ENHANCING RAILWAY SYSTEM MAINTENANCE:

AN AI-DRIVEN APPROACH TO IMPROVE THE

PREVENTIVE MAINTENANCE OF RAIL TRACKS

AYOMIKUN MOSAKU

BSC. MECHATRONICS & COMPUTER SYSTEMS ENGINEEEING

2023/2024

24th April 2024

ACKNOWLEDGEMENTS

I thank my project supervisor, Dr Rasha Mohammad, for her continuous invaluable guidance and support throughout this project. Her expertise and counsel are assisting.

ABSTRACT

This research addresses leaf accumulation on railway tracks, a significant cause of train delays and safety hazards due to reduced wheel-rail adhesion. This study was motivated by the need to increase rail safety and efficiency, particularly during autumn when leaves fall, which significantly impacts rail operations. This research successfully developed a system capable of identifying leaf accumulation on railway tracks in real-time using advanced machine-learning models and computer vision techniques. The study involved the deployment of convolutional neural networks (CNNs), specifically the VGG19 model, which demonstrated exceptional performance with a test accuracy of 99.67%. This approach quantifies the extent of leaves falling during autumn and offers a technological solution that surpasses traditional methods in precision and applicability. The result of this study has an impact both in practice and theory. Integrating such a system into the rail network could mitigate the risks associated with low adhesion conditions, thereby enhancing rail services' safety, reliability, and customer satisfaction.

Theoretically, this research contributes to the ongoing discourse on rail network efficiency and customer satisfaction, aligning with strategic visions for Rail and the Rail Capability Plan. It lays a foundation for future research on rail maintenance and safety approaches based more on proactive than reactive strategy. The benefits of this research are manifold. It offers a framework for developing better real-time rail track monitoring and maintenance solutions. This study has the potential to transform the management of rail networks. However, despite the promising outcomes, certain aspects still need to be solved, such as the full-scale implementation of the system and its adaptation across different geographical and environmental conditions

Keywords**: Leaf Accumulation, Rail Tracks, Machine Learning, Computer Vision, Rail Safety, Real-time Detection, CNN, Preventive Maintenance, Transfer Learning, Railway, AI, Image Classification**

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

CNN - Convolutional Neural Network

AI - Artificial Intelligence

GB - Great Britain

VGG19 - Visual Geometry Group 19

MobileNetV2 - Mobile Network Version 2

ResNet50V2 - Residual Network 50 Version 2

InceptionV3 - Inception Version 3

CHAPTER 1: INTRODUCTION

1.1 INTRODUCTION

Rail travel is the most fuel and carbon-efficient mode of transport and is a critical part of the strategy to achieve net zero carbon emissions in Europe (Folorunso et al., 2023). It has become a popular transportation method in the United Kingdom and has increased in use over the years. From October to December 2023, 417 million rail passenger journeys were recorded in Great Britain (GB) in the latest quarter. This value presents a 20% increase on the 348 million journeys in the same quarter of the previous year (October to December 2022) (of Rail, 2024).

The UK currently has one of the most congested rail networks in Europe, with utilisation of UK railways being 60% higher than in the rest of Europe and expected to double in the next 25 years (Louise & Kirby, 2019). However, specific barriers to railway use threaten its wider adoption; hence, these challenges must be tackled to accommodate the increased rail system use.

Some obstacles to the UK railway include rail delays, train cancellations, crumbling infrastructure challenges, and operational and service issues. Research on the causes of railway barriers is strongly needed, and effective methods for mitigation and control are paramount.
Network Rail, the UK rail infrastructure operator, has stated that delay reduction is a top priority to reduce delays by 15% by 2024 (Louise & Kirby, 2019) as a delay on the route in one part of the country has the potential to delay a fleet hundreds of miles away. Therefore, understanding the cause of one delay that may cause several other issues along the rail network and making any attempts to mitigate the initial delay has the potential to exponentially delay across the country (Louise & Kirby, 2019).

Network Rail (2024) identified critical reasons for delays along a rail network, including buckled rails, engineering works, fatalities, flooding, knock-on delays, leaves on the line, landslips, winter weather—snow and ice, vandalism and trespass, broken Rail, and storms. The list from Network Rail shows that train delays can be attributed to the weather and the environment, with some weather-related delays occurring in the autumn and early winter (Network Rail, 2024).

During autumn, the leading cause of delays on the rail network is low adhesion between the wheel and the rail track due to the leaves falling on the rail track (Louise & Kirby, 2019). Field observations and laboratory tests indicated that these fallen leaves are crushed in the wheel-rail contact, leaving a black, well-bonded, slippery layer on the rails. In wet conditions, this black layer can reduce the maximum friction coefficient between the wheel and the Rail to 0.01 (Watson et al., 2020). Adverse weather conditions such as drizzles and rain also enhance this layer's formation.

There are numerous adverse effects of low adhesion. Firstly, the slippery layer makes train movement on the track complex, challenging the train wheels to maintain grip with the Rail, thus leading to train slip or derailment. As a result, trains must slow down on the slippery sections and adjust their braking and acceleration to a slow pace to avoid slipping. This situation often results in delays and disruptions, with extreme cases of trains stopping entirely until the tracks are cleared and safe for travel. For instance, in October 2022, a freight train halted on a busy TransPennine route due to leaves, causing a nine-hour delay before the track was cleared and the train could move again (Network Rail, 2023).

Another adverse effect of low adhesion is station overruns. It occurs when a train fails to stop at a station or the designated point along a platform because its braking performance is compromised when entering the station due to low adhesion. Due to the slippery conditions, the driver applies the brakes earlier and more gently to avoid wheel slides, which may lead to the train stopping too far along the platform or, in more severe cases, not stopping until it has passed the platform entirely, causing delay. An example was the low adhesion event at Chester Station in November 2013, where a Virgin Trains empty class 221 approached Chester Station under initially overcast and dry conditions. However, as the train neared the station, rain began to fall, which increased the risk of low adhesion on the rail head due to the wet conditions; the train was unable to stop in time and overran its intended stopping point, resulting in a collision with the buffer stop (Louise & Kirby, 2019).

Signals passed at danger (SPADs) is another adverse effect of low adhesion, and it occurs when a train passes a signal displaying 'stop' without authorisation due to the train's inability to brake effectively (Chen & Fukagai, 2018).

Another known effect of low adhesion is Station Stop Incidents (SSI). An example of an SSI incident is station stop shorts, which occur when a train has applied its brakes too early and not all the doors can be opened to allow passengers to alight the train. SSI can have numerous consequences, ranging from operational risks to safety for alighting passengers to delay, repair costs, and knock-on delay effects impacting the entire rail network.

Finally, delays also occur when the accumulation of leaves on the railway obstructs the train wheels and the track's electrical components. This prevents the accurate determination of the position of the trains. For cases of uncertainty on the precise location of a train, subsequent trains must wait at red signals until the position of the leading train is confirmed to ensure that the control room can maintain a safe separation between the trains.

These delays disrupt train operations, lead to train delays across the network, and cost the UK rail industry over £350 million annually (Folorunso et al., 2023). Results have also proven that delays caused by low adhesion are the highest cause of customer dissatisfaction with rail services (Folorunso et al., 2023).

Countermeasures like spraying sand and ceramic particles onto the wheel/rail contact zone will not do (Chen & Fukagai, 2018). This indicates a need for more innovative solutions to tackle the problem of rail delays and improve the reliability of rail services

The research successfully developed a real-time system to identify leaf falls on railway tracks by deploying advanced machine-learning models and computer vision techniques. This approach demonstrated high accuracy, with the VGG19 model showing exceptional performance, achieving a test accuracy of 99.67%1. This study's results quantify the extent of the problem and offer a technological solution that surpasses traditional methods in precision and applicability. The findings suggest that integrating such a system into the rail network could significantly mitigate the risks associated with low adhesion conditions, enhancing rail services' safety, reliability, and customer satisfaction.

By focusing on the limitations of current practices, this research contributes to the ongoing discourse on rail network efficiency and customer satisfaction, aligning with the Department for Transport's 'Strategic Vision for Rail' and the 'Rail Capability Plan'. It provides a foundation for future studies to explore alternative methods and technologies that could offer more effective solutions to the problem of leaf-induced low adhesion on rail tracks. The implications of this work for future research are significant. It suggests a substantial opportunity for developing novel approaches that can preemptively detect and address leaf accumulation, potentially through advanced machine learning algorithms and computer vision. Such technologies could lead to more targeted and efficient maintenance operations, ultimately enhancing the safety and punctuality of rail services.

Beyond immediate operational applications, this work also establishes the foundation of future research for rail maintenance and safety, offering a shift for predictive maintenance over reactive measures. Machine learning and computer vision create an avenue for new predictive maintenance strategies, which can transform how rail networks are monitored and maintained.

Given the detailed analysis of leaf accumulation on railway tracks and its effect on rail adhesion, which results in operational and safety concerns, the research question for this study is: **"How can the integration of computer vision and machine learning technologies improve the detection and mitigation of leaf accumulation on railway tracks, thereby enhancing rail adhesion and reducing operational disruptions?"**

The structure of this paper is designed to address the research question. Firstly, the introduction sets the context and rationale for the research, followed by a literature review, which reviews existing methods for detecting and mitigating the challenge of leaf accumulation on rail tracks and details the limitations of existing technologies and practices. Next is the methodology and materials section, which describes the development of a machine-learning model using computer vision and procedural details related to gathering the data used to train and evaluate the model. The results section presents a comparative analysis of several deep learning architectures on a benchmark test dataset. The discussion section interprets the result in light of earlier research and summarises the implications for new research and the development of modern technologies. Finally, the paper concludes by summarising the study's key insights and contributions to rail maintenance and safety.

1.2 AIM OF STUDY

The primary aim of this study is to enhance the reliability and efficiency of the UK rail network by integrating advanced machine learning and computer vision technologies. This integration seeks to improve the detection and mitigation of leaf accumulation on railway tracks, a major cause of train delays and safety hazards during autumn. By leveraging these technologies, the study aims to develop a real-time detection system that can accurately identify leaf accumulation and enable prompt maintenance actions.

1.3 OBJECTIVES OF STUDY

The specific objectives of the study are outlined as follows:

- To design and train a machine learning model using convolutional neural networks (CNNs) that can effectively recognise and quantify leaf accumulation on railway tracks from visual data.
- To implement a real-time monitoring system that integrates the trained model to continuously assess the condition of rail tracks and provide timely alerts for maintenance action.
- To evaluate the accuracy and efficiency of different CNN architectures, including VGG19, MobileNetV2, ResNet50V2, InceptionV3, and Xception, in order to select the most effective model for detecting leaf accumulation on railway tracks.
- To evaluate the impact of the real-time detection system on reducing delays, improving train punctuality, and enhancing overall customer satisfaction within the rail network

CHAPTER 2: LITERATURE REVIEW

2.1 LITERATURE REVIEW

Studies have analysed the conditions where leaves on the line cause adhesion, leading to countermeasures to control or prevent slipping and sliding (Folorunso et al., 2023) . Such countermeasures include

2.1.1 Leaf-busting Trains

Introduced by Network Rail in 2019, its leaf-busting trains move around the railway, cleaning the top of the rail by spraying it with a high-pressure water jet to blast away leaf mulch accumulating on the rails (Network Rail, 2021). After cleaning, the trains also apply a gel of sand and steel grains to improve traction by increasing the friction between the train wheels and the tracks (Network Rail, 2021). In addition to the trains, manual leaf-busting teams are deployed across key locations to scrub the rails by hand with a sand-based treatment (Network Rail, 2021).

Fig 2.1 Leaf Bursting Trains

2.1.2 Leaf-busting Detection Of Leaves

The leaf-busting trains do not detect leaves on the lines but do so through weather monitoring and forecasting (Network Rail, 2021).
Network Rail employs specialist weather forecasters who monitor weather conditions and provide forecasts, including leaf fall estimates (Network Rail,

2021). Between October and December, the specialists receive forecasts twice daily and highlight locations that may need more attention due to leaf fall. Based on these forecasts, Network Rail deploys leaf-busting trains and teams to the areas most in need, ensuring the most effective use of resources (Network Rail, 2021).

In addition, some train operators publish special autumn timetables for areas with heavy leaf fall that allow extra time for journeys. This accounts for the more cautious driving required during heavy leaf-fall periods (Network Rail, 2021).

2.1.3 Water-trak System

Network Rail further researched to counter low adhesion, which led to the discovery of the water-trak system (Network Rail, 2023). However, as of the research for this paper, this innovation was still in the testing phase, with multiple positive trial results suggesting that water-trak is an effective way to deal with leaves on the track (Network Rail, 2023).

The water-trak system works based on the principle of simulating the conditions of a rainy day by applying lesser amounts of water from the underside of the train onto the railhead during low adhesion conditions (Network Rail, 2023). Though paradoxical, research showed that while the low adhesion condition is formed by forming a liquid-like slippery gel on the track, the situation can be improved by making the rail head wet in low adhesion conditions so trains can once more run smoothly along the line. This is evident in actual world circumstances where train wheels glide more easily along the rail during and after heavy rainfall despite leafy conditions, and the trains also brake faster, preventing train delays. (Network Rail, 2023).

Figure 2.2 Water Trak System

2.1.4 How The Water-trak System Differs from The Leaf-busting System

The water-trak and leaf-busting systems are similar, but there are a few distinctions in their principles of operation and method of tackling low adhesion.

Firstly, the water-Trak system improves adhesion by wetting the rail head, simulating rainy conditions that naturally wash away leaves. In contrast, leaf-busting trains remove leaves with high-pressure water jets and then apply an adhesion modifier to the cleaned rail head.

The water-trak system operates based on the assumption that low adhesion conditions are always present on the tracks during autumn. It wets the rail head continuously as the train glides along the track due to the absence of integrating a leaf detection and low adhesion system with the Water-Trak System. On the other hand, based on forecasts by specialist weather forecasters, the Leaf-Busting trains are deployed to the areas most in need.

The Water-Trak system was trialled on two electric trains in northwest England in autumn 2021 and two diesel trains in Yorkshire (Network Rail, 2023). Both trials turned out positive, indicating that the Water-Trak system's design allows for its installation on various train models regardless of their power source, therefore leading to the conclusion that this system has the future potential to be applied across a wide range of trains operating in regions prone to leaf fall and other contaminants that reduce rail adhesion. On the other hand, the leaf-busting trains are separate specialised trains deployed specifically for track cleaning. As of when this research paper was written, there has yet to be a publication on whether a leaf-busting system can be integrated into numerous types of trains.

2.1.5 A Machine Learning-based Approach

Although studies have analysed the conditions where leaves on the line cause adhesion, which in turn have led to measures to mitigate the dangers of low adhesion on tracks and prevent delay, these solutions, such as the leaf-busting trains and water-track systems, while innovative, have their limitations, primarily due to their broad-scale application and the inherent uncertainties in weather predictions affecting their precision and effectiveness (Folorunso et al., 2023). Hence, there is a great need for studies on technologically advanced methods for predicting and preventing low adhesion between trains and rails. Unfortunately, there are limited studies on this.

(Folorunso et al., 2023) 's study, "A Machine Learning Approach for Real-Time Wheel/Rail Interface Friction Estimation," introduces an advancement in addressing the issue of low adhesion on railway tracks. Machine learning was used to predict wheel/rail interface friction in real-time with remarkable accuracy, as evidenced by an R^2 value of 0.97, indicating high precision in friction prediction.

Current forecasts in the rail industry give regional-level predictions. However, it is well-known that friction conditions can change dramatically over a few hundred meters (Folorunso et al., 2023). The primary advancement in Folorunso et al. 's work thus lay in its ability to predict friction conditions at a precise location, thereby overcoming the limitations of regional-level forecasts that fail to account for the dramatic variations in friction conditions over short distances. This approach can significantly enhance the rail industry's ability to predict and mitigate low adhesion conditions caused by leaves on the track, thereby reducing delay.

2.2 THE GAP

In 2016, the low adhesion and leaf fall-related issues on the Piccadilly Underground line cost £20 million in lost customer hours (Louise & Kirby, 2019). This prompted an independent incident report to be carried out to investigate the severity of the delays (Louise & Kirby, 2019). The report stated that there are four critical elements of an effective adhesion regime:

- Prevention – to retain adhesion at safe levels through vegetation control and early warning systems.
- Prediction – use of reports to trigger early warning indicators of poor adhesion days in the future.
- Containment – short-term action to reduce low adhesion, such as railhead cleaning techniques.

- Recovery – which is reducing the effects of a low adhesion event such as changes in train timetable(Louise & Kirby, 2019).

The leaf-busting trains and water-trak system described in this literature fall under prediction and containment measures, while (Folorunso et al., 2023) study focuses on prediction measures. However, there are more containment methods than prediction methods.

The disadvantage of the leaf fall, or low adhesion prediction approach, lies in its subjection to the inherent uncertainties and inaccuracies of weather predictions. Leaf fall prediction is never 100% due to the unpredictable nature of weather and its impact on vegetation. Also, while effective on a broad scale, leaf-busting trains may only target some areas with high precision. Given the scale of railway networks and the focus on critical locations, some areas may be overlooked or receive less frequent treatment.

While the water-trak system can be applied to a wide range of trains without an integrated mechanism for detecting leaves or low adhesion conditions in real-time, its application may also need to be more targeted and efficient. Research is needed on a more novel approach that shifts from a predictive, weather-based approach to a real-time, data-driven method.

2.3 PREVENTIVE MAINTENANCE USING COMPUTER VISION

Preventive maintenance refers to the systematic execution of regularly planned maintenance tasks to proactively avoid unforeseen problems in the future. By leveraging machine learning, engineers can enhance maintenance practices and mitigate risks associated with the dependability of facility operations using predictive maintenance software.

In the context of this research paper, predictive maintenance involves swiftly implementing mitigation strategies in specific areas where leaves fall before the formation of the slippery layer, thereby preventing low adhesion, enhancing safety, and reducing delays.

Innovation and investment in the UK rail industry remain a vital focus of the UK Government. The UK railway industry has a solid government-led vision to create a fully integrated smart transport system that can be fully operational across the country (Louise & Kirby, 2019). By focusing on the detection and prevention of leaf-induced low adhesion conditions, this research can contribute to the development of more effective, real-time solutions for rail track monitoring and maintenance that align with the vision of the UK government.

Contrary to predicting leaf fall using specialised weather forecasts, employing preventive maintenance to detect leaf fall offers a more direct, technologically advanced, and more accurate method of identifying and addressing delays caused by low adhesion.

With the advent of technology and advancements in Artificial Intelligence, applying Deep Learning Algorithms in Computer Vision for Pattern Recognition has become a viable approach (Louise & Kirby, 2019). This implies that computer vision can be utilised to identify leaves on cluttered rail tracks and differentiate them from clean rail tracks by analysing images captured by specialised Rail Cameras (Chollet, 2017; Devlin et al., 2019).

Computer vision's capability to process real-time imagery from these dedicated cameras enables the precise detection of localised leaf accumulations that might not warrant deploying a specialised system such as a leaf-busting train, thereby enabling more timely, accurate, and efficient maintenance operations (Louise & Kirby, 2019).

Another advantage of this method lies in its non-reliability on predictions and speculations but on the actual conditions observed on the rail. Additionally, computer vision could offer a more comprehensive monitoring solution for areas that might be overlooked due to the scale of railway networks, thereby ensuring that all areas are addressed (Chollet, 2017).

CHAPTER 3: METHODOLOGY AND MATERIALS

3.1 OVERVIEW

By applying computer vision, this project intends to develop a machine-learning model that predicts and prevents low adhesion conditions on railway tracks by identifying leaf falls. These low adhesion conditions are known to create significant delays and represent major threats to the safety of the network.

The design behind this study was built on the need for a novel predictive and preventive approach that surpasses the limitation of current demand and preventive measures, which are often hindered by weather forecast uncertainties and the application of the solutions at scale. This study draws from Folorunso et al.'s (2023) paper, which proposed the potential of using machine learning to estimate real-time wheel-rail interface friction. Using computer vision, this study intends to develop an accurate real-time detection and monitoring system to detect leaf fall on railway lines. This will address some of the limitations of the current detection methods, such as the leaf-busting trains and the Water-Trak system.

This study's theoretical framework is based on machine learning and computer vision principles. It applies Convolutional Neural Networks (CNNs) to the image processing task, as CNNs have proven superior to traditional image analysis techniques (O'Mahony et al., 2019). The research employs a quantitative approach to processing and analysing high-resolution images of railway tracks.

The study utilised real-world data from rail track monitoring systems and augmented it with field data from high-resolution cameras mounted on inspection vehicles. The images were obtained under various circumstances, including weather conditions, corresponding leaf accumulations, and different times of the day. They included details of the railhead's state, including leaves and other debris. A convolutional neural network (CNN) model for computer vision was then trained based on the collected data to respond to the presence or absence of leaf accumulation on the tracks.

3.2 THEORETICAL FRAMEWORK

3.2.1 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are the backbone of deep learning systems, especially in computer vision. Inspired by the animal visual cortex, their architecture enables the deep learning of spatial hierarchies of features from images, which is appropriate for identifying leaf accumulation on train tracks. The grid-like layout of rail track images made CNNs suitable for processing and evaluating the data for regions of interest (LeCun et al., 2015). CNNs consist of convolutional, pooling, and fully connected layers, each performing various modifications on the input data. Convolutional layers with learnable kernels record spatial dependencies and encode basic image features, such as edges and textures. Pooling layers reduce spatial dimensions, which decreases the computing load and also controls overfitting, while fully connected layers synthesise the learned features to recognise complex patterns (Krizhevsky et al., 2012).

These layers work in unison to distinguish leaves by the textural and colour differences they exhibit against the rail track backdrop for rail track leaf detection. The convolutional layers detect the precise shapes and edges of leaves, while the pooling layers make the detection resilient to leaf size and orientation variations. The fully connected layers then interpret these features to determine leaf accumulation.

CNN architectures have also seen significant improvements that enhance their performance. Alex Net, VGGNet, Google Net (Inception), and ResNet are examples of architectures that have pushed the limit of image recognition. These architectures introduce deeper networks, innovative modules, and residual learning, all of which facilitate improved feature extraction and classification.

The depth and richness of neural networks are necessary for the vast range of features available in images of rail tracks in autumn. This learning ability is essential for designing a system to identify leaf accumulation on rail tracks under various environmental conditions, at different times of day, and with different camera angles.

CNN's adaptability goes beyond image classification: they have been successfully employed in applications such as video analysis, natural language processing, medical image analysis, and autonomous vehicles. Underlying this broad adaptability is the principle of transfer learning: the notion that learning in one domain can be used to facilitate learning in a related one. In the case of rail track leaf detection, transfer learning means the application of CNNs allows the use of CNNs pre-trained on large datasets, such as ImageNet, to be fine-tuned for the specific task.

3.2.2 Transfer Learning

Transfer learning – the reuse of information learned in one task to improve performance in a related but distinct task. It is a powerful tool in computer vision. It is beneficial when the target task's data is limited, as is often the case in specialised applications, such as detecting leaf accumulation on rail tracks.

With transfer learning models trained on ImageNet, a reference database of 14 million images and 22,000 categories can be fine-tuned for new tasks such as object identification or image segmentation. This benefits the study's aim of identifying leaves on rail tracks. Transfer learning is based on two approaches: fine-tuning and feature extraction (Yosinski et al., 2014).

Fine-tuning involves tweaking the parameters of a pre-trained model and further training it on the data to adapt the model to the new task. Alternatively, feature extraction involves using the extracted features from the pre-trained model as input to a new model explicitly trained for the target task.

Transfer learning offers substantial advantages for identifying leaves on rails: It speeds up training and alienates the need to begin model building from scratch. It also gives the model a better chance of generalising from training to real-world data. Furthermore, transfer learning is beneficial when dealing with a small amount of domain-specific data since the pre-trained model offers a solid foundation for development.

Additionally, transfer learning exploits the hierarchical nature of the features learned by deep neural networks. Lower-level features (edges and textures) are more transferrable across domains, whereas higher-level features are more task-specific (Yosinski et al., 2014). Using these transferrable lower-level characteristics allows the model to focus on learning the higher-level, task-relevant features necessary to detect leaves on rails.

Transfer learning has disadvantages. Its success lies in the similarity between the source and target tasks. If they differ substantially, information transmitted from the source task might prove irrelevant and limit the learned model's performance on the target task. Furthermore, fine-tuning a pre-trained model involves careful evaluation of the model architecture and training parameters to avoid overfitting the target dataset.

3.3 MODELS USED FOR TRANSFER LEARNING

3.3.1 VGG19

The VGG19 model is a 19-layer deep convolutional network with 16 convolutional, 3 fully connected, 5 MaxPooling, and 1 SoftMax layer. It employs 3x3 convolutional filters and 2x2 MaxPooling to handle 224x224 pixel pictures, classifying them into 1000 categories based on ImageNet training. It is versatile for various tasks, allows C/C++ code generation for deployment, and is available in pre-trained or untrained forms (Simonyan & Zisserman, 2014).

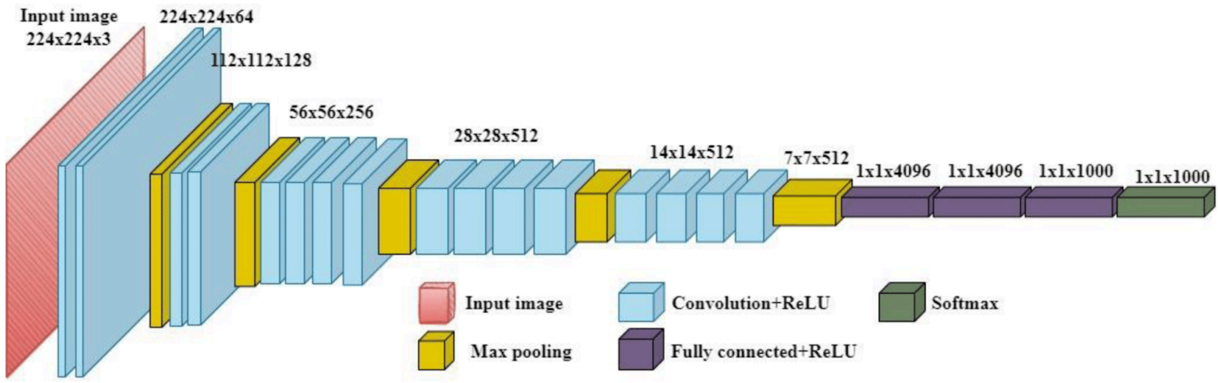


Fig 3.1. A depiction of the VGG19 neural network architecture

Advantages of VGG19

- High Accuracy: One of VGG19's advantages is its high accuracy in image classification tasks (Simonyan & Zisserman, 2014). Its accuracy makes it the ideal option for locating leaf accumulation sites precisely.

- Simple Architecture: The model facilitates comprehension and adjustment (Simonyan & Zisserman, 2014), enabling it to adapt to the peculiarities of rail track images.
- Effective Feature Extraction: The model was trained to extract detailed image features (Simonyan & Zisserman, 2014). This capacity ensured that the model accurately identified leaf accumulation sites under varied environmental conditions.

Limitations of VGG19

- Large Model Size: One downside of the VGG19 model was its size, which made it computationally demanding and memory-intensive. Its size challenges real-time applications where swift processing is vital (Simonyan & Zisserman, 2014).
- Image Classification Focus: The model's architecture was built primarily for image classification tasks, restricting its performance in non-classification tasks like segmentation (Simonyan & Zisserman, 2014).
- Fixed Architecture: While VGG19's pre-defined structure could be suitable for several applications, considerable modifications to the architecture might be required to achieve the best results for other applications (Simonyan & Zisserman, 2014).

3.3.2 MobileNetV2

Sandler et al. created MobileNetV2, an improvement on the original MobileNet designed to be more effective in embedded and mobile applications. It incorporates depth-wise separable convolutions and a novel inverted residual structure for enhanced performance. At its core, MobileNetV2 incorporates inverted residual blocks with 1x1 convolutions for extending features, depth-wise convolutions for spatial filtering, and successive 1x1 convolutions for lowering dimensionality. Linear bottlenecks and shortcut connections reinforce this model. Using depth-wise separable convolutions reduces the model's parameters and computational demand, making it highly efficient. Linear bottlenecks and inverted residuals facilitate smooth information flow and minimise representational bottlenecks. Designed to handle 224x224 pixel inputs, MobileNetV2 offers a fully connected layer and SoftMax for classification, typically trained on ImageNet. It is well-suited for various vision tasks, including object detection and segmentation (Sandler et al., 2018).

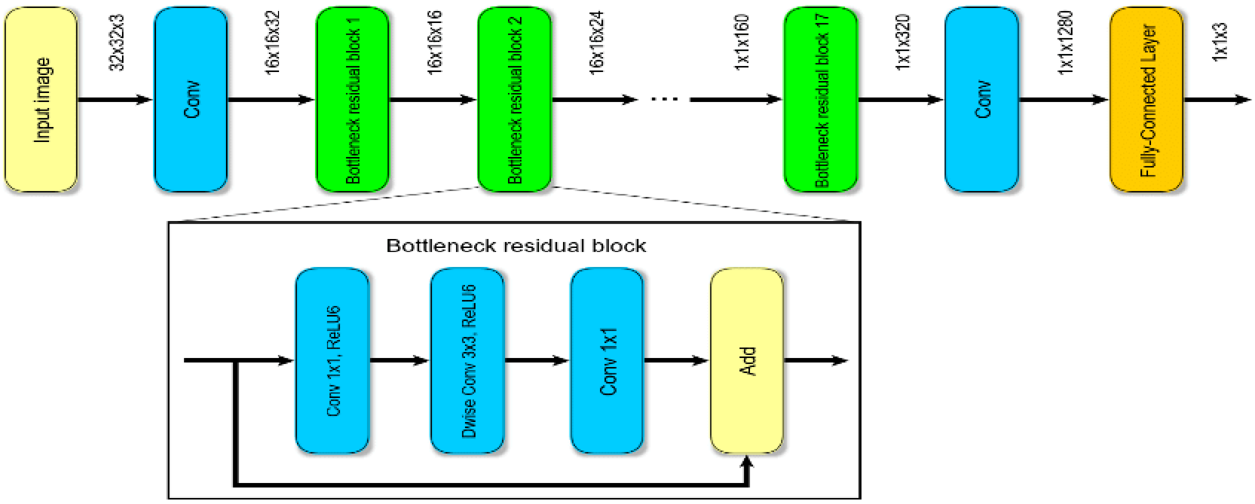


Fig 3.2 MobileNetV2 Architecture

Advantages of MobileNetV2

- High Efficiency: The model's low computational overhead (Sandler et al., 2018) makes it compatible with real-time mobile applications requiring a deployable, on-site rail track monitoring system.
- Versatility: MobileNetV2's proficiency in a range of vision tasks, including object detection and segmentation (Sandler et al., 2018), adapted to detecting leaf accumulation on rail tracks.
- Adaptability: The model's adjustable architecture allows straightforward modifications (Sandler et al., 2018), enabling the system to adapt to the visual characteristics of leaf accumulation on rail tracks.

Limitations of MobileNetV2

- Hyperparameter Sensitivity: To maximise MobileNetV2 performance, hyperparameters require careful tuning (Sandler et al., 2018).

3.3.3 ResNet50V2

This model evolves from the original ResNet (Residual Networks) architecture, which introduced the concept of residual learning to facilitate the training of deeper networks. ResNet50V2, with its 50 layers, incorporates improvements over the original ResNet50, such as applying batch normalisation before activations (pre-activation) and modifying the original residual block structure for enhanced performance. ResNet50V2 overcomes the vanishing/exploding gradient problem, a typical issue in intense networks, by leveraging skip connections or shortcuts to jump over some layers. These shortcuts allow the gradient to flow directly through the network without encountering too many layers simultaneously, preserving the gradient's strength and enabling more profound, more effective models. The model's potential to classify images into 1000 categories, like VGG19, originates from its training on the extensive ImageNet dataset. This capacity, in combination with its architectural advances, makes ResNet50V2 extraordinarily successful for an extensive range of image recognition and computer vision applications (He et al., 2016).

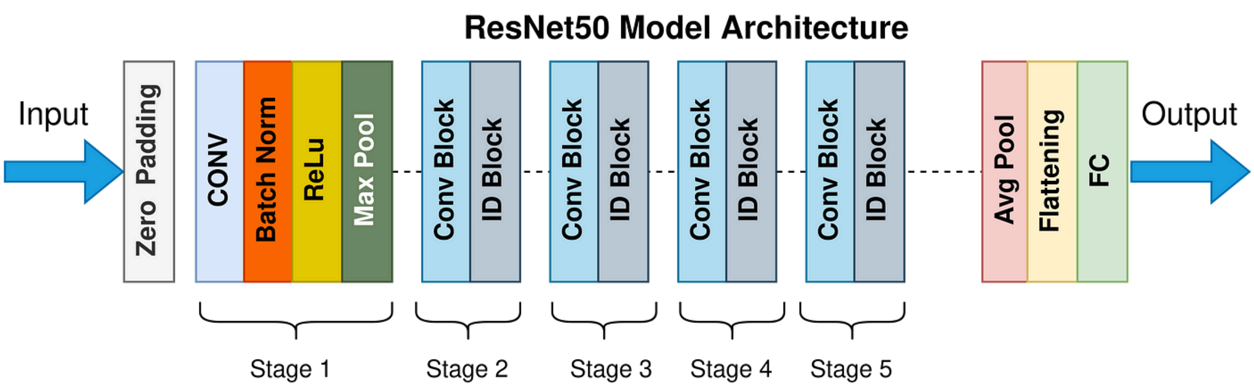


Fig 3.3 Resnet50V2 architecture

Advantages of ResNet50V2

- Deep Network Capabilities: This model provided more profound network training without performance loss (He et al., 2016). Its thoroughness enabled exact leaf accumulation detection in diverse situations
- Improved Performance: The pre-activation and modifications to the residual block structure in ResNet50V2 resulted in seamless gradient flow across layers. This improvement provides reliable problem

identification for deep network training regardless of complex visual input (He et al., 2016).

- Versatility: ResNet50V2's transfer learning abilities enable it to be fine-tuned for various vision tasks, such as image recognition and computer vision (He et al., 2016). Due to its customisable nature, this model was ideal for leaf detection on rail tracks.
- Efficiency: ResNet50V2's architectural optimisations made it a more resource-efficient leaf detection system (He et al., 2016), enhancing its implementation on current rail track monitoring systems without requiring significant hardware upgrades.

Limitations of ResNet50V2

- Complexity: ResNet50V2's deep and convoluted architecture, while beneficial for learning complex patterns, can also be complex to understand and execute, particularly for individuals new to deep learning (He et al., 2016).
- Computational Demands: ResNet50V2's computational demands remain high despite architectural optimisations, which poses a problem for real-time applications (He et al., 2016).
- Adaptability: To improve model accuracy, it is necessary to fine-tune it for new tasks and different data types (He et al., 2016).

3.3.4 Inceptionv3

The InceptionV3 model is a revised version of the original Inception architecture, which is notable for its innovative approach to convolutional neural network building. It was designed to enhance efficiency and performance in image recognition jobs without significantly increasing computing costs. InceptionV3's architecture is defined by its 'Inception modules', a revolutionary idea that allows the model to pick from various filter sizes within the same layer. This architecture enables the model to capture information at various scales, making it exceptionally adept at recognising patterns in photos regardless of their size or position. One of the significant breakthroughs in InceptionV3 is the inclusion of factorised convolutions. This technique breaks down more extensive convolutions into smaller, more manageable processes, which minimises the number of parameters and computations needed, boosting the model's efficiency without compromising its depth or capacity. The model also includes asymmetric convolutions, which entail breaking n x n convolutional filters into a mixture of 1 x n and an x 1 filters. These convolutions significantly decrease computational requirements while maintaining the network's effectiveness. Additionally, InceptionV3 incorporates advanced regularisation techniques, including label smoothing, which helps limit overfitting and boost the model's generalisation capabilities (Szegedy et al., 2016).

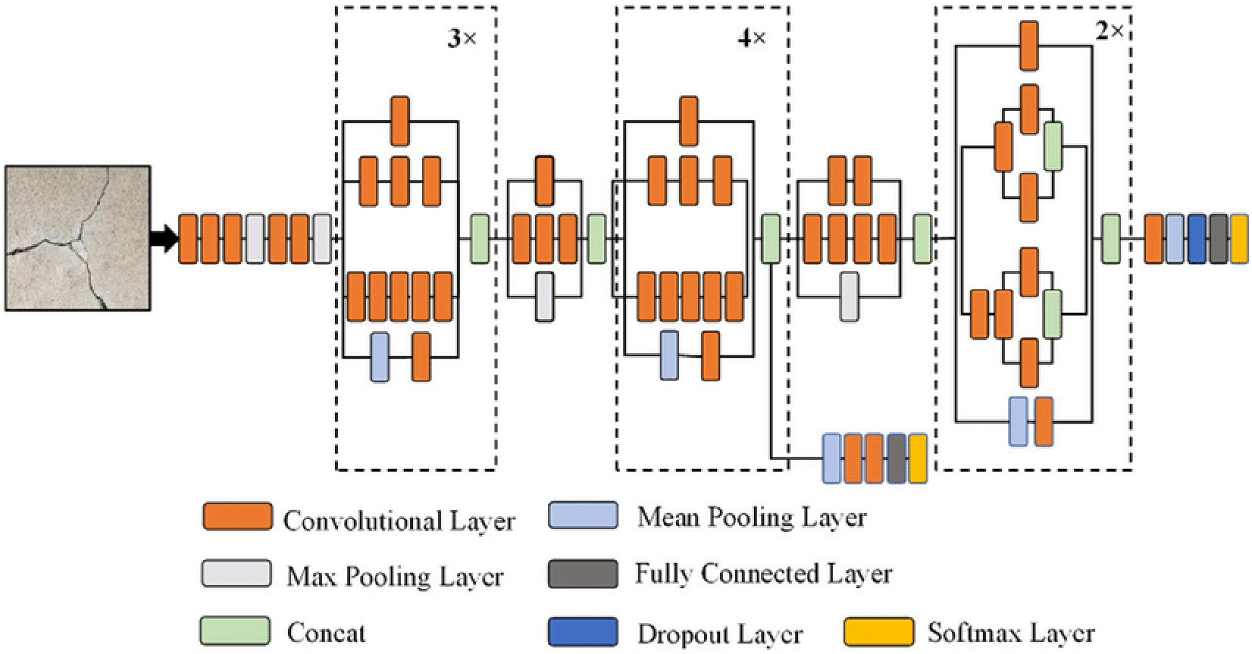


Fig 3.4 INCEPTIONV3 Architecture

Advantages of InceptionV3

- Advanced Architecture: InceptionV3's improved performance is attributed to factorised convolutions and expanded inception modules (Szegedy et al., 2016). Its performance led to its success in leaf recognition on tracks by correcting leaves' shifting sizes and locations, thereby ensuring improved accuracy in determining leaf accumulation sites.

- Efficient Computing: Due to its asymmetric convolutions (Szegedy et al., 2016), InceptionV3 can seamlessly integrate into railway track monitoring systems without enormous computational demands.
- Regularisation Techniques: InceptionV3 leverages label smoothing to limit overfitting and increase generalisation (Szegedy et al., 2016), strengthening the reliability of detecting leaf accumulation.

Limitations of InceptionV3

- Model Complexity: While InceptionV3's complex structure contributes to its performance, it can also be challenging to understand and modify (Szegedy et al., 2016).
- Resource Intensity: Despite its efficiency gains, InceptionV3 still demands high computational resources (Szegedy et al., 2016).
- Fine-tuning Required: InceptionV3 needs significant adjustments for new tasks or data types (Szegedy et al., 2016).

3.3.5 Xception Model

The Xception model is an advanced deep convolutional neural network that extends the Inception design. It is designed around depth-wise separable convolutions, a form of factorised convolutions that decreases the model's computational burden while enhancing efficiency. The Xception model consists of 36 convolutional layers, forming the foundation of its architecture, which are organised into 14 modules, all constructed using depth-wise separable convolutions. Unlike typical convolutions, depth-wise separable convolutions split the process into two layers: a depth-wise convolution and a pointwise convolution, enabling the model to capture more complicated patterns with fewer parameters. This model is particularly effective at processing high-resolution photos. It is trained to classify them into 1000 categories based on the ImageNet dataset, making it suitable for a wide range of image recognition tasks (Chollet, 2017).

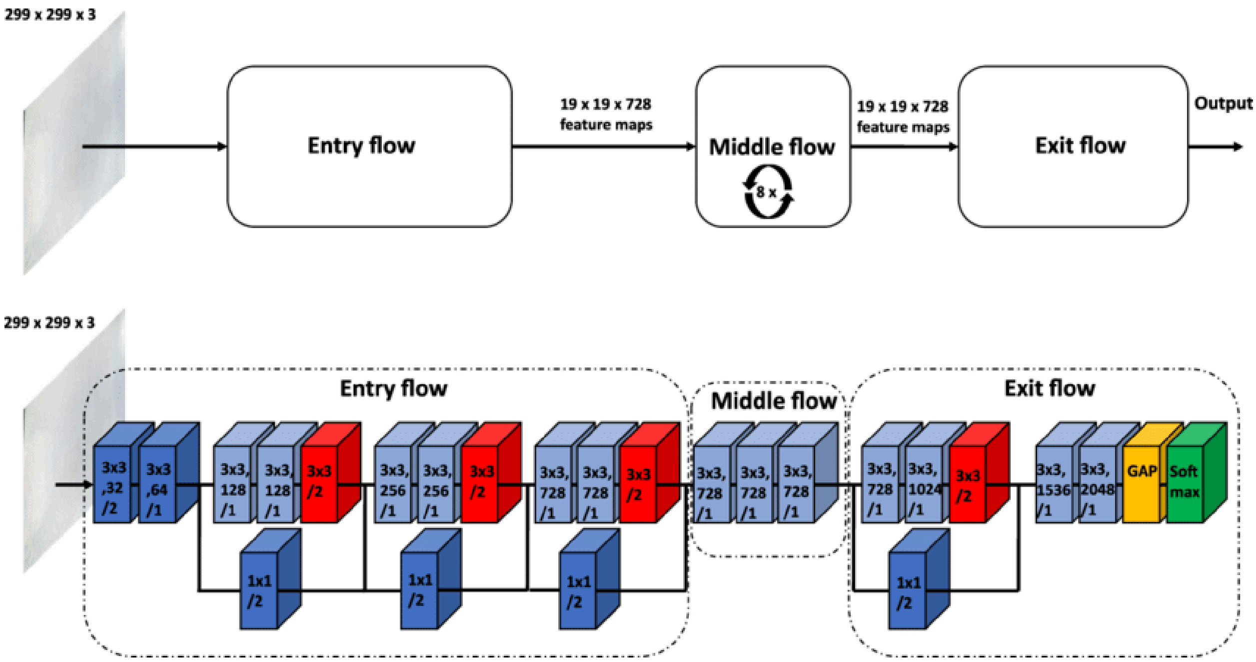


Fig. 3.5 Xception Model Architecture

Advantages of Xception

- High Performance: Xception's parameter efficiency exceeds that of other models in the benchmark (Chollet, 2017), allowing accurate leaf accumulation detection on rail tracks.
- Computational Efficiency: Xception leverages depthwise separable convolutions to cut computational costs while maintaining accuracy (Chollet, 2017), thus facilitating its integration into rail track monitoring systems without significant computational demands.
- Adaptability: The Xception model's flexible architecture facilitates transfer learning for diverse applications (Chollet, 2017), including leaf detection on railway tracks.
- Feature Extraction: The Xception model excels in complex vision tasks due to its thorough feature extraction (Chollet, 2017). This ability informed its choice for the study's goal of detecting leaf accumulation under diverse conditions.

Limitations of Xception

- Model Complexity: Although the Xception model's depth and architecture improve its efficiency, they make it challenging to understand and modify (Chollet, 2017).

- Training Requirements The model demands substantial computational resources for training (Chollet, 2017).
- Fine-tuning for Specific Tasks: The Xception model may require notable fine-tuning to attain optimal performance on tasks different from image classification (Chollet, 2017).
- Resource-Intensive: Despite its efficiency, the Xception model can be resource-intensive, mainly when working with massive images or datasets (Chollet, 2017).

3.3.6 Custom-Built TensorFlow Model

The model, a convolutional neural network (CNN), is designed with Keras' Sequential API for binary classification tasks with image data. Its architecture is straightforward, comprising layers specifically crafted to abstract and capture the features of an input image of size 256x256 with 3 colour channels (RGB).

Architecture Overview

Initially, the model was designed to receive an image input with (256, 256, 3) shape, implying a 256 by 256-pixel image with 3 color channels. The image data passed through the model and encountered five convolutional layers. These layers are the building blocks of the model, each with a specific number of filters—32, 64, 128, 256, and 256 and applied across the image with a kernel size of (3, 3), allowing the model to learn spatial hierarchies and features from the input data. The Rectified Linear Unit (ReLU) activation function was used within these layers to introduce non-linearity, enabling the model to learn more complex patterns. 'Same' padding was used to ensure the spatial dimensions of the output feature map were consistent with the input unless modified by the stride parameter.

Following each convolutional layer, batch normalisation was applied to stabilise learning and normalise the activations of the previous layer. After the batch normalisation and activation, a max pooling operation with a window of (2, 2) was carried out on each feature map to reduce the spatial dimension of the two-dimensional array by half, hence reducing the number of parameters of the network and, in turn, reducing the computational load on the network.

The output from this final pooling layer was then fed to the flatten layer, which converted the multidimensional feature maps into a one-dimensional feature vector. This extracted feature vector was then used in the model's final part, which consisted of two fully connected layers, also known as dense layers.

In the final part of the model, there were two dense layers. The first dense layer consisted of 256 units and served as the fully connected hidden layer in the neural network that processes the flatten feature vector. The second layer was the output layer, consisting of a single unit with a sigmoid activation function, making it suitable for binary classification tasks. The dropout technique was applied after each dense layer with a rate of 0.3 to mitigate the risk of overfitting when training with a small dataset. A random fraction of the input units was set to 0 during each update in the training phase, which helped create a more generalised model that can perform reliably on new data.

Advantages

- Efficiency in Parameter Usage: The batch normalisation and pooling layers help reduce the required parameters and computational costs, making the model trainable.
- Robustness to Overfitting: The dropout layers mitigate overfitting when training on small datasets.
- Flexibility and Scalability: The model is scalable, with the ability to add more layers or change the existing layers to suit the complexity of the data.

Limitations

- Limited Contextual Awareness: The consistent use of small (3, 3) kernels may limit the network's ability to capture higher-level, abstract features visible at larger receptive fields.
- Potential for Vanishing Features: The successive use of pooling layers may result in the loss of critical features, mainly if the nuanced and subtle features are spread across a large spatial region in the input image
- Fixed Input Size Requirement: The model requires a fixed input size (256x256), which makes it challenging to adapt to resize or pad the image to fit the specification.

3.4 METHODOLOGICAL APPROACH

This study evaluated the performance of different pre-trained convolutional neural networks (CNNs), including MobileNet, ResNet, Inception, VGG19, and Xception, for detecting leaf accumulation on railways. The objective was to assess the accuracy and efficiency of these models in categorising images from a specific dataset, employing TensorFlow and Keras for implementation and fine-tuning the models.

3.4.1 The Building Environment – Google Colab

This study employed Google Colab, an integrated Python environment in a cloud-based hands-on computing notebook. It allows users to work collaboratively, share, run, and work on any Python notebook with access to GPU computing power and Google's Cloud Platform.

Google Colab is simple to use, integrates with Google Drive, and is a popular tool among academics, data scientists, and educators for sharing and opening notebooks similar to Jupyter Notebooks. This research used Google Colab to train and analyse the performance of the Convolution Neural Network architectures.

One advantage of Google Colab is its GPU computing power. GPUs significantly accelerate the training of deep learning models and the processing of large datasets, particularly for tasks that involve image processing and complex computational operations.

Additionally, using Google Colab's GPU resources drastically reduced the time required for training the convolutional neural networks (CNNs).

The project required a lot of computational power to process the high-quality images; hence, a Subscription was made to Colab to access the premium GPUS, which made computation and training time faster. Also, it provided a cloud environment for working, thereby avoiding file loss cases.

3.4.2 Data Collection

Diverse data collection methods used both primary and secondary sources to represent the rails fully. Primary data collection involved directly capturing images of rail tracks to allow for precise images tailored to the study's objectives. The data included up-to-date images of the current state and conditions of rail tracks, essential for a high degree of model accuracy and relevance.

Secondary data collection combined abundant online resources with rail track data to create a diverse dataset. The online sources included Railcam.uk, Google Images, and Unsplash Images.

Railcam.uk, which specialises in railway images, provided an extensive repository of rail-related content that significantly enhanced the diversity and specificity of the dataset. Google Images provided railway track images from different contexts and locations that contributed to the generalisability of the dataset. Unsplash Images, specialising in royalty-free high-quality images, contributed aesthetically appealing and high-resolution images useful in detailed image analysis.

Merging images from various sources contributed to a dataset with multiple track conditions, environmental settings, and perspectives from different camera angles. This diversity was essential for training the models to recognise and analyse rail tracks under various conditions and from different perspectives, improving model adaptability and performance.

Ethical and legal implications around copyright and usage rights for images sourced from the Internet were considered during data collection. The terms of use of each data source and the creator's legal rights for each image were analysed to ensure compliance with the legal requirements.

3.4.3 Preprocessing And Augmentation

The images were loaded using the image_dataset_from_directory function from TensorFlow's Keras API, facilitating the creation of training and validation datasets. The datasets were configured with a batch size of 32 and an 80-20 split ratio, designating 80% for training and 20% for validation purposes. Further splitting within the training data set allocated 75% of the images for training and 25% for testing.

The preprocessing phase began with resizing the images to a uniform dimension of 512x512 pixels. This standardisation is crucial for CNNs requiring uniform input sizes (Goodfellow et al., 2016). The resizing was executed using the Python Imaging Library (PIL), which is renowned for opening, manipulating, and saving various image file formats. The preprocessing phase also involved converting images to the RGB colour space when necessary, ensuring that all images conformed to a standard colour representation.

Image augmentation is designed to increase the diversity training dataset by generating altered versions of existing images (Chollet, 2017). Techniques such as random rotations, shifts, shears, zooms, and flips were applied to the dataset to simulate different perspectives, angles, and environmental conditions. The goal of augmentation is to build models that are resilient to overfitting and capable of generalising to new images that has not been encountered during training (Perez & Wang, 2017). The preprocessing and augmentation of images were performed using a custom function process_and_augment_images.

Following resizing, the images underwent augmentation. The augmentation process was conducted using the ImageDataGenerator class from TensorFlow's Keras API, which applied various transformations to the images to broaden the dataset’s diversity. These transformations included:

- Rescaling the pixel values to the range.
- Applying random rotations within a range of 30 degrees.
- Implementing random horizontal and vertical shifts up to 20% of the image dimensions.
- Introducing shear transformations with an intensity factor of 0.15.
- Executing random zoom operations up to 20%.
- Flipping images horizontally
- Utilising the nearest neighbour method as the fill mode for areas outside the image boundaries.

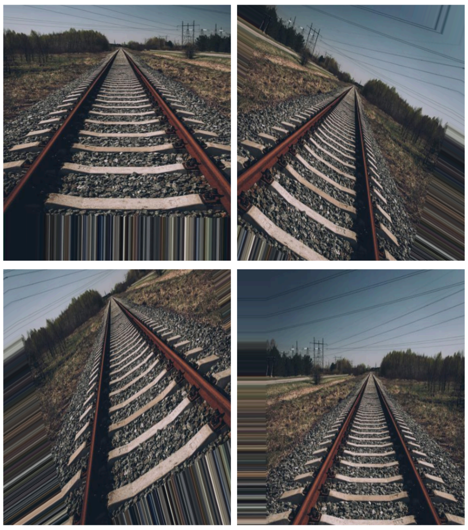


Fig 3.6 Image Augmentation Results

After resizing, each original image was saved, and augmented images were generated and saved with the prefix 'aug_' in JPEG format. The augmentation factor, set to 5, determined the number of augmented images produced per original image.

3.4.4 Renaming

The image files within each directory were renamed. This process, executed by the rename_images_in_directory function, replaced the original filenames with a sequential numbering system, retaining the file extensions. This helped to simplify dataset management and indexing, thus facilitating easier access and manipulation of the data (Witten et al., 2016).

3.4.5 Image Hashing and Removing Similar Images

Perceptual hashing is a technique that enhances the dataset's quality by deleting visually similar images. This process reduces redundancy, thereby enhancing the dataset's quality. By eliminating similar images, hashing aids in preventing overfitting and ensuring better performance across diverse visual scenarios. The delete_similar_images function was used to automate this process. It scanned each image, calculated a hash, and compared it to existing hashes. Images deemed similar based on a defined threshold were deleted, ensuring only unique images remained.

3.4.6 Data Splitting

The last step in the data preparation phase was splitting the dataset into training, validation, and testing sets using the split dataset function. This division ensured that the model was trained on a diverse subset of the data, validated on a separate subset to tune the hyperparameters, and evaluated on a distinct subset to assess its generalisation capability (James et al., 2013). The division followed a conventional distribution, allocating 60% of the data to the training set and 20% each to the validation and testing sets.

3.4.7 Model Implementation

Five pre-trained models were utilised in this study: MobileNet, ResNet, Inception, VGG19, and Xception. Each model was loaded with its pre-trained weights and configured to be non-trainable to preserve the learned features. A global average pooling layer followed by a dense layer with 1024 units and ReLU activation was appended to each pre-trained model. A dropout layer with a rate of 0.2 was included to reduce overfitting. The final layer was a dense layer with a single unit for binary classification. A custom-built TensorFlow model was also used to make the prediction.

3.4.8 Training and Evaluation

The models were compiled with the Adam optimiser, a learning rate 0.0001, and binary cross-entropy loss. Training was conducted over 32 epochs with early stopping based on validation loss to prevent overfitting. The performance of each model was evaluated based on training and validation accuracy and loss metrics.

3.4.9 Visualization

The accuracy and loss of the training process were visualised using Matplotlib to compare the models' performance visually. Plots were generated to display the training and validation accuracy and loss for each model across the epochs.

3.4.10 Prediction app

The best model was implemented in the sample prediction app, and the correct classification for 9/10 pictures was detected. The pictures that were sent were entirely new, which further cements the results.

```
model_monet.summary()

Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 mobilenetv2_1.00_224 (Func  (None, 8, 8, 1280)        2257984
 tional)

 global_average_pooling2d (  (None, 1280)              0
 GlobalAveragePooling2D)

 dense (Dense)               (None, 1024)              1311744

 dropout (Dropout)           (None, 1024)              0

 dense_1 (Dense)             (None, 1)                 1025

=================================================================
Total params: 3570753 (13.62 MB)
Trainable params: 1312769 (5.01 MB)
Non-trainable params: 2257984 (8.61 MB)
_________________________________________________________________
```

Fig 3.7: Model Summary

3.4.11 Code

The codes for the implementation are attached below:

Data Preparation:

https://colab.research.google.com/drive/1tN3KoXjtHH1QtRe-_XGqkNEHy3MSSqQK#scrollTo=0jK4sJf5yz5O

Model Building:

https://colab.research.google.com/drive/1VPhN8f39BSH8ydwOcbNJIWtKvbPTeK32#scrollTo=Z_TAQIlp8Hd7

CHAPTER 4: RESULTS

4.1 RESULTS

The study aimed to build a machine learning model based on computer vision to identify leaf falls on railway tracks, predict and prevent low adhesion conditions that can lead to significant delays and safety concerns within the rail network, and improve safety.

The study employed convolutional neural networks (CNNs) for image processing. Five pre-trained models were used: MobileNetV2, ResNet50V2, InceptionV3, VGG19, and Xception. These models were loaded with their pre-trained weights and configured to be non-trainable. A global average pooling layer, a dense layer with 1024 units, and a ReLU activation function were added after the pre-trained model. A dropout layer with a rate of 0.2 was also incorporated. Finally, a dense layer with a single unit was inserted for binary classification. The models were compiled using the Adam optimiser with a learning rate 0.0001 and binary cross-entropy loss. The training was performed over 32 epochs with early stopping based on validation loss to prevent overfitting.

The results are presented below.

4.2 COMPARATIVE MODEL PERFORMANCE

Six different deep-learning architectures were evaluated on a standard test dataset. The models evaluated included Xception, VGG19, ResNet50, Inception, MobileNet, and the custom TensorFlow Model.

Table 4.1 Comparative Model Performance

	Xception	**VGG19**	**ResNet50**	**Inception**	**MoBilenet**	**TF**
Test Loss	0.058961	0.012111	0.711462	0.568759	0.064665	1.036726
Test Accuracy	0.980263	0.996711	0.963816	0.904605	0.973684	0.993421

The model evaluation results are as follows: The VGG19 model demonstrated superior performance, with the lowest test recorded loss of 0.010535 and the highest test accuracy of 99.67%. This indicates a better fit for the data and a high accuracy in classifying the presence or absence of leaf accumulations on the tracks. The Custom TensorFlow Model also showed promising results, with a high test accuracy of 99.01% but with the highest recorded loss among the models at 1.040758.

Furthermore, training and validation analyses revealed a consistent trend. While some models exhibited signs of overfitting, the VGG19 consistently achieved high accuracy in both training and validation datasets, demonstrating good generalisation capabilities.

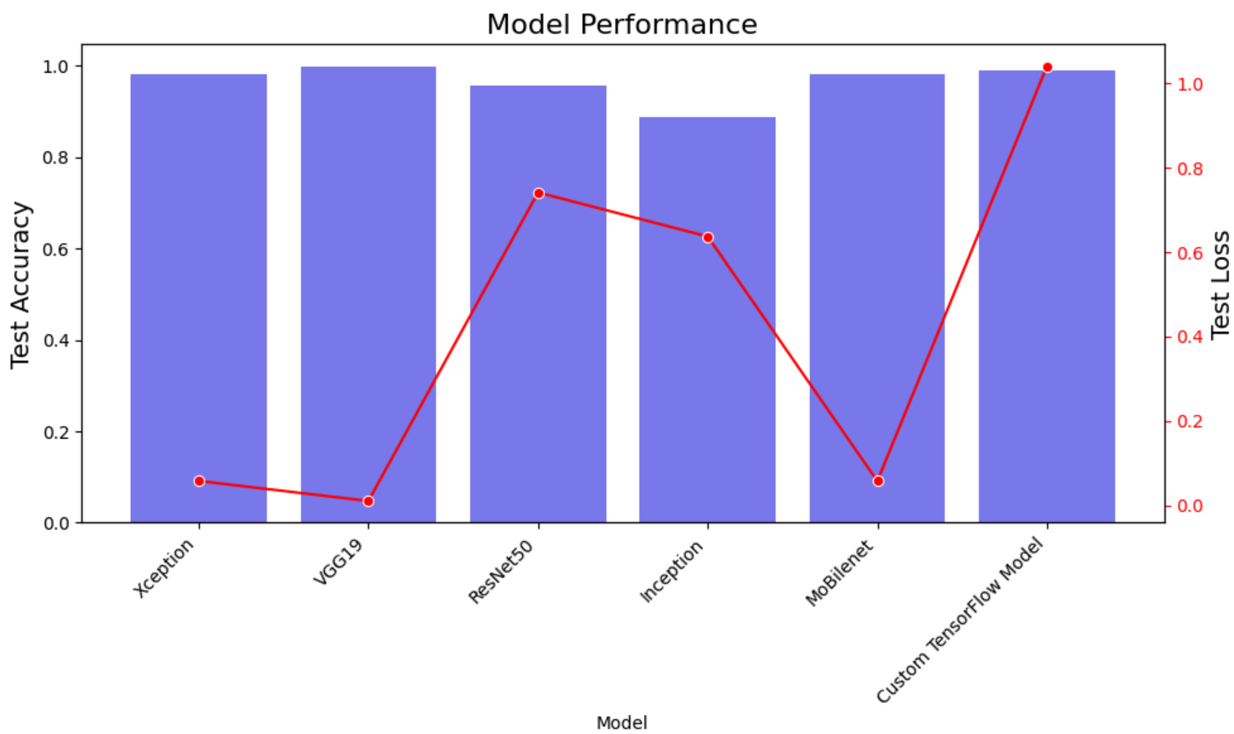


Fig 4.1 Model Performance Plot

4.3 TRAINING AND VALIDATION ANALYSIS

The training and validation accuracy and loss for each model were plotted against the epochs to visualise the models' learning behaviour and generalisation ability.

Custom TensorFlow Model: The Custom TensorFlow Model, despite displaying a high training accuracy that suggests a potential overfit to the training dataset, also shows promise. Despite some fluctuations, the validation accuracy remained high, and the validation loss shows a general downward trend, indicating the model's potential effectiveness in generalising to new data.

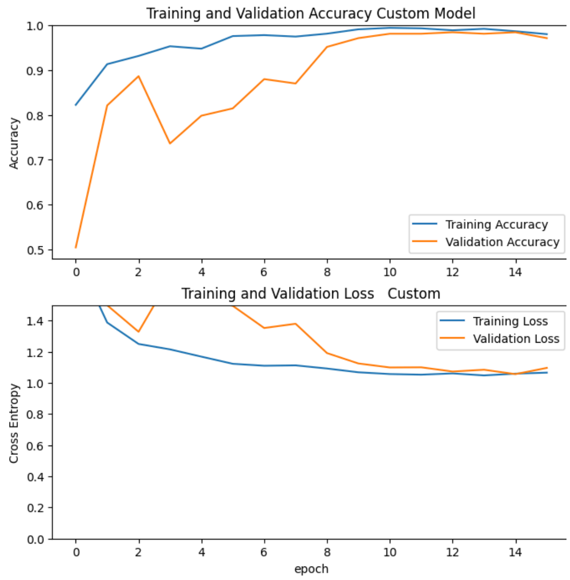
Xception: The Xception model showed optimal learning and generalisation with minimal overfitting. Its training and validation accuracies were closely aligned and plateaued around 95%.

Inception: From the beginning of training, InceptionV3 revealed a gap between training and validation accuracy, suggesting overfitting. The validation loss curve began to increase after intersecting with the training loss around the tenth epoch, further suggesting a diminishing generalisation as training progressed.

ResNet: ResNet50V2 displayed high training accuracy contrasted with volatile validation accuracy. The latter dipped significantly around the sixth epoch, suggesting potential overfitting. The validation loss was erratic, indicating challenges in model training dynamics.

MobileNet: MobileNetV2's accuracy displayed consistent learning and good generalisation, with closely matched training and validation accuracy curves. The loss metrics decreased uniformly, reinforcing the model's stable learning capability.

VGG19 VGG19 outperformed other models in both accuracy and loss metrics. The graph for the VGG19 model showed a steady increase in training accuracy and a corresponding decrease in training loss over the epochs. The training and validation accuracies converged closely, with the model achieving the highest accuracy of 99.67% on the test dataset. The loss metrics steadily decreased, indicating the model's good fit and generalisation capabilities.



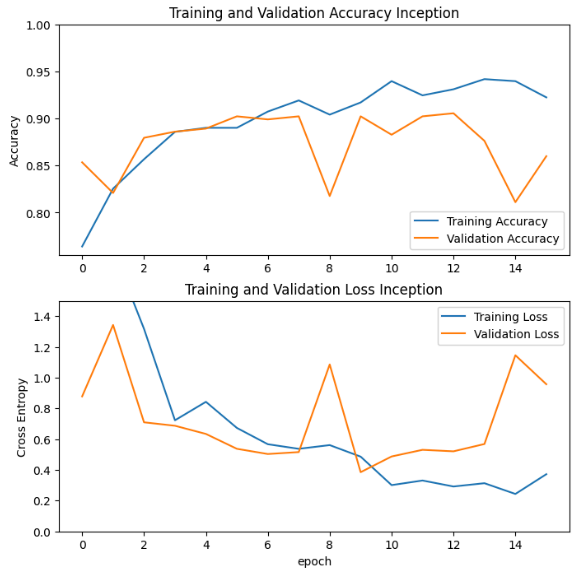


Fig 4.2 Training and Validation Accuracy and Loss Plot for Custom-Built Model

Fig 4.4 Training and Validation Accuracy and Loss Plot for Inception Model

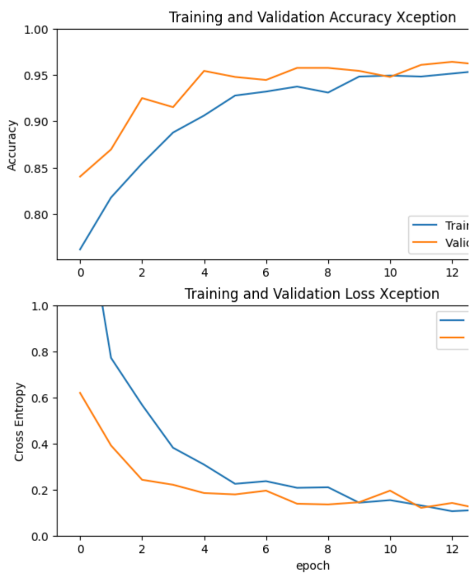


Fig 4.3 Training and Validation Accuracy and Loss Plot for Xception Model

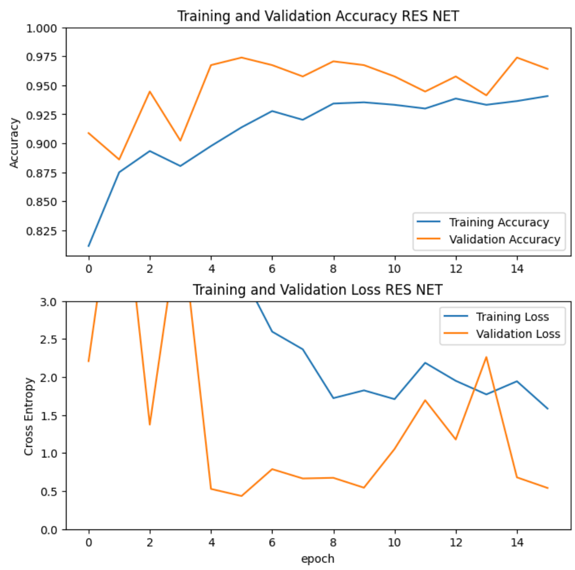


Fig 4.5 Training and Validation Accuracy and Loss Plot for RESNET (Residual Networks) Model

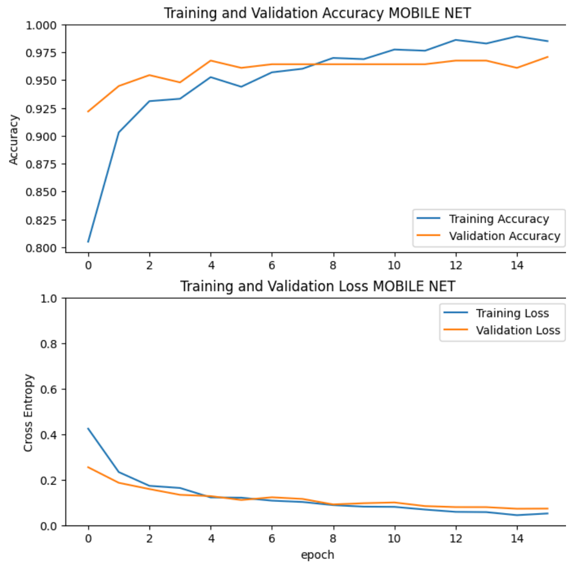


Fig 4.6 Training and Validation Accuracy and Loss Plot for MobileNet Model

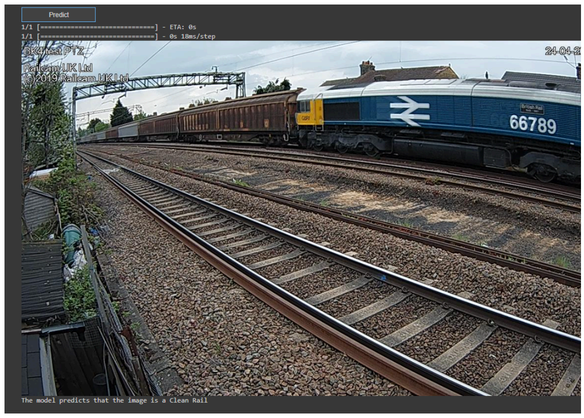


Fig 4.8 Sample prediction

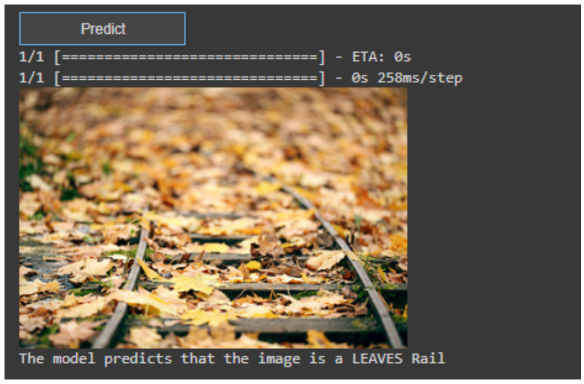


Fig 4.9 Sample Prediction for Leaves

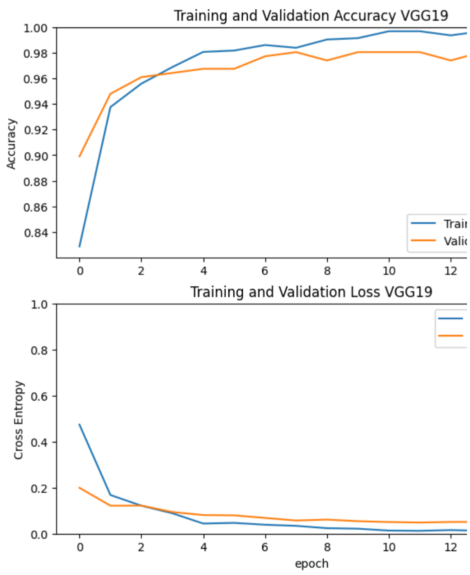


Fig 4.7 Training and Validation Accuracy and Loss Plot for VGG19 Model

The model was tested with realtime images of rail tracks and it predicted the state of the rail track, it was inaccurate in some cases and was retrained to be more accurate but the model cannot be 100% accurate. There were cases of inaccurate predictions also but with research, it can get better.

CHAPTER 5: DISCUSSION

5.1 DISCUSSION

The research findings reinforce the potential of advanced machine learning models in railway safety through computer vision for detecting leaf accumulation on rail tracks. VGG19 emerged as the most effective among the models tested, outperforming the custom-built model. Notably, VGG19 demonstrated remarkable consistency in its predictions and did not exhibit signs of overfitting, maintaining an impressive 97% accuracy rate. This performance suggests the capability of computer vision technologies to significantly enhance predictive maintenance strategies by identifying critical maintenance needs in real time.

The VGG19 model's real-time detection capability significantly surpasses existing systems, which rely on lagging indicators or indirect data to predict track conditions. By directly analysing images of the tracks, the VGG19 model enables immediate action, thereby reducing downtime and substantial costs associated with delayed maintenance responses. Unlike traditional systems that might only be activated once the problematic conditions had already impacted rail operations, the application of VGG19 could preemptively identify and address issues before they escalated into more severe problems.

Integrating such a model into applications with access to real-time image data of rail tracks could transform how rail operators manage track maintenance. These applications would not only predict and identify leaf accumulation effectively but could also alert maintenance crews immediately, enhancing the responsiveness of the maintenance processes. However, these integrations' full potential and performance remained areas ripe for further research. Future research is suggested to validate the model's effectiveness in diverse operational contexts and explore the scalability of deploying such technology across different regions and rail network conditions.

The integration with existing leaf-busting and Water-Trak systems could further enhance the effectiveness of the AI-driven leaf detection system. The leaf-busting system, which cleans the rails using high-pressure water jets and applies sand and steel grain gel for traction, could be deployed more strategically by targeting areas identified by the AI system as having significant leaf accumulation. This targeted approach could optimize resource use and ensure the tracks are cleaned precisely where needed without running unnecessarily.

Similarly, the Water-Trak system, which applies water to the railhead to simulate rainy conditions that help wash away leaves, could be activated based on real-time data provided by the AI system. This would ensure that the Water-Trak system is only active when necessary, reducing water usage and wear and tear on the equipment, extending its operational life, and reducing maintenance costs.

Despite the promising outcomes, this study faced several significant limitations that must be acknowledged. The data collection process proved challenging, primarily due to stringent regulations surrounding acquiring aerial images of rail tracks and the limited viability of the images obtained through internet scraping. Although enhanced through data augmentation, the resultant dataset consisted of only 1,500 images, which might not fully represent the variability in real-world conditions. Access to a more extensive and diverse dataset could improve the model's accuracy and generalizability.

Furthermore, the computational intensity of the model training, reliant on high-performance GPUs available through platforms like Google Colab, posed challenges for deployment in environments with limited technological infrastructure. This limitation underscored the need for research to optimise model efficiency without compromising performance, making implementing these solutions on less capable devices feasible.

Additional limitations included the model's dependence on high-quality, high-resolution images and its untested performance under extreme weather conditions, which must be adequately represented in the test datasets. It's important to note that the model's performance might be affected by extreme weather conditions, such as heavy rain or snow, might affect the model's performance, which could lead to suboptimal leaf detection. These factors could affect the model's effectiveness in actual operational settings, highlighting the necessity for ongoing testing and refinement.

Given the outlined limitations and the demonstrated potential of the machine learning approach, several recommendations are proposed:

5.2 RECOMMENDATIONS

- Future studies should explore the practical integration of these models into existing railway safety systems to facilitate continuous monitoring and real-time predictive capabilities. This integration could significantly enhance the responsiveness and effectiveness of maintenance operations, improving rail service safety and reliability.
- Expanding the training datasets to encompass a broader range of environmental conditions and geographical variations is crucial. This expansion would enhance the model's adaptability and robustness and ensure that the system is effective across different regions and rail network conditions.
- Railway operators are encouraged to adopt these AI-driven systems within their safety infrastructures, particularly in pilot areas known for significant leaf-related disruptions. This would allow for evaluating the systems' effectiveness in operational environments and provide valuable insights into their practical deployment and scalability.
- Collaboration with technology developers to refine and customise these models is recommended to ensure that the solutions are both scalable and economically viable. Tailoring these models to meet the specific needs of diverse rail networks would make them more applicable and beneficial across the industry.

CHAPTER SIX: CONCLUSION

6.1 CONCLUSION

This research developed and implemented a machine learning model using computer vision that has successfully identified the degree of leaf accumulation on railway tracks with high accuracy. The study suggests an innovative and ideal approach for the future in dealing with the problem of the negative impact of leaf fall on railways. Using novel state-of-the-art machine learning models and computer vision techniques, the method developed discovers the presence of leaf fall for accurate and real-time decision-making to execute the appropriate chores. Among the models employed, the performance of the VGG19 model proved best in predicting leaf fall with a test accuracy of 99.67%. For the first time, the method not only quantifies the severity of the problem but also adapts a technological resolution that goes way beyond the existing legacy methods and their limitations in precision and applicability, marking a paradigm shift way forward in the maintenance and safety of railways.

Integrating this system into the rail network is a significant step towards mitigating the emerging risks associated with low adhesion conditions. This advancement enhances rail service customers' safety, reliability, and satisfaction. Our research not only identifies the limitations of current practices but also contributes to the ongoing discourse on improving the efficiency of the rail network and customer satisfaction. This aligns with the Department for Transport's 'Strategic Vision for Rail' and the 'Rail Capability Plan'. Moreover, it opens up new avenues for future research on alternative methods and technologies to address the issue of leaf-affected low adhesion on rail tracks, thereby enhancing the safety and reliability of future rail services.

This work has implications that extend beyond immediate operational efficiencies. It lays the foundation for future potential research into maintenance and safety for rail services, and it is crucial that we adopt a proactive approach. The use of machine learning and computer vision presents new opportunities for a progressive maintenance strategy. This could potentially revolutionise the way we monitor and maintain our rail networks, inspiring a new era of rail network safety and efficiency.

However, the work noted some limitations. Specifically, data collection is constrained by what images are available for a given condition and lighting, and it needs to be clarified to what extent these factors could impact modelling and generalisability. At the time, the computing power required to extract sparrows from the background of high-resolution images and process the resulting analysis also tied optimising the network to only some operational contexts.

Moving forward, researchers should increase the sample size for environmental conditions and track types to increase the robustness and applicability of the model in a broad range of rail networks. Furthermore, researchers could explore the computational efficiency of the model. If the computational burden is reduced, the model could subsequently be used for real-time applications on a larger scale. Lastly, investigating the implementation of this technology into the existing rail maintenance and monitoring systems could allow for a more integrated and holistic solution to the issue of leaves on the track.

REFERENCES

Chen, H., & Fukagai, S. (2018). Influence of Leaves on the Adhesion between Wheel and Rail, 59(1).

Chollet, F. (2017). Deep Learning with Python. Manning Publications.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.

Dramatic freight train derailment causes delays along Toronto-Montreal line | Globalnews.ca. https://globalnews.ca/news/8161693/only-minor-injuries-reported-after-2-cargo-trains-crash-in-prescott-ont-opp/

Folorunso, M. O., Watson, M., Martin, A., Whittle, J. W., Sutherland, G., & Lewis, R. (2023). A Machine Learning Approach for Real-Time Wheel-Rail Interface Friction Estimation. Journal of Tribology, 145(9). https://doi.org/10.1115/1.4062373

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 770-778).

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). An Introduction to Statistical Learning: with Applications in R. Springer.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing Systems (pp. 1097-1105).

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

Louise, J., & Kirby, A. (2019). High-resolution leaf fall monitoring using hemispherical NIR imagery.

Network Rail. (2023). Leaves on the line. Retrieved from https://www.networkrail.co.uk/running-the-railway/looking-after-the-railway/delays-explained/leaves/

O' Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Velasco Hernandez, G., Krpalkova, L., Riordan, D., & Walsh, J. (2019). Deep Learning vs. Traditional Computer Vision.

Of Rail, O. (2024). Passenger rail usage - October to December 2023.

Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering, 22(10), 1345-1359.

Perez, L., & Wang, J. (2017). The Effectiveness of Data Augmentation in Image Classification using Deep Learning. Convolutional Neural Networks Vis. Recognit.

Pyke, D. (2022). Don't fall for leaf delays. Retrieved from https://www.linkedin.com/pulse/dont-fall-leaf-delays-daniel-pyke/

Sanjar, K., Bekhzod, O., Kim, J., Kim, J., Anand, P., & Kim, J. (2020). Improved U-Net: Fully Convolutional Network Model for Skin-Lesion Segmentation. Applied Sciences, 10(10), 3658.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. In 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1-9).

Watson, M., White, B., Lanigan, J., Slatter, T., & Lewis, R. (2020). The composition and friction-reducing properties of leaf layers. Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 476(2239). https://doi.org/10.1098/rspa.2020.0057

Wikipedia contributors. (2024). Rail transport in Great Britain. Retrieved from https://en.wikipedia.org/wiki/Rail_transport_in_Great_Britain.

Witten, I. H., Frank, E., Hall, M. A., & Pal, C. J. (2016). Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann.

Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? In Advances in Neural Information Processing Systems (pp. 3320-3328).

APPENDIX

Rail transport in Great Britain: **Oldest railway system in the world**
Infrastructure Company: Network **Rail (until 2024)**
Major operators: **National Rail franchisees, independent operators, State-owned operators**
Ridership: **1.738 billion (2019/20)**
Passenger km: **41.5 billion mi (66.8 billion km) (2019/20)**
Total track length: **9,824 mi (15,811 km)**
Electrified track length: **3,339 mi (5,374 km)**
Number of stations: **2576**
(Source: Wikipedia contributors (2024))

CODES

The codes for the implementation are attached below:

Data Preparation:
https://colab.research.google.com/drive/1tN3KoXjtHH1QtRe-_XGqkNEHy3MSSqQK#scrollTo=0jK4sJf5yz5O

Model Building:
https://colab.research.google.com/drive/1VPhN8f39BSH8ydwOcbNJIWtKvbPTeK32#scrollTo=Z_TAQIlp8Hd7