Article

Exploring the Efficacy of Sparse Feature in Pavement Distress Image Classification: A Focus on Pavement-Specific Knowledge

Ye Yuan, Jiang Chen, Hong Lang * and Jian (John) Lu

The Key Laboratory of Road and Traffic Engineering, Ministry of Education, Tongji University, Shanghai 201080, China; yey@ieee.org (Y.Y.); jiangchen@tongji.edu.cn (J.C.)
* Correspondence: honglang@tongji.edu.cn

Abstract: Road surface deterioration, such as cracks and potholes, poses a significant threat to both road safety and infrastructure longevity. Swift and accurate detection of these issues is crucial for timely maintenance and user security. However, current techniques often overlook the unique characteristics of pavement images, where the small distressed areas are vastly outnumbered by the background. In response, we propose an innovative road distress classification model that capitalizes on sparse perception. Our method introduces a sparse feature extraction module using dilated convolution, tailored to capture and combine sparse features of different scales from the image. To further enhance our model, we design a specialized loss function rooted in domain-specific knowledge about pavement distress. This loss function enforces sparsity during feature extraction, guiding the model to align precisely with the sparse distribution of target features. We validate the strength and effectiveness of our model through comprehensive evaluations of a diverse dataset of road images containing various distress types and conditions. Our approach exhibits significant potential in advancing traffic safety by enabling more efficient and accurate detection and classification of road distress.

Keywords: machine learning; pavement distress; intelligent transportation systems

1. Introduction

Road infrastructure is pivotal for the socio-economic development and general prosperity of societies, acting as the backbone of connectivity and transportation. Maintaining the integrity and longevity of these road networks is imperative for transportation authorities globally. A major impediment faced by these bodies is the swift detection and subsequent classification of road anomalies like potholes, cracks, and other deformities. Such irregularities, if left unchecked, can result in accidents, amplify vehicular wear and tear, and escalate road maintenance expenditures.

Historically, the detection of such road maladies predominantly depended upon human inspection, a method that, despite its earnest intent, is laborious, time-intensive, and susceptible to oversights. The evolution of camera and laser technologies [1] has ushered in a new era, allowing for the capture of road images in high resolution, even while in transit. This has resulted in the accumulation of expansive datasets of road imagery. However, the true challenge surfaces in the meticulous identification and categorization of road distress within these extensive datasets [2].

Addressing the intricacies of pavement management, Jooste et al. architected a model harnessing the prowess of multi-classification machine learning algorithms, aiming for the judicious application of treatments to distinct road segments. This model showcases formidable accuracy, predicting both untreated and treated segments with high precision [3].

Numerous innovators have further enhanced this domain. Hanandeh introduced a novel intersection of mathematical modeling with Genetic Algorithm and Artificial Neural Networks, aspiring to deduce the Pavement Quality Index of flexible pavements [4]. Zhao

and collaborators ventured into machine learning, aiming to gauge the repercussions of excessive traffic weight on asphalt pavement's life, employing a random survival forest algorithm [5]. In yet another groundbreaking endeavor, Kaloop et al. presented a hybrid model, demonstrating superior accuracy in estimating the International Roughness Index of rigid pavements compared to traditional techniques [6].

In this expansive realm, researchers like Qureshi et al. probed the potential of intelligent image analysis techniques for pavement assessment, elucidating that these avant-garde systems could substantially economize both cost and time [7]. Dong and his team showcased a fusion LSTM-BPNN model that surpassed its predecessors in forecasting pavement performance [8]. Moreover, a slew of reviews and studies from researchers like Yang et al. and Damirchilo et al. spotlighted the myriad applications and predictive capacities of machine learning and artificial intelligence within pavement engineering [9–11].

Despite these advancements, road distress identification, with its intricate nuances, often remains elusive to conventional image classification models. A pivotal reason is the sparse distribution of pixels that characterize road distress juxtaposed against a predominant background—creating a stark imbalance and a unique challenge in image analysis.

To counter these challenges, this paper details a novel model rooted in sparse perception for road distress classification. The model's architecture rests on two foundational pillars: a sparse feature extraction module and a domain-specific loss function. The sparse feature extraction module, leveraging the potential of dilated convolution, is adept at capturing sparse features across various scales, creating a comprehensive representation of road conditions. Meanwhile, our loss function, conceived with a meticulous understanding of the pavement domain, ensures the consistent sparsity of the feature space.

Rigorous validation on an authentic dataset affirmed the potency of our model, recording remarkable $F1$ scores and signifying a profound edge over conventional methods.

In encapsulation, our contributions are as follows:

- We design a pioneering model emphasizing sparse perception tailored for road distress classification.
- We design an intricate sparse feature extraction module, capitalizing on dilated convolution, to distill and amalgamate sparse features from diverse feature scales.
- We crafting a loss function underpinned by a thorough of the pavement, ensuring consistent feature space sparsity.
- We empirical validation of the model on an authentic dataset, underscoring its superiority.

2. Method

The maintenance and preservation of transportation infrastructure systems hinge crucially upon the effective identification and classification of pavement distress. The health and safety of these systems are paramount, necessitating reliable and precise methods for identifying potential hazards. In the current study, we introduce a novel approach explicitly crafted for the classification of pavement image distress. This innovative method addresses the shortcomings of existing image classification models in the field of computer vision, which often fail to account for features unique to pavement distress images.

At the heart of our method lies a bespoke sparse feature module. This module is designed to optimize the extraction of image features, thus significantly enhancing the overall performance of the model. Pavement images typically present a unique challenge, wherein the background occupies the majority of the image, and the distresses, which are the primary targets of our analysis, are sparsely distributed. Figure 1 presents an instance of a road image dataset. The visual display comprises two facets: a grayscale image to the left, which is employed in our model due to its optimal fitting characteristics, and a depth map appearing concurrently on the right for clarity and comprehension. A critical aspect to note in the figure is a specific region of interest, demarcated by a red bounding box. This portion, though relatively diminutive in comparison to the entirety of the image, captures the focus of our model due to the significant information it carries about the road's condition. The highlighted area demonstrates the inherent sparse nature of the distress

features in road imagery—it constitutes only a small proportion of the total image pixels, yet represents the vital information necessary for effective distress classification.
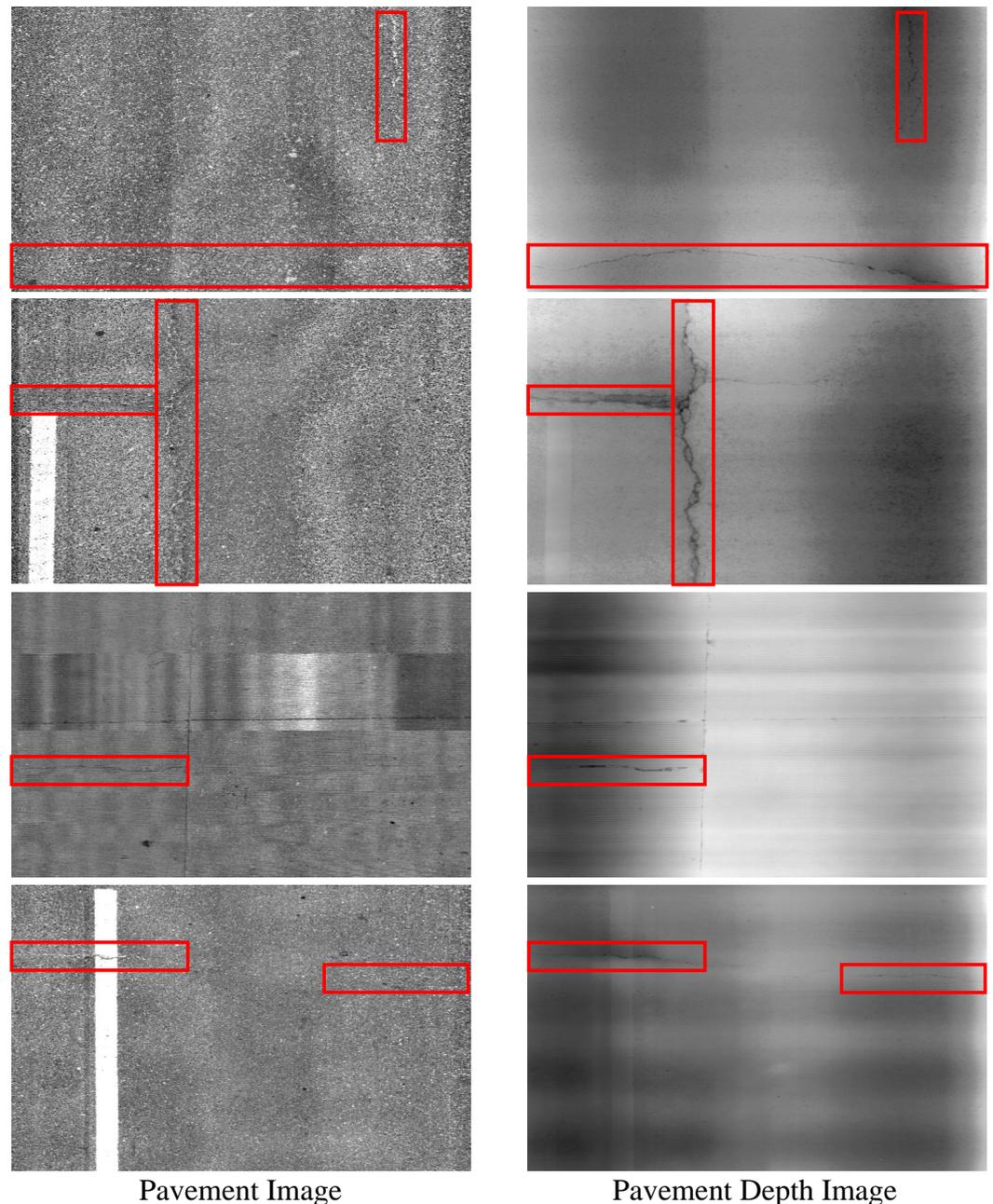

Figure 1. Samples of the pavement image data, with the depth image shown simultaneously for clarity considerations (only the left grayscale image is used in the model for applicable fitness considerations). It can be seen that the target shown in the red box (distress) is only a small part of the whole image pixels, reflecting the sparse feature.

By centering our attention on these sparse regions, we ensure that no critical data points are overlooked. Further, we have developed a loss function tailored to imposing a sparsity constraint, ensuring that the features extracted align with the sparse nature of the target areas. This illustrative example, therefore, underlines our model's core proposition: focusing on sparse, significant features rather than the largely monotonous background. It encapsulates the challenges of pavement distress classification and how our proposed method addresses them, serving as a testament to the potential effectiveness of our sparse perception-based road disease classification model.

2.1. BoneNetwork

Our method leverages the strengths of the High-Resolution Network (HRNet) [12], an established architecture, or BoneNetwork, widely employed for diverse tasks within the realm of computer vision. HRNet has heralded significant progress in the field by maintaining high-resolution representations throughout the model. Its multi-scale, parallel, and high-resolution design effectively tackles the issue of information loss at lower resolutions, thereby facilitating the capture of fine-grained details. However, when HRNet is applied to pavement image data, certain limitations become evident. The model fails to effectively isolate and highlight the sparse target features against the pervasive background. Our work is designed to address these limitations, focusing on the critical, sparse target areas to yield better results.

2.2. Sparse Module

The second key component of our method is the Sparse Module. This module serves a critical role in our proposed model. It is crafted to focus on the sparse targets amidst the monotonous backgrounds, a feature often overlooked in conventional methods. The module is designed to be adaptable and scalable, ensuring compatibility with diverse model architectures. It harnesses the power of sparse representation theory, integrating sparse paradigm formulas to enhance the extraction of sparse features from pavement images.

Figure 2 representation portrays the comprehensive architecture of our model, meticulously designed to address the challenges of pavement distress classification. This schematic illustrates a holistic workflow that initiates with an input pavement image and ends with the extraction and sharing of sparse image features then getting classification results.
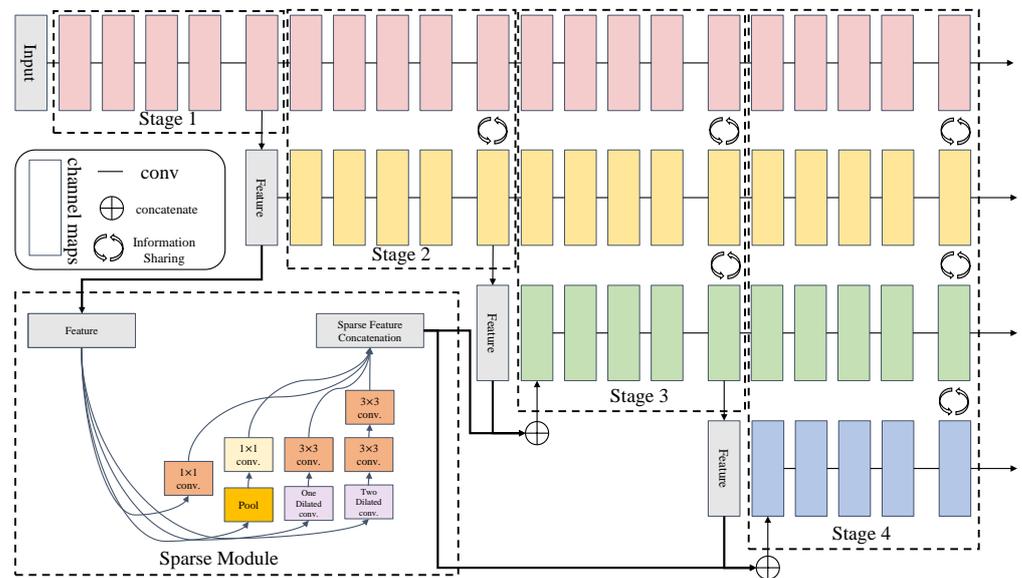

Figure 2. The overall architecture of our model. Given a pavement image, four feature extractors of different sizes extract image features and iterate forwards, sharing information between the different sizes. Meanwhile, we design a sparse module to take features from the max size features, while using a 1×1 convolution to keep the features, a pooling layer to read the sparse points and expand the sensory field using two dilated convolutions with different sizes, and finally concatenate the outputs up to participate in the sharing of different sizes as sparse features.

The pavement image serves as the initial input, and the image is processed through four distinctive feature extractors. Each of these extractors is unique in terms of size, offering a multi-scale method to feature extraction. Simultaneously, as part of this intricate process, we introduce a bespoke sparse module. This component is designed to extract features from the largest size features, further enhancing the ability of the model to identify

sparse areas of interest. The utilization of a 1×1 convolution aids in the maintenance of these features. In the following step, a pooling layer is employed to identify and highlight sparse points within the image, consequently expanding the sensory field of the model. The sparse module also leverages dilated convolutions of varying sizes to further scrutinize the image, focusing on the comprehensive coverage and extraction of relevant features. The model then amalgamates these diverse outputs, effectively compiling them into a singular, coherent format. This concatenated output serves to participate in multi-dimensional information exchange among different sizes, thus enhancing the overall feature extraction process.

Figure 3 visually encapsulates a comparative exploration of feature extraction methodologies for pavement distress. The graphical representation utilizes color coding to distinguish between different regions of interest within the pavement image. Notably, the red box delineates the target area, indicative of pavement distress, while the blue box demarcates the surrounding background region.
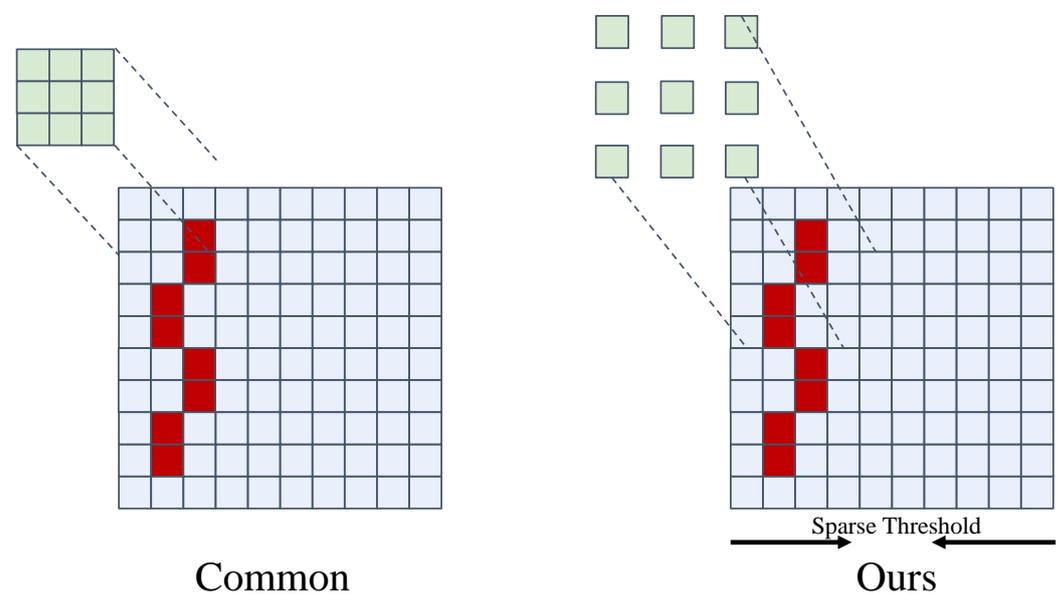

Figure 3. The figure demonstrates the difference in feature extraction for pavement distress, where the red box represents the target,the green box represents the convolutional kernel and the blue box represents the background. It can be seen that we have improved the qualitative feature extraction by expanding the receptive field and setting the sparse threshold (0.1).

The crux of our enhancement lies in the ingenious alteration of the conventional feature extraction process. As the illustration suggests, we have enhanced the qualitative aspect of feature extraction by expanding the receptive field. This expansion allows the model to grasp a broader perspective of the image, effectively incorporating a larger, more diverse set of contextual details. By doing so, the model transcends the limitations of a narrow field, gaining a more holistic understanding of the spatial and relational intricacies embedded within the image.

Furthermore, the illustration elucidates our decision to implement a sparse threshold set at 0.1. This sparse threshold plays a pivotal role in fine-tuning the feature extraction process, ensuring that the model maintains a focus on areas with significant information density. The sparse threshold serves as a guiding parameter, allowing the model to prioritize extraction from the key areas of interest—the sparse distress points—while effectively reducing the noise from the relatively monotonous background. Through the expanded receptive field and the incorporation of a sparse threshold, we have been able to enhance the qualitative feature extraction, leading to a more accurate and effective representation of pavement distress.

2.3. Domain Sparse Constrained Loss Function

Building upon the foundational concepts of computer vision and image classification, the third component of our research introduces a uniquely tailored mechanism—the Domain Sparse Constrained Loss Function. This innovative approach to handling the intricacies of pavement distress classification is borne out of the need to explicitly incorporate the sparse nature of pavement distresses into our model's learning process.

Figure 4 portrays the distinct difference in feature space distribution between conventional image classification tasks and the specialized domain of pavement images. The left shows the feature distribution pattern of usual computer vision tasks, the middle shows the sparse feature distribution specific to the road distress domain, and the right shows our proposed loss function dedicated to optimizing the feature extraction by exploiting the distance between sparse clusters. The key objective of our function is to discern and distinguish these sparse, yet critical, features in the pavement images. The loss function, by imposing a distance-based separation, helps to accentuate the sparse features, thereby facilitating their recognition and classification.
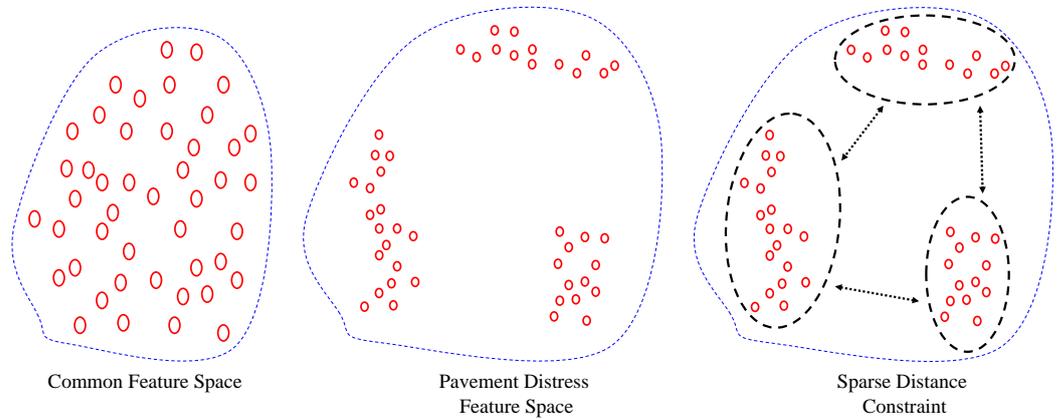

Common Feature Space   Pavement Distress Feature Space   Sparse Distance Constraint

Figure 4. The figure illustrates the feature space for the conventional task as well as the feature space in the pavement image. In general, the feature distribution is more even, whereas the pavement domain is more sparse and concentrated, and the loss function we designed aims to clearly separate sparse features with distance.

Unlike traditional loss functions, our design takes into account the sparsity of target features in pavement image data. This tailoring to the specifics of pavement distresses ensures the optimal extraction and classification of the sparse features. In essence, we modify the original loss function by incorporating a term for $L1$ regularization, which helps enforce sparsity in the extracted features:

$$L_{final} = L_{CrossEntropy} + \lambda \times L1\_loss \tag{1}$$

where $L_{CrossEntropy}$ is the cross entropy loss function typically used in classification tasks, and $L1_{loss}$ is the $L1$ regularization term. The λ factor in this equation is a hyperparameter that controls the degree of regularization, striking a balance between the original loss and the imposed sparsity constraint (is 1 in this case). The $L1$ regularization term is defined as the sum of the absolute values of the weights in our model's dilated convolution modules, a formulation known for its property of promoting sparsity:

$$L1\_loss = \sum_{Sparse\ Module} ||weight||_1 \tag{2}$$

In this equation, $||weight||_1$ represents the $L1$ norm of the weight matrix in the convolutional layer of each dilated convolution module.

Moreover, the Domain Sparse Constrained Loss Function addresses the common challenges in handling sparse data. Conventional models often fall short in dealing with

sparse data due to the risk of overfitting or inadequate feature extraction. Our custom loss function mitigates this by imposing a constraint that not only aligns with the data's natural distribution but also aids in maintaining a robust model by discouraging overfitting.

To recap, the Domain Sparse Constrained Loss Function serves a dual purpose: Firstly, it improves the quality of feature extraction from sparse data by incorporating the nature of the data directly into the model's learning process. Secondly, it improves the overall robustness and performance of the model by mitigating the common challenges associated with handling sparse data.

3. Results

In this section, we present an empirical evaluation of our proposed model and compare its performance with six contemporary, state-of-the-art methods in image classification. The primary aim of this study was to ascertain the performance and efficiency of our model in addressing image classification tasks with a particular focus on pavement distress categories. A dataset embodying an array of pavement distress categories was meticulously analyzed as part of the evaluation process. Extensive experimentation underscored the superior performance of our sparse model in terms of classification accuracy as compared to the other methods.

- ResNet (Residual Network) [13]: A type of convolutional neural network (CNN) that uses shortcut connections (also known as residual connections) to make deeper models possible. These shortcut connections help to solve the vanishing gradient problem, which can occur when training very deep neural networks. ResNet models can have hundreds or even thousands of layers, and they have achieved state-of-the-art performance on many computer vision tasks.
- VGGNet (Visual Geometry Group Network) [2]: The model is characterized by its simplicity, using only 3×3 convolutional layers stacked on top of each other in increasing depth. The depth of the network, which is a defining characteristic of VGGNet, contributes to learning more complex features at higher layers. Despite its simplicity and depth, it is quite resource-intensive.
- DenseNet (Densely Connected Convolutional Networks) [14]: This is a type of CNN that connects each layer to every other layer in a feed-forward fashion, different from ResNet which only involves a connection from the previous layer. This architecture enhances the vanishing gradient problem, strengthens feature propagation, encourages feature reuse, and significantly reduces the number of parameters, thus making the model more efficient.
- EfficientNet [15]: This model uses a simple yet effective compound coefficient. The EfficientNet models, with this compound scaling method, have achieved state-of-the-art accuracy on ImageNet while being up to more efficient than ConvNets.
- HRNet (High-Resolution Networks) [16]: The HRNet maintains high-resolution representations throughout the entire network, different from other networks that might reduce the resolution in the earlier stages. It has parallel multi-resolution convolutions which exchange information across different resolutions, making it particularly effective for semantic segmentation tasks where high-resolution spatial information is important.
- RexNet [12]: This model proposes a simple yet effective channel configuration that can be parameterized by the layer index. As a result, RexNet achieves remarkable performance on ImageNet classification and transfer learning tasks including COCO object detection, COCO instance segmentation, and fine-grained classifications.
- SparseModule refers to the ablation experiment of our method, using only SparseModule without adjusting the loss function.

Dataset. Our dataset comprises over 100,000 km of road data from various Chinese regions, such as Jiangsu, Henan, Xinjiang, and Shanxi, and features a variety of pavement conditions. Data were collected using a 3D line laser scanner mounted on a vehicle, with a 1 mm sampling accuracy and a 100 km/h acquisition speed, scanning a 4 m wide lane

at a 28 kHz frequency. The raw data were converted to images, and the brightness was homogenized. Images were resized to 512 × 512 pixels for model input as shown in Figure 5. Road engineering professionals performed pixel-level mask labeling, and domain experts confirmed accuracy. The dataset was split into 80% training, 10% validation, and 10% test sets. The Road dataset includes diverse cracks and healthy pavement, with no overlap or excess images from the same road segment, ensuring a high-resolution and robust dataset.
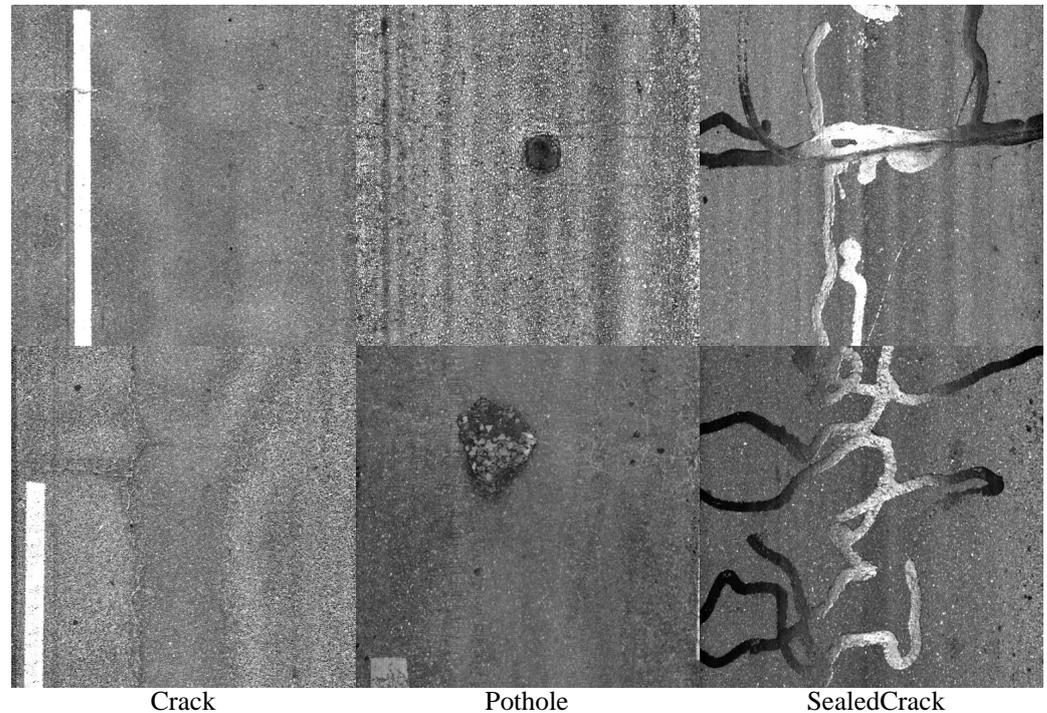

| Crack | Pothole | SealedCrack |

Figure 5. A sample of pavement distress types.

Metrics. The *F*1 score is a commonly used metric in machine learning, particularly in tasks that deal with binary or multi-class classification. It is a balanced measure of precision and recall, providing a single metric that encapsulates model performance in terms of both false positives and false negatives. Precision measures the correctness of the positive predictions. It calculates the ratio of correctly predicted positive observations to the total predicted positives. Recall measures the ability of a model to find all the relevant cases in a dataset. It calculates the ratio of correctly predicted positive observations to all the observations in the actual positive class.

$$\text{Precision} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)} + \text{False Positive (FP)}} \tag{3}$$

$$\text{Recall} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)} + \text{False Negative (FN)}} \tag{4}$$

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{5}$$

where True Positive (TP) means images correctly classified as positive, False Positive (FP) means images are incorrectly classified as positive and False Negative (FN) means positive images are incorrectly classified as negative.

Tables 1 and 2 outlines the performance metrics of the various models tested against three classes of road damage, namely, Crack, Pothole, and SealedCrack, and their average performance across these classes. In terms of the average performance, our method significantly outperforms all other models with average *F*1 scores of 0.94, an improvement over

the closest competitors, EfficientNet and REXNet, which both achieved an average *F*1 score of 0.92.

Table 1. Precision/recall comparison with state-of-the-art methods.

Method	Crack	Pothole	SealedCrack	Avg
ResNet	0.92/0.99	0.92/0.79	0.88/0.95	0.90/0.87
VGGNet	0.93/1.00	0.92/0.86	0.87/0.87	0.90/0.91
DenseNet	0.93/1.00	0.92/0.79	0.92/0.92	0.92/0.90
EfficientNet	0.96/0.99	1.00/0.79	0.92/0.92	0.96/0.90
HRNet	0.97/0.99	0.92/0.79	0.89/0.89	0.92/0.89
RexNet	0.98/1.00	1.00/0.71	0.93/0.97	0.97/0.89
SparseModule	0.96/1.00	1.00/0.79	0.90/0.95	0.95/0.91
Ours	0.98/1.00	1.00/0.86	0.88/0.95	0.95/0.93

Table 2. *F*1 score comparison with state-of-the-art methods.

Method	Crack	Pothole	SealedCrack	Avg
ResNet	0.96	0.85	0.91	0.90
VGGNet	0.96	0.89	0.87	0.90
DenseNet	0.97	0.85	0.92	0.91
EfficientNet	0.98	0.88	0.92	0.92
HRNet	0.98	0.85	0.89	0.90
REXNet	0.99	0.83	0.95	0.92
SparseModule	0.98	0.88	0.92	0.92
Ours	0.99	0.92	0.91	0.94

Our model also excelled in the identification of Crack and Pothole defects, recording top *F*1 scores of 0.99 and 0.92, respectively. Given the inherent sparsity of cracks and potholes in pavement images, our model has showcased impressive performance gains in these categories. Cracks and potholes present unique challenges due to their infrequent occurrence and diverse characteristics, calling for a model that can effectively isolate and recognize these sparse yet critical data points. Our proposed model, designed with a focus on the unique sparsity of such features, has excelled in the accurate detection of both crack and pothole defects, recording top *F*1 scores of 0.99 and 0.92, respectively. The effective capture and identification of these sparse and intricate patterns in road images demonstrate the adaptability and efficacy of our method. The superior performance in detecting these specific distress types underscores the model's ability to address the critical aspect of sparsity in pavement distress detection.

These findings robustly affirm the efficacy of our proposed model in recognizing a wide array of road damage types. The superior performance of our model can be largely attributed to the innovative design of our sparse feature module and the domain-specific sparse constraint, which contribute to the robustness of the model in a variety of detection scenarios.

4. Conclusions

In this study, we have presented a novel method for pavement distress classification, designed specifically to handle the unique challenges posed by the sparse distribution of features in road images. The traditional image classification models used in computer vision often fail to capture these sparse features effectively, leading to suboptimal performance when applied to tasks such as road distress detection and classification. Our model, equipped with a specially designed sparse feature module and a sparse constrained loss function, shows superior performance in tackling this problem.

The experimental results indicate that our model outperforms other state-of-the-art models in terms of classification accuracy for various road distress types, especially for sparsely distributed features such as cracks and potholes. The improved performance can

be attributed to the effective extraction and focus on these sparse yet crucial features in the image.

Additionally, the sparse constrained loss function provides a more refined control over the learning process, guiding the model to focus on the sparse target areas and ignore the largely monotonous background. This results in a more accurate and effective distress classification, as demonstrated by the high average $F1$ score achieved by our model.

This research provides paving the way for further optimization of image classification models for similar tasks. Future work can extend this methodology to other sparse feature problems within the domain of road detection.

The success of our model underlines the significance of incorporating domain-specific knowledge in the design of image classification models, encouraging continued exploration of specialized models to tackle domain-specific challenges effectively. This work also holds practical implications for transportation authorities, offering a more reliable, efficient, and scalable method for road maintenance and safety assessment.

5. Future Work

While our study has achieved notable progress in the domain of pavement distress classification, there are several promising directions for future research that can extend and enhance the proposed methodology. Improving model performance through augmented data is a crucial consideration. Developing specialized data augmentation techniques that capture the nuances of real-world road conditions and distress instances could lead to enhanced model generalization. Extending the methodology to analyze distress patterns across extensive road networks could provide insights into broader maintenance strategies. Exploring ways to incorporate geospatial information and analyzing distress trends on a regional scale could guide informed decision making by transportation authorities. Researching ways to optimize the algorithm's efficiency and enable real-time distress assessment would be valuable. In conclusion, our study lays the foundation for further research that addresses both technical intricacies and real-world applications in pavement distress classification. By exploring these potential directions, we aim to contribute to the advancement of computer vision methodologies and the effectiveness of road maintenance practices.

Author Contributions: Methodology, Y.Y.; Software, Y.Y.; Resources, H.L.; Data curation, J.C.; Funding acquisition, J.L. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data available on request due to restrictions eg privacy or ethical The data presented in this study are available on request from the corresponding author. The data are not publicly available due to business reasons.

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

References

1. Rodríguez-Garavito, C.; Ponz, A.; García, F.; Martin, D.; de la Escalera, A.; Armingol, J.M. Automatic laser and camera extrinsic calibration for data fusion using road plane. In Proceedings of the IEEE 17th International Conference on Information Fusion (FUSION), Salamanca, Spain, 7–10 July 2014; pp. 1–6.
2. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
3. Jooste, F.; Costello, S.; Rainsford, S. Prediction of network level pavement treatment types using multi-classification machine learning algorithms. *Road Mater. Pavement Des.* **2022**, *24*, 410–426. [CrossRef]

4. Hanandeh, S. Introducing mathematical modeling to Estimate Pavement Quality Index of Flexible Pavements based on Genetic Algorithm and Artificial Neural Networks. *Case Stud. Constr. Mater.* **2022**, *16*, e00991. [CrossRef]

5. Zhao, J.; Wang, H.; Lu, P. Machine learning analysis of overweight traffic impact on survival life of asphalt pavement. *Struct. Infrastruct. Eng.* **2021**, *19*, 606–616. [CrossRef]

6. Kaloop, M.; El-Badawy, S.; Hyuck Ahn, J.; Sim, H.B.; Hu, J.; El-Hakim, R.A.A. A hybrid wavelet-optimally pruned extreme learning machine model for the estimation of international roughness index of rigid pavements. *Int. J. Pavement Eng.* **2020**, *23*, 862–876. [CrossRef]

7. Qureshi, W.S.; Hassan, S.I.; Mckeever, S.; Power, D.; Mulry, B.; Feighan, K.; O'Sullivan, D. An Exploration of Recent Intelligent Image Analysis Techniques for Visual Pavement Surface Condition Assessment. *Sensors* **2022**, *22*, 9019 . [CrossRef] [PubMed]

8. Dong, Y.; Shao, Y.; Li, X.; Li, S.; Quan, L.; Zhang, W.; Du, J. Forecasting Pavement Performance with a Feature Fusion LSTM-BPNN Model. In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, Beijing, China, 3–7 November 2019; pp. 1953–1962. [CrossRef]

9. Yang, X.; Jinchao, G.; Ling, D.; You, Z.; Lee, V.C.; Hasan, M.; Cheng, X. Research and applications of artificial neural network in pavement engineering: A state-of-the-art review. *J. Traffic Transp. Eng.* **2021**, *8*, 1000–1021. [CrossRef]

10. Damirchilo, F.; Hosseini, A.; Parast, M.; Fini, E. Machine Learning Approach to Predict International Roughness Index Using Long-Term Pavement Performance Data. *J. Transp. Eng. Part B Pavements* **2021**, *147*, 04021058. [CrossRef]

11. Xu, Y.; Zhang, Z. Review of Applications of Artificial Intelligence Algorithms in Pavement Management. *J. Transp. Eng. Part B Pavements* **2022**, *148*, 03122001. [CrossRef]

12. Han, D.; Yun, S.; Heo, B.; Yoo, Y. Rethinking channel dimensions for efficient model design. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 732–741.

13. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.

14. Huang, G.; Liu, Z.; Van Der Maaten, L.; 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.

15. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, PMLR, Long Beach, CA, USA, 9–15 June 2019 ; pp. 6105–6114.

16. Wang, J.; Sun, K.; Cheng, T.; Jiang, B.; Deng, C.; Zhao, Y.; Liu, D.; Mu, Y.; Tan, M.; Wang, X.; et al. Deep high-resolution representation learning for visual recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *43*, 3349–3364.