

Enhancing Autonomous Driving in Adverse Weather: Road Surface Classification and Image Restoration Evaluation

Prayag Gaonkar *

Dougherty Valley High School, San Ramon, CA 94582 USA.

World Journal of Advanced Research and Reviews, 2025, 28(03), 030-041

Publication history: Received 21 October 2025; revised on 29 November 2025; accepted on 01 December 2025

Article DOI: <https://doi.org/10.30574/wjarr.2025.28.3.3992>

Abstract

The integration of artificial intelligence into transportation has pushed for the development of autonomous vehicles, promising improved safety and convenience. Despite numerous advancements, however, autonomous vehicles remain unreliable in adverse weather conditions including rain, fog, snow, and darkness. Current methods of scene mapping rely on LiDAR technology, which tends to fail when exposed to weather that distorts its signals. This issue emphasizes the need for robust image-based systems for suboptimal driving conditions. Two models, CLIP and ResNet, were selected, both of which offer unique approaches toward image-based weather classification compared to conventional models such as convolutional neural networks (CNNs). Using a self-compiled custom dataset of 4800 road images in varying conditions, each model was trained and tested. ResNet-50 was the most effective model, reaching an accuracy of 0.95 on the testing data set. The predictions of road surface weather from the models were compared to empirically determined data to estimate the coefficient of friction, which can be used to maximize safety. Furthermore, an image restoration model—a model that removes weather effects such as raindrops—was analyzed and its performance was measured quantitatively. Algorithmically measuring the confidence of object detection numerically showed the improvement from the original images to the restored images, serving as a reliable evaluation technique. This concept could be used as an optimizer for these models to maximize their performance. Overall, this research reveals the potency of previously unused techniques for the development of autonomous driving and serves as a foundation for future developments involving adverse weather.

Keywords: Autonomous Vehicle Safety; Image Classification; Image Restoration; Object Detection

1. Introduction

The rapid advancement of technology, particularly in artificial intelligence (AI), has fostered the diversification of various sectors, including transportation. Though the development of autonomous automobiles has promised improved safety, accessibility, and efficiency, there remain glaring challenges that need to be addressed in order to make the usage of autonomous vehicles a lifestyle, rather than a liability. One of the most pressing challenges is their vulnerability to adverse weather conditions. According to the US Department of Transportation, approximately 21% of automobile crashes (~1.2 million/year) in the US are weather-related [1]. Meta-analyses show that crash rates climb by ~71% during rain and ~84% during snow [2]. As the world shifts closer towards fully autonomous transportation, it is essential that self-driving vehicles can detect and respond accordingly to adverse weather conditions, such as snow, rain, and fog. Failure to appropriately account for these conditions compromises safety and efficiency, preventing organizations such as the National Highway Traffic Safety Administration (NHTSA) from approving the deployment of autonomous vehicles on public roads [3]. Though millions of dollars and hours have been put into the development of autonomous vehicles, relatively little has been done to prepare these vehicles for abnormal weather conditions.

* Corresponding author: Prayag Gaonkar

Weather-related safety and efficiency concerns will undoubtedly hinder the widespread use of these automobiles as it mitigates its practicality.

This research aims to address key limitations in current vehicle sensing and decision-making frameworks under the unfavorable weather conditions of rain, fog, snow, and darkness. These circumstances require different decisions from the autonomous system, depending on the severity and type of weather—a factor often underrepresented in prior research. Most existing weather-detection models tend to classify weather based on environmental features (e.g., sky conditions) rather than the actual road surface—an approach that can be dangerously misleading. For example, while the sky may appear clear, the road could still be wet and slippery. Additionally, unusual weather circumstances call for more advanced object detection systems due to poor visibility in cameras due to precipitation. With unclear images, autonomous systems will be less aware of the surroundings, potentially compromising safety.

Current autonomous vehicles in production and development mostly rely on LiDAR (Light Detection and Ranging) technology, which does not address either of these two problems. LiDAR systems create high-resolution 3D maps of environments, which are used by the vehicles to make decisions [4]. The technology functions by emitting laser pulses and tracking the time it takes to hit an object and reflect back, allowing it to determine the shape and distance of objects around it [4]. However, precipitation and harsh weather involving raindrops and snow distort the lasers emitted and thus significantly interfere with the resulting 3D map produced, preventing the vehicle from sensing the objects around it [5]. Finally, LiDAR is unable to determine the surface condition of the road (e.g., dry, wet, or icy) and therefore does not provide an autonomous system with sufficient information to appropriately adjust vehicle speed in response to reduced traction and friction [6]. LiDAR's struggles in poor weather result in the need and increased reliance on image-based systems for autonomous vehicle development.

One of the primary objectives of this study is to develop a robust image-based classification model capable of distinguishing the five road conditions (rain, fog, snow, clear, night) with high accuracy. To improve upon past work, the models were tailored to two key aspects that were previously overlooked:

- Specificity: the ability of the model to only consider the condition of the road and not external factors.
- Adaptability: the models should be robust for nearly any scenario in any location, so they should not be overfit to any particular trait.

Once the most effective model is developed and tested, the outputs of the classifier can be integrated with physical models of tire–road interaction: by applying empirically determined coefficients of friction for each road surface category, the system could compute and recommend proportional reductions in vehicle speed to maintain safe stopping distances. These adaptive recommendations could be sent to the longitudinal controller of an autonomous vehicle to make real-time adjustments. In this study, two major groups of models were trained and tested to classify road surface condition, namely ResNet and CLIP.

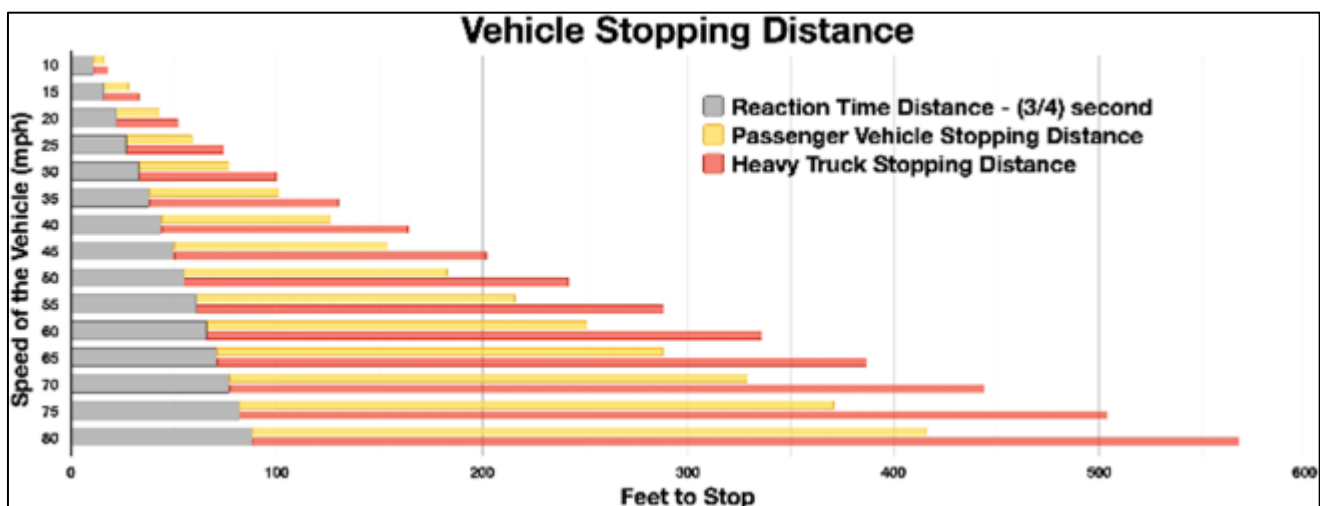


Figure 1 Chart of vehicle stopping distance based on speed; fails to consider surface of road and its friction coefficient

To support the development of these models, a diverse and curated dataset of road weather images was compiled from various publicly available sources. However, despite the widespread research and data availability in the field of

autonomous vehicles, there is little publicly available data in the form of images of weather conditions specifically tailored to vehicles or roads. Therefore, the data used to train the models was collected and combined from a variety of sources, possibly leading to biases in the classification of the images. For example, one dataset might classify a road during twilight time as normal lighting, while another might classify a similar image as nighttime. The limitation of having a lack of standardized and reliable data will reduce model accuracy and will make it less robust to new, unfamiliar weather images and scenes. However, to mitigate geographic and infrastructural bias, the dataset incorporates samples from multiple regions, encompassing a wide range of lighting, weather intensities, and road environments. Moreover, the compiled dataset was hand-filtered to ensure that there are no major mis categorizations. The data compiled from the different sources was shuffled and used to train and test the various models for image-based weather detection.

The other primary objective of this study is to establish a standardized evaluation framework for existing image-cleaning (de-weathering) models. These models aim to restore scene visibility and enhance the clarity of images captured by vehicle-mounted cameras, thereby improving downstream tasks such as object detection, lane following, and semantic segmentation. The lack of consistent benchmarking criteria makes it difficult to compare models objectively or identify the most effective methods for specific weather conditions. The proposed evaluation framework seeks to address this gap by introducing a quantitative metric that assesses not only image restoration quality (e.g., PSNR, SSIM, LPIPS) [7,8], but also task-specific performance, such as object detection accuracy on restored frames. In addition to providing a fair basis for comparison, the framework will be designed as an optimization utility for researchers developing new de-weathering models. By integrating the evaluation algorithm with automated hyperparameter tuning tools (e.g., grid search, Bayesian optimization, or reinforcement learning-based tuning), developers can iteratively adjust network architecture and training parameters to maximize real-world performance under targeted environmental conditions. This approach supports both model-level refinement and broader research progress by enabling reproducible, data-driven improvement of perception systems in autonomous vehicles operating under visually degraded conditions.

Together, these proposed frameworks of comprehensive object detection and context-aware speed adaptation will equip autonomous platforms with the necessary means to perceive and respond to challenging weather scenarios more reliably, leading to safer and more efficient driving behaviors.

2. Methods

The first component of this research involves a comparative analysis of underutilized models for image-based weather classification. Although originally designed for general image classification tasks, these models were reprogrammed and optimized to address the specific requirements of weather classification. The second component is exploratory in nature, aiming to establish a method for evaluating image restoration models, given the lack of standardized quantitative performance metrics. Developing a reliable evaluation technique is essential for the effective optimization and practical deployment of both classification and restoration models.

2.1. Data collection and image processing

Sufficient and usable data is critical to maximizing the efficiency and accuracy of any machine learning model. However, at the time of research, there was no publicly available dataset of road images with image labels corresponding to varying weather or visibility conditions. As a result, it was key to filter through and preprocess the images in the dataset to maximize the performance of the models.

- **Dataset Creation:** Compiled a diverse dataset of 4,800 images (1000 images for the normal, fog, and night classes; 900 images for the rain and snow classes) by collecting images from a variety of open-source databases. Images that did not fit with the intended classes (i.e. dust storm) or lacked usability (i.e. hard for even a human to classify) were removed from the dataset and replaced with another. This technique of image cleaning is important to improve model prediction accuracy.

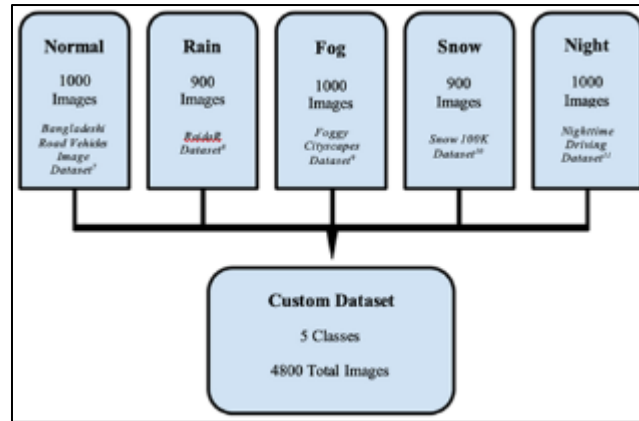


Figure 2 Distribution of Images in Custom Dataset [16,17,18,19,20]

- **Data Segmentation:** The dataset was programmatically shuffled and then split into a 85:15 ratio of training to testing images using the `train_test_split()` function of the Scikit-learn Python library. The resulting split datasets were saved to separate JSON files to ensure that each model was trained and tested on the same sets of images, minimizing bias.
- **Image Preprocessing:** The images were normalized to the same shape, a term that generally refers to three components of an image: the height, width, and channels. All images for standard neural networks were converted to (256, 256, 3), which indicates a 256x256 pixel ratio and three channels, converting all images to RGB. Standardizing the images is critical for the models to extract features and patterns.

2.2. Weather model selection and implementation

2.2.1. CLIP (Contrastive Language-Image Pretraining):

CLIP is a pretrained neural network (trained on large image-caption datasets using a contrastive objective that maximizes the similarity between correct image-text pairs while minimizing the similarity of mismatched pairs) developed by OpenAI that learns visual distinctions and similarities in an unsupervised manner [21]. It combines Natural Language Processing (NLP) techniques with traditional Computer Vision (CV) techniques to make class names/descriptors hold significance in the prediction algorithms. CLIP is a dual-encoder multimodal model that uses one encoder for images (e.g., a Vision Transformer or ResNet) and one for text (a Transformer-based language model), trained jointly to align image and text embeddings the shared latent space [21]. This enables “zero-shot” capabilities, recognizing objects or concepts without task-specific training based on descriptive text prompts. Given an image and a set of text descriptions, CLIP can output a similarity score of each description to the image without any additional training [21].

CLIP’s versatility compared to standard supervised models makes it extremely suitable for autonomous vehicles and weather classification, as it is essential that a model can adapt to new, unforeseen circumstances. To improve performance from the pre-trained version, the CLIP model was fine-tuned by training it on the custom dataset that was created. The images were preprocessed using the CLIP ViT-B/32 Preprocessor (an image encoder) and set to 512 “features”, the dimensionality of each image embedding. In other words, this is a tensor of 512 numbers that correspond to a specific image, which is produced from a series of manipulations to the original multidimensional vector representing the image. Then, these image embeddings are compared to the text embeddings (also a 1x512 tensor) of descriptors of each of the five classes (normal, rain, snow, fog, night). An integer value one to five is outputted, corresponding to the class with the text embedding tensor that has the highest similarity to the image embedding tensor [21]. This value is compared to the correct image label, and the CLIP model then receives feedback regarding the correctness of its prediction.

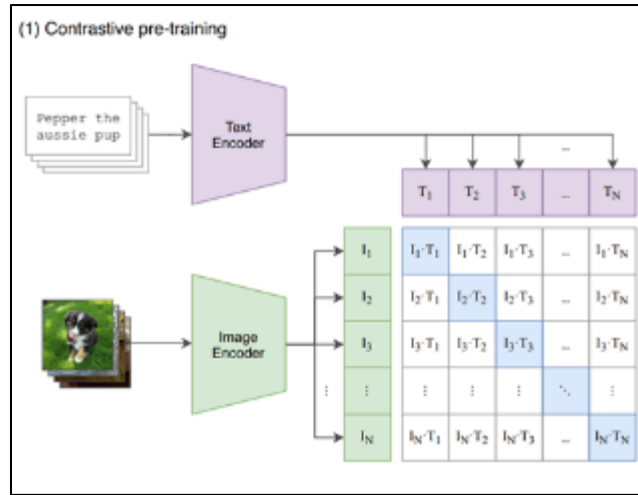


Figure 3 Diagram of CLIP's Shared Text and Image Embedding Space

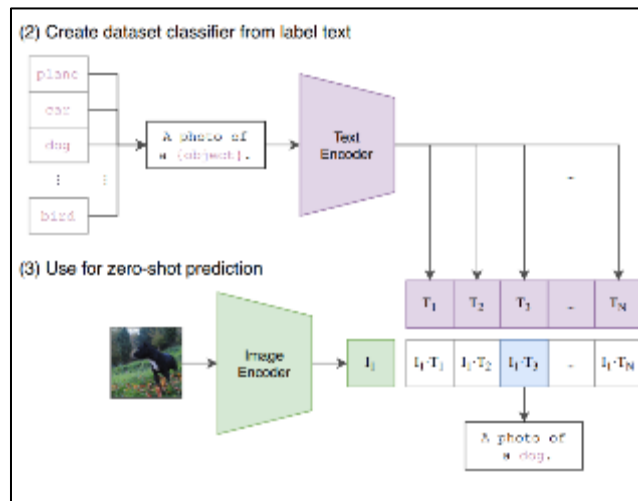


Figure 4 Methodology of CLIP's Zero-shot Prediction Capabilities

2.2.2. ResNet (Residual Network):

ResNet is a type of Convolutional Neural Network (CNN) and is pretrained on the ImageNet dataset, with over 14 million images. CNNs are especially effective with image classification but tend to struggle with two key problems.

- **Vanishing Gradients:** Gradients, which are indicators of a model's loss, become very small as layers to the network are added, resulting in little to no learning.

During backpropagation, small numbers outputted from functions are multiplied to the gradients and cause their value to diminish over time. After a certain point, the gradient value stops changing, causing the optimizer to stop learning, since there is no change in loss [22].

- **Degrading Accuracy:** Contrary to intuition, many CNNs have lower accuracy when there are too many layers in the model. This problem is partially caused by overfitting, the concept that an excess number of parameters causes the network to simply “memorize” the training data rather than extracting patterns, resulting in poor performance on new data. Having a surplus of layers can also hurt optimization efficiency [22].

ResNet models, however, tackle both problems. The key mechanism of residual networks is that they learn to compute the change between the input (image) and output (road condition) [22], making it more efficient than in a normal CNN, where the model must learn how to produce the entire output from a given input. Therefore, in contrast to regular CNNs, residual networks have layers that can add directly to the final output by using residual (skip) connections, the ability

to bypass layers to change the output [22]. In a ResNet model, each layer can contribute to the difference (residual) individually, but in a normal CNN, layers “work together” to produce the output. This feature ensures that each layer is learning and directly contributing to the output, preventing any override by another layer down the line [22]. In other words, each layer is independent, so having a sophisticated model with numerous layers will not cause either of the two problems mentioned above.

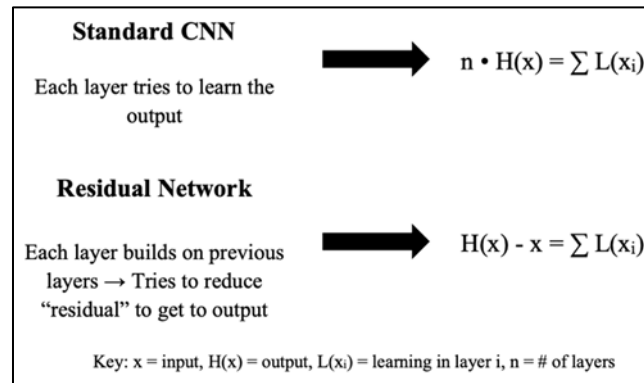


Figure 5 Structural Differences Between Residual Networks and Standard CNN Models

2.2.3. Model learning and evaluation

Loss functions provide quantitative data regarding a model’s performance, guiding the optimizer to improve predictions by minimizing the difference between true values and predictions. The cross-entropy loss function is particularly strong with probabilistic models and is gradient-friendly, making it suitable for both CLIP and ResNet models. Furthermore, the cross-entropy loss function is strong with nonbinary classification, which is required for this task [23]. Specifically, the cross-entropy loss function is defined by

$$H(p, q) = - \sum q_k \log(p_k),$$

where q_k is the one-hot encoded true label and p_k is the predicted probability from the SoftMax output [23]. Minimizing cross-entropy encourages the model to assign high probability to the correct class and is especially effective in deep networks due to its smooth, well-behaved gradients.

The cross-entropy loss function was paired with the AdamW optimizer, which has many key features tailored to sophisticated, multi-layer image classification in comparison to the standard Adam optimizer [24]. For instance, the AdamW features decoupled weight decay, which prevents weights from growing too large, thereby improving generalization [24]. Since it is “decoupled” from the optimization step, it can compute adaptive learning rates which helps minimize overfitting [24].

The models were evaluated primarily on two metrics, accuracy and loss. Each metric was measured individually for each model on both the testing and training sets. The values were computed at regular intervals of each epoch (one epoch is one complete iteration of the training data through a model), for a total of ten measurements (ten epochs were run for each model). Then, the models were compared and ranked for accuracy based on the validation data set of 720 images.

2.2.4. Mapping speed control to road surface condition

To ensure safe and adaptive driving under adverse weather conditions, the output of a road weather classification model is integrated with a physics-based framework for speed control, which could potentially be used in a longitudinal controller in an autonomous vehicle. Each predicted weather class (normal, rain, snow, fog, night) is mapped to an estimated coefficient of friction (μ) based on empirical values reported in transportation safety literature.

Table 1 Mapping of Road Surface Conditions to Plausible Ranges for the coefficient of friction (μ)

Road Surface Condition	Additional Driving Info	Coefficient of Friction (μ)
Normal (Dry)	Asphalt/Concrete	0.75-0.85
Rain	Wet/Slick Road	0.25-0.50
Snow/Ice	Covered	0.05-0.25
Fog	Lower Visibility	0.60-0.75*
Night	Lower Visibility	0.65-0.80*

*Friction is the same as normal conditions for fog and night, but the stopping distance is higher due to reduced visibility. For calculation, μ is reduced proportionally.

Based on the estimated friction coefficients, the system can compute and recommend proportional reductions in vehicle speed to maintain safe stopping distances by the equation

$$d = \frac{v^2}{2\mu g},$$

where d is the braking distance, v is the vehicle speed, μ is the friction coefficient, and g is the gravitational constant. Ideally, additional factors, such as turning radius (if the vehicle is turning) and tire condition (wear and age), should be considered by the system for speed adjustments to maximize safety whilst maintaining efficiency.

2.2.5. Evaluating image restoration model performance

raindrops and snow, from images. However, currently, there are no proven methods to quantitatively measure the performance of these models because it is impossible to produce images in the exact same scenes without the poor weather that the restored images can be compared to. Therefore, it is critical to go back to the original purpose of the de-weathering model, which is to make scene recognition from images more accurate. In other words, the de-weathering model's purpose was to improve object detection of images taken with "weathered" lenses.

An image restoration model for weather-affected images by Zhu et al. [25] was implemented to evaluate. Since the model was pretrained for rain, snow, and haze (fog) images, images from the validation dataset classified into those categories by the ResNet-50 model were used to evaluate the de-weathering model. It was impractical to de-weather and compare all images in the dataset due to memory and GPU constraints. After implementing and running Zhu et al.'s model, the de-weathered (restored) images were paired with the original images.

To evaluate the effectiveness of image restoration models in the context of autonomous driving under adverse weather conditions, an object detection-based performance assessment framework was developed. The DetrForObjectDetection model from the HuggingFace transformers library [26] was implemented due to its capability to output a numerical confidence score for each detected object. This confidence score, a continuous value ranging from 0 to 1, represents the model's certainty in correctly identifying and localizing an object within the scene. This feature makes DETR particularly suitable for evaluating restoration quality, as it provides quantifiable insight into the impact of de-weathering processes on downstream perception tasks.

The evaluation technique consists of applying the object detection model to both the original degraded image (e.g., containing rain, snow, fog) and its corresponding restored image from the de-weathering model. For each detected object, the confidence score generated by the DETR model on the restored image was compared to the score obtained from the original weather-degraded image. These differences in confidence values were then averaged across all detected objects within a single scene to produce an overall net confidence gain for a particular image. Repeating this procedure across a large set of restored and original image pairs produced a statistically meaningful metric of performance attributable to the restoration model, which could be used for further development or analysis.

Similar to the training and fine-tuning of weather classification models using adaptive optimizers like AdamW, the proposed evaluation framework also supports the iterative improvement of image restoration models. By using the average confidence gain from object detection as a feedback signal, these models can be adjusted more effectively to enhance both visual clarity and object detectability, both of which are essential for the safe and reliable operation of autonomous vehicles in challenging weather conditions.

3. Results

3.1. Weather classification

Four separate models—CLIP, ResNet-34, ResNet-50, and ResNet-101—were trained and tested on the same custom dataset of approximately 4800 road images. Ten epochs were run on each model, producing final validation accuracies as follows: CLIP (0.79), ResNet-34 (0.92), ResNet-50 (0.95), and ResNet-101 (0.93). The loss after each epoch (the complete pass of the training data through the model) was also tracked, as shown in the graphs below. The data below shows the performance of the models under a batch size of 64 and base learning rate of 0.001, both of which seem to be best for the given conditions.

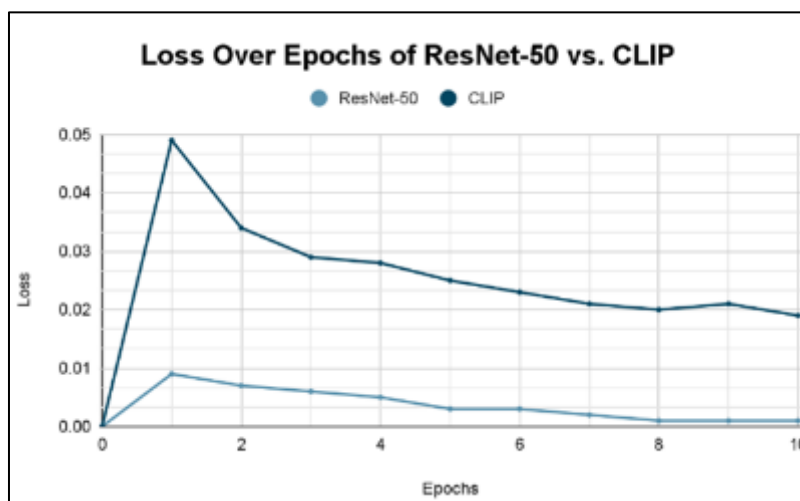


Figure 6 Performance (accuracy) of CLIP vs. ResNet-50 during training on validation data set (720 images)

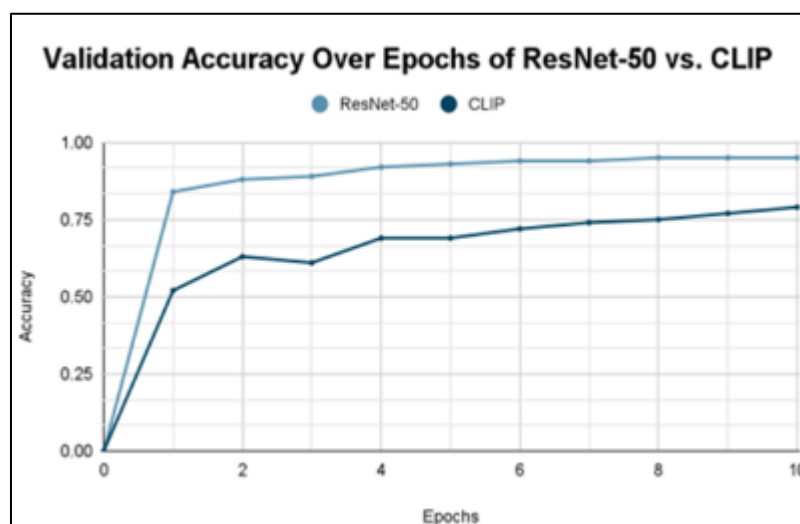


Figure 7 Gradual optimization of CLIP vs. ResNet-50 through loss function over iterations of training data (4080 images)

Both models, ResNet and CLIP, experienced a significant increase in accuracy and a decrease in loss as more epochs were run. This indicates that the models were learning from the training data, showing the potency of using a custom dataset to refine these models. Although CLIP significantly underperformed in comparison to ResNet-50, it still could be practical to use in the future. Unlike ResNet-50, CLIP's accuracy did not plateau after ten epochs and steadily continued to increase. Thus, with more data and epochs run, it is likely that CLIP will be able to achieve performance similar to ResNet, if not better. Testing on both the AdamW and Adam optimizers, AdamW outperformed the latter (data for models above) showing better standardization and consistency.

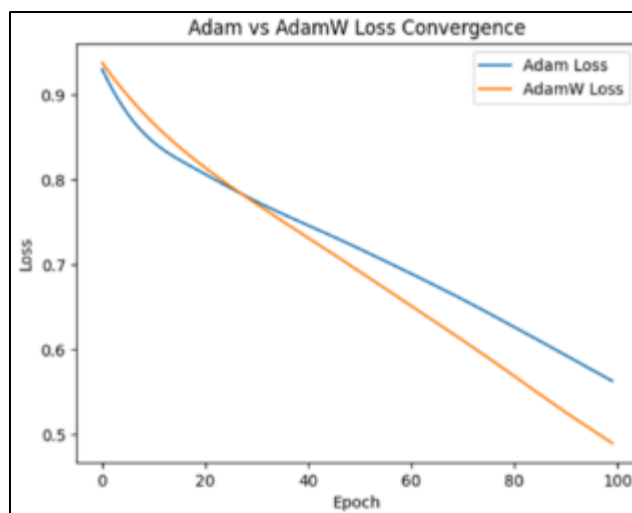


Figure 8 Comparison of loss between AdamW and Adam optimizers; AdamW better for generalization due to standardization and adaptability

3.2. Image restoration (de-weathering) analysis classification

To determine and evaluate the performance of Zhu et al.'s de-weathering model, the confidence in object detection with the original and restored images was measured and compared. Averaging these values produced a proportion that represents the effectiveness of the model. An example of this process for a single image is shown below.

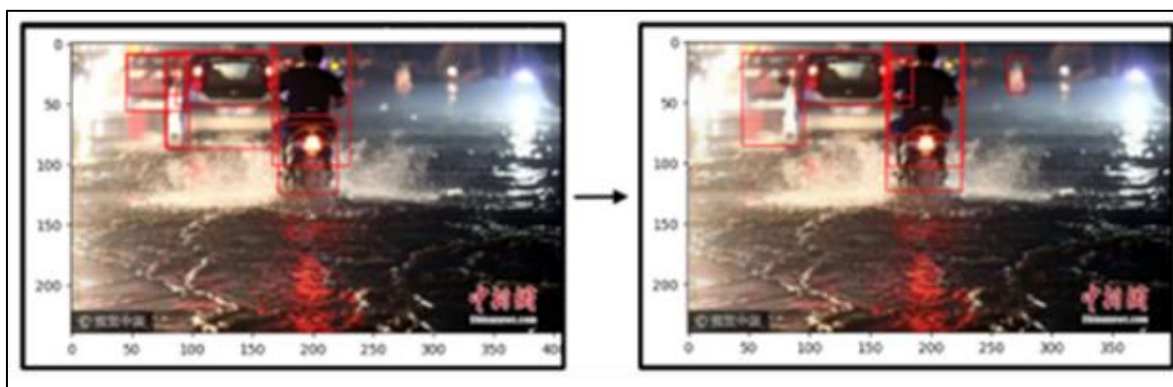


Figure 9 Effect of Zhu et al.'s Image Restoration Model on Object Detection: Rain (527)

Table 2 Difference in confidence of object detection between original and restored image: Rain (527)

Object Number	Original Image (O)	Restored Image (R)	Difference (R - O)
1	0.848	0.856	+0.008
2	0.890	0.967	+0.077
3	0.710	0.784	+0.074
4	0.879	0.901	+0.022
5	0.814	0.966	+0.152
6	0.995	0.982	-0.013
7	0.984	0.994	+0.010
8	0.654	0.766	+0.112
Average	0.847	0.902	0.055

For this image in particular, Rain (527), the restoration model tends to improve object detection by an average of 5.5% confidence for each object. This process was repeated algorithmically for each rain/fog/snow image in the validation dataset, yielding an average improvement in confidence of 0.063 (6.3%) for each object. This indicates that Zhu's model can undoubtedly help in poor weather, but further optimization needs to be done to maximize its performance. Perhaps this algorithm can be modified into an optimizer for the model.

4. Discussion

To summarize, ResNet models tended to perform better than the CLIP model with weather image classification. In particular, the ResNet-50 model performs best among the ResNet models for training time, feature extraction, and performance, reaching an accuracy of 0.95 for the testing data. Although the CLIP model performed significantly worse than ResNet models, attaining an accuracy of just 0.79, this approach is still very promising. Since CLIP can compare images to text-based descriptors [6], it is far more versatile than other models, which is critical for tasks such as autonomous driving, where the system must be able to adapt to unseen and unexpected situations.

Nevertheless, predictions from either of the models should be used to adjust the style of driving from the vehicle. For example, in rainy conditions, the vehicle should slow down but avoid sharp turns and sudden braking, either of which could lead to hydroplaning. Similarly, in icy or snowy conditions, the vehicle might tell the passenger that chains are required for the tires to ensure safety. The performance and usability of both models not only met the objectives for this research but also helped fill a critical gap in current-day autonomous vehicles: driving in harsh, unsafe weather. In particular, previous work has not been focused on the conditions of the roads themselves, which are often different from the conditions of the rest of the environment. Having refined, specific models tailored to this task will greatly improve the safety of autonomous vehicles. However, the predictions by the models are limited by the severity of each weather condition. For instance, a light drizzle and pouring rain would be filtered into the same class of rainy weather, which is not ideal because the decisions that a vehicle would have to make in these conditions are significantly different. In the future, researchers can build upon these findings by developing a numerical scale to analogous to the severity of weather. Using this scale, a vehicle can make adjustments proportionally.

Another major piece of this research was finding a way to measure the performance of image de-weathering models. Without a reliable technique to do so, these models will be rendered as unusable because there will be no systematic way to evaluate them. To solve this, an object detection model that could determine the "confidence" of recognizing an object in a scene was integrated into the de-weathering model. Then, the confidence (in object detection) of the original and de-weathered images was compared, which shows the improvement in clarity caused by the de-weathering model. This method can be used by researchers to create an optimizer/loss function, allowing the model to attain maximum performance by having data to learn from. However, this approach for optimization requires significant computing power that is not sustainable for a regular computer on a whole dataset. Thus, an object detection optimizer will require considerable funding to implement in the model.

Regardless, image de-weathering capabilities are essential for autonomous driving because the presence of water or snow on image lenses is an issue that is bound to occur during adverse weather. Since LiDAR has proven to be minimally effective in conditions interfering with its signals, it is paramount that image-based systems can handle these conditions appropriately. Optimized de-weathering techniques solve this issue, allowing autonomous vehicles to make decisions with as much data as possible.

Overall, although various autonomous driving features, such as auto lane-changes, have already been deployed in commercial vehicles, the primary factor holding back the widespread usage of completely autonomous vehicles is their inability to adapt to unanticipated scenarios. This research hopes to tackle the key obstacle behind weather-related conditions in order to improve the most important aspect of any vehicle: safety.

5. Conclusion

The main goal of this study was to advance autonomous vehicle safety by improving their performance in adverse weather, such as rain and snow. In these conditions, precipitation either distorts the signals of sensors or blurs cameras, resulting in inadequate data for decision-making. Using the ResNet-50 model for weather classification enables researchers and autonomous systems to reliably adjust speed and turning based on mapped friction values. Thus, issues such as hydroplaning and skidding will be avoided, resulting in safer driving in suboptimal conditions.

The algorithm to evaluate image restoration models holds significant value for future research. By converting qualitative improvements in image quality to quantitative metrics, the algorithm can be used to optimize and refine existing image restoration models. Thus, self-driving vehicles will be better able to identify objects, such as pedestrians or traffic signs, resulting in faster decisions even when visibility is low.

Currently, this camera-based system can be used if integrated with a dashcam to advise drivers to adjust their speed or warn for unseen objects. However, it cannot be fully implemented into an autonomous vehicle without further testing and approval from governing bodies. This research serves as the foundation for future weather-based safety developments in autonomous vehicles by offering new perspectives using cameras instead of LiDAR sensors as the primary navigation tool.

Compliance with ethical standards

Acknowledgments

I would like to express my immense gratitude to Akshay Gopalkrishnan, a master's student at University of California San Diego (UCSD), for his invaluable guidance and support through the course of this research. His experience and knowledge have contributed to this project.

References

- [1] US Department of Transportation. How Do Weather Events Impact Roads? https://ops.fhwa.dot.gov/weather/q1_roadimpact.htm (2024).
- [2] N. Becker, H.W. Rust, U. Ulbrich. Weather impacts on various types of road crashes: a quantitative analysis using generalized additive models. <https://etrr.springeropen.com/articles/10.1186/s12544-022-00561-2> (2022).
- [3] National Highway Traffic Safety Administration (NHTSA). Automated Driving Systems. <https://www.nhtsa.gov/vehicle-manufacturers/automated-driving-systems>.
- [4] Synopsys. What is LiDAR and How Does it Work? <https://www.synopsys.com/glossary/what-is-lidar.html>.
- [5] L. Moussy. Is LiDAR compatible with rainy or foggy weather? <https://www.yellowscan.com/knowledge/is-lidar-compatible-with-rainy-or-foggy-weather> (2021).
- [6] E. Gent. Camera Crushes Lidar, Claims Startup. <https://spectrum.ieee.org/camera-crushes-lidar> (2023).
- [7] J. Li, H. Wang, Y. Li, H. Zhang. A Comprehensive Review of Image Restoration Research Based on Diffusion Models. <https://www.mdpi.com/2227-7390/13/13/2079> (2025).
- [8] Y. Ma, M. Wang, Q. Feng, Z. He, M. Tian. Current Non-Contact Road Surface Condition Detection Schemes and Technical Challenges. <https://www.mdpi.com/1424-8220/22/24/9583> (2022).
- [9] J. Casselgren, S. Rosendahl, M. Sjö Dahl, P. Jonsson. Road condition analysis using NIR illumination and compensating for surrounding light. <https://www.sciencedirect.com/science/article/abs/pii/S014381661500192X> (2016).
- [10] T. Wang, X. Yang, K. Xu, S. Chen, Q. Zhang, R. Lau. Spatial Attentive Single-Image Deraining With a High Quality Real Rain Dataset. https://openaccess.thecvf.com/content_CVPR_2019/html/Wang_Spatial_Attentive_Single-Image_Deraining_With_a_High_Quality_Real_Rain_CVPR_2019_paper.html (2019).
- [11] L. Rahadiani, A.Y. Azizah, H. Deborah. Evaluation of the quality indicators in dehazed images: Color, contrast, naturalness, and visual pleasingness. [https://www.cell.com/heliyon/fulltext/S2405-8440\(21\)02141-1](https://www.cell.com/heliyon/fulltext/S2405-8440(21)02141-1) (2021).
- [12] R. Zhang, P. Isola, A.A. Efros, E. Shechtman, O. Wang. The Unreasonable Effectiveness of Deep Features as a Perceptual Metric. https://openaccess.thecvf.com/content_cvpr_2018/html/Zhang_The_Unreasonable_Effectiveness_CVPR_2018_paper.html (2018).
- [13] Y. Fan, Y. Wang, M. Wei, F.L. Wang, H. Xie. FriendNet: Detection-Friendly Dehazing Network. <https://arxiv.org/html/2403.04443v1> (2024).
- [14] S. Sun, W. Ren, T. Wang, X. Cao. Rethinking Image Restoration for Object Detection. https://proceedings.neurips.cc/paper_files/paper/2022/file/1cac8326ce3f8e79171db9754211530c-Paper-Conference.pdf (2022).

- [15] H. Zhang, Y. Ba, E. Yang, V. Mehra, B. Gella, A. Suzuki, A. Pfahnl, C. Chandrappa, A. Wong, A. Kadambi. WeatherStream: Light Transport Automation of Single Image Deweathering. https://openaccess.thecvf.com/content/CVPR2023/papers/Zhang_WeatherStream_Light_Transport_Automation_of_Single_Image_Deweathering_CVPR_2023_paper.pdf (2023).
- [16] A. Yeafi. Road Vehicle Images Dataset. <https://www.kaggle.com/datasets/ashfakyeafi/road-vehicle-images-dataset> (2023).
- [17] J. Jin, A. Fatemi, W. Lira, F. Yu, B. Leng, R. Ma, A. Mahdavi-Amiri, H. Zhang. RaidaR: A Rich Annotated Image Dataset of Rainy Street Scenes. <https://raidar-dataset.com/download> (2021).
- [18] C. Sakaridis, D. Dai, L. V. Gool. Semantic Foggy Scene Understanding with Synthetic Data. <https://paperswithcode.com/paper/semantic-foggy-scene-understanding-with> (2017).
- [19] Y. Liu, D. Jaw, S. Huang, J. Hwang. Snow100K Dataset. <https://paperswithcode.com/dataset/snow100k> (2017).
- [20] D. Dai, L. V. Gool. Dark Model Adaptation: Semantic Image Segmentation from Daytime to Nighttime. <https://paperswithcode.com/paper/dark-model-adaptation-semantic-image> (2018).
- [21] A. Radford, I. Sutskever, J. W. Kim, G. Krueger, S. Agarwal. CLIP: Connecting text and images. <https://openai.com/index/clip/> (2021).
- [22] K. He, X. Zhang, S. Ren, J. Sun. Deep Residual Learning for Image Recognition. <https://arxiv.org/abs/1512.03385v1> (2015).
- [23] C. Hughes. A Brief Overview of Cross Entropy Loss. <https://medium.com/@chris.p.hughes10/a-brief-overview-of-cross-entropy-loss-523aa56b75d5> (2024).
- [24] F. M. Graetz. Why AdamW matters. <https://towardsdatascience.com/why-adamw-matters-736223f31b5d> (2018).
- [25] Y. Zhu, T. Wang, X. Fu, X. Yang, X. Guo, J. Dai, Y. Qiao, X. Hu. Learning Weather-General and Weather-Specific Features for Image Restoration Under Multiple Adverse Weather Conditions. https://openaccess.thecvf.com/content/CVPR2023/papers/Zhu_Learning_Weather-General_and_Weather-Specific_Features_for_Image_Restoration_Under_Multiple_CVPR_2023_paper.pdf (2023).
- [26] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, S. Zagoruyko. End-to-End Object Detection with Transformers. <https://arxiv.org/abs/2005.12872> (2020).