sustainability

MDPI

Article

Classification of Asphalt Pavement Defects for Sustainable Road Development Using a Novel Hybrid Technology Based on Clustering Deep Features

Jia Liang [1], Qipeng Zhang [2,*] and Xingyu Gu [3]

[1] School of Automotive and Traffic Engineering, Jiangsu University, Zhenjiang 212013, China; liangjiahs@ujs.edu.cn
[2] Department of Architecture and Civil Engineering, City University of Hong Kong, Hong Kong, China
[3] School of Transportation, Southeast University, Nanjing 211189, China; guxingyu1976@seu.edu.cn
* Correspondence: qipzhang@cityu.edu.hk

Abstract: In the rapid development of urbanization, the sustained and healthy development of transportation infrastructure has become a widely discussed topic. The inspection and maintenance of asphalt pavements not only concern road safety and efficiency but also directly impact the rational allocation of resources and environmental sustainability. To address the challenges of modern transportation infrastructure management, this study innovatively proposes a hybrid learning model that integrates deep convolutional neural networks (DCNNs) and support vector machines (SVMs). Specifically, the model initially employs a ShuffleNet architecture to autonomously extract abstract features from various defect categories. Subsequently, the Maximum Relevance Minimum Redundancy (MRMR) method is utilized to select the top 25% of features with the highest relevance and minimal redundancy. After that, SVMs equipped with diverse kernel functions are deployed to perform training and prediction based on the selected features. The experimental results reveal that the model attains a high classification accuracy of 94.62% on a self-constructed asphalt pavement image dataset. This technology not only significantly improves the accuracy and efficiency of pavement inspection but also effectively reduces traffic congestion and incremental carbon emissions caused by pavement distress, thereby alleviating environmental burdens. It is of great significance for enhancing pavement maintenance efficiency, conserving resource consumption, mitigating environmental pollution, and promoting sustainable socio-economic development.

Keywords: sustainable roadways; intelligent detection; hybrid learning; feature extraction; resource conservation

check for updates

Citation: Liang, J.; Zhang, Q.; Gu, X. Classification of Asphalt Pavement Defects for Sustainable Road Development Using a Novel Hybrid Technology Based on Clustering Deep Features. *Sustainability* **2024**, *16*, 10145. https://doi.org/10.3390/su162210145

Academic Editors: Davide Lo Presti, Giuseppe Sollazzo and Kelvin CP Wang

Received: 10 October 2024
Revised: 18 November 2024
Accepted: 18 November 2024
Published: 20 November 2024

1. Introduction

In the context of accelerating globalization and the continuous development of road transportation, the inspection and maintenance of highways have become significant issues of widespread international concern. As a core component of the transportation network, asphalt pavements, as direct carriers of transportation, directly impact road safety and traffic efficiency. During their service life, asphalt pavements are exposed to a combination of loads and local climatic conditions, and their surface inevitably produces crack defects due to aging [1]. Without timely and effective maintenance measures, these initial cracks further deteriorate and peel off and eventually evolve into more serious pavement defects, such as potholes [2]. As a common road distress, asphalt pavement cracks not only damage the smoothness and aesthetics of the road but also negatively affect regional economic development [3]. Existing research shows that delays in maintenance can lead to increasing maintenance costs over time [4]. Conventional pavement maintenance methods, such as manual inspections, are time-consuming, labor-intensive, and have low recognition accuracy and poor consistency, making them inadequate for the demands of modern transportation infrastructure management [5]. Therefore, the development of an efficient, precise, and highly automated pavement crack identification technology holds immense value for accelerating pavement maintenance processes, achieving efficient resource utilization, reducing environmental pollution, and promoting sustainable socio-economic development.

In the early stages, researchers primarily relied on basic image processing techniques and threshold segmentation methods to perform asphalt pavement defect detection tasks [6]. These techniques mainly depended on road engineers to collect pavement defect images on site. However, on-site image collection could lead to increased labor, material, and additional costs due to road closures, subsequently causing detection and maintenance costs to exceed expectations [7]. As computer science and technology rapidly evolve, conventional machine learning algorithms have significantly improved the efficiency and quality of pavement defect detection. A series of conventional machine learning algorithms, strongly supported by digital image technology, such as decision trees [8], random forests [9], SVMs [10], and Bayesian learning [11], have found extensive application in pavement defect detection [12–14], and have even been used to predict the development of pavement distresses [15]. This marks a significant leap in the advancement of intelligent pavement defect detection technology. After comprehensively evaluating the performance of various machine learning algorithms, it is not difficult to find that machine learning has become an indispensable means for maintaining road infrastructure [16], especially Artificial Neural Networks (ANNs), which have shown great potential in pavement crack classification tasks [17]. Additionally, SVM, as a classic classification algorithm, has delivered impressive results in the field of pavement defect identification due to its excellent classification performance and low computational complexity [18,19], demonstrating good application prospects. The widespread application of machine learning in these fields is primarily due to its robust predictive capabilities. However, as detection demands escalate, machine learning faces challenges, such as limited accuracy and efficiency in feature extraction, which are limitations that have hindered their widespread application in the field of road engineering to some extent [20].

In recent years, the rapid development of deep learning technologies, particularly the widespread application of convolutional neural networks (CNNs), has provided new solutions for pavement crack identification. CNNs, with their powerful feature extraction capabilities, have achieved significant results in fields such as image classification and object detection. Because they possess the ability to autonomously learn features and correlations between tasks, they demonstrate near-human potential [21]. Researchers have meticulously designed a series of high-performance CNN models, including AlexNet [21], GoogLeNet [22], VGGNet [23], ResNet [24], and DenseNet [25], leveraging large-scale general datasets such as ImageNet [26]. These advancements have opened new horizons in the field of pavement defect detection, bringing unprecedented possibilities and broad prospects. Preliminary investigations revealed that by constructing shallow CNN architectures to train low-resolution crack images, CNNs demonstrated superior predictive capabilities compared to conventional machine learning [27]. Furthermore, as the depth of these networks increased, the prediction performance of the models was further elevated [28]. With the relentless progression of technology, more sophisticated CNN models, characterized by deeper layers, have been devised and successfully implemented in practical engineering contexts [29,30]. These models have significantly improved the predictive performance of crack detection by integrating multi-level and multi-scale feature fusion strategies [31,32]. To facilitate the quantitative assessment of asphalt pavement crack defects, researchers have further introduced attention mechanisms into CNNs, effectively achieving precise localization of crack boundaries [33,34]. Additionally, Vision Transformer, a novel deep learning architecture, has provided a new perspective and direction for asphalt pavement crack detection through its applications in pavement defect classification [35] and semantic segmentation [36] while also reducing carbon emissions by 30.80% and maintenance costs by 20.81% [37]. Notably, the cornerstone to achieving such commendable outcomes in these studies is the employment of extensive training datasets. Through these data, deep learning models can autonomously learn and extract pivotal feature information from images, thereby facilitating the precise identification of defects, such as cracks.

However, in the field of road engineering practices, there are challenges in two aspects. On the one hand, obtaining large-scale pavement images often faces numerous constraints, and using limited datasets often leads to issues of model overfitting. On the other hand, relying solely on deep learning models may encounter problems such as overfitting, high computational costs, and insufficient generalization ability for specific tasks. Because of their excellent portability, safety, and efficiency, unmanned aerial vehicles (UAVs) have greatly promoted the progress and development of pavement defect detection, thus enhancing the construction and utilization of pavement datasets [38,39]. To effectively employ deep learning models under conditions of limited samples and fulfill detection requirements, researchers have adopted a transfer learning methodology. This involves transferring the learned weights from pre-trained deep learning models to fresh detection tasks and subsequently fine-tuning the parameters of the pertinent layers [40]. This strategy significantly improves the model's effectiveness in scenarios with limited data and attains commendable prediction outcomes. Furthermore, to overcome the challenge of limited sample sizes more proficiently, researchers have also concentrated on developing adaptive and lightweight CNN models [41,42]. These models not only maintain high prediction accuracy while minimizing computational expenses but also facilitate efficient pavement defect detection within constrained resources. Nevertheless, it is worth noting that the efficient operation of data-driven models often heavily relies on high-performance computing equipment, which inadvertently poses obstacles to further optimization of the models.

To overcome these challenges, this study, based on the dual constraints of limited sample size and computational resources, combines the powerful feature extraction capabilities of CNNs with the superior classification performance of SVMs in machine learning algorithms to propose a CNN-SVM model. The goal is to effectively improve the efficiency and accuracy of road defect detection through integrated and optimized model design without significantly increasing computational burden. Specifically, firstly, a vehicle-mounted line-scan industrial camera is used to capture images of pavement defects, which eliminates the need to close roads separately for data collection, thereby avoiding traffic congestion. Secondly, a hybrid learning model that combines CNNs and SVMs is proposed instead of adopting the time-consuming conventional manual analysis method. Finally, through algorithmic integration and innovation, efficient allocation and utilization of resources are achieved, thereby optimizing road management and maintenance processes and enhancing overall efficiency and service quality. This method not only helps overcome current technical bottlenecks but also provides a new perspective and practical path for advancing the development of intelligent road detection technology in resource-constrained environments.

This study aims to provide a more efficient and accurate solution for pavement crack identification through technological innovation while simultaneously promoting the development of intelligent transportation systems and smart cities and achieving sustainable development in road transportation. The main contributions of this study are as follows:

(1) By fine-tuning the ShuffleNet model, a novel hybrid method for identifying and classifying defects in pavement cracks is proposed, which overcomes the shortcomings of large computation in the deep learning model.
(2) The MRMR feature selection method is employed to select the most relevant features, which effectively overcomes the problem of insufficient detection performance of asphalt pavement defects with limited samples and computing resources.
(3) By exploring the role of intelligent detection technology in promoting sustainable road development, the green and low-carbon development of transportation infrastructure is promoted from the perspective of saving resources and improving efficiency.

The organizational structure of this study is as follows: In Section 2.1, the image dataset used in this study is introduced. The methods adopted in this study are presented in Sections 2.2–2.4. The implementation details of the DCNN are described in Section 3.1, including parameter settings, training procedures, etc. The experimental results and detailed analysis of the proposed method are presented in Section 3.2. Section 4 compares different categories of models and analyzes the performance of the proposed model in detail. Additionally, the promotional effects of pavement intelligent detection technology on sustainable road development were also analyzed. Finally, Section 5 summarizes the research findings and provides insights for future research directions.

2. Methodology

2.1. Dataset

Conventional or UAV-based image acquisition methods may lead to traffic congestion due to road closures, thereby increasing fuel consumption and carbon emissions. A multi-functional road inspection vehicle was utilized to systematically collect asphalt pavement image data from Lanhua Road in Nanjing, China, in this study. These images meticulously captured typical visible defects, including pavement cracks, as well as some irregularities like potholes. The resolution of all collected images remained consistent at 4096 × 2000 pixels.

To overcome the computational challenges inherent in processing high-resolution images with the existing computing resources, the sliding window technique was employed to divide the original images into 224 × 224-pixel patches to meet the input requirements of the pre-trained model, as illustrated in Figure 1. This strategy efficiently alleviates the computational burden while ensuring that the image data conforms to the specific input resolution specifications of the transfer learning model. Subsequently, road engineering experts classified these image patches into four distinct categories: background, cracks, potholes, and landmarks. The consolidated dataset, termed NKLHData, provides enhanced efficiency and accuracy in detecting pavement defects, thereby laying a robust foundation for future research works. To verify the reliability of the model, this study employed three public datasets for comparative experiments, including CRACK500 [43] and GAPs384 [44]. In all the datasets utilized, 80% of the data are allocated for model training, 10% for validation, and the remaining 10% are reserved for testing. Table 1 describes the distribution of pavement images in each classification in every dataset.

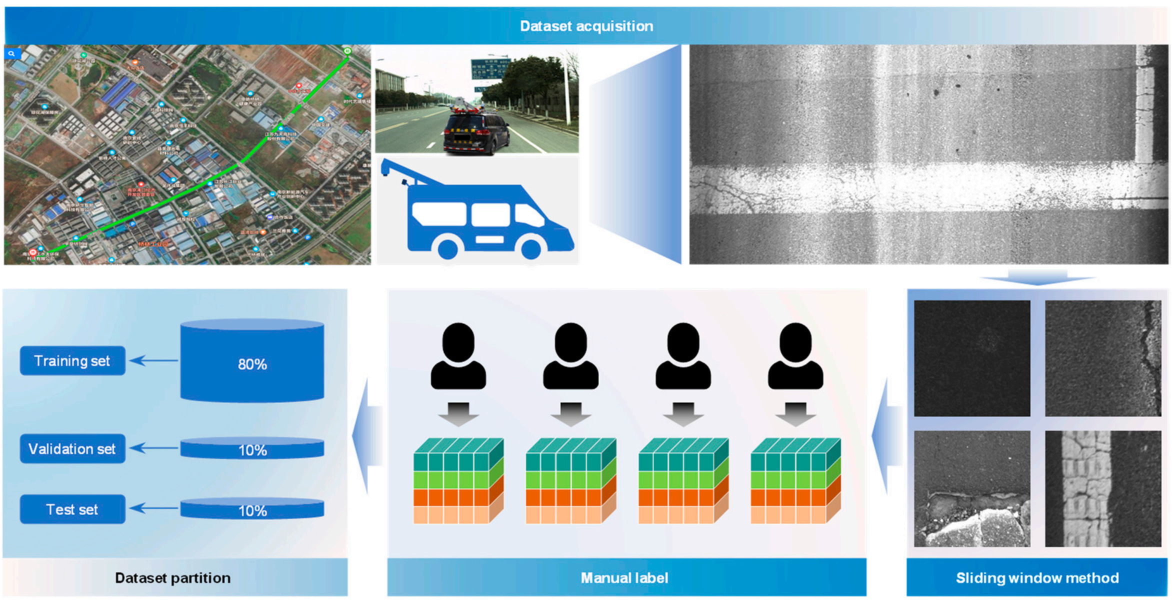

Figure 1. Flow of image data acquisition and processing.

Table 1. Distribution of different datasets employed in this study.

Datasets		Number of Images	Background	Crack	Landmark	Pothole
NKLHData	Training set	9272	2318	2318	2318	2318
	Validation set	1160	290	290	290	290
	Test set	1160	290	290	290	290
	All	11,592	2898	2898	2898	2898
Crack500	Training set	6462	3231	3231	/	/
	Validation set	806	403	403	/	/
	Test set	806	403	403	/	/
	All	8074	4037	4037	/	/
GAPs384	Training set	12,780	6390	6390	/	/
	Validation set	1598	799	799	/	/
	Test set	1598	799	799	/	/
	All	15,976	7988	7988	/	/

2.2. Pavement Defect Identification Based on Deep Learning

Considered a pivotal subset of deep learning, CNNs have garnered extensive application in various fields, attributed to their remarkable prowess in image recognition tasks. The objective of this section is to construct a streamlined CNN model grounded in established deep learning principles that is specifically tailored for training the NKLHData dataset.

The architecture of our proposed CNN model encompasses four key phases, each consisting of a convolutional layer, a batch normalization layer to stabilize the training process and mitigate gradient vanishing or exploding issues [24], and a nonlinear activation layer. To minimize the number of parameters and computational cost, a max pooling layer follows each phase, performing downsampling to decrease the feature map resolution. The comprehensive architecture of the proposed CNN network is illustrated in Figure 2. At the conclusion of the model, a global max pooling layer is utilized to extract significant features from the final convolutional layer. These features are then fed into a fully connected layer for enhanced information integration. Ultimately, the model computes and outputs the probability distribution across classes via a softmax layer.

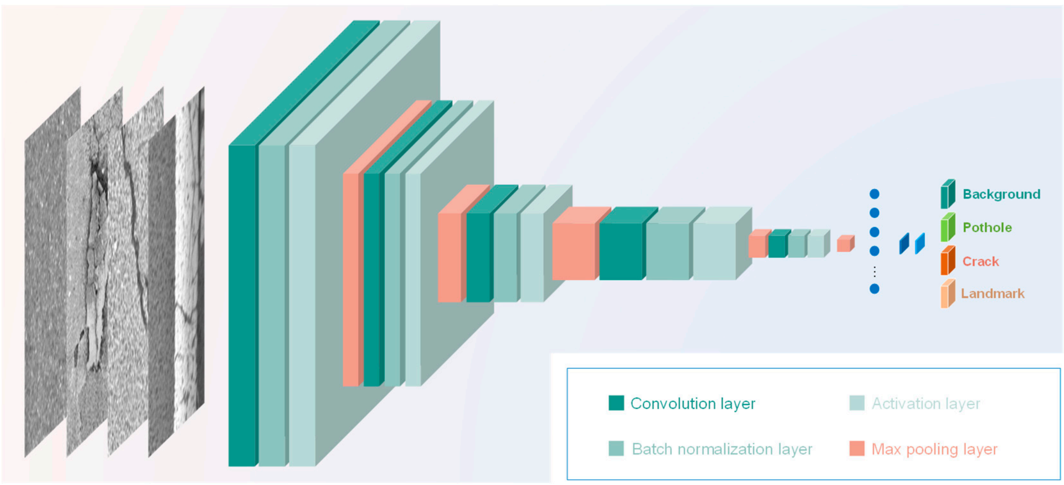

Figure 2. Proposed CNN network architecture.

In the course of the model training process, this study employs the cross-entropy loss function to quantify the discrepancy between the predicted outcomes and the actual labels. Additionally, to mitigate potential overfitting concerns, a dropout layer with a 10% dropout rate is incorporated into the model, thereby enhancing its generalization capabilities.

2.3. Pavement Defect Identification Based on Transfer Learning

Due to the scarcity of large-scale pavement defect image datasets for this study, training large models directly could result in overfitting challenges. Consequently, this study adopts transfer learning for pavement defect identification. As a highly effective machine learning approach, transfer learning harnesses the weights of pre-trained models to provide robust assistance for new tasks that suffer from insufficient data and limited computational resources, thereby enhancing the performance of the target task.

Within the field of transfer learning, numerous transfer strategies exist, encompassing feature-based, instance-based, parameter-based, and relation-based transfers [45]. Considering the unique scenario and data constraints in this study, a parameter-based transfer learning strategy was employed. Guided by this strategy, the ShuffleNet model [46] was chosen as the source for transfer learning, attributed to its minimal parameters and superior performance, making it an ideal candidate for this purpose. During the fine-tuning phase, the ShuffleNet model's entire architecture, encompassing convolutional layers, batch normalization layers, ReLU activation functions, and more, was preserved. The trained weights and biases of these layers were transferred to the new task. Furthermore, the parameters of these layers were frozen during training, with only select parameters undergoing further training.

Given the disparity in the number of sample categories between the NKLHData dataset and the pre-trained network, a new fully connected layer was introduced to accommodate the 4-class sample dataset, replacing the existing one. This novel layer employed the He weight initialization method [47] for initializing its weights, which is tailored for the ReLU activation function. This method can mitigate the issues of gradient vanishing and gradient explosion, thereby expediting the model training process. Figure 3 illustrates the fine-tuning process and network architecture based on the ShuffleNet model in detail. By fine-tuning the ShuffleNet model, this study aims to effectively identify pavement defects while minimizing the risk of overfitting.

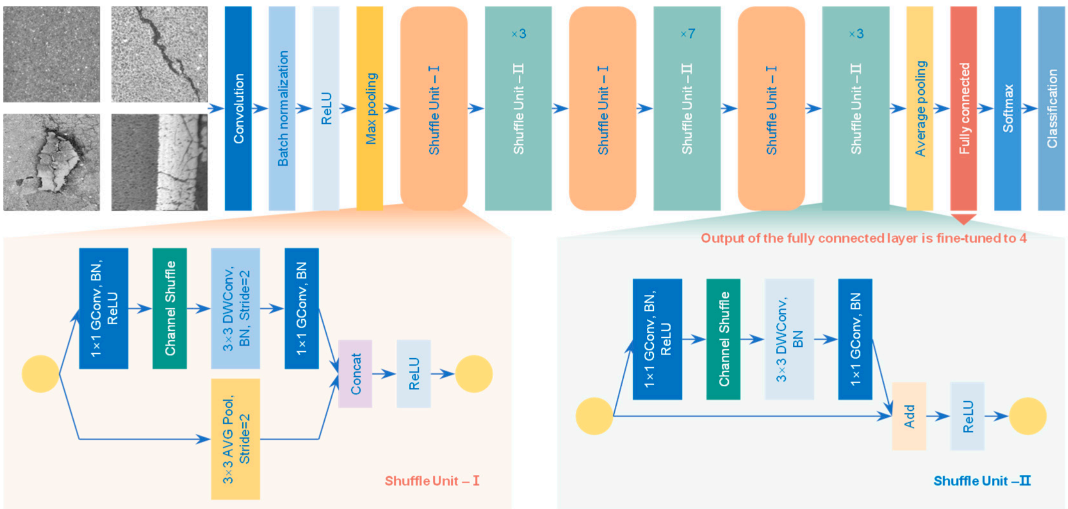

Figure 3. ShuffleNet network architecture of transfer learning.

2.4. Pavement Defect Identification Based on Hybrid Learning

The extensive hyperparameters in the fully connected layers of deep learning models often lead to overfitting. Conversely, conventional machine learning algorithms, such as SVMs, successfully address overfitting issues by incorporating regularization techniques. SVM stands as a prominent supervised learning algorithm in machine learning, renowned for its strong generalization capabilities and robust mathematical theoretical framework grounded in statistical learning theory and the principle of structural risk minimization. The core idea of SVM is to identify an optimal hyperplane within the feature space, serving as a decision boundary to accurately differentiate sample points belonging to different categories.

In contrast to deep learning models, the parameter tuning of SVM is more streamlined, minimizing the dependence on complex network architectures and extensive parameters, which enhances training efficiency. When managing small-scale datasets, SVM not only boasts a swift training duration but also demonstrates superior performance. Notably, the computational complexity of SVM primarily depends on the count of support vectors rather than the dimensionality of the data. This distinctive benefit renders SVM exceptionally remarkable in tackling high-dimensional data classification tasks. Furthermore, the classification outcomes furnished by SVM are highly understandable, distinctly outlining the boundaries separating various categories of data.

This study introduces the ShuffleNet-SVM model, aiming to synthesize the distinct strengths of CNN and SVM. This model capitalizes on CNN's robust feature extraction capabilities while incorporating SVM's edge in the interpretability of results. In this configuration, ShuffleNet's final three fully connected layers are omitted, and SVM is integrated to undertake the classification role. The network's architecture is illustrated in Figure 4. During implementation, the ShuffleNet model's fully connected layer is leveraged to derive 1000 deep features from NKLHData. Subsequently, the MRMR feature selection method [48] is applied to sieve out the top 25% of these features, serving as the classifier's input. These refined, high-level abstract features are then processed by the SVM classifier for accurate classification. To attain optimal prediction performance, this study employs SVM models equipped with linear, polynomial, and Gaussian kernels to classify NKLHData. A 5-fold cross-validation strategy is utilized to fine-tune the kernel and penalty parameters within the kernel functions.

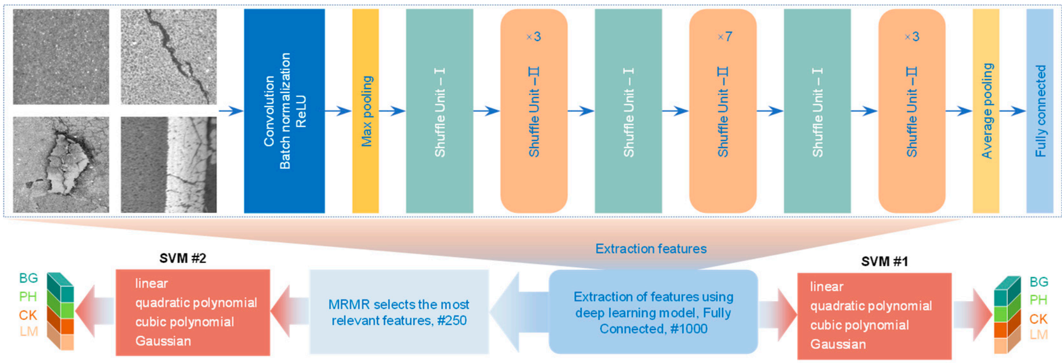

Figure 4. Network architecture of ShuffleNet-SVM model.

Maximum Relevance Minimum Redundancy

The Maximum Relevance Minimum Redundancy (MRMR) approach for feature selection represents a classical technique aimed at optimizing feature selection. Its objective is to boost the correlation between chosen features and the target variable while simultaneously decreasing the redundancy among these features. The ultimate goal is to construct an optimal subset S of features, which maximizes the relevance metric vs. and minimizes the redundancy metric *W*s. The formulas for computing relevance and redundancy are outlined in Equations (1) and (2), respectively. This strategy plays a crucial role in the data preprocessing and feature engineering stages, effectively enhancing model accuracy and efficiency.

$$V_s = \frac{1}{|S|}\sum\nolimits_{x\in S} I(X,Y) \quad (1)$$

$$W_s = \frac{1}{|S|^2}\sum\nolimits_{x,z\in S} I(X,Z) \quad (2)$$

where $|S|$ represents the count of features contained within the subset S. The significance of each feature's relevance and redundancy is assessed through their mutual information with the target variable. By determining the mutual information between each feature and the target and then prioritizing the features according to their importance, the most influential features are selected. The computation of mutual information follows Equation (3):

$$I(X,Y) = \sum_{i,j} P(X = x_i, Y = y_j) \log \frac{P(X = x_i, Y = y_j)}{P(X = x_i)P(Y = y_j)} \tag{3}$$

Within the MRMR feature selection method, the prioritization of the most relevant features is accomplished through the computation of the Mutual Information Quotient (*MIQ*), as detailed in Equation (4):

$$MIQ_x = \frac{I(X,Y)}{\frac{1}{|S|}\sum_{z \in S} I(X,Z)} \tag{4}$$

Utilizing this methodology, the present study endeavors to develop a hybrid learning model that is more streamlined, effective, and precise. This model demonstrates exceptional performance during subsequent experimental validations, thereby furnishing innovative concepts and approaches for the creation and refinement of hybrid learning models.

2.5. Performance Evaluation

To evaluate the predictive capability of the model, three metrics are employed: precision (p), recall (r), and F_1-score. Specifically, precision (p) quantifies the fraction of samples the model classifies as cracks that genuinely correspond to cracks. Recall (r), on the other hand, reflects the percentage of actual crack samples that are accurately identified by the model. The F_1-score provides a more comprehensive and balanced perspective by harmonizing precision and recall. These metrics follow Equation (5):

$$\begin{aligned} p &= \tfrac{TP}{TP+FP} \\ r &= \tfrac{TP}{TP+FN} \\ F_1 &= \tfrac{2 \times p \times r}{p+r} \end{aligned} \tag{5}$$

where *TP* (True Positive) represents the count of images accurately identified as cracks, *FN* (False Negative) represents the number of images mistakenly labeled as cracks despite being non-cracks, *TN* (True Negative) represents the correctly classified non-crack images, and *FP* (False Positive) represents the images incorrectly identified as non-cracks while actually being cracks. A 2 × 2 confusion matrix, as shown in Table 2, provides a visual depiction of these evaluation metrics.

Table 2. Confusion matrix of binary prediction.

		Prediction	
		Predicted as Positive	**Predicted as Negative**
True	positive	True positive (*TP*)	False negative (*FN*)
	negative	False positive (*FP*)	True negative (*TN*)

Furthermore, this study also introduces another evaluation metric, AUC (Area Under the Curve). AUC represents the area surrounded by the ROC curve and coordinate axis. The calculation of AUC is governed by Equation (6).

$$AUC = \frac{1}{2}\sum_{i=1}^{m-1} (x_{i+1} - x_i)(x_i + x_{i+1}) \tag{6}$$

In these curves, the horizontal axis represents the False Positive Rate (*FPR*), whereas the vertical axis represents the True Positive Rate (*TPR*). TPR quantifies the fraction of all actual positive samples accurately recognized as positive by the model, whereas FPR describes the percentage of all actual negative samples incorrectly labeled as positive.

3. Experimental Results

3.1. Implementation Details

The experiments were carried out utilizing the MATLAB 2023b platform (925512), specifically leveraging the Statistics and Machine Learning Toolbox and the Deep Learning Toolbox. The training of all models occurred on a workstation featuring an Intel Core i5-12600KF CPU, 32 GB RAM, and Windows 11 operating system. Additionally, an NVIDIA GeForce RTX 3060 GPU with 12 GB GDDR6 memory was employed to accelerate processing tasks. The training process was configured with a maximum of 30 epochs, with the minibatch size adjusted to 16 due to GPU memory constraints. The initial learning rate was established at 0.001, while the regularization coefficient was set at 0.003. To promote network convergence and mitigate overfitting [49], the learning rate was decreased to one-tenth of its initial value every 20 epochs. For model training, an SGD optimizer with a momentum factor of 0.9 was utilized [50]. After each training iteration, the validation set was assessed.

3.2. Model Training

This section delves into three exemplary implementations for identifying defects in asphalt pavements using innovative methods. Given the scarcity of NKLHData, the initial strategy in this study was to construct a basic CNN model grounded in the core principles of CNNs to tackle this issue. The second approach involved fine-tuning a ShuffleNet model pre-trained on ImageNet. Specifically, we tuned the number of outputs in the fully connected layer while keeping the weights and biases of the remaining layers frozen, employing a stabilization strategy to ensure the model's robustness and efficiency. The third approach, while also relying on the pre-trained ShuffleNet model for feature extraction, further integrated these features into the training and prediction processes of an SVM. This hybrid method merges the strengths of deep learning-based feature extraction with the efficacy of SVM classifiers, aiming to improve the overall effectiveness of defect identification. For this third method, this study employed the MRMR feature selection method to select 250 of the most relevant features and utilized various kernel functions of the SVM, including linear, quadratic, cubic, and Gaussian, to classify defects, ultimately striving for more precise and efficient identification of asphalt pavement defects.

3.2.1. Basic CNN Model

In Figure 5, a visualization of the training process of the basic CNN model utilizing the NKLHData dataset is provided. Specifically, Figure 5 illustrates the fluctuations in the model's accuracy throughout the training phase. As the training epochs accumulate, the accuracy curve demonstrates a consistent upward trend, signifying a notable enhancement in the model's recognition capabilities through relentless learning and refinement. Notably, during the early phases of training, the swift surge in accuracy underscores the model's proficient aptitude for adapting to and learning from the dataset's features.

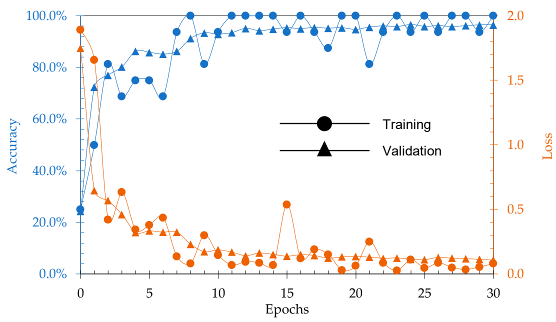

Figure 5. Training process of shallow CNN model on NKLHData dataset.

Additionally, Figure 5 offers an in-depth examination of how the loss function evolves over the course of the training epochs. Serving as a vital metric of model performance, the loss function directly mirrors the discrepancy between the model's predicted and actual values. Noticeably, as the number of iterations accumulates, the loss value exhibits a steady downward trend, indicating that the model is progressively minimizing prediction errors and improving prediction accuracy. Early in the training process, particularly within the initial 10 epochs, the loss function experiences a particularly steep decline, hinting at the model's swift acquisition and extraction of essential data features. However, as the training advances, the rate of reduction in the loss function progressively tapers off, indicating that the model is nearing its performance ceiling, and further performance gains through mere iterations become increasingly challenging.

Upon completion of adequate training, it was observed that by the 25th epoch, the model's classification accuracy on the training dataset had reached a stabilized high of 96.19%. This notable and consistent performance enhancement demonstrates the model's outstanding generalization capabilities and recognition accuracy while also confirming the rationality and efficacy of the hyperparameters (including the learning rate and batch size) employed during training. These meticulously chosen parameters guarantee that the model attains optimal performance despite resource constraints.

3.2.2. Fine-Tune ShuffleNet

Figure 6 delves into the dynamic fluctuations of accuracy and loss during the ShuffleNet model's training process, leveraging transfer learning. It is evident that during the initial stages, there is a substantial surge in model accuracy due to a sharp decline in loss. Specifically, within the first three epochs, the loss reduction is particularly prominent, highlighting the model's rapid adaptability and robust learning capability when confronted with new data. This rapid learning rate implies that the model efficiently captures key data features during the early stages, setting a firm foundation for subsequent learning. As the training epochs advance, the improvement in model performance gradually reaches a plateau. During this phase, the rate of loss reduction begins to decelerate, and the accuracy growth stabilizes. This shift suggests that the model is progressively nearing its performance ceiling. However, the model persists in refining its internal parameters to further enhance classification accuracy. After 20 epochs, it becomes apparent that the loss value exhibits minor fluctuations around 0.6. This stable fluctuation clearly signals that the training process has stabilized and the model's performance has attained a relatively optimal state. At this stage, the model has thoroughly grasped the features within the training data and is capable of delivering precise predictions on new, unseen data.

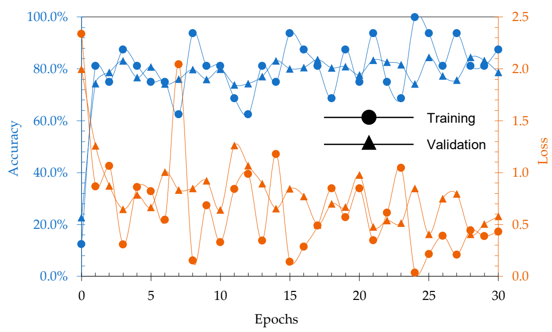

Figure 6. Dynamic variations of accuracy and loss of ShuffleNet model based on transfer learning.

After meticulously training for 30 epochs, the ShuffleNet model, leveraging transfer learning, ultimately attained an accuracy of 79.92%. This achievement not only marks the pinnacle of its classification performance but also fully demonstrates the feasibility of the transfer learning strategy and the ShuffleNet model in image classification tasks. By harnessing pre-trained model parameters and employing fine-tuning strategies, this study has successfully improved the classification accuracy of traditional machine learning methods, paving the way for extensive application scenarios of the model.

3.2.3. ShuffleNet-SVM

To address the complexity of linearly inseparable data, SVMs have proven effective in mapping the original space to a feature space by incorporating kernel functions. The objective of this study is to conduct a comprehensive assessment of the efficacy of six prevalent kernel functions in SVMs. These include the linear kernel, quadratic polynomial kernel, cubic polynomial kernel, and Gaussian kernels with varying scales, specifically, 8, 32, and 64. Figure 7 offers a visual representation of the confusion matrices of these kernel functions in practical applications, providing a quantitative foundation for assessing their effectiveness in resolving linearly inseparable data challenges. In this matrix, the predicted classification outcomes are plotted along the x-axis, whereas the true sample categories are represented on the y-axis. The diagonal elements signify the count of accurately classified samples for each category, while the off-diagonal elements depict the number and distribution of misclassified samples.

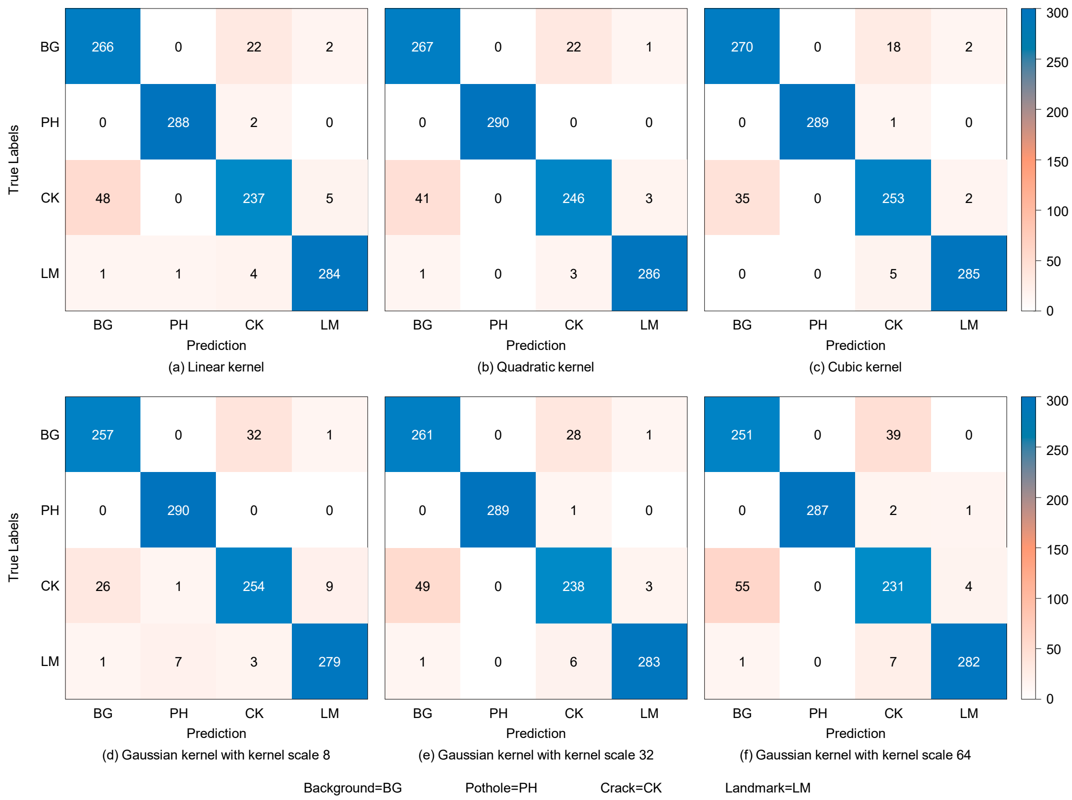

Figure 7. Confusion matrix of SVM model with six different kernel functions.

Upon meticulous comparison and analysis of the data presented in Table 3, it becomes evident that the model utilizing 1000 features demonstrates a subtle yet discernible superiority over the model with just 250 features across three pivotal performance metrics: precision, recall, and F_1-score. This finding emphasizes that expanding the feature set enables the model to encapsulate a broader spectrum of informational dimensions, thereby augmenting its overall predictive capability. However, it is noteworthy that the discrepancy in AUC—a metric assessing the overall classification effectiveness—between these two feature scales does not attain statistical significance. This revelation suggests that while augmenting the number of features may optimize the model's localized performance to a certain extent, it does not significantly substantially elevate its comprehensive classification capability on a global level.

Table 3. Comparison of SVM prediction results with different numbers of features.

Models	Kernal	P/%	R/%	F_1/%	AUC/%
SVM with 1000 features	Linear	92.82	92.67	92.74	98.81
	Quadratic	94.08	93.97	94.02	99.22
	Cubic	94.37	94.31	94.34	99.04
	Gaussian kernel with scale 8	77.96	75.78	76.85	96.31
	Gaussian kernel with scale 32	94.27	94.22	94.24	99.23
	Gaussian kernel with scale 64	92.87	92.76	92.81	98.89
SVM with 250 features	Linear	92.78	92.67	92.72	98.71
	Quadratic	93.95	93.88	93.91	99.15
	Cubic	94.62	94.57	94.59	99.22
	Gaussian kernel with scale 8	93.06	93.10	93.08	99.24
	Gaussian kernel with scale 32	92.43	92.33	92.38	98.83
	Gaussian kernel with scale 64	90.70	90.60	90.65	98.38

This study aims to systematically identify the most representative top 25% feature subset as the pivotal input for enhancing the model's performance and accuracy, utilizing the MRMR methodology. Through an exhaustive evaluation and comparison of SVM models employing various kernel function configurations, the study found that the SVM equipped with a cubic polynomial kernel and the SVM featuring a Gaussian kernel (specifically with a kernel scale of 8) demonstrated superior prediction performance. This superiority can be attributed to the polynomial kernel's robustness in capturing intricate nonlinear relationships within datasets, which is particularly crucial in road engineering. In this field, tackling challenging tasks like detailed classification and precise prediction of road defects significantly depends on the model's proficiency in analyzing complex data patterns. Consequently, the findings of this study not only validate the efficacy of the MRMR feature selection method but also emphasize the potential and significance of polynomial kernel functions in augmenting the performance of predictive models applied to road engineering issues.

It is worth highlighting that the inference speed of an SVM equipped with a Gaussian kernel in Table 4 is profoundly influenced by the configuration of the scale parameter, underscoring the necessity of meticulous parameter tuning during the model optimization phase. In stark contrast, the linear kernel SVM, characterized by its remarkably low memory footprint (a mere 0.26 MB), demonstrates exceptional versatility in environments with limited resources, rendering it particularly well-suited for real-time data processing tasks that necessitate swift processing speeds. While SVMs with polynomial and Gaussian kernels incur relatively higher memory usage (spanning from 3 MB to 13 MB), this remains manageable and within practical boundaries.

Table 4. Comparison of model prediction performance after feature selection.

Models	Kernal	Inference Speed/Obs/s	Model Memory/MB
SVM with 250 features	Linear	6000	0.26
	Quadratic	2700	3
	Cubic	2800	3
	Gaussian kernel with scale 8	500	13
	Gaussian kernel with scale 32	1900	3
	Gaussian kernel with scale 64	11,000	4

Considering the significant disparities in SVM model performance across various kernel configurations, meticulously selecting the appropriate kernel for specific applications is paramount. Notably, the linear kernel, distinguished by its swift inference speed and minimal resource consumption, stands as an optimal choice for real-time analysis tasks involving vast datasets. Conversely, the polynomial kernel is better aligned with applications demanding a thorough exploration of intricate nonlinear features within inherently complex relationships, albeit necessitating cautious evaluation of the balance between heightened computational expenses and slower inference rates. As for the Gaussian kernel, it offers superior modeling flexibility; however, optimizing model performance and computational efficiency hinges on the precise setting of the scale parameter.

In summary, this study adopts the MRMR method as a feature selection strategy, aiming to refine the quality of the model's input features. The pioneering ShuffleNet-SVM model cleverly integrates the automated feature learning capabilities of deep learning with the exceptional predictive accuracy of machine learning, substantially bolstering the model's robustness and generalization prowess in tackling complex data. This fusion not only adeptly addresses the challenges posed by redundancy and noise interference during the training process but also appreciably accelerates the model's inference speed for classification tasks.

4. Discussion and Analysis

4.1. Comparative Analysis of the Results of Three Methods

In this study, we contrast the precision of classification among three distinct models, with Table 5 offering a comprehensive breakdown of the findings. Notably, the SVM model with the Cubic kernel utilizing artificial features exhibits commendable forecasting accuracy; however, there remains considerable room for enhancement when benchmarked against three deep learning approaches: CNN, ShuffleNet, and ShuffleNet-SVM. Precisely, the SVM's predictive accuracy lags behind CNN, ShuffleNet, and ShuffleNet-SVM by 16.57%, 12.33%, and 17.68%, respectively. This comparative analysis underscores a significant conclusion: for the purpose of identifying defects in asphalt pavements, deep learning-driven models surpass standalone machine learning models in terms of efficacy.

Table 5. Comparison of classification accuracy of different models.

Model	P/%	R/%	F_1/%
SVM with artificial features (Cubic)	76.94	76.05	76.49
CNN	93.51	93.45	93.48
ShuffleNet	89.27	89.48	89.37
ShuffleNet-SVM	94.62	94.57	94.59

While parameter-based transfer learning tackles the CNN models' demand for extensive computational resources and extensive datasets during the training phase and efficiently reduces the likelihood of overfitting due to smaller datasets, the ShuffleNet model employing this strategy in our study failed to exhibit superior predictive accuracy on the test set compared to the CNN model; it actually experienced a decline of 4.24%. One plausible explanation could be the substantial divergence between ShuffleNet's pre-training task and the pavement crack classification task. Furthermore, our study revealed that the ShuffleNet model's predictions with transfer learning were 5.35% inferior to those of the ShuffleNet-SVM model. This disparity is likely attributed to SVM's ability to enhance the model by projecting input features into a high-dimensional space via kernel functions, thereby maximizing the classification margin and achieving impressive classification results on unseen test samples.

The findings of this study underscore the exceptional classification prowess of the ShuffleNet-SVM model in detecting asphalt pavement crack defects, with a remarkable accuracy rate of 94.62%. When juxtaposed against traditional CNN and ShuffleNet models, ShuffleNet-SVM demonstrates substantial performance gains of 1.11% and 5.35%, respectively. This notable superiority unmistakably confirms the efficacy of the ShuffleNet-SVM model in tackling such challenges. The exceptional performance of ShuffleNet-SVM is largely credited to its clever integration of SVM. By decreasing the parameter count in ShuffleNet's fully connected layers, ShuffleNet-SVM adeptly circumvents overfitting issues and drastically cuts down on training time and computational resource usage. Moreover, ShuffleNet-SVM fully embraces the regularization and structural risk minimization principles of SVM, thereby further bolstering the model's robustness and generalization abilities.

In summary, the ShuffleNet-SVM model not only capitalizes on SVM's strengths in classification endeavors but also attains substantial overall performance enhancements, furnishing an efficient and precise method for detecting asphalt pavement crack defects. To delve deeper into the model's performance across specific sample categories, Figure 8 presents the confusion matrices of the three models on the test set. It is worth highlighting that the confusion between cracks and backgrounds stands out prominently in the image classification process. This phenomenon is fundamentally rooted in the cracks' relatively minute features, which result in a considerable degree of similarity with the background. This pivotal insight not only reveals the current models' limitations in distinguishing subtle differences but also provides invaluable perspectives and directives for our ongoing model optimization and the enhancement of classification accuracy.

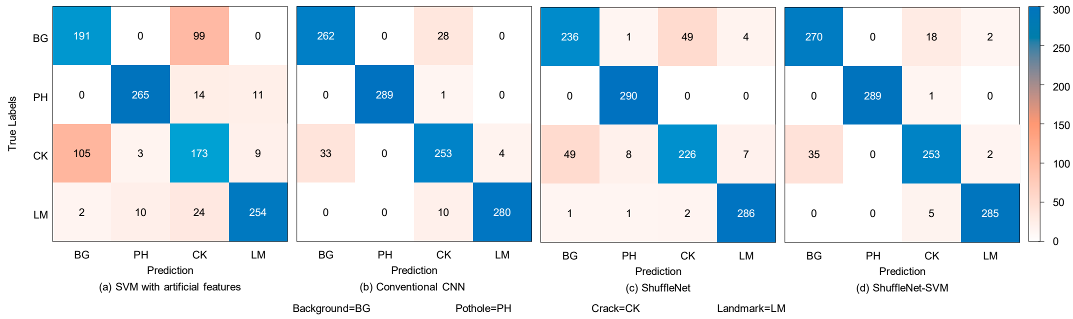

Figure 8. Confusion matrices of the three compared models.

Although the proposed hybrid model has demonstrated significant advantages in prediction performance, the specific mechanisms by which it perceives and processes input data remain unclear. Feature visualization serves as a potent analytical tool that not only fosters a deeper understanding of the model's prediction behavior but also dissects its decision-making logic and aids in identifying potential biases or shortcomings.

Figure 9 presents detailed visualization results of features extracted using two different methods, which profoundly reveal the intrinsic differences in the way the model processes information. Specifically, Figure 9b displays features obtained through manual feature extraction techniques, simplified into a single one-dimensional vector. While this method is intuitive and easy to understand, it may be constrained by human prior knowledge and the subjectivity of feature design, hindering its ability to fully capture the complexity and diversity of the data. In contrast, deep learning models exhibit superior feature extraction capabilities. Deep learning enables the gradual and layered extraction of features from input data, with a large number and diverse types of features that capture complex patterns and intricate structures within the data, as illustrated in Figure 9c–e. As the network layers deepen, the extracted features become increasingly abstract and high-level, encompassing not only the low-level details of the input data but also rich, deep semantic information.

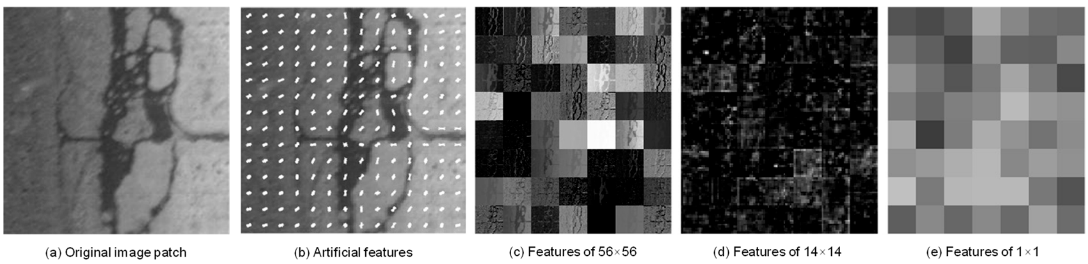

Figure 9. Visualization of feature maps extracted by different methods.

This layered extraction and progressive abstraction of features allow deep learning models to better adapt to complex and diverse data environments, ultimately resulting in significant improvements in prediction performance.

To systematically evaluate and compare the predictive performance of various models, this study employed the t-SNE [51], which aims to map high-dimensional feature maps into a low-dimensional space for visual analysis through effective dimensionality reduction. Figure 10 intuitively demonstrates the feature distribution overview in the low-dimensional space of 290 selected image samples from various categories in the test set.

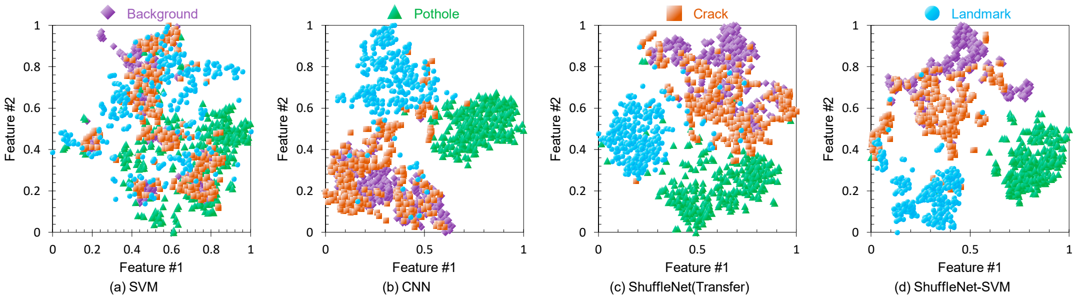

Figure 10. Feature scatter plots extracted by different feature extractors.

The comparative analysis demonstrates that the features automatically extracted based on deep learning, as shown in Figure 10b,c, exhibit higher linear separability compared to the traditional artificial features depicted in Figure 10a. This finding profoundly unveils the primary reason behind the limited classification performance in traditional methods, namely, their relatively insufficient feature representation capabilities. In contrast, features extracted through CNNs not only possess remarkable recognition abilities but also present clear decision boundaries, a characteristic that significantly simplifies classification tasks, enabling even simpler classifiers to achieve exceptional classification results. Thus, this study further substantiates the notable advantages of deep learning in feature extraction and classification tasks.

However, the training process of deep learning models is frequently constrained by the demand for vast amounts of data and the substantial consumption of computational resources, posing a significant challenge in practical applications. To address this issue, the meticulously designed model in this study demonstrates prominent advantages, cleverly overcoming the limitations of limited data and computational resources and providing an effective solution for the application of deep learning in resource-constrained environments.

4.2. Superiority Analysis of ShuffleNet-SVM

4.2.1. Compared with Mainstream Models

In this study, various classical recognition models were employed as comparative benchmarks to further validate the superiority of the proposed ShuffleNet-SVM model. These models are not only classics in the field of image classification; they also incorporate a myriad of enhancement strategies, thereby providing a solid foundation for a comprehensive evaluation of the ShuffleNet-SVM model's performance. To facilitate the training process, the initial weights of all comparative models were adopted from ImageNet-based pre-trained weights. Additionally, all compared models were implemented using the same experimental settings as the proposed ShuffleNet-SVM model.

(1) VGG19: This is a classic deep convolutional neural network architecture designed for image classification tasks. This network comprises 19 weighted layers, including 16 convolutional layers and 3 fully connected layers. It utilizes stacks of multiple 3×3 convolutional kernels to increase the network's depth, enabling it to extract rich features from images and achieve remarkable classification performance on large-scale datasets like ImageNet.

(2) ResNet50: This is a deep residual network architecture designed to address the issue of performance degradation that occurs during the training of deep neural networks. ResNet50 contains 50 convolutional layers and introduces shortcut connections, enabling the network to learn identity mappings and making it easier to optimize deep networks.

(3) Inceptionv3: This is a high-performing convolutional neural network architecture that utilizes a module called "Inception" to optimize the network structure and enhance performance. Inceptionv3 captures multi-scale features in images by parallelly utilizing convolutional kernels of different sizes and pooling operations, and it reduces the amount of computation and number of parameters through bottleneck layers.

(4) DarkNet19: This is a convolutional neural network architecture specifically designed for real-time object detection tasks. DarkNet19 achieves efficient computation speed and a small model size by utilizing a simple stack of convolutional layers without complex modules or connections.

(5) DenseNet201: This is an efficient deep convolutional neural network architecture that significantly reduces the number of parameters while enhancing feature propagation and reuse. DenseNet201 consists of 201 layers, where dense connections enable each layer to be directly connected to all preceding layers, thus achieving effective feature reuse.

(6) MobileNetv2: This is a lightweight convolutional neural network architecture specifically designed for mobile and embedded vision applications. MobileNetv2 introduces depthwise separable convolution blocks with linear bottlenecks and inverted residual connections, which significantly reduces the model's parameters and computations while maintaining high performance.

The comprehensive test results are presented in Table 6. The proposed model has demonstrated exceptional performance in the classification task on the NKLHData dataset, achieving a classification accuracy of 94.57%, precision of 94.62%, recall of 94.57%, and an F_1-score of 94.59%. In contrast, classical image classification models failed to exhibit a significant competitive advantage on the same dataset. Notably, among all the classification categories, the identification of crack images remains the most challenging task within the NKLHData dataset.

Table 6. Classification precision of various compared models.

Metrics \ Model		VGG19	ResNet50	InceptionV3	DarkNet19	DenseNet201	MobileNetv2	ShuffleNet-SVM
All categories	P/%	68.54	91.35	85.46	86.50	89.82	87.01	94.62
	R/%	53.69	90.78	85.26	85.52	88.36	87.16	94.57
	F_1/%	60.21	91.06	85.36	86.01	89.08	87.08	94.59
Background	P/%	95.12	91.77	74.92	70.18	72.56	78.45	88.52
	R/%	13.45	73.10	85.52	91.72	94.83	80.34	93.10
	F_1/%	23.57	81.38	79.87	79.52	82.21	79.38	90.75
Pothole	P/%	78.19	100.00	96.25	96.99	97.97	94.77	100.00
	R/%	95.17	100.00	97.24	100.00	100.00	100.00	99.66
	F_1/%	85.85	100.00	96.74	98.47	98.47	97.31	99.83
Crack	P/%	64.15	77.65	82.91	82.76	91.88	77.29	91.34
	R/%	11.72	91.03	66.90	57.93	62.41	72.76	87.24
	F_1/%	19.82	83.81	74.05	68.15	74.33	74.96	89.24
Landmark	P/%	36.71	95.99	87.75	96.06	96.88	97.54	98.62
	R/%	94.42	98.97	91.38	92.41	96.21	95.52	98.28
	F_1/%	52.87	97.46	89.53	94.20	96.54	96.52	98.45

Furthermore, through the analysis of Table 7, it becomes apparent that VGG19 stands out as the most intricate model, boasting a substantial number of parameters (139.5 M), whereas ShuffleNet-SVM underscores its strength in lightweight design with a minimal number of parameters (0.86 M). In terms of processing speed, ShuffleNet-SVM surpasses all other models significantly, attaining an impressive frame rate of 371 FPS, highlighting its exceptional capability for real-time processing. This underscores the suitability of lightweight models such as ShuffleNet-SVM for devices with limited resources, including embedded systems and mobile devices, as well as applications demanding high-speed processing, such as real-time detection requirements.

Table 7. Differences in prediction performance of various models.

Model	Parameters	Frame Per Second (FPS)
VGG19	139.5 M	75
ResNet50	23.5 M	164
InceptionV3	20.80 M	162
DarkNet19	19.8 M	266
DenseNet201	18.1 M	40
MobileNetv2	2.2 M	224
ShuffleNet-SVM	0.86 M	371

4.2.2. Compared with Various Datasets

Table 8 presents a detailed overview of the predictive performance of the proposed model across four datasets: NKLHData, Crack500, GAPs384, and METU [52]. It systematically quantifies this performance through key metrics, including precision, recall, and F_1-score. The analysis results indicate that the model's performance varies significantly across different datasets, and within the same dataset, there is also an imbalance in predictive performance among different categories. Specifically, the model achieves precision, recall, and F_1-scores extremely close to 100% on the METU dataset, demonstrating exceptional performance. In contrast, while its performance on the Crack500 and GAPs384 datasets is relatively lower, it is still maintained above 90%. Furthermore, a comparative analysis of the NKLHData dataset under different brightness conditions reveals that the proposed model exhibits similar performance, suggesting that the CNN-SVM model has robust adaptability to images captured under varying lighting conditions. Further exploration indicates that these performance differences and imbalances may be attributed to the complex interplay of various underlying factors. Firstly, the quality of the datasets is a crucial factor directly related to the model's predictive performance. High-quality datasets often significantly enhance the model's prediction accuracy. Secondly, the imbalance in category distribution may also be a key factor limiting the model's performance improvement in certain categories.

Table 8. Prediction results based on various datasets.

Metrics	Datasets	NKLHData (↓50%)	NKLHData	NKLHData (↑100%)	Crack500	GAPs384	METU
All categories	P/%	93.28	94.62	94.04	90.80	90.16	99.83
	R/%	93.10	94.57	93.97	90.57	90.00	99.83
	F_1/%	93.19	94.59	94.00	90.68	90.08	99.83
Background	P/%	84.06	88.52	87.06	87.76	87.65	99.90
	R/%	92.76	93.10	92.76	94.29	93.13	99.75
	F_1/%	88.20	90.75	89.82	90.91	90.30	99.82
Crack	P/%	90.77	91.34	90.84	93.83	92.67	99.75
	R/%	81.38	87.24	85.52	86.65	86.88	99.90
	F_1/%	85.52	89.24	88.10	90.21	89.68	99.83
Pothole	P/%	100.00	100.00	100.00	/	/	/
	R/%	99.31	99.66	99.31	/	/	/
	F_1/%	99.65	99.83	99.65	/	/	/
Landmark	P/%	98.29	98.62	98.28	/	/	/
	R/%	98.97	98.28	98.28	/	/	/
	F_1/%	98.63	98.45	98.28	/	/	/

Note: ↓50% means that the brightness of the data set is reduced by 50%, and ↑100% means that the brightness of the data set is doubled.

4.3. Advantages of Intelligent Detection Technology

Road intelligent inspection technology, as a pivotal technological advancement in modern traffic engineering, has profoundly impacted the sustainable development of highway engineering. By merging the Internet of Things, extensive data analytics, and advanced artificial intelligence algorithms, this technology facilitates efficient and precise monitoring of highway statuses. It not only markedly augments detection accuracy but also effectively minimizes human errors, thus furnishing more dependable data support for highway maintenance work.

In the field of highway engineering inspection and maintenance, road intelligent inspection technology demonstrates numerous advantages over traditional inspection methods. From an economic perspective, traditional inspection techniques are costly due to inefficient use of materials and labor, with frequent material waste and construction delays further exacerbating the economic burden [53]. In contrast, intelligent inspection technology, through efficient data collection and analysis, enables the precise localization of damaged areas, facilitating accurate material usage and optimal labor allocation [54]. This drastically reduces maintenance costs and enhances economic efficiency.

On the societal front, the prolonged implementation of traditional inspection techniques often leads to extended construction periods, disrupting community daily life, increasing traffic accident risks, and lowering residents' quality of life [55]. Road intelligent inspection technology, with its rapid response and precise localization capabilities, effectively shortens construction durations and minimizes traffic disruptions, thereby mitigating negative societal impacts and enhancing public satisfaction and quality of life [56].

Regarding environmental protection, traditional inspection techniques consume substantial natural resources and generate significant waste, exacerbating environmental pollution [57]. Road intelligent inspection technology, however, reduces resource consumption and waste production by minimizing unnecessary material use and construction activities [58]. Additionally, it shortens construction durations and mitigates additional vehicle emissions due to traffic disruptions, demonstrating notable environmental protection advantages.

Moreover, road intelligent detection technology can collect and examine real-time information, including road statuses, traffic volumes, and weather conditions, and swiftly issue advance warnings for potential hazards like traffic collisions and congestion. This assists traffic management agencies in prompt responses, safeguarding highway safety, maintaining smooth traffic flow, and minimizing accident frequencies. Furthermore, data obtained from intelligent detection technology can be utilized to develop models for assessing highway technical conditions that predict the residual service life of highways and their maintenance requirements. This enables the refinement of maintenance management strategies and the development of scientific and logical maintenance schedules, which, in turn, mitigate damage resulting from excessive use or inadequate maintenance, thereby extending the service life of highways.

In summary, road intelligent inspection technology exhibits more positive impacts in terms of saving economic costs, improving societal quality of life, and protecting the environment. With continuous technological advancements and deeper application promotion, road intelligent inspection technology is poised to inject new vitality and momentum into the sustainable development of the road maintenance sector, driving the highway transportation industry towards a smarter and more sustainable direction.

5. Conclusions

To address the issues posed by the high computational resource consumption requirements and large data volumes required for deep learning-based pavement defect recognition models in sustainable roads, this study proposes a hybrid methodology that integrates a convolutional neural network (ShuffleNet) with an SVM for automatic classification of a self-constructed, limited pavement defect image dataset. Firstly, features are extracted from the fully connected layer of a pre-trained ShuffleNet model. Secondly, the

MRMR method is employed to select the 250 most relevant features from the extracted 1000 features, which serve as the dataset's characteristics. Then, SVM classifiers utilizing various kernel functions are employed to classify these features and compare them with current classical image classification models for analysis. Finally, the significant advantages of intelligent detection technology for sustainable road development are analyzed from three perspectives. The findings of this study support the following conclusions:

(1) After optimizing the proposed hybrid model with the MRMR feature selection method, it achieved a high classification accuracy of 94.62% on a self-constructed asphalt pavement image dataset and exhibited robustness on public datasets.
(2) By effectively mitigating traffic delays and reducing additional vehicle emissions resulting from untimely pavement maintenance or incorrect detections, this technology is of significant importance in alleviating environmental pressures and advancing sustainable development in transportation and the environment.
(3) The widespread promotion and application of this technology will further elevate the level of intelligence and sophistication in pavement maintenance management, positioning it as a pivotal strategy for driving sustainable development in highway engineering.

This study aims to explore a novel method for pavement defect identification suitable for limited datasets and computational resources. According to the experimental findings, the recognition of cracks in asphalt pavements remains a significant challenge, primarily due to the inherent variability in pavement crack defect recognition across diverse regions, road classifications, and climatic conditions. However, methodologies such as transfer learning and domain adaptation can empower models to rapidly adapt to various contexts while sustaining high levels of performance. These technologies pave new pathways for sustainable technologies in road maintenance by reducing carbon emissions during the production of repair materials and lowering resource and energy consumption, embodying the concept of sustainable development. Therefore, future research will focus on further enhancing the robustness of sustainable road detection technologies and conducting in-depth quantitative assessments of the practical contributions of smart detection technologies in advancing the sustainable development of highway engineering based on specific maintenance schemes. The goal is to provide more solid support for theoretical research and practical applications in related fields.

Author Contributions: The authors confirm contributions to the paper as follows: J.L. contributed to conceptualization, resources, validation, methodology, software, writing—original draft; Q.Z. contributed to investigation, validation, writing—review and editing, and project administration; X.G. contributed to supervision, validation, software, and writing—original draft. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the project funded by China Postdoctoral Science Foundation, grant number 2023M731369.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data will be made available on request.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Hu, Y.; Sreeram, A.; Xia, W.; Wang, H.; Zhou, L.; Si, W.; Airey, G.D. Use of Hansen Solubility Parameters (HSP) in the selection of highly effective rejuvenators for aged bitumen. *Road Mater. Pavement Des.* **2024**, 1–19. [CrossRef]
2. Hu, Y.; Ryan, J.; Sreeram, A.; Allanson, M.; Pasandín, A.R.; Zhou, L.; Singh, B.; Wang, H.; Airey, G.D. Optimising the dosage of bio-rejuvenators in asphalt recycling: A rejuvenation index-based approach. *Constr. Build. Mater.* **2024**, *433*, 136761. [CrossRef]
3. Ozer, H.; Al-Qadi, I.L.; Singhvi, P.; Bausano, J.; Carvalho, R.; Li, X.; Gibson, N. Prediction of pavement fatigue cracking at an accelerated testing section using asphalt mixture performance tests. *Int. J. Pavement Eng.* **2018**, *19*, 264–278. [CrossRef]
4. Vaitkus, A.; Čygas, D.; Motiejūnas, A.; Pakalnis, A.; Miškinis, D. Improvement of Road Pavement Maintenance Models and Technologies. *Balt. J. Road Bridge Eng.* **2016**, *11*, 242–249. [CrossRef]

5. Jiang, C. A Crack Detection and Diagnosis Methodology for Automated Pavement Condition Evaluation. Ph.D. Thesis, Georgia Institute of Technology, Atlanta, GA, USA, 2016.
6. Peng, L.; Chao, W.; Shuangmiao, L.; Baocai, F. Research on Crack Detection Method of Airport Runway Based on Twice-Threshold Segmentation. In Proceedings of the 2015 Fifth International Conference on Instrumentation and Measurement, Computer, Communication and Control (IMCCC), Qinhuangdao, China, 18–20 September 2015; pp. 1716–1720. [CrossRef]
7. Dekker, R.; Plasmeijer, R.P.; Swart, J.H. Evaluation of a new maintenance concept for the preservation of highways. *IMA J. Manag. Math.* **1998**, *9*, 109–156. [CrossRef]
8. Chen, D.; Montano, V.; Huo, L.; Song, G. Depth detection of subsurface voids in concrete-filled steel tubular (CFST) structure using percussion and decision tree. *Measurement* **2020**, *163*, 107869. [CrossRef]
9. Shi, Y.; Cui, L.; Qi, Z.; Meng, F.; Chen, Z. Automatic Road Crack Detection Using Random Structured Forests. *IEEE Trans. Intell. Transp. Syst.* **2016**, *17*, 3434–3445. [CrossRef]
10. Hoang, N.-D. An Artificial Intelligence Method for Asphalt Pavement Pothole Detection Using Least Squares Support Vector Machine and Neural Network with Steerable Filter-Based Feature Extraction. *Adv. Civ. Eng.* **2018**, *2018*, 7419058. [CrossRef]
11. Fang, F.; Li, L.; Gu, Y.; Zhu, H.; Lim, J.-H. A novel hybrid approach for crack detection. *Pattern Recognit.* **2020**, *107*, 107474. [CrossRef]
12. Mei, Q.; Gül, M.; Azim, M.R. Densely connected deep neural network considering connectivity of pixels for automatic crack detection. *Autom. Constr.* **2020**, *110*, 103018. [CrossRef]
13. Yang, X.; Li, H.; Yu, Y.; Luo, X.; Huang, T.; Yang, X. Automatic Pixel-Level Crack Detection and Measurement Using Fully Convolutional Network. *Comput. Aided Civ. Infrastruct. Eng.* **2018**, *33*, 1090–1109. [CrossRef]
14. Dorafshan, S.; Thomas, R.J.; Maguire, M. Comparison of deep convolutional neural networks and edge detectors for image-based crack detection in concrete. *Constr. Build. Mater.* **2018**, *186*, 1031–1045. [CrossRef]
15. Jung, D.; Lee, J.; Baek, C.; An, D.; Yang, S. Predicting Concrete Pavement Condition for Sustainable Management: Unveiling the Development of Distresses through Machine Learning. *Sustainability* **2024**, *16*, 573. [CrossRef]
16. Li, H.; Nyirandayisabye, R.; Dong, Q.; Niyirora, R.; Hakuzweyezu, T.; Zardari, I.A.; Nkinahamira, F. Crack damage prediction of asphalt pavement based on tire noise: A comparison of machine learning algorithms. *Constr. Build. Mater.* **2024**, *414*, 134867. [CrossRef]
17. Kaseko, M.S.; Lo, Z.P.; Ritchie, S.G. Comparison of Traditional and Neural Classifiers for Pavement-Crack Detection. *J. Transp. Eng.* **1994**, *120*, 552–569. [CrossRef]
18. Hoang, N.-D.; Nguyen, Q.-L.; Bui, D.T. Image Processing–Based Classification of Asphalt Pavement Cracks Using Support Vector Machine Optimized by Artificial Bee Colony. *J. Comput. Civ. Eng.* **2018**, *32*, 04018037. [CrossRef]
19. Wang, S.; Qiu, S.; Wang, W.; Xiao, D.; Kelvin, C.P.W. Cracking Classification Using Minimum Rectangular Cover–Based Support Vector Machine. *J. Comput. Civ. Eng.* **2017**, *31*, 04017027. [CrossRef]
20. Mei, Q.; Gül, M. A cost effective solution for pavement crack inspection using cameras and deep neural networks. *Constr. Build. Mater.* **2020**, *256*, 119397. [CrossRef]
21. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [CrossRef]
22. Szegedy, C.; Wei, L.; Yangqing, J.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9. [CrossRef]
23. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
24. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 30 June 2016; pp. 770–778. [CrossRef]
25. Huang, G.; Liu, Z.; Laurens, V.D.M.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708. [CrossRef]
26. Deng, J.; Dong, W.; Socher, R.; Li, L.; Kai, L.; Li, F.-F. ImageNet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255. [CrossRef]
27. Zhang, L.; Yang, F.; Zhang, Y.D.; Zhu, Y.J. Road crack detection using deep convolutional neural network. In Proceedings of the 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; pp. 3708–3712. [CrossRef]
28. Cha, Y.J.; Choi, W.; Büyüköztürk, O. Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks. *Comput. Aided Civ. Infrastruct. Eng.* **2017**, *32*, 361–378. [CrossRef]
29. Pauly, L.; Peel, H.; Luo, S.; Hogg, D.; Fuentes, R. Deeper networks for pavement crack detection. In Proceedings of the ISARC 2017—The 34th International Symposium on Automation and Robotics in Construction, Taipei, Taiwan, 28 June–1 July 2017; pp. 479–485. [CrossRef]
30. 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]
31. Xu, Y.; Bao, Y.; Chen, J.; Zuo, W.; Li, H. Surface fatigue crack identification in steel box girder of bridges by a deep fusion convolutional neural network based on consumer-grade camera images. *Struct. Health Monit.* **2018**, *18*, 653–674. [CrossRef]

32. Wang, Q.; Gao, T.; He, Q.; Liu, Y.; Wu, J.; Wang, P. Severe rail wear detection with rail running band images. *Comput. Aided Civ. Infrastruct. Eng.* **2023**, *38*, 1162–1180. [CrossRef]
33. Liang, J.; Zhang, Q.; Gu, X. Lightweight convolutional neural network driven by small data for asphalt pavement crack segmentation. *Autom. Constr.* **2024**, *158*, 105214. [CrossRef]
34. Liang, J.; Gu, X.; Jiang, D.; Zhang, Q. CNN-based network with multi-scale context feature and attention mechanism for automatic pavement crack segmentation. *Autom. Constr.* **2024**, *164*, 105482. [CrossRef]
35. Chen, Y.H.; Gu, X.Y.; Liu, Z.; Liang, J. A Fast Inference Vision Transformer for Automatic Pavement Image Classification and Its Visual Interpretation Method. *Remote Sens.* **2022**, *14*, 1877. [CrossRef]
36. Guo, F.; Liu, J.; Lv, C.; Yu, H. A novel transformer-based network with attention mechanism for automatic pavement crack detection. *Constr. Build. Mater.* **2023**, *391*, 131852. [CrossRef]
37. Chong, D.; Liao, P.; Fu, W. Multi-Objective Optimization for Sustainable Pavement Maintenance Decision Making by Integrating Pavement Image Segmentation and TOPSIS Methods. *Sustainability* **2024**, *16*, 1257. [CrossRef]
38. Shan, J.; Jiang, W.; Huang, Y.; Yuan, D.; Liu, Y. Unmanned Aerial Vehicle (UAV)-Based Pavement Image Stitching Without Occlusion, Crack Semantic Segmentation, and Quantification. *IEEE Trans. Intell. Transp. Syst.* **2024**, *25*, 1–16. [CrossRef]
39. Pranav, R.T.P.; Puppala, H.; Kim, B. Pavement Monitoring Using Unmanned Aerial Vehicles: An Overview. *J. Transp. Eng. Part B Pavements* **2023**, *149*, 03123002. [CrossRef]
40. Gao, Y.; Mosalam, K.M. Deep Transfer Learning for Image-Based Structural Damage Recognition. *Comput. Aided Civ. Infrastruct. Eng.* **2018**, *33*, 748–768. [CrossRef]
41. Hou, Y.; Li, Q.; Han, Q.; Peng, B.; Wang, L.; Gu, X.; Wang, D. MobileCrack: Object Classification in Asphalt Pavements Using an Adaptive Lightweight Deep Learning. *J. Transp. Eng. Part B Pavements* **2021**, *147*, 04020092. [CrossRef]
42. Liang, J.; Zhang, Q.; Gu, X. Small-sample data-driven lightweight convolutional neural network for asphalt pavement defect identification. *Case Stud. Constr. Mater.* **2024**, *21*, e03643. [CrossRef]
43. Yang, F.; Zhang, L.; Yu, S.; Prokhorov, D.; Mei, X.; Ling, H. Feature Pyramid and Hierarchical Boosting Network for Pavement Crack Detection. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 1525–1535. [CrossRef]
44. Eisenbach, M.; Stricker, R.; Seichter, D.; Amende, K.; Debes, K.; Sesselmann, M.; Ebersbach, D.; Stoeckert, U.; Gross, H.M. How to get pavement distress detection ready for deep learning? A systematic approach. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; pp. 2039–2047. [CrossRef]
45. Atasever, S.; Azginoglu, N.U.H.; Terzi, D.S.; Terzi, R. A comprehensive survey of deep learning research on medical image analysis with focus on transfer learning. *Clin. Imaging* **2023**, *94*, 18–41. [CrossRef]
46. Zhang, X.; Zhou, X.; Lin, M.; Sun, J. ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. *arXiv* **2017**, arXiv:1707.01083. [CrossRef]
47. He, K.; Zhang, X.; Ren, S.; Sun, J. Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification. *arXiv* **2015**, arXiv:1502.01852. [CrossRef]
48. Ding, C.; Peng, H. Minimum redundancy feature selection from microarray gene expression data. In Proceedings of the 2003 IEEE Bioinformatics Conference, CSB2003, Stanford, CA, USA, 11–14 August 2003; Computational Systems Bioinformatics. pp. 523–528. [CrossRef]
49. Alipour, M.; Harris, D.K. Increasing the robustness of material-specific deep learning models for crack detection across different materials. *Eng. Struct.* **2020**, *206*, 110157. [CrossRef]
50. Murphy, K.P. *Machine Learning: A Probabilistic Perspective*; The MIT Press: Cambridge, MA, USA, 2012; ISBN 978-0-262-01802-9.
51. Laurens, V.D.M.; Hinton, G. Visualizing Data using t-SNE. *J. Mach. Learn. Res.* **2008**, *9*, 2579–2605.
52. Sorguç, A.G. Performance Comparison of Pretrained Convolutional Neural Networks on Crack Detection in Buildings. In Proceedings of the 35th International Symposium on Automation and Robotics in Construction (ISARC), Berlin, Germany, 20–25 July 2018; pp. 693–700. [CrossRef]
53. Polder, R.B.; Peelen, W.H.A.; Courage, W.M.G. Non-traditional assessment and maintenance methods for aging concrete structures—Technical and non-technical issues. *Mater. Corros.* **2012**, *63*, 1147–1153. [CrossRef]
54. Gan, T.; Dong, S.; Wang, S.; Li, J. Distributed Resource Allocation in Dispersed Computing Environment Based on UAV Track Inspection in Urban Rail Transit. *Comput. Mater. Contin.* **2024**, *80*, 643–660. [CrossRef]
55. Retallack, A.E.; Ostendorf, B. Current Understanding of the Effects of Congestion on Traffic Accidents. *Int. J. Environ. Res. Public Health* **2019**, *16*, 3400. [CrossRef] [PubMed]
56. Gangwani, D.; Gangwani, P. Applications of Machine Learning and Artificial Intelligence in Intelligent Transportation System: A Review. In *Applications of Artificial Intelligence and Machine Learning*; Choudhary, A., Agrawal, A.P., Logeswaran, R., Eds.; Springer: Berlin/Heidelberg, Germany, 2021; pp. 203–216. [CrossRef]

57. Grădinaru, C.M.; Muntean, R.; Șerbănoiu, A.A.; Ciocan, V.; Burlacu, A. Sustainable Development of Human Society in Terms of Natural Depleting Resources Preservation Using Natural Renewable Raw Materials in a Novel Ecological Material Production. *Sustainability* **2020**, *12*, 2651. [CrossRef]
58. Cho, J.; Kang, J.; Song, Y.; Lee, S.; Yeon, J. Innovative Imaging and Analysis Techniques for Quantifying Spalling Repair Materials in Concrete Pavements. *Sustainability* **2024**, *16*, 112. [CrossRef]