IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

# RAILWAY TRACK CRACK DETECTION WITH YOLOv5 AND GEOSPATIAL LOCALIZATION

**Appaji Guruvu**
Assistant Professor, Information Technology, GVPCEW
Visakhapatnam, Andhra Pradesh, India
appajeeguruvu@gvpcew.ac.in

**K. Bhagya Sree**
Information Technology, GVPCEW
Visakhapatnam, Andhra Pradesh, India
20jg1a1222.bhagya@gvpcew.ac.in

**V. Ommika Sai**
Information Technology, GVPCEW
Visakhapatnam, Andhra Pradesh, India
20jg1a1257.tejaswini@gvpcew.ac.in

**T. Kavya Sri**
Information Technology, GVPCEW
Visakhapatnam, Andhra Pradesh, India
20jg1a1251.kavya@gvpcew.ac.in

**M. Bhagya Sri**
Information Technology, GVPCEW
Visakhapatnam, Andhra Pradesh, India
20jg1a1236.bhagyasri@gvpcew.ac.in

**Abstract-** Railway infrastructure is a vital part of global transportation networks, enabling the efficient movement of people and goods. However, the safety and reliability of railway tracks are crucial for maintaining these networks. The research proposes a Deep learning-based method for identifying and localizing railway track cracks, a major concern in railway maintenance. The method uses two advanced Deep learning models, YOLOv5 for object identification and Efficient Net for classification tasks. The diversified dataset allows the model to perform better in real world situations and learn and generalize faster. The system classifies and finds cracks with great accuracy. And plots the locations of discovered cracks using geospatial coordinates for better comprehension and visualization. Transfer learning approaches improve the model's resilience and adaptability to new and untested data. The system's performance is evaluated through extensive trails and comparisons, showing significant improvements in precision, effectiveness and dependability, underscoring its potential for improving railway track maintenance and security.

**Keywords:** Deep learning, YOLOv5, Efficient Net, Classification, Crack identification, Coordinate Mapping.

## I.INTRODUCTION

The integrity of railway tracks is paramount for ensuring the safe operation of train systems. However, the presence of cracks in railway tracks poses a significant challenge to safety and operational efficiency. These cracks can propagate over time due to various factors such as cyclic loading and environmental conditions, leading to potential derailments or structural failures. Traditional methods of rail track inspection conducted by trained personnel. While these methods have been the standard approach for track maintenance, they come with inherent limitations. Manual inspections are often time-consuming, labor intensive, and subject to human error, making them less effective for detecting early-stage defects line hairline cracks that can be evolve into more significant issues over time. The proposed system comes with an advancement in computer vision and deep learning technologies having promising avenues for enhancing railway track inception and maintenance practices. Deep learning models, equipped with the capability to analyze large dataset classified into two classes defective and non-defective leverage advanced image processing techniques. These models can analyze high-resolution images of railway tracks and automatically detect cracks with high precision, even under challenging environmental conditions.

Fig 1: Defective track

Fig 2: Non-Defective track

## II.LITERATURE SURVEY

Railway track maintenance and safety have always been critical concerns in the transportation sector. Over the years, various studies have explored the use of advanced technologies and methodologies to enhance railway track inspection and defect detection processes. The following literature survey highlights key research works and developments related to automated defect detection systems, deep learning applications, and computer vision techniques in railway track maintenance.

[1] In the paper "Review on Deep Learning Techniques for Railway Infrastructure Monitoring" written by Maria Di Summa, Marielena Griseta, Nicola Mosca, Cosimo Patruno, Massimiliano Nitti, Vito Reno, and Ettore Stella. The authors worked on multiple techniques CNN, RNNSs and LSTM. They introduced the deep learning applications to identify potential gaps, or areas, in their existing system.

[2] In the paper "Deep Learning and Machine Vision-Based Inspection of Rail Surface Defects" written by Hongfei, Yanzhang Wang, Jiatang He, Zongwei Yao, and Qiuishi Bi. The authors worked on deep learning approach for precise defect localization and detection using 'You only look once'(YOLOv2) object detection and CNNs.

[3] In the paper "Advanced Automatic Detection of Cracks in Railway Tracks" written by B. S. Satish, P. Ganesan, A. Ranganayakulu, Sanjay S. Dola, S. Jagan Mohan Rao. The authors worked on deep neural network-based architecture to accurately detect track failures and hazards.

[4] In the paper "Analysis of Railroad Track Crack Detection using Computer Vision" written by Sumit Saha, Samit Karmakar, Debasmita Manna. The authors worked on machine learning algorithms, supervised learning approaches, to train computer vision model for accurate crack detection in rail road tracks.

[5] In the paper "Computer Vision System for Railway Track Crack Detection using Deep Learning" written by R. Thendral, A. Ranjeeth. The authors worked on algorithm Gabor transform and Deep Learning Neural Network. These techniques collectively contribute to the development of efficient computer vision system for automatic detection of cracks in railway tracks achieving high accuracy rate and low overall error rate.

## III.METHODOLOGY

The dataset we used for this problem statement is "Railway Track Fault Images". Then we preprocessed the images like rotation, rescaling width shifting, shear range and flipping. This dataset moves into model for training and then tested on test data. We also assess the performance of the model by using evaluation metrics.

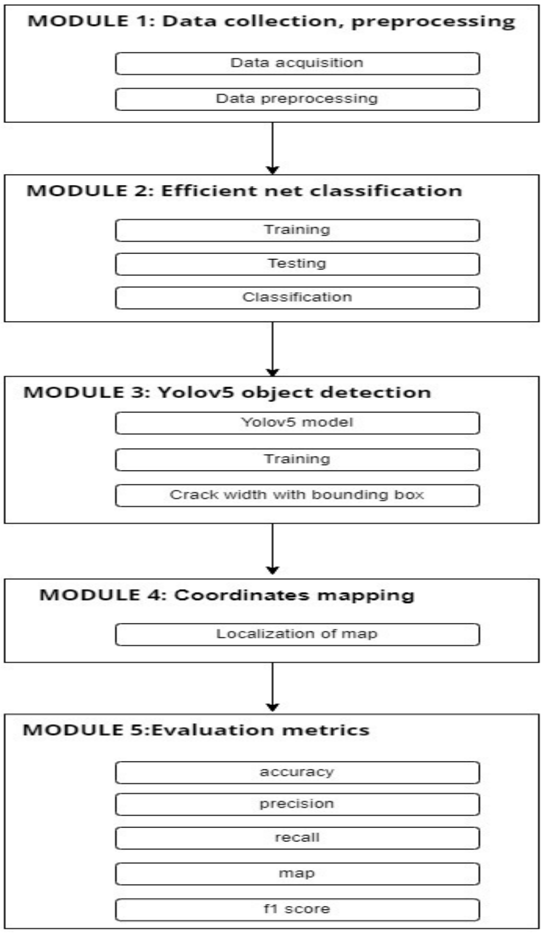


Fig 3: Data Flow of Proposed system

### 1.Data Acquisition:

We have used the dataset named as "Railway track fault detection". The dataset is made up of track images divided into 2 classes they are Defective and Non-Defective containing 1074 images. The images are represented in JPEG format.

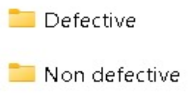


Fig 4: Dataset classes

### 2. Data Preprocessing:

Data Preprocessing is the process of converting raw data into format suitable for machine learning algorithms. It is done to improve the quality and enhance the performance of data. In the context of preprocessing, rescaling ensures consistent image sizes, rotation and shear introduces variability, flipping augments data, width shifting diversifies features. This process prepares the input data for training.

### 3. Efficient Net Classification:

Efficient Net is a convolutional neural network built upon a concept called compound scaling. This concept addresses the longstanding trade-off between model size, accuracy, and computational efficiency. Efficient Net uses Mobile Inverted Bottleneck (MBConv) layers, which are a combination of depth-wise separable convolution and inverted residual blocks. Additionally, the model architecture uses the Squeeze and Excitation optimization to further enhance the model's performance. These layers allow the model to learn channel-wise feature dependencies and create attention weights that are multiplied with the original feature map, emphasizing important information.

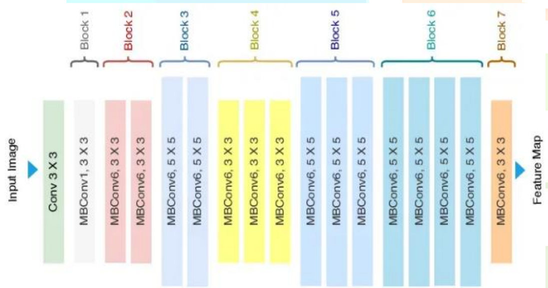


Fig 5: Efficient Net Architecture

[https://blog.roboflow.com/content/images/2024/04/image-1081.webp]

### A. Importing library – "TensorFlow"

TensorFlow is a free and open-source machine learning framework developed by Google Brain Team. It is used to build and train deep learning models. TensorFlow offers flexibility, scalability for performing tasks like image classification, natural language processing and reinforcement learning. To import TensorFlow library, simply write a command "import tensor flow as tf".

### B. Input Layer

Accepts a 124x124 RGB image.

### C. Convolutional Layer

These layers apply filters to produce feature maps.

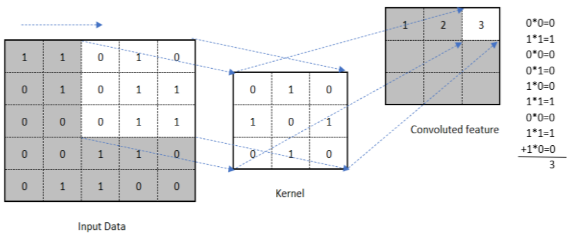


Fig 6: Convolutional Operation

### D. MB Convolution Layer

The Mobile Inverted Residual Bottleneck Convolution (MBConv) layer is a key building block in Efficient Net architectures, designed for efficient and effective feature extraction. It comprises several components: depth wise convolution, point wise convolution, batch normalization, and activation functions, typically swish. Firstly, depth wise convolution applies convolution separately to each input channel. Then, pointwise convolution combines these channels using 1x1 convolutions to adjust the number of output channels. The SE(Squeeze-and-Excitation) layer adaptively recalibrates important features and suppressing less relevant ones.

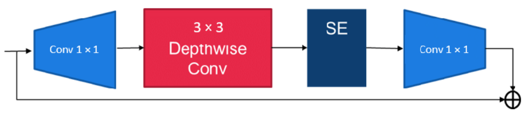


Fig 7: MB Convolution operations

[https://www.researchgate.net/publication/363701733/figure/fig5/AS:11431281085427560@1663763529155/The-architecture-of-the-MB-convolution.png]

### F. SoftMax Activation Function

Activation Functions are the mathematical operations applied over the neurons in the neural networks introducing non-linearity to the model for extracting complex patterns from the images. SoftMax is commonly used but also many other depending upon the problem. SoftMax is typically used in the final dense layer to produce probability distributions over multiple classes, making it suitable for classification tasks.

$$f(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}$$

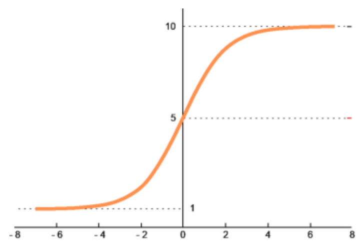

Fig 8: SoftMax Activation Function

**H. Batch Normalization**

Improves training stability and speed.

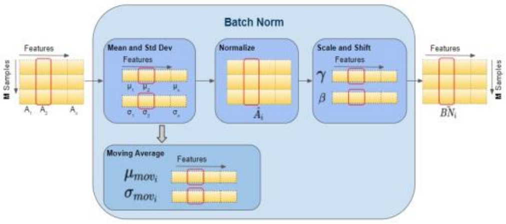


Fig 9: Sigmoid-activation function

**I. Dropout Layer**

Prevents overfitting randomly dropping neurons.

**J. Output Layer**

Outputs the class with highest probability.

**4. YOLOv5 Object Detection:**

YOLOv5 is a state-of-art object detection algorithm known for its speed and precision. It divides the input into a grid and predicts bounding boxes, class probabilities, and objectless scores for each grid. The model is trained on a dataset of annotated images containing examples of railway track cracks.

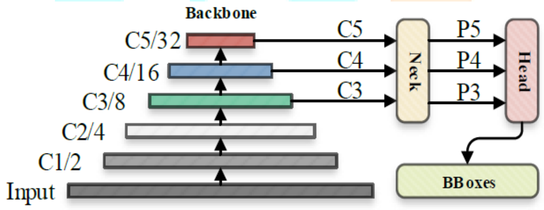


Fig 10: YOLOv5 Architecture

[https://miro.medium.com/v2/resize:fit:640/format:webp/1*VbSsY7414ogzpHmHnP7EVQ.png]

**A. Backbone: CSP Darknet Layer**

The Backbone employs CSPDarknet53, combining CBS and C3 modules for feature extraction. An SPPF modules enhances feature expression by optimizing max-pooling operations, improving both speed and efficiency in the network.

**B. Neck: PANet Layer**

YOLOv5 incorporates FPN and PAN methods. FPN up-samples feature maps from different layers to create new ones, like P3, P4, P5, enabling the detection of targets at various scales.

**C. Head: Yolo Layer**

The head of YOLOv5 predicts bounding boxes, class probabilities and objectness scores.

**5. Coordinates Mapping:**

In this research, we focus on mapping image coordinates to geographical locations using spatial transformation techniques. This process enables us to localize points of captured images onto a geographical map, enhancing spatial understanding and visualization through an interactive map-based interface.

**5.Model Execution:**

In this phase, we merge the capabilities of Efficient Net for image classification with YOLOv5 object detection expertise to create an integrated analysis pipeline. Initially, YOLOv5, utilizing its best.pt file, identifies objects within the images and delineates their bounding boxes with precision. Leveraging its model.h5 file to classify the localized object detection and classification, yielding a comprehensive and detailed understanding of the image's content, enhancing both accuracy and interpretability of image tasks.

**6. Model testing and Evaluation:**

When evaluating the effectiveness of a model intended for railway track crack detection and geospatial localization, model testing is a crucial stage. This stage involves measuring the accuracy, precision, recall and F1-score and mean average precision (mAP) of the trained model using “Railway track fault detection” dataset. Where in classification model 0 represents defective and 1 represents non-defective. Preprocessing is done before entering into the classification and detection module. The model testing and evaluation is done separately.

Formulas for evaluation metrics:

$$Accuracy: \quad \frac{TP + TN}{TP + TN + FP + FN}$$

$$Precision: \quad \frac{TP}{TP + FP}$$

$$F1\text{-}score: \quad \frac{2TP}{2\mathrm{TP} + FP + FN}$$

$$Recall: \quad \frac{TP}{TP+FN}$$

$$\mathrm{mAP} = \frac{1}{N}\sum_{i=1}^{N} \mathrm{AP}_i$$

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.60 | 0.27 | 0.37 | 11 |
| 1 | 0.53 | 0.82 | 0.64 | 11 |
| accuracy | | | 0.55 | 22 |
| macro avg | 0.56 | 0.55 | 0.51 | 22 |
| weighted avg | 0.56 | 0.55 | 0.51 | 22 |

| Class | Images | Instances | P | R | mAP50 |
|---|---|---|---|---|---|
| all | 223 | 219 | 0.681 | 0.682 | 0.678 |
| defective | 223 | 163 | 0.665 | 0.65 | 0.649 |
| non-defective | 223 | 56 | 0.697 | 0.714 | 0.707 |

## IV.SYTEM ARCHITECTURE

The suggested system design enhances gathering images data relevant to railway track cracks, preprocessing it via efficient net layers, and then categorizing the classification of class types based on features like crack color, texture, shape and structures. Users provide input data, such as track image to classify them. We use detection to determine the intensity of the crack by identifying the crack width, to access how sever the crack is. Various techniques are utilized to improve the quality and diversity of the data, such as resizing, rescaling, flipping, and rotation. The input data travels through multiple layers for feature extraction. Fig 11 represents how the proposed system works and flow of the input image through layers for classification and detection and localization of map.

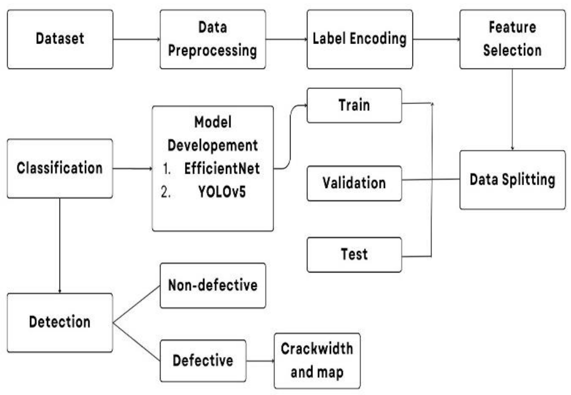


Fig 11: System Architecture

## V.IMPLEMENTATION DETAILS

### A. Python

We have developed this system using python. To work with DL models, python is efficient. Python offers wide range of libraries like numpy, keras, and pandas. The platform we used is google Collaboratory.

### B. Google Collaboratory

Google Collaboratory (Collab) is a cloud-based platform offering a jupyter notebook environment, allowing seamless code execution and collaborative research. It provides access to GPUs, enabling efficient development and execution of deep learning models.

## VI.RESULTS

**Efficient Net:**

We gained an accuracy of 91% by training the model with 30 epochs using the images from railway track fault detection dataset.

The fig.11 represents the visualization of train and validation set accuracy by plotting over accuracy and epochs.


Fig 11: Accuracy of Efficient Net

The below Fig.12 represents the heatmap illustrating the distribution of certain predictions. Basically, used in the context of classifications and object detection.

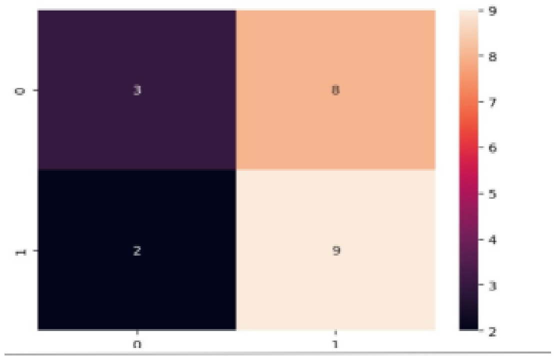

Fig 12: Heat map for classes

**YOLOv5:**

The graph from Fig.13 illustrates the progression of various metrics, such as training/validation loss, precision, recall, and F1 score, during the training process of a deep learning model. Monitoring these metric curves across iterations provides insights into model's learning dynamics, data fit, and generalization ability, enabling performance assessment and potential issue detection.

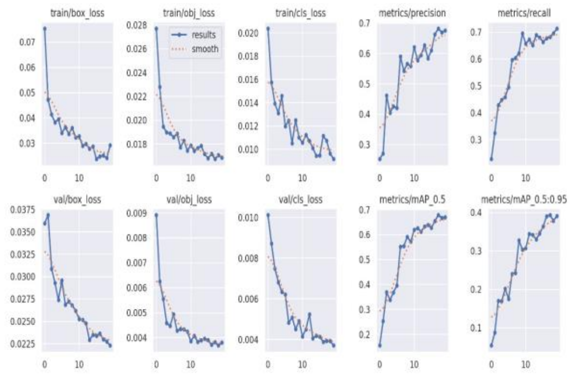


Fig 13: YOLOv5 evaluation metrics

The below Fig.14 represents the confusion matrix where the detections are on test set. The no. of true positive (TP), true negative (TN), false positive (FP) and false negative (FN) predictions that the model made is displayed in the confusion matrix for classes.

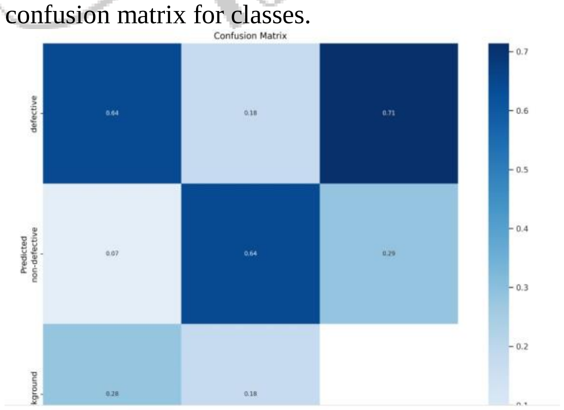


Fig:14 Confusion Matrix for classes

## VII.CONCLUSION

This project aims to develop a robust method for utilizing deep learning techniques, our railway track detection project demonstrates the potential of artificial intelligence in enhancing railway safety. By integrating YOLOv5 for object detection and Efficient Net for classification and integrating it with geospatial localization, this allows accurate localization of defects on a map. Streamlining maintenance prioritization and enabling timely interventions. The integrated approach enhances defect detection and classification accuracy, learning technologies, our project lays the groundwork for revolutionizing railway safety protocols, promising increased efficiency, reduced maintenance costs, and ultimately, safer travel for passengers and railway personnel.

## REFERENCES

[1] Maria Di Summa, Mariaelena Griseta 2, Nicola Mosca, Cosimo Patruno, Massimiliano Nitti, Vito Reno, and Ettore Stells, 29 August 2023, data of current version 20 October 2023, "Review on Deep Learning Techniques for Railway Infrastructure Monitoring"

[2] Hongfei Yang, Yanzhang Wang, Jiyong Hu, Jiatang He, Zongwei Yao, and Qiushi Bi, "Transactions on Instrumentation and Measurement", Vol-71, 2022, "Deep Learning and Machine Vision Based Inspection of Rail Surface Defects"

[3] B. S. Sathish, P. Ganesan, A. Ranganayakulu, S.S. Dola, and S.J.M. Rao, "Advanced Automatic Detection of Cracks in Railway Tracks," 2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS), Coimbatore, India, 2019, pp. 98-101, doi: 10.1109/ICACCS.2019.8728451

[4] S. Saha, S. Karmakar, and D. Manna, "Analysis of Railroad Track Crack Detection using Computer Vision," 2022 Interdisciplinary Research in Technology and Management, Kolkata, pp.1-4, doi: 10.1109/IRTM54583.2022.9791808

[5] R. Thendral and A. Ranjeeth, "Computer Vision System for Railway Track Crack Detection using Deep Learning Neural Network," 2021 3rd International Conference On Signal Processing and Communication (ICPSC), Coimbatore, pp,193-196, doi: 10.1109/ICSPC51351.2021.9451771