SCALABLE AND EFFICIENT MULTI-WEATHER CLASSIFICATION FOR AUTONOMOUS DRIVING WITH CORESETS, PRUNING, AND RESOLUTION SCALING

Tushar Shinde & Avinash Kumar Sharma
Indian Institute of Technology Madras Zanzibar shinde@iitmz.ac.in

ABSTRACT

Autonomous vehicles require robust perception systems capable of operating in diverse weather conditions, including snow, rain, fog, and storms. In this work, we present a scalable and efficient approach for multi-weather classification in autonomous driving, leveraging the WEDGE (WEather images by DALL-E GEneration) dataset. Our study investigates three complementary techniques to enhance classification performance and efficiency: Coreset Selection, Resolution Scaling, and Model Compression via Adaptive Pruning and Quantization. Specifically, we evaluate the impact of coreset selection methods (random and margin-based) at varying data fractions (e.g., 1, 0.75, 0.5, 0.25, 0.1), assess model robustness under low-resolution settings (224×224, 112×112, 56×56), and demonstrate that adaptive pruning combined with 8-bit quantization can reduce model size by up to 85% while maintaining competitive classification accuracy. Experimental results validate the effectiveness of our integrated approach, providing a scalable and robust solution for multi-weather classification. This work advances the feasibility of deploying perception models in real-world autonomous driving systems operating under adverse weather conditions and limited computational resources.

1 INTRODUCTION AND RELATED WORK

Autonomous driving systems rely heavily on computer vision to perceive their environment and make real-time decisions Yurtsever et al. (2020). However, performance degradation under extreme weather conditions such as rain, snow, fog, and storms poses significant challenges, impairing both visibility and sensor readings Eigen et al. (2013); Marathe et al. (2022). While substantial progress has been made in developing systems for ideal weather conditions, robustness under adverse weather remains an open problem. The scarcity of diverse and annotated datasets for extreme weather conditions intensifies the challenge, as collecting such data is both expensive and hazardous.

To address this challenge, synthetic data generation methods have become invaluable, providing a feasible alternative to real-world data collection. Recent studies have explored synthetic datasets such as WEDGE Marathe et al. (2023), that offer an alternative by simulating diverse weather conditions, enabling the development of robust vision models. WEDGE, generated using a vision-language model (VLM), consists of 3,360 images spanning 16 distinct weather conditions, providing a scalable resource for training and evaluating models. This dataset is particularly valuable for classification tasks under diverse weather scenarios.

Weather conditions like rain, snow, and fog introduce considerable challenges for visual perception, especially in tasks like classification and object detection Li et al. (2023); Mehra et al. (2020). While these methods have shown promise, they often require extensive fine-tuning and may not generalize effectively to out-of-distribution (OOD) data Filos et al. (2020). In contrast, our work leverages WEDGE for scalable training and introduces coreset selection Lin et al. (2024), a method to reduce the size of training data while maintaining generalization performance. This approach has received limited attention in the context of multi-weather classification.

Efficient deployment of deep learning models in autonomous vehicles also requires addressing computational and memory constraints Liu et al. (2021); Ople et al. (2022). Model compression techniques such as pruning Su et al. (2024); Li et al. (2016); Liu et al. (2018) and quantization Jacob et al. (2018); Cui et al. (2022) have been widely adopted to address these challenges. Pruning removes less important parameters, while quantization reduces the precision of weights and activations, significantly lowering storage and computation costs. Another critical factor in deploying vision models is the impact of input resolution on performance Sun et al. (2020); Song & Chandraker (2014). Low-resolution images often degrade accuracy due to the loss of fine-grained details. Previous studies Ding et al. (2023) have proposed methods to mitigate this impact.

Despite significant advances in weather condition robustness Almalioglu et al. (2022), coreset selection Lu et al. (2024); Verwimp et al. (2023), resolution scaling Wang et al. (2022), and model compression Tukan et al. (2022), no prior work has integrated these techniques into a unified framework. Our work fills this gap by combining these approaches into a scalable solution for multi-weather classification, improving both accuracy and efficiency. Experimental results demonstrate the effectiveness of this approach for developing robust vision models suitable for autonomous driving. The remainder of this paper is organized as follows: Section 2 details the methodology, Section 3 presents datasets and experimental setup, Section 4 presents experimental results, and Section 5 concludes the paper with future directions.

2 METHODOLOGY

We propose an efficient framework for multi-weather classification in autonomous driving, integrating coreset selection, resolution scaling, and model compression to optimize accuracy and computational efficiency.

Coreset Selection. To reduce training complexity while preserving classification accuracy, we employ coreset selection techniques to reduce the training dataset size Lee et al. (2024). Coreset selection aims to identify a subset of the most informative samples from the original training set $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{N}$, where x_i represents the input sample and y_i the corresponding label. The goal is to minimize computational overhead by selecting a representative subset $\mathcal{C} \subseteq \mathcal{D}$ such that the coreset retains the model's classification performance. This allows us to perform training on a significantly smaller subset while ensuring that the resulting model's performance is not substantially degraded.

Resolution Scaling. To assess model robustness across varying resolutions, images are resized to 224×224, 112×112, and 56×56 pixels. We analyze the trade-off between computational efficiency and classification accuracy to determine the optimal resolution for real-time deployment.

Model Compression. We apply *pruning* and *quantization* to optimize model size and inference speed. Extending Shinde (2024), we introduce Layer Importance Metric (LIM), incorporating the *Layer Parameter Ratio* and *Layer Sparsity Ratio* to guide compression. Layers with a higher number of parameters contribute more to the model size and thus are prioritized for compression. The parameter ratio of layer i is computed as:

$$P_i = \frac{\text{Number of parameters in layer } i}{\text{Total number of parameters in the model}} \tag{1}$$

The sparsity ratio of a layer, defined as the proportion of near-zero parameters relative to the total number of parameters in the layer, reflects its contribution to the model's representational capacity. The layer sparsity ratio is computed as:

$$S_i = \frac{\text{Number of near-zero parameters in layer } i}{\text{Total number of parameters in layer } i} \tag{2}$$

The LIM for each layer is then computed as a weighted combination of these two factors:

$$LIM_i = \beta \cdot P_i + (1 - \beta) \cdot S_i \tag{3}$$

where $\beta \in [0, 1]$ is a hyperparameter empirically determined to balance the impact of parameter ratio and sparsity ratio on classification performance. **Pruning.** Adaptive layer-wise pruning removes less important weights based on LIM, using a layer-specific threshold for structured weight reduction. **Quantization.** Post-pruning, weights are quantized to 8-bit precision, significantly reducing memory and computation requirements for edge deployment. By combining coreset selection, resolution scaling, and model compression, our framework ensures scalable, efficient, and accurate multi-weather classification for autonomous driving.

3 Datasets and Experimental Setup

Dataset: We conduct all experiments using the WEDGE dataset Marathe et al. (2023), a synthetic dataset specifically designed for weather classification and object detection tasks in autonomous driving. The dataset contains 3360 images representing 16 distinct weather conditions: snowing, raining, dusty, foggy, sunny, lightning, cloudy, hurricane, night, summer, spring, winter, fall, tornado, day, and windy. The dataset is split into 80% for training and 20% for testing, ensuring robust evaluation and model generalization.

Experimental Setup: *Configuration.* All experiments were conducted on the Kaggle platform, utilizing NVIDIA Tesla P100 and G4 GPUs to accelerate coreset selection, model training, and adaptive model compression. We used Python implementations with PyTorch to ensure reproducibility and ease of deployment. We evaluate several CNN architectures for image classification, including *ResNet18* He et al. (2016) and *EfficientNetB0.* Models were trained with a learning rate of 0.001, a batch size of 64, and for 25 epochs. To prevent exploding gradients, gradient clipping was applied with a value of 1.0. A 3-epoch warm-up phase was used to stabilize training in its initial stages, and early stopping was employed to terminate training if validation accuracy did not improve over three consecutive epochs. The Adam optimizer was used for its robust convergence properties.

Coreset Selection Setup. To explore the effects of dataset size on both performance and computational cost, we evaluated various coreset selection strategies, including Random and Margin-based methods, across dataset fractions of $\{1.0, 0.75, 0.5, 0.25, 0.1, 0.05\}$.

Resolution Scaling Setup. To identify the optimal resolution for the weather classification task, we evaluated different image resolutions (224x224, 112x112, and 56x56) to assess their impact on both model performance and computational efficiency.

Pruning Setup. Pruning experiments employed both standard pruning strategies and Layer-wise Adaptive Pruning (LAP), with pruning levels $k_i = \{2, 1.5, 1, 0.5\}$, representing pruning rates of approximately 95%, 86%, 68%, and 38%, respectively. This setup allows us to investigate how pruning impacts both performance and compression, particularly when combined with coreset selection methods that reduce training data size.

Quantization Setup. Quantization experiments used fixed 8-bit precision for model quantization post pruning. These were evaluated alongside coreset selection methods (Random and Margin-based) across different dataset fractions on the ResNet18 architecture.

Evaluation Metrics: We assess both model performance and compression efficiency using the following metrics. *Accuracy* denotes the percentage of correctly classified test images. *Compression Ratio* quantifies compression effectiveness by measuring the reduction in model size, calculated as the ratio of the original model size to the compressed model size.

4 Results and Analysis

We evaluate the performance of our scalable multi-weather classification framework using classification accuracy (Acc.) and compression ratio (CR). Table 1 and Table 2 compare different image resolutions (224x224, 112x112, 56x56), coreset selection strategies (Random and Margin), and pruning levels (P levels), assessing both baseline and adaptive methods, including Layer Adaptive Pruning (LAP) and LAP with 8-bit quantization (LAP + 8-bit Q).

Baseline Performance Across Models. At 224x224 resolution, EfficientNetB0 achieves the highest accuracy of 77.83%, followed by ResNet18 at 75.60%. These models excel due to efficient parameter use and advanced architectures like compound scaling. ResNet18 perform well at 75.60%, offering a balance between efficiency and performance.

Impact of Resolution Scaling. Reducing resolution from 224x224 to 112x112 causes a slight drop in performance, with EfficientNetB0 experiencing a notable decline. However, lightweight models like EfficientNetB0 maintain competitive results, with EfficientNetB0 at 73.21%. At 56x56, all models show significant degradation, though EfficientNetB0 retain better accuracy (68.01%).

Effect of Coreset Selection Strategies. The Margin-based coreset selection outperforms Random selection, especially at lower resolutions and smaller dataset fractions. For example, at 224x224

Table 1: Performance Comparison of Multi-Weather Classification Across Models, Resolution Levels, and Coreset Selection Methods (Ran. Random and Marg. Margin).

	ResNet18						EfficientNetB0					
Fraction	**224x224**		**112x112**		**56x56**		**224x224**		**112x112**		**56x56**	
	Ran.	Marg.	Ran.	Marg.	Ran.	Marg.	Ran.	Marg.	Ran.	Marg.	Ran.	Marg.
1.00	75.60	-	73.21	-	64.88	-	77.83	-	73.21	-	68.01	-
0.75	73.51	72.62	68.01	69.49	62.50	62.95	76.93	75.74	72.92	69.05	63.10	63.69
0.50	70.83	70.98	68.45	64.58	61.76	60.57	73.36	71.88	71.43	69.64	60.27	63.54
0.25	65.48	65.33	61.46	61.46	54.46	55.21	68.90	68.60	64.73	63.54	54.02	55.51
0.10	53.72	56.10	48.81	46.58	42.26	46.58	59.97	59.08	53.57	54.76	43.75	41.82
0.05	38.54	42.71	33.48	36.01	32.44	25.89	45.68	47.47	40.18	44.20	29.17	33.04

Table 2: Comparison of Accuracy and Compression Ratio (CR) across resolutions, coreset methods, and P levels on ResNet18 model.

Resolution	**224x224**				**112x112**				**56x56**			
Coreset	**Random**		**Margin**		**Random**		**Margin**		**Random**		**Margin**	
Method	**Acc.**	**CR**	**Acc.**	**CR**	**Acc.**	**CR**	**Acc.**	**CR**	**Acc.**	**CR**	**Acc.**	**CR**
P level 2	6.25%	21.3	6.25%	21.3	6.25%	21.2	6.25%	21.2	6.25%	20.8	6.25%	20.7
P level 1.5	6.55%	8.3	6.40%	8.3	7.74%	8.3	6.70%	8.3	6.70%	8.3	8.33%	8.3
P level 1	9.52%	3.5	15.92%	3.5	24.55%	3.5	27.53%	3.5	29.91%	3.6	21.13%	3.6
P level 0.5	69.64%	1.7	63.10%	1.7	69.49%	1.7	68.45%	1.7	64.43%	1.7	61.90%	1.7
Baseline	76.49%	1.0	74.55%	1.0	72.62%	1.0	71.88%	1.0	66.82%	1.0	64.29%	1.0
LAP	76.51%	1.2	74.40%	1.3	73.07%	1.5	70.09%	1.9	65.48%	1.4	63.84%	2.4
LAP + 8-bit Q	76.64%	5.5	74.11%	6.4	74.26%	8.3	72.02%	8.2	66.07%	6.7	65.48%	7.3

resolution with 10% dataset fraction, ResNet18 achieves 56.10% accuracy with Margin-based selection, compared to 53.72% with Random. The Margin-based strategy helps preserve crucial information, improving model performance with fewer samples. EfficientNetB0 maintain 59.97% accuracy at a 10% dataset fraction, showcasing their robustness even with reduced data.

Effect of Model Compression.

Pruning Levels (P levels). Increasing the pruning level (lowering P) enhances compression but reduces accuracy. For example, at 224x224 resolution, accuracy drops from 76.49% (Baseline, CR=1.0) to 6.25% at P level 2 (CR=21.3), highlighting the trade-off between compression and accuracy. *Performance of Pruning and 8-bit Quantization.* The LAP and LAP + 8-bit Q methods improve compression without significant accuracy loss. At 224x224 resolution, LAP achieves 76.51% accuracy with a CR of 1.2, while LAP + 8-bit Q reaches a CR of 5.5 with 76.64% accuracy, demonstrating their efficiency for resource-constrained applications. *Impact of Resolution Scaling on Compression.* Reducing resolution from 224x224 to 56x56 enhances compression (e.g., P level 0.5, CR=1.7) but decreases accuracy (from 69.64% to 64.43%), suggesting that resolution scaling affects accuracy more than model size. *Effect of Coreset Selection on Compression.* The Margin-based coreset selection outperforms Random selection, achieving better accuracy (e.g., 68.45% vs. 67.49% at $112\text{x}112$, P level 0.5) while maintaining the same CR. This improves performance by focusing on more informative samples.

Overall, adaptive methods like LAP and LAP + 8-bit Q balance accuracy and compression effectively, ideal for resource-limited environments. Resolution scaling and pruning provide flexible trade-offs, making them suitable for autonomous driving models on edge devices.

5 Conclusion

We presented a scalable and efficient framework for multi-weather classification in autonomous driving, incorporating margin-based coreset selection, resolution scaling, and model compression. Our approach effectively reduces training data and model size, achieving up to compression 85%, while maintaining competitive accuracy, making it suitable for edge deployment. Despite minor performance drops with low-resolution inputs, the framework offers a favorable trade-off between accuracy and efficiency. Future directions include adapting the model to dynamic environments and exploring advanced compression techniques such as mixed precision quantization and dynamic pruning for improved real-time performance.

ACKNOWLEDGMENT

We thank Sai Krishna for his assistance in exploring simple coreset methods.

REFERENCES

Yasin Almalioglu, Mehmet Turan, Niki Trigoni, and Andrew Markham. Deep learning-based robust positioning for all-weather autonomous driving. *Nature machine intelligence*, 4(9):749–760, 2022.

Shaohua Cui, Yongjie Xue, Maolong Lv, Baozhen Yao, and Bin Yu. Cooperative constrained control of autonomous vehicles with nonuniform input quantization. *IEEE Transactions on Vehicular Technology*, 71(11):11431–11442, 2022.

Xinpeng Ding, Jianhua Han, Hang Xu, Wei Zhang, and Xiaomeng Li. Hilm-d: Towards high-resolution understanding in multimodal large language models for autonomous driving. *arXiv preprint arXiv:2309.05186*, 2023.

David Eigen, Dilip Krishnan, and Rob Fergus. Restoring an image taken through a window covered with dirt or rain. In *Proceedings of the IEEE international conference on computer vision*, pp. 633–640, 2013.

Angelos Filos, Panagiotis Tigkas, Rowan McAllister, Nicholas Rhinehart, Sergey Levine, and Yarin Gal. Can autonomous vehicles identify, recover from, and adapt to distribution shifts? In *International Conference on Machine Learning*, pp. 3145–3153. PMLR, 2020.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

Benoit Jacob, Skirmantas Kligys, Bo Chen, Menglong Zhu, Matthew Tang, Andrew Howard, Hartwig Adam, and Dmitry Kalenichenko. Quantization and training of neural networks for efficient integer-arithmetic-only inference. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2704–2713, 2018.

Hojun Lee, Suyoung Kim, Junhoo Lee, Jaeyoung Yoo, and Nojun Kwak. Coreset selection for object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 7682–7691, 2024.

Alexander C Li, Mihir Prabhudesai, Shivam Duggal, Ellis Brown, and Deepak Pathak. Your diffusion model is secretly a zero-shot classifier. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 2206–2217, 2023.

Hao Li, Asim Kadav, Igor Durdanovic, Hanan Samet, and Hans Peter Graf. Pruning filters for efficient convnets. *arXiv preprint arXiv:1608.08710*, 2016.

Jinpeng Lin, Zhihao Liang, Shengheng Deng, Lile Cai, Tao Jiang, Tianrui Li, Kui Jia, and Xun Xu. Exploring diversity-based active learning for 3d object detection in autonomous driving. *IEEE Transactions on Intelligent Transportation Systems*, 2024.

Hantao Liu, Ying He, F Richard Yu, and Jeremy James. Flexi-compression: a flexible model compression method for autonomous driving. In *Proceedings of the 11th ACM Symposium on Design and Analysis of Intelligent Vehicular Networks and Applications*, pp. 19–26, 2021.

Zhuang Liu, Mingjie Sun, Tinghui Zhou, Gao Huang, and Trevor Darrell. Rethinking the value of network pruning. *arXiv preprint arXiv:1810.05270*, 2018.

Han Lu, Xiaosong Jia, Yichen Xie, Wenlong Liao, Xiaokang Yang, and Junchi Yan. Activead: Planning-oriented active learning for end-to-end autonomous driving. *arXiv preprint arXiv:2403.02877*, 2024.

Aboli Marathe, Pushkar Jain, Rahee Walambe, and Ketan Kotecha. Restorex-ai: A contrastive approach towards guiding image restoration via explainable ai systems. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3030–3039, 2022.

Aboli Marathe, Deva Ramanan, Rahee Walambe, and Ketan Kotecha. Wedge: A multi-weather autonomous driving dataset built from generative vision-language models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 3318–3327, 2023.

Aryan Mehra, Murari Mandal, Pratik Narang, and Vinay Chamola. Reviewnet: A fast and resource optimized network for enabling safe autonomous driving in hazy weather conditions. *IEEE Transactions on Intelligent Transportation Systems*, 22(7):4256–4266, 2020.

Jose Jaena Mari Ople, Shang-Fu Chen, Yung-Yao Chen, Kai-Lung Hua, Mohammad Hijji, Po Yang, and Khan Muhammad. Controllable model compression for roadside camera depth estimation. *IEEE Transactions on Intelligent Transportation Systems*, 24(12):15478–15485, 2022.

Tushar Shinde. Adaptive quantization and pruning of deep neural networks via layer importance estimation. In *Workshop on Machine Learning and Compression, NeurIPS 2024*, 2024.

Shiyu Song and Manmohan Chandraker. Robust scale estimation in real-time monocular sfm for autonomous driving. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1566–1573, 2014.

Wensheng Su, Zhenni Li, Minrui Xu, Jiawen Kang, Dusit Niyato, and Shengli Xie. Compressing deep reinforcement learning networks with a dynamic structured pruning method for autonomous driving. *IEEE Transactions on Vehicular Technology*, 2024.

Pei Sun, Henrik Kretzschmar, Xerxes Dotiwalla, Aurelien Chouard, Vijaysai Patnaik, Paul Tsui, James Guo, Yin Zhou, Yuning Chai, Benjamin Caine, et al. Scalability in perception for autonomous driving: Waymo open dataset. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 2446–2454, 2020.

Murad Tukan, Loay Mualem, and Alaa Maalouf. Pruning neural networks via coresets and convex geometry: Towards no assumptions. *Advances in Neural Information Processing Systems*, 35: 38003–38019, 2022.

Eli Verwimp, Kuo Yang, Sarah Parisot, Lanqing Hong, Steven McDonagh, Eduardo Pérez-Pellitero, Matthias De Lange, and Tinne Tuytelaars. Clad: A realistic continual learning benchmark for autonomous driving. *Neural Networks*, 161:659–669, 2023.

Hai Wang, Yanyan Chen, Yingfeng Cai, Long Chen, Yicheng Li, Miguel Angel Sotelo, and Zhixiong Li. Sfnet-n: An improved sfnet algorithm for semantic segmentation of low-light autonomous driving road scenes. *IEEE Transactions on Intelligent Transportation Systems*, 23(11):21405–21417, 2022.

Ekim Yurtsever, Jacob Lambert, Alexander Carballo, and Kazuya Takeda. A survey of autonomous driving: Common practices and emerging technologies. *IEEE access*, 8:58443–58469, 2020.