infrastructures

MDPI

Article

Asphalt Pavement Damage Detection through Deep Learning Technique and Cost-Effective Equipment: A Case Study in Urban Roads Crossed by Tramway Lines

Marco Guerrieri [1,*], Giuseppe Parla [2], Masoud Khanmohamadi [1] and Larysa Neduzha [3]

[1] Department of Civil, Environmental and Mechanical Engineering (DICAM), University of Trento, Via Mesiano 77, 3812 Trento, Italy; masoud.khanmohamadi@unitn.it
[2] ISMETT, Via E. Tricomi, 5, 90127 Palermo, Italy; peppeparla@yahoo.it
[3] Department of Technical Mechanics, Ukrainian State University of Science and Technology, Lazaryan 2, 49010 Dnipro, Ukraine; nlorhen@i.ua
* Correspondence: marco.guerrieri@unitn.it

Abstract: Asphalt pavements are subject to regular inspection and maintenance activities over time. Many techniques have been suggested to evaluate pavement surface conditions, but most of these are either labour-intensive tasks or require costly instruments. This article describes a robust intelligent pavement distress inspection system that uses cost-effective equipment and the 'you only look once' detection algorithm (YOLOv3). A dataset for flexible pavement distress detection with around 13,135 images and 30,989 bounding boxes of damage was used during the neural network training, calibration, and validation phases. During the testing phase, the model achieved a mean average precision of up to 80%, depending on the type of pavement distress. The performance metrics (loss, precision, recall, and RMSE) that were applied to estimate the object detection accuracy demonstrate that the technique can distinguish between different types of asphalt pavement damage with remarkable accuracy and precision. Moreover, the confusion matrix obtained in the validation process shows a distress classification sensitivity of up to 98.7%. The suggested technique was successfully implemented in an inspection car. Measurements conducted on urban roads crossed by tramway lines in the city of Palermo proved the real-time ability and great efficacy of the detection system, with potentially remarkable advances in asphalt pavement examination efficacy due to the high rates of correct distress detection.

Keywords: asphalt pavement; distress detection; deep learning; urban roads

Citation: Guerrieri, M.; Parla, G.; Khanmohamadi, M.; Neduzha, L. Asphalt Pavement Damage Detection through Deep Learning Technique and Cost-Effective Equipment: A Case Study in Urban Roads Crossed by Tramway Lines. *Infrastructures* **2024**, *9*, 34. https://doi.org/10.3390/infrastructures9020034

Academic Editor: Rui Micaelo

Received: 4 January 2024
Revised: 5 February 2024
Accepted: 6 February 2024
Published: 16 February 2024

1. Introduction

Most state road operators supervise road pavement distress to support their asset management systems. An efficient pavement management system (PMS) necessitates the integration of modules for pavement inspection, condition assessment, condition prediction, optimisation, and decision-making regarding maintenance actions. The initial stage involves determining pavement conditions, which can be performed using a range of techniques, from manual to fully automated, to reduce subjectivity and increase efficiency. The implementation of an efficient pavement management system (PMS) requires accurate pavement inspections [1]. Correctly identifying pavement distress is of huge importance for maintaining high levels of safety and performance for road transportation systems. Road pavement degradation is predominantly due to light and heavy vehicle traffic, weather conditions, and sunlight. Pavements can be classified into four main categories according to the materials used, namely asphalt (also known as flexible pavement), concrete, gravel, brick, and block. Over 90% of the total European road network has a flexible pavement [2]. Flexible pavements are composed of different layers: surface, binder, base, and subbase

courses. Distresses can be grouped into two different types (i.e., cracking and non-cracking), as shown in Figure 1 [3,4].

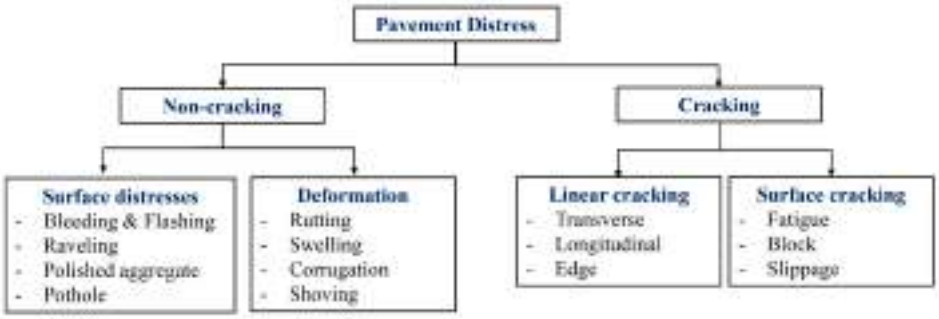


Figure 1. Pavement distress types.

Other classifications can be found in the literature. For instance, according to [5], distress can be categorised into six groups. The shapes of some common distresses are depicted in Figure 2. In addition, cracks can be subdivided based on the crack width (Table 1) [6].

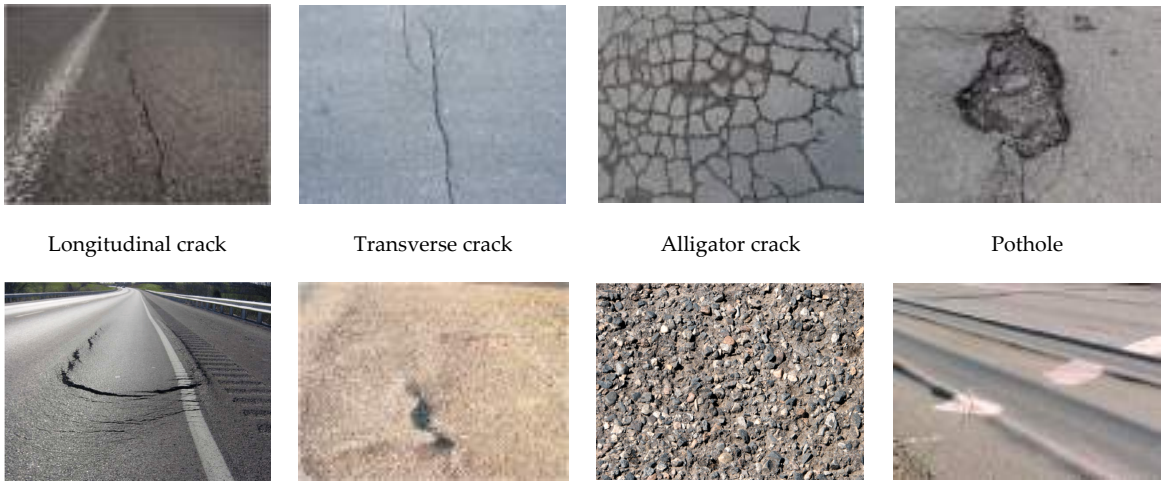


Figure 2. Examples of distress [1–7].

Table 1. Examples of distress severity (adapted from [6]).

Crack Type	Standards/Guidelines	Severity Based on Crack Width		
Longitudinal/Transverse	AASHTO R 55–10 ASTM D6433-16 FHWA LTPP	Level 1 ≤ 3 mm Low ≤ 10 mm Low ≤ 6 mm	3 mm ≤ Level 2 ≤ 6 mm 10 mm ≤ Medium ≤ 75 mm 19 mm ≤ Medium ≤ 75 mm	Level 3 > 6 mm High > 19 mm High > 75 mm
Block	AASHTO R 55–10 ASTM D6433-16 FHWA LTPP	Level 1 ≤ 3 mm Low ≤ 13 mm Low ≤ 6 mm	3 mm ≤ Level 2 ≤ 6 mm 13 mm ≤ Medium ≤ 75 mm 19 mm ≤ Medium ≤ 75 mm	Level 3 > 6 mm High > 50 mm High > 19 mm

The quality of road pavement can be assessed by applying several standards, as follows [2] (Qureshi et al., 2023):

- Pavement Surface Evaluation Rating (PASER); rating from 10 to 1 (9–10 excellent condition, 2–1 extremely poor).
- Pavement Condition Index (PCI) (ASTM 2011, ASTM 2020); rating from 100 to 0 (100–85 good condition, 10–0 completely deteriorated).
- Pavement Surface Condition Index (Irish PSCI); rating from 1 to 10.
- Road Condition Indicator (RCI).

Damage to road pavement mainly results from wear and tear, defects in the materials, or issues during the construction phases.

The pavement distress survey provides a helpful basis for suggesting corrective actions that allow for the scheduling of effective maintenance interventions with clear financial benefits for the road operator [7]. Traditional methods for pavement distress identification are complex and ineffective at dealing with large quantities of images to be inspected. Numerous automated procedures have been developed to identify pavement distress. However, currently, the most widely used and most reliable technique for evaluating flexible pavement distress requires manual or semi-automated data collection by specialised technicians [7], which is time-consuming and labour-intensive [6].

Although some phases of pavement inspection have been moderately automated through image and video data acquisition, the following main problems persist:

- Damage detection is still a difficult activity [7].
- Current automated systems are often expensive to acquire and operate, and they are not simple to use [8]. Consequently, distress is generally detected through visual inspections or manual measurement instruments [5]. Manual techniques aim to identify and classify pavement cracks on the basis of shape, dimensions, and other parameters.
- Manual procedures have several restrictions, such as modest precision, subjectivity, and inconsistency in analysis outcomes.

To overcome these and other limitations, several automated criteria have been proposed, starting with the evidence that applications of artificial intelligence (AI) are becoming common in transportation and pavement engineering. In particular, deep learning (DL) algorithms were applied in previous studies [9–12]. Recently, deep learning has shown several potential applications in many real-world domains [13], including target detection in pavement engineering. In deep learning, a computer model learns how to perform classification tasks based on different types of information, such as texts, sounds, or images. The models used in deep learning are built from a large amount of labelled data and neural network structures that contain different layers. The word 'deep' refers to the number of hidden layers in the neural network. The assessment of pavement conditions can be conducted through the utilisation of deep learning models, which can analyse pavement images and videos to evaluate the state of the pavement, encompassing such aspects as cracks, potholes, and other defects. This appears to be a beneficial approach for prioritising maintenance and repair endeavours.

Although PMS was originally designed for motorway and highway management, it can also be adapted for urban roads. Artificial intelligence techniques to identify pavement distress are increasingly used to solve the inadequacies of manual techniques (Figure 3). AI detects pavement defects quickly and in real-time; it is more efficient than manual methods, handles large-scale tasks with many defects, improves the accuracy of detection, and has high efficiency and strong scalability. On the other hand, AI detection methods can be expensive in their initial implementation, require technical expertise to be operated and maintained, may need frequent updates to be adapted to new pavement distress, and entail a risk of error if not properly trained or in unexpected conditions.

In the present article, we describe a deep learning-based technique for monitoring asphalt pavement health and decreasing the global time necessary for pavement evaluation. The detection of pavement damage is obtained using the YOLOv3 algorithm. The most recent editions of the YOLO family are YOLOv5 and YOLOv8. YOLOv8 was released in 2023. It uses a deep convolutional neural network (CNN) architecture like its predecessors but with some variations, comprising a new backbone architecture called CSPNet, a new neck architecture termed FPN+PAN, a new head architecture termed PANet, and a new training procedure. Because of its recent release, there are only a few YOLOv8 applications in the field of pavement engineering.

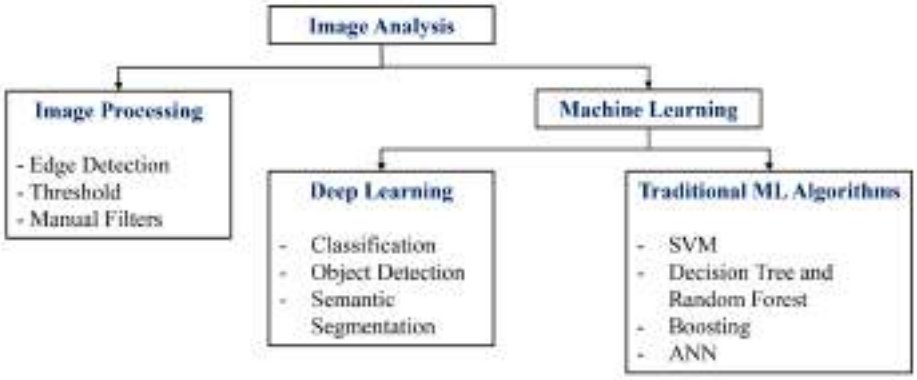


Figure 3. Image analysis method.

YOLOv3 possesses numerous advantages and disadvantages. One considerable benefit is its swiftness, thus preserving the same detection speed as in the YOLO family. Another advantage is its ability to detect small objects, which has become better than the previous versions. Nevertheless, YOLOv3 has its own limitations. One is its difficulty in handling variations in scale, particularly when confronted with exceedingly minuscule or colossal objects. In short, YOLOv3 ensures fast speed and high detection accuracy for small entities, yet it may find it difficult to deal with scale variations and certain specialised detection tasks. These characteristics allow YOLOv3 to be applied successfully for object detection. Overall, its balance between speed and precision makes it the obvious choice for various applications. Consequently, the YOLOv3 algorithm was used for detecting pavement damages in this research.

The present article explores case studies of flexible pavements on urban roads crossed by tramway lines. The proposed technique requires a simple vehicle-mounted camera system.

The experiments proved that a simple video recording device, together with the use of a deep learning-based approach, can successfully detect several types of pavement damage. The main results also showed that the proposed technique is both affordable and accurate.

The article is structured as follows: Section 2 explains the deep learning algorithms applied here. Section 3 briefly describes the known datasets for road pavement damage detection and, more specifically, those used in this research, as well as the required survey equipment. Section 4 explains the neural network training and the main outcomes in terms of loss, precision, recall, and RMSE. Section 5 illustrates the case study, results, and discussion. The main achievements, challenges, contributions, and limitations are summarised in Section 6.

2. Algorithms for Crack Detection

In recent times, deep learning (DL) has awakened great interest in several fields of highway and pavement engineering. In this regard, Mohan and Poobal [14] reviewed 50 scientific articles and collected several procedures for automated distress detection through image processing. Two-stage object detection algorithms (namely convolutional neural networks (CNNs)) have proven excellent performance on segmenting pavement cracks, but the calculation time is unreasonable for real-time applications [14].

Such a problem can be solved by adopting faster 'one-stage' (Figure 4) object detection [15,16]. One-stage models include the single-shot multibox detector (SSD) [17], the retinanet [18], and 'you only look once' (YOLO).

The YOLO ('you only look once') algorithm was created by Redmon et al. [19]. It can classify and locate objects in only one step. YOLOv2 is an enhanced edition of the original YOLO model. It implements the concept of anchor priors in single-shot multibox detectors (SSD). YOLOv2 markedly differs in the lower number of layers, with only 19 convolution layers composed of 3×3 and 1×1 filters.

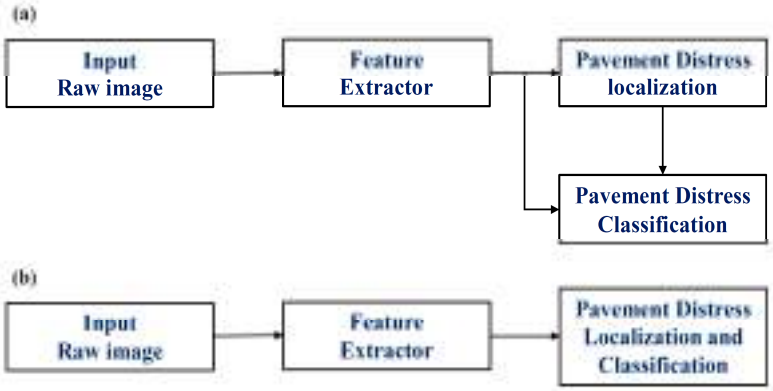


Figure 4. One-stage and two-stage pavement distress detection flows: (**a**) two-stage distress detection; (**b**) one-stage distress detection.

YOLOv3 uses the new Darknet-53 architecture based on successive 3 × 3 and 1 × 1 filters and a residual block inspired by ResNet [20,21]. It showed considerable progress in detecting small objects in real-time.

YOLOv8 is the last variant of the YOLO family, first released in May 2023. It uses a deep convolutional neural network (CNN) architecture like its predecessors but with some changes, including a new backbone architecture called CSPNet, a new neck architecture called FPN+PAN, a new head architecture called PANet, and a new training procedure.

Despite being launched recently, YOLOv8 has not yet been much used in the field of pavement engineering. Therefore, this research has applied only the YOLOv3 algorithm to detect pavement damages.

Figures 5 and 6 show the YOLOv3 architecture. For more details about this algorithm, the interested reader may consult [22,23].

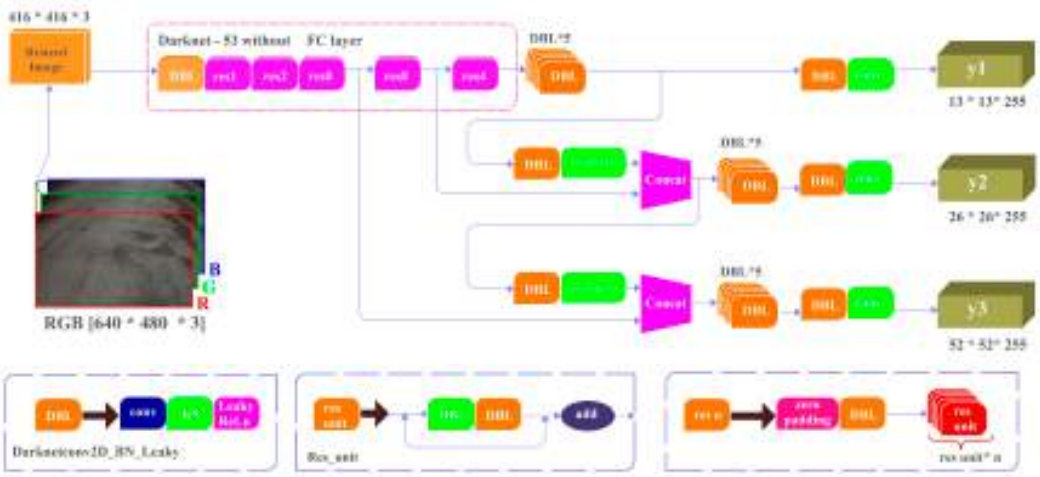


Figure 5. YOLOv3 network structure (adapted from [21]).

The image of interest is partitioned into S × S grids; each grid determines whether the centre of the focused object is located within it.

The grid evaluates B bounding boxes and the confidence of each box, i.e., the C(Ob), as follows [22]:

$$C(Ob) = P(Ob) \times IOU_{pred}^{truth} \quad (1)$$

$$P(Ob) = 1 \text{ (no target in the cell); } P(Ob) = 0 \text{ (there are targets in the cell);} \quad (2)$$

$$\mathrm{IoU}_{\mathrm{pred}}^{\mathrm{truth}} = \frac{\mathrm{B}_{\text{ground truth}} \cap \mathrm{B}_{\text{predicted}}}{\mathrm{B}_{\text{ground truth}} \cup \mathrm{B}_{\text{predicted}}} \quad (3)$$

where $\mathrm{IoU}_{\mathrm{pred}}^{\mathrm{truth}}$ indicates the intersection over the union (Figure 7).

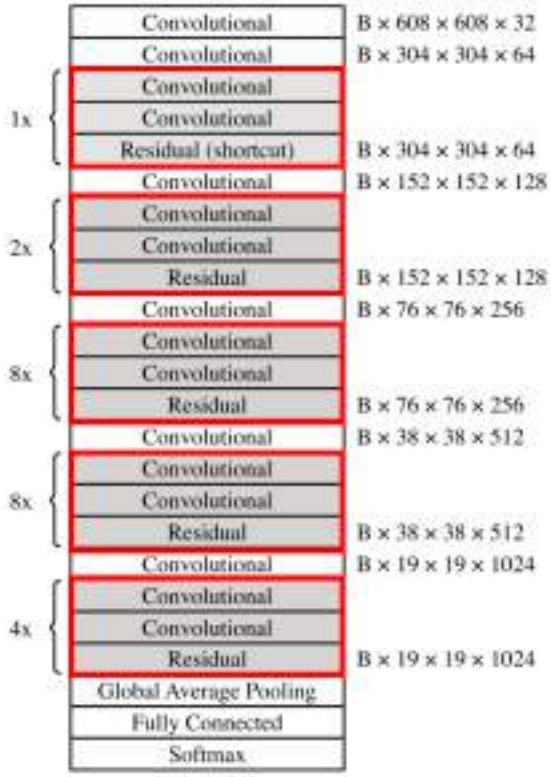


Figure 6. Darknet-53 architecture ([24]).

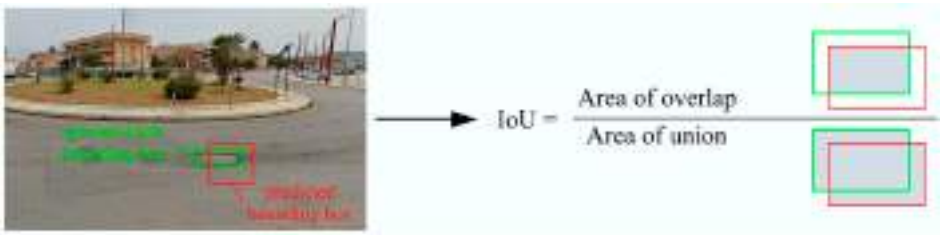


Figure 7. IoU determination.

The following variables are also calculated:

- (x, y): position of the centre of the bounding box;
- (w, h): height and width of the bounding box;
- $P(\mathrm{Class}_i \mid \mathrm{Ob})$: probability that the centre of the i-th object falls into the grid.

2.1. Detector Loss Function

The loss function comprises the following components [25]:

- Classification loss:

$$\sum_{i=0}^{S^2} \mathrm{I}_{ij}^{\mathrm{obj}} (\mathrm{p}_i(c) - \hat{\mathrm{p}}_i(c))^2 \quad (4)$$

- Localisation loss [26]:

$$\lambda_{coord}\sum_{i=0}^{S^2}\sum_{j=o}^{B} I_{ij}^{obj}[(x-\hat{x}_i)^2+(y-\hat{y}_i)^2]+\lambda_{coord}\sum_{i=0}^{S^2}\sum_{j=o}^{B} I_{ij}^{obj}[(\sqrt{\omega_i}-\sqrt{\hat{\omega}_i})^2+\sqrt{h_i}-\sqrt{\hat{h}_i})^2] \quad (5)$$

- Confidence loss [26]:

$$\sum_{i=0}^{S^2}\sum_{j=o}^{B} I_{ij}^{obj}(C-\hat{C}_i)^2 \quad (6)$$

When an object is not detected, it results in the following [25]:

$$\lambda_{noobj}\sum_{i=0}^{S^2}\sum_{j=o}^{B} I_{ij}^{noobj}(C-\hat{C}_i)^2 \quad (7)$$

Therefore, the final loss is as follows:

$$\begin{aligned}&\lambda_{coord}\sum_{i=0}^{S^2}\sum_{j=o}^{B} I_{ij}^{obj}[(x-\hat{x}_i)^2+(y-\hat{y}_i)^2]+\lambda_{coord}\sum_{i=0}^{S^2}\sum_{j=o}^{B} I_{ij}^{obj}[(\sqrt{\omega_i}-\sqrt{\hat{\omega}_i})^2+\sqrt{h_i}-\sqrt{\hat{h}_i})^2]+\\&\sum_{i=0}^{S^2}\sum_{j=o}^{B} I_{ij}^{obj}(C-\hat{C}_i)^2+\lambda_{noobj}\sum_{i=0}^{S^2}\sum_{j=o}^{B} I_{ij}^{noobj}[(C-\hat{C}_i)^2+\sum_{i=0}^{S^2} I_{ij}^{obj}(p_i(c)-\hat{p}_i(c))^2\end{aligned} \quad (8)$$

Finally, the location prediction is calculated as follows (Figure 8) [27]:

$$\begin{cases} b_x=\sigma(t_x)+c_x \\ b_y=\sigma(t_y)+c_y \end{cases} \quad (9)$$

$$\begin{cases} b_w=p_w e^{t_w} \\ b_h=p_h e^{t_h} \end{cases} \quad (10)$$

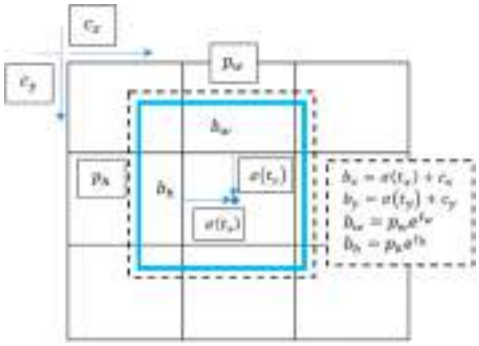


Figure 8. Location prediction in a bounding box (adapted from [28]).

2.2. Performance Metrics

The performance of the model in crack detection and classification can be calculated using the following parameters [29,30]:

$$Pr=\frac{\alpha}{\alpha+\gamma} \quad (11)$$

$$Rec=\frac{\alpha}{\alpha+\delta} \quad (12)$$

$$Acc=\frac{\alpha+\beta}{\alpha+\beta+\gamma+\delta} \quad (13)$$

$$F1 = \frac{2}{\frac{1}{Pr} + \frac{1}{Rec}} = \frac{2\alpha}{2\alpha + \gamma + \delta} \quad (14)$$

Symbol definitions are summarised in Table 2 [29].

Table 2. Meaning of the symbols in Equations (11)–(14).

Symbol	Description
Pr	precision
Rec	recall
Acc	accuracy
α	true positive
β	true negative
γ	false positive
δ	false negative

Finally, RMSE is obtained with the following relationship:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i}^{n}(f_i - y_i)^2} \quad (15)$$

RMSE allows the error to be assessed between the ground-truth distress numbers y_i and the predicted distress numbers f_i.

3. Survey Equipment

Nowadays, a lot of data sources for flexible pavement distress are openly available to be used, and the most popular among them are [23,31–35] (Table 3): CFD dataset, AigleRN dataset, CRACK500 dataset, GAPs dataset, CrackTree200 dataset, Road Damage dataset 2018, and Road Damage dataset 2019.

Table 3. Main properties of some public pavement distress datasets.

Dataset	Task Type	Distress	Device	N° of Images	Resolution
Aigle-RN	segmentation	crack	professional camera	38	991 × 462, 311 × 462
CFD	segmentation	crack	smartphone	118	480 × 320
Crack500	segmentation	crack	LG-H345	500	2000 × 1500
Road surface damage	-	-	-	18,345	-
Pavement image dataset	-	-	-	7237	-
GAPs v1	object detection	severe distress	professional camera	1969	1920 × 1080
GaMM	segmentation	crack	professional camera	42	1920 × 480
Cracktree200	segmentation	crack	-	206	800 × 600
CrackIT	segmentation	crack	optical	84	1536 × 2048
EdmCrack600	segmentation	crack	GoPro 7	600	1920 × 1080
GAPs v2	object detection	several distress	-	2468	1920 × 1080
Road damage dataset 2018	object detection	several distress	smartphone	9053	600 × 600
Road damage dataset 2019	object detection	several distress	smartphone	13,135	600 × 600
CQU-BPDD	-	-	-	60,059	-

The model was trained on the Road Damage dataset. Its initial version was made available in 2018, while the most recent version was released in 2019 [33]. Compared to the first version, in the Road Damage dataset 2019, the total number of annotated images increased from 9053 to 13,135 and the number of annotations increased from 15,435 to 30,989. Table 4 lists the damage categories, their definitions, and their class names.

Table 4. Classes of pavement distress.

Damage Description	Flexible Pavement	Class Name
Longitudinal linear crack, wheel mark part	✓	D00
Lateral linear crack, equal interval	✓	D10
Alligator crack	✓	D20
Bump, rutting, separation, pothole		D40
Crosswalk blur	✓	D43
White line blur	✓	D44
Manhole	✓	D50

Therefore, the present study takes 13,135 images of pavement cracks into consideration; the ratio of training, validation, and testing sets was set at 7:1.5:1.5 (i.e., 9195, 1970, and 1970 images, respectively).

Survey Vehicle

In the experimental phase, the input data for pavement distress detection and classification are videos of flexible pavements on urban roads in the city of Palermo. The videos are obtained through a camera installed on the rear windscreen of a car (Figure 9) by means of a gripper suction device. In order to reduce the motion blur due to vehicle speed and road conditions, the equipment was only experimented with at speeds $v \leq 50$ km/h (the maximum speed limit of urban roads in Palermo). In addition, a proper camera angle of around forty-five degrees with respect to the road pavement surface was adopted. The detected images were 640 × 480 pixels in size, which corresponds to a ground truth measurement area of about 1080 mm × 1447 mm. The accuracy of crack detection was ensured by a strong camera calibration performed as shown in Figure 10. More details and an example of this process can be found in [23].

Figure 9. Survey vehicle.

The first step of this research was the camera calibration obtained from Zhang's algorithm [35]. The calibration was then performed by using several images of a chessboard in the outdoor environment (Figure 10). The extrinsic parameters estimated by the calibration process are depicted in Figure 11. Finally, the model was validated by comparing the real and predicted distances of some objects [36]. Figure 12 shows an example of distress annotation [34].

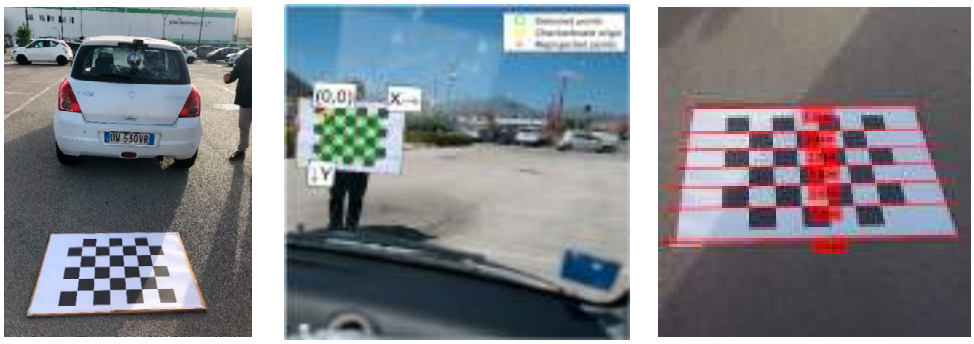


Figure 10. Some phases of camera calibration.

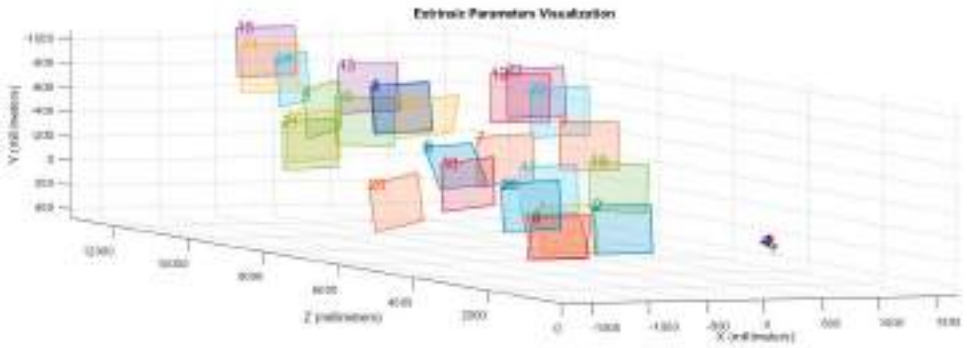


Figure 11. Camera parameter representation.

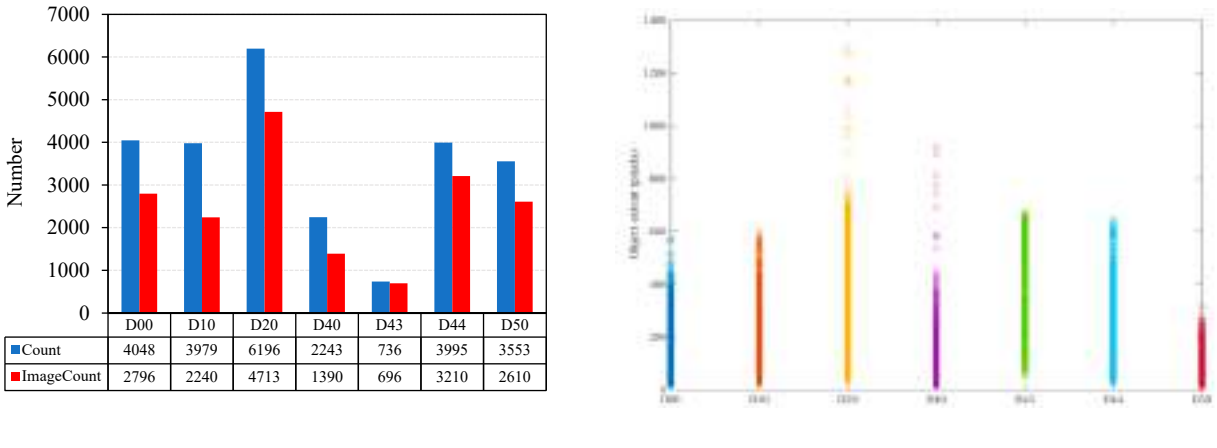
Figure 12. An example of distress annotation.

4. Neural Network Training

The crucial elements for an accurate model are the image characteristics, in terms of quantity and quality, which are selected for the training. In this study, the Road Damage dataset 2019 was used (Table 4). Therefore, the pre-trained model was able to identify such types of distress in the pavement surface as those given in Table 4.

The initial learning rate was fixed to 0.001, and the 'rmsprop' algorithm was adopted. RMSprop is a popular optimisation algorithm used in machine learning. It is designed to improve the speed of convergence and find the minimum of the loss function quickly. In addition, the following parameters were fixed: minimum batch size 8, total epochs 20 (38,840 iterations). For the training dataset, Figure 13a summarises the total number of bounding boxes for each damage class and the total images in which a given damage was detected. As expected, the bounding boxes have relatively small variance in the pixel dimensions for the damage types ID00, ID10, ID43, and ID44 and high variations for ID20 and ID40, as shown in Figure 13b.

Figures 14–17 represent the base learning rate, loss, RMSE, and precision–recall curves as a function of the iterations obtained in the training process. It is immediately clear how the model can detect flexible pavement damages accurately once around 8000 iterations are reached. The precision–recall curves demonstrate that the average precision ranges from 0.5 to 0.8, depending on the distress type (Figure 17).



(**a**) (**b**)

Figure 13. (**a**) Sample of images used for different types of damage; (**b**) object size variance across classes.

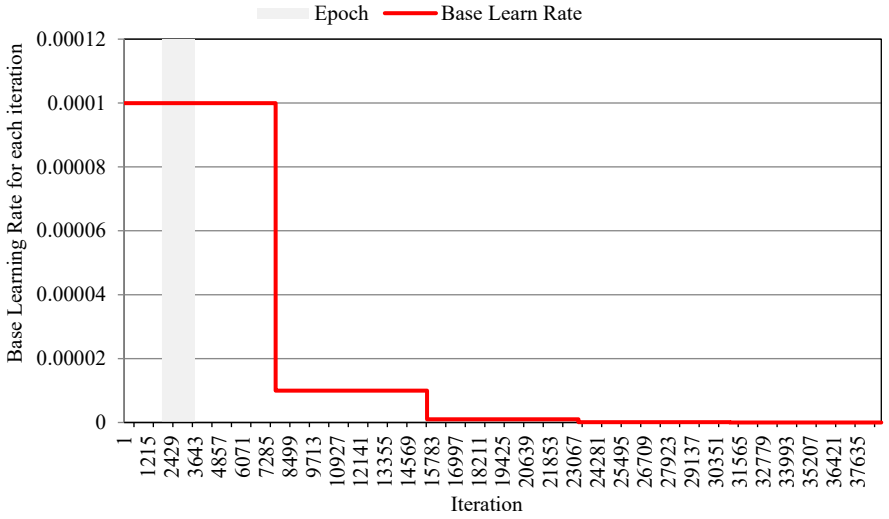


Figure 14. Base learning rate related to the number of iterations.

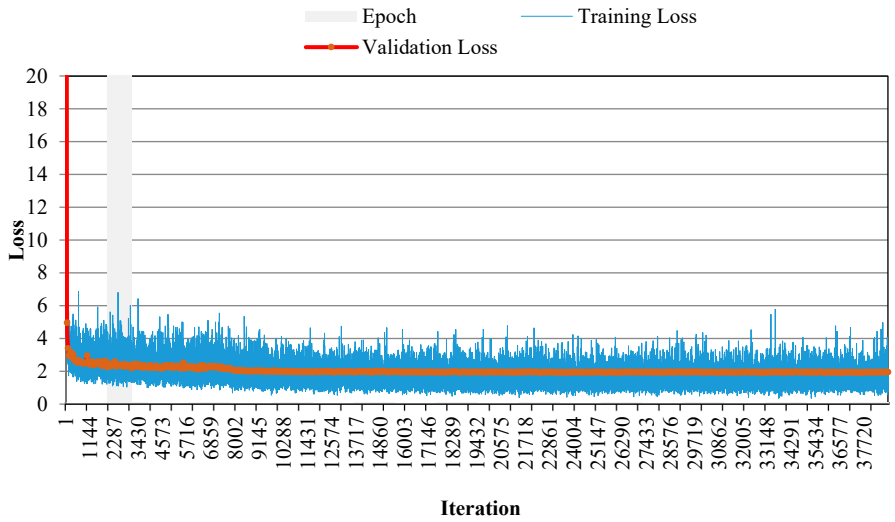


Figure 15. Loss related to the number of iterations.

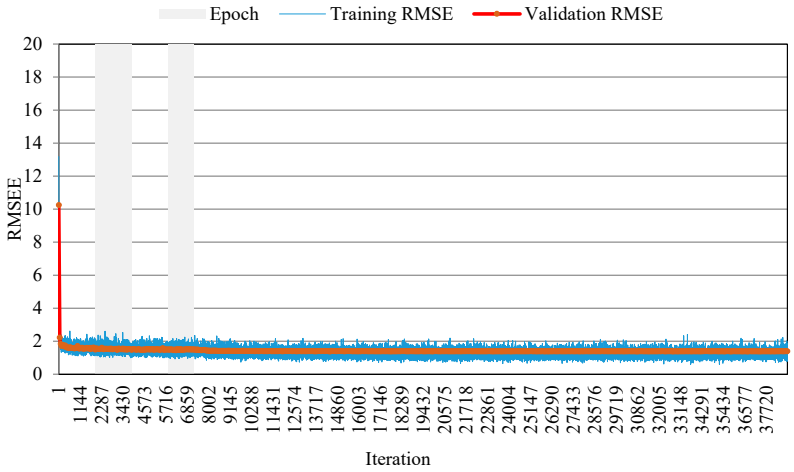


Figure 16. RMSE related to the number of iterations.

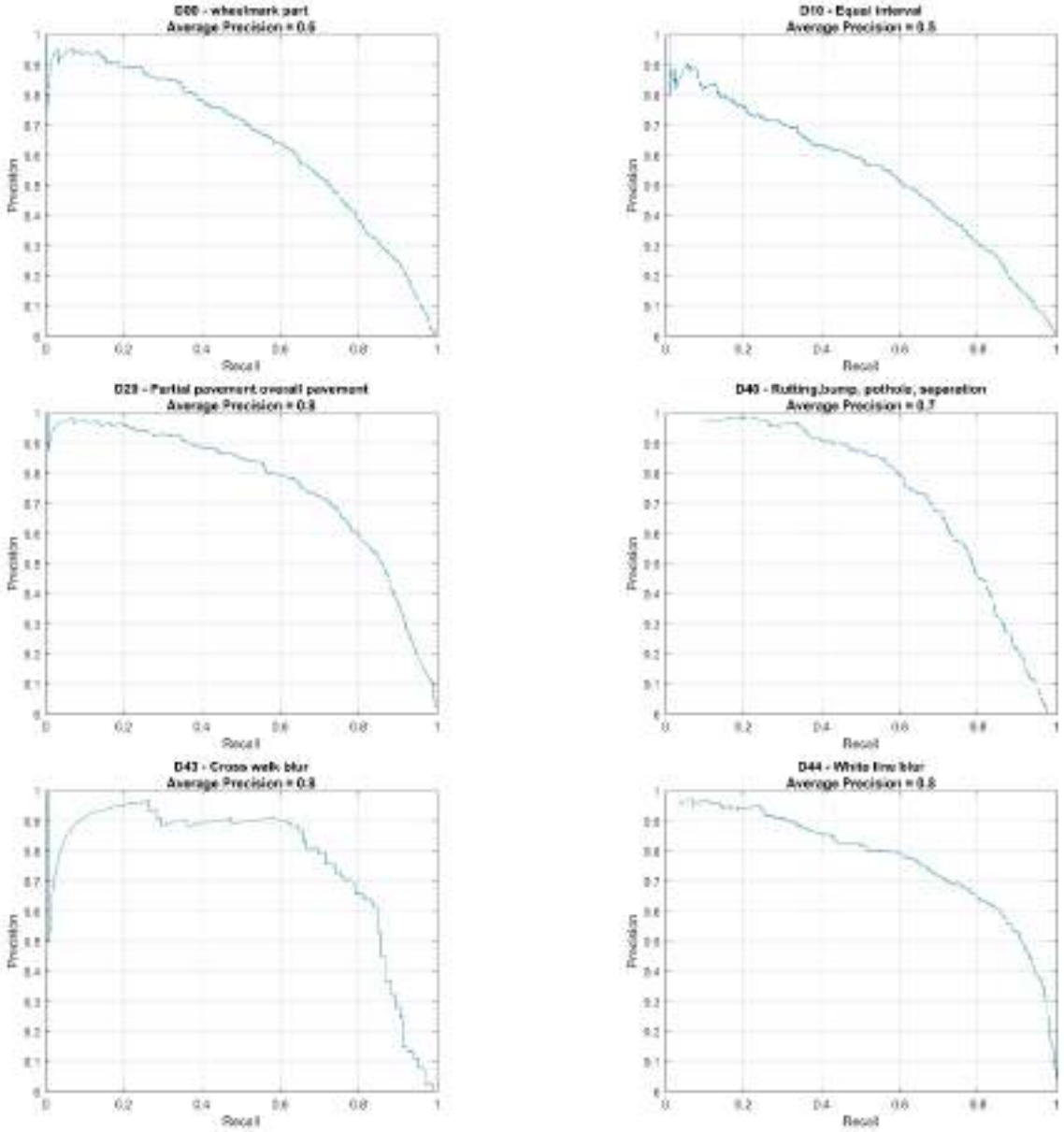


Figure 17. Precision–recall curves for six types of pavement damage (cf. Table 4).

For a more in-depth evaluation of the flexible pavement distress classification results, the confusion matrix was calculated (Figure 18). In deep learning applications, the confusion matrix is used for a demonstration of the classification model's performance. In this matrix, the columns show the predicted values, while the rows show the actual values. The cell where the row and column for a certain pavement distress class intersect indicates the true positive values for that class. We can observe that the pavement distress class with the lowest sensitivity in the validation data is D43, with a 86.8% sensitivity. On the contrary, the class with the highest sensitivity is D20, with a 98.7% sensitivity.

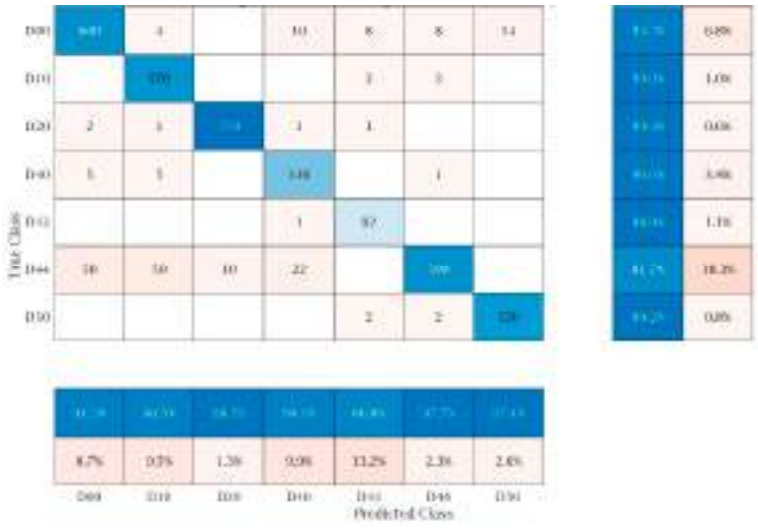


Figure 18. Confusion matrix (zero values omitted for clarity). The diagonal shows the true positive values for each class (i.e., those labelled and classified as that class).

Distress Tracking and Surface Evaluation

The correspondence between a point in an image and its projection on a 2D image is determined with the use of a geometric model. The real dimension of several distresses can be estimated by the inverse perspective mapping (IPM) method [35–37]. The IPM method eliminates the perspective effect in images by converting them to a bird's-eye view. This method corrects image distortion caused by tilt through a mathematical transformation deriving from the vanishing point, image plane, and slope. Multiple transformed images are stitched together to create a panoramic image. Thanks to the IPM, a top-down view of the damage to road pavements can be obtained (Figure 19) by means of the following equations:

$$ {}^{g}_{i}T = h \begin{bmatrix} -\frac{1}{f_u}c_2 & -\frac{1}{f_v}s_1s_2 & \frac{1}{f_u}c_uc_2 - \frac{1}{f_v}c_vs_1s_2 - c_1s_2 & 0 \\ \frac{1}{f_u}s_2 & \frac{1}{f_v}s_1c_1 & -\frac{1}{f_u}c_us_2 - \frac{1}{f_v}c_vs_1c_2 - c_1c_2 & 0 \\ 0 & \frac{1}{f_v}c_1 & -\frac{1}{f_v}c_vc_1 + s_1 & 0 \\ 0 & -\frac{1}{hf_v}c_1 & \frac{1}{hf_v}c_vc_1 - \frac{1}{h}s_1 & 0 \end{bmatrix} \quad (16) $$

$$ {}^{i}_{g}T = \begin{bmatrix} f_uc_2 + c_uc_1s_2 & c_uc_1c_2 - s_2f_u & -c_us_1 & 0 \\ s_2(c_vc_1 - f_vs_1) & c_2(c_vc_1 - f_vs_1) & f_vc_1 - c_vs_1 & 0 \\ c_1s_2 & c_1c_2 & -s_1 & 0 \\ c_1s_2 & c_1c_2 & -s_1 & 0 \end{bmatrix} \quad (17) $$

where the projection on the pavement surface of the generic point is denoted with ${}^{i}P = \{u, v, 1, 1\}$ and the point placed on the road pavement surface is denoted with ${}^{g}P = \{x_g, y_g, -h, 1\}$.

A bird's-eye view allows us to project the coordinates of each distress from the input image onto the pavement surface and then determine the information of interest (e.g., the length or surface of each distress type). Therefore, the proposed technique can automatically detect asphalt pavement distress coordinates starting from video recording, in that it implements a specific procedure for tracking pavement damage present in subsequent

frames. The proposed tracking algorithm is divided into the phases illustrated in Figure 20. During detection, the noise was reduced by applying the linear Kalman filter. The use of this filter is necessary to estimate the coordinates of the points on the perimeter of any damage present on the road pavement. It is a recursive filter [38] that estimates the state of a dynamic system of relationships [39,40] as follows:

$$x_{n+1} = A_n x_n + B_n u_n \tag{18}$$

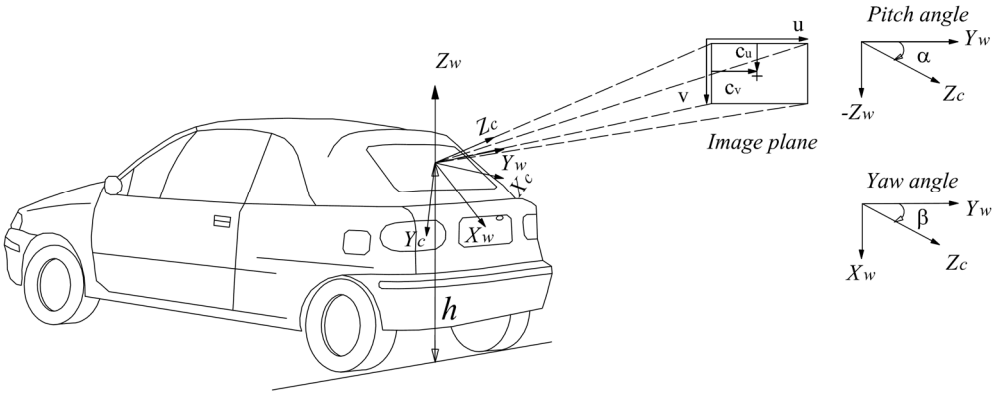


Figure 19. The IPM method used in this research.

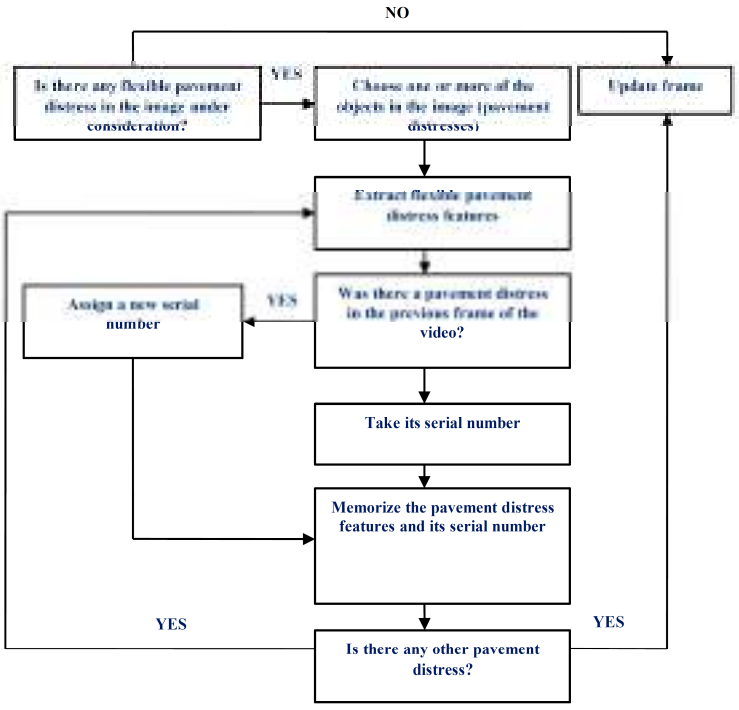


Figure 20. Algorithm for tracking pavement distress.

Taking the error covariance [40,41] into consideration, as follows:

$$P_{n+1} = A_n P_n A_n{}^T + Q_n \tag{19}$$

where x_n denotes the state value at phase n, A_n denotes the state transition matrix, and u_n is the measurement and the input at phase n. Q_n is the white noise covariance [41]. This is the 'prediction step' because it estimates the n + 1 state. Kalman gain value is obtained with the following relation [41]:

$$K_n = P_n C^T \left(C P_n C^T + R_n \right)^{-1} \tag{20}$$

in which C denotes the measurement matrix and R is the measurement noise.

The actual measurement value is:

$$P_n = (I - K_n H)P_n \quad (21)$$

where K_n and H are the measurement value and the mapping matrix from the true state, respectively.

The combination of the IPM and tracking algorithm procedures allows us to determine the type and area of the surface for each damage. These data are essential for estimating numerous performance factors (e.g., PCI, RQI, RDI, and SRI [42–44]).

5. The Case Study: Results and Discussions

The case study concerns several urban road sections in Palermo. We selected only roads crossed by tramway lines. The tramway transportation system of the city was opened on 30 Dec 2015 and comprises four lines for a total of 23.3 km. According to Bieberc's classifications [45,46], some sections of these tramway lines belong to class E (common corridor) and others to class B (exclusive protected corridor). Figure 21 depicts the tramway track details for both ballasted tracks and slab tracks. Figure 22 shows the construction phases of the tramway slab track in Palermo. Finally, Figures 23 and 24 illustrate a few images of some tramway lines in operation.

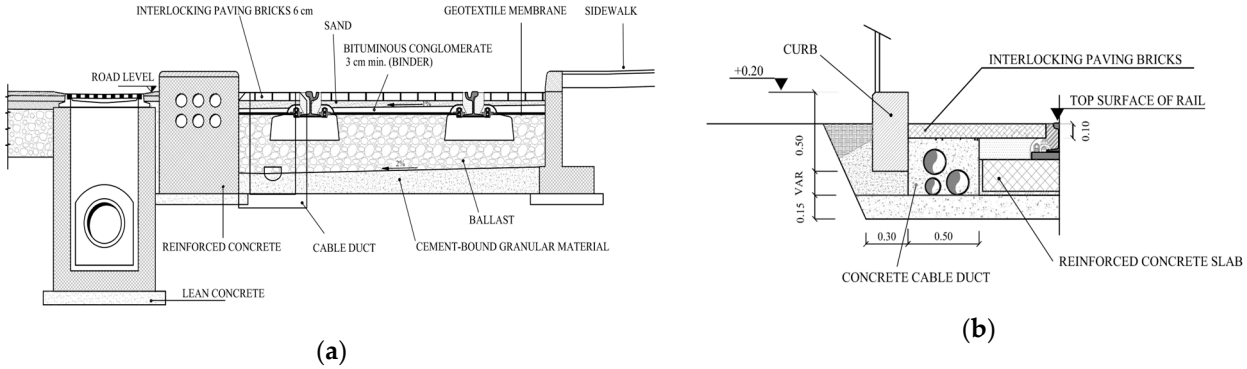


Figure 21. Tramway track details: (**a**) ballasted track and (**b**) slab track.

Figure 22. Construction phases of the tramway track on the analysed urban roads. (**a**) casting of a lean concrete layer and curb construction; (**b**) laying of steel reinforcement bars; (**c**) laying of the flat framework; (**d**) casting of the concrete slab and paving.

Figure 23. Photos of the analysed road pavement (via Antonio Laudicina)—class E tramway line.

Figure 24. Photos of the analysed road pavement (via Carlo Gulì)—class B tramway line.

In this research, the proposed technique was used for evaluating distress in lateral lanes adjacent to the tramway lines [47].

The inspection videos were obtained by the rear camera outside the test vehicle (Figure 9). As demonstrated before, the loss and the RMSE values obtained in the training phase of the neural network prove that the model can detect flexible pavement distresses with high precision [48–50]. Some test outcomes are shown in Figure 25. The proposed method can accurately detect, classify, and measure different pavement cracks and other damage types in the recorded videos.

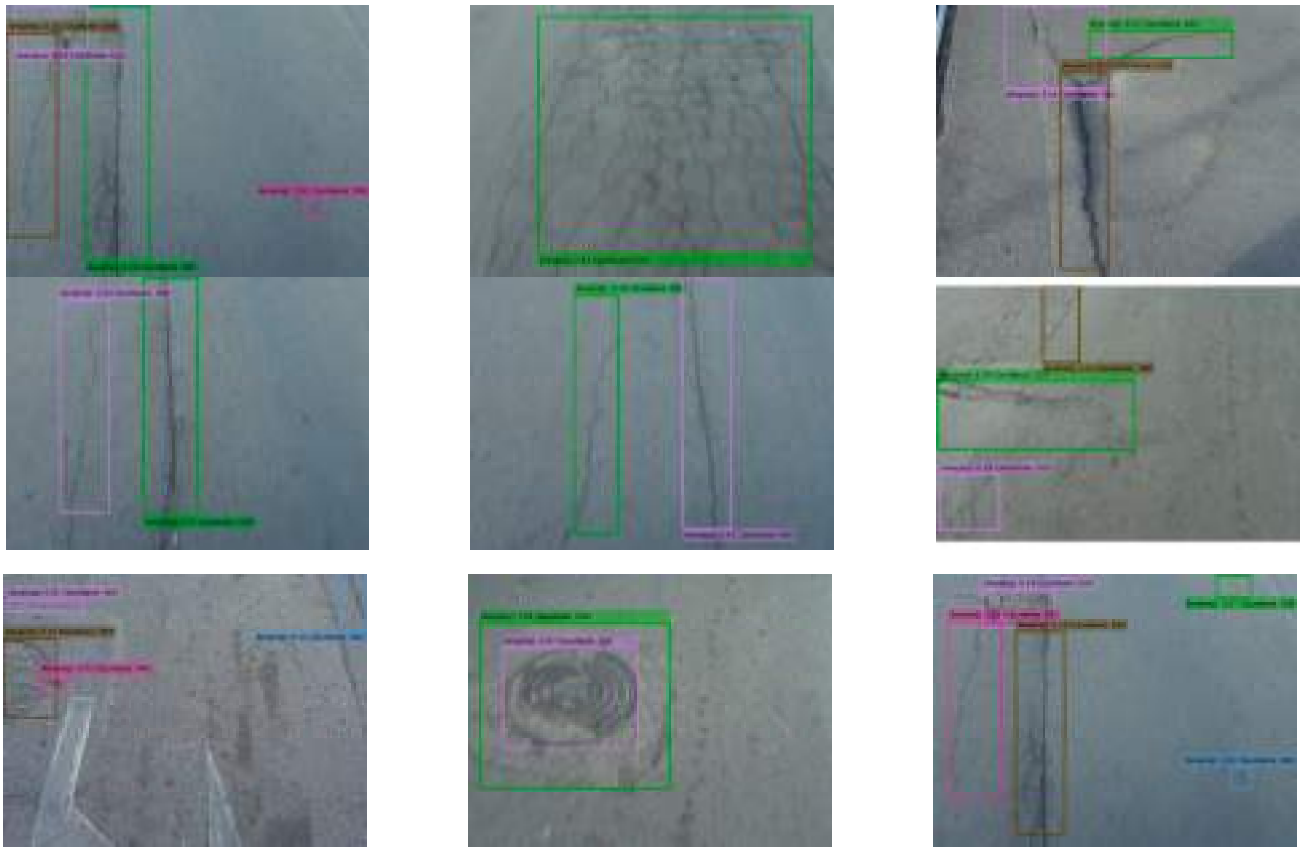
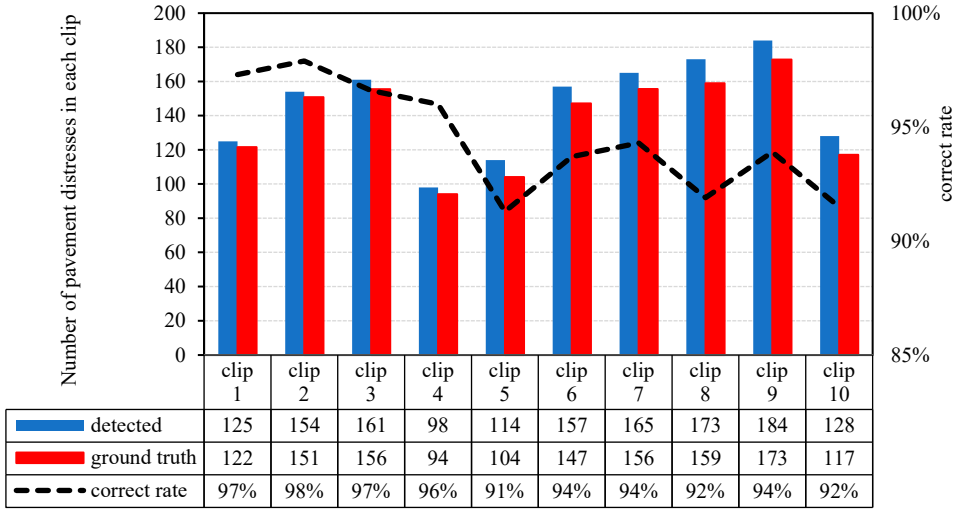


Figure 25. Examples of damage detection and surface estimation.

The procedure described in the previous sections was applied to carry out an error analysis in the case study. Numerous images extracted from 10 video clips were analysed, and damages were hand-labelled.

A total of 459 labelled frames contain 1378 boundary boxes of damages. The ID50 class is not taken into consideration because it refers to manholes (cf. Table 4), which are not considered damage. A comparison was made between the observed and real distress.

The outcomes summarised in Figure 26 demonstrate the correct detection rate to range from 91% to 98% in distress detection.



	clip 1	clip 2	clip 3	clip 4	clip 5	clip 6	clip 7	clip 8	clip 9	clip 10
detected	125	154	161	98	114	157	165	173	184	128
ground truth	122	151	156	94	104	147	156	159	173	117
correct rate	97%	98%	97%	96%	91%	94%	94%	92%	94%	92%

Figure 26. The correct detection rate for a sample of ten clips.

Therefore, the empirical analyses demonstrate that the algorithm used is sufficiently precise in detecting and measuring pavement damages. This is true, although some errors are inevitable in the detection phase, especially due to the irregularities in the pavement, which generate oscillations in the vehicle and therefore in the camera [51–54].

To summarise, the practical importance of the proposed approach is related to the fact that the timely identification of damages helps the city administration or the road operator in the decision-making process to choose the appropriate technology for pavement detection and repair [55]. On the other hand, automated pavement monitoring makes it possible to choose in advance an environmentally friendly technology (in terms of LCA) for the maintenance phase, resulting in a significant social benefit.

6. Conclusions

The accurate detection and classification of road pavement distress need to acquire three-dimensional depth pavement pictures, but this process requires expensive dedicated survey vehicles, sensors, and other devices. Therefore, alternative techniques should be considered to analyse road pavement carefully and at a low cost.

This research presents cost-effective equipment for asphalt pavement surface condition evaluation that has been applied for assessing urban road surface condition. The potential benefits of this equipment were evaluated by means of experiments on several urban road sections in the city of Palermo (Italy). Such road sections crossed by tramway lines were analysed because of their numerous distress types, extensions, and spreading. The procedure is founded on examining videos of asphalt pavement taken by a test car equipped with a rear camera and then applying the YOLOv3 algorithm for the detection and measurement of different damage categories. A dataset with about 13,135 images and 30,989 bounding boxes was used. Damages were classified into seven different types. The outcomes of the analyses prove that the pre-trained network can classify distress into several categories, i.e., longitudinal, lateral, and alligator cracks, rutting, bumps, potholes, etc. The loss and the RMSE guarantee that the algorithm can detect damages with very good accuracy and speed. In addition, as shown by the confusion matrix calculated during the validation process, the pavement distress classification sensitivity reaches up to 98.7%.

The experiments showed that a simple videorecording device, together with the implementation of a deep learning-based procedure, can successfully assess pavement quality, detect several damage types, and meet the real-time requirements.

Even though the method needs to be validated with additional experiments in other road conditions, sample data from 10 video clips from this research proves that the correct detection rate ranges from 91% to 98%.

However, there are still limitations that require future studies for the following main reasons:

- Due to vehicle vibrations and visibility conditions, the algorithm is unable to identify some pavement distress.
- Use of a public dataset.
- In this research, YOLOv3 was applied even though it is a less-performing version of the YOLO family (e.g., YOLOv5 and YOLOv8). Despite this choice, it is possible to obtain excellent results in both the detection and classification of road surface damage using low-cost detection devices. Therefore, the application of deep learning algorithms in pavement engineering can produce enormous benefits even with less-performing versions of YOLO and cost-effective equipment.

Future studies may be dedicated to the application of a new generation of YOLO algorithms (i.e., YOLOv8 or later) and more expensive and higher-resolution cameras. Finally, it is worth highlighting that with the rapid development of autonomous vehicles, which will be equipped with numerous cameras and other sensors, the proposed technique will likely provide details about the state of flexible pavements and useful information for road operators.

Author Contributions: Conceptualisation, M.G. and G.P.; methodology, M.G. and G.P.; software, G.P.; validation, M.G.; investigation, M.G. and G.P.; resources, M.G.; data curation, G.P.; writing—original draft preparation, M.G.; writing—review and editing, M.G.; visualisation, M.K. and L.N.; supervision, M.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All data used in this research can be provided upon request.

Acknowledgments: The first author acknowledges the Italian Ministry of Universities and Research (MUR) in the framework of the project DICAM–EXC (Departments of Excellence 2023–2027, grant No. L232/2016).

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

References

1. Donev, V.; Hoffmann, M. Optimisation of pavement maintenance and rehabilitation activities, timing and work zones for short survey sections and multiple distress types. *Int. J. Pavement Eng.* **2020**, *21*, 583–607. [CrossRef]
2. Qureshi, W.S.; Power, D.; Ullah, I.; McKeever, S.; O'Sullivan, D. Deep learning framework for intelligent pavement condition rating: A direct classification approach for regional and local roads. *Autom. Constr.* **2023**, *153*, 104945. [CrossRef]
3. Chu, C.; Wang, L.; Xiong, H. A review on pavement distress and structural defects detection and quantification technologies using imaging approaches. *J. Traffic Transp. Eng.* **2022**, *9*, 135–150. [CrossRef]
4. Miller, J.S.; Bellinger, W.Y. Distress Identification Manual for the Long-Term Pavement Performance Program, Georgetown Pike. 2014. Available online: https://highways.dot.gov/sites/fhwa.dot.gov/files/docs/research/long-term-pavement-performance/products/1401/distress-identification-manual-13092.pdf (accessed on 1 January 2024).
5. Baduge, S.K.; Thilakarathna, S.; Perera, J.S.; Ruwanpathirana, G.P.; Doyle, L.; Duckett, M.; Lee, J.; Saenda, J.; Mendis, P. Assessment of crack severity of asphalt pavements using deep learning algorithms and geospatial system. *Constr. Build. Mater.* **2023**, *401*, 132684. [CrossRef]
6. Miller, J.S.; Bellinger, W.Y. *Distress Identification Manual*; FHWA-HRT-13-092; U.S. Department of Transportation: Washington, DC, USA, 2014.
7. Apeagyei, A.; Ademolake, T.E.; Adom-Asamoah, M. Evaluation of deep learning models for classification of asphalt pavement distresses. *Int. J. Pavement Eng.* **2023**, *24*, 2180641. [CrossRef]
8. Vavrik, W.; Evans, L.; Sargand, S.; Stefanski, J. *PCR Evaluation—Consider Transition from Manual to Semi-Automated Pavement Distress Collection and Analysis*; Ohio Department of Transportation Office of Statewide Planning and Research: Columbus, OH, USA, 2013.
9. Zhang, A.; Wang, K.C.P.; Fei, Y.; Liu, Y.; Tao, S.; Chen, C.; Li, J.Q.; Li, B. Deep Learning-Based Fully Automated Pavement Crack Detection on 3D Asphalt Surfaces with an Improved CrackNet. *J. Comput. Civ. Eng.* **2018**, *32*, 04018041. [CrossRef]

10. Zhang, A.; Wang, K.C.P.; Fei, Y.; Liu, Y.; Chen, C.; Yang, G.; Li, J.Q.; Yang, E.; Qiu, S. Automated Pixel-Level Pavement Crack Detection on 3D Asphalt Surfaces with a Recurrent Neural Network. *Comput. Aided Civ. Infrastruct. Eng.* **2019**, *34*, 213–229. [CrossRef]
11. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder- Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef]
12. Liu, Z.; Gu, X.; Yang, H.; Wang, L.; Chen, Y.; Wang, D. Novel YOLOv3 Model with Structure and Hyperparameter Optimization for Detection of Pavement Concealed Cracks in GPR Images. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 22258–22268. [CrossRef]
13. Sha, J.; Wang, J.; Hu, H.; Ye, Y.; Xu, G. Development of an Accurate and Automated Quality Inspection System for Solder Joints on Aviation Plugs Using Fine-Tuned YOLOv5 Models. *Appl. Sci.* **2023**, *13*, 5290. [CrossRef]
14. Mohan, A.; Poobal, S. Crack detection using image processing: A critical review and analysis. *Alex. Eng. J.* **2018**, *57*, 787–798. [CrossRef]
15. Qu, Z.; Chen, W.; Wang, S.-Y.; Yi, T.-M.; Liu, L. A crack detection algorithm for concrete pavement based on attention mechanism and multi-features fusion. *IEEE Trans. Intell. Transp. Syst.* **2021**, *23*, 11710–11719. [CrossRef]
16. Wang, T.; Fangfang, Y.; Tsui, K.L. Real-Time Detection of Railway Track Component via One-Stage Deep Learning Networks. *Sensors* **2020**, *20*, 4325. [CrossRef]
17. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. In Proceedings of the Advances in Neural Information Processing Systems 28 (NIPS 2015), Montreal, QC, Canada, 7–12 December 2015; Volume 28, pp. 91–99.
18. Shang, L.; Yang, Q.; Wang, J.; Li, S.; Lei, W. Detection of rail surface defects based on CNN image recognition and classification. In Proceedings of the 20th International Conference on Advanced Communication Technology (ICACT), Chuncheon-si, Republic of Korea, 11–14 February 2018; pp. 45–51.
19. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
20. Ge, L.; Dan, D.; Li, H. An accurate and robust monitoring method of full-bridge traffic load distribution based on YOLO-v3 machine vision. *Struct. Control Health Monit.* **2020**, *27*, e2636. [CrossRef]
21. Liu, L.; Zhou, F.; He, Y. Vision-based fault inspection of small mechanical components for train safety. *IET Intell. Transp Syst.* **2016**, *10*, 130–139. [CrossRef]
22. Jin, Z.-Z.; Zheng, Y.-F. Research on application of improved YOLO V3 algorithm in road target detection. *J. Phys. Conf. Ser.* **2020**, *654*, 012060.
23. Guerrieri, M.; Parla, G. Flexible and stone pavements distress detection and measurement by deep learning and low-cost detection devices. *Eng. Fail. Anal.* **2022**, *141*, 106714. [CrossRef]
24. Redmon, J.; Farhadi, A. Yolov3: An incremental improvement. *arXiv* **2018**, arXiv:abs/1804.02767.
25. Hui, J. Real-time Object Detection with YOLO, YOLOv2 and Now YOLOv3, 2018. Available online: https://jonathan-hui.medium.com/real-time-object-detection-with-yolo-yolov2-28b1b93e2088 (accessed on 1 January 2021).
26. Lechgar, H.; Bekkar, H.; Rhinane, H. Detection of cities vehicle fleet using YOLO V2 and aerial images, International Archives of the Photogrammetry. *Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2019**, *42*, 121–126.
27. Li, Y.; Han, Z.; Xu, H.; Li, X.; Zhang, K. YOLOv3-lite: A lightweight crack detection network for aircraft structure based on depthwise separable convolutions. *Appl. Sci.* **2019**, *9*, 3781. [CrossRef]
28. Dewi, C.; Chen, R.-C.; Yu, H. Weight analysis for various prohibitory sign detection and recognition using deep learning. *Multimed. Tools Appl.* **2020**, *79*, 32897–32915. [CrossRef]
29. Hu, G.X.; Hu, B.L.; Yang, Z.; Huang, L.; Li, P. Pavement Crack Detection Method Based on Deep Learning Models. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 5573590. [CrossRef]
30. Xiao, X.; Cao, S.; Wang, L.; Cheng, S.; Yuan, E. Deep hashing image retrieval based on hybrid neural network and optimized metric learning. *Knowledge-Based Systems* **2024**, *284*, 111336. [CrossRef]
31. Ranyal, E.; Sadhu, A.; Jain, K. Road Condition Monitoring Using Smart Sensing and Artificial Intelligence: A Review. *Sensors* **2022**, *22*, 3044. Available online: https://etd.ohiolink.edu/apexprod/rws_etd/send_file/send?accession=osu1523978995395614&disposition=inline (accessed on 10 August 2021). [CrossRef] [PubMed]
32. Kanaeva, I.A.; Ivanova, J.A. Road Pavement Crack Detection Using Deep Learning with Synthetic Data. *Inf. Commun. Technol.* **2019**, *1019*, 012036. [CrossRef]
33. Maeda, H.; Kashiyama, T.; Sekimoto, Y.; Seto, T.; Omata, H. Generative adversarial network for road damage detection. *Comput. Aided Civ. Infrastruct. Eng.* **2021**, *36*, 47–60. [CrossRef]
34. García-Segura, T.; Montalbán-Domingo, L.; Llopis-Castelló, D.; Sanz-Benlloch, A.; Pellicer, E. Integration of deep learning techniques and sustainability-based concepts into an urban pavement management system. *Constr. Build. Mater.* **2023**, *401*, 132684. [CrossRef]
35. Zhang, Z. A flexible new technique for camera calibration. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*; IEEE: Piscataway, NJ, USA, 2002; Volume 22, pp. 1330–1334.
36. Bi, F.; Yang, J. Target Detection System Design and FPGA Implementation Based on YOLO v2 Algorithm. In Proceedings of the International Conference on Imaging Signal Processing and Communication, ICISPC, Singapore, 27–29 July 2019; pp. 10–14.

37. Dorj, B.; Lee, D.J. A Precise Lane Detection Algorithm Based on Top View Image Transformation and Least-Square Approaches. *J. Sens.* **2016**, *2016*, 4058093. [CrossRef]
38. Aly, M. Real time Detection of Lane Markers in Urban Streets. In Proceedings of the IEEE Intelligent Vehicles Symposium, Proceedings, Eindhoven, The Netherlands, 4–6 June 2008; pp. 7–12.
39. Kalman, E. A new approach to linear filtering and predictions problems. *J. Basic Eng.* **1960**, *82*, 35–45. [CrossRef]
40. Welch, G.; Bishop, G. An Introduction to the Kalman Filter. Department of Computer Science University of North Carolina at Chapel Hill Chapel Hill, NC 27599-3175. 2006. Available online: https://www.cs.unc.edu/~welch/media/pdf/kalman_intro.pdf (accessed on 1 August 2021).
41. Niu, M. Object Detection and Tracking for Autonomous Driving by MATLAB Toolbox. Ph.D. Thesis, Ohio State University, Columbus, OH, USA, 2018.
42. *ASTMD6433-09*; Standard Practice for Roads and Parking Lots Pavement Condition Index Surveys. ASTM: West Conshohocken, PA, USA, 2011. Available online: https://www.astm.org/d6433-09.html (accessed on 1 September 2023).
43. ASTM. Standard Test Method for Airport Pavement Condition Index Surveys. In *Book of Standards*; ASTM: West Conshohocken, PA, USA, 2020; Volume 04.03.
44. Li, J.; Yin, G.; Wang, X.; Yan, W. Automated decision making in highway pavement preventive maintenance based on deep learning. *Autom. Constr.* **2022**, *135*, 104111. [CrossRef]
45. Guerrieri, M. Catenary-Free Tramway Systems: Functional and Cost–Benefit Analysis for a Metropolitan Area. *Urban. Rail. Transit.* **2019**, *5*, 289–309. [CrossRef]
46. Bieberc, C.A. Les choix techniques pour les transport collectifs. *Ec. Natl. Ponts Chaussèes Paris* **1986**.
47. Guerrieri, M. Tramways in Urban Areas: An Overview on Safety at Road Intersections. *Urban. Rail. Transit.* **2018**, *4*, 223–2331. [CrossRef]
48. Wang, D.; Liu, Z.; Gu, X.; Wu, W.; Chen, Y.; Wang, L. Automatic Detection of Pothole Distress in Asphalt Pavement Using Improved Convolutional Neural Networks. *Remote Sens.* **2022**, *14*, 3892. [CrossRef]
49. Du, Y.; Pan, N.; Xu, Z.; Shen, Y.; Kang, H. Pavement distress detection and classification based on YOLO network. *Int. J. Pavement Eng.* **2012**, *22*, 1659–1672. [CrossRef]
50. Ghosh, R.; Smadi, O. Automated detection and classification of pavement distresses using 3d pavement surface images and deep learning. *Transp. Res. Rec.* **2021**, *2675*, 1359–1374. [CrossRef]
51. Guerrieri, M.; Parla, G.; Corriere, F. A new methodology to estimate deformation of longitudinal safety barriers. *ARPN J. Eng. Appl. Sci.* **2013**, *8*, 763–769.
52. Guerrieri, M.; Corriere, F.; Parla, G.; Ticali, D. Estimation of pollutant emissions from road traffic by image processing techniques: A case study in a suburban area. *ARPN J. Eng. Appl. Sci.* **2013**, *8*, 668–676.
53. Guerrieri, M.; Ticali, D. Design Standards for Converting Unused Railway Lines into Greenways. In *ICSDC 2011, Proceedings of the International Conference on Sustainable Design and Construction, Kansas City, Missouri, 23-25 March 2011*; ASCE: Reston, VA, USA; University of Kansas: Lawrence, KS, USA, 2012; Volume 2012, pp. 654–660.
54. Guerrieri, M.; Parla, G.; Ticali, D. A theoretical and experimental approach to reconstructing the transverse profile of worn-out rails. *Ing. Ferrov.* **2012**, *67*, 23–37.
55. Lima, M.S.S.; Buttgereit, A.; Queiroz, C.; Haritonovs, V.; Gschösser, F. Optimizing Financial Allocation for Maintenance and Rehabilitation of Munster's Road Network Using the World Bank's RONET Model. *Infrastructures* **2022**, *7*, 32. [CrossRef]