Enhancing Traffic Counting in Rainy Conditions: A Deep Learning Super Sampling and Multi-ROI Pixel Area Approach

Elly Warni
Department of Informatics, Faculty of Engineering, Hasanuddin University, Indonesia
elly@unhas.ac.id (corresponding author)

A. Ais Prayogi Alimuddin
Department of Informatics, Faculty of Engineering, Hasanuddin University, Indonesia
aisprayogi@unhas.ac.id

A. Ejah Umraeni Salam
Department of Electrical Engineering, Faculty of Engineering, Hasanuddin University, Indonesia
ejah@unhas.ac.id

Moch Fachri
Department of Informatics, Faculty of Engineering, Krisnadwipayana University, Indonesia
moch.fachri@unkris.ac.id

Muhammad Rizal H.
Department of Informatics Engineering, Universitas Teknologi Akba Makassar, Indonesia
rizal@unitama.ac.id

Received: 6 November 2024 | Revised: 12 December 2024 | Accepted: 22 December 2024

 DOI: https://doi.org/10.48084/etasr.9515

ABSTRACT

In Intelligent Transportation Systems (ITS), adaptive traffic control relies heavily on precise, real-time traffic data. Controllers use information such as vehicle count, vehicle density, traffic congestion, and intersection wait times to optimize traffic flow and improve efficiency. Traffic cameras collect and process this data, but environmental factors like rain can degrade the performance of data retrieval systems. We propose a vehicle detection method that integrates pixel area analysis with Deep Learning Super Sampling (DLSS) to enhance performance under rainy conditions. Our method achieved an accuracy of 80.95% under rainy conditions, outperforming traditional methods, and performing comparably to specialized methods such as DCGAN (93.57%) and DarkNet53 (87.54%). However, under extreme conditions such as thunderstorms, the method's accuracy dropped to 36.58%, highlighting the need for further improvements. These results, evaluated using the AAU RainSnow Traffic Surveillance Dataset, demonstrate that our method improves traffic data collection in diverse and challenging weather conditions while identifying areas for future research.

Keywords-deep learning super sampling; digital image processing; intelligent transportation system; pixel area; traffic counter

I. INTRODUCTION

Contemporary urbanization and the resulting increase in vehicular traffic have posed significant challenges to city planners and traffic management bodies, necessitating the evolution of Intelligent Transportation Systems (ITS) to improve traffic flow and alleviate congestion through the use of cutting-edge technologies [1]. Real-time traffic data collection and analysis are essential for ITS to facilitate adaptive traffic control mechanisms. These data include metrics such as the number of vehicles passing a point, vehicle density, congestion levels, and intersection waiting times [2]. Traffic cameras are essential for collecting these data. These devices collect real-time traffic information by capturing and processing footage [3] The quality of the captured images has a significant impact on the performance of traffic data retrieval systems. Rain, fog, snow, and varying lighting conditions adversely affect traffic monitoring and vehicle detection accuracy by introducing noise and degrading image quality [4].

Machine learning and deep learning have led to substantial progress in the development of traffic monitoring and vehicle detection systems. Traditionally, traffic monitoring relied on basic image processing methods such as edge detection, background subtraction, and optical flow analysis [5]. These methods paved the way for vehicle detection and traffic flow analysis, revealing important essential traffic patterns and congestion levels. In recent years, more advanced traffic monitoring systems have emerged. Convolutional Neural Networks (CNNs) have become a cornerstone in this field [6], with models such as You Only Look Once (YOLO) [7] and Single Shot MultiBox Detector (SSD) [8, 9] achieving high accuracy and real-time performance in vehicle detection tasks. These models can swiftly and accurately identify and classify vehicles from traffic camera video frames, but typically work best when image quality is not affected by external factors such as weather and lighting changes. The integration of multi-camera systems and sensor fusion represents a significant advancement in traffic monitoring technology [10]. These methods involve fusing data from multiple cameras or multiple sensors, like LiDAR and radar, to generate a complete understanding of traffic conditions. Overlapping fields of view from multiple cameras improve the precision of vehicle detection and monitoring. Sensor fusion combines the strengths of various sensing modalities for improved robustness and reliability. Although these methods significantly enhance traffic data collection, their intricate installation and increased cost hinder their widespread implementation. In addition, advancements in image enhancement techniques have contributed to improved performance in traffic monitoring systems. Techniques such as histogram equalization, contrast enhancement, and super-resolution have been applied to improve the visibility and detail of traffic images [11]. These methods improve the image quality of traffic cameras, allowing for more precise vehicle detection. Adverse environmental conditions often challenge the effectiveness of image enhancement techniques, which typically do not address these specific issues. Many existing traffic monitoring methods overlook the effects of non-ideal environmental conditions. Deep learning models and image processing techniques are usually designed under the assumption of clear visibility and stable lighting. This assumption reduces the effectiveness of these methods when dealing with real-world conditions such as rain, fog, snow, and varying lighting.

Our work fills a critical gap in the current literature: the integration of deep learning-based image enhancement techniques with specialized traffic analysis methods that can handle the challenges posed by non-ideal weather conditions. Specifically, the proposed method combines Deep Learning Super Sampling (DLSS) with multi-Region of Interest (multi-ROI) pixel area analysis to enhance vehicle detection in rainy conditions. Our approach not only improves the visual quality of traffic camera footage but also ensures reliable vehicle counting and identification in complex traffic scenarios, even under adverse environmental conditions. By explicitly addressing this gap, our work contributes a novel solution that enhances the robustness and accuracy of traffic monitoring systems under real-world conditions, where traditional methods often fail. DLSS has emerged as a promising method for image enhancement in the context of computer graphics and gaming [12]. DLSS uses deep learning to increase the resolution of low-resolution images, resulting in enhanced image detail and clarity. Its potential for improving traffic camera footage under non-ideal conditions is increasingly recognized, especially in real-time rendering applications. Studies suggest that DLSS significantly improves the accuracy and reliability of vehicle detection by reducing the influence of environmental noise [13]. The concept of multi-ROI analysis has been utilized in diverse image processing tasks such as medical imaging and remote sensing [14]. Multi-ROI methods help narrow the focus of traffic monitoring analysis by selectively zeroing in on areas of interest within images. This technique is most effective in intricate scenarios with numerous lanes and irregular traffic flows. Despite recent advances, there is limited research on the application of DLSS and multi-ROI pixel area analysis to traffic monitoring. Previous research has primarily addressed the issues of improving image quality or developing advanced detection algorithms separately. Our method, which integrates DLSS with multi-ROI analysis, aims to address the challenge of vehicle detection in non-ideal environmental conditions by offering a more robust solution.

This study introduces a novel method for traffic counting under adverse weather conditions, specifically focusing on rainy scenarios where traditional vehicle detection methods face significant challenges. The key contribution of this research lies in the integration of DLSS with advanced image processing techniques, which enhances the image quality of traffic camera footage and improves vehicle detection accuracy even under low visibility caused by rain. DLSS, which has been primarily used in the field of computer graphics, is adapted here for traffic monitoring to address the degradation in image clarity due to environmental factors, such as rain and fog. In addition, we introduce a unique approach that combines DLSS with multi-ROI pixel area analysis. This combination allows for more precise vehicle identification and counting, particularly in complex traffic scenarios where multiple lanes and irregular traffic flows must be considered. By focusing on specific regions of interest within the image, our method can better detect and track vehicles in congested areas, improving the granularity and accuracy of traffic data collected by cameras. Another important contribution is our comprehensive evaluation of the proposed method using the AAU RainSnow Traffic Surveillance Dataset. This dataset provides real-world data under various weather conditions, including rain, which allows us to demonstrate the robustness of our approach in realistic settings. The method not only outperforms traditional vehicle detection techniques under rain conditions, but also shows comparable performance to methods designed specifically for adverse weather conditions, making it a versatile solution for real-world traffic monitoring applications.

II. METHOD

The traffic analysis data processing relies on NVIDIA's tensor cores and is executed on a Graphics Processing Unit (GPU) [15]. The accumulation and averaging of pixel weights from successive input images are the steps employed in the background subtraction algorithm via the accumulate-weighted method [16]. Figure 1 shows the algorithm utilized for the vehicle detection system. The input data are initially processed using background reconstruction and subtraction techniques. For each frame in the input video, the background image is subtracted. Background subtraction is performed using the accumulate-weighted method, where pixel weights from the input digital images are accumulated and averaged over time [17]. The average weights of the input frames are used to acquire the background image. The subtraction process produces an image showing only the objects.

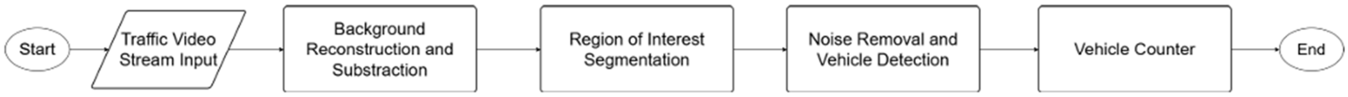

Fig. 1. Proposed method for vehicle detection and counter.

Following background subtraction, ROI segmentation, noise removal, and vehicle detection are performed. Segmentation isolates objects in an image so that their number in a frame can be counted. Otsu's thresholding method, which determines the optimal threshold value based on the intensity histogram of the input image, is employed for image segmentation [18]. DLSS applies super-sampling and deep learning-based denoising to the segmented image to eliminate or reduce weather-induced noise.

The input noisy image $\mathbf{X}$ can be transformed into the clean image $\mathbf{Y}$. Minimize the objective function $(F[h(X)])$ directly using multiple images via (1) with (F) represents.

$$\mathcal{L} = \sum_i \|h(\mathbf{X}_i) - \mathbf{Y}_i\|_F^2 \tag{1}$$

To improve network learning, the mapping range is compressed to reduce the solution space [19]. For simplicity, we assume that the images $\mathbf{X}$ and $\mathbf{Y}$, each containing D pixels, are normalized to the range [0, 1]. When a regression function is applied to map each D pixels, the residual values of the noisy image are generally smaller than those of a clean image within this range. Integrating these residuals can enhance the network's mapping capability. The residual is extracted from the parameter layers for the output, and the skip connection ensures the seamless flow of information across the network, facilitating precise estimation of the denoised image. In this context, noise typically introduces negative differences between the pixel values of $\mathbf{Y}$ and $\mathbf{X}$, often appearing as white streaks. This phenomenon is referred to as negative residual mapping. Equation (2) in the revised objective function reflects this concept.

$$\mathcal{L} = \sum_i \|h(\mathbf{X}_i) + \mathbf{X}_i - \mathbf{Y}_i\|_F^2 \tag{2}$$

We used the ResNet architecture [20], specifically the negative residual mapping method, for its enhanced ability to distinguish noise streaks from object details in images. During training, the detailed layer transmits information to the parameter layers. First, we created a noisy image model using (3).

$$\mathbf{X} = \mathbf{X}_{\text{detail}} + \mathbf{X}_{\text{base}} \tag{3}$$

The base layer can be obtained using low-pass filtering of $\mathbf{X}$, and the detail layer can be obtained by $\mathbf{X}_{\text{detail}} = \mathbf{X} - \mathbf{X}_{\text{base}}$. By subtracting the base layer, only weather noise streaks and object structures remain in the detail layer. We incorporated the detail layer $\mathbf{X}_{\text{detail}}$ and the negative residual mapping $\mathbf{Y} - \mathbf{X}$ into the parameter layers of ResNet. We call the network trained on the detail layer a deep detail network. The denoiser system takes the noisy image $\mathbf{X}$ as its input and generates an approximation of the clean image $\mathbf{Y}$ as its output. In light of our previous discussion, (4) outlines the objective function.

$$\mathcal{L} = \sum_{i=1}^{N} \left\|f\left(\mathbf{X}_{i,\text{detail}}, \mathbf{W}, \mathbf{b}\right) + \mathbf{X}_i - \mathbf{Y}_i\right\|_F^2 \tag{4}$$

ResNet network parameters $\mathbf{W}$ and $\mathbf{b}$ are to be learned from N training images using $f(\cdot)$. We applied guided filtering as a low-pass filter to generate base and detail layers for $\mathbf{X}_{\text{detail}}$ .The basic network structure can be represented as in (5), as the image indexing is eliminated in this network structure.

$$\begin{aligned}
\mathbf{X}_{\text{detail}}^{0} &= \mathbf{X} - \mathbf{X}_{\text{base}} \\
\mathbf{X}_{\text{detail}}^{1} &= \sigma\left(\mathrm{BN}(\mathbf{W}^{1} * \mathbf{X}_{\text{detail}}^{0} + \mathbf{b}^{1})\right) \\
\mathbf{X}_{\text{detail}}^{2l} &= \sigma\left(\mathrm{BN}\left(\mathbf{W}^{2l} * \mathbf{X}_{\text{detatil}}^{2l-1} + \mathbf{b}^{2l}\right)\right) \\
\mathbf{X}_{\text{detail}}^{2l+1} &= \sigma\left(\mathrm{BN}\left(\mathbf{W}^{2l+1} * \mathbf{X}_{\text{detail}}^{2l} + \mathbf{b}^{2l+1}\right)\right) + \mathbf{X}_{\text{detail}}^{2l-1} \\
\mathbf{Y}_{\text{approx}}^{2l} &= \mathrm{BN}(\mathbf{W}^{L} * \mathbf{X}_{\text{detail}}^{L-1} + \mathbf{b}^{L}) + \mathbf{X}
\end{aligned} \tag{5}$$

For $l = 1, \ldots, \frac{L-2}{2}$, the model involves applying batch normalization, convoluting weights and biases $[\mathbf{W}, \mathbf{b}]$, and repeating until the final layer. $\sigma(\cdot)$ is a rectified linear unit for non-linearity, whereas $\mathrm{BN}(\cdot)$ indicates batch normalization to alleviate internal covariate shift. The proposed network retains spatial information by eliminating all pooling operations. For the initial layer, we generate a1 feature maps by implementing filters of size $c \times s1 \times s1 \times a1$, where c is the number of image channels (1 for grayscale, 3 for color). Filters for layers 2 to L-1 have dimensions $a1 \times s2 \times s2 \times a2$. The last layer was defined with filters of size $a2 \times s3 \times s3 \times c$ for the estimation of the negative residual. The denoised image was formed by adding the estimated residual to the noisy image $\mathbf{X}$.

The subsequent step involves vehicle detection within the ROIs. The first process is to calculate the area of white pixels within the ROI resulting from Otsu's thresholding, which represents objects in the denoised image. The pixel count is then compared to a predefined vehicle threshold value. If the threshold is exceeded, it indicates the presence of a vehicle. This method employs multi-ROI to detect vehicles in various lanes, providing better detection accuracy compared to using a single ROI. The results of vehicle detection, the post-DLSS processed images, and the vehicle counts for each lane are stored in a database.

III. RESULTS AND DISCUSSION

We assess our technique by employing the AAU RainSnow Traffic Surveillance Dataset [4]. This dataset consists of traffic camera recording from seven intersections under varying rain conditions. In our approach, vehicle detections in the videos are evaluated using a confusion matrix consisting of True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). We calculate accuracy metrics by comparing the number of true positives and true negatives to the overall detection results. To benchmark the performance of our method, we also evaluate several other methods on the same dataset. Some of these methods, such as pixel area [21], pixel area with multi-ROI [22], CNN [23], and CNN with feature concatenation [24], are not optimized for rain conditions. In addition, we evaluate test methods for handling rain conditions such as DCGAN [4], YOLOv3 [25], YOLO-UA [26], and DarkNet53 [27]. By comparing our proposed method to these existing approaches, we demonstrate its effectiveness in accurately detecting vehicles under challenging weather conditions. Our method outperforms existing methods in detecting objects in rainy traffic surveillance videos. The results of all methods are shown in Table I.

Our method outperformed non-weather specialized methods at all intersections and outperformed specialized methods at certain intersections. It achieved the highest accuracy of 93.44% at the Hobrovej intersection and had the worst result among all specialized methods with a score of 64.87% at the Hasserisvej intersection, but it was still superior to non-specialized methods. At the remaining intersections, satisfactory but not optimal results are achieved. At the Egensevej intersection, YOLOv3, DarkNet53, and DCGAN achieved higher accuracy than our method (81.31%). At the Hadsundvej intersection, our method achieved better performance than YOLOv3, Dark Net53, and DCGAN, although it did not reach the YOLO-UA benchmark of 87.46%. At the Hjorringvej intersection, our method achieved a result of 73.12%, which was only lower compared to DCGAN's 77.39%. At the Ostre intersection, it had a lower accuracy (68.12%) than YOLOv3 (84.59%) and DarkNet53 (77.43%). At the Ringvej intersection, our method achieved a higher accuracy of 92.78%, surpassing all but DarkNet53's 94.06%.

Figure 2 illustrates the challenges posed by adverse weather conditions, such as rain and snow, on traffic camera footage and demonstrates the effectiveness of the proposed method in mitigating these challenges. The upper part of the figure shows raw images captured under rain and snow conditions, highlighting the significant degradation in image quality caused by environmental factors such as water droplets, reflections, and reduced visibility. These issues obscure critical details necessary for accurate vehicle detection and counting. The lower part of the figure shows the enhanced images after processing with our proposed DLSS and multi-ROI pixel area analysis method. Notably, the refined images exhibit improved clarity and detail, with reduced noise and enhanced visibility of vehicles. This improvement enables more reliable vehicle detection and counting, even under adverse environmental conditions. By comparing the raw and processed images, the figure underscores the ability of the proposed method to overcome the limitations of traditional traffic monitoring techniques in non-ideal weather scenarios.

TABLE I. ACCURACY RESULTS OF THE PROPOSED METHOD COMPARED TO OTHER METHODS ON THE AAU RAINSNOW TRAFFIC SURVEILLANCE DATASET

No	Method	Intersection						
		Egensevej (%)	Hadsundvej (%)	Hasserisvej (%)	Hjorringvej (%)	Hobrovej (%)	Ostre (%)	Ringvej (%)
1	Proposed method	81.31	75.23	64.87	73.12	93.44	68.12	92.78
2	Pixel area	19.82	41.71	27.74	30.34	25.25	41.90	18.85
3	Multi-ROI pixel area	24.77	32.87	16.25	17.96	25.69	41.16	36.74
4	CNN with feature concatenation	50.49	51.38	60.00	57.43	25.89	34.92	38.49
5	DCGAN	87.24	58.69	95.75	77.39	71.42	62.84	78.41
6	YOLOv3	85.21	68.75	78.87	66.02	90.49	84.59	73.22
7	YOLO-UA	68.94	87.46	66.04	70.37	73.30	59.63	72.81
8	DarkNet53	84.70	49.19	72.59	65.61	60.64	77.43	94.06

Fig. 2. Comparison of image quality and detection performance before and after applying the proposed method under adverse weather conditions. First row is the original image, second row is the image after applying the proposed method.

We took a closer look at the results of our experiments and saw the influence of certain weather types at each set of intersections on the obtained results. Therefore, we regrouped the results based on weather types, which consisted of rain and thunderstorm, as shown in Table II. The results obtained are really intriguing because the reliability of each method is not evenly distributed for all poor weather conditions. In rain conditions, our method demonstrates a notable performance, achieving an accuracy of 80.95%. This result reflects the effectiveness of our approach in dealing with moderate adverse weather conditions, where DLSS plays a crucial role in enhancing the quality of the input data, leading to improved detection accuracy. The ability of DLSS to upscale and denoise images is particularly beneficial in rain, where visual noise is less severe and more predictable, allowing the model to focus on refining key features and minimizing errors. Our method performs worst in thunderstorms, with an accuracy of 36.58%. This is even worse than one of the methods not designed for adverse weather conditions, CNN with feature concatenation, which achieved an accuracy of 37.01%. The significant drop in accuracy under thunderstorm conditions could be attributed to the severe challenges posed by these conditions, including heavy rain, lightning, and potentially low visibility, which significantly degrade the quality of the input data. DLSS, the technique utilized in our method, is designed to enhance resolution and detail, especially in more common or moderate scenarios like rain. However, in extreme weather conditions such as thunderstorms, where noise and visual artifacts are more complex and pronounced, DLSS may struggle to effectively reconstruct the fine details necessary for accurate object detection. The variability and intensity of visual disturbances in thunderstorms likely exceed the capacity of our current DLSS implementation to correct and enhance the input data, leading to the observed drop in performance.

TABLE II. ACCURACY RESULTS OF THE PROPOSED METHOD COMPARED TO OTHER METHODS ON THE AAU RAINSNOW TRAFFIC SURVEILLANCE DATASET: RESULTS FOR BAD WEATHER CONDITION TYPES

No	Method	Weather Condition	
		Rain (%)	Thunderstorm (%)
1	Proposed method	80.95	36.58
2	Pixel area	62.37	56.88
3	Multi-ROI pixel area	54.36	30.97
4	CNN	62.61	26.29
5	CNN with feature concatenation	52.64	37.01
6	DCGAN	93.57	74.37
7	YOLOv3	72.51	57.15
8	YOLO-UA	69.16	68.46
9	DarkNet53	87.54	80.92

Moreover, while DLSS is powerful for refining images, its performance is highly dependent on the quality and type of training data. If the model was not adequately trained on data that simulated the specific challenges of thunderstorm conditions, it would naturally struggle to generalize to these scenarios, resulting in lower accuracy. This highlights the need to further enhance the model's robustness, possibly by incorporating more diverse and challenging weather data during training, or by developing additional pre-processing techniques specifically aimed at mitigating the unique disturbances present in thunderstorm conditions. Thunderstorms can cause rapid and unpredictable changes in the environment, such as sudden flashes of lightning, shifts in rain intensity, and variations in visibility. These rapid changes create inconsistencies in the input data that can be challenging for DLSS to handle effectively. Unlike some of the compared methods, which have certain resistance to such inconsistencies, DLSS is known to make highly overconfident predictions in the presence of uncertainty caused by these inconsistencies. This overconfidence can lead to significant errors in detection, as the model may misinterpret noise as meaningful features, resulting in lower accuracy. Under rain conditions, our method achieves an accuracy of 80.95%, which, while respectable, still falls short of the performance achieved by DCGAN (93.57%) and DarkNet5 (87.54%). This suggests that while DLSS improves performance in moderate conditions, there is room for optimization, particularly in how the model handles the varying intensities and types of noise introduced by adverse weather.

The proposed method demonstrates significant improvements in vehicle detection under rain conditions, but several limitations must be acknowledged. Its performance under extremely adverse weather conditions, such as thunderstorms, remains limited. Thunderstorms introduce dynamic noise elements like lightning flashes and heavy rainfall, causing significant inconsistencies in the input data. These challenges highlight the need for advanced preprocessing techniques and training data augmentation tailored to such conditions. Additionally, the method relies heavily on the quality and diversity of the training data. A limited representation of rare or extreme weather scenarios in the dataset may restrict the model's ability to generalize effectively, emphasizing the importance of developing more extensive datasets that encompass a broader range of environmental conditions. Furthermore, the computational complexity of integrating DLSS with multi-ROI analysis can be challenging for real-time applications, especially in large-scale traffic networks. Optimization techniques or hardware acceleration strategies, such as leveraging edge computing devices, could help address this issue. Finally, the method's reliance on visual data alone can be a significant drawback in scenarios with near-zero visibility scenarios.

While the proposed method demonstrates strong performance in moderate adverse weather conditions, its limitations under thunderstorm scenarios highlight areas for future research. Future research to address these limitations can begin by exploring temporal noise filtering. This technique leverages the temporal consistency across successive frames to identify and suppress transient noise, such as flashes of lightning, that often disrupt vehicle detection. By focusing on patterns that persist over time, the detection system can reduce false positives caused by brief but intense disturbances and improve overall reliability. Another promising direction is the integration of additional sensing modalities, such as LiDAR or radar, which are less affected by the visibility issues common in thunderstorms. LiDAR, for instance, provides precise distance and object detection capabilities, while radar offers robustness to environmental noise like heavy rain. Combining these modalities with the existing DLSS framework could create a complementary system where weaknesses in one sensor type are offset by the strengths of others, ensuring more reliable vehicle detection.

Expanding the training dataset with augmented samples that simulate extreme weather scenarios, including thunderstorms, can significantly improve model robustness. Techniques such as applying synthetic rain effects, lightning patterns, and varying visibility levels to training data can better prepare the deep learning model for real-world conditions. This approach helps the model learn to distinguish noise from meaningful patterns, reducing the risk of misclassification under adverse weather conditions. Adversarial training approaches using Generative Adversarial Networks (GANs) can be employed to create realistic thunderstorm scenarios. GAN-generated data can mimic complex environmental factors such as varying rain intensity and dynamic lighting changes, offering a unique opportunity to fine-tune the model. By training the model to overcome these adversarial examples, it may develop enhanced resilience and generalization capabilities for handling extreme weather conditions.

IV. CONCLUSIONS

We introduced a method that combines Deep Learning Super Sampling (DLSS) and multi-Region of Interest (multi-ROI) pixel analysis for vehicle detection under non-ideal environmental conditions. Our method demonstrated superior performance under rainy weather, showcasing its robustness compared to traditional approaches. While the method performed well under moderate adverse conditions, challenges remain in extreme scenarios, such as thunderstorms, where dynamic noise and visibility issues significantly affect accuracy.

Future research could explore techniques such as temporal noise filtering to address transient noise like flashes of lightning in thunderstorms. In addition, the integration of complementary sensing modalities, such as LiDAR or radar, could mitigate the limitations of visual data in extreme weather conditions. Expanding the training dataset to include more diverse and challenging weather scenarios would further enhance model robustness and adaptability. By pursuing these future research directions, this study provides a foundation for developing more resilient and adaptable traffic monitoring systems. These advancements have the potential to significantly enhance Intelligent Transportation Systems (ITS) by improving traffic flow and safety under a wide range of environmental conditions.

ACKNOWLEDGMENT

The researchers would like to thank the Ministry of Research, Technology and Higher Education of the Republic of Indonesia for funding this research under the Fundamental Research Scheme in 2024 with Grand Number 02035/UN4.22.2/PT.01.03/2024.

REFERENCES

[1] Q. Zhu, Y. Liu, M. Liu, S. Zhang, G. Chen, and H. Meng, "Intelligent Planning and Research on Urban Traffic Congestion," *Future Internet*, vol. 13, no. 11, Nov. 2021, Art. no. 284, https://doi.org/10.3390/fi13110284.

[2] M. Alam, J. Ferreira, and J. Fonseca, "Introduction to Intelligent Transportation Systems," in *Intelligent Transportation Systems: Dependable Vehicular Communications for Improved Road Safety*, M. Alam, J. Ferreira, and J. Fonseca, Eds. Cham, Switzerland: Springer International Publishing, 2016, pp. 1–17.

[3] J. Lin, Y. Huang, X. Su, Z. Su, and P. Zhao, "An In-vehicle Camera Based Traffic Estimation in Smart Transportation," in *2019 IEEE 5th International Conference on Computer and Communications*, Chengdu, China, 2019, pp. 2186–2192, https://doi.org/10.1109/ICCC47050.2019.9064182.

[4] C. H. Bahnsen and T. B. Moeslund, "Rain Removal in Traffic Surveillance: Does it Matter?," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 8, pp. 2802–2819, Aug. 2019, https://doi.org/10.1109/TITS.2018.2872502.

[5] N. K. Jain, R. K. Saini, and P. Mittal, "A Review on Traffic Monitoring System Techniques," in *Soft Computing: Theories and Applications, Proceedings of SoCTA 2017*, Jhansi, India, 2019, pp. 569–577, https://doi.org/10.1007/978-981-13-0589-4_53.

[6] T. Pamula, "Road Traffic Conditions Classification Based on Multilevel Filtering of Image Content Using Convolutional Neural Networks," *IEEE Intelligent Transportation Systems Magazine*, vol. 10, no. 3, pp. 11–21, 2018, https://doi.org/10.1109/MITS.2018.2842040.

[7] I. C. Amitha and N. K. Narayanan, "Object Detection Using YOLO Framework for Intelligent Traffic Monitoring," in *Machine Vision and Augmented Intelligence—Theory and Applications: Select Proceedings of MAI 2021*, Jabalpur, India, 2021, pp. 405–412, https://doi.org/10.1007/978-981-16-5078-9_34.

[8] K. Yan and Z. Zhang, "Automated Asphalt Highway Pavement Crack Detection Based on Deformable Single Shot Multi-Box Detector Under a Complex Environment," *IEEE Access*, vol. 9, pp. 150925–150938, 2021, https://doi.org/10.1109/ACCESS.2021.3125703.

[9] D. Biswas, H. Su, C. Wang, A. Stevanovic, and W. Wang, "An automatic traffic density estimation using Single Shot Detection (SSD) and MobileNet-SSD," *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 110, pp. 176–184, Apr. 2019, https://doi.org/10.1016/j.pce.2018.12.001.

[10] S. Zhang, Y. Guo, P. Zhao, C. Zheng, and X. Chen, "A Graph-Based Temporal Attention Framework for Multi-Sensor Traffic Flow Forecasting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 7743–7758, Jul. 2022, https://doi.org/10.1109/TITS.2021.3072118.

[11] C. Hu, L. Xu, Y. Guo, X. Jing, X. Lu, and P. Liu, "HSV-3S and 2D-GDA for High-Saturation Low-Light Image Enhancement in Night Traffic Monitoring," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 12, pp. 15190–15206, Dec. 2023, https://doi.org/10.1109/TITS.2023.3308894.

[12] W. Zhang, X. Sui, G. Gu, Q. Chen, and H. Cao, "Infrared Thermal Imaging Super-Resolution via Multiscale Spatio-Temporal Feature Fusion Network," *IEEE Sensors Journal*, vol. 21, no. 17, pp. 19176–19185, Sep. 2021, https://doi.org/10.1109/JSEN.2021.3090021.

[13] J. Shi and K. Yang, "An Improved Histogram Equalization Method in the Traffic Monitoring Image Processing Field," *Journal of Computer and Communications*, vol. 3, no. 11, pp. 25–32, Nov. 2015, https://doi.org/10.4236/jcc.2015.311005.

[14] C.-H. Hu *et al.*, "Joint Image-to-Image Translation for Traffic Monitoring Driver Face Image Enhancement," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 8, pp. 7961–7973, Aug. 2023, https://doi.org/10.1109/TITS.2023.3258634.

[15] X. Ji, J. Cheng, J. Bai, T. Zhang, and M. Wang, "Real-time enhancement of the image clarity for traffic video monitoring systems in haze," in *2014 7th International Congress on Image and Signal Processing*, Dalian, China, 2014, pp. 11–15, https://doi.org/10.1109/CISP.2014.7003741.

[16] K. Mondal, R. Rabidas, and R. Dasgupta, "Single image haze removal using contrast limited adaptive histogram equalization based multiscale fusion technique," *Multimedia Tools and Applications*, vol. 83, no. 5, pp. 15413–15438, Feb. 2024, https://doi.org/10.1007/s11042-021-11890-0.

[17] S. Basak and S. Suresh, "Vehicle detection and type classification in low resolution congested traffic scenes using image super resolution," *Multimedia Tools and Applications*, vol. 83, no. 8, pp. 21825–21847, Mar. 2024, https://doi.org/10.1007/s11042-023-16337-2.

[18] K. Guo *et al.*, "Video Super-Resolution Based on Inter-Frame Information Utilization for Intelligent Transportation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 11, pp. 13409–13421, Nov. 2023, https://doi.org/10.1109/TITS.2023.3237708.

[19] Y. Jin, Y. Zhang, Y. Cen, Y. Li, V. Mladenovic, and V. Voronin, "Pedestrian detection with super-resolution reconstruction for low-quality image," *Pattern Recognition*, vol. 115, Jul. 2021, Art. no. 107846, https://doi.org/10.1016/j.patcog.2021.107846.

[20] S. Nousheen and S. P. Kumar, "Novel Fog-Removing Method For The Traffic Monitoring Image," *International Journal of Innovative Technology and Research*, vol. 4, no. 5, pp. 4159–4162, Aug. 2016.

[21] S. Markidis, S. W. D. Chien, E. Laure, I. B. Peng, and J. S. Vetter, "NVIDIA Tensor Core Programmability, Performance & Precision," in *2018 IEEE International Parallel and Distributed Processing Symposium Workshops*, Vancouver, Canada, 2018, pp. 522–531, https://doi.org/10.1109/IPDPSW.2018.00091.

[22] R. Chowdhury, F. Silvestri, and F. Vella, "A Computational Model for Tensor Core Units," in *Proceedings of the 32nd ACM Symposium on Parallelism in Algorithms and Architectures*, Virtual Event, USA, 2020, pp. 519–521, https://doi.org/10.1145/3350755.3400252.

[23] C. M. Bautista, C. A. Dy, M. I. Mañalac, R. A. Orbe, and M. Cordel, "Convolutional neural network for vehicle detection in low resolution traffic videos," in *2016 IEEE Region 10 Symposium*, Bali, Indonesia, 2016, pp. 277–281, https://doi.org/10.1109/TENCONSpring.2016.7519418.

[24] F. Zhang, C. Li, and F. Yang, "Vehicle Detection in Urban Traffic Surveillance Images Based on Convolutional Neural Networks with Feature Concatenation," *Sensors*, vol. 19, no. 3, Jan. 2019, Art. no. 594, https://doi.org/10.3390/s19030594.

[25] L. Ge, D. Dan, and H. Li, "An accurate and robust monitoring method of full-bridge traffic load distribution based on YOLO-v3 machine vision," *Structural Control and Health Monitoring*, vol. 27, no. 12, Dec. 2020, Art. no. e2636, https://doi.org/10.1002/stc.2636.

[26] C.-Y. Cao, J.-C. Zheng, Y.-Q. Huang, J. Liu, and C.-F. Yang, "Investigation of a Promoted You Only Look Once Algorithm and Its Application in Traffic Flow Monitoring," *Applied Sciences*, vol. 9, no. 17, Sep. 2019, Art. no. 3619, https://doi.org/10.3390/app9173619.

[27] M. Hassaballah, M. A. Kenk, K. Muhammad, and S. Minaee, "Vehicle Detection and Tracking in Adverse Weather Using a Deep Learning Framework," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 7, pp. 4230–4242, Jul. 2021, https://doi.org/10.1109/TITS.2020.3014013.