



Article

# Towards Low-Cost Pavement Condition Health Monitoring and Analysis Using Deep Learning

Ronald Roberts \* , Gaspare Giancontieri, Laura Inzerillo  and Gaetano Di Mino

DIING—Department of Engineering, University of Palermo, Viale delle Scienze ed.8, 90128 Palermo, Italy; gaspare.giancontieri@unipa.it (G.G.); laura.inzerillo@unipa.it (L.I.); gaetano.dimino@unipa.it (G.D.M.)

\* Correspondence: ronaldanthony.roberts@unipa.it; Tel.: +39-348-442-8206

Received: 13 December 2019; Accepted: 28 December 2019; Published: 1 January 2020



**Abstract:** Governments are faced with countless challenges to maintain conditions of road networks. This is due to financial and physical resource deficiencies of road authorities. Therefore, low-cost automated systems are sought after to alleviate these issues and deliver adequate road conditions for citizens. There have been several attempts at creating such systems and integrating them within Pavement management systems. This paper utilizes replicable deep learning techniques to carry out hotspot analyses on urban road networks highlighting important pavement distress types and associated severities. Following this, analyses were performed illustrating how the hotspot analysis can be carried out to continuously monitor the structural health of the pavement network. The methodology is applied to a road network in Sicily, Italy where there are numerous roads in need of rehabilitation and repair. Damage detection models were created which accurately highlight the location and a severity assessment. Harmonized distress categories, based on industry standards, are utilized to create practical workflows. This creates a pipeline for future applications of automated pavement distress classification and a platform for an integrated approach towards optimizing urban pavement management systems.

**Keywords:** pavement management system; road pavement distresses; automated detection; low-cost technologies; deep learning; pavement condition monitoring

## 1. Introduction

### 1.1. Pavement Management Systems

Today, Road authorities worldwide are facing increasingly daunting challenges with regards to Pavement Maintenance and rehabilitation (M&R) programs. This is largely due to deficient budgets which have faced even further reductions over the last few years [1]. These reduced costs place significant stress on Pavement Management Systems (PMS), which try to balance budgets with the optimal road user conditions [2]. The PMS is highly data-dependent and the acquisition of this data can be costly and time-consuming. As a result of this, agencies are generally relegated to using conventional manual conditional surveys to detect and monitor the conditions of road networks [3]. This can lead to inefficient practices and strategies for interventions. As there is a direct relationship between accident rates and surface conditions [4], it is, therefore, vital to have effective systems in place to ensure the health of the road structure is kept in an optimal state. For this to be possible, authorities must have systems in place for the acquisition of road condition data, specifically: the identification of pavement distresses in the network, their location, classification and importantly their severity.

### 1.2. The Need for Automated Detection Systems for Condition Monitoring

Given the importance of detecting, classifying and analyzing distresses, there have been numerous attempts exploring different methodologies for detecting pavement distresses [5]. The methods researched can generally be grouped in terms of the equipment utilized, which are namely laser technologies, camera and imagery tools, in situ sensors, radar and other less researched methods such as pressure-based, acoustic and vibration methods. Of these technologies, the two most researched and utilized in industry are laser and image-based methods with the laser-based ones generally having a higher accuracy but also carrying significantly higher costs. The in situ sensors have also shown value in their use given that sensors can be deployed within the pavement allowing for remote condition monitoring [6,7]. Sensors can be deployed in various forms allowing sensor information, such as strain and deflection, to be collected for interpretation using structural health monitoring algorithms. Many of these applications involve wired sensors but these have difficulties for deployment given their fixed nature and they can also be costly [8]. Others have looked at the use of wireless smart sensors [9]. There are merits to these systems but there are still significant studies needed to understand the data from these sources and verifying the life expectancy of the sensors.

Laser systems have been used to develop systems such as the Laser Crack Measurement System (LCMS) [10], Laser Road Imaging System (LRIS) [11] and other vehicular systems developed in Australia [12] by combining lasers systems with imagery as well. The laser profiling systems have been developed based on this and are seen to be components in several commercial road systems such as ARAN, ROMDAS and Dynatest [5] and they offer accurate assessments of the state of the road but at a higher cost. With respect to the image-based systems, there has been a lot of research done on the use of images to detect and analyze specific distresses with systems being developed around Digital cameras, Line-scan cameras and 3D Imaging with methods achieving suitable accuracies in replicating distress patterns and classifying them [13]. These systems offer a cheaper alternative to carrying out distress surveys as opposed to laser-based systems.

Recent work has also included research on low-cost assessments of 3D models for pavement distress analysis wherein the accuracy of models created using the techniques of photogrammetry and 3D modelling have been validated [14]. Whilst this and similar methods show accuracy in replicating the distress, it is also important to quickly pinpoint where the distresses are occurring. Therefore, as a precursor to more sophisticated methods, it is useful to have a low-cost system able to determine the specific locations of the distresses which would then be followed by detailed assessments. Together these would form a low-cost tool ready to be used by road authorities. To quickly perform a hotspot analysis of a road network would be very useful for any road authority. To this end, the use of Artificial Intelligence and deep learning technologies have gained traction in this automated process and towards this type of pavement distress hotspot analysis.

This paper aims at utilizing deep learning techniques namely the use of object detection within imagery collected, using low-cost smartphones, to gain a rapid assessment of the condition of a road network based on standardized distress techniques and severity determinations and to demonstrate the ability to use this process to continuously monitor the condition of the pavement structure. The research explores the requirements needed for a deep learning model for this type of work which can be applied in different environments and regions. The study utilizes the use of harmonized distress categories based on industry standardized distress types with associated specified severities so that the final model can give the user a better understanding of the level of distresses along a network and the required interventions. This differs from previous work as the model developed not only seeks to highlight the location of the distress and type but also a quick but accurate severity assessment. Using this, a pipeline for future integrations within the PMS is proposed.

## 2. Materials and Methods

### 2.1. Background of Deep Learning Techniques

Recently, there has been a universal rise in the research and use of deep learning applications to solve complex and different problems across varying research fields [15]. This has been due to the increases in the accuracy of these methodologies. The accuracies in some instances are now considered even better than those of human abilities as demonstrated by the ImageNet large scale visual recognition challenge [16] in which the Human benchmark for recognizing objects was beaten in 2015 by a developed deep learning framework [17] and whose accuracy mark has continued to be improved upon every year of the challenge as new frameworks and algorithms have been developed.

The basic concept of Machine learning is one in which a computer system can automatically carry out tasks such as object detection, image classification and speech recognition based on supplying it with datasets centered around the individual task. Deep Learning is a subset of Machine learning which utilizes neural networks rather than traditional handcrafted features. Deep Learning involves the use of computationally based models which are made of several processing layers which learn data representations with several layers of generalization [18]. Artificial Neural Networks (ANN) are based on the biological neural network and consist of input layers, hidden layers and finally output layers. Deep learning networks figure out the complex structures within datasets by the utilization of backpropagation algorithms which alert the computer how to alter its internal parameters to yield the best representation of the next layer and eventually the final model. Neurons within the network receive inputs then process them and feed it forward to other neurons in successive layers. The hidden neurons take outputs from previous layers, compute new outputs with the use of activation functions which are then fed forward to the next layer with the use of different weights applied during the applications. A simple graphic of this is demonstrated in Figure 1.

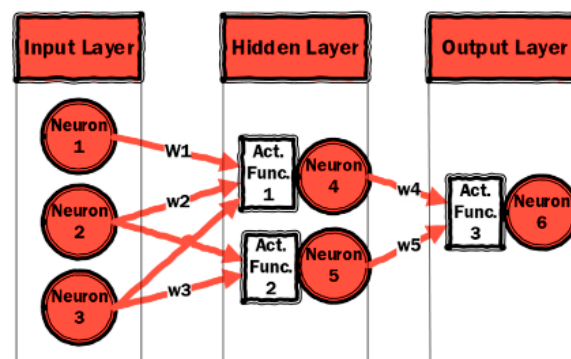


Figure 1. Basic structure within a neural network.

They essentially work to understand how the data is built up and thus how to identify, predict or classify future similar but unknown datasets. For image processing, convolutions are utilized. Convolutions utilize filters which help extract different features from the image such as edge information. In neural networks, features are extracted using filters with weights that undergo learning during the training process. The retrieved features are then summed together to make choices. The spatial relationship of pixels within an image is also considered in the convolution which helps to identify particular objects that have defined spatial relationships with other objects. The most common type of system is called supervised learning wherein a convolutional neural network is fed annotated datasets and learns to make similar predictions based on the learning data. Convolutional neural networks (CNN) have been vastly utilized for this and they are designed in such a way to process data in the form of arrays. CNN's are generally utilized for detection, image segmentation and classification and identifying objects and regions within an image. A typical depiction of the workflow for deploying one of these models and network is shown in Figure 2.

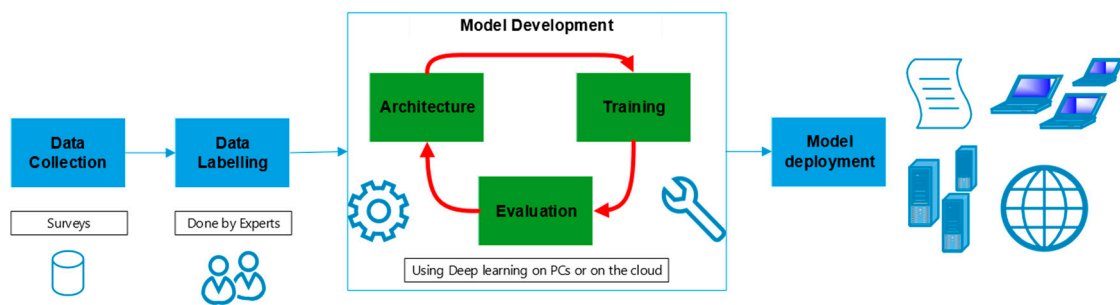


Figure 2. Typical model development setup.

There are existing networks which have been trained on millions of images and utilizing high-performance PC systems which are very deep with complex convolutional layers. Using these as building blocks, Transfer learning has been utilized as a quicker way to develop models without starting over the process, thus reducing time and computational costs. It is an approach wherein knowledge developed in one task is transferred and used to improve the learning of different target tasks using a pre-trained model as a baseline [19]. Within the CNN, each layer extracts features from an image and successive layers further extract more complex features. As the initial layers are essentially used for low-level features such as curves and edges these layers can be used for new tasks with training being done on the classifier of the network or the fully connected layer specifically for the new task. This process has been readily used over the last few years given its efficiency [20].

## 2.2. The Use of Deep Learning in Pavement Engineering

Given the advances of deep learning, there has been significant research using these techniques for Pavement Engineering applications [21–24]. These applications can be assigned to the following areas: Pavement condition and performance predictions [25–28], Pavement management systems [29–31], pavement performance forecasting [32–34], structural evaluations [35–37], modelling pavement materials [38–40] and pavement image analysis and classification [22,41–44]. Pavement Image analysis and classification is the most researched area, where the focus has been split between image classifications, where images are classified based on the distress occurring in the image; and object detection, where distresses are located within bounding boxes or masks within the image. There are however issues with image classification as distresses regularly occur in a grouped manner thereby making it difficult to label a particular image with just one distress type label. This is important for a road engineer as information on connected distresses and the locations of areas where there are multiple distresses is vital for the asset management system to accurately monitor the road conditions. With object detection, it is possible to have multiple overlapping objects within an image and therefore it possible to detect multiple overlapping pavement distresses.

A significant proportion of previous studies has been focused on developing models and networks to determine whether there is a presence of a distress or not and also the general detection of pavement distresses [41]. Of the distress types, the main focus has been crack detection and analysis. This is due to the fact that pavement cracks are seen as the most predominant distress type [45] and they are also easier to measure, with the typical requirements being simply to measure the crack's width and length. There are a tremendous number of studies on developing specific neural networks for crack detection and analysis using both 2D and 3D imagery [41,42,46–52] and with comparisons made to results from image-based toolboxes for crack detection and analysis such as CrackIT [53]. While the detection and monitoring of cracks are important to road agencies, this represents only one main category of distress. There has not been a lot of focus on standardized distress categories as accredited by international manuals on the subject [39]. There have been a few that have tried to analyze multiple distress categories and generate datasets of multiple types. A research team in Germany developed a CNN for pavement distress application based on imagery obtained through surveys across the German road network using a mobile mapping system attached to a vehicle [54,55]. Their team developed,



the German Asphalt Pavement Distress (GAPS) dataset which was utilized to generate classifications based on six different distress categories based on the German road manuals with research ongoing utilizing their developed Neural network, ASVINOS.

Studies have also been carried out in Italy in which fourteen different categories of distresses were analyzed with the application of semantic segmentation and object detection algorithms on a dataset within Naples, Italy [56]. This was done to formulate a decision support system based on the occurrence of the predicted distresses within the datasets which further highlights the importance of the detection of multiple distresses. There is also a large dataset of road surfaces which exists called the KITTI dataset [57] but this dataset was created primarily for the purpose of assisting with automated driving research. There has also been the development of a database of road distresses in Japan through a mobile application [58] in which eight distresses were annotated within the dataset. The work has led to technical challenges such as an IEEE Big Data challenge based around the database where different models were submitted to obtain higher accuracies based on different network and hyperparameter configurations. This led to several different network configurations using different base networks and models for the same goal of detecting the distresses within the dataset [59–61]. Whilst these developed models represent a significant step forward in pavement distress detection and analysis the models do not yield any information on the severity of the distress, which would provide an understanding of not only the distress type present but also a trigger for interventions. This study further explores this, wherein different distress assessments and model configurations were used to develop a low-cost methodology and tool to enable road agencies to monitor the road structure and enable the establishment of points for road maintenance intervention.

### 3. Materials and Methods

Given the state of the research field and the importance of automating pavement detection systems, to allow for effective condition monitoring, the most integral part of the work was the development of the object detection model which was done within the open-source TensorFlow environment [62]. The setup was done to ensure compatibility with this environment. The workflow to do this is shown in Figure 3, wherein the steps are shown from data collection to final model deployment. This workflow is further explained in Sections 3.1–3.5. It is fully replicable for other datasets and can be utilized for generating models for different cities or regions to enable creating a model based on particular conditions that exist within the road authorities' environment.

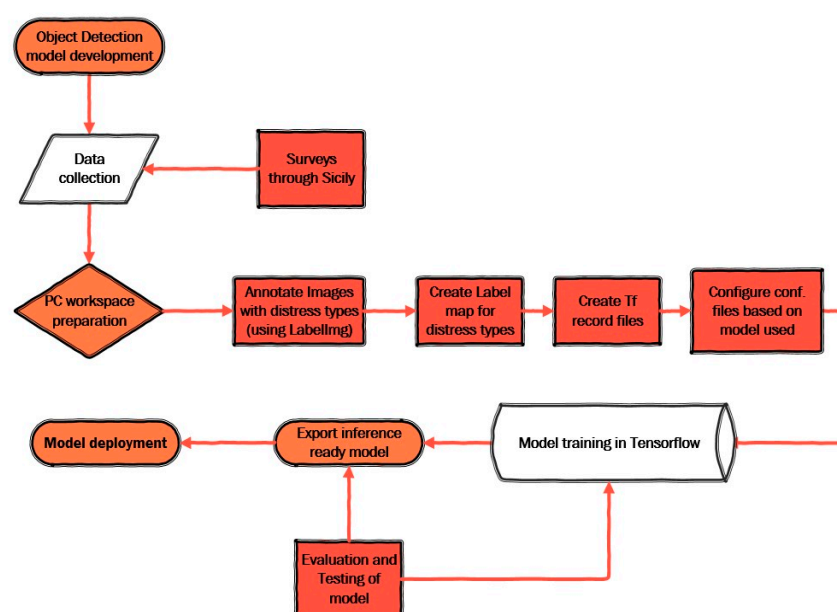


Figure 3. Pipeline utilized for the development of object detection model.

### 3.1. Data Collection

It was important for the exercise to establish a model that could be used in the specific local conditions in Sicily, Italy. Given this challenge, it was necessary to obtain a collection of images from the Sicilian region. To overcome this challenge, the application, MyCityReport [58], was utilized to capture images from a smartphone which was mounted in a car driven along the Sicilian urban roadways. This setup is depicted in Figure 4. The application has the ability to capture images at an approximate distance of 10 m ahead of the positioned phone with photographs taken every second. The application also has the ability to classify 8 types of distresses as previously mentioned. This option was not utilized however as the application was relied on solely for data collection purposes. For the purpose of this study, several trips were made across urban road networks in Sicily generating over 7000 images. The weather and area were diverse within the dataset so as to offer a robust training dataset of urban road networks within the region. However, only the images that were taken during the day and when there was no rain were used. This does identify a limitation of the process as it is difficult to accurately identify distresses during inclement weather conditions. Additionally, only images of flexible pavements were used. The same camera phone, the Google Pixel 2XL, was utilized for all trips to ensure all of the images had similar image qualities and dimensions.



Figure 4. Image depicting smartphone setup during surveys.

### 3.2. Data Annotation

For this study, the open-source labelling software, LabelImg [63], was utilized to individually manually label images based on the type of distresses present. An example of this work is shown in Figure 5. This was done by trained civil engineers with experience in asphalt pavement engineering and conditional surveys, which are typically used to define and detect pavement distresses.

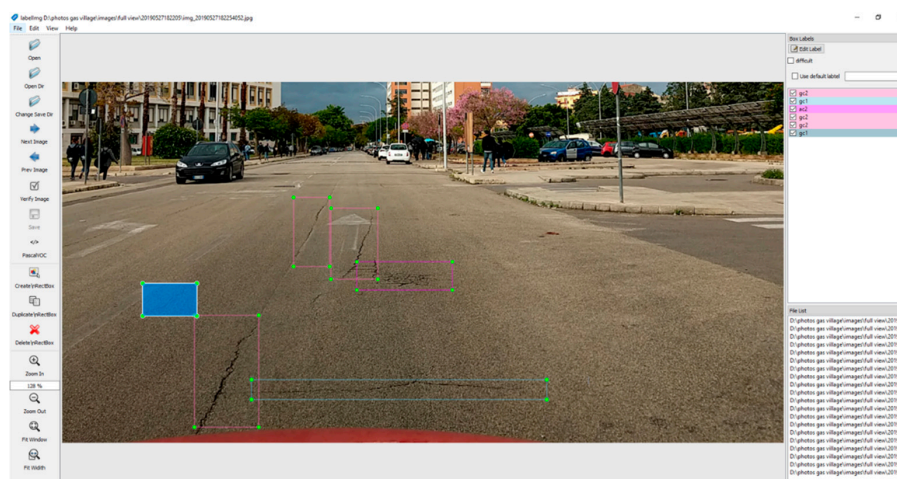


Figure 5. Annotation of images using LabelImg.

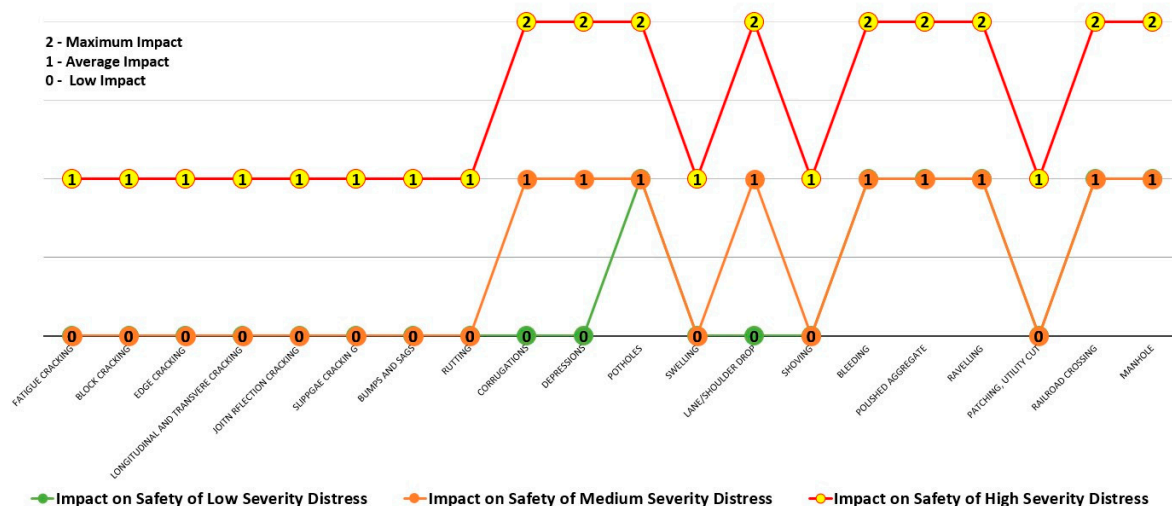
This software utilizes the PASCAL VOC format [64] for the labelling to create .xml files for each image. The critical and novel step in this process was the development of the distress categories for the model. Typically, pavement distresses can be broken down into four major categories namely: cracking, visco-plastic deformation, surface defects and other miscellaneous types [13]. The grouping is shown in Table 1.

**Table 1.** Generalized pavement distress categories.

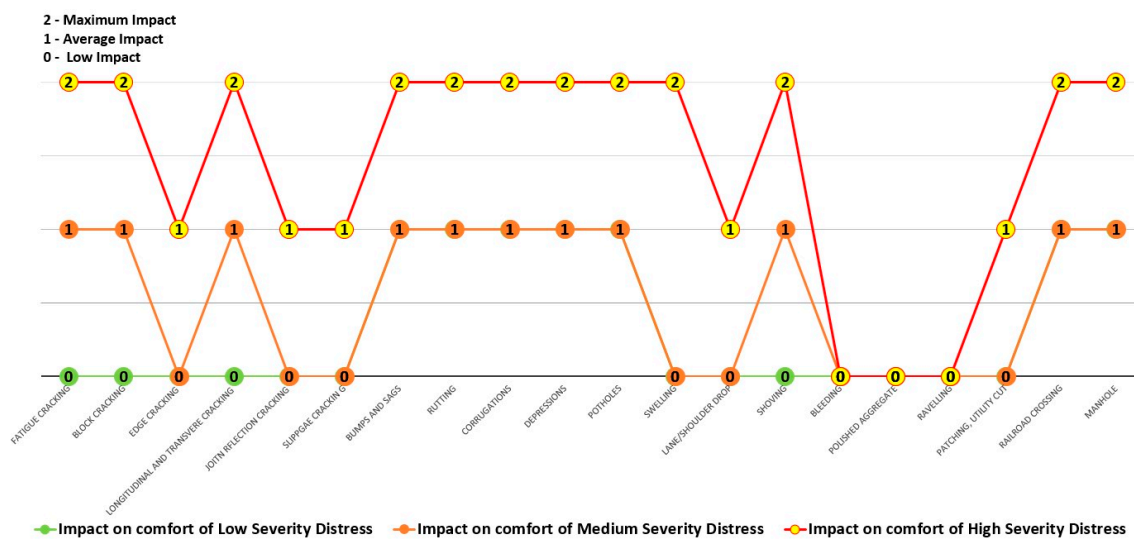
Distress Category	Distresses
Cracking	Fatigue cracking, Block cracking, Edge Cracking, Longitudinal and Transverse cracking, Joint reflection cracking, Slippage cracking
Visco-plastic deformations	Bumps and Sags, Rutting, Corrugations, Depressions, Potholes, Swelling, Lane/Shoulder drop off, Shoving, Stripping
Surface defects	Bleeding, Polished Aggregate, Raveling
Others	Patching/Utility cut patching, Railroad crossing, Manholes

Of these categories, other studies have shown that there is a direct relationship between the impact each distress has upon safety and comfort per level of severity [13]. Based on this each distress has an impact based on the severity of the occurring distress. This is further depicted in Figures 6 and 7. From these figures, it can be seen that the most impactful severity level, as expected, was the High severity cases whilst the medium and low severities in most cases have fewer impacts upon safety and comfort.

The severity level is usually then derived based on data from manual surveys and based on regulations implemented by pavement distress manuals such as [65–67]. Different ratings are then utilized to determine the overall condition of the roads based on the presence of the distresses and other factors such as roughness which are correlated in typical performance indices such as International Roughness Index (IRI) [68] and Pavement Condition Index (PCI) [65].



**Figure 6.** Impacts of the severity of pavement distresses on safety.

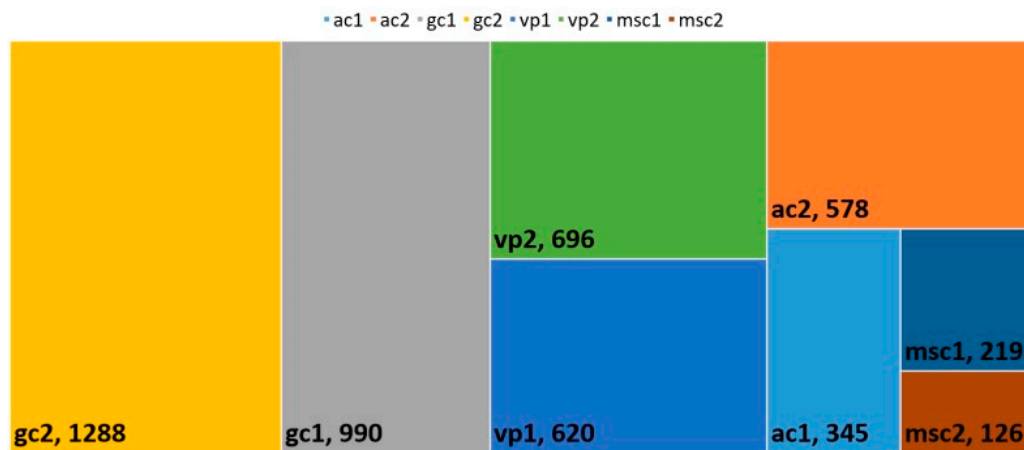


**Figure 7.** Impacts of the severity of pavement distresses on comfort.

Based on this, the decision was chosen to have only two severity levels (level 1 and 2), wherein the first represents situations wherein remedial action is not a necessity and the second where it should be done. Using the four general groups as a base case, annotations were made for each with the exemption of Surface defects (Bleeding, Polished aggregate and Raveling) where it is difficult to accurately pinpoint the distress and its associated severity level from a 2D image. The cracking group was also split between general cracking (gc) and area cracking (ac) (as defined by cracking made over a section, such as alligator cracking, as opposed to instances of cracks that are formed at specific points on the road surface or along its surface). With these categories, the developed model should be able to predict different instances of cracking, instances of visco-plastic deformations (vp) and miscellaneous distresses (msc) such as manholes whilst also providing a quick analysis of their severities. There are limitations using this approach as no precise metric measurements would be done on the distresses and having only 2 severity levels does not provide information on cases where intervention action may be required in the near future as could be interrupted by a medium level severity assessment.

However, gaining a general understanding of where these groups of distresses occur and their frequency can provide practitioners with valuable information and allow for an adequate resource for continuously monitoring the overall health of the road structure. Future work will focus on how to integrate other types of groupings as well as surface defects. From these groupings, annotations were manually made only on images where there were was a clear view of the distress and it could also be clearly marked. This resulted in a total of 4862 distress annotations to be used for the model as split as shown in Figure 8.

Within this distribution, the most observed distresses are general cracking followed by viscoplastic deformations and then area cracking, with the least being the miscellaneous category. This is expected given the nature of distresses on most urban Italian road networks [45]. Using these distress types, a label map was generated for the previously identified TensorFlow pipeline. Subsequently, the annotated files were converted to the record format to be used within TensorFlow and the datasets were randomly split in the ratio of 80%:20% for training the model and for testing to ensure the model is not overfitted to the dataset and would therefore not be able to effectively perform on unseen real-world data. Data augmentation was also applied wherein each training image horizontally flipped with a probability of 0.5. This enables using a smaller dataset as well.



**Figure 8.** Treemap diagram showcasing the distribution of annotated distresses. ac: Area cracking, gc: general cracking, vp: visco-plastic deformations, msc: miscellaneous distresses 1: Low severity, 2: High severity requiring intervention.

### 3.3. Model Setup

For the neural network setup, it was decided to use transfer learning given the size of the dataset and the available base networks. Several different base models were considered given the available networks within the TensorFlow Object Detection API (application program interface) [69], which is an open-source framework developed by Google, which was built upon TensorFlow [62] that allows easy construction, training and deployment of object detection models. Within the Google API, there are prebuilt architectures and weights such as the Single Shot Multi-box detector (SSD) [70] using MobileNet [71], Inception V2 [72], Region-based fully convolutional networks (R-FCN) [73] as well as the Faster R-CNN networks (region based convolutional networks) [74].

For the purpose of this study, the following base models were considered: Faster R-CNN using Inception V2, based on COCO (common objects in context) dataset [75], Single Shot Detector (SSD) using InceptionV2 model based on the COCO dataset and the SSD using MobileNetV2 also based on the COCO dataset. These were chosen as they've shown accuracies in previous model evaluations [69]. The properties of the chosen models are given in Table 2.

**Table 2.** Properties of object detection models.

Model Name	Speed (milliseconds, ms)	COCO Mean Average Precision (mAP)	Outputs
faster_rcnn_inception_v2_coco	58	28	Boxes
ssd_inception_v2_coco	42	24	Boxes
ssd_mobilenet_v2_coco	31	22	Boxes

These models are all publicly available through the TensorFlow object detection API zoo. Each of these models provides a quick model creation pipeline which can be developed without the use of heavy computational resources. The two main networks utilized for the work are the Faster R-CNN model and the SSD model with the utilization of the inceptionv2 and mobilenetv2 CNNs. Within the Faster R-CNN base model, the same convolutional network is used for both region proposal generation and the object detection task. This model essentially proposes regions, extracts features from these regions and classifies the regions based on the features. This enables the detection to be faster. This is depicted in Figure 9 below based on the architecture of the model [74].



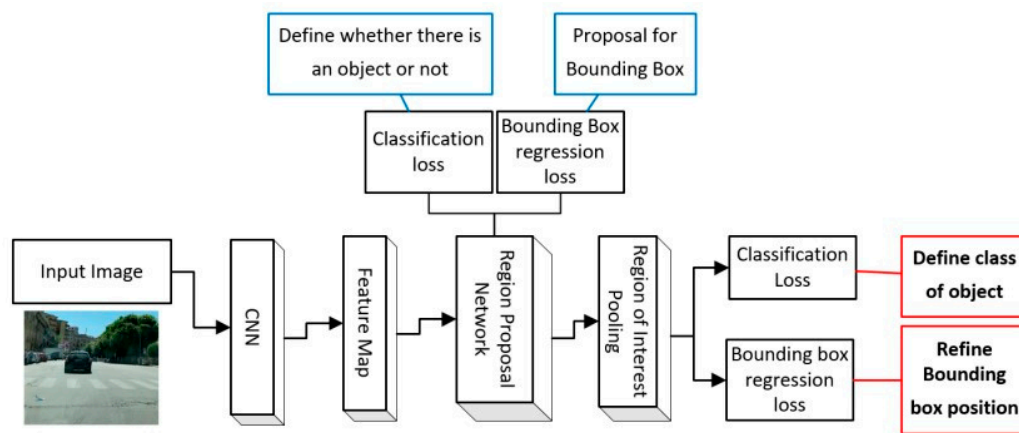


Figure 9. The architecture of Faster R-CNN model.

For the Single Shot Detector (SSD) base network, only one single shot is required to detect multiple objects within an image. It is faster when compared to Region proposed networks which require two shots, one for region proposal generation and then the second for the object detection. Within the SSD, the input images are passed through several convolutional layers to produce several feature maps at varying scales. Then for each location in each map, a convolutional filter evaluates a small set of bounding boxes and for each bounding box, it predicts the offset and the class probabilities. The SSD is a feed-forward CNN which yields a static size of bounding boxes and scores which is followed by a non-maximum suppression step that yields the final detections of the model. This framework for the network is depicted in Figure 10 which shows the architecture of the model [70]. It is similar to the other model but essentially avoids the region proposal step and considers all possible bounding boxes in every location in the image whilst simultaneously carrying out the classification.

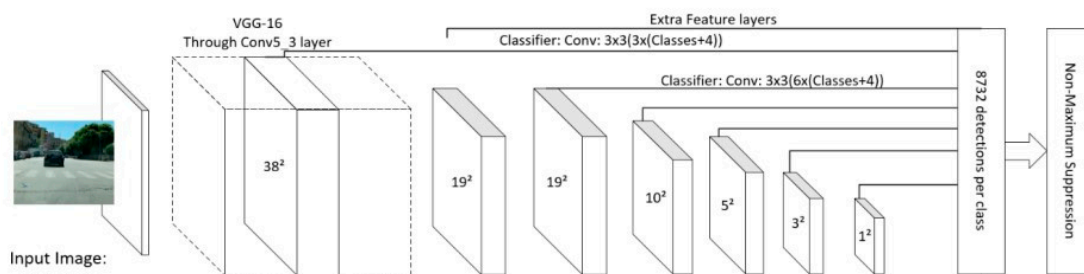


Figure 10. The architecture of SSD model.

### 3.4. Object Detection Model

For the Faster R-CNN with Inception v2, the parameters used were as follows. An initial learning rate of 0.001 was used and then reduced by a decay of 0.95 every 10,000 steps. The learning rate was in line with that used by previous studies on pavement distress images [58] whilst other rates were experimented with but this rate proving to be effective. A decay was also utilized for the learning rate which helps the model develop momentum and create a quicker convergence as well as reducing opportunities for overfitting with one constant rate. The input images were also resized to  $300 \times 300$  pixels. For the SSD using Inception V2, an initial learning rate of 0.002 was used and then reduced by a decay of 0.95 every 10,000 steps. The same approach of a time decay for the learning rate was utilized. The input images were also resized to  $300 \times 300$  pixels. The same hyperparameters were established for the SSD using the MobileNetv2 model for comparative purposes. These configurations were set within the configuration files of each prebuilt model.

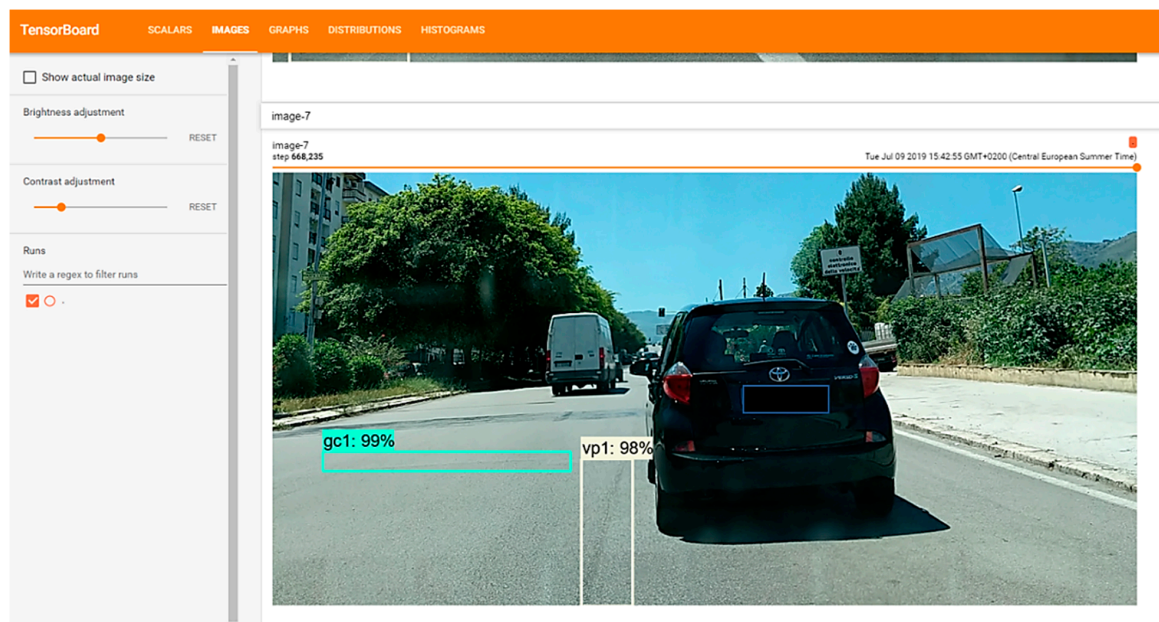
### 3.5. Experimental Setup

For the training and evaluation, a Windows 10 PC was utilized with an NVIDIA Quadro P4000 GPU (8 GB ram) and total CPU memory of 32 GB. Within the evaluation state, the Intersection Over Union (IOU) evaluation metric was used. This metric is defined by dividing the area of overlap between the bounding boxes by the area of union between them and for this exercise the threshold was set to 0.5. The IOU essentially provides an estimation of the accuracy of the bounding box as compared to the ground truth defined by labels that are kept separate from the training data used.

## 4. Results and Discussions

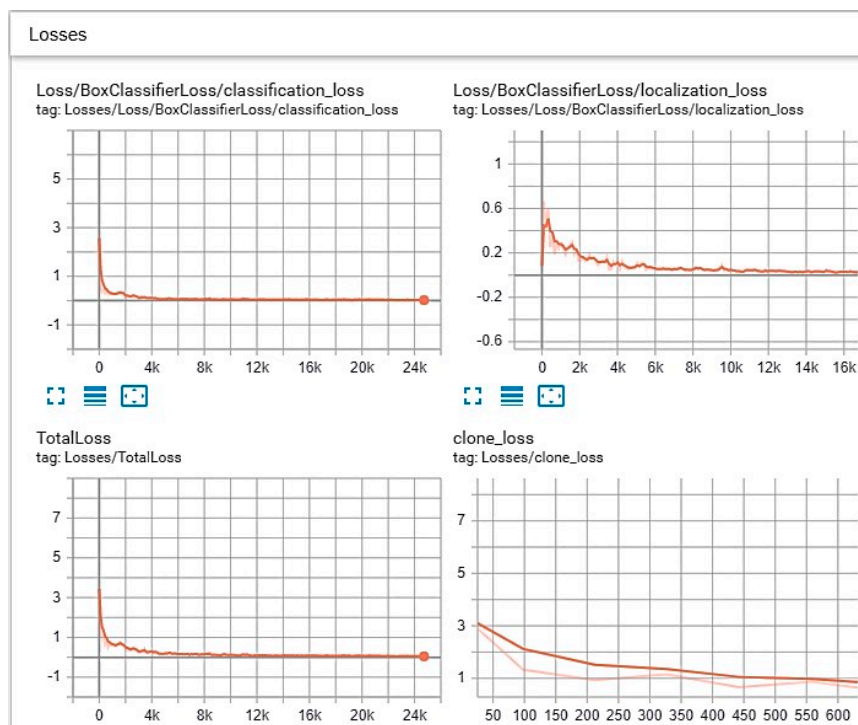
### 4.1. Trained Models

As highlighted in Section 3, each model was run using the TensorFlow environment and during the training, the evaluations were observed and followed until the model achieved an acceptable loss level below 1. This, however, does not represent an accuracy of the model and evaluation needed to be carried out on the model regardless of the loss value. This was observed through the TensorBoard environment on as shown in Figure 11.



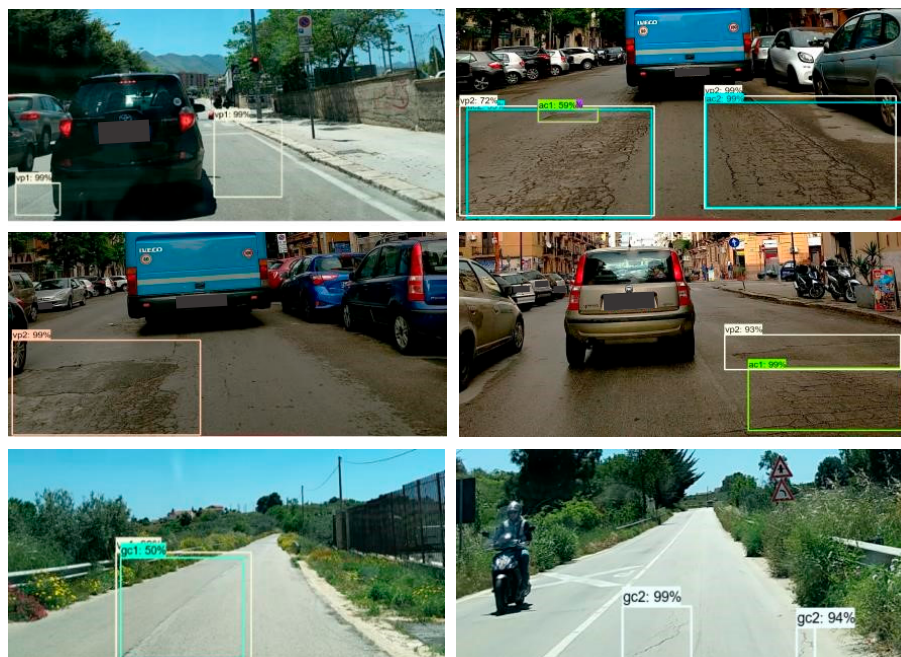
**Figure 11.** TensorBoard environment utilized during the training of the network.

By utilizing the TensorBoard system, it allowed the user to monitor the progress of the model throughout its training. Of particular importance to monitor in this case was the loss which is shown in Figure 12. The graphs in Figure 12 display the loss over the training iterations time with respect to different characteristics of the model i.e., object detection, localization, and classification. Within the figure, it can be observed that this loss reduces over time and it is important for the model to have a sufficiently low value for the total loss. The platform also allowed the user to monitor the progress of the model on test data during the training which is also helpful as the training could be stopped if it was seen that the model was not producing appropriate results.



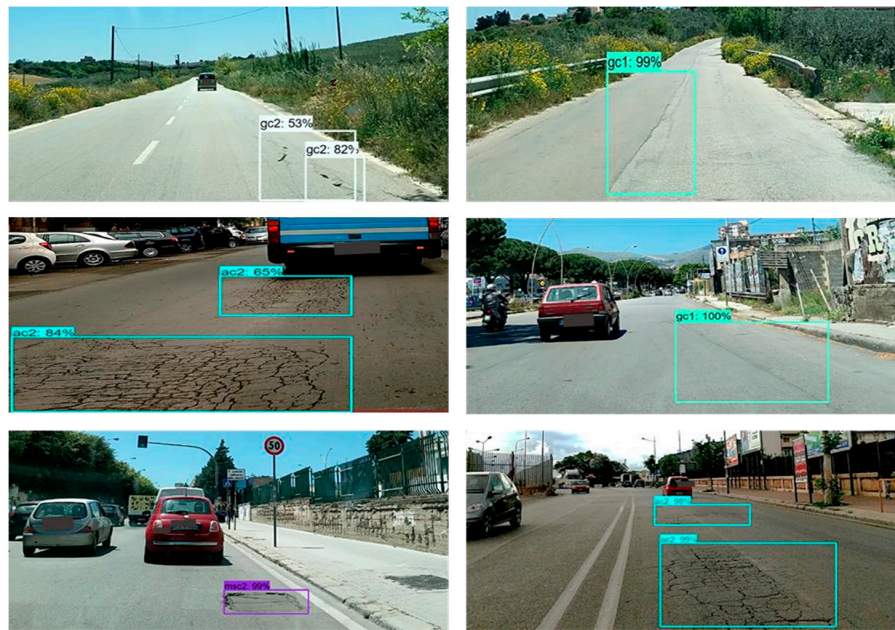
**Figure 12.** The use of TensorBoard to monitor the loss and check on accuracies within the models during evaluation.

Each model was run on the same dataset. After training, the models produced were tested on prediction images run from the test data set and examples of these can be seen in Figures 13–15. The model produces a bounding box on the image of the distress type and severity along with a percentage assumption of how accurate this bounding box is based on the calibrations of the model. This value would provide users with an overall assessment of how good the hotspot analysis is based on the sum of the total possible errors on a survey over a road network.

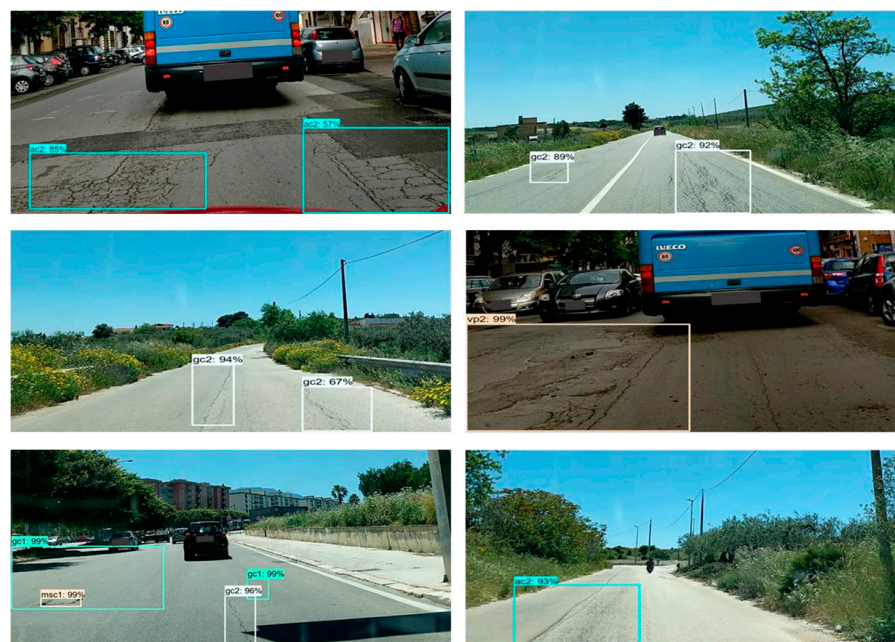


**Figure 13.** Faster R-CNN with inceptionv2 model predictions on test data.





**Figure 14.** SSD with Inceptionv2 model predictions on test data.



**Figure 15.** SSD with mobilenetv2 model predictions on test data.

## 4.2. Inference

Each model was exported to create a graph file that is able to run inference and possibly be deployed to an application or mobile app. This was successfully done for each model at the last given checkpoints of the models. This is important and this step is needed for the model to run on a mobile device and thus be ready for use in a mobile application.

## 5. Discussions

### 5.1. Performance of Models

Based on the outputs of each model, complete evaluations were done on the final graph. As the base models used for the development of the new model were generated on the 'COCO dataset' it

was, therefore, instructive to utilize the ‘COCO metrics’ for evaluating the performance of the model. These metrics involve the use of Precision, denoted as Average Precision (AP) and recall, denoted as Average Recall (AR) with regards to the bounding boxes within test cases. Precision can be classified as the results within a test that are relevant to the classification problem whilst Recall refers to the percentage of relevant results which are in turn correctly identified by the model. In the evaluations, the Intersection Over Union threshold was set to 0.5. The two parameters are determined by the following equations:

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (1)$$

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (2)$$

These values are obtained by running the eval.py scripts on the model using the test dataset for the study. These results for all the models are shown in Table 3 below.

**Table 3.** Average Precision and Recall values for models utilized.

	SSD with Inception v2	Faster R-CNN with Inception v2	SSD with MobileNet v2
Average Precision	0.909	0.933	0.880
Average Recall	0.929	0.938	0.867

Given the numbers showcased within this table, it can be surmised that all 3 models are capable of solving the problem at hand with the model utilizing the Faster R-CNN model being the most effective of the three. For the purposes of this study, however, what is critical is not only the accuracy of the model but also the simple capacity to recognize that the type of distress and its associated distress category are present within the image and subsequently within the network or road section under analysis. This creates a tool that allows for low-cost hot spot analyses of points on a given road network where there are structural defects and as a result can provide an overall condition estimation of the road network being assessed.

Once a survey is carried out using the model it would also provide a baseline which can then be utilized for future continuous overall condition monitoring of the pavement structure by simply rerunning the model and comparing the defects between the different surveys. This would be without the need for expensive equipment or elaborate survey mechanisms. This is key for road managers and allows for integration within the PMS system. The road authority would be able to determine the location of the specific types of distresses covered by the model, its severity (and thereby an idea of whether an intervention is necessary) and the location and frequency of specific and important types of distresses. This can also then be integrated with techniques that can provide the specific or more detailed measurements required for further actions.

## 5.2. Pipeline for Utilizing Model with Hotspot Analysis

Given the results of Section 5.1, a pipeline was developed where the model can be integrated into the assessment of pavement distresses along with 3D Modelling techniques utilized in previous work [14]. This pipeline is depicted in Figure 16, with the final goal being to produce decisions on which M&R activities are necessary for a given road network. Within this pipeline, the integration is made using the Deep learning network along with 3D Modelling techniques for the overall assessment of the road network and its distresses. In the framework, the assessment can be done by utilizing smartphones. The Neural network model can be integrated easily into an application and used in rapid surveys to produce locations (GPS based) of the distresses (which are detected in a box on the images), an understanding of the structural defects at the locations; and when this survey is carried out over a network, the frequency of these two parameters would be produced. It also showcases how Road Authorities can help solve the issue of monitoring the health of the road network in order to determine which roadways to rehabilitate and when. This would be possible as the data in their asset



management system would be easily kept up to date, with the east access to sufficient monitoring data to make these decisions continuously over the pavement's life cycle.

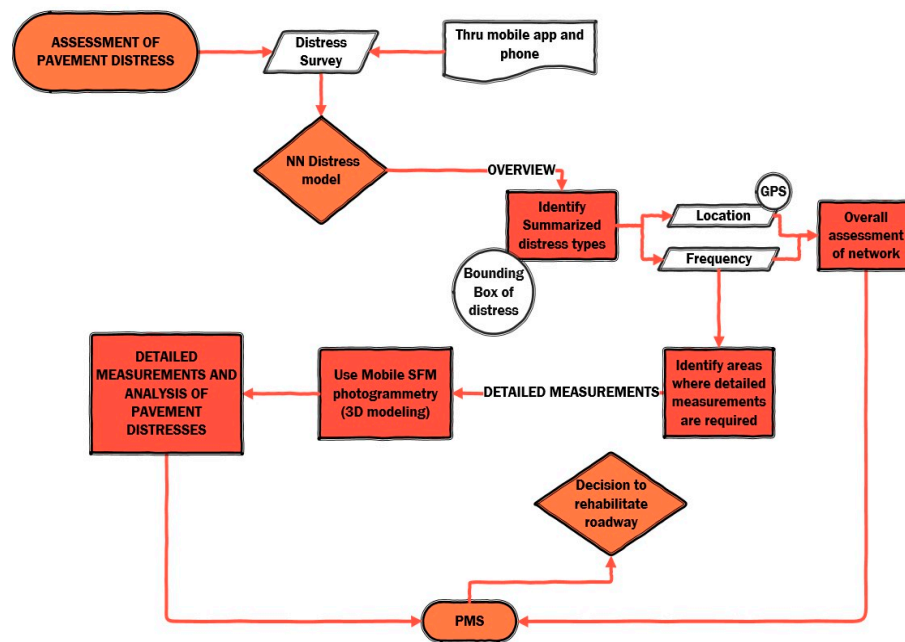


Figure 16. Pipeline for hotspot analysis and future integration.

### 5.3. Case Study Showing an Application of the Model for Monitoring the Health of a Road Network

Additionally, whilst previous works have been focused on the creation of deep learning models for the purpose of detecting road pavement distresses, it is also critical to have an idea of how the models can be practically applied in the real world and within practical management systems. To this end, this study also created a pipeline to show how the information from the deep learning model can effectively be utilized and integrated within the asset database and subsequently the management system. From the discussions of the study, it was demonstrated that the model developed generates a bounding box for each observed distress within the image as well as a judgement on the severity based on the annotation parameters (a level of 1 or 2 for each distress type). This is coupled with the accuracy of the detection. With this information, a database displaying and storing the location of the distress per image along a road network can be generated and be kept for monitoring comparisons and purposes. For this to be done, each image captured during a survey has to be accounted for, highlighting not only the detection of the distress but also a relationship of this to the length and area of road segment being surveyed. For this to be possible, a pipeline for this to be done is illustrated in Figure 17 below.

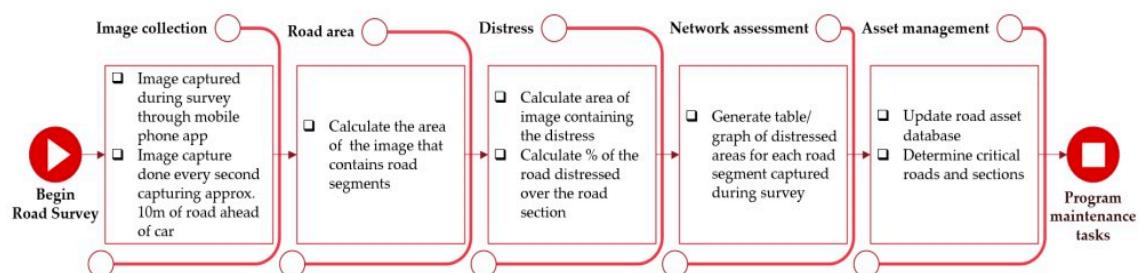
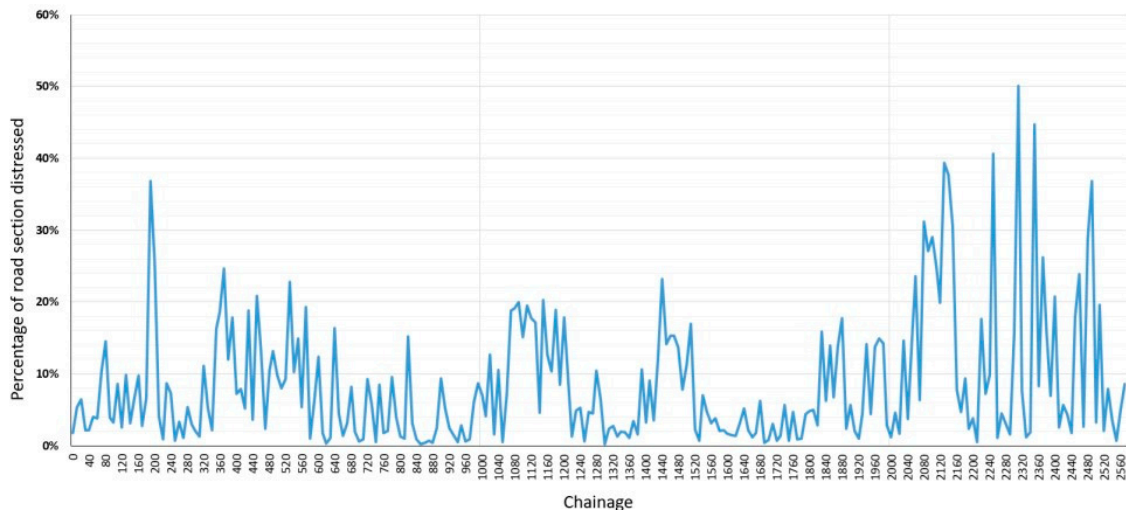


Figure 17. Pipeline for applying the model in real-world conditions for health monitoring.

In the pipeline, the area of the road being surveyed is taken into account to have an approximation of the area that the distress covers and as a result, the total distressed areas along a road segment.

This is an approximate calculation based on the length of the road segment which is clearly visualized within an imager during the survey (10 m utilized for the survey and case study). Given this pipeline, an example road section during the survey was utilized as a case study to demonstrate the practicality of carrying out these tasks. A random section of approximately 2.6 km was identified to test the pipeline. The area of the road distressed at each 10 m interval was determined and this was plotted for the full section. This is depicted in Figure 18.

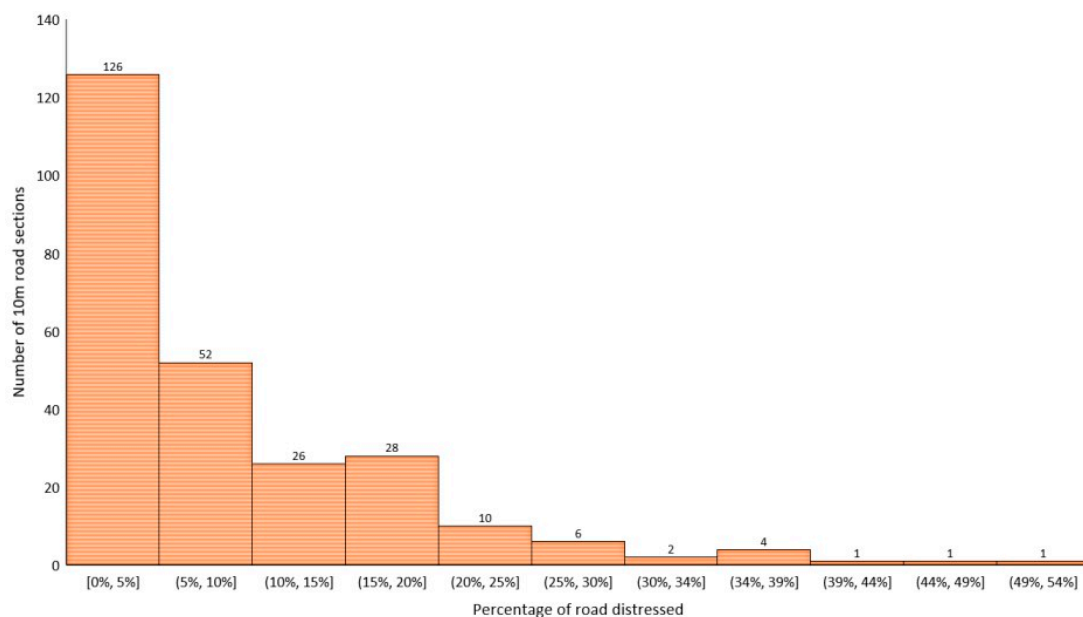


**Figure 18.** Level of the distress on the road across the test section.

From this figure, you can visually identify the sections of the analyzed road that are more distressed. For the section under analysis, it appears that between the road chainage of 2000 m to 2600 m the distressed area is greater than the other sections of the road which means that this section could be attributed with a classification of ‘critical’ health. Therefore, this gives the road practitioner a quick overview of the section under analysis and can allow the agency to pinpoint the area most in need of rehabilitation and maintenance. More critical examinations can be made over this section to determine metric analyses of the distresses and the pipeline utilizing more advanced visual and metric modeling displayed in Figure 16 can be utilized for this. Furthermore, the sections with less distressed areas can be noted within the database and monitored over time. This creates the possibility of monitoring the health of the road network as the survey with the mobile phone can easily be replicated over time creating a timeline of events to monitor.

Another use of this information is to create comparisons of different roads along a network with the goal being to determine which roads should be prioritized for interventions. This is also critical for road agencies as it is typical for an agency to be faced with a scenario where there are several roads that have many distressed areas but there are only sufficient funds for rehabilitations to be done on a selected number of the roads. To this end, the data can be utilized to create a histogram showcasing the number of sections with different levels of distress. Using the same distress information from Figure 18, a histogram highlighting the frequency of distresses along the road section was plotted and this is visualized in Figure 19. For the road section under analysis in Figure 19, most of the intervals have damaged areas with distressed of 0 to 5% with only a few intervals have sections where most of the road has suffered damage. Similar representations can thus be made for each road section that is surveyed over a network. After this, comparisons could be made between them to see which roads have more distressed sections and therefore which should be prioritized for interventions. This will lead to a further channel for monitoring the health of road networks and optimizing maintenance activities. For the methodology proposed, the computational power would be relatively high at the point of model creation as dictated by the deep learning methodologies developed as illustrated in Figure 3 and the computational power needed to run the Tensorflow model. Once the model has

been validated for a given network, an application can then easily record the points of interest, their location and the percentage of damage incurred on the pavement. This is as a result of the application producing grouped georeferenced data on the damaged points and sections with the definitions of severity already defined within the model. This data can essentially be put in the form on a logged CSV file which is not computationally difficult to interpret. Whilst it should be noted that this could produce a very large file based on the size of the network, the data can be manipulated within a simple statistical programming environment to produce figures such as Figures 18 and 19 which visually depict the points of interest along the network and thus the points for maintenance. The data analysis can be more streamlined to highlight other practical examinations considering other environmental factors and testing on larger networks. The case study presented and illustrated in Figures 18 and 19 represents a snapshot of what is possible.



**Figure 19.** Histogram displaying of sections of the road and the respective percentages distressed.

## 6. Conclusions and Future Work

This paper was aimed at presenting a deep learning pipeline and framework for the purpose of carrying out a low-cost hotspot analysis of the pavement distresses that are present on an urban road network for the purpose of continually monitoring the health of the pavement structure. The intention of this was to provide road authorities and engineers with a low-cost platform for quickly understanding not only the type of distresses that are present on the network but also provide an idea of the severity of the distresses which can be stored in an asset database for decision making and monitoring purposes. This would, therefore, lead to better planning for interventions to restore the health of the pavement to an acceptable standard which is critical for any city or town.

Artificial Neural Networks were developed based on local imagery in the Sicilian region in Italy, to set up a model capable of carrying out this task. The imagery was captured using a low-cost smartphone. For the model, a harmonized severity classification was developed which grouped distresses based on the type of damage and their perceived safety and comfort impact. These classifications allowed the model to produce predictions that adequately can carry out the hot-spot analysis required for the study to enable a rapid assessment of the damages to the road. Future work will focus on how to integrate other types of groupings as well as the surface defects as these distresses as they are important to the overall determination of road conditions. Future work will also aim to establish which distresses are more important to be assessed based on the area under analysis. The models produced showed great accuracy in identifying the distress categories developed and the

associated severities. An essential assessment of the models was simply their ability to identify the distress type and severity. However, it must be noted that whilst the models created were found to be suitable for the case study, a new model development should be done for a different city or region. Therefore, what is critical is the pipeline needed to create such a model and this is replicable regardless of the scenario as depicted in the study's methodology and by the ease of use of the available opensource platforms and deep learning techniques. Consequently, this process can be replicated in other regions at a similar low-cost. The data collection processes used also can lead to crowd-sourcing platforms that could be developed on cloud-based systems using smartphones in which datasets are generated from commuters along the network and not the road authority itself.

Once the model is created it can be deployed for use by authorities for the hot-spot analysis. This can be then be complemented by other low-cost techniques such as 3D Imaging with smartphone and drone imagery to provide detailed quantitative assessments of the areas which are considered in need of intervention. This combination provides a step further for the low-cost automation of the system and allows the authority to receive the data necessary for their PMS optimization schedules. The information obtained through the hotspot analysis is also very useful and this study demonstrated this by applying the information over a test section. The results of this case study showcased how useful the data is, not only for the asset database but also for the purpose of making critical maintenance intervention decisions and creating a viable channel for monitoring the health of the pavement or pavement network being investigated. It should be noted that the main goal of a PMS is to treat a pavement before serious structural damage has occurred, as the system tries to avoid excessive and corrective maintenance practices; However, given the current state of road networks worldwide and the budgetary issues as exposed by the study, it is simply not possible to ensure this is done on every road in every network. Budgets and time constraints have a major say in this being possible and therefore there must be alternatives within the PMS to ensure that the planning of maintenance and rehabilitation activities are kept optimal to ensure the most value of the available resources. To this end, it is critical that authorities have enough data to make critical decisions on which roads do they carry out maintenance and which should be prioritized. The methodologies and technologies discussed in the study help to bridge this gap providing low-cost but important data for the authority. Whilst it involves monitoring of sections that are already in many cases structurally damaged it does add significantly to the road asset database held by the authority. Additionally, the processes discussed within the paper also provide for hotspot indications of places where there is only minimal severity of distresses (level 1) and this can be monitored over time to give the road authorities data on the sections most likely to suffer significant structural failure in the near future.

This allows for structural health monitoring of the pavement to be done and to be effective at a low cost. For assessments of the layers beneath the surface and the structural health of these layers, embedded technologies would have to be utilized. Future work will expand on this and try to provide more points of indication of differing levels of the structural damage to help further bolster this.

This analysis is part of an ongoing study which is developing the automation pipeline for the optimization of the PMS. Future work will focus on a more thorough assessment of the pavement distresses with regards to their occurrence, impacts on safety and comfort and the level of traffic in the network. The development of a metric for describing these distresses based on these factors will further aid in the effectiveness of the developed model. This will then take another key step towards the automation of the pavement distress surveys and asset management system.

**Author Contributions:** Conceptualization, R.R.; methodology, R.R., L.I. and G.D.M.; software, R.R.; validation, R.R. and G.D.M.; formal analysis, R.R. and G.D.M.; investigation, R.R. and G.G.; resources, G.G.; data curation, G.G.; writing—original draft preparation, R.R.; writing—review and editing, R.R., G.D.M., L.I.; visualization, R.R. and G.D.M.; supervision, R.R. and G.D.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** The research presented in this paper was carried out as part of the H2020-MSCA-ETN-2016. This project has received funding from the European Union's H2020 Program for research, technological development and demonstration under grant agreement number 721493.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References and Note

1. International Road Federation (IRF). *IRF World Road Statistics 2018 (Data 2011–2016)*; International Road Federation: Brussels, Belgium, 2018.
2. Peterson, D. *National Cooperative Highway Research Program Synthesis of Highway Practice Pavement Management Practices*, 135th ed.; Transportation Research Board: Washington, DC, USA, 1987; ISBN 0309044197.
3. Radopoulou, S.C.; Brilakis, I. Improving road asset condition monitoring. *Transp. Res. Procedia* **2016**, *14*, 3004–3012. [[CrossRef](#)]
4. Zwerling, C.; Peek-Asa, C.; Whitten, P.S.; Choi, S.W.; Sprince, N.L.; Jones, M.P. Fatal motor vehicle crashes in rural and urban areas: Decomposing rates into contributing factors. *Inj. Prev.* **2005**, *11*, 24–28. [[CrossRef](#)] [[PubMed](#)]
5. Coenen, T.B.J.; Golroo, A. A review on automated pavement distress detection methods. *Cogent Eng.* **2017**, *4*, 1374822. [[CrossRef](#)]
6. Xue, W.; Wang, D.; Wang, L. A review and perspective about pavement monitoring. *Int. J. Pavement Res. Technol.* **2012**, *5*, 295–302.
7. Merenda, M.; Praticò, F.G.; Fedele, R.; Carotenuto, R.; Corte, F.G. Della a real-time decision platform for the management of structures and infrastructures. *Electronics* **2019**, *8*, 1180. [[CrossRef](#)]
8. Arun Sundaram, B.; Ravisankar, K.; Senthil, R.; Parivallal, S. Wireless sensors for structural health monitoring and damage detection techniques. *Curr. Sci.* **2013**, *104*, 1496–1505.
9. Alavi, A.H.; Hasni, H.; Lajnef, N.; Chatti, K. Continuous health monitoring of pavement systems using smart sensing technology. *Constr. Build. Mater.* **2016**, *114*, 719–736. [[CrossRef](#)]
10. Laurent, J.; Hébert, J.F.; Lefebvre, D.; Savard, Y. Using 3D Laser Profiling Sensors for the Automated Measurement of Road Surface Conditions. In *Proceedings of the 7th RILEM International Conference on Cracking in Pavements*; Scarpas, A., Kringos, N., Al-Qadi, I., Loizos, A., Eds.; Springer: Dordrecht, The Netherlands, 2012; pp. 157–167.
11. Oliveira, H.; Correia, P.L. Automatic road crack segmentation using entropy and image dynamic thresholding. In *Proceedings of the 17th European Signal Processing Conference*, Glasgow, Scotland, 24–28 August 2009; pp. 622–626.
12. Wix, R.; Leschinski, R. Cracking—A tale of four systems. In *Proceedings of the 25th Australian Road Research Board Conference*, Perth, Australia, 23–26 September 2012; pp. 1–20.
13. Ragnoli, A.; De Blasiis, M.; Di Benedetto, A. Pavement distress detection Methods: A review. *Infrastructures* **2018**, *3*, 58. [[CrossRef](#)]
14. Inzerillo, L.; Di Mino, G.; Roberts, R. Image-based 3D reconstruction using traditional and UAV datasets for analysis of road pavement distress. *Autom. Constr.* **2018**, *96*, 457–469. [[CrossRef](#)]
15. Liu, W.; Wang, Z.; Liu, X.; Zeng, N.; Liu, Y.; Alsaadi, F.E. A survey of deep neural network architectures and their applications. *Neurocomputing* **2017**, *234*, 11–26. [[CrossRef](#)]
16. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [[CrossRef](#)]
17. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE International Conference on Computer Vision*, Washington, DC, USA, 7–13 December 2015; pp. 1026–1034.
18. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
19. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [[CrossRef](#)]
20. Canziani, A.; Culurciello, E.; Paszke, A. An analysis of deep neural network models for practical applications. *arXiv* **2016**, arXiv:1605.07678.
21. Ceylan, H.; Bayrak, M.B.; Gopalakrishnan, K. Neural networks applications in pavement engineering: A recent survey. *Int. J. Pavement Res. Technol.* **2014**, *7*, 434–444.



22. Gopalakrishnan, K. Deep Learning in data-driven pavement image analysis and automated distress detection: A review. *Data* **2018**, *3*, 28. [\[CrossRef\]](#)
23. Zantalis, F.; Koulouras, G.; Karabetsos, S.; Kandris, D. A review of machine learning and Iot in smart transportation. *Futur. Internet* **2019**, *11*, 94. [\[CrossRef\]](#)
24. Koch, C.; Georgieva, K.; Kasireddy, V.; Akinci, B.; Fieguth, P. A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure. *Adv. Eng. Inform.* **2015**, *29*, 196–210. [\[CrossRef\]](#)
25. Singh, A.P.; Sharma, A.; Mishra, R.; Wagle, M.; Sarkar, A.K. Pavement condition assessment using soft computing techniques. *Int. J. Pavement Res. Technol.* **2018**, *11*, 564–581. [\[CrossRef\]](#)
26. Attoh-Okine, N.O. Analysis of learning rate and momentum term in backpropagation neural network algorithm trained to predict pavement performance. *Adv. Eng. Softw.* **1999**, *30*, 291–302. [\[CrossRef\]](#)
27. Eldin, N.N.; Senouci, A.B. A pavement condition-rating model using backpropagation neural networks. *Comput. Civ. Infrastruct. Eng.* **1995**, *10*, 433–441. [\[CrossRef\]](#)
28. Hamdi, H.; Sigit, P.; Correia, A.G.; Pereira, P.; Cortez, P. Prediction of surface distress using neural networks. *AIP Conf. Proc.* **2017**, *1855*, 1–8.
29. Amin, M.S.R.; Amador-Jiménez, L.E. Pavement management with dynamic traffic and artificial neural network: A case study of Montreal. *Can. J. Civ. Eng.* **2015**, *43*, 241–251. [\[CrossRef\]](#)
30. Elbagalati, O.; Elseifi, M.A.; Gaspard, K.; Zhang, Z. Development of an enhanced decision-making tool for pavement management using a neural network pattern-recognition algorithm. *J. Transp. Eng. Part B Pavements* **2018**, *144*, 04018018. [\[CrossRef\]](#)
31. Janani, L.; Dixit, R.K.; Sunitha, V.; Mathew, S. Prioritisation of pavement maintenance sections deploying functional characteristics of pavements. *Int. J. Pavement Eng.* **2019**, *8*, 1–8. [\[CrossRef\]](#)
32. Terzi, S. Modeling the pavement serviceability ratio of flexible highway pavements by artificial neural networks. *Constr. Build. Mater.* **2007**, *21*, 590–593. [\[CrossRef\]](#)
33. Bianchini, A.; Bandini, P. Prediction of pavement performance through neuro-fuzzy reasoning. *Comput. Civ. Infrastruct. Eng.* **2010**, *25*, 39–54. [\[CrossRef\]](#)
34. Gu, F.; Luo, X.; Zhang, Y.; Chen, Y.; Luo, R.; Lytton, R.L. Prediction of geogrid-reinforced flexible pavement performance using artificial neural network approach. *Road Mater. Pavement Des.* **2018**, *19*, 1147–1163. [\[CrossRef\]](#)
35. Sollazzo, G.; Fwa, T.F.; Bosurgi, G. An Ann model to correlate roughness and structural performance in asphalt pavements. *Constr. Build. Mater.* **2017**, *134*, 684–693. [\[CrossRef\]](#)
36. Plati, C.; Georgiou, P.; Papavasiliou, V. Simulating pavement structural condition using artificial neural networks. *Struct. Infrastruct. Eng.* **2016**, *12*, 1127–1136. [\[CrossRef\]](#)
37. Yao, L.; Dong, Q.; Jiang, J.; Ni, F. Establishment of prediction models of asphalt pavement performance based on a novel data calibration method and neural network. *Transp. Res. Rec.* **2019**, *2673*, 66–82. [\[CrossRef\]](#)
38. Rakesh, N.; Jain, A.K.; Reddy, M.A.; Reddy, K.S. Artificial neural networks—Genetic algorithm based model for backcalculation of pavement layer moduli. *Int. J. Pavement Eng.* **2006**, *7*, 221–230. [\[CrossRef\]](#)
39. Goktepe, A.B.; Agar, E.; Lav, A.H. Advances in backcalculating the mechanical properties of flexible pavements. *Adv. Eng. Softw.* **2006**, *37*, 421–431. [\[CrossRef\]](#)
40. Shafabakhsh, G.H.; Ani, O.J.; Talebsafa, M. Artificial neural network modeling (ANN) for predicting rutting performance of nano-modified hot-mix asphalt mixtures containing steel slag aggregates. *Constr. Build. Mater.* **2015**, *85*, 136–143. [\[CrossRef\]](#)
41. Gopalakrishnan, K.; Khaitan, S.K.; Choudhary, A.; Agrawal, A. Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection. *Constr. Build. Mater.* **2017**, *157*, 322–330. [\[CrossRef\]](#)
42. Chatterjee, S.; Saeedfar, P.; Tofangchi, S.; Kolbe, L. Intelligent road maintenance: A machine learning approach for surface defect detection. *ECIS 2018 Proc. Res. Pap.* **2018**, *194*, 1–16.
43. Nhat-Duc, H.; Nguyen, Q.L.; Tran, V.D. Automatic recognition of asphalt pavement cracks using metaheuristic optimized edge detection algorithms and convolution neural network. *Autom. Constr.* **2018**, *94*, 203–213. [\[CrossRef\]](#)
44. Zhang, L.; Yang, F.; Daniel Zhang, Y.; Zhu, Y.J. Road crack detection using deep convolutional neural network. In Proceedings of the International Conference on Image Processing, ICIP, Phoenix, AR, USA, 25–28 September 2016; pp. 3708–3712.

45. Loprencipe, G.; Pantuso, A. A specified procedure for distress identification and assessment for urban road surfaces based on PCI. *Coatings* **2017**, *7*, 65. [\[CrossRef\]](#)
46. Saar, T.; Talvik, O. Automatic asphalt pavement crack detection and classification using neural networks. In Proceedings of the BEC 2010—2010 12th Biennial Baltic Electronics Conference, Tallinn, Estonia, 4–6 October 2010; pp. 345–348.
47. Li, B.; Wang, K.C.P.; Zhang, A.; Yang, E.; Wang, G. Automatic classification of pavement crack using deep convolutional neural network. *Int. J. Pavement Eng.* **2018**, 1–7. [\[CrossRef\]](#)
48. Fan, Z.; Wu, Y.; Lu, J.; Li, W. Automatic pavement crack detection based on structured prediction with the convolutional neural network. *arXiv* **2018**, arXiv:1802.02208.
49. Liu, Y.; Yao, J.; Lu, X.; Xie, R.; Li, L. DeepCrack: A deep hierarchical feature learning architecture for crack segmentation. *Neurocomputing* **2019**, *338*, 139–153. [\[CrossRef\]](#)
50. Cha, Y.J.; Choi, W.; Büyüköztürk, O. Deep learning-based crack damage detection using convolutional neural networks. *Comput. Civ. Infrastruct. Eng.* **2017**, *32*, 361–378. [\[CrossRef\]](#)
51. Zhang, A.; Wang, K.C.P.; Fei, Y.; Liu, Y.; Chen, C.; Yang, G.; Li, J.Q.; Yang, E.; Qiu, S. Automated pixellevel pavement crack detection on 3d asphalt surfaces with a recurrent neural network. *Comput. Civ. Infrastruct. Eng.* **2019**, *34*, 213–229. [\[CrossRef\]](#)
52. Tong, Z.; Gao, J.; Han, Z.; Wang, Z. Recognition of asphalt pavement crack length using deep convolutional neural networks. *Road Mater. Pavement Des.* **2018**, *19*, 1334–1349. [\[CrossRef\]](#)
53. Oliveira, H.; Correia, P.L. CrackIT—An image processing toolbox for crack detection and characterization. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 798–802.
54. Eisenbach, M.; Stricker, R.; Seichter, D.; Amende, K.; Debes, K.; Sesselmann, M.; Ebersbach, D.; Stoeckert, U.; Gross, H.M. How to get pavement distress detection ready for deep learning? A systematic approach. In Proceedings of the International Joint Conference on Neural Networks, Anchorage, AK, USA, 14–19 May 2017.
55. Seichter, D.; Eisenbach, M.; Stricker, R.; Gross, H.M. How to improve deep learning based pavement distress detection while minimizing human effort. In Proceedings of the 2018 IEEE 14th International Conference on Automation Science and Engineering (CASE), Munich, Germany, 20–24 August 2018; pp. 63–70.
56. Ciaparrone, G.; Serra, A.; Vito, C.; Finelli, P.; Scarpato, C.A.; Tagliaferri, R. A deep learning approach for road damage classification. In *Advanced Multimedia and Ubiquitous Engineering*; Springer: Singapore, 2018; pp. 655–661.
57. Geiger, A.; Lenz, P.; Stiller, C.; Urtasun, R. Vision meets robotics: The KITTI dataset. *Int. J. Robot. Res.* **2013**, *32*, 1231–1237. [\[CrossRef\]](#)
58. Maeda, H.; Sekimoto, Y.; Seto, T.; Kashiyama, T.; Omata, H. Road damage detection and classification using deep neural networks with smartphone images. *Comput. Civ. Infrastruct. Eng.* **2018**, *33*, 1127–1141. [\[CrossRef\]](#)
59. Wang, Y.J.; Ding, M.; Kan, S.; Zhang, S.; Lu, C. Deep proposal and detection networks for road damage detection and classification. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 5224–5227.
60. Alfarrarjeh, A.; Trivedi, D.; Kim, S.H.; Shahabi, C. A Deep learning approach for road damage detection from smartphone images. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 5201–5204.
61. Singh, J.; Shekhar, S. Road damage detection and classification in smartphone captured images using mask R-CNN. *arXiv* **2018**, arXiv:1811.04535.
62. Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; Isard, M.; et al. Tensorflow: A system for large-scale machine learning. In Proceedings of the 12th USENIX Conference on Operating Systems Design and Implementation, Savannah, GA, USA, 2–4 November 2016.
63. Tzutalin. 2015. LabellImg (version 1.8.3). Gitcode.
64. Everingham, M.; Eslami, S.M.A.; Van Gool, L.; Williams, C.K.I.; Winn, J.; Zisserman, A. The pascal visual object classes challenge: A retrospective. *Int. J. Comput. Vis.* **2014**, *111*, 98–136. [\[CrossRef\]](#)
65. ASTM. *ASTM D 6433-18 Standard Practice for Roads and Parking Lots Pavement Condition Index Surveys*; ASTM International: West Conshohocken, PA, USA, 2018.

66. British Columbia Ministry of Transportation and Infrastructure Construction Maintenance Branch. *Pavement Surface Condition Rating Manual*, 5th ed.; British Colombia Ministry of Transportation and Infrastructure: Victoria, BC, Canada, 2016.
67. Lombardia, R.; e Mobilità, D.G.I. *Catalogo Dei Dissesti Delle Pavimentazioni Stradali*; Direzione Generale Infrastrutture e Mobilita: Milan, Italy, 2005.
68. Paterson, W. International roughness index: Relationship to other measures of roughness and riding quality. *Transp. Res. Rec.* **1986**, *1084*, 49–59.
69. Huang, J.; Rathod, V.; Sun, C.; Zhu, M.; Korattikara, A.; Fathi, A.; Fischer, I.; Wojna, Z.; Song, Y.; Guadarrama, S.; et al. Speed/accuracy trade-offs for modern convolutional object detectors. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, 22–25 July 2017.
70. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. SSD: Single shot multibox detector. In Proceedings of the European Conference on Computer Vision 2016, Amsterdam, The Netherlands, 8–16 October 2016; Springer: Cham, Switzerland, 2016.
71. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
72. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.
73. Dai, J.; Li, Y.; He, K.; Sun, J. R-FCN: Object detection via region-based fully convolutional networks. In Proceedings of the Advances in Neural Information Processing Systems, Barcelona, Spain, 5–10 December 2016; pp. 379–387.
74. 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, Montreal, QC, Canada, 7–12 December 2015; Volume 39, pp. 1137–1149.
75. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft COCO: Common objects in context. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; Springer: Cham, Switzerland, 2015; Volume 8693, pp. 740–755.



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).