Journal of
Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7765578

# IMPROVE CUSTOM OBJECT DETECTION OF ROAD SURFACES BY SMOOTHING OF LABELS AND IMAGE ENHANCEMENT

**R. Palani**
Research Scholar, Department of Computer and Information Science, Annamalai University, Chidambaram, Tamil Nadu, India

**Dr. N. Puviarasan**
Professor and Head, Department of Computer and Information Science, Annamalai University, Chidambaram, Tamil Nadu, India

**Dr. A. Rama Prasath**
Associate Professor, School of Computing Sciences, Hindustan University, Chennai, Tamil Nadu, India

**Abstract**:
Roads play a vital role in transporting the passengers/goods for short and medium distances. The objective of this research is to use video or image sources to detect road surface damage. Utilizing Image Augmentation and Annotation Label Smoothing techniques can increase the effectiveness and precision of spotting damage to the road's surface. This research paper includes investigating data preparation techniques, annotating data labels, and providing a road damage dataset with a training model. Tamil-Nadu-Road-Surface-Dataset-2023(TNRSD2023) is a road surface dataset that was compiled in the area surrounding the Cuddalore district in Tamil Nadu, India by the author.
Keywords: Road Surface damage detection, Image Augmentation, Label Smoothing

## 1. Introduction

Roads are significant in India for economic development, connectivity, tourism, employment, and safety. They play a crucial role in connecting cities, towns, and villages, facilitating the movement of goods and people, promoting trade and commerce, and bringing remote areas into the mainstream. Additionally, roads are crucial for the tourism industry and provide employment opportunities to millions of people in India. Good roads also provide safe transportation, reducing accidents and fatalities. As a result, the government is investing heavily in building new roads and upgrading existing ones to support India's growth and development.

### 1.1 Possibilities for Road Damages

Roads can suffer damage due to various factors, including natural disasters such as floods, earthquakes, and landslides. Other factors that can cause damage to roads include heavy traffic, overloading of vehicles, poor road design, inadequate maintenance, and environmental factors such as extreme heat, cold, or moisture. The damage to roads can range from minor wear and tear to significant structural damage, such as cracks, potholes, and pavement failures. When left unaddressed, road damage can worsen over time, resulting in more costly repairs, safety hazards, and increased travel time for commuters. Therefore, it is essential to maintain and repair roads regularly to ensure safe and efficient transportation.

## 1.2 Road damages in India with statistical references

Road damages in India have significant impacts on various aspects of the economy and society. Here is a summary of the effects of road damages in India with statistical information and references. Impact denotes the immediate and direct results of an activity, whereas effect denotes the longer lasting or indirect effects that follow from the action.

Impact on account of Road damages:

- Traffic congestion: Traffic congestion caused by poor road conditions in India's major cities results in an annual cost of about $22 billion to the Indian economy. (Source: Boston Consulting Group, "Unlocking India's Potential: The Economic Impact of Improved Infrastructure," 2016)
- Road accidents: Road damages contribute to an increase in road accidents. In 2019, there were over 4.5 lakh road accidents in India, resulting in more than 1.5 lakh deaths and over 4.5 lakh injuries. (Source: National Crime Records Bureau, "Accidental Deaths and Suicides in India," 2019)

Effects on account of Road damages:

- Economic cost: Poor Road conditions in India result in an estimated economic cost of around 3.5% of the country's GDP. (Source: Ministry of Road Transport and Highways, "Economic Cost of Bad Roads," 2018)
- Health impact: Dust and pollution from damaged roads can lead to respiratory problems and other health issues, particularly in densely populated urban areas. (Source: Hindustan Times, "Craters in Roads Pose Health Hazards for Delhiites," 2017)
- Vehicle maintenance: Poor Road conditions lead to increased wear and tear on vehicles, resulting in higher maintenance costs for vehicle owners. (Source: The Economic Times, "Potholes Cost Indian Car Owners Billions Every Year," 2019)
- These statistics highlight the significant economic, social, and health impacts of road damages in India, emphasizing the need for regular maintenance and repairs to ensure safe and efficient transportation.

## 1.3 Aim of this Research

There are currently established methods for identifying and measuring the various forms of road damages. The purpose of this study is to improve current practices by employing Computer Vision technology.

Computer vision is a field of artificial intelligence and computer science that focuses on enabling machines to interpret and analyse visual information from the world. It involves the development of algorithms and techniques for tasks such as object detection, facial recognition, image classification, and gesture recognition.

1.4 Object Detection

Object detection is a computer vision task that involves detecting and localizing objects within an image or video. It typically involves using pre-trained models that can detect common objects such as people, cars, animals, and so on. These pre-trained models are trained on large datasets, such as ImageNet, that contain a wide range of objects and scenes.

Custom object detection, on the other hand, involves training a machine learning model to detect specific objects or classes of objects that are not present in pre-trained models. This is often necessary in scenarios where the objects of interest are uncommon or have unique features that are not well-represented in the pre-trained models.

Object detection in road damage surfaces refers to the process of automatically identifying and localizing different types of damage on road surfaces, such as potholes, cracks, and other types of defects. This is typically done using computer vision techniques and algorithms that can analyse images or videos of the road surface captured by cameras mounted on vehicles.

Object detection algorithms for road damage surfaces typically involve several stages, including image pre-processing, feature extraction, object localization, and classification. In the pre-processing stage, the image is prepared by removing noise and enhancing features of interest. Then, features are extracted from the image, such as edges or corners, which can be used to localize potential damage locations.

Next, the algorithm performs object localization, which involves identifying the precise location of damage in the image. This is usually done by segmenting the image and determining the boundaries of the damaged area. Finally, the algorithm classifies the type of damage based on its features, such as its size, shape, and colour.

Custom Object detection in road damage surfaces can be used to improve road maintenance and repair, as it enables more accurate and efficient detection of damage. This can help reduce accidents and improve road safety for drivers, cyclists, and pedestrians.

1.5 Pros and Cons of Stage detectors

One-stage and two-stage object detection are two different approaches to object detection in computer vision.

One-stage detectors, such as YOLO (You Only Look Once) and SSD (Single Shot Detector), are faster and simpler models that detect objects in a single pass through the image. They generally use a convolutional neural network (CNN) to generate a set of anchor boxes at different scales and aspect ratios, and then predict object scores and bounding box coordinates for each anchor box.

Two-stage detectors, such as Faster R-CNN (Region-based Convolutional Neural Network), are more complex models that use a two-stage process to detect objects. They first generate object proposals using a region proposal network (RPN) and then classify and refine the proposals using a second stage network. This approach is more accurate but slower than one-stage detectors.

Some studies have shown that two-stage detectors like Faster R-CNN and its variants have higher accuracy for object detection tasks, while one-stage detectors like YOLO and SSD are faster and better suited for real-time applications with speed requirements.
Overall, choosing the best detection method involves considering factors such as accuracy, speed, complexity, and hardware requirements. It is often a trade-off between these factors and depends on the specific application.

## 2. Related works

During the last few years, many datasets of road damages were released and made publicly available; some of these datasets are being used extensively by other researchers, such as the RDD2018 dataset released by Maeda et al. [4], which contained around 9000 images of eight types of road damage in Japan. The RDD2020 by Arya et al. [5] included around 26,000 images of eight classes of damage collected from the roads of India, Japan, and the Czech Republic. In general, road damage datasets can be divided into two types based on their uses. Detection datasets that are used primarily for binary classification such as the AigleRN dataset collected from France by Amhaz et al. [6], CFD collected from China released by Shi et al. [7], Temple University dataset collected from the United States by Zhang et al. [8], Denish Technological Institute dataset collected from Denmark by Silva et al. [9], and METU gathered from Turkey released by Özgenel [10]. The other type is the classification datasets which usually contain more than two positive classes, such as the Taiwan and Japan datasets released by Chen et al. [11], the Czech and Slovak datasets released by Mraz et al. [12], and many others.

Machine learning (ML) algorithms have been used for more than two decades to detect road damage, such as the work completed by Hoang et al. [13], Song et al. [14], and Hoang et al. [15]. Numerous studies have been proposed to detect potholes and cracks using edge detection and image thresholding, for instance, Otsu et al. [16], Ayenu-Prah et al. [17], and Koch et al. [18]. Many studies developed methods to detect cracks utilizing random structured forests, e.g., Shi et al. [7], and an unsupervised method based on Otsu's thresholds and photo-metric information, e.g., Akagic et al. [19]. The ML methods used to detect potholes detection included unsupervised fuzzy c-means clustering and morphological reconstruction, e.g., Ouma et al. [20], and Support Vector Machine (SVM), e.g., Marques et al. [21]. Furthermore, Ahmadi et al. [22] and Cubero-Fernandez et al. [23] implemented several ML classification methods to classify four types of road damage, including K-nearest neighbors (KNN), Bagged Trees, SVM, and Decision Tree.

In recent years, deep learning (DL) has been widely used to detect road damage, for instance, Biçici et al. [24], Stricker et al. [25], and Zhang et al. [26]. These algorithms and techniques are being used now in self-driving cars to avoid obstacles and ensure road safety while driving. Most research uses detection tasks that only discover the damage. In many studies presented by Zhang et al. [8], Silva and Lucena [9], Rao et al. [27], and Fan et al. [28], several Convolutional Neural Network (CNN) models were presented to detect cracks in road images. Despite the good results that were achieved, the detection method only determines the presence of the damage; still, it does not classify its type. As a result, in recent studies, classification methods based on DL algorithms have been used on input images to classify them into various types of damage that can assist municipalities in accurately identifying and classifying damage. For instance, Ebenezer et al. [29] and Elghaish et al. [30] presented a method to detect four types of damage to the road using several CNN models.

Despite the good results that were achieved in the previous studies, the method detects only four types of road damage. However, a number of methods were developed to classifiy more than six types of road damage using different techniques. Maeda et al. [31] and Mraz et al. [12] proposed a method for road damage detection and classification based on Convolutional Neural Networks (CNNs). The model was trained using the SSD MobileNet and SSD Inception V2 frameworks. Singh et al. [32], Vishwakarma et al. [33], Kortmann et al. [34] and Wang et al. [35] proposed an automatic road damage detection and classification based on Convolutional Neural Network (CNN) using an R-CNN model. In the last two years, some studies have used advanced algorithms such as a single stage that uses a single CNN, such as YOLO, to predict the class and location of damage directly: for instance, Doshi and Yilmaz [36], Alfarrarjeh et al. [37], Jeong [38], Hegde et al. [39], Pena-Caballero [40] and Al-Shaghouri [41].

Although many studies have proposed approaches to automate the detection of road damages, several problems remain, and there is still room for more improvement. New methods could be proposed to classify more types of road damage more precisely, especially those present in the Kingdom of Saudi Arabia, whose nature of damage may differ from other countries from which the available datasets were collected in terms of geography and climate. Table 1 summarizes the studies conducted to detect and classify road damage using ML and DL methods, and as shown, the most used ML method was SVM. CNN was the most commonly used for DL.

| Research | Method | Dataset | Country | Classes | Result |
|---|---|---|---|---|---|
| [12] | ANN | 3923 images | Czech, Solvak | Cracks, potholes, rutting, bump, separation, and deteriorated markings | F1 score 67 |
| [32] | Mask-RCNN | RDD2018 (9053 images) | Japan | Cracks, rutting, bump, pothole, separation, and blur | F1 score 0.528 |
| [33] | Faster R-CNB | RDD2020 (26,322 images) | Japan, Czech, India | Longitudinal cracks, transverse Cracks, alligator cracks and potholes | F1 score of 0.542–0.536 |
| [34] | FRCNN | RDD2020 (26,322 images) | Japan, Czech, India | Longitudinal cracks, transverse cracks, alligator cracks and potholes | F1 score of 0.487 |
| [35] | Faster R-CNN | RDD2018 (9053 images) | Japan | Cracks, rutting, bump, pothole, separation, and blur | Mean F1 score 0.6255 |

| | | | | | |
|---|---|---|---|---|---|
| [41] | YOLOv4 | Udemy website, Lebanon images + Internet images | Lebanon, undefined | Potholes/Non potholes | Recall 81% Precision 85% |
| [36] | YOLOv4 | RDD2020 (26,336 images) | India, Czech, and Japan | Longitudinal cracks, transverse cracks, alligator cracks and potholes | F1 scores 0.628 and 0.6358. |
| [37] | YOLO | RDD2018 (9053 images) | Japan | Cracks, rutting, bump, pothole, separation, and blur | 1 score 0.62 |
| [39] | u-YOLO | RDD2020 (26,336 images) | India, Czech, and Japan | Longitudinal cracks, transverse cracks, alligator cracks and potholes | F1 score of up to 0.67 |
| [40] | YOLO | 4000 images | USA | Longitudinal, lateral, and alligator crack, pothole, manhole, blurred streed line, blurred crosswalk | mAP 97.98% |
| [38] | YOLOv5 | RDD2020 (26,336 images) | India, Czech, and Japan | Longitudinal cracks, transverse cracks, alligator cracks and potholes | F1 score of 0.58 |

**Table 1.** Summary of ML and DL methods.

**Data Collection**

The road surface video was taken on roads along Tamil Nadu's Cuddalore and Chengalpattu districts' National Highways (NH) and Tamil Nadu State Highways (TNSH). The video was captured with a Redmi5 and an iPhone12 mobile phones. The video file play time is restricted to ten minutes in order to prevent play mistake. Also, some of the pictures were acquired from Google Images and other road datasets [54]. The ethics of filming video are addressed. The video /images are stored in Google Drive.

The images were in JPG format with various resolutions, and they were scaled to 640 by 640 to create square images and maintain the dataset uniform. Also, a total of 1640 images were taken in a variety of weather conditions, including sunny, rainy, and winter. The images were also captured at various angles. Just 1129 images were found suitable for the study. The remaining images were rejected because they were either poorly coloured or did not clearly show road damage.

Figure 1. State Highway Road Damage Dataset Sample.

### 3.1 Damage Classification

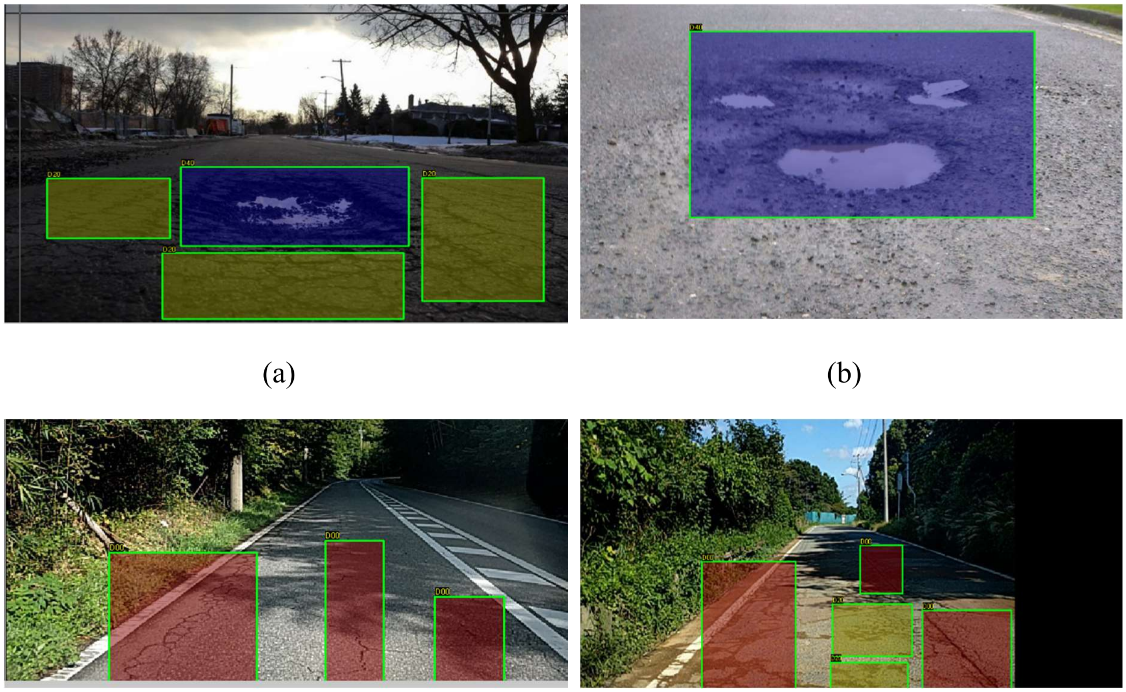
In an international study report [13], the different types of road damage are categorized. Only four categories of damages, which are listed in table 1, are detected in this research article. which is significant road damage common on Indian highways. This research does not address certain damage kinds, such as longitudinal construction joint part (D01), lateral construction joint part (D11), Cross walk blur (D43), and White line blur (D43). The damages are represented by image annotation labels which are discussed below.

| Damage Type | | | Detail | Class Name |
|---|---|---|---|---|
| Crack | Linear Crack | Longitudinal | Wheel mark part | D00 |
| | | Lateral | Equal interval | D10 |
| | Allegator | | Partial Pavement, All Pavement' | D20 |
| Other Corruption | | | Rutting, bump, pothole, Separation | D40 |

Table 2: Classification of Road surface Damages

### 3.2 Data Annotation

Data annotation is the process of adding metadata or labels to a dataset to make it understandable and useful for machine learning algorithms. The dataset had to be checked to make sure all the images were valid, free of errors, and named properly before the annotation could begin, which took a lot of time. The images were annotated, except for unclear images, and the appropriate categories were selected; Table 1 lists four different damage categories. To make the images more accessible and usable, they were organized into folders and saved to Google Drive.



(a) (b)

( c) (d)

Figure 2. Damage class are marked with different colours.

3.3 Data statistics

The statistics for the Tamil Nadu Road datasets are shown in Figure 2. It should be noticed that datasets have an uneven distribution of occurrences for different damage classes.  image augmentation techniques were used to create a balanced representation that can be used to train deep learning models.

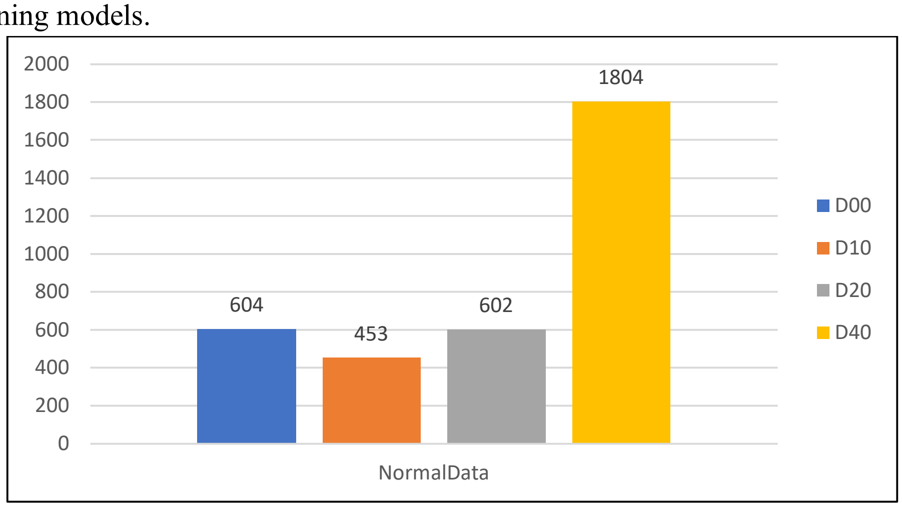



**Figure 3.** Statistics about the total of damage cases inside the underlying datasets.

## 3. Methodology

The proposed classification model was divided into five primary stages, as illustrated in Figure 3: data collection, data pre-processing and augmentation, implementation of several CNN models, experiment, and assessment. The damages were classified according to their types, pre-training models, and the proposed self-designed CNN model. Finally, this work evaluates and analyses the obtained results using the state-of-the-art performance metrics; the results are presented in Section 6.

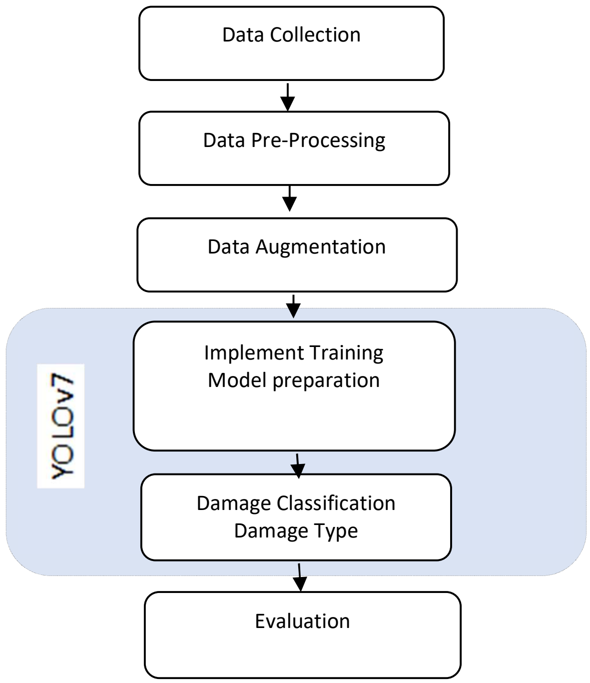



**Figure 4.** Steps for proposed methods.

4.1 Data Pre-processing

The sources images are various dimension which are mentioned below. Image size refers to the physical dimensions of an image. Scaling refers to the process of changing the size of an image while preserving its aspect ratio. Scaling can be done by either increasing or decreasing the number of pixels in the image. This paper includes a dataset of images of various sizes, as shown in Table 3 .

| Image sources | sizes |
|---|---|
| Redmi5 - mobile | 854 x 480 |
| Iphone12 -Video file | 1280 x 780 |
| Google images | 268 X 188, 262 x 192, 271 x 186 |
| GRDD2022 - images | 600 x 600 |

Table 3. Dataset video file / image size

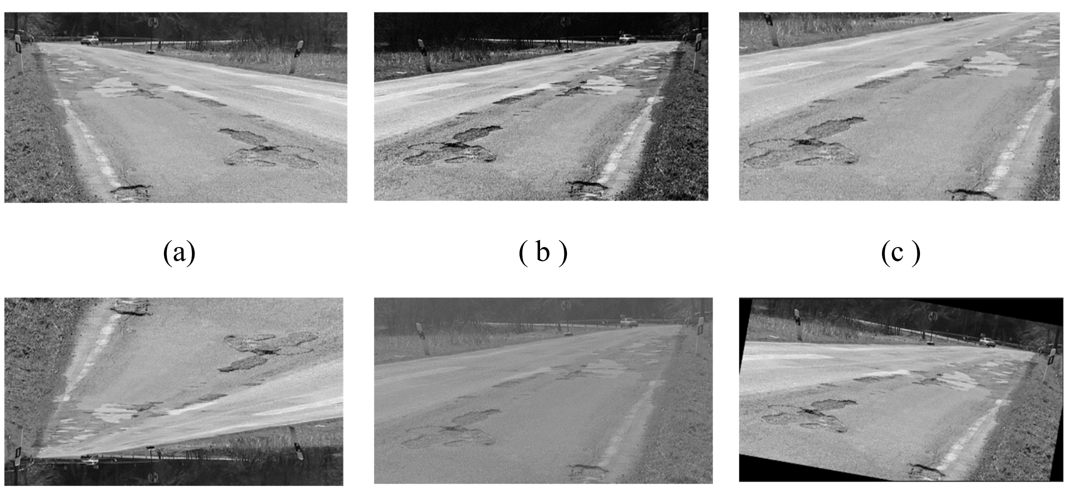
4.2 Data Augmentation

Deep networks require a large amount of training data to work well. Image augmentation is usually required to build a robust image classifier with very little training data [46]. In image augmentation, training images are artificially created using various processing methods or a combination of them, such as rotation, noise, translation, mirroring, and blurring [47]. The original dataset included 1129 images. After applying the augmentation techniques, it was divided into two sets; the first set contained 80% of the images and was used for training, and the second set contained 20% of the images and was used for testing. In this work, we used four different types of augmentation techniques to create new images, including rotate, flip, zoom, and brightness. Image rotation is one of the most used augmentation techniques.

| Classes | # Before Augment | # After Augment |
|---|---|---|
| **D00** | 726 | 1316 |
| **D10** | 526 | 1013 |
| **D20** | 707 | 1450 |
| **D40** | 1908 | 3382 |
| **No. of Total Images** | 1129 | 2159 |
| **No. Rejected Images** | 538 | 711 |

Table 4. Statistical of Normal and Augmented Images

This research has challenges during the augmentation, for classes D00 & D10, some augmentations, such as flip rotation and left-to-right rotation, are invalid. The results are distorted when class D00 is turned to become class D10. Class D20 and D40 are exempt from image augmentation, nevertheless. All classes can benefit from contrast augmentation, which is a component of image augmentation. By enhancing 1129 base images, 2159 augmented images with 1791 bounding boxes are generated. As illustrated in Figure 4, these techniques were applied to each image to obtain a new training and testing sample.



(a) ( b ) (c ) ( d ) ( e ) ( f )

**Figure 5.** Example of augmented images: (**a**) Contrast, (**b**) Contrast sigmoid (**c**) Crop (**d**) Left to right (**e**) linear contrast (f) rotation. In degree.

### 4.3 Splitting Data

Splitting data is an essential element of data science, especially for building accurate models based on machine learning to avoid overfitting. Typically, data is split into two or more subsets. When split into two parts, one part is used to evaluate or test the models and the other is used to train the models. Therefore, to evaluate the performance of a deep neural network model in an unbiased manner, a dataset properly divided into training and testing sets is required. In this proposed work, the data was split according to the standard 80:20 ratio, where 80% of the dataset was used for training and 20% for evaluation (testing and validation), using python script. Consequently, the performance of the model was measured based on 20% of the data that was neither used in training nor previously seen by the model to ensure that the analysis was fair. The training data statistical show in table 3.

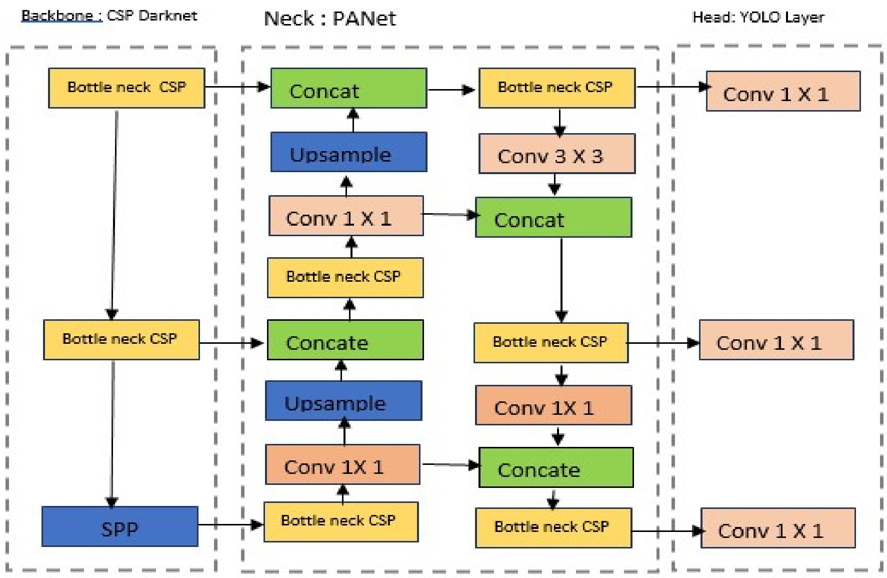
### 4.4 Architecture of YOLO

In the YOLO object detection architecture, the backbone, neck, and head are important components that work together to detect objects in an image.

The backbone is the main feature extractor in the YOLO architecture. It is usually a deep convolutional neural network (CNN) that is used to extract features from the input image. The backbone network is typically pre-trained on a large dataset such as ImageNet, and then fine-tuned on the object detection task.

The neck is a network that connects the backbone to the head. It typically consists of a few layers of convolutional neural networks that are used to merge features from the backbone and prepare them for object detection. The neck is responsible for refining the feature maps from the backbone by applying various techniques such as skip connections, pooling, and upsampling.

The head is the final part of the YOLO architecture, which is responsible for detecting objects in an image. The head usually consists of a few convolutional layers that are used to predict bounding boxes and class probabilities for each object in the image. The head takes the features from the neck and applies a set of fully connected layers to make the final predictions.




CSP: Cross Stage Partial Network Conv: Convolutional layer SPP: Spatial Pyramid Pooling Concat: Concatenate function

Figure 6. Network architecture of YOLO

4.5 Object Training

YOLO is an end-to-end architecture that can detect objects in an image and simultaneously classify them into predefined categories, all in a single pass. YOLO divides an input image into a grid of cells (8 X 8) and predicts bounding boxes 'B' and class probabilities for each cell. Each bounding box prediction includes the coordinates of the object's centre, width($B_w$), and height($B_h$). The class probabilities($P_c$) indicate the likelihood of the object belonging to each of the predefined classes(C1, C2, C3 and C4).

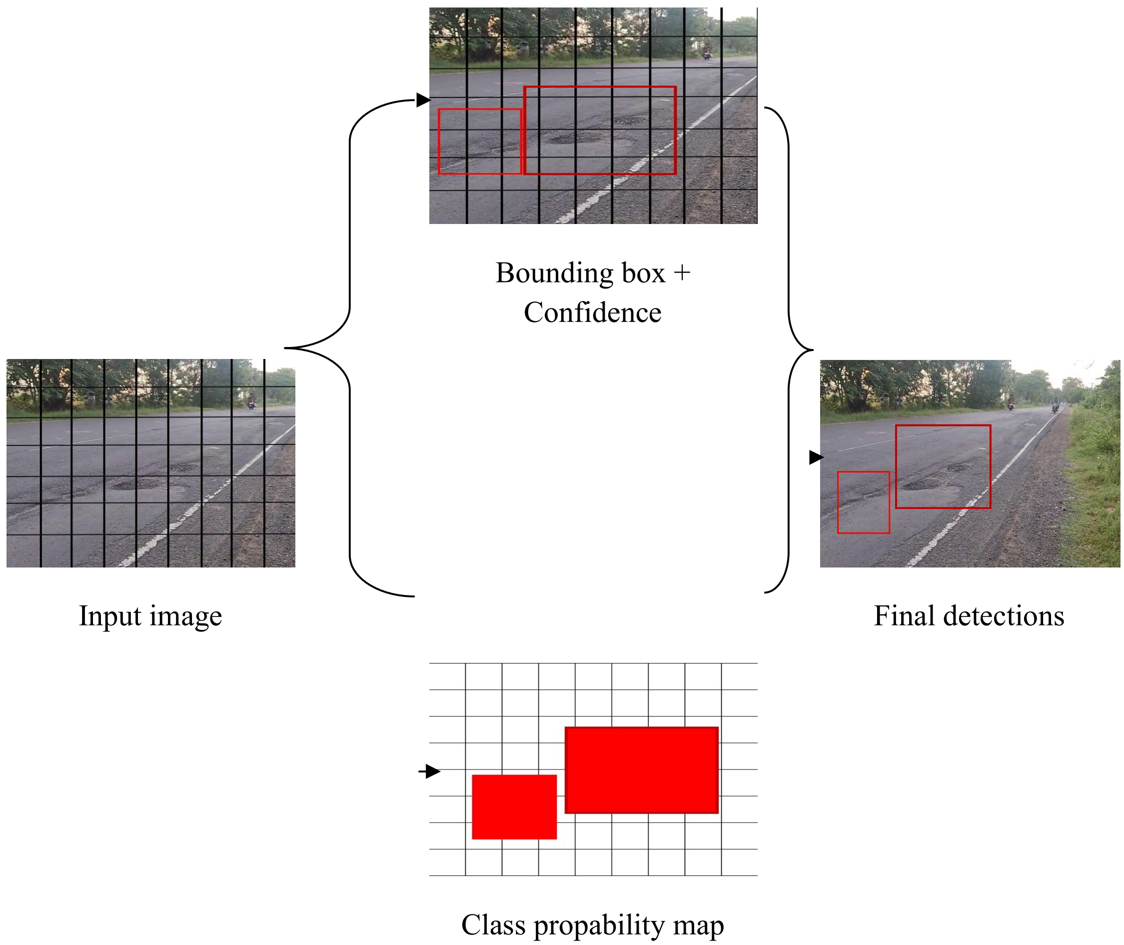



Figure 7. Steps of Object Detection

YOLO uses a single convolutional neural network (CNN) to process the entire image, which makes it very fast compared to other object detection architectures.

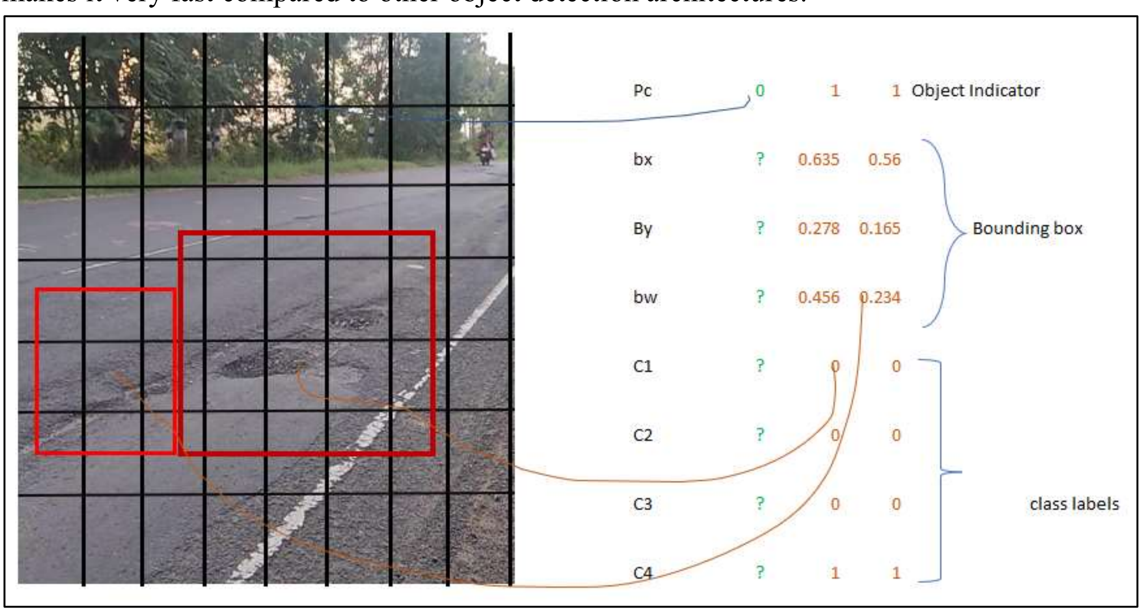


Figure 8. Network architecture of YOLO

4.6 Training model

Developed two training models for TNRSD2023, augmented images of TNRSD2023. This training model was prepared in the Colab Research CUDA environment. In section 5, the environmental setup is described. A total of 5 models have been created in this study based on varying parameters for machine learning rate and layer models, which are listed in table 5. Training model metrics are captured in wandb which is metrics analytics database. The model precision, recall, F1 score, bounding box loss, learning rate etc., details are available in Git repository with public access for reference or anyone utilising further. URL: https://github.com/PalaniRamu/TNRSDD2023.git

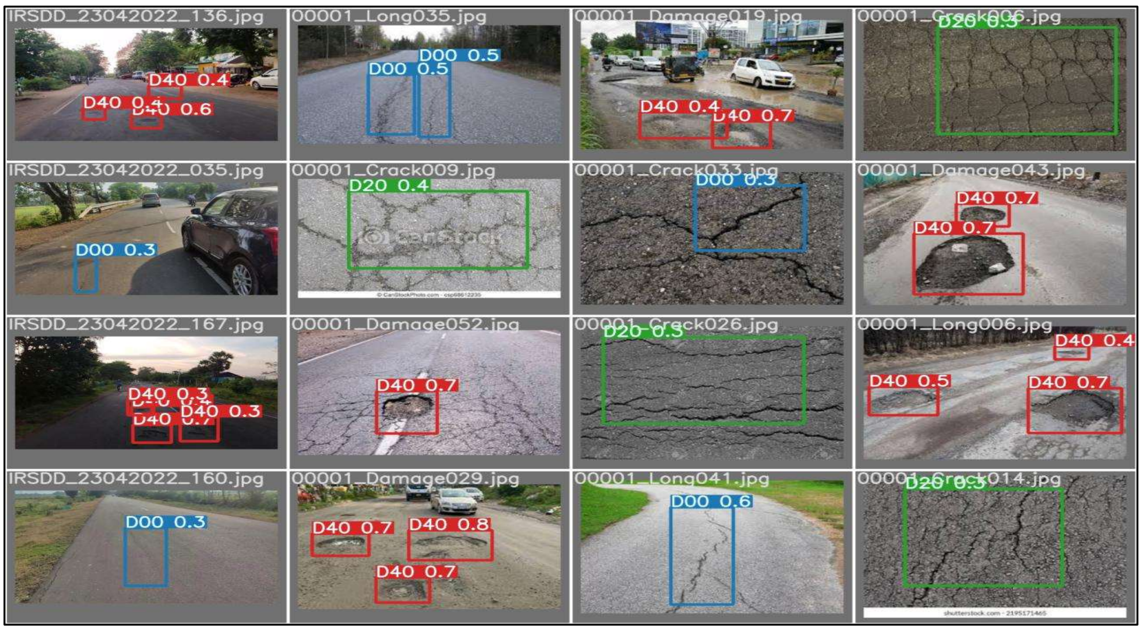



Figure 9 Class Objects prediction during the model training in epoch 5

The metrics data's are in wandb, it can re-play anytime anywhere. The trained models were metrics are compared in visual form.


Figure 10. Augmented images by python script for Object detection

4.7 Label smoothing

The annotated labels are not entirely accurate for several reasons. For instance, many traverse cracks may be mistaken for longitudinal cracks when viewed from different angles or with a slight camera rotation. Additionally, traverse, and longitudinal cracks are frequently mistaken for alligator cracks. Therefore, in this situation, the label smoothing [17] technique aids in performance improvement. Specifically, cross entropy is used to calculate classification loss from classification heads (classifying road damages into one of the four considered crack types) using the following formula:

$$\text{Loss} = \sum_{c=1}^{C} \left(-y_c \log(\hat{y}c) - (1 - y_c)\ \log(1 - \hat{y}_c)\right) \qquad (1)$$

where C is the number of classes, $\hat{y}c$ is the predicted probability of class label c and $y_c$ is calculated as:

$$y_c = \begin{cases} 1\ \textit{if the item is of class C} \\ 0\ \textit{Otherwise} \end{cases} \qquad (2)$$

This way of setting the value for $y_c$ is called hard label (100% sure that the label is of class c). However, label smoothing allows having a smoothing constant ∈ that softens this confidence:

$$y_c = \begin{cases} 1-\in+\in/C \text{ if the item is of class c} \\ \in/C\ \textit{Otherwise} \end{cases} \qquad (3)$$

The common value for ∈ is 0.1, which is experimented with in this project. Notably, the values of all $y_c$ over all classes still add up to one ($\sum_{c}^{C} y_c = 1$).

4.7.1 Experimental label smoothing:

There are two types of label smoothing like Hard and Soft.

- **Soft label: A soft label is a score that is associated with a probability or possibility. Eg: (0.1 0.2 0.8,0.9)**

- **Hard label: A hard label fits within one of the two classifications. It has a binary structure (0 or 1)**

**Label smoothing formula is**

New_labels = original_labels * (1 – label_smoothing) + label_smoothing / num_classes

The original_labels are [0 1 2 3], label_smoothing is 0.3 and num_classes is 4

new_labels = [ 0 1 2 3 ] * (1 – 0.3) + 0.3 / 4
= [ 0 1 2 3 ] *0.7 + 0.075
= [ 0.075 0.775 1.475 2.175 ]

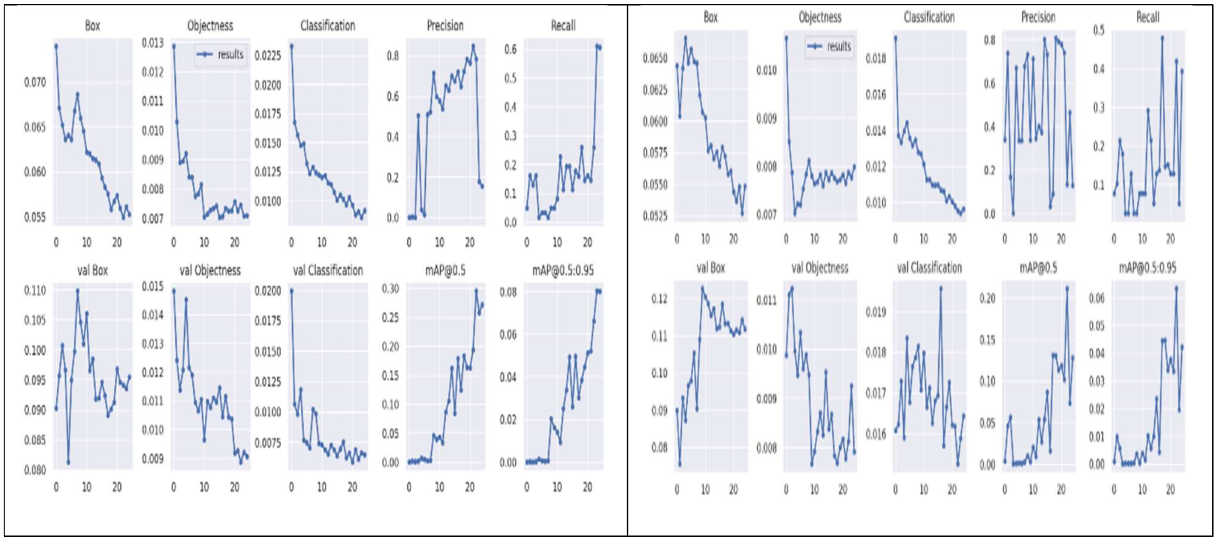
Now, the new labels will be [0.075 0.775 1.475 2.175] instead of [0 1 2 3].
The model becomes less confident with extremely confident labels. That is exactly what wanted to avoid. Now, the penalty given to a model due to an incorrect prediction will be slightly lower than using hard labels which would result in a smaller gradient.

## 4. Experimental setup

This research's data was analysed using the GPU environment of the NVIDIA Jetson nano hardware device. Since the CPU was unable to process the image data, the GPU environment was used. Most of the time training model prepared on Colab Research environment, which is another environment that is also available. Depending on availability, Google Colab gives users access to one of several different NVIDIA GPUs. This study's processing was done on a Tesla T4 processor running Torch 1.10.2+cu111 CUDA:0 with 12 GB of RAM. used Google Drive, which has a default 15 GB storage restriction, for storing files.

### 5.1 Evaluation measures

This study prepared, trained model prepared for Road surface object detection, the model has validated. The same model prepared for augmented image also. The training model result are shown in PNG format below.




( a) Normal images ( b) Augmented images

Figure 11. Result of trained model in PNG format.

Both models were verified with real time video / image files. Many metrics were used to evaluate the models' performance for the tests. Following is a definition of the formulae:

**Precision**: The formula for precision is True Positive divided by the sum of True Positive and False Positive create a class detection model and give it 1000 images to detect. If it successfully detects 500 images (TP), incorrectly detects 300 images (FP), and fails to detect 200 images (FN): The precision of 500/800, or 62.5%.

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

**Recall**: The formula for recall is True Positive divided by the sum of True Positive and False Negative

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

Using the above image example from earlier, our model would have a recall of 500/700, or 71%.

**F1 score** is the evaluation metric that is used to evaluate the performance of the machine learning model. It uses both precision and Recall, that makes it best for unbalanced dataset.

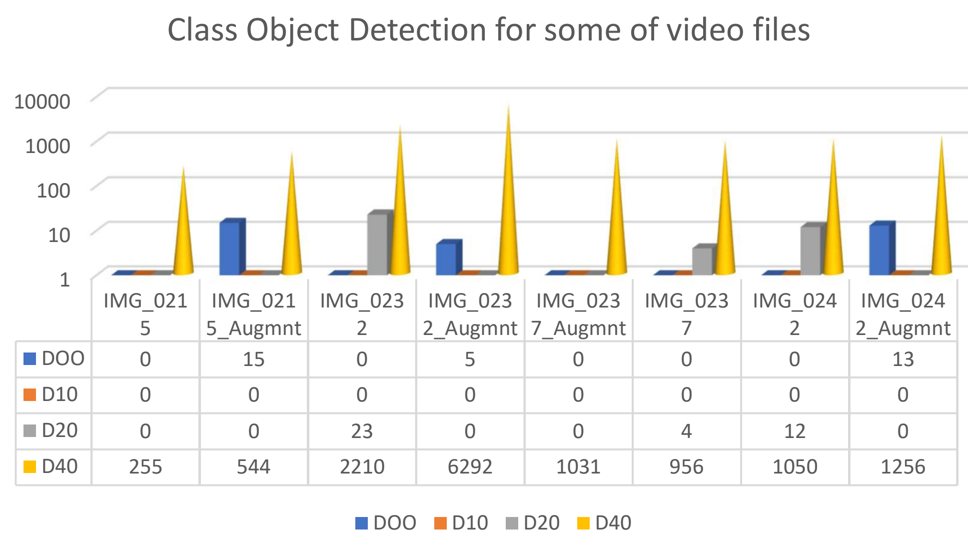
$$F1\ Score = \frac{2\ x\ Precision\ x\ Recall}{Precision + Recall} \quad \text{----} \quad (3)$$

- The value of the F1 score ranges from 0 to 1, 0: Worst Case ; 1: Best Case
- The F1 score of a model will be high/ low if both Precision and Recall have high/low values.

- The F1 score will be medium if one of the values (either Precision or Recall) is low and the other has a high value.

6. Result and discussion

In this research, the following information is suggested, the video's focus is all directions; this contrasts with the earlier suggestion[57]. Some of the class objects (D00 and D10) are not using rotation, ups, and downs enhancement. Small size of bounding boxes is not efficient in object detection. All digital device types are acceptable because object detection does not take image size into account. One of the best ways to enhance the number of source images that are useful for training the model is through image augmentation. The example When the before and after enhanced models were tested on some video clips, the after augmented model correctly predicted more class objects, as shown in figure 12.




| | IMG_0215 | IMG_0215_Augmnt | IMG_0232 | IMG_0232_Augmnt | IMG_0237_Augmnt | IMG_0237 | IMG_0242 | IMG_0242_Augmnt |
|---|---|---|---|---|---|---|---|---|
| DOO | 0 | 15 | 0 | 5 | 0 | 0 | 0 | 13 |
| D10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| D20 | 0 | 0 | 23 | 0 | 0 | 4 | 12 | 0 |
| D40 | 255 | 544 | 2210 | 6292 | 1031 | 956 | 1050 | 1256 |

Figure 13. class object detection counts for sample video using Before Augment model and After Augment model.

**7. Conclusion**

The methodology used in this paper for creating final road surface damage images can be used by the municipal government authorities for targeted road repairs. This research observation concludes that the created dataset TNRSD2023, and training models are available in git repository with public access for further investigation or usage.

**References:**

1. Khan, K.; Zaidi, S.B.; Ali, A. Evaluating the nature of distractive driving factors towards road traffic accident. Civ. Eng. J. 2020, 6, 1555–1580. [CrossRef]

2. Ministry of Transport and Logistic Services. Road Maintenance. Available online: https://mot.gov.sa/en/Roads/Pages/RoadsMaintenance.aspx (accessed on 15 September 2022).

3. Department of Local Infrastructure Development and Agricultural Roads (DoLIDAR). Road Maintenance Groups (RMG). Available online: https://www.ilo.org/dyn/asist/docs/F1259778069/rmg.pdf (accessed on 1 March 2016).

4. Maeda, H.; Sekimoto, Y.; Seto, T.; Kashiyama, T.; Omata, H. Road damage detection and classification using deep neural networks with smartphone images. Comput.-Aided Civ. Infrastruct. Eng. 2018, 33, 1127–1141. [CrossRef]

5. Arya, D.; Maeda, H.; Ghosh, S.K.; Toshniwal, D.; Sekimoto, Y. RDD2020: An annotated image dataset for automatic road damage detection using deep learning. Data Brief 2021, 36, 107133. [CrossRef] [PubMed]

6. Amhaz, R.; Chambon, S.; Idier, J.; Baltazart, V. Automatic crack detection on two-dimensional pavement images: An algorithm based on minimal path selection. IEEE Trans. Intell. Transp. Syst. 2016, 17, 2718–2729. [CrossRef]

7. Shi, Y.; Cui, L.; Qi, Z.; Meng, F.; Chen, Z. Automatic road crack detection using random structured forests. IEEE Trans. Intell. Transp. Syst. 2016, 17, 3434–3445. [CrossRef]

8. Zhang, L.; Yang, F.; Zhang, Y.D.; Zhu, Y.J. Road crack detection using deep convolutional neural network. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 3708–3712.

9. Silva,W.R.L.D.; Lucena, D.S.D. Concrete cracks detection based on deep learning image classification. Multidiscip. Digit. Publ. Inst. Proc. 2018, 2, 489.

10. Özgenel, Ç.F. Concrete Cracks Detection Based on Deep Learning Image Classification Proceedings 2018, 2, 489. [CrossRef]

11. Chen, D.R.; Chiu, W.M. Deep-Learning-Based Road Crack Detection Frameworks for Dashcam-captured Images under Different Illumination Conditions. Available online: https://europepmc.org/article/ppr/ppr369116 (accessed on 15 September 2022) . [21] R. Palani[1], Dr. N. Puviarasan[2] and Dr. A. Rama Prasath[3], Literature Review of Road Damage Detection with Repairing Cost Estimation, International Journal of Mechanical, Engineering, e-ISSN 0974-5823| Feb 2022

12. Mraz AM, A.; Sekimoto YS, Y.; Kashiyama TK, T.; Maeda HM, H. Development of the Localized Road Damage Detection Model Using Deep Neural Network. In Proceedings of the 2020 3rd International Conference on Sensors, Signal and Image Processing, Virtual, 9–11 October 2020; pp. 1–6.

13. Hoang, N.D.; Nguyen, Q.L.; Tien Bui, D. Image processing–based classification of asphalt pavement cracks using support vector machine optimized by artificial bee colony. J. Comput. Civ. Eng. 2018, 32, 04018037. [CrossRef]

14. Song, H.; Baek, K.; Byun, Y. Pothole detection using machine learning. Adv. Sci. Technol. 2018, 151 –155. [CrossRef]

15. Hoang, N.D. An artificial intelligence method for asphalt pavement pothole detection using least squares support vector machine and neural network with steerable filter-based feature extraction. Adv. Civ. Eng. 2018, 2018, 7419058. [CrossRef]

16. Otsu, N. A threshold selection method from gray-level histograms. IEEE Trans. Syst. Man, Cybern. 1979, 9, 62–66. [CrossRef]

17. Ayenu-Prah, A.; Attoh-Okine, N. Evaluating pavement cracks with bidimensional empirical mode decomposition. Eurasip J. Adv. Signal Process. 2008, 2008, 1–7. [CrossRef]

18. Koch, C.; Brilakis, I. Pothole detection in asphalt pavement images. Adv. Eng. Inform. 2011, 25, 507–515. [CrossRef]

19. Akagic, A.; Buza, E.; Omanovic, S.; Karabegovic, A. Pavement crack detection using Otsu thresholding for image segmentation. In Proceedings of the 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), Opatija, Croatia, 21–25 May 2018; pp. 1092–1097.

20. Ouma, Y.O.; Hahn, M. Pothole detection on asphalt pavements from 2D-colour pothole images using fuzzy c-means clustering and morphological reconstruction. Autom. Constr. 2017, 83, 196–211. [CrossRef]

21. Marques, A.; Correia, P.L. Automatic Road Pavement Crack Detection Using SVM. Lisbon, Portugal: Dissertation for the Master of Science Degree in Electrical and Computer Engineering at Instituto Superior Técnico. Available online: https://api.semanticscholar.org/CorpusID:111379079 (accessed on 10 February 2022).

22. Ahmadi, A.; Khalesi, S.; Golroo, A. An integrated machine learning model for automatic road crack detection and classification in urban areas. Int. J. Pavement Eng. 2022, 23, 3536–3552. [CrossRef]

23. Cubero-Fernandez, A.; Rodriguez-Lozano, F.; Villatoro, R.; Olivares, J.; Palomares, J.M. Efficient pavement crack detection and classification. EURASIP J. Image Video Process. 2017, 2017, 1–11. [CrossRef]

24. Biçici, S.; Zeybek, M. An approach for the automated extraction of road surface distress from a UAV-derived point cloud. Autom. Constr. 2021, 122, 103475. [CrossRef]

25. Stricker, R.; Eisenbach, M.; Sesselmann, M.; Debes, K.; Gross, H.M. Improving visual road condition assessment by extensive experiments on the extended gaps dataset. In Proceedings of the 2019 International Joint Conference on Neural Networks (IJCNN), Budapest, Hungary, 14–19 July 2019; pp. 1–8.

26. Zhang, A.; Wang, K.C.; Li, B.; Yang, E.; Dai, X.; Peng, Y.; Fei, Y.; Liu, Y.; Li, J.Q.; Chen, C. Automated pixel-level pavement crack detection on 3D asphalt surfaces using a deep-learning network. Comput.-Aided Civ. Infrastruct. Eng. 2017, 32, 805–819. [CrossRef]

27. Rao, A.S.; Nguyen, T.; Palaniswami, M.; Ngo, T. Vision-based automated crack detection using convolutional neural networks for condition assessment of infrastructure. Struct. Health Monit. 2021, 20, 2124–2142. [CrossRef]

28. Fan, R.; Bocus, M.J.; Zhu, Y.; Jiao, J.;Wang, L.; Ma, F.; Cheng, S.; Liu, M. Road crack detection using deep convolutional neural network and adaptive thresholding. In Proceedings of the 2019 IEEE Intelligent Vehicles Symposium (IV), Paris, France, 9–12 June 2019; pp. 474–479.

29. Shamila Ebenezer, A.; Deepa Kanmani, S.; Sheela, V.; Ramalakshmi, K.; Chandran, V.; Sumithra, M.; Elakkiya, B.; Murugesan, B. Identification of Civil Infrastructure Damage Using Ensemble Transfer Learning Model. Adv. Civ. Eng. 2021, 2021, 5589688. [CrossRef]

30. Elghaish, F.; Talebi, S.; Abdellatef, E.; Matarneh, S.T.; Hosseini, M.R.; Wu, S.; Mayouf, M.; Hajirasouli, A.; Nguyen, T.-Q. Developing a new deep learning CNN model to detect and classify highway cracks. J. Eng. Des. Technol. 2021, 20, 993–1014. [CrossRef]

31. Maeda, H.; Sekimoto, Y.; Seto, T.; Kashiyama, T.; Omata, H. Road damage detection using deep neural networks with images captured through a smartphone. arXiv 2018, arXiv:1801.09454.

32. Singh, J.; Shekhar, S. Road damage detection and classification in smartphone captured images using mask r-cnn. arXiv 2018, arXiv:1811.04535.

33. Vishwakarma, R.; Vennelakanti, R. Cnn model & tuning for global road damage detection. In Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 10–13 December 2020; pp. 5609–5615.

34. Kortmann, F.; Talits, K.; Fassmeyer, P.; Warnecke, A.; Meier, N.; Heger, J.; Drews, P.; Funk, B. Detecting various road damage types in global countries utilizing faster r-cnn. In Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 10–13 December 2020; pp. 5563–5571.

35. Wang,W.;Wu, B.; Yang, S.;Wang, Z. Road damage detection and classification with Faster R-CNN. In Proceedings of the 2018 IEEE International Conference on Big Data (Big Data), Seattle, WA, USA, 10–13 December 2018; pp. 5220–5223.

36. Doshi, K.; Yilmaz, Y. Road damage detection using deep ensemble learning. In Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 10–13 December 2020; pp. 5540–5544.

37. 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.

38. Jeong, D. Road damage detection using yolo with smartphone images. In Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 10–13 December 2020; pp. 5559–5562.

39. Hegde, V.; Trivedi, D.; Alfarrarjeh, A.; Deepak, A.; Kim, S.H.; Shahabi, C. Yet another deep learning approach for road damage detection using ensemble learning. In Proceedings of the 2020 IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 10–13 December 2020; pp. 5553–5558.

40. Pena-Caballero, C.; Kim, D.; Gonzalez, A.; Castellanos, O.; Cantu, A.; Ho, J. Real-time road hazard information system. Infrastructures 2020, 5, 75. [CrossRef]

41. Shaghouri, A.A.; Alkhatib, R.; Berjaoui, S. Real-Time Pothole Detection Using Deep Learning. arXiv 2021, arXiv:2107.06356.

42. Authority, Q.S. The General Census of Population and Housing, and Establishment, April 2010. Available online: https://www.stats.gov.sa/en/13 (accessed on 11 May 2022).

43. Ministry of Municipal and Rural Affairs. Evaluation Guide for Road Paving Layers and Facilities (2019), PDF. Available online: https://www.scribd.com/document/475078318/ (accessed on 11 May 2022).

44. Cunha, B.S.D. Development of Computer Vision Based Models for Automated Crack Detection. Master's Thesis, Universidade Federal de Pernambuco, Recife, Brazil, 2020.

45. Ma, N.; Fan, J.; Wang, W.; Wu, J.; Jiang, Y.; Xie, L.; Fan, R. Computer vision for road imaging and pothole detection: A state-of-the-art review of systems and algorithms. arXiv 2022, arXiv:2204.13590.

46. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. J. Big Data 2019, 6, 1–48. [CrossRef]

47. Taylor, L.; Nitschke, G. Improving deep learning with generic data augmentation. In Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence (SSCI), Bangalore, India, 18–21 November 2018; pp. 1542–1547.

48. Marcelino, P. Transfer learning from pre-trained models. Towards Data Sci. 2018, 10, 23.

49. Feng, C.; Zhang, H.; Wang, S.; Li, Y.; Wang, H.; Yan, F. Structural damage detection using deep convolutional neural network and transfer learning. Ksce J. Civ. Eng. 2019, 23, 4493–4502. [CrossRef]

50. Nie, M.;Wang, K. Pavement distress detection based on transfer learning. In Proceedings of the 2018 5th International Conference on Systems and Informatics (ICSAI), Nanjing, China, 10–12 November 2018; pp. 435–439.

51. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 2012, 25, 89–90. [CrossRef]

52. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv 2014, arXiv:1409.1556.

53. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

54. C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," arXiv preprint arXiv:2207.02696, 2022.

55. Alqethami, S.; Alghamdi, S.; Alsubait, T.; Alhakami, H. RoadNet: Efficient Model to Detect and Classify Road Damages. Appl. Sci. 2022, 12, 11529. https://doi.org/10.3390/app122211529

[56] R. Palani[1], Dr. N. Puviarasan[2] and Dr. A. Rama Prasath[3],, COLLECTION OF DISTINCT IMAGES FROM VIDEO FILE USING STRUCTURAL SIMILARITY INDEX MEASURE, ISSN: 2096-3246, Volume 55, Issue 01, January, 2023.

[57] R. Palani[1], Dr. N. Puviarasan[2] and Dr. A. Rama Prasath[3], Literature Review of Road Damage Detection with Repairing Cost Estimation, International Journal of Mechanical, Engineering, e-ISSN 0974-5823| Feb 2022