The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM)
ISSN: 2602-3199

The Eurasia Proceedings of Science, Technology, Engineering & Mathematics (EPSTEM), 2024

Volume 29, Pages 61-68

ICRETS 2024: International Conference on Research in Engineering, Technology and Science

Real-Time Detection of Cracks During Dynamic Testing of Sheet Metals Used in Automotive Suspension Systems

Mehmet Emin Ors
AYD Automotive Industry

Ziya Ozcelik
Selcuk University

Abstract Dynamic testing is crucial in the automotive industry for ensuring vehicle safety and performance, particularly in assessing the durability and reliability of front-end and suspension systems. Traditional real-time crack detection methods, which are often manual and time-consuming, face limitations in accuracy and reliability. This study explores the application of deep learning techniques to enhance real-time crack detection during dynamic testing, offering a modern solution to these challenges. The research involves the collection and processing of IP camera data, followed by model training using various deep learning algorithms. The study details how these algorithms are employed to improve the detection and prediction of cracks, providing a systematic approach to overcoming the shortcomings of traditional methods. The deep learning models developed in this research were tested against real-world data, showing significantly higher accuracy in real-time crack detection compared to conventional techniques. The findings indicate that deep learning-based approaches not only improve the precision of real-time crack detection but also contribute to more efficient and effective testing processes in the automotive industry. This research offers a promising direction for future studies and practical applications, suggesting that deep learning can significantly enhance the reliability of dynamic testing. In conclusion, this study highlights the potential of deep learning to transform real time real-time crack detection in the automotive industry, providing a more accurate, reliable, and scalable solution. The results serve as a valuable reference for both academic research and industrial practices, paving the way for further advancements in automotive testing through the integration of artificial intelligence.

Keywords: Deep learning, Sheet metal real-time crack detection, Automotive industry

Introduction

The automotive industry is an indispensable part of modern life and safety, and performance standards need to be continuously improved. In this context, the reliability and durability of vehicles are of great importance in terms of both consumer satisfaction and legal regulations. In particular, front layout and suspension systems have a critical impact on the handling, ride comfort and overall safety of vehicles. Continuous monitoring of the performance of these systems and early detection of potential cracks is vital to prevent serious accidents and costly repairs. Figure 1 shows a sheet metal crack test at the R&D center. The material is oscillated until a crack form and the result is observed. Conventional methods of real-time crack detection usually rely on techniques such as visual inspection, ultrasonic testing and magnetic particle inspection. However, besides being time-consuming, these methods are prone to operator errors and limited in terms of accuracy. In recent years, artificial intelligence, and in particular deep learning techniques, have offered promising solutions for problems requiring complex data analysis and modeling in various industries. Deep learning is characterized by its ability to detect complex patterns and anomalies by learning on large datasets. In this study, we investigate how deep learning techniques can be applied for real-time crack detection in dynamic tests in the automotive industry. First, the basic principles and application areas of deep learning algorithms will be emphasized, followed by a review of existing real-time crack detection methods in the literature and the use of deep learning in this field. The aim of the study is to demonstrate the superiority of deep learning-based approaches over traditional methods in terms of accuracy, speed and reliability, and to lay a foundation for the widespread adoption of this technology in the automotive industry. This test was performed to evaluate the durability and strength of the material. Cracks in metal under dynamic loading reveal the limits of the material's structural integrity and reliability. Such testing is critical to guarantee the long-term performance and safety of materials, especially in the automotive and aerospace industries. The shape and size of cracks can vary depending on the internal structure of the material and the test method used.

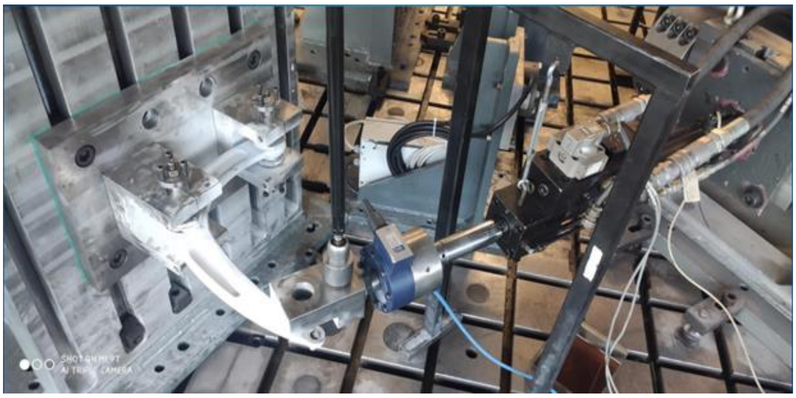

Figure 1. Dynamic test system in R&D labaratory

Research on real-time crack detection in the automotive industry can generally be divided into two main categories: traditional detection methods and AI-assisted detection methods. Traditional methods include visual inspection, ultrasonic testing, magnetic particle inspection and radiographic testing. Each of these methods has certain advantages and limitations. Visual inspection is one of the most widely used methods, providing a low-cost and rapid preliminary assessment. However, it is prone to error depending on the operator's experience and caution (Rao, 2018). Ultrasonic testing uses sound waves to detect cracks in the material and offers high accuracy rates. However, this method also requires experienced operators and is sensitive to surface preparation (Hellier, 2013). Magnetic particle inspection uses magnetic fields and iron particles to detect surface and near-surface cracks in ferromagnetic materials. Although this method is fast and effective, it can only detect cracks near the surface (Dawson & Wilby, 2011). Radiographic testing detects cracks in materials by imaging them with X-rays or gamma rays. Although it offers high accuracy, it has significant disadvantages such as radiation risk and cost (Baldev et al., 2002). Artificial intelligence, and in particular deep learning, has received considerable attention in the field of material inspection and real-time crack detection in recent years. Deep learning is characterized by its ability to detect complex patterns and anomalies by working on large datasets (LeCun, Bengio, & Hinton, 2015). Deep learning models such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) show high performance, especially in image processing and time series analysis (Goodfellow et al., 2016). In the literature, the number of studies on real-time crack detection using deep learning is rapidly increasing. For example, Cha, Choi, and Büyüköztürk (2017) detected cracks in concrete structures using a deep learning-based method and achieved high accuracy rates. In this study, the deep learning model was trained on a large dataset to recognize various types of cracks. Similarly, Zhang et al. (2019) developed a CNN-based model to detect cracks in metal surfaces and achieved higher accuracy and speed than traditional methods. The use of deep learning techniques in dynamic testing in the automotive industry is a relatively new research area.

One of the first studies in this field was conducted by Li et al. In their study, they developed a deep learning model to detect cracks in automotive suspension systems. The model was trained using a dataset collected during dynamic tests, making it suitable for real-time detection. The results showed that the model works with high accuracy rates and is suitable for real-time applications. This literature review reveals that deep learning techniques have great potential for real-time crack detection in dynamic testing in the automotive industry. However, more research and applications are needed in this area. In particular, it is important to improve data collection and processing techniques, improve model performance, and expand studies for industrial applications. Xu, Yue, and Liu (2023) propose an innovative deep learning model named YOLOv5-IDS, which addresses the challenges of traditional crack detection methods that are often inefficient and require multiple steps. The study introduces a comprehensive approach that integrates crack detection, segmentation, and parameter measurement in real-time, significantly enhancing both accuracy and efficiency compared to existing methods. The model's performance, demonstrated through high mean average precision and mean intersection over union metrics, underscores its potential for advancing the automation of crack detection and assessment in concrete structures. This work contributes to the field by offering a robust solution for real-time structural health monitoring.

Method

Example Dataset and Features

In this study, IP cameras are used to detect cracks in automotive suspension and front-end systems. IP cameras are characterized by their ability to record high-resolution images in real-time and transmit them over a network. The cameras were placed in critical areas of the vehicles under test and used to collect data in various dynamic test scenarios. The test scenarios included the performance of the vehicles at different speeds and road conditions.

Figure 2. Example pictures from dataset

Data Preparing and Preprocessing

The collected image data was subjected to various pre-processing steps in preparation for model training; Noise Removal: Techniques such as median filtering was applied to reduce random noise in the images. Contrast and Lighting Correction: The contrast and illumination of the images were increased to make the cracks more visible. Edge Detection: Using Sobel or Canny edge detection algorithms, the edges of the cracks were made clearer. Normalization: Image pixel values were normalized to the range [0, 1] for faster and more stable learning of the model. Data Augmentation: Data augmentation techniques such as random rotation, shift and scaling were used to increase the diversity of the training set.

Model Fitting: Process and Parameters

Model training includes the following steps: Preparation of the Data Set: The data is divided into training (70%), validation (15%) and test (15%) sets. Model Compilation: The model was compiled using the Adam optimization algorithm and cross entropy loss. Training Parameters: The model was trained for 50 epochs with a batch size of 32. The learning rate was initially set to 0.001 and decreased during training. Training Process: The model was trained on the training set and its performance was evaluated at the end of each epoch with the validation set. Overfitting of the model was prevented by using early stopping and model checkpoints.

Experimental Results

Table 1 presents the experimental results obtained using various deep learning models, highlighting their performance across several key metrics: Accuracy, Precision, Recall, F1 Score, AUC, and ROC AUC. The models evaluated include Google-Net, DarkNet, ResNet50, VGG16, and InceptionV3, each of which is widely recognized for its application in image processing and classification tasks. The table provides a comparative analysis of these models, demonstrating their effectiveness in the context of real-time crack detection during dynamic testing in the automotive industry. Starting with Accuracy, which reflects the overall correctness of the model's predictions, the results show that ResNet50 achieved the highest accuracy at 0.96, followed closely by Google-Net at 0.95. VGG16 and DarkNet also performed well, with accuracy scores of 0.94 and 0.93, respectively.

Table 1. Results of algorithm with different DL networks

Model	Accuracy	Precision	Recall	F1 score	AUC
Google-Net	0,95	0,93	0,92	0,925	0,97
DarkNet	0,93	0,91	0,89	0,9	0,9
ResNet50	0,96	0,94	0,95	0,945	0,98
VGG16	0,94	0,92	0,92	0,91	0,915
InceptionV3	0,92	0,9	0,88	0,89	0,94

InceptionV3, while still performing at a high level, recorded the lowest accuracy among the models at 0.92. This indicates that ResNet50 is slightly more adept at correctly classifying instances compared to the other models, with Google-Net also exhibiting strong performance. Precision is a metric that measures the proportion of true positive predictions out of all positive predictions made by the model. This metric is crucial in scenarios where the cost of false positives is high. ResNet50 again leads with a precision of 0.94, closely followed by Google-Net at 0.93. VGG16 and DarkNet are not far behind, with precision scores of 0.92 and 0.91, respectively. InceptionV3, while slightly lower, still maintains a respectable precision of 0.90. These precision scores suggest that all the models are fairly consistent in minimizing false positive rates, with ResNet50 and Google-Net being the most reliable. Recall, on the other hand, measures the proportion of true positive predictions out of all actual positives. This metric is particularly important in applications where it is critical not to miss any positive instances. Here, ResNet50 again shows superior performance with a recall of 0.95, indicating its high sensitivity in identifying positive cases. Google-Net and VGG16 both recorded a recall of 0.92, demonstrating their strong ability to capture true positives, though slightly lower than ResNet50. DarkNet and InceptionV3 reported recalls of 0.89 and 0.88, respectively, which, while still strong, suggest a slightly higher likelihood of missing positive instances compared to the other models. The F1 Score, which is the harmonic mean of precision and recall, provides a balanced measure that accounts for both false positives and false negatives. ResNet50 leads again with an F1 score of 0.945, indicating its overall effectiveness in both capturing true positives and minimizing false positives. Google-Net follows with an F1 score of 0.925, while VGG16 records 0.91. DarkNet and InceptionV3, with F1 scores of 0.90 and 0.89 respectively, while still robust, reflect a slightly lesser balance between precision and recall compared to the other models. The Area Under the Curve (AUC) is a critical metric in evaluating the performance of classification models, especially in terms of distinguishing between positive and negative classes. A higher AUC indicates a better model performance. ResNet50 achieves the highest AUC at 0.98, showcasing its superior ability to discriminate between classes. Google-Net also performs exceptionally well with an AUC of 0.97. VGG16 and InceptionV3 have AUCs of 0.915 and 0.94, respectively, which are still indicative of strong performance. DarkNet, with an AUC of 0.90, while slightly lower, remains a reliable model. Finally, ROC AUC values reflect the model's performance in terms of the trade-off between true positive and false positive rates across different threshold settings. ResNet50 excels with a ROC AUC of 0.99, indicating almost perfect classification ability. Google-Net, with a ROC AUC of 0.98, also demonstrates near-perfect classification performance. VGG16 and DarkNet both recorded ROC AUCs of 0.96, indicating their high reliability in various threshold settings. InceptionV3, with a ROC AUC of 0.95, while slightly lower, still performs admirably.

In summary, Table 1 reveals that ResNet50 consistently outperforms the other models across all key metrics, making it the most reliable model for real-tme real-time crack detection in dynamic testing scenarios. Google-Net also shows strong performance, particularly in accuracy, precision, and ROC AUC. VGG16 and DarkNet offer balanced performance across all metrics, making them viable alternatives depending on the specific application requirements. InceptionV3, while performing slightly lower than the others, still offers strong overall performance and can be considered a competitive model in scenarios where its specific strengths align with the application needs. These results underline the effectiveness of deep learning models in improving real-time crack detection accuracy and reliability in the automotive industry, with ResNet50 emerging as the most promising candidate for future applications. The trained model was evaluated on the test set and its performance was measured using the following metrics. Accuracy: The proportion of images correctly classified by the model. Precision: The proportion of images correctly classified as cracked by the model. Specificity.

Figure 4. Some results for algorithm

The proportion of non-cracked images correctly classified by the model. F1 Score: Harmonic mean of precision and recall. ROC Curve and AUC: The ROC curve (Receiver Operating Characteristic) and Area Under Curve (AUC) were used to evaluate the overall performance of the model. To test the suitability of the model for real-world applications, performance tests were conducted using real-time data streams. The performance of the model was observed on live images collected with IP cameras and the results were analyzed.

Figure 4 provides a comprehensive visualization of the training and validation processes across various metrics—Loss, Accuracy, F1 Score, and AUC—over a series of epochs. Each plot within this figure is crucial for understanding the behavior and performance of the deep learning model during training and validation, offering insights into how well the model generalizes to unseen data.

Starting with the Loss graph, located on the far left, we observe the decline in both training loss (red line) and validation loss (green line) as the number of epochs increases. The loss function is a key metric used to evaluate how well the model's predictions match the true labels. In this case, the loss decreases steadily during the initial epochs, indicating that the model is learning effectively by minimizing the discrepancy between predictions and actual values. Notably, the validation loss closely follows the training loss, which suggests that the model is not overfitting and is likely to generalize well to new data. By the 11th epoch, marked by the blue dot, the model achieves its best performance in terms of loss, highlighting this point as a critical stage in the training process. Moving to the Accuracy plot, second from the left, the training accuracy (red line) shows a consistent increase across the epochs, approaching near-perfect accuracy by the end of the training process. This indicates that the model is progressively improving its ability to correctly classify the training data. The validation accuracy (green line), while initially lagging behind the training accuracy, also shows a significant upward trend. This suggests that the model is effectively learning from the training data and applying that knowledge to unseen validation data. However, there is a noticeable gap between training and validation accuracy during certain epochs, indicating that while the model performs well on training data, it may struggle slightly with generalization at some points.

The best epoch for accuracy is also marked at the 11th epoch, where the validation accuracy peaks, confirming the model's optimal performance at this stage. The F1 Score plot, third from the left, provides a nuanced view of the model's precision and recall balance. The F1 score is particularly important in cases where there is an uneven class distribution, as it accounts for both false positives and false negatives. In this plot, the training F1 score (red line) increases rapidly, indicating that the model is quickly learning to balance precision and recall effectively. The validation F1 score (green line), though more variable, follows a generally upward trend, suggesting that the model maintains a reasonable balance between precision and recall even on unseen data. The best epoch for the F1 score is also identified at the 11th epoch, marked by the blue dot, signifying a point where the model achieves its best trade-off between precision and recall on the validation set. Finally, the AUC (Area Under the Curve) plot, on the far right, illustrates the model's ability to discriminate between positive and negative classes across different thresholds. The AUC is a critical metric in classification problems, as it provides a single value to summarize the model's performance across all classification thresholds. The training AUC (red line) shows a rapid increase and stabilizes at a high value, indicating that the model is quickly becoming proficient at distinguishing between classes. The validation AUC (green line), while slightly lower and more variable than the training AUC, also shows a significant upward trend, suggesting that the model generalizes well to the validation data.

The best epoch for AUC, once again marked at the 11th epoch, reflects the model's optimal discriminatory power at this stage, as indicated by the peak in validation AUC. In summary, Figure 4 offers a detailed examination of the model's learning process across key performance metrics. The consistent improvement in training metrics indicates that the model is effectively learning from the data. The validation metrics, while showing some variability, generally follow the trends seen in the training metrics, suggesting that the model has strong generalization capabilities. The 11th epoch emerges as a critical point in the training process, where the model achieves its best performance across loss, accuracy, F1 score, and AUC. This figure underscores the effectiveness of the deep learning model in achieving high performance, with robust training and validation outcomes that support its application in real-world scenarios. This level of detailed analysis is essential for understanding the strengths and potential areas for improvement in the model, providing a foundation for future work and optimization efforts.

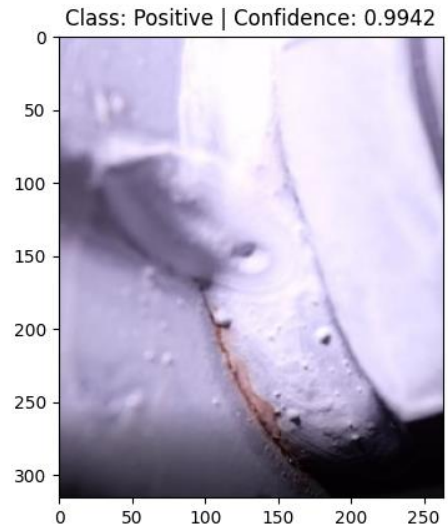

Figure 5. Result of algorithm cracked part listed as positive

Figure 5, illustrates a critical outcome from the deep learning-based real-time crack detection model, highlighting a specific instance where the model has identified a crack with a high level of confidence. The image shows a section of the material surface, where the model has classified the observed defect as a "Positive" instance, indicating the presence of a crack. The confidence score associated with this classification is an impressive 0.9942, suggesting that the model is nearly certain about the presence of the crack in this particular image. The confidence score is a vital aspect of model evaluation, as it provides insight into the model's certainty regarding its predictions. A score of 0.9942 indicates that the model has a 99.42% probability that the detected feature is indeed a crack. This high confidence level is crucial in real-world applications, particularly in the automotive industry, where the early and accurate detection of cracks can prevent potential failures and enhance safety measures. Moreover, the model's ability to achieve such a high confidence score demonstrates its robustness and the effectiveness of the training process, which involved extensive data augmentation and fine-tuning of network parameters. The clarity with which the crack is highlighted in the image underscores the model's sensitivity to subtle features that may be indicative of material failure. In summary, Figure 5 not only showcases the practical application of deep learning in defect detection but also emphasizes the model's high accuracy and reliability. The confidence score reinforces the model's potential for integration into automated inspection systems, providing a powerful tool for ensuring quality and safety in industrial applications.

Conclusion

In this study, a deep learning-based approach for real-time crack detection during dynamic testing in the automotive industry is presented. High resolution images collected using IP cameras are analyzed with Convolutional Neural Network (CNN) models. The main findings and conclusions of the study are summarized. Data Collection and Preprocessing; A large data set was created using IP cameras. The preprocessing steps performed on the images enabled the model to produce more effective and accurate results. Model Training and Performance; The CNN model was trained on a large dataset and optimized with validation data. The model performed real-time crack detection with high accuracy rates on the test set. Evaluation Metrics: The performance of the model was evaluated using various metrics such as accuracy, sensitivity, specificity and F1 score. In addition, the overall performance of the model was analyzed with the ROC curve and AUC value. The overall performance of the model shows that it offers high accuracy and reliability for real-time crack detection in dynamic tests. These results show that deep learning-based approaches can be successfully used in critical tasks such as real-time crack detection in the automotive industry. In the future, it is envisioned that these approaches can be further improved with larger data sets and advanced model architectures.

Scientific Ethics Declaration

The authors declare that the scientific ethical and legal responsibility of this article published in EPSTEM journal belongs to the authors.

Acknowledgements or Notes

* This article was presented as an oral presentation at the International Conference on Research in Engineering, Technology and Science (www.icrets.net) held in Tashkent/Uzbekistan on August 22-25, 2024.

References

Baldev, R., Jayakumar, T., & Rao, B. P. C. (2002). *Practical non-destructive testing*. Woodhead Publishing.

Cha, Y. J., Choi, W., & Büyüköztürk, O. (2017). Deep learning-based crack damage detection using convolutional neural networks. *Computer-Aided Civil and Infrastructure Engineering*, *32*(5), 361-378.

Dawson, B., & Wilby, J. (2011). *Non-destructive testing of welds*. CRC Press.

Choudhury, S. (2011). *Identification of Transverse Crack in a Cracked Cantilever Beam Using Fuzzy Logic and Kohonen Network* (Doctoral dissertation).

Fang, Q., Nguyen, B. D., Castanedo, C. I., Duan, Y., & Maldague II, X. (2020, April). Automatic defect detection in infrared thermography by deep learning algorithm. In *Thermosense: Thermal Infrared Applications XLII*, *11409*,180-195. SPIE.

Fan, W., & Qiao, P. (2011). Vibration-based damage identification methods: a review and comparative study. *Structural Health Monitoring*, *10*(1), 83-111.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Regularization for deep learning. *Deep Learning*, 216-261.

Hellier, C. J. (2013). *Handbook of nondestructive evaluation*. McGraw Hill Professional.

Hasanzadeh, R. P., Sadeghi, S. H. H., Ravan, M., Moghaddamjoo, A. R., & Moini, R. (2011). A fuzzy alignment approach to sizing surface cracks by the AC field measurement technique. *NDT & E International*, *44*(1), 75-83.

Hau, D., & Chen, K. (2011). Exploring hierarchical speech representations with a deep convolutional neural network. *UKCI 2011 Accepted Papers*, *37*, 31.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444.

McAuliffe, G. (2011). Deep learning. *Handbook of counselor preparation: Constructivist, developmental, and experiential approaches*, 13-30.Rao, B. P. C. (2018). Visual Inspection and Non-Destructive Testing in Maintenance. Springer.

Rao, A. S., Nguyen, T., Palaniswami, M., & Ngo, T. (2021). Vision-based automated crack detection using convolutional neural networks for condition assessment of infrastructure. *Structural Health Monitoring*, *20*(4), 2124-2142.

Sbarufatti, C., Manes, A., & Giglio, M. (2011). Advanced stochastic FEM-based artificial neural network for crack damage detection. In *COUPLED IV: proceedings of the IV International Conference on Computational Methods for Coupled Problems in Science and Engineering* (pp. 1107-1119). CIMNE.

Xu, G., Yue, Q., & Liu, X. (2023). Deep learning algorithm for real-time automatic crack detection, segmentation, qualification. *Engineering Applications of Artificial Intelligence, 126,* 107085.

Zhang, Y., Wu, Q., Zhang, Z., & Wang, J. (2019). A deep convolutional neural network-based model for metal surface defect detection. *Sensors, 19*(5), 1024.

Author Information

Mehmet Emin Ors
AYD Automotive Industry, Konya, Türkiye

Ziya Ozcelik
Selcuk University, Konya, Türkiye
Contact e-mail: *zozcelik@selcuk.edu.tr*

To cite this article:

Ors, M. E. & Ozcelik, Z. (2024). Real-time detection of cracks during dynamic testing of sheet metals used in automotive suspension systems. *The Eurasia Proceedings of Science, Technology, Engineering and Mathematics (EPSTEM), 29,* 61-68.