Building Application for Detection of Damaged Road by using Deep learning.

**简介:**

It is impossible to exaggerate the significance of roads to a nation's economy, as they are crucial to people's daily lives. They're crucial for getting people and goods where they need to go, which in turn stimulates the economy and helps it flourish. Roads are vital to people's day-to-day life because they allow them to get to work, school, the hospital, and other places they need. They allow individuals to get to and from work, meet friends and family, and take part in other social events. The condition of the roads in an area greatly affects how quickly and easily people can go where they need to go.

Roads are also important to a country's economy. They play a crucial role in facilitating the smooth flow of goods and services between manufacturers, retailers, and end users. Businesses benefit from a well-developed road network because it lowers their transportation costs, allowing them to ship goods to new markets more cheaply. Productivity rises, production rises, and the economy grows as a result of this efficiency. Putting people to work is another benefit of building and maintaining roads. Thousands of people are employed directly and indirectly by road building projects. The building of roads is one source of direct employment, while the supply of materials and labor, such as petrol, tools, and repair services, is another. Consistent spending on roads boosts the economy, creates new jobs, and raises people's standard of living.

Nonetheless, the normal lives of people and the economy of a country can be severely disrupted by damaged roadways. Productivity takes a hit when accidents, traffic jams, and longer commute times are caused by roads that aren't up to par. Repairing and maintaining broken roads may be quite expensive, taking money that could be used for other necessities. In addition, poor road conditions might restrict investment and make firms less competitive. Companies' bottom lines might take a hit due to higher transportation costs, longer delivery delays, and lower profit margins caused by poor road conditions. Economic growth and progress may be stunted if this continues. In addition, deteriorated roads may have social repercussions, particularly in rural regions. Communities can become isolated due to bad road conditions, making it difficult for residents to access vital services like healthcare, education, and career prospects. This situation may lead to fewer opportunities, a lower standard of living, and higher levels of poverty.

Moreover, damaged roads can affect the environment. Potholes and cracks in the road surface can cause vehicles to emit more exhaust fumes, leading to air pollution. Water can also collect in potholes and cause erosion, leading to soil and water contamination. There are numerous classifications for damaged roads. The type and intensity of damage is one typical categorization [27]. Here are a couple of such examples:

**1.** **Potholes:** These are tiny to medium-sized holes in the road caused by traffic wear and tear, weather conditions, and inadequate maintenance.

**2.** **Cracks:** These are minor fractures in the road surface that, if not corrected immediately, might grow into bigger holes.

**3.** **Rutting:** It is a depression or groove in the road surface produced by large loads, repetitive driving, and poor drainage.

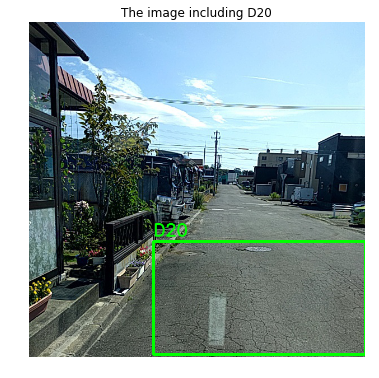
**4.** **Roughness:** This refers to the unevenness of the road surface, which can cause passenger discomfort and vehicle damage.

**5.** **Sinkholes:** These are massive, unexpected depressions on the road surface caused by subsurface infrastructure failure or soil erosion.

**6.** **Flooding:** During strong rains or flooding occurrences, this is a typical concern that can damage the road surface, wash out bridges, and hinder transportation.

In this project, we focused on particularly 4 types of damage. Three different types of cracks based on the view, Rutting, bump, pothole, and separation are classes of this detection project [1]. The purpose of this project is to develop a technology that can automatically detect various types of road damage, such as potholes, cracks, and other defects, using advanced computer vision algorithms. The basic idea is to use deep learning algorithms, modeled after the human brain, to recognize patterns in images or videos of roads with known damage and use this information to identify similar damage in new images or videos or real-time footage. Deep learning and computer vision have revolutionized the field of object detection. Object detection is the process of identifying objects of interest within an image or video stream, and it has numerous practical applications, such as surveillance, autonomous driving, and medical imaging. Deep learning, which is a subset of machine learning, has enabled the development of sophisticated object detection algorithms that can accurately detect and localize objects of interest, even in complex and cluttered scenes. Deep learning models such as convolutional neural networks (CNNs) are particularly effective at learning features and patterns from images, which allows them to generalize well to new data. Computer vision techniques such as image segmentation and object tracking also complement deep learning methods by providing additional context and information about the objects being detected. Together, deep learning and computer vision have opened up new possibilities for object detection, with applications in a wide range of fields, from healthcare to transportation.

Human and object detection technology has had a big influence on detecting damaged roadways. Historically, human inspectors visually inspected roadways for damage indicators such as cracks, potholes, and other problems. Unfortunately, this method was time-consuming and costly, and the quality and consistency of the inspections varied depending on the inspector's experience and knowledge. Object detection technology has made it feasible to automate road inspections using cameras and other sensors to detect and identify indicators of deterioration. It is possible to teach machine learning algorithms to recognize the patterns and characteristics of damaged roadways, allowing them to detect even the most minute indicators of deterioration. This technology has the potential to enhance the precision and efficiency of road inspections, save expenses, and increase driver safety. In addition, the combination of human experience with object detection technology can further increase the accuracy of inspections, since human inspectors are able to comprehend and evaluate the output of machine learning algorithms to make educated recommendations regarding road maintenance and repair. Using object detection technologies for road inspections can save costs. Conventional methods of road inspection can be time-consuming and costly, as they need inspectors to go to the site, physically evaluate the roads, and record their findings. In contrast, object detection technology may be automated by installing cameras and other sensors on vehicles or infrastructure to record data in real time. This can lessen the need for human inspectors and cut inspection costs and time. In addition, the application of machine learning algorithms to evaluate the data can produce more precise and consistent findings, allowing road maintenance and repair to be targeted more efficiently and lowering the likelihood of costly and needless repairs. In addition, early diagnosis of road damage can save money by preventing more serious damage and decreasing the need for future repairs or rebuilding that are more costly. The influence of human and object detection technologies on detecting damaged roads has considerable potential for enhancing the safety and longevity of our transportation infrastructure. Hence, the use of object detection technology for road inspections has the potential to save money by increasing the efficiency and accuracy of inspections and by more efficiently focusing road maintenance and repair.



***Figure 1:*** (a) The samples of damaged roads according to the CRDDC-2022 dataset [1].

In this project, we are proposing YOLO (You Only Look Once) algorithm to detect and classify the damaged types of roads. The goal of this technology is to improve road safety by enabling faster and more accurate detection of road damage. It can also help to reduce maintenance costs and prolong the lifespan of roads by identifying damage earlier, allowing for timely repairs. The key technology used in this project is deep learning algorithms, which are trained on large datasets of road images and videos with annotated road damage. These algorithms can recognize specific types of damage by analyzing thousands of images and videos of roads with known damage. Once trained, the algorithms can be deployed in the field to detect road damage in real-time, using cameras mounted on vehicles or drones, or by analyzing existing video footage of roads.

The main technical index of road damage detection based on deep learning projects is its accuracy in detecting different types of road damage. This is typically measured by evaluating the algorithm's performance on a separate dataset of images and videos with annotated damage. The higher the accuracy of the algorithm, the better it is at detecting and localizing road damage in real-world conditions. Other technical indices may include processing speed, memory usage, and energy efficiency of the algorithm.

The novelty of this project lies in its ability to simplify and enhance transportation, while also aiding relevant departments in identifying and repairing damaged sections of roadways. The primary objective of this work is to develop a user-friendly detection model that can be implemented on both mobile phones and computer systems. The proposed model can be further employed with the use of security cameras, drones, or mobile phone cameras. Another primary goal of this research is to improve upon existing detection models by reducing computational complexity and increasing accessibility. It is expected that this proposed model will require less memory even when processing large images, making it more effective in addressing decision-making and recording concerns. The technical indices of this work include accuracy, processing speed, memory usage, and energy efficiency. These indices will be evaluated to determine the effectiveness and efficiency of the proposed model. Moreover, to make this detection model user convenient we are also developing a computer-aided application for this project. By using this application, the relevant department can use our detection model in real-time work, and also in videos and footage individually. This application will also help reduce the time consumed in calculating and evaluating the result from our detection model.

**详细介绍:**

Road damage detection using deep learning techniques is an area of research that has gained significant attention in recent years. This area is characterized by its scientific nature and advanced technological innovations. Deep learning models have been employed to detect various types of road damage, including potholes, cracks, and other defects. These models are based on neural networks, which are designed to learn from large datasets of images and videos of roadways. The use of deep learning algorithms has allowed for the development of more accurate and efficient detection models, which can identify road damage with high precision and speed. Moreover, the advanced nature of this project is evident in the use of various technologies for collecting data, including drones, mobile cameras, and security cameras. The captured data is then processed by deep learning algorithms, which are capable of handling large amounts of data and making accurate predictions. Additionally, researchers are exploring new techniques for improving the accessibility and usability of these models, such as developing user-friendly interfaces that can be implemented on mobile phones and other devices.

## **Methodology:**

The following are the technical prospects of this project:

(1) Determine and choose the optimal number of images to utilize as a training dataset. The dataset known as the Crowdsensing-based Road Damage Detection Challenge (CRDDC2022) was used for the objectives of this study [1-7]. This dataset contains around 50,000 images, both for training and testing purposes, from a wide range of countries. The major purpose of using this dataset is to acquaint the detector model with the myriad of conceivable road damage scenarios. This is done to assure the model's applicability in any country or territory of the globe.

(2) Prior to commencing the work, the images contained in this dataset must be preprocessed and then processed. To achieve this, the images and annotations must first be transformed into a JSON file while following the PASCAL VOC datasets code of conduct, depending on which format is suitable for the models. Image augmentation is a procedure that may improve the suitability of a dataset for training a certain model by making it more comprehensive. Image optimization techniques such as the Gaussian blur augmentation approach, rotating the objects in images at various angles, adjusting the sheer values, and many other comparable techniques are incredibly important. Typically, this is one of the more effective ways of extracting features from images, which is important for detector models.

(3) Recurrent neural networks [8], convolutional neural networks [9], and long short-term memory networks are the three varieties of neural networks that are used the most often in the incorporation of models for use in deep learning. In particular, a convolutional neural network is a potent tool that can be used for the analysis of images as well as the development of prediction models. There are numerous distinct algorithms available for locating objects, including R-CNN, EfficientNet [10], ImageNet [11], SSD [12], YOLO [13], Detectron2 [14], and plenty of others besides. In contrast to YOLO, SSD use CNN to detect immediately, rather than waiting until after the completely linked layer is completed. SSD will extract feature maps at varying scales for the purpose of detection. The large-scale feature map, which was created earlier, may be used to identify little items, while the small-scale feature map, which was created later, is utilized to detect huge objects. The SSD makes use of earlier boxes in a variety of sizes and aspect ratios. SSD has a higher detection speed than R-CNN, even though its detection accuracy is not quite as great as that of R-CNN. Yet, YOLO is vice versa. Although R-CNN's speed is slower than that of other methods, it is nevertheless quite effective at finding large objects. In this scenario, Faster R-CNN [15] or Mask R-CNN [16] may be used for experimental purposes. Hence, we selected the YOLO algorithm as our final outcome.

(4) The following will be used for the technical tasks:

|  |  |
| --- | --- |
| **Operation** | **Library** |
| Computation | NumPy |
| Dataset handling | ElementTree, Pandas, Json |
| Image Processing and visualization | OpenCV, Matplotlib |
| Image Augmentation | Keras |
| Model Architecture design, Model training, Model fine tuning | PyTorch |
| Others | OS, BeautifulSoup, pickle |

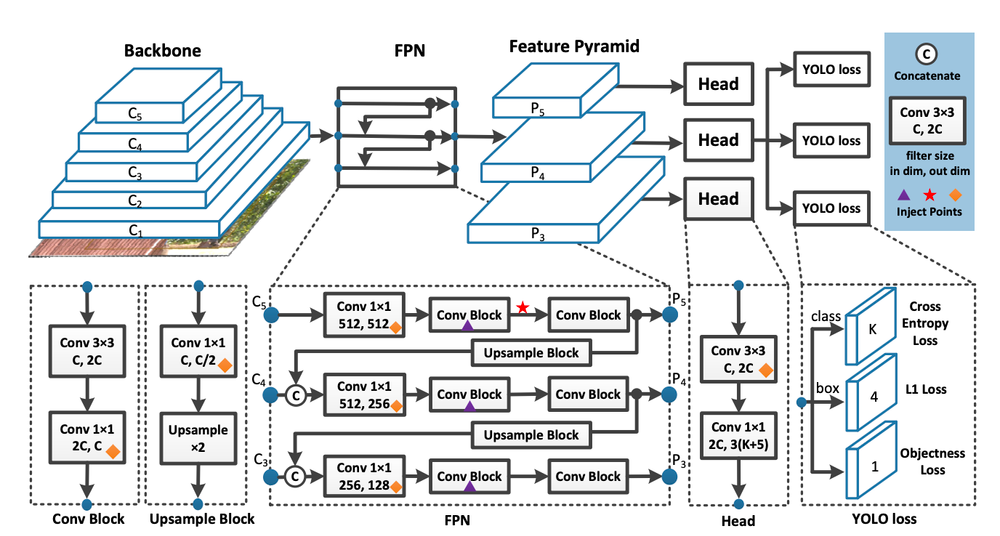
Additionally, Windows 10 will also be employed as the platform for the computer's operating system. The programming language will be Python, and the integrated development environment will be Jupyter Notebook (IDE) or Anaconda. This will make deployment and visualization much easier.

## **Proposed Model:**

Efficient road maintenance requires a reliable monitoring system and a straightforward method is the human visual inspection; however, it is infeasible due to being expensive, laborious, and time-consuming. Therefore, researchers have developed various solutions for automatic road damage inspection, including vibration-based [17], laser-scanning-based [18], and image-based [19-22] methods. While detection by vibration methods is limited to the contacted parts of the road, laser-scanning methods provide accurate information about the status of roads; however, such methods are expensive and require road closure. Meanwhile, image processing methods are inexpensive but may suffer from a lack of accuracy. In spite of its immaturity, recent advancements in image analysis techniques have been producing impressive results and thus increasing their usage for various applications (e.g., street cleanliness [23], traffic flow analysis [24], situation awareness of disasters [25], and image search [26]). A few researchers developed image-based approaches for road surface inspection using state-of-the-art deep learning methods. In particular, some works focus on detecting only the existence of the damage regardless of its type [19]. Other works focus on classifying road damage into a few types. For example, Zhang et al. [20] devised an approach for detecting two directional cracks (i.e., horizontal and vertical), while Akarsu et al. [21] developed another approach for detecting three categories of damages, namely horizontal, vertical, and crocodile. Due to the fact that differentiating among damage types is critical for proper road maintenance planning, Maeda et al. [22] have implemented an approach for a thorough classification of road damage types.

In this project, we are proposing the YOLOv7 model for detecting and classifying the damaged parts of the roads. The YOLOv7 algorithm is a cutting-edge object identification model built to outperform its predecessors in the You Only Look Once (YOLO) family of algorithms in both accuracy and speed. This algorithm improves upon prior iterations by capitalizing on its strengths while overcoming its weaknesses; it has found widespread use in fields as diverse as surveillance, autonomous vehicle technology, and robotics.

The YOLOv7 algorithm's strength lies in its ability to effectively recognize and categorize objects in real-time video feeds. To do this, a deep neural network is utilized which has been pre-trained on big annotated picture datasets. There are several tiers of linked nodes in the network, and each one serves a distinct purpose in the overall task of object detection. In order to forecast the existence and position of objects in an image or video frame, these layers are built to recognize different characteristics of such items, such as their shape, color, and texture. The YOLOv7 algorithm is versatile in that it can process objects of varying shapes and sizes. This is accomplished with the help of a feature pyramid network trained to recognize objects across a range of sizes and granularities. As a result, the algorithm can reliably detect objects of varying sizes, even in scenes with many different kinds of objects in close proximity to one another. The YOLOv7 method is also very low on the processing resources it needs to run. The computational complexity of the model is reduced through the use of a lightweight backbone network and other optimizations. Due to its low power requirements, the algorithm is suitable for use in a variety of contexts, including on smartphones and embedded systems, where it may operate in real-time. Overall, the YOLOv7 algorithm represents a significant advance in the field of object detection, and its high accuracy, speed, and efficiency make it a popular choice for a wide range of applications [29-31].



**Figure 2:** The model architecture of YOLOv7.

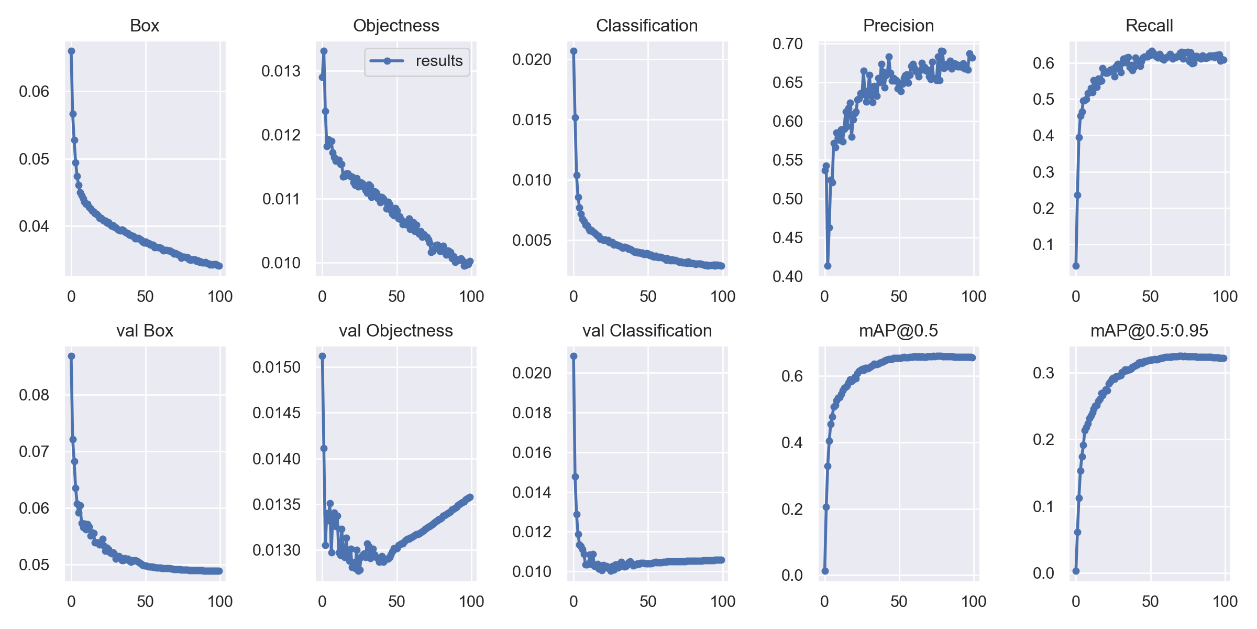
There are various models in this version of the YOLO algorithm. These different models are based on the number of layers and total parameters. In this project, we proposed to use the yolov7 model which is generous and has comparatively less computation than the other models of this YOLO version. The average time of this model during the training is 2.8*ms* for batch size 32 whereas other models such as yolov7-x is 4.3*ms*, and yolov7-E6E is 18.7*ms* [32,33].

## **Dataset:**

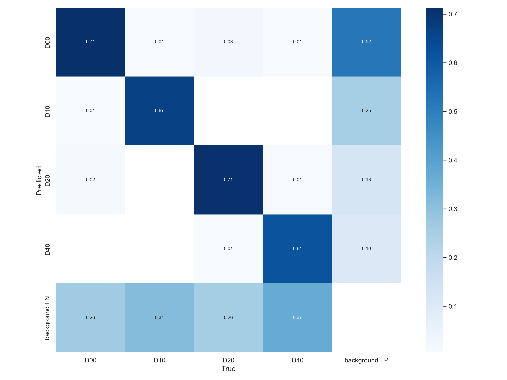
In our project, we follow the Crowdsensing-based Road Damage Detection Challenge (CRDDC-2022) [1] dataset. They follow the standard of the Road Maintenance and Repair Guidebook 2013 JRA (2013) in Japan. There are several kinds of cracks in their datasets. In the base of the view, there are 2 kinds of cracks, one is Linear Crack and the other is Alligator Crack. In linear crack there are two classes of labels of cracks, one class is D00 for longitudinal crack and D10 is for lateral crack. For alligator crack class is D20. We included another class as the damage type is D40.



## **Performance:**

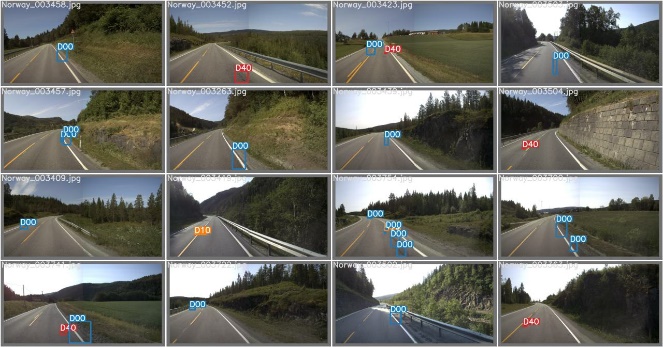


**Figure 3:** Loss values of bounding boxes, objects, and classifications for training and validation datasets. The precision and recall curves during the training.



**Figure 4:** Confusion matrix on predicting the test data.

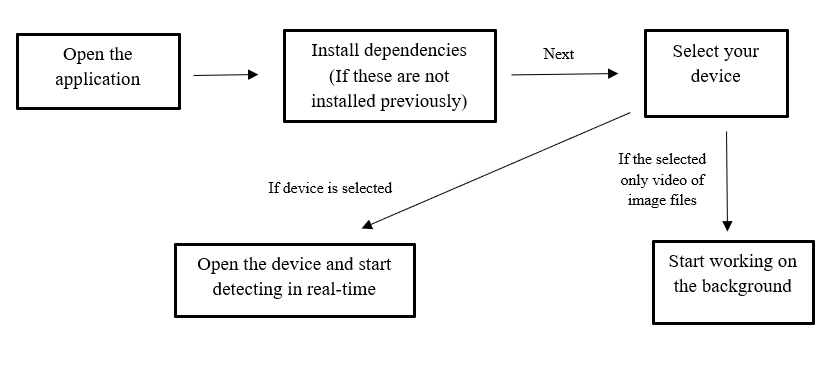
In figure 3,4 we can notice that our model is continuously reducing the loss values of predicting objects, bounding boxes, and classifications. We continued this training for 100 epochs with fewer hyperparameters which helped our model to be trained in more specifically with fewer epochs. Due to our shortage of GPU computation devices, we ran only 100 epochs for this training. However, this result gives us a bigger hope to get a higher and more precise and accurate model with more epochs with a higher configured GPU device.

**Figure 5:** Evaluating the trained model on different footage.

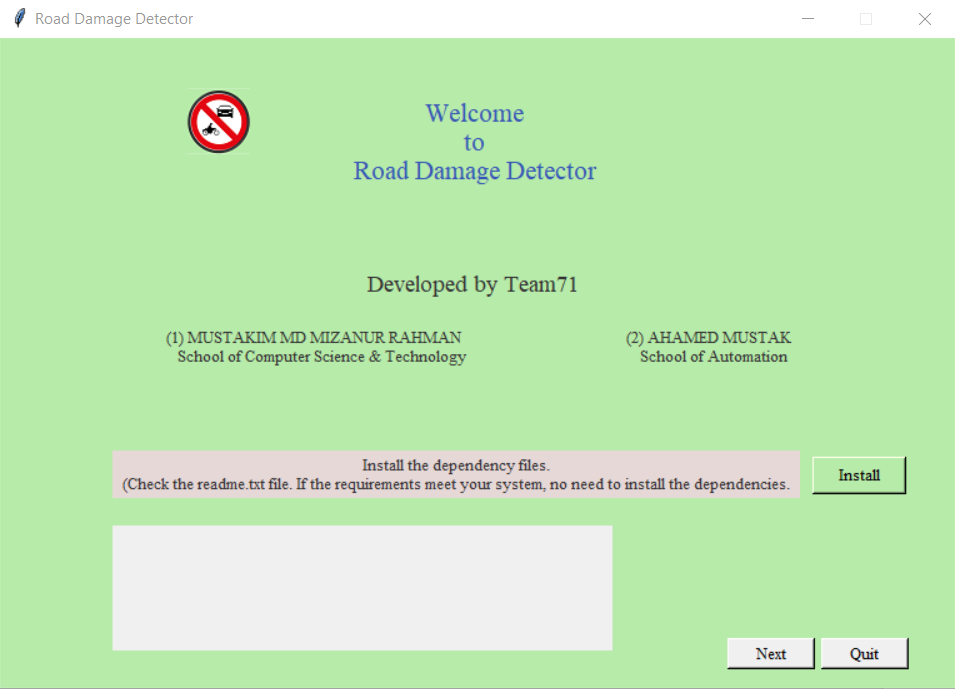
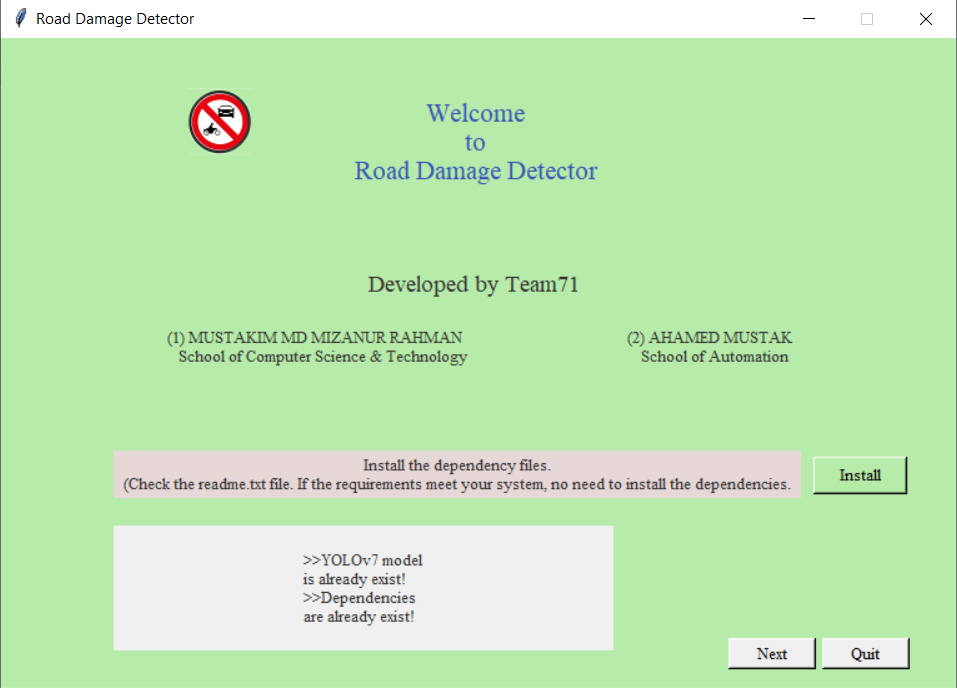
After training, we made an experiment with different images which our model never meets during the training. In figure 5, we can see that almost above 85% of the damaged roads can be detected by our model.

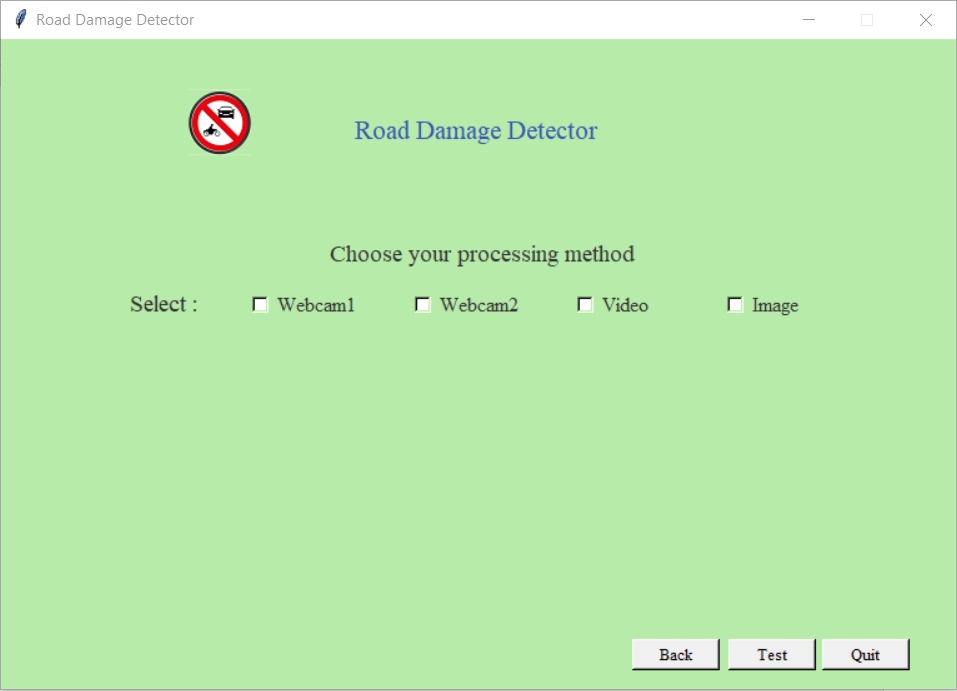
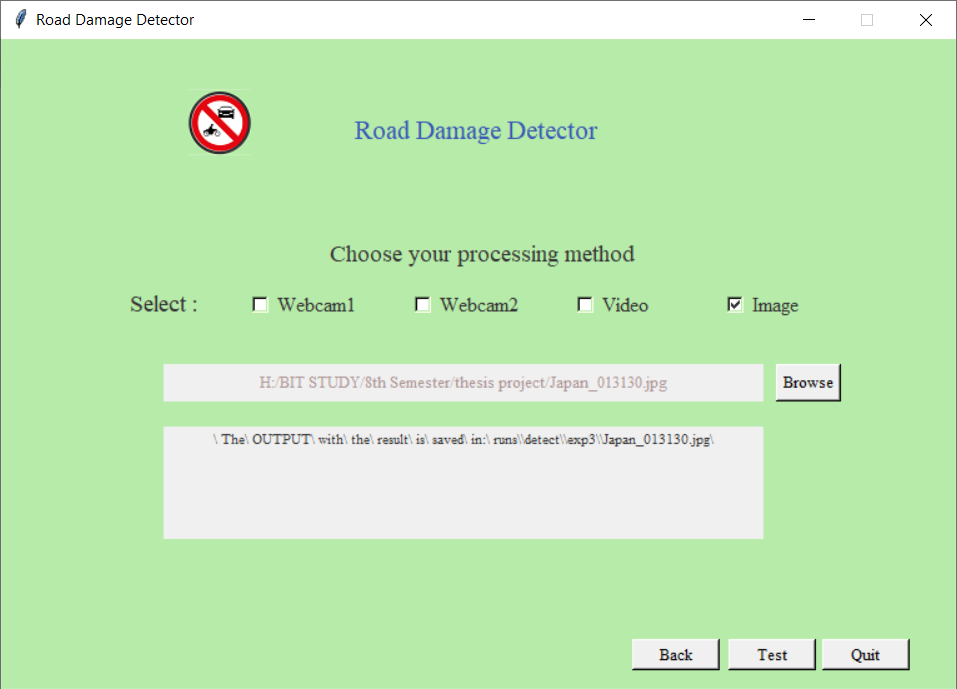
**Application Pipeline & Documentation:**



**Figure 6:** Application Pipeline

Our application pipeline is way specific. First, the users have to run the application on their computers. And then they will require to install the dependencies if the dependencies are not installed on their device. It can easily be done by pressing the ‘Install’ button over the interface. After getting the successful installation message in the message prompt, the users will move to the next window by pressing the ‘Next’ button. In the new window, it is visible to have some choices to specify the detection method. If the user wants to do real-time detection, they can choose either ‘Webcam1’ or ‘Webcam2’, also the users can choose the video or the photo checkbox if they want to use their previously captured videos or images. After choosing the detecting gateway, the detection process will be started after the user presses the ‘Test’ button.

**Figure 7:** The user interface of our model application.

**Advantages of this project:**

The technologies being used in China to identify damaged roads include intelligent transport systems (ITS) and road monitoring systems [34]. ITS uses sensors and other advanced technologies to collect real-time data on road conditions and traffic patterns, which can be used to identify areas of congestion and prioritize repairs. Road monitoring systems involve installing sensors and other equipment on roads to collect data on traffic flow and road conditions, which can then be analyzed to identify areas of damage. However, it’s a matter of cost and a setup of many hardware, whereas our system can be used by any surveillance camera, drone, or high-definition camera, or even by computer webcam. This application doesn’t require any kind of sensor.

In addition, our proposed model and also the application is built with the latest technology. In recent technologies, YOLOv7 is the most talkable topic for developing high powerful object detection model. Overall, the scientific and advanced nature of road damage detection research is driving significant progress in the field of transportation infrastructure management. By developing more accurate and efficient detection models, this research has the potential to improve road safety and reduce the costs associated with road maintenance and repair.

## **Analysis:**

Defect detection has highly become technology-driven, based on the constant developments in big data and artificial intelligence. The utilization of smart cameras and associated AI-driven systems is already assisting manufacturers to provide a super-quality examination in shorter cycles, decreasing latency and costs, and establishing new mandates that are not feasible to be performed even by highly experienced human inspectors. The Global Defect Detection Market attained a market size of $3,394.6 million in 2020 and it is expected to reach $5,113.4 million by 2027, growing at a CAGR of 6.6% during 2021 -2027. Hence, defect detection is a major component of any manufacturing quality control and assurance process. Conventionally, defect detection was performed manually by human beings who are highly susceptible to fatigue, biases, and inattentiveness. However, currently, manual inspection is enhanced by rule-based machine vision technologies.

Over the course of the last several decades, China has enjoyed amazing economic growth, and the country's transportation infrastructure has been a significant factor in the country's ability to support this expansion. Nonetheless, the degrading roads in China have had a substantial negative effect on the country's economy. Inadequate road conditions can result in higher transportation costs, longer travel times, and decreased efficiency, all of which can have a detrimental impact on the competitiveness of Chinese companies and drive up the price of goods and services. According to research published by the World Bank, the Chinese economy loses around 5% of its GDP every year due to insufficient infrastructure, particularly poor road conditions [28].

Road damage detection based on deep learning has a wide range of potential uses and advantages. Some of them are:

**Early Detection:** The detection of road damage in its early stages can help prevent accidents and save lives.

**Cost Savings:** Early detection can also help save costs associated with extensive road repairs, which can be much more expensive than small-scale repairs.

**Improved Maintenance:** Road damage detection can lead to better and more efficient maintenance of road infrastructure, helping to ensure road safety and reduce traffic congestion.

**Automated Detection:** Automated Road damage detection systems can help reduce the need for manual inspections, which can be time-consuming, costly, and may not always identify all forms of damage.

**Increased Accuracy:** Deep learning algorithms used for road damage detection can achieve high levels of accuracy, reducing the risk of false positives and false negatives.

**Accessibility:** User-friendly interfaces for road damage detection models can make the technology more accessible to a wider range of users, including government agencies, private contractors, and individual users.

**Scalability:** Road damage detection models based on deep learning can be easily scaled up or down, depending on the size and complexity of the road network being monitored.

**Integration with Other Technologies:** Road damage detection can be integrated with other technologies such as autonomous vehicles, to ensure safe and reliable transportation.

Moreover, road damage detection based on deep learning can improve road safety, reduce costs, and ensure more efficient maintenance of road infrastructure, making it an important area of research and development in transportation engineering.

## **Conclusion:**

In conclusion, the road damage detection model is a key piece of technology that may have a substantial influence on both the economy and people's everyday lives. The prompt and precise identification of damaged roadways can decrease accidents, improve traffic flow, and ultimately save lives. Additionally, this technology can result in considerable cost savings for governments and road authorities by enabling them to more efficiently prioritize and schedule road maintenance and repair. The advancement of machine learning and computer vision technology has substantially enhanced the precision and speed of road damage detection models. Using deep learning methods such as YOLOv7 has considerably increased the detection accuracy of these models, while also decreasing their computing complexity and enhancing their efficiency. In addition, the combination of object identification technology with satellite images and LiDAR data can give a more comprehensive and thorough view of road conditions, enabling more accurate and efficient planning of road maintenance and repair.

Overall, the model for road damage identification is a promising technology that can provide multiple benefits to society. Its precision, speed, and efficiency make it an essential tool for managing and maintaining road networks by governments and road authorities. This technology has the potential to transform how we manage our roadways and transportation networks, making them safer and more efficient for everybody. Thus, it can be asserted that this model is widely acknowledged as an indispensable instrument for road maintenance and safety.

**References**

1. Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., & Sekimoto, Y. (2022). Crowdsensing based Road Damage Detection Challenge (CRDDC-2022). arXiv preprint arXiv:2211.11362.
2. Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., & Sekimoto, Y. (2022). RDD2022: A multi-national image dataset for automatic Road Damage Detection. arXiv preprint arXiv:2209.08538
3. Arya, D., Maeda, H., Sekimoto, Y., Omata, H., Ghosh, S. K., Toshniwal, D., Sharma, M., Pham, V. V., Zhong, J., et al. (2022). RDD2022 - The multi-national Road Damage Dataset released through CRDDC'2022 (Version 1). figshare. <https://doi.org/10.6084/m9.figshare.21431547.v1>
4. Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., Mraz, A., Kashiyama, T., & Sekimoto, Y. (2021). Deep learning-based road damage detection and classification for multiple countries. Automation in Construction, 132, 103935. 10.1016/j.autcon.2021.103935
5. Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., & Sekimoto, Y. (2021). RDD2020: An annotated image dataset for Automatic Road Damage Detection using Deep Learning. Data in Brief, 107133. <https://doi.org/10.1016/j.dib.2021.107133>.
6. Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., Omata, H., Kashiyama, T., & Sekimoto,Y. (2020). Global Road Damage Detection: State-of-the-art Solutions. IEEE International Conference on Big Data (Big Data), Atlanta, GA, USA, 2020, pp. 5533-5539, doi: 10.1109/BigData50022.2020.9377790.
7. Maeda, H., Kashiyama, T., Sekimoto, Y., Seto, T., & Omata, H. (2021). Generative adversarial network for road damage detection. Computer‐Aided Civil and Infrastructure Engineering, 36(1), 47-60.
8. Medsker, L.R. and Jain, L.C., 2001. Recurrent neural networks. *Design and Applications*, *5*, pp.64-67.
9. Kim, P. and Kim, P., 2017. Convolutional neural network. *MATLAB deep learning: with machine learning, neural networks and artificial intelligence*, pp.121-147.
10. Tan, M. and Le, Q., 2019, May. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International conference on machine learning* (pp. 6105-6114). PMLR.
11. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K. and Fei-Fei, L., 2009, June. Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248-255). Ieee.
12. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y. and Berg, A.C., 2016. Ssd: Single shot multibox detector. In *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14* (pp. 21-37). Springer International Publishing.
13. Redmon, J., Divvala, S., Girshick, R. and Farhadi, A., 2016. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).
14. Wu, Y., Kirillov, A., Massa, F., Lo, W.Y. and Girshick, R., 2019. Detectron2.
15. Ren, S., He, K., Girshick, R. and Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, *28*.
16. He, K., Gkioxari, G., Dollár, P. and Girshick, R., 2017. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961-2969).
17. B. X. Yu and X. Yu, “Vibration-based system for pavement condition evaluation,” in AATT, 2006, pp. 183–189.
18. Q. Li, M. Yao, X. Yao, and B. Xu, “A real-time 3d scanning system for pavement distortion inspection,” MST, vol. 21, no. 1, p. 015702, 2009.
19. A. Zhang, K. C. Wang, B. Li, E. Yang, X. Dai, Y. Peng, Y. Fei, Y. Liu, J. Q. Li, and C. Chen, “Automated pixel-level pavement crack detection on 3d asphalt surfaces using a deep-learning network,” CACAIE, vol. 32, no. 10, pp. 805–819, 2017.
20. L. Zhang, F. Yang, Y. D. Zhang, and Y. J. Zhu, “Road crack detection using deep convolutional neural network,” in ICIP. IEEE, 2016, pp. 3708–3712.
21. B. Akarsu, M. KARAKOSE, K. PARLAK, A. Erhan, and A. SARI- ¨ MADEN, “A fast and adaptive road defect detection approach using computer vision with real time implementation,” IJAMEC, vol. 4, no. Special Issue-1, pp. 290–295, 2016.
22. H. Maeda, Y. Sekimoto, T. Seto, T. Kashiyama, and H. Omata, “Road damage detection and classification using deep neural networks with smartphone images,” CACAIE.
23. A. Alfarrarjeh, S. H. Kim, S. Agrawal, M. Ashok, S. Y. Kim, and C. Shahabi, “Image classification to determine the level of street cleanliness: A case study,” in BigMM. IEEE, 2018.
24. S. H. Kim, J. Shi, A. Alfarrarjeh, D. Xu, Y. Tan, and C. Shahabi, “Realtime traffic video analysis using intel viewmont coprocessor,” in DNIS. Springer, 2013, pp. 150–160.
25. A. Alfarrarjeh, S. Agrawal, S. H. Kim, and C. Shahabi, “Geo-spatial multimedia sentiment analysis in disasters,” in DSAA. IEEE, 2017, pp. 193–202.
26. A. Alfarrarjeh, C. Shahabi, and S. H. Kim, “Hybrid indexes for spatialvisual search,” in ACM MM Thematic Workshops. ACM, 2017, pp. 75–83.
27. Kamaruddin, I., Singh, L., & Sudin, I. (2019). Classification of Road Surface Defects: A Review. Journal of Engineering Science and Technology, 14(3), 1473-1495.
28. World Bank. (2013). China 2030: Building a Modern, Harmonious, and Creative Society. Washington, DC: World Bank.
29. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M. (2021). YOLOv7: Real-time object detection revisited. arXiv preprint arXiv:2103.09460.
30. Redmon, J., & Farhadi, A. (2018). YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.
31. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).
32. Wang, C., Liao, H., & Yeh, I. (2022). Designing Network Design Strategies Through Gradient Path Analysis. ArXiv. /abs/2211.04800
33. Wang, C., Bochkovskiy, A., & Liao, H. (2022). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. ArXiv. /abs/2207.02696
34. China Daily. (2019). China seeks smart solutions to fix crumbling roads. Retrieved from https://www.chinadaily.com.cn/a/201908/06/WS5d49b0f6a310cf3e35563f75.html