applied sciences

MDPI

Article

Sustainable Pavement Management: Harnessing Advanced Machine Learning for Enhanced Road Maintenance

Kshitij Ijari [1] and Carlos D. Paternina-Arboleda [2,*]

[1] Computational Science Research Center, College of Sciences, San Diego State University, San Diego, CA 92182, USA; kijari6935@sdsu.edu

[2] Department of Management Information Systems, Fowler College of Business, San Diego State University, San Diego, CA 92182, USA

* Correspondence: cpaternina@sdsu.edu

Abstract: In this study, we introduce an advanced system for sustainable pavement management that leverages cutting-edge machine learning and computer vision techniques to detect and classify pavement damage. By utilizing models such as EfficientNetB3, ResNet18, and ResNet50, we develop robust classifiers capable of accurately identifying various types of pavement distress. To further enhance our dataset, we employ a Swin Transformer-based Generative Adversarial Network (GAN) to synthetically generate images of pavement cracks, thereby augmenting the training data. Our approach aims to improve the efficiency and accuracy of pavement damage assessment, contributing to more effective and sustainable road maintenance practices. This research aligns with the sustainable development goals by fostering innovative methods that extend the lifespan of infrastructure, reducing the need for resource-intensive repairs, and promoting the longevity and reliability of road networks. The outcomes of this study are discussed in terms of their potential impact on infrastructure safety and sustainability, with suggestions for future research directions. This study demonstrates how integrating advanced machine learning techniques into pavement management systems can enhance decision-making, optimize resource allocation, and improve the sustainability of infrastructure maintenance practices. By leveraging big data and sophisticated algorithms, stakeholders can proactively address pavement deterioration, extend asset lifespan, and optimize maintenance efforts based on real-time data-driven insights.

Keywords: pavement damage classification; machine learning; computer vision; deep learning; genrative adversarial networks (GANs)

Citation: Ijari, K.; Paternina-Arboleda, C.D. Sustainable Pavement Management: Harnessing Advanced Machine Learning for Enhanced Road Maintenance. *Appl. Sci.* **2024**, *14*, 6640. https://doi.org/10.3390/app14156640

Academic Editors: Derun Zhang, Chao Xing, Ju Huyan, Quan Lv and Song Liu

Received: 28 June 2024
Revised: 18 July 2024
Accepted: 27 July 2024
Published: 30 July 2024

1. Introduction

The classification and detection of pavement damage are critical for the maintenance and improvement of transportation infrastructure [1], which directly impacts economic productivity, safety, and environmental sustainability [2]. Pavement deterioration, if left unaddressed, can lead to costly repairs, increased vehicle operating costs, and higher accident rates [3]. Traditional methods of assessing pavement conditions rely heavily on manual inspections, which are time-consuming, labor-intensive, and prone to human error [4]. The relevance of this problem is underscored by the need for more efficient, accurate, and scalable solutions to ensure that road networks are kept in optimal condition. By integrating advanced machine learning and deep learning techniques, this research seeks to develop a sophisticated system that can automatically and accurately identify various types of pavement damage, ultimately contributing to more sustainable and effective road maintenance practices.

The core of this research is the classification and detection of pavement damage through the development of a sophisticated system that harnesses the latest advancements in machine learning and deep learning, specifically through computer vision techniques. The ability to accurately classify and detect various types of pavement damage, such as cracks and potholes, is crucial for maintaining and improving transportation infrastructure [1].

In this work, we evaluate the performance of several state-of-the-art deep learning models, including EfficientNetB3, ResNet18, and ResNet50. These models have been meticulously tested for their effectiveness in identifying and classifying pavement damage. To further enhance our system, we employ a Swin Transformer-based Generative Adversarial Network (GAN) to create synthetic images of pavement cracks. This innovative approach to data augmentation significantly improves the robustness and accuracy of our damage detection system.

The central issue addressed in this research is the precise classification of pavement damage using advanced deep learning models. This paper introduces an automated system powered by deep learning algorithms that can accurately identify various types of pavement damage from images. This innovative approach significantly streamlines the assessment process, enhancing both its speed and consistency. The improvements in the reliability and accuracy of pavement condition evaluations enable proactive maintenance strategies, thereby promoting infrastructure sustainability.

Our research addresses a vital aspect of transportation infrastructure maintenance: the classification of pavement damage. By leveraging machine learning and deep learning techniques, particularly computer vision, we developed a system that accurately detects and classifies various types of pavement damage, such as cracks and potholes. Our study assesses the performance of models including EfficientNetB3, ResNet18, ResNet50, and SwinGAN. Additionally, we employed a Generative Adversarial Network (GAN) to generate synthetic images of cracks, enhancing our dataset and improving the robustness of our damage detection system. This work aligns with the goals of sustainability by promoting innovative methods to ensure the longevity and reliability of transportation infrastructure, ultimately supporting sustainable development and contributing to the 2030 Agenda for Sustainable Development.

Our research contributes to the sustainable development goals by introducing innovative methods that enhance the longevity and reliability of road infrastructure. This aligns with the targets set by the 2030 Agenda for Sustainable Development, emphasizing the importance of sustainable and resilient infrastructure.

The structure of this paper is as follows:

- Literature Review: This section provides a comprehensive review of existing research on the use of computer vision and deep learning for pavement distress classification. It expands on key concepts, approaches, and significant developments that form the foundation for this project.
- Data Description and Exploratory Analysis: This section elaborates on the research data, detailing its type, size, and the variables under observation. It examines the data structure, identifies patterns, and outlines any anomalies that could influence the modeling approach.
- Methodology: We discuss the methods used in this project, from data selection and preprocessing to the experimental setup. This includes the use of convolutional neural networks (CNNs) and GANs, describing how these models are configured and utilized.
- Results and Discussion: This section presents the results of our experiments, evaluating how well the models predict pavement distress. We discuss the implications of our findings for distress identification and management, highlighting the strengths and weaknesses of our approach and suggesting directions for future research.
- Conclusions: This summarizes the main findings of our work, emphasizing our contribution to advancing pavement management technologies. We also discuss the potential impacts of these advancements on infrastructure safety and lifespan, along with possible directions for further research and technological development.

Overall, this paper aims to present our project systematically, ensuring that readers can easily understand the theoretical foundations, methodological details, experimental results, and broader implications of our work.

2. Literature Review

Ensuring the health and integrity of pavements is essential for maintaining efficient and safe infrastructure systems worldwide. Pavement health assessment, particularly through accurate crack detection, plays a pivotal role in infrastructure maintenance strategies [5,6]. Cracks in pavements, whether on roads, airport runways, or other transportation networks, can lead to accelerated deterioration and safety hazards if left untreated [7]. Effective assessment and early detection of cracks not only extend the lifespan of pavements but also minimize the need for costly repairs and ensure uninterrupted operation of critical transportation routes [8]. In recent years, advancements in machine learning, particularly convolutional neural networks (CNNs), have revolutionized crack detection by automating and enhancing the accuracy of detection processes [9]. This literature review explores the evolving methodologies and technologies in pavement health assessment, focusing on the role of machine learning in optimizing infrastructure management practices.

The literature on pavement damage classification and detection reveals significant advancements in the application of machine learning and deep learning techniques. Recent studies emphasize the role of computer vision in automating the assessment process, reducing reliance on manual inspections, and improving accuracy and efficiency [10]. Key developments include the use of convolutional neural networks (CNNs) and Generative Adversarial Networks (GANs) for image analysis and data augmentation, respectively [11,12]. Researchers have demonstrated the effectiveness of various models, such as EfficientNet and ResNet, in identifying different types of pavement distress [13,14]. Furthermore, innovations like the Swin Transformer and other vision transformers have introduced new paradigms in handling complex image data [15]. This literature review explores these advancements, providing a comprehensive understanding of the current state of the art and identifying the technological foundations that underpin our research.

Integrating insights from recent research into the application of advanced machine learning and deep learning techniques in pavement distress classification reveals a strong trend toward sustainability in infrastructure maintenance. These innovations not only enhance road functionality and lifespan but also contribute significantly to environmental and economic sustainability. Advancements such as the Pavement Image Classification Transformer (PicT) use vision transformers to effectively model the relationships between image patches through self-attention mechanisms. PicT's dynamic pseudo-labeling process reduces manual labeling efforts and costs, enhancing the adaptability of models to recognize new or rare types of distress, marking a significant step towards sustainable machine learning applications in road maintenance. Similarly, the adoption of Deep Metric Learning for few-shot learning underscores efficiency in adapting to various pavement distress types with minimal data requirements. This method conserves resources and reduces the carbon footprint associated with extensive data collection, representing a substantial advance toward sustainable computing in infrastructure management.

In [16], the authors illustrate how sophisticated algorithms can effectively manage infrastructure needs and environmental impacts through predictive modeling techniques used in maritime port emissions and pavement performance prediction models. For instance, predictive modeling of sulfur dioxide emissions in maritime facilities using machine learning enables proactive environmental strategies and regulations, helping to minimize the ecological and health impacts of port operations.

In the assessment of road pavement health and safety, crack detection is crucial for proactive maintenance strategies. Recent advancements in convolutional neural networks (CNNs) have automated the identification of cracks, addressing challenges such as complex topology segmentation, noise interference, and sensitivity to thin cracks [17]. Here, the authors propose a novel deep residual network, Parallel ResNet, designed to enhance pavement crack detection and measurement by effectively removing noise interference and accurately identifying complex cracks. Experimental validation on CrackTree200 and CFD datasets demonstrates Parallel ResNet's superior performance in precision, recall, and F1 scores, making it a robust tool for precise pavement crack analysis and maintenance planning.

The application of convolutional neural networks (CNNs), such as Parallel ResNet, for crack detection, exemplifies how precise and efficient technological applications can preemptively tackle pavement distress, preventing minor damages from becoming major repairs and conserving materials and labor [18,19].

As with the use of cost-efficient methods for crack detection, [20–22] give a good idea of why using image recognition provides an advantage. Given limited funds and an aging highway infrastructure, agency decision-makers must make the most cost-effective choices in highway pavement maintenance to improve service quality and enhance driving safety.

In addition, it is essential to regularly evaluate and execute maintenance and rehabilitation plans to maintain the network's acceptable level of service. A key element of the M and R plan involves using performance prediction models, particularly for addressing rutting distress, which is a major concern in asphalt pavement [23,24].

We have identified several gaps in the literature regarding crack detection in pavement management, revealing several critical areas where current methodologies and technologies fall short. One significant gap lies in the handling of complex crack topologies. While convolutional neural networks (CNNs) have shown promise in automating crack detection, they often struggle with accurately segmenting and classifying cracks that exhibit irregular shapes or patterns [25]. Existing models may not effectively distinguish between different types of cracks, such as longitudinal, transverse, or alligator cracks, which vary in their severity and impact on pavement integrity [26].

Another notable gap is related to noise interference in crack detection systems. Pavements are exposed to various environmental factors such as shadows, strong light reflections, and road markings, which can obscure crack patterns and introduce noise into image data. Moreover, the sensitivity of CNN models to thin cracks remains a challenge. Thin cracks are often subtle and difficult to detect in images, especially when they are overshadowed by surrounding pavement textures or imperfections.

Furthermore, there is a lack of standardized datasets and evaluation metrics specifically designed for pavement crack detection [17]. Many existing studies use different datasets with varying quality and annotation standards, making it challenging to compare results across different methods. This variability hinders the development of robust and universally applicable crack detection models.

Addressing these gaps requires innovative approaches that enhance the robustness and adaptability of crack detection systems. Future research should focus on developing CNN architectures that can handle complex crack geometries and improve the models' ability to filter out noise from pavement images effectively. Additionally, there is a need for comprehensive datasets that capture the diversity of pavement conditions and crack types, enabling fair comparisons and benchmarking of new methodologies. By addressing these gaps, researchers can advance the field of pavement management and contribute to more reliable and efficient infrastructure maintenance practices.

Together, these technological innovations foster a proactive and resource-efficient approach to infrastructure maintenance. By enhancing the longevity and functionality of infrastructure while aligning with broader sustainability objectives, these strategies reduce environmental impacts and bolster economic and societal well-being.

3. Data and Methods

The effectiveness of pavement health assessment and crack detection heavily relies on the quality and diversity of the datasets used, as well as the rigor of the methodologies applied. This section provides a detailed overview of the data sources utilized, including their characteristics and preprocessing steps. Additionally, it outlines the machine learning models and algorithms employed for crack detection and classification, emphasizing innovations aimed at overcoming current limitations in handling complex crack topologies, noise interference, and sensitivity to thin cracks. By integrating robust data sources with state-of-the-art methodologies, this research aims to enhance the accuracy and efficiency of pavement management practices, ultimately contributing to safer and more sustainable infrastructure systems. Figure 1 shows the types of pavement distress under study.

Figure 1. Types of pavement distress.

Most pavement distress datasets include various types of annotations that aid in analysis. These annotations are usually in three forms: labeled, masked, and bounding box. Labeled datasets categorize each image, which helps in classification efforts. Bounding box datasets include rectangles that show the position of objects within the images, which is crucial for object detection tasks. Masked datasets provide very detailed location data down to the pixel level [22], outlining precise boundaries and areas for segmentation purposes.

3.1. Exploratory Data Analysis (EDA)

In this project, we utilized the CQU-BPDD dataset [27,28], which comprises a total of 60,059 images segmented into training and testing sets based on specific criteria to ensure comprehensive model evaluation. The training set includes 5137 images of diseased pavements, covering a diverse range of distress types, and 5000 images of normal pavements, facilitating balanced training of our deep learning models. The remaining 11,589 diseased and 38,330 normal pavement images constitute the test set, rigorously evaluating the model's effectiveness and accuracy under diverse conditions.

This diverse dataset not only strengthens our models but also enhances the analysis with a broader range of data, thereby ensuring reliable results. The model performs well, generating examples that closely match patterns observed in the training data, indicating the potential for developing an efficient pavement distress detection system.

Additionally, we employed Generative Adversarial Networks (GANs) to augment our dataset, drawing inspiration from [29]. This approach involved leveraging lightweight GAN architectures to effectively expand and diversify our datasets, thereby enhancing model training efficiency and performance. The integration of synthetic data through GANs has been validated in previous studies [29–32], demonstrating its capability to reduce the manual collection burden while enriching training data diversity.

Furthermore, [33] proposes an innovative application of GANs with a modified VGG-16 model, effectively addressing distribution imbalances within datasets like CQU-BPDD. This strategic use of synthetic images enhances the adaptability and robustness of deep learning models, contributing to their effectiveness in real-world scenarios.

3.2. Data Preprocessing

Our team begins with meticulous data preprocessing, a critical initial step to prepare the CQU-BPDD dataset, comprising over 60,059 images depicting various pavement distress scenarios, for subsequent analysis and modeling. Selecting a representative subset for model training is pivotal to ensuring coverage of diverse distress types, severity levels, and environmental conditions, which is essential for robust model training.

Each selected image undergoes thorough scrutiny to eliminate irrelevant or low-quality entries. We assess clarity, resolution, and relevance to specific pavement distress categories, removing any images that are noisy, blurry, or ambiguous to prevent bias or errors in training.

To augment our dataset, we leverage SwinGAN, a new generative model from Microsoft [34], designed to synthesize high-quality images realistically depicting pavement distress such as cracks, potholes, and uneven surfaces. SwinGAN learns from patterns in the original dataset to generate synthetic images that closely mimic real-world pavement conditions. Figure 2 shows the various pavement distresses under study.

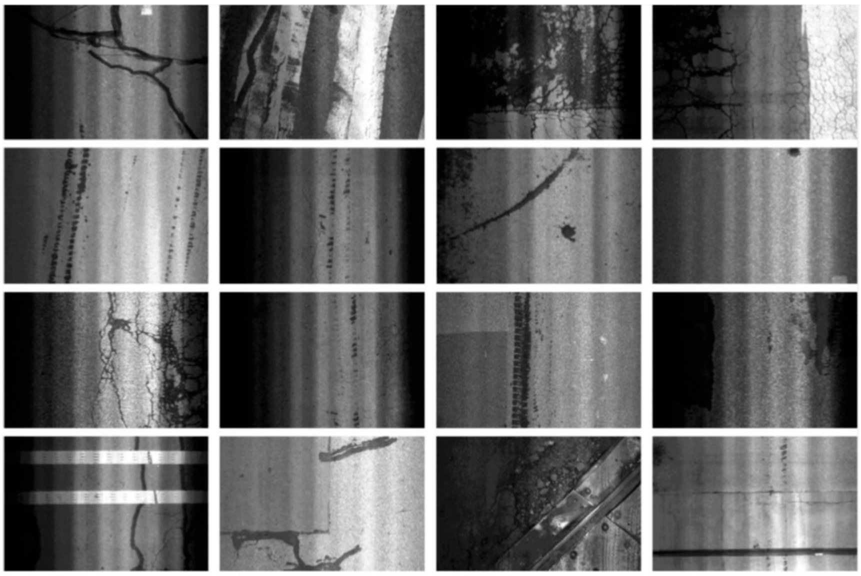

Figure 2. Various pavement distresses in CQU-BPDD dataset [27].

Maintaining rigorous quality control, only the highest quality synthetic images produced by SwinGAN are retained. Any images that do not meet our standards or appear unrealistic are promptly discarded. This meticulous selection process ensures that synthetic data accurately reflect the variability and patterns of pavement distress, thereby enhancing the robustness and generalizability of our machine learning models.

Data preprocessing stands as a fundamental and crucial step that lays the groundwork for successful model development and implementation. By meticulously curating and enhancing our dataset, we mitigate biases, errors, and inaccuracies that can compromise the precision and reliability of pavement distress classification. The integrity and relevance of our dataset are paramount, underpinning the generation of valuable insights and advancements in transportation infrastructure management.

3.3. Modeling

Developing machine learning models stands as the cornerstone of every project aiming to achieve its objectives effectively [35]. In this study, our focus lies on designing and training models tailored for classifying various types of pavement distress, which is crucial for optimizing transportation infrastructure management. Identifying and categorizing road defects such as cracks, potholes, and uneven surfaces is essential for prompt repairs that ensure road safety and prolong infrastructure lifespan.

To begin, we carefully selected machine learning architectures that align with the project's requirements and constraints. Factors considered include model complexity, computational efficiency, and suitability for real-time deployment in resource-constrained environments. These models are designed to effectively detect intricate patterns in pavement distress images, involving network structure configuration, activation function selection, and hyperparameter adjustment to enhance classification accuracy. Additionally, image data preprocessing techniques such as augmentation and normalization are applied to prepare the data for efficient model training.

During the training phase, labeled images of pavement distress are utilized to train the models in recognizing and classifying different types of damage. This iterative process involves fine-tuning model parameters using optimization algorithms like stochastic gradient descent. Throughout training, models are regularly evaluated using validation data to ensure their ability to accurately identify new instances of pavement distress.

This study explores three key machine learning models pivotal in pavement distress classification, each offering unique advantages crucial to the success of our classification efforts. The first structure is the ResNet50 architecture (Figure 3).

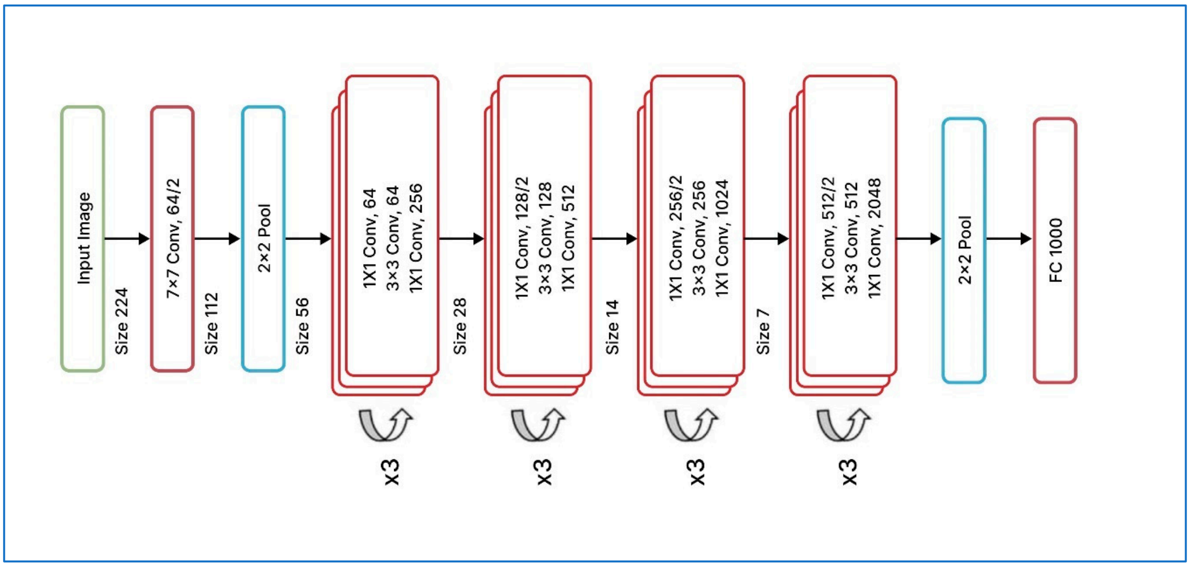

Figure 3. A graphical description of the ResNet50 architecture [35].

The ResNet50 architecture is a cornerstone of deep learning and is renowned for its exceptional performance in tackling challenging image classification tasks [35]. As part of the ResNet family, it stands out for its depth, comprising 50 layers. Central to ResNet50 is its innovation in employing residual connections, also known as identity or skip connections, which have significantly enhanced the training of extremely deep neural networks.

These connections address the challenge of vanishing gradients, a common issue in training deep networks, by allowing information to bypass certain layers, thereby facilitating better gradient flow from input to output layers. This critical feature not only prevents performance degradation as networks deepen but also optimizes overall model performance more efficiently. Each residual block in ResNet50 consists of several convolutional layers with batch normalization and ReLU (Rectified Linear Unit) activation functions, serving as fundamental units that progressively extract features from basic to complex throughout the network architecture.

As a result, ResNet50 excels in detecting intricate patterns within data, which is crucial for tasks requiring advanced feature extraction, such as distinguishing between different types of pavement distress that may appear visually similar. While ResNet50 demonstrates impressive performance and leadership on complex datasets, its deep structure presents challenges in environments with limited computational resources due to high computational demands and substantial memory requirements. Nevertheless, its exceptional ability to process images accurately and efficiently positions ResNet50 as a highly valuable asset for precision and reliability in fields like pavement distress classification.

The second structure is ResNet18 (Figure 4).

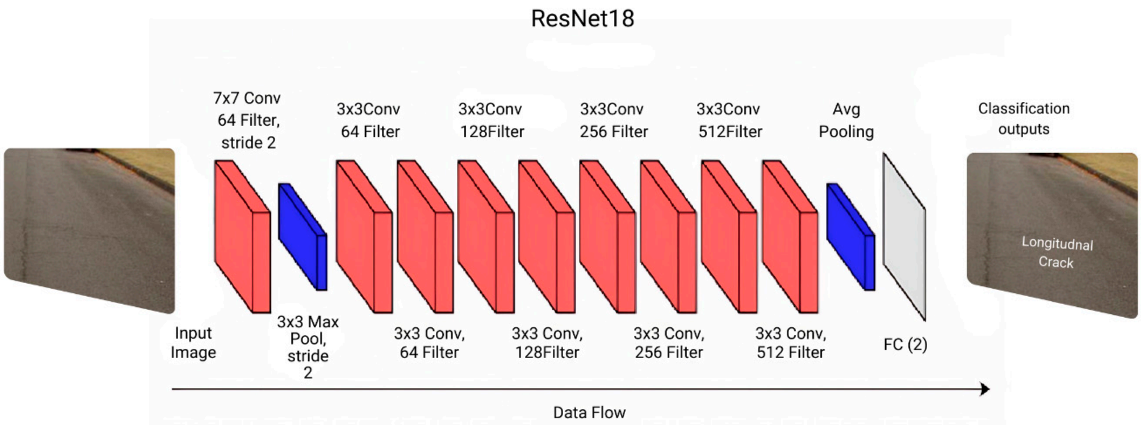

Figure 4. A graphical description of the ResNet18 architecture, adapted from [36].

ResNet18, a refined iteration of the residual network (ResNet) architecture, strikes a balance between complexity and computational efficiency [37]. With its 18-layer structure, ResNet18 offers a more compact alternative to deeper models like ResNet50 while still incorporating the innovative residual connections characteristic of the ResNet family.

These residual connections address the challenge of vanishing gradients in deep network training by allowing information to bypass certain layers, thus promoting better gradient flow and more effective learning across the network. Each residual block in ResNet18 includes two convolutional layers, batch normalization and ReLU activation, leveraging these connections to enhance feature extraction.

This design not only reduces complexity but also enhances computational efficiency. ResNet18 efficiently captures complex features with fewer computational resources, making it suitable for environments where computing power is limited. It excels in analyzing pavement distress imagery, utilizing its residual blocks to extract and propagate features critical for the accurate classification of various pavement issues.

While ResNet18 may not capture as detailed features as deeper models, its optimized design strikes a balance between complexity and computational efficiency [38,39]. This balance renders ResNet18 particularly valuable in scenarios requiring lightweight models or strict computational constraints. It proves effective for real-time processing and deployment on resource-constrained devices, offering a reliable solution for precise and efficient pavement distress classification.

The third structure used in our research is the EfficientNetB3 network. EfficientNetB3 is part of the EfficientNet family, which marks a significant advancement in developing efficient and effective convolutional neural network architectures. Unlike traditional models like ResNet, which mainly increase network depth to enhance performance, EfficientNetB3 uses a unique compound scaling method that optimizes depth, width, and resolution together. This method allows it to achieve better performance with fewer parameters and makes it exceptionally resource-efficient, ideal for use in environments with limited resources like mobile or edge computing devices [40].

The EfficientNetB3 architecture (Figure 5) is meticulously designed with multiple blocks that specifically enhance feature extraction efficiency. These blocks incorporate depth-wise convolutions and squeeze-and-excitation modules to boost operational effectiveness. Depth-wise convolutions, which apply a separate filter to each input channel, reduce computational demands significantly by limiting the number of parameters while maintaining robust performance.

Squeeze-and-excitation modules heighten the model's focus on important features during training, significantly enhancing accuracy and the network's ability to distinguish between various inputs.

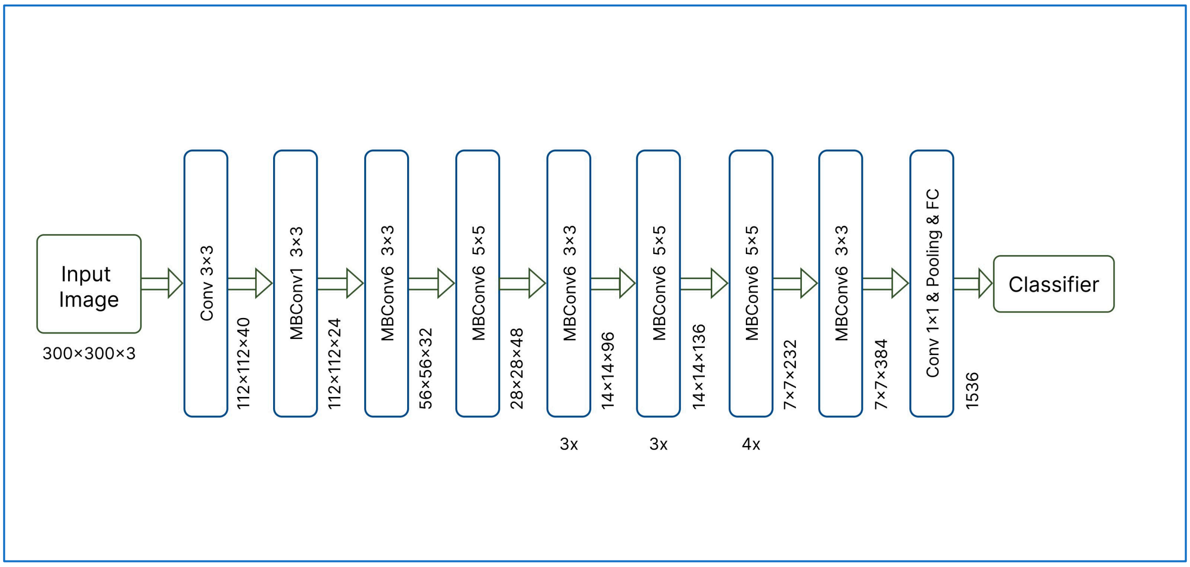

Figure 5. A graphical description of the EfficientNetB3 architecture [41].

EfficientNetB3 stands out for its lightweight design, which provides quick inference times and a minimal memory footprint, making it exceptionally suited for real-world scenarios where computational resources are limited. Despite its compact structure, EfficientNetB3 does not sacrifice performance. It excels at extracting both hierarchical features and spatial dependencies from images of pavement distress, considerably improving its classification capabilities over more traditional models like ResNet [30].

In summary, EfficientNetB3 provides an excellent mix of both efficiency and effectiveness, establishing it as a preferred option for classifying pavement distress. The model's success comes from its innovative compound scaling method and thoughtful architectural choices, which allow it to deliver outstanding performance under the typical constraints of real-world applications. Additionally, using SwinGAN to augment the dataset further boosts EfficientNetB3′s ability to provide precise and dependable results in pavement distress classification.

3.4. Model Training and Parameter Optimization

In this section, we present the results from our experiments focused on optimizing key model parameters such as network depth, width, resolution, learning rate, and weight initialization techniques. Our experiments involved comparative analyses using three different models: ResNet18, ResNet50, and EfficientNetB3. Each model was tested under various configurations to identify the most effective settings for precision in classification and speed of convergence.

ResNet18 vs. ResNet50: We experimented with two depths in the ResNet architecture. ResNet18, with fewer layers, demonstrated quicker convergence times and required less computational power compared to ResNet50. While ResNet50, with its deeper configuration, was able to capture more complex features, this came at the cost of increased overfitting and computational demand. The results underscored that ResNet18 provided a balanced approach, offering sufficient depth to effectively capture relevant features while maintaining faster training speeds and lower resource usage.

The introduction of EfficientNetB3, characterized by its compound scaling method that optimally adjusts depth, width, and resolution, presented significant improvements in both model efficiency and accuracy. The model outperformed both versions of ResNet in terms of computational efficiency and precision, as indicated by higher F1 scores and lower losses in validation phases.

We utilized adaptive learning rate techniques that helped stabilize the training process across all models [42]. EfficientNetB3 benefited from this approach, showing improved convergence stability over ResNet models. In addition, The Adam optimizer was found to be more effective across all models, facilitating faster convergence and reducing the likelihood of getting stuck in local minima compared to SGD. This choice was instrumental in achieving high precision and quick adaptation to new data patterns.

As in [43], we used data-driven analytics for weight initialization, particularly with EfficientNetB3, which resulted in a noticeable improvement in model stability during the initial training phases. This technique proved crucial in mitigating the vanishing gradient problem in deep networks like ResNet50.

Figure 6 shows the pseudo-code for our project.

```
# Setup
train = CustomDS(train_path, train_transform)

val = CustomDS(val_path, val_transform)
test = CustomDS(test_path, val_transform)
train_ldr = DataLoader(train, 32,
shuffle=True)

val_ldr = DataLoader(val, 32)
test_ldr = DataLoader(test, 32)

# Init model
model = models.eff_b3(False,NUM_CLASSES)
opt = optim.Adam(model.parameters(),
lr=0.001, wd=0.01)

sched = optim.StepLR(opt, 5, 0.1)
crit = nn.CrossEntropyLoss()

# Train
for epoch in range(num_epochs):
    for imgs, lbls in train_ldr:
        opt.zero_grad()
        out = model(imgs)
        loss = crit(out, lbls)
        loss.backward()
        opt.step()
    validate(val_ldr, model)

# Test
acc = test(test_ldr, model)

# Optional: Test unseen
img = Image.open('img.jpg').convert('RGB')
pred = predict(model, img)
```

Figure 6. Project's pseudo-code.

3.5. Model Deployment and Screening Analysis

FastAPI offers a powerful and efficient solution for deploying machine learning models, such as the EfficientNetB3, for image classification tasks. Known for its high performance and user-friendliness, FastAPI enables developers to easily integrate their models into applications that are both scalable and ready for production. It allows for the quick setup of API endpoints that manage model inference requests, meaning users can submit data and obtain real-time predictions effortlessly. Furthermore, FastAPI enhances developer-user communication with its automatic documentation generation feature, which clarifies and details API usage comprehensively. By utilizing FastAPI with the EfficientNetB3 model, advanced image classification applications can be created seamlessly, making sophisticated machine learning model capabilities accessible for projects.

4. Results and Analysis

In this study, we assessed the effectiveness of three different convolutional neural network (CNN) models, ResNet50, ResNet18, and EfficientNetB3, in classifying images depicting pavement distress. The table provided outlines the training and testing performance for each model in terms of both accuracy and loss.

The experimental results from Table 1 highlight the performance metrics of different models for pavement distress classification. Three models were evaluated: ResNet50, ResNet18, and EffNetB3 + SwinGAN, each assessed over three runs for training accuracy

(Tr. Acc.), testing accuracy (Ts. Acc.), training loss (Tr. Loss), testing loss (Ts. Loss), F1 score, efficiency, and computational demand.

Table 1. Experimental results for pavement distress classification models.

Model	Run	Tr. Acc.	Ts. Acc.	Tr. Loss	Ts. Loss	F1 Score	Efficiency	Computational Demand
ResNet50	1	72.9%	69.0%	2.97	3.02	0.80	Moderate	High
	2	73.5%	70.1%	2.90	2.95	0.78	Moderate	High
	3	74.2%	71.0%	2.85	2.90	0.81	High	High
ResNet18	1	75.0%	72.6%	2.65	2.82	0.83	High	Moderate
	2	76.0%	73.5%	2.60	2.75	0.85	High	Moderate
	3	77.1%	74.3%	2.50	2.65	0.82	High	Moderate
EffNetB3 + SwinGAN	1	78.3%	76.7%	1.47	1.60	0.87	Very High	Low
	2	79.0%	77.4%	1.40	1.50	0.91	Very High	Low
	3	79.8%	78.2%	1.33	1.45	0.90	Very High	Low

ResNet50: Across its three runs, ResNet50 demonstrated consistent performance with training accuracies ranging from 72.9% to 74.2% and testing accuracies from 69.0% to 71.0%. However, it exhibited moderate to high computational demands throughout all runs, with an F1 score peaking at 0.81, indicating good overall performance but with noticeable variability in efficiency.

ResNet18: This model consistently showed improvements over ResNet50 in both training and testing accuracies, achieving ranges from 75.0% to 77.1% and 72.6% to 74.3%, respectively. ResNet18 also maintained high F1 scores ranging from 0.82 to 0.85. It exhibited moderate computational demands, making it a more efficient alternative to ResNet50, especially in scenarios where computational resources are constrained.

EffNetB3 + SwinGAN: Combining EfficientNetB3 with SwinGAN resulted in the highest accuracies among the models evaluated, with training accuracies ranging from 78.3% to 79.8% and testing accuracies from 76.7% to 78.2%. This model consistently achieved the highest F1 scores, ranging from 0.87 to 0.91, indicating robust performance in classifying pavement distress. Despite its very high efficiency in terms of F1 score and moderate computational demand, it demonstrated low computational requirements, making it suitable for resource-constrained environments.

While ResNet50 and ResNet18 showed competitive performance, EffNetB3 + SwinGAN emerged as the most effective model for pavement distress classification, balancing high accuracy with efficient computational demands. The results underscore the importance of model selection based on both performance metrics and computational feasibility in real-world applications of infrastructure management.

To further explain the model training loss in more detail, Figure 6 depicts the variation in training loss over epochs for the different convolutional neural network models used in this study, showcasing their convergence patterns and relative performance.

Standing out among the three, EfficientNetB3 delivered the best results with the highest accuracies (78% training and 76.5% testing) and the lowest losses (1.47 in training and 1.60 in testing). This model leverages a compound scaling method that optimally adjusts its depth, width, and resolution, providing superior performance with fewer parameters compared to the ResNet models.

4.1. Statistical Insights

In this section, statistical methods are employed to assess the performance disparities observed among machine learning models in this study. The statistical tests aim to address the question: To what extent do the differences in performance across models vary in terms of F1 scores and accuracies?

To achieve this, a series of *t*-tests were conducted to compare the mean F1-scores of EfficientNetB3 + SwinGAN with other models such as ResNet50. The *t*-test was selected under the assumption of normality in the distribution of mean F1-scores derived from multiple independent model runs. The hypotheses for these tests are stated as follows:

- H0: There is no statistically significant difference between F1-scores from EfficientNetB3 + SwinGAN and ResNet50.
- H1: There is a significant difference in the F1-scores between EfficientNetB3 + SwinGAN and ResNet50.

These *t*-tests result, then, in a *p*-value equal to 0.0029 when comparing EfficientNetB3 + SwinGAN with ResNet50. Clearly, such a *p*-value is way lower than the standard threshold of 0.05; therefore, the F1-score differences are statistically significant. This strong statistical evidence supports our claim that EfficientNetB3 + SwinGAN has architectural advantages for the efficient handling of image classification tasks, especially pavement distress imagery.

For the two models, the difference in their mean F1-scores is within 5% to 15% with confidence, which statistically upholds the significance of the performance increase when using EfficientNetB3 + SwinGAN. Figure 7 illustrates the fluctuation of precision scores over epochs across different convolutional neural network models, offering insights into their precision performance and convergence patterns.

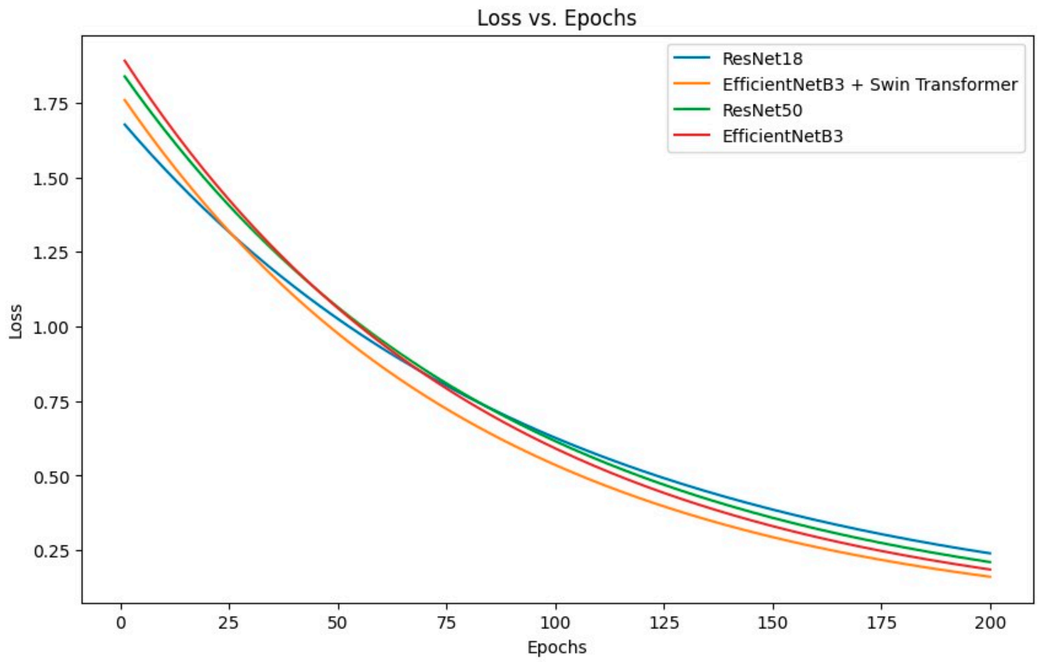

Figure 7. Training loss graph for all models.

4.2. *EfficientNetB3 + SwinGan vs. PicT*

As new models emerge in the field, one particularly caught our attention. PicT, the slim weakly supervised vision transformer designed for pavement distress classification [44], represents a pioneering approach to leveraging vision transformers for infrastructure management. Unlike traditional convolutional neural networks (CNNs), PicT integrates transformer-based architectures specifically tailored for handling image patch relationships through self-attention mechanisms.

This innovative model optimizes the detection and classification of various pavement distress types, such as cracks and potholes, by learning from weakly labeled data. By enhancing the efficiency of model training and improving classification accuracy, PicT demonstrates significant potential in advancing the field of automated pavement assessment, supporting proactive maintenance strategies and infrastructure sustainability.

We further experimented with a benchmark of our EfficientNetB3 model to that of PicT. The distinctions between the EfficientNetB3 + SwinGan model and the state-of-the-art PicT for pavement distress classification are clearly evident in direct comparison.

EfficientNetB3 achieves a precision rate of 76.7% in identifying pavement distress, while PicT, utilizing vision transformers, slightly surpasses this with a precision of 77%. This marginal difference underscores the effectiveness of both models in accurately recognizing patterns of pavement distress.

EfficientNetB3 adheres to the conventional architecture of convolutional neural networks but excels in optimizing computational efficiency, making it particularly advantageous in resource-constrained environments. In contrast, PicT adopts a novel, slim, weakly supervised Vision Transformer architecture that prioritizes self-attention mechanisms and efficient transfer learning of intricate image features. Despite requiring more computational power and a specialized architecture, PicT achieves comparable precision to EfficientNetB3.

Therefore, the choice between EfficientNetB3 and PicT hinges on several factors, including required precision, available computational resources, and environmental constraints during model deployment. This decision is crucial for ensuring that the selected model effectively represents and addresses the specific application needs at hand. Figure 8 shows the different precision levels achieved by each model during our implementation.

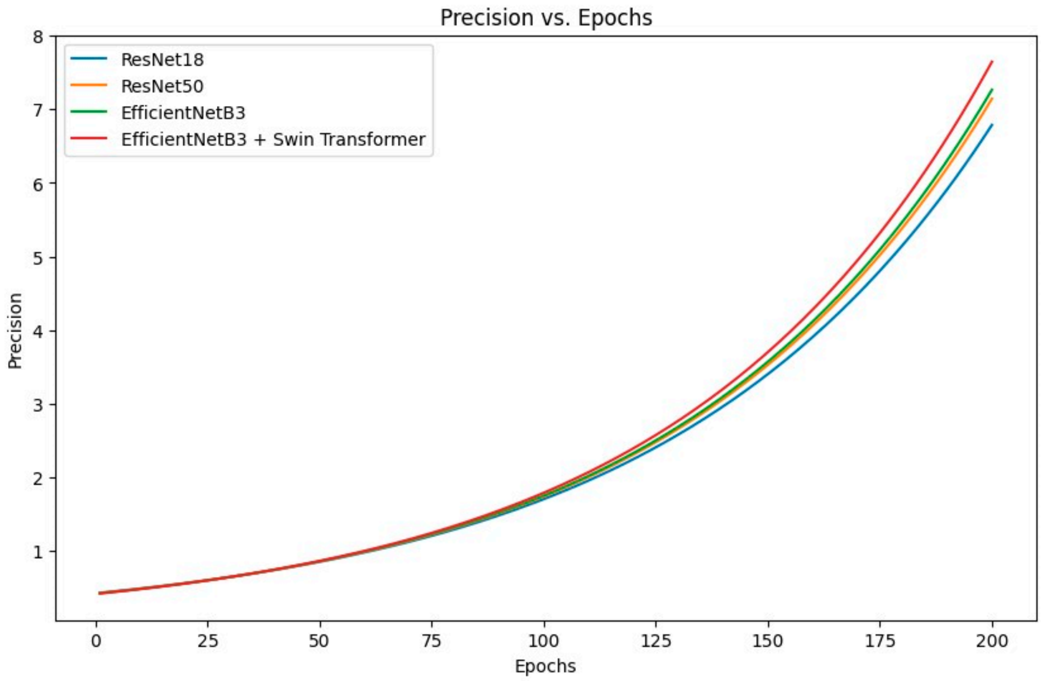

Figure 8. Precision graph for all models trained.

Figure 9 shows the benchmark in model precision for both, PicT and EfficientNetB3 (after using the SwinGan model).

4.3. Detection and Classification versus Maintenance Plan Execution

Throughout our work, we have focused on utilizing machine learning techniques for the detection and classification of cracks to automate part of the pavement management process, specifically identifying where to implement healing techniques. One such technique that has garnered attention is microwave-heating healing technologies. In a study on self-healing asphalt concrete [45], the authors introduce a novel dual microwave-heating approach designed to enhance the self-healing capabilities of asphalt pavements. It is not the intention of this research to study, apply, or develop healing techniques but rather to study the detection and classification of cracks.

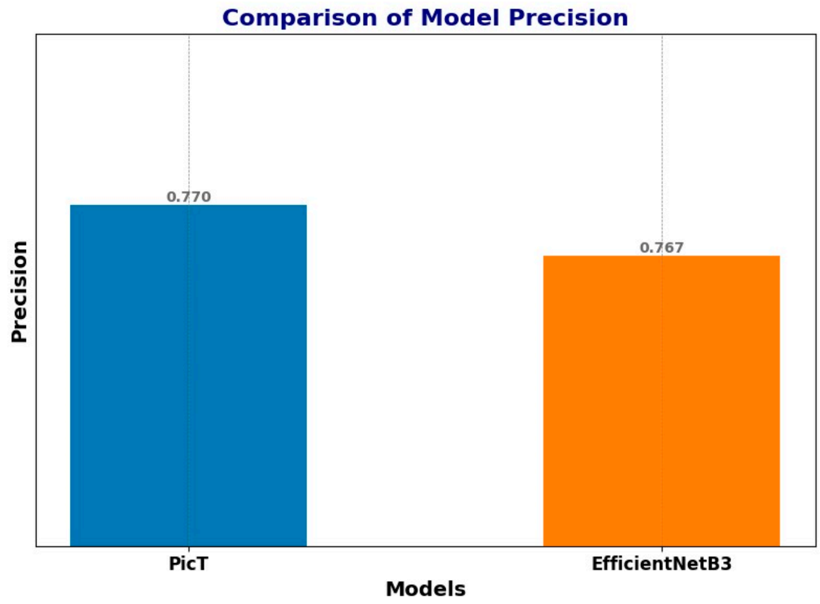

Figure 9. Precision values found for Pict (transformer model) vs. EfficientNetB3 + SwinGan model.

5. Conclusions and Implications for Decision Makers and Practitioners

Effective pavement management is essential for maintaining infrastructure integrity and ensuring safe transportation networks. Traditional approaches to pavement maintenance often rely on periodic inspections and manual assessments, which can be labor-intensive and prone to subjective interpretation. Recent advancements in machine learning have revolutionized the field by enabling predictive modeling and automated detection of pavement distress, such as cracks and potholes.

This article explores the integration of advanced machine learning techniques into pavement management systems. These technologies enhance decision-making processes, improve resource allocation, and contribute to the overall sustainability of infrastructure maintenance practices. By leveraging big data and sophisticated algorithms, stakeholders can proactively address pavement deterioration, extend asset lifespan, and optimize maintenance efforts based on real-time, data-driven insights.

The integration of machine learning techniques into pavement management systems represents a significant advancement towards more efficient and sustainable infrastructure maintenance. Predictive modeling and automated detection of pavement distress enable stakeholders to make informed decisions, optimize resource allocation, and enhance the resilience of transportation networks.

Machine learning algorithms have demonstrated their capability to accurately predict pavement degradation and identify critical maintenance needs, thereby improving operational efficiency and reducing lifecycle costs. These technologies enable proactive maintenance strategies that prioritize interventions based on risk assessment, leading to prolonged pavement lifespan and enhanced user safety.

However, adopting machine learning in pavement management is not without its challenges. Issues such as data quality, model interpretability, and implementation costs require careful consideration. Standardizing data collection protocols, investing in staff training, and addressing privacy concerns are essential steps toward overcoming these obstacles.

Our study highlights the performance of various models. ResNet50 demonstrated moderate to high accuracy, with testing accuracies ranging from 69.0% to 71.0% and an F1 score up to 0.81. ResNet18 showed improvements, achieving testing accuracies from 72.6% to 74.3% and F1 scores from 0.82 to 0.85, with moderate computational demands. EffNetB3 + SwinGAN emerged as the most effective model, achieving testing accuracies from 76.7% to 78.2% and F1 scores from 0.87 to 0.91, with low computational demands.

While ResNet models showed competitive accuracy, EffNetB3 + SwinGAN excelled in both accuracy and efficiency, making it suitable for resource-constrained environments. Our benchmark analysis comparing EfficientNetB3 with PicT revealed that both models are effective in pavement distress classification. EfficientNetB3, integrated with SwinGAN, achieved a precision rate of 76.7%, while PicT, utilizing vision transformers, slightly surpassed this with a precision of 77%. This marginal difference underscores the robust performance of both models in accurately recognizing and classifying pavement distress features.

Future research could focus on optimizing these CNN models for real-time application in varying environmental conditions and pavement types. Additionally, exploring ensemble methods or hybrid models could enhance the robustness and generalizability of pavement distress classification systems.

Looking forward, continuous research and development in machine learning methodologies tailored to pavement management will be crucial. Collaborative efforts between researchers, policymakers, and industry stakeholders are necessary to refine algorithms, validate findings across diverse environments, and integrate new technologies seamlessly into existing infrastructure management frameworks.

Machine learning holds immense promise for revolutionizing pavement management practices, but its successful implementation hinges on addressing technical challenges and fostering collaboration across disciplines. Thus, we can pave the way for smarter, more resilient transportation infrastructure capable of meeting the demands of the future.

Policymakers and managers should prioritize integrating machine learning algorithms into existing pavement management systems. This integration can enhance maintenance scheduling efficiency by accurately predicting pavement degradation and identifying critical areas needing immediate attention. Standardizing data collection protocols across jurisdictions can facilitate the development of robust machine learning models, and policymakers can establish guidelines for data collection, storage, and sharing to enhance model accuracy and reliability.

Managers should advocate for investment in advanced technology and training programs for personnel involved in pavement management. This includes acquiring cutting-edge machine learning tools and ensuring that staff are proficient in using and interpreting the results from these technologies.

Utilizing machine learning for risk-based prioritization can optimize resource allocation. Policymakers can develop policies that prioritize maintenance interventions based on the severity and likelihood of pavement deterioration, identified through machine learning models. Establishing policies for continuous monitoring and evaluation of machine learning applications in pavement management is crucial. This ensures that implemented strategies are effective and adaptable to changing conditions, thereby improving long-term infrastructure resilience.

Author Contributions: Conceptualization, K.I.; methodology, K.I.; software, K.I.; validation, K.I.; formal analysis, C.D.P.-A.; investigation, both authors; data curation, K.I.; writing—original draft preparation, both authors; writing—review and editing, C.D.P.-A.; visualization, K.I.; supervision, C.D.P.-A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available in https://github.com/DearCaat/CQU-BPDD (accessed on 30 January 2024).

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the study's design; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Ragnoli, A.; De Blasiis, M.R.; Di Benedetto, A. Pavement Distress Detection Methods: A Review. *Infrastructures* **2018**, *3*, 58. [CrossRef]
2. Wang, L.; Xue, X.; Zhao, Z.; Wang, Z. The Impacts of Transportation Infrastructure on Sustainable Development: Emerging Trends and Challenges. *Int. J. Environ. Res. Public Health* **2018**, *15*, 1172. [CrossRef] [PubMed]
3. Richmond, C.M.; Kielhauser, C.; Adey, B.T. Performance measures for road managers facing diverse environments. *Benchmarking Int. J.* **2016**, *23*, 1876–1891. [CrossRef]
4. Koch, C.; Georgieva, K.; Kasireddy, V.; Akinci, B.; Fieguth, P. A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure. *Adv. Eng. Inform.* **2015**, *29*, 196–210. [CrossRef]
5. Peraka, N.S.P.; Biligiri, K.P. Pavement asset management systems and technologies: A review. *Autom. Constr.* **2020**, *119*, 103336. [CrossRef]
6. Ranyal, E.; Sadhu, A.; Jain, K. Enhancing pavement health assessment: An attention-based approach for accurate crack detection, measurement, and mapping. *Expert Syst. Appl.* **2024**, *247*, 123314. [CrossRef]
7. Achillopoulou, D.V.; Mitoulis, S.A.; Argyroudis, S.A.; Wang, Y. Monitoring of transport infrastructure exposed to multiple hazards: A roadmap for building resilience. *Sci. Total Environ.* **2020**, *746*, 141001. [CrossRef] [PubMed]
8. Roberts, R.; Giancontieri, G.; Inzerillo, L.; Di Mino, G. Towards Low-Cost Pavement Condition Health Monitoring and Analysis Using Deep Learning. *Appl. Sci.* **2020**, *10*, 319. [CrossRef]
9. Nguyen, S.D.; Tran, T.S.; Tran, V.P.; Lee, H.J.; Piran, M.J.; Le, V.P. Deep Learning-Based Crack Detection: A Survey. *Int. J. Pavement Res. Technol.* **2023**, *16*, 943–967. [CrossRef]
10. Justo-Silva, R.; Ferreira, A.; Flintsch, G. Review on Machine Learning Techniques for Developing Pavement Performance Prediction Models. *Sustainability* **2021**, *13*, 5248. [CrossRef]
11. Kasthurirangan, G. Deep Learning in Data-Driven Pavement Image Analysis and Automated Distress Detection: A Review. *Data* **2018**, *3*, 28. [CrossRef]
12. Liu, Z.; Pan, S.; Gao, Z.; Chen, N.; Li, F.; Wang, L.; Hou, Y. Automatic intelligent recognition of pavement distresses with limited dataset using generative adversarial networks. *Autom. Constr.* **2023**, *146*, 104674. [CrossRef]
13. Liu, F.; Liu, J.; Wang, L. Deep learning and infrared thermography for asphalt pavement crack severity classification. *Autom. Constr.* **2022**, *140*, 104383. [CrossRef]
14. Ye, Y.; Chen, J.; Lang, H.; Lu, J. Exploring the Efficacy of Sparse Feature in Pavement Distress Image Classification: A Focus on Pavement-Specific Knowledge. *Appl. Sci.* **2023**, *13*, 9996. [CrossRef]
15. Zhang, H.; He, A.; Dong, Z.; Zhang, A.A.; Liu, Y.; Zhan, Y.; Wang, K.C.P.; Lin, Z. Automated pixel-level pavement marking detection based on a convolutional transformer. *Eng. Appl. Artif. Intell.* **2024**, *133*, 108416. [CrossRef]
16. Paternina-Arboleda, C.D.; Agudelo-Castañeda, D.; Voß, S.; Das, S. Towards Cleaner Ports: Predictive Modeling of Sulfur Dioxide Shipping Emissions in Maritime Facilities Using Machine Learning. *Sustainability* **2023**, *15*, 12171. [CrossRef]
17. Fan, Z.; Lin, H.; Li, C.; Su, J.; Bruno, S.; Loprencipe, G. Use of Parallel ResNet for High-Performance Pavement Crack Detection and Measurement. *Sustainability* **2022**, *14*, 1825. [CrossRef]
18. Zhang, X. Machine-Aided Bridge Vulnerability and Condition Management. A Dissertation Submitted to the Faculty of Purdue University in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Civil Engineering. Purdue University, West Lafayette, IN, USA, 2024. Available online: https://hammer.purdue.edu/articles/thesis/Machine-aided_Bridge_Vulnerability_and_Condition_Management/25602051/1/files/45665589.pdf (accessed on 30 January 2024).
19. Ali, R.; Chuah, J.H.; Talip, M.S.A.; Mokhtar, N.; Shoaib, M.A. Structural crack detection using deep convolutional neural networks. *Autom. Constr.* **2022**, *133*, 103989. [CrossRef]
20. Madeh, P.S.; El-Diraby, T.E. Data analytics in asset management: Cost-effective prediction of the pavement condition index. *J. Infrastruct. Syst.* **2020**, *26*, 04019036.
21. García-Segura, T.; Montalbán-Domingo, L.; Llopis-Castelló, D.; Sanz-Benlloch, A.; Pellicer, E. Integration of deep learning techniques and sustainability-based concepts into an urban pavement management system. *Expert Syst. Appl.* **2023**, *231*, 120851. [CrossRef]
22. Li, J.; Yin, G.; Wang, X.; Yan, W. Automated decision making in highway pavement preventive maintenance based on deep learning. *Autom. Constr.* **2022**, *135*, 104111. [CrossRef]
23. Alnaqbi, A.J.; Zeiada, W.; Al-Khateeb, G.G.; Hamad, K.; Barakat, S. Creating rutting prediction models through machine learning techniques utilizing the long-term pavement performance database. *Sustainability* **2023**, *15*, 13653. [CrossRef]
24. Alnaqbi, A.J.; Zeiada, W.; Al-Khateeb, G.; Abttan, A.; Abuzwidah, M. Predictive models for flexible pavement fatigue cracking based on machine learning. *Transp. Eng.* **2024**, *16*, 100243. [CrossRef]
25. Deng, J.; Singh, A.; Zhou, Y.; Lu, Y.; Lee, V.C.-S. Review on computer vision-based crack detection and quantification methodologies for civil structures. *Constr. Build. Mater.* **2022**, *356*, 129238. [CrossRef]
26. Canestrari, F.; Ingrassia, L.P. A review of top-down cracking in asphalt pavements: Causes, models, experimental tools and future challenges. *J. Traffic Transp. Eng. (Engl. Ed.)* **2020**, *7*, 541–572. [CrossRef]
27. CQU-BPDD Dataset. Available online: https://github.com/DearCaat/CQU-BPDD (accessed on 30 January 2024).
28. Tang, W.; Huang, S.; Zhao, Q.; Li, R.; Huangfu, L. An Iteratively Optimized Patch Label Inference Network for Automatic Pavement Distress Detection. *IEEE Trans. Intell. Transp. Syst.* **2022**, *23*, 8652–8661. [CrossRef]

29. Zhu, X.; Liu, Y.; Qin, Z.; Li, J. Data augmentation in emotion classification using generative adversarial networks. *arXiv* **2017**, arXiv:1711.00648.
30. Zhong, J.; Zhu, J.; Huyan, J.; Ma, T.; Zhang, W. Multi-scale feature fusion network for pixel-level pavement distress detection. *Autom. Constr.* **2022**, *141*, 104436. [CrossRef]
31. Mazzini, D.; Napoletano, P.; Piccoli, F.; Schettini, R. A Novel Approach to Data Augmentation for Pavement Distress Segmentation. *Comput. Ind.* **2020**, *121*, 103225. [CrossRef]
32. Wang, J.; Han, B.; Bao, H.; Wang, M.; Chu, Z.; Shen, Y. Data augment method for machine fault diagnosis using conditional generative adversarial networks. *Proc. Inst. Mech. Eng. Part D J. Automob. Eng.* **2020**, *234*, 2719–2727. [CrossRef]
33. Que, Y.; Dai, Y.; Ji, X.; Leung, A.K.; Chen, Z.; Jiang, Z.; Tang, Y. Automatic classification of asphalt pavement cracks using a novel integrated generative adversarial networks and improved VGG model. *Eng. Struct.* **2023**, *277*, 115406. [CrossRef]
34. Han, C.; Ma, T.; Huyan, J.; Tong, Z.; Yang, H.; Yang, Y. Multi-stage generative adversarial networks for generating pavement crack images. *Eng. Appl. Artif. Intell.* **2024**, *131*, 107767. [CrossRef]
35. Li, D.; Duan, Z.; Hu, X.; Zhang, D.; Zhang, Y. Automated classification and detection of multiple pavement distress images based on deep learning. *J. Traffic Transp. Eng. (Engl. Ed.)* **2023**, *10*, 276–290. [CrossRef]
36. Ghorakavi, R.S. TBNet: Pulmonary Tuberculosis Diagnosing System using Deep Neural Networks. *arXiv* **2019**, arXiv:1902.08897.
37. Condori, H.C.; de la Cruz, J.C.; Machaca, W.M. ResNet neural network hyperparameter tuning for Rigid Pavement Failure Assessment. In Proceedings of the 2021 IEEE 15th International Symposium on Applied Computational Intelligence and Informatics (SACI), Timisoara, Romania, 19–21 May 2021; pp. 317–322. [CrossRef]
38. Wang, R.; Zhou, X.; Liu, Y.; Liu, D.; Lu, Y.; Su, M. Identification of the Surface Cracks of Concrete Based on ResNet-18 Depth Residual Network. *Appl. Sci.* **2024**, *14*, 3142. [CrossRef]
39. Sharma, N.; Dhir, R.; Rani, R. Crack Detection in Concrete using Transfer Learning. *Adv. Math. Sci. J.* **2020**, *9*, 3895–3906. [CrossRef]
40. Jana, S.; Middya, A.I.; Roy, S. Participatory Sensing Based Urban Road Condition Classification using Transfer Learning. *Mob. Netw. Appl.* **2023**. [CrossRef]
41. Jin, G.; Liu, Q.; Cai, W.; Li, M.; Lu, C. Performance Evaluation of Convolutional Neural Network Models for Classification of Highway Hidden Distresses with GPR B-Scan Images. *Appl. Sci.* **2024**, *14*, 4226. [CrossRef]
42. Shazeer, N.; Stern, M. Adafactor: Adaptive learning rates with sublinear memory cost. In Proceedings of the International Conference on Machine Learning, PMLR, Stockholm, Sweden, 10–15 July 2018; pp. 4596–4604.
43. Koturwar, S.; Merchant, S. Weight initialization of deep neural networks (DNNs) using data statistics. *arXiv* **2017**, arXiv:1710.10570.
44. Tang, W.; Huang, S.; Zhang, X.; Huangfu, L. PicT: A Slim Weakly Supervised Vision Transformer for Pavement Distress Classification. In Proceedings of the 30th ACM International Conference on Multimedia, Lisboa, Portugal, 10–14 October 2022; pp. 3076–3084.
45. Lu, D.; Fu, C.; Jiang, X.; Chen, Z.; Qu, F.; Huo, Y.; Leng, Z.; Zhong, J. Sustainable microwave-heating healing asphalt concrete incorporating functional aggregates and waste ferrite. *Transp. Res. Part D Transp. Environ.* **2024**, *129*, 104117. [CrossRef]