv8 are used to carry out pavement pattern recognition research. On this basis, in order to further improve recognition performance, the pavement pattern recognition model based on YOLOv8 is improved in various ways, and the pavement pattern recognition model based on improved YOLOv8 is tested for recognition performance. The following conclusions are obtained:

(1) The experimental system of pavement image shooting is set up, and 3178 pavement images of three different weather conditions, such as dry pavement, wet pavement, and waterlogged pavement, five different pavement materials, such as asphalt road, concrete road, soil road, sand road, and masonry road, and three different damage degrees, such as

good pavement, slightly damaged pavement, and heavy damaged pavement, are jointly collected to form 21 different modes of pavement image datasets.

(2) Based on the study of the network structure mechanism of Resnet 18, a pavement pattern recognition model based on Resnet 18 is constructed, 2550 photos are randomly selected as the training set, and the remaining 628 photos are used as the test set. The accuracy of pavement pattern recognition is found to be 88.4%.

(3) On the basis of studying the mechanism of the YOLO v8 target detection model, a pavement pattern recognition model based on YOLOv8n is constructed, and the recognition model is trained according to different conditions of pavement images. The test recognition accuracy of 21 groups of pavement images with different modes is 91.5%. Compared with the pavement pattern recognition model based on Resnet 18, the recognition accuracy is improved by 3.1%. Recognition accuracy and response speed are both improved, and the recognition effect is better.

(4) In order to further improve the recognition ability of the pavement pattern recognition model based on YOLOv8n, the C2f-ODConv module, the AWD adaptive weight downsampling module, and the EMA attention mechanism, respectively, are introduced to optimize, replace, or supplement some modules of the pavement pattern recognition model based on YOLOv8n. Three improved YOLO v8 pavement recognition models based on single module changes are obtained, namely an improved YOLO v8 road surface recognition model based on the C2f-ODConv module, an improved YOLO v8 road surface recognition model based on the AWD adaptive weight downsampling module, and an improved YOLO v8 road surface recognition model based on the EMA attention mechanism. And an improved YOLO v8 road surface recognition model based on multimodule coordination optimized by multiple modules is obtained. The pavement pattern recognition results show that the accuracy of the four optimized pavement pattern recognition models reaches more than 92%. Compared with the pavement pattern recognition model based on YOLO v8n, recognition accuracy is improved.

(5) Among the four different optimization recognition models, the improved YOLO v8 pavement recognition model based on multimodule collaboration has the highest recognition accuracy, reaching more than 93%. The improved YOLO v8 pavement recognition model based on the C2f-ODConv module has the highest recognition accuracy among the three improved YOLO v8 pavement pattern recognition models with single-module improvement.

The research provides the pavement recognition model based on improved YOLO v8 for pavement pattern recognition, provides the reference for the performance optimization and improvement of the YOLO v8 algorithm, and provides theoretical guidance and the research basis for the accurate recognition and navigation control of the environment and the application of technology for visual navigation robots.

Author Contributions: Conceptualization, methodology, investigation, experiment, data curation, writing—original draft preparation, writing—review and editing, X.Z. writing—original draft preparation, writing—review and editing, project management, Y.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Innovation Ability Improvement Project of Science and Technology smes in Shandong Province (Grant No. 2023TSGC0334), and the Jining City key research and development project (Grant No. 2023KJHZ001).

Data Availability Statement: Data are within the paper and will be made available on request.

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

References

1. Liu, J.; Liu, H.; Chen, H.; Mao, F. Road Types Identification Method of Unmanned Tracked Vehicles Based on Fusion Features. *Acta Armamentarii* **2023**, *44*, 1267–1276.
2. Yamaguchi, T.; Mizutani, T. Road crack detection interpreting background images by convolutional neural networks and a self-organizing map. *Comput. Aided Civ. Infrastruct. Eng.* **2023**, 1–25. [CrossRef]

3. Oum, Y.O.; Hahn, M. Wavelet-morphology based detection of incipient linear cracks in asphalt pavements from RGB camera imagery and classification using circular Radon transform. *Adv. Eng. Inform.* **2016**, *30*, 481–499. [CrossRef]
4. Tedeschi, A.; Benedetto, F. A real-time automatic pavement crack and pothole recognition system for mobile Android-based devices. *Adv. Eng. Inform.* **2017**, *32*, 11–25. [CrossRef]
5. Wu, W.; Tian, S.; Zhang, Z.; Jin, B.; Qiu, Z. Research on Surface Geometry Parameter Recognition and Model Reconstruction of Uneven Road. *Automot. Eng.* **2023**, *45*, 273–284.
6. Wu, W.; Tian, S.; Zhang, Z.; Zhang, B. Research on Semantic Segmentation of Uneven Features of Unpaved Road. *Automot. Eng.* **2023**, *45*, 1468–1478.
7. Maeda, H.; Sekimoto, Y.; Seto, T.; Kashiyama, T.; Omata, H. Road Damage Detection and Classification Using Deep Neural Networks with Smartphone Images. *Comput. Aided Civ. Infrastruct. Eng.* **2018**, *33*, 1127–1141. [CrossRef]
8. Zhao, J.; Liu, Y.; Zhu, B.; Li, Y.; Li, Y.; Kong, D.; Jiang, H. Research on Road Recognition Algorithm of Off-Road Vehicle Based on Shap-Rf Framework. *Chin. J. Theor. Appl. Mech.* **2022**, *54*, 2922–2935.
9. Yang, F.; Zhang, L.; Yu, S.; Prokhorov, D.; Mei, X.; Ling, H. Feature Pyramid and Hierarchical Boosting Network for Pavement Crack Detection. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 1525–1535. [CrossRef]
10. Chao, F.; Ping, S. Multi-scale Feature Fusion Network for Few-Shot Bridge Pavement Crack Segmentation. *Radio Eng.* **2023**. Available online: https://link.cnki.net/urlid/13.1097.TN.20230904.0843.002 (accessed on 27 June 2023).
11. Wang, S.; Kodagoda, S.; Shi, L.; Wang, H. Road-Terrain Classification For Land Vehicles. *IEEE Veh. Technol. Mag.* **2017**, *12*, 34–41. [CrossRef]
12. Liu, H.; Liu, C.; Han, L.; He, P.; Nie, S. Road Information Recognition Based on Multi-Sensor Fusion in Off-Road Environment. *Trans. Beijing Inst. Technol.* **2023**, *43*, 783–791.
13. Shi, W.; Li, Z.; Lv, W.; Wu, Y.; Chang, J.; Li, X. Laplacian Support Vector Machine for Vibration-Based Robotic Terrain Classification. *Electronics* **2020**, *9*, 513. [CrossRef]
14. Jiang, C.; Ye, Y.; Hong, L.; Tian, W.; Shuo, D.; Jian, L. Multi-distress detection method for asphalt pavements based on multi-branch deep learning. *J. Southeast Univ. Nat. Sci. Ed.* **2023**, *53*, 123–129.
15. Dimastrogiovanni, M.; Cordes, F.; Reina, G. Terrain Estimation for Planetary Exploration Robots. *Appl. Sci.* **2020**, *10*, 6044. [CrossRef]
16. Wang, Q.; Xu, J.; Su, J.; Zong, G.; Xue, M. Study on Pavement Condition Recognition Method Based on Improved ALexNet Model. *J. Highw. Transp. Res. Dev.* **2023**, *40*, 209–218.
17. Qin, Y.; Xiang, C.; Wang, Z.; Dong, M. Road excitation classification for semi-active suspension system based on system response. *J. Vib. Control.* **2017**, *24*, 2732–2748. [CrossRef]
18. Wang, X.; Li, S.; Liang, X.; Li, S.; Zheng, J. Fast Identification Model for Complex Pavement based on Structural Reparameterization and Adaptive Attention. *China J. Highw. Transp.* **2023**. Available online: https://link.cnki.net/urlid/61.1313.U.20231127.0947.002 (accessed on 27 June 2023).
19. Wang, W.; Zhang, B.; Wu, K.; Chepinskiy, S.A.; Zhilenkov, A.A.; Chernyi, S.; Krasnov, A.Y. A visual terrain classification method for mobile robots' navigation based on convolutional neural network and support vector machine. *Trans. Inst. Meas. Control.* **2021**, *44*, 744–753. [CrossRef]
20. Xu, T.; Jiang, Z.; Liang, Y.; Chen, Z.; Sun, L. Pavement Distress Detection Based on Historical Information. *J. Tongji Univ. Nat. Sci.* **2022**, *50*, 562–570.
21. Chen, C.; Chandra, S.; Han, Y.; Seo, H. Deep Learning-Based Thermal Image Analysis for Pavement Defect Detection and Classification Considering Complex Pavement Conditions. *Remote Sens.* **2022**, *14*, 106. [CrossRef]
22. Kou, F.; He, J.; Li, M.; Xu, J.; Wu, D. Adaptive Fuzzy Control of an Electromagnetic Hybrid Suspension Based on Road Recognition. *J. Vib. Shock.* **2023**, *42*, 303–311.
23. Dewangan, D.K.; Sahu, S.P. RCNet: Road classification convolutional neural networks for intelligent vehicle system. *Intell. Serv. Robot.* **2021**, *14*, 199–214. [CrossRef]
24. Zhang, L.; Guan, K.; Ding, X.; Guo, P.; Wang, Z.; Sun, F. Tire-Road Friction Estimation Method Based on Image Recognition and Dynamics Fusion. *Automot. Eng.* **2023**, *45*, 1222–1234.
25. Yousefzadeh, M.; Azadi, S.; Soltani, A. Road profile estimation using neural network algorithm. *J. Mech. Sci. Technol.* **2010**, *24*, 743–754. [CrossRef]
26. Du, Z.; Zhang, W.; Zhu, X. Road Roughness Assessment based on Fusion of Connected-Vehicles Data. *China J. Highw. Transp.* **2024**, 1–26. Available online: http://kns.cnki.net/kcms/detail/61.1313.U.20230627.0943.004.html (accessed on 27 June 2023).
27. Bai, C.; Guo, J.; Guo, L.; Song, J. Deep Multi-Layer Perception Based Terrain Classification for Planetary Exploration Rovers. *Sensors* **2019**, *19*, 3102. [CrossRef]
28. Yang, Y.; Wang, M.; Liu, C.; Xu, H.; Zhang, X. Intelligent identification of asphalt pavement cracks based on semantic segmentation. *J. Zhejiang Univ. Eng. Sci.* **2023**, *57*, 2094–2105.
29. Šabanovič, E.; Žuraulis, V.; Prentkovskis, O.; Skrickij, V. Identification of Road-Surface Type Using Deep Neural Networks for Friction Coefficient Estimation. *Sensors* **2020**, *20*, 612. [CrossRef] [PubMed]
30. Li, L.; Bao, Y.; Yang, W.; Chu, Q.; Wang, G. Standardized constructing method of a roadside multi-source sensing dataset. *J. Jilin Univ. Eng. Technol. Ed.* **2024**, 1–7. [CrossRef]

31. Cheng, C.; Chang, J.; Lv, W.; Wu, Y.; Li, K.; Li, Z.; Yuan, C.; Ma, S. Frequency-Temporal Disagreement Adaptation for Robotic Terrain Classification via Vibration in a Dynamic Environment. *Sensors* **2020**, *20*, 6550. [CrossRef]
32. Chen, J.; Ji, X.; Que, Y.; Dai, Y.; Jiang, Z. Classification Recognition of Pavement Disaster with Small Sample Size Based on Improved VGG Algorithm. *J. Hunan Univ. Nat. Sci.* **2023**, *50*, 206–216.
33. Andrades, I.S.; Aguilar, J.J.C.; García, J.M.V.; Carrillo, J.A.C.; Lozano, M.S. Low-Cost Road-Surface Classification System Based on Self-Organizing Maps. *Sensors* **2020**, *20*, 6009. [CrossRef] [PubMed]
34. Xiao, L.; Li, W.; Yuan, B.; Cui, Y.; Gao, R.; Wang, W. A Pavement Crack Identification Method Based on Improved Instance Segmentation Model. *Geomat. Inf. Sci. Wuhan Univ.* **2023**, *48*, 765–776. [CrossRef]
35. Yousaf, M.H.; Azhar, K.; Murtaza, F.; Hussain, F. Visual analysis of asphalt pavement for detection and localization of potholes. *Adv. Eng. Inform.* **2018**, *38*, 527–537. [CrossRef]
36. Zhao, J.; Wu, H.; Chen, L. Road Surface State Recognition Based on SVM Optimization and Image Segmentation Processing. *J. Adv. Transp.* **2017**, *2017*, 6458495. [CrossRef]
37. Bonfitto, A.; Feraco, S.; Tonoli, A.; Amati, N. Combined regression and classification artificial neural networks for sideslip angle estimation and road condition identification. *Veh. Syst. Dyn.* **2019**, *58*, 1766–1787. [CrossRef]
38. Oumaa, Y.O.; Hahn, M. Pothole detection on asphalt pavements from 2D-colour pothole images using fuzzy c-means clustering and morphological reconstruction. *Autom. Constr.* **2017**, *83*, 196–211. [CrossRef]
39. Wang, M.; Ye, L.; Sun, X. Adaptive online terrain classification method for mobile robot based on vibration signals. *Int. J. Adv. Robot. Syst.* **2021**, *18*, 1–14. [CrossRef]
40. Liang, G.; Zhao, T.; Shangguan, Z.; Li, N.; Wu, M.; Lyu, J.; Du, Y.; Wei, Y. Experimental study of road identification by LSTM with application to adaptive suspension damping control. *Mech. Syst. Signal Process.* **2022**, *177*, 1–20. [CrossRef]
41. Yiğit, H.; Köylü, H.; Eken, S. Estimation of road surface type from brake pressure pulses of ABS. *Expert Syst. Appl.* **2023**, *212*, 1–11. [CrossRef]
42. Zhou, Y.; Zhu, W.; He, Y.; Li, Y. YOLOv8-based Spatial Target Part Recognition. In Proceedings of the 2023 IEEE 3rd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA), Chongqing, China, 26–28 May 2023; Volume 3, pp. 1684–1687.
43. Su, Z.; Huang, Z.; Qiu, F.; Guo, C.; Yin, X.; Wu, G. Weld defect detection of Aviation Aluminum alloy based on improved YOLOv8. *J. Aerosp. Power* **2024**, *39*, 20230414.
44. Li, C.; Zhou, A.; Yao, A. Omni-dimensional dynamic convolution. *arXiv* **2022**, arXiv:2209.07947.
45. Zhang, X.; Liu, C.; Yang, D.; Song, T.; Ye, Y.; Li, K.; Song, Y. RFAConv: Innovating Spatital Attention and Standard Convolutional Operation. *arXiv* **2023**, arXiv:2304.03198.
46. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*. Available online: https://proceedings.neurips.cc/paper_files/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf (accessed on 27 June 2023). [CrossRef]
47. Ouyang, D.; He, S.; Zhang, G.; Luo, M.; Guo, H.; Zhan, J.; Huang, Z. Efficient Multi-Scale Attention Module with Cross-Spatial Learning. In Proceedings of the ICASSP 2023–2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 4–10 June 2023.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.