RESEARCH-ARTICLE

Reflection Removal and Facial Detection of Individuals in Vehicles

WORAPOB KEATKONGSANG, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

WATANAI MAYTHAMALUANG, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

ABDULHAKIM MAHA, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

ORACHAT CHITSOBHUK, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

Published: 15 January 2024

ICCDE 2024: 2024 10th International Conference on Computing and Data Engineering
January 15 - 17, 2024
Bangkok, Thailand

ICCDE '24: Proceedings of the 2024 10th International Conference on Computing and Data Engineering (January 2024)
https://doi.org/10.1145/3641181.3641193
ISBN: 9798400709319

Reflection Removal and Facial Detection of Individuals in Vehicles

Worapob Keatkongsang
King Mongkut's Institute of Technology Ladkrabang, Thailand
63010841@kmitl.ac.th

Watanai Maythamaluang
King Mongkut's Institute of Technology Ladkrabang, Thailand
63010846@kmitl.ac.th

Abdulhakim Maha
King Mongkut's Institute of Technology Ladkrabang, Thailand
63011075@kmitl.ac.th

Orachat Chitsobhuk*
King Mongkut's Institute of Technology Ladkrabang, Thailand
orachat.ch@kmitl.ac.th

ABSTRACT

The research, "Reflection Removal and Facial Detection of Individuals in Vehicles," utilizes Single Image Reflection Removal (SIRR) technology and Face Detection to remove reflections and reduce glare caused by automotive glass and film. This enables the capture of facial images of individuals inside vehicles. SIRR technology enhances image quality by removing reflections from the surfaces of glass that might obscure objects. In this research, we explore the use of three models specialized in SIRR and YOLOv7 for Face Detection. However, the pre-trained models for reflection removal failed to effectively remove reflections and reduce glare from films. In this paper, we propose an approach to enhance the efficiency of removing reflections and reducing glare caused by automotive glass and film with opacities set at 40% and 60%, achieving an impressive improvement in Peak Signal-to-Noise Ratio (PSNR) by approximately 42.96% and Structural Similarity Index (SSIM) by approximately 34.16% compared to the pre-trained models.

ACM Reference Format:
Worapob Keatkongsang, Watanai Maythamaluang, Abdulhakim Maha, and Orachat Chitsobhuk. 2024. Reflection Removal and Facial Detection of Individuals in Vehicles. In *2024 10th International Conference on Computing and Data Engineering (ICCDE 2024), January 15–17, 2024, Bangkok, Thailand*. ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3641181.3641193

1 INTRODUCTION

In the domain of Computer Vision and Image Processing, the research on "Reflection Removal and Facial Detection of Individuals in Vehicles" addresses a critical challenge – the presence of reflections and glare on automotive glass and film that prohibit the clear capture of facial images of individuals inside vehicles. This research leverages two cutting-edge technologies: Single Image Reflection Removal (SIRR) and Face Detection.

SIRR technology, renowned for its ability to enhance image quality, plays a crucial role in the elimination of reflections that arise from the glass surfaces that often obscure objects. The reflections caused by the refraction of light off the front-facing glass can significantly impede visibility, thus emphasizing the importance of utilizing more advanced solutions. In the framework of our investigation for the effective reflection removal and glare reduction, we assessed the performance of 3 SIRR models: "Your trash is my treasure" (YTMT) [1], "Perceptual Reflection Removal" (PRR) [2], and "Enhanced Reflection Removal Network" (ERRNet) [3]. These models were integrated with the "You Only Look Once version 7" (YOLOv7) [4] models for Face Detection system. However, all the pre-trained models for SIRR exhibited certain constraints in effectively handling the issues caused by reflections and excessive glare on automotive glass and film within our specific scenario.

To address these challenges, our research presents a novel strategy to improve the performance of the SIRR models in removing reflections and reducing glare, particularly from automotive glass and film with opacities set at 40% and 60%. This is accomplished through our dataset preparation process, utilizing synthetic and real-world images, as well as model optimization techniques, including fine-tuning and fine-tuning with a freeze discriminator [5]. Our research not only highlights the significance of dealing with reflection-related challenges but also demonstrates the potential of leveraging state-of-the-art technologies to address real-world issues effectively. In the following sections, we will explore the techniques, experiments, and results, to illustrate the efficacy of our proposed approach compared to the pretrained models.

2 RELATED WORK

The removal of reflections from images has attracted considerable attention in computer vision and image processing. It plays a significant role in enhancing the quality and usability of images in various applications. We highlight three notable works in this domain:

2.1 Single Image Reflection Removal (SIRR)

The removal of reflections from images has garnered significant attention in computer vision and image processing. It plays a pivotal role in enhancing the quality and usability of images in various applications. We highlight three notable works in this domain:

2.1.1 Single Image Reflection Removal Exploiting Misaligned Training Data and Network Enhancements [3]. This research work brings forward a significant contribution in the field of image reflection

*Correspondence: orachat.ch@kmitl.ac.th (O.C.)

ICCDE 2024, January 15–17, 2024, Bangkok, Thailand

ACM ISBN 979-8-4007-0931-9/24/01
https://doi.org/10.1145/3641181.3641193

removal. Notably, it introduces a novel strategy in dataset preparation process for training purposes, which demonstrates the ability to effectively employ misaligned data for training. A key innovation of this work lies in the calculation of loss for misaligned data.

The model architecture employed in this study incorporates a straightforward Residual block, with the removal of the Batch Normalize Layer. Furthermore, in the final stages of the model, Pyramid pooling is employed to collect multi-scale spatial information from each channel within the CNN structure. This addition aids in capturing global-scene-level features from the input images, enhancing the model's ability to handle complex reflection removal scenarios.

2.1.2 Single Image Reflection Separation with Perceptual Losses [2]. In this research study, it is emphasized that a model aiming to effectively separate reflections from images must possess a deep understanding of the content within those images. The network in this research work employs Hypercolumn Features, extracted from selected layers of VGG19, pre-trained on the ImageNet dataset. These features are then fed into a Fully Convolutional Network (FCN).

The research also incorporates Feature loss, calculated by comparing the predicted transmission with the ground truth transmission. Additionally, an Adversarial loss is utilized to address issues related to unrealistic coloration and irrelevant details in the predicted transmission. Notably, this research model does not focus on the loss associated with predicting reflections because it employs a dataset collected from real-world scenarios. As indicated in the paper, model evaluation is conducted through a Blind Test, comparing the model mentioned above with CEILNet [6] and Li and Brown's models [7]. These models are trained on the same dataset used in this research, and they are evaluated by the community in general. The evaluation involves pairing reflection-free predictions from all three models with corresponding reflection-free ground truth images. The results demonstrate that the model from this research performs better in removing reflections.

2.1.3 Trash or Treasure? An Interactive Dual-Stream Strategy for Single Image Reflection Separation [1]. In this research, the innovative concept of "Your trash is my treasure" is introduced. The primary objective of this study is to address data loss issues caused by the deficiency of the regular ReLU activation functions. To alleviate the issues, the research adopts Negative ReLU, facilitating rapid reduction in training errors.

The algorithm in this research boasts a Dual-Stream architecture, Notably, these two streams exchange information with each other, facilitating the transfer of non-relevant data from one stream to another using Negative ReLU. The backbone of the architecture is structured as a U-Net. Additionally, this research utilizes a two-stage training approach. In the first stage, the model is trained until convergence, after which the parameters are frozen. The second stage of training involves utilizing the result of the first stage as its input and conducting additional training until reaching its convergence.

These three significant researches in the field of Single Image Reflection Removal have demonstrated various approaches to address the challenges posed by reflections in images. Each work contributes valuable insights into the field, ranging from innovative loss calculations, feature extraction techniques, to training strategies. The methodologies presented in these works serve as a foundation for our research, "Reflection Removal and Facial Detection of Individuals in Vehicles," as our objective to build upon their four successes and tailor them to our specific application domain. In the subsequent sections, we will detail the relevant research in the area of face detection, which complements our primary focus on reflection removal.

2.2 Face Detection

2.2.1 YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors [4]. YOLOv7, a novel real-time object detection system that excels in both speed and accuracy. YOLOv7's architecture is founded on extended efficient layer aggregation networks. This architectural design carefully balances parameters, computation volume, computational density, input/output channel ratios, the number of architecture branches, and element-wise operations to enhance network inference speed. YOLOv7 presents an effective solution for real-world applications, accompanied by open-source access to its codebase. This work significantly advances the field of real-time object detection, offering valuable insights and methodologies for future research endeavors.

3 METHOD

In this section, we present an overview of our system for effectively removing reflections and detecting faces as illustrated in Fig. 1

3.1 Reflection Removal

We begin by inputting images with reflections caused by automotive glass and film into our SIRR model to remove reflections and reduce glare. The goal is to generate high-quality reflection free image that could obstruct the view of objects inside the vehicle, which will be used in the subsequent face detection step.

Our specific circumstances are the cases where reflection effects are considered in a variety of glass reflections, especially with film conditions. This poses a unique set of challenges, as the presence of films can introduce additional complexities since these films can modify the visual apparent of reflections in images. To address this challenge, we explore and potentially develop novel methodologies, building upon the success of established methods like YTMT [1], PRR [2], and ERRNet [3], which have demonstrated remarkable performance in reflection removal. The approach and key features can be summarized in Table 1.

3.2 Face Detection

The Face Detection model, YOLOv7, has gained considerable recognition for its exceptional accuracy in detecting objects. However, the severe conditions on input images such as reflections may deter the performance of face detection model. A preprocess, which removes these visual effects, can greatly enhance the detection performance. Consequently, we utilize the outcomes from the reflection reduction stage to supply the transmission layer. The utilization of the face detection model's bounding boxes serves the purpose of determining the boundaries of the detected individuals' faces. We then extract the facial regions from the images using the bounding boxes obtained in the face detection step. The objective of this step is to obtain clear and accurate images of individuals'

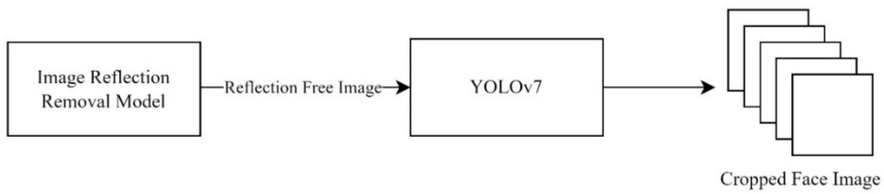

Figure 1: The overview of the process for effectively removing reflections and face detection.

Table 1: A summary of Approach and Key Features of the Reflection Removal Models

Model	Approach and Key Features
PRR [2]	Deep neural network with perceptual losses, feature loss for comparing images in feature space, adversarial loss for realistic image refinement, and exclusion loss to separate reflection and transmission layers. Utilizes hyper column features from a pre-trained VGG-19 network for semantic understanding. Fully convolutional network with a large receptive field. 5000 random pairs of images from Flickr: one outdoor image and one indoor image for each pair. A synthetic dataset with a Gaussian smoothing with a random kernel size, 110 real image pairs: image with reflection and its corresponding ground-truth transmission image.
YTMT [1]	A general interactive dual stream/branch strategy to determine exchanging features between two streams. Activation functions (e.g., ReLU) to select and exchange features. Addressing information loss and dead ReLU problems. Simple and flexible implementation on plain and UNet architectures. Performing studies and experiments on a fusion of synthetic from PASCALVOC dataset [8] and 90 real-world images from [2].
ERRNet [3]	Focus on early layers of CNNs for reflection removal. Utilizes channel-wise context and multi-scale spatial context. Channel attention mechanism for feature weighting. Aggregation of information across different scales for global contextual consistency. Performing studies and experiments on the same dataset as [1].

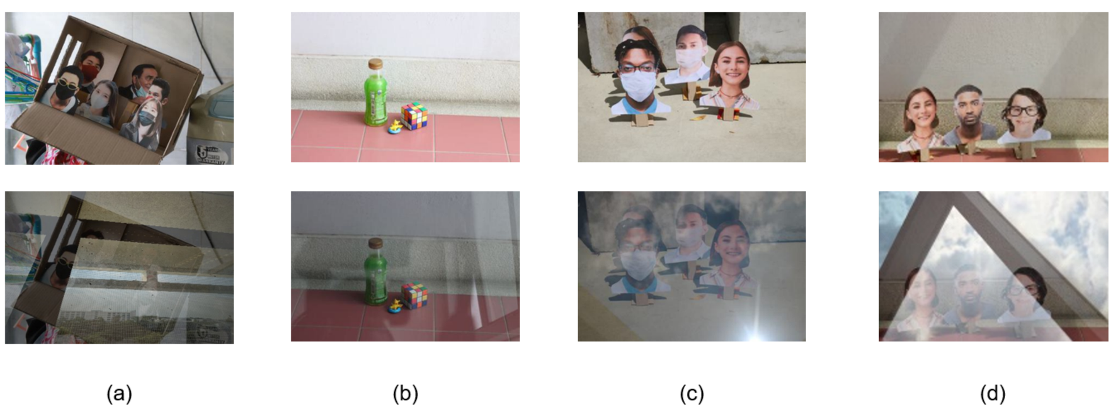

Figure 2: Examples from Dataset used in our experiments. (a) synthesis image (b) – (d) our self-taken images where the top row represents the ground truth images and the bottom row depicts the image with the reflection.

faces that have been detected. The extracted facial images are used to evaluate the effectiveness and quality of the developed system.

4 EXPERIMENT

In this section, we present the experiments conducted in this research, focusing on two main aspects: reflection removal and face detection.

4.1 Reflection Removal Experiments

4.1.1 Data Preparation. In this section, we detail the data preparation for training our reflection removal model. The dataset comprises of two main components: synthetic data and real images used for training the model, as shown in Figure 2.

For synthetic data. We utilize the same dataset as the [1] for synthetic data. The data generation process is based on the methods in [2], with our additional gamma values randomly adjusted within the range of 0.3 to 0.7 throughout the experiments. This adjustment helps simulate the level of light reflection typically encountered in automotive glass and film with opaque automotive glass films set at 40% and 60% reflectivity.

For the real images data, we collected a total of 144 images for the training set and 50 images for the test set. These real images were obtained through our self-taken images, following a data collection approach similar to [2]. However, we modified the process by using glass with reflective films at two different levels, 40% and 60% reflectivity, instead of regular clear glass. These real-world images closely resemble the scenarios encountered in practical use cases. Data preparation for both synthetic and real-world data is a crucial step in training the model to achieve accuracy in removing reflections, especially in scenarios that closely mimic real-world conditions.

4.1.2 Training. In the part of training the reflection removal model, we proceeded with various steps as follows:

Fine-Tuning with Our Dataset: We started with the utilization of pre-trained models as references (YTMT, PRR, and ERRNet). During this stage, we applied the default parameters of each model and conducted fine-tuning using our self-taken image dataset. This fine-tuning process enabled us to adapt the model's parameters specifically for the task of reflection removal.

Fine-Tuning with Freezing Lower Discriminator Layers: In the second experiment, we implemented a comparable fine-tuning procedure utilizing pre-trained models, while integrating the concept from reference [5]. In particular, the lower layers of the discriminator model were frozen. This freezing process helped the model focus on learning and improving the crucial aspects of reflection removal efficiently, while ensuring that the parameters inside these layers remained unaffected.

Reducing the Blurriness of the Reflection Layer in Synthetic Data: In the third experiment, we observed that the level of blurriness of reflection layer in the synthetic data was significantly high. Consequently, we decided to reduce the level of blur and continued fine-tuning the pre-trained models while still freezing the lower layers of the discriminator [5]. This adjustment resolved the issue of excessive blurriness in the synthetic data, resulting in improved model's accuracy for reflection removal.

We conducted model training using various approaches to continuously develop and enhance the model's ability for reflection removal. Techniques such as utilizing pre-trained models and freezing discriminator layers at lower levels were utilized in order to improve the learning process of the model in the desired direction. In addition, lowering the blur level of the reflection layer in the synthetic data assisted the model in accurately performing the removal of reflections, which aligned with the scenarios presented in our research work.

4.2 Face Detection Experiment

To ensure efficient face detection in the context of vehicles, we conducted experiments using the YOLOv7 model, known for its accuracy in object detection. Initially, we employed the WIDER FACE dataset [9], a comprehensive face detection benchmark containing over 32,203 images and a total of 393,703 annotated faces, for Transfer Learning of the pre-trained YOLOv7 model. This Transfer Learning process allowed the YOLOv7 model to learn and improve its ability to accurately detect faces especially in reflection circumstances efficiently and effectively.

4.3 Experimental Results

4.3.1 Face Detection. In Face Detection experiments, we illustrate the performance of our fine-tuned YOLOv7 model using WIDER FACE dataset and default hyperparameters for face detection. Our training configuration included a batch size of 4, training over 200 epochs, and using image dimensions of 640 x 640. Following the training process, we assessed the model's performance on the WIDER FACE test set, utilizing a confidence threshold of 0.5 and an Intersection over Union (IoU) threshold of 0.5. The detailed results are listed in Table 2.

From Table 2, we observe that the Precision metric exhibits a relatively high value of 92.6%, indicating the model's accuracy in correctly identifying actual faces. On the other hand, the Recall value, at 0.665, signifies that the model successfully detects around 66.5% of all the ground truth faces. Furthermore, the mAP@.5 (Mean Average Precision at IoU 0.5) at 0.661, reflects the average detection accuracy when using an IoU threshold of 0.5. Meanwhile, the mAP@.5:.95 (Mean Average Precision from IoU 0.5 to 0.95) yields a value of 0.379, which represents the average detection accuracy over a range of IoU values from 0.5 to 0.95.

In general, the model exhibits a respectable performance, even though certain metrics, like Recall, may not fully meet our expected benchmarks for comprehensive face detection. However, it remains sufficiently effective for practical applications in face detection and image cropping, making it a valuable asset for performance assessment in subsequent phases of our study.

4.3.2 Reflection Removal. In our reflection removal study, we incorporated a set of our 50 self-taken images that featured diverse scenarios, including situations involving human faces and objects behind reflective surfaces. We also measured the reflection removal specifically in the facial region of the 191 ground truth images. This custom dataset served as a comprehensive testing ground for evaluating the effectiveness of the strategies.

By comparing the results of our experiments with those obtained using pre-trained models from related research, we were able to thoroughly assess the performance of each method we employed. This analysis provided valuable insights into the efficacy of our approaches. To provide further understanding of the results of reflection removal, we present a visual comparison between the results obtained from the reference models and our approaches to fine tuning the reference models in Figure 3.

In addition to visual comparisons, we quantitatively evaluated the performance of our methods using well-known metrics such as PSNR and SSIM. These metrics provide quantitative insights into the quality of the resulting images, especially inside facial region, compared to ground truth data.

Table 3 summarizes the results of our reflection removal experiments. In these experiments, we utilized three models: ERRNet, PRR, and YTMT. All these models were fine-tuned or had their discriminator frozen to enhance their performance. The "Pre-trained" row serves as a baseline measurement, representing the initial

Table 2: The results of YOLOv7 transfer learning with WIDER FACE dataset

	Precision	Recall	mAP@.5	mAP@.5:.95
YOLOv7	0.926	0.665	0.661	0.379

Figure 3: Visual Comparison of the Results obtained from Reference Pretrained Model and Our Developed Methods for Reflection Removal.

Table 3: The results of Reflection Removal Experiments

	Model	Metric (Entire Image)		Metric (Facial Region)	
		PSNR	SSIM	PSNR	SSIM
Pre-trained	ERRNet	13.29	0.493	14.32	0.404
	PRR	12.68	0.539	13.69	0.356
	YTMT	13.29	0.524	14.19	0.419
Fine-tune	ERRNet	17.79	0.666	17.52	0.607
	PRR	17.41	0.752	17.29	0.590
	YTMT	18.78	0.703	18.33	0.623
Freeze Discriminator	ERRNet	17.68	0.672	16.92	0.589
	PRR	17.39	0.754	17.29	0.590
	YTMT	18.31	0.696	17.89	0.612
Freeze Discriminator& Reducing the Blurriness	ERRNet	17.91	0.678	17.41	0.601
	PRR	17.45	0.759	17.52	0.597
	YTMT	19.00	0.703	18.23	0.616

performance of each model before any specific experiments were conducted.

For the fine-tuned models, we observed significant improvements in both PSNR and SSIM values. Notably, the YTMT model achieved the highest PSNR of 18.78 and an SSIM of 0.703 for the entire images, and the PSNR of 18.33 and an SSIM of 0.623 inside facial region, demonstrating its effectiveness in enhancing image quality. Even with the freezing of the discriminator, we continued to observe improvements in PSNR and SSIM values in comparison to the other pretrained models with remarkable performance, boasting the highest PSNR of 18.31 and an SSIM of 0.696 for the entire image and the PSNR of 17.89 and an SSIM of 0.612 inside facial region.

In the last experiment, with a reduction in the blurriness of the synthetic data, we continued to observe performance enhancements. The YTMT model reached a PSNR of 19.00 and an SSIM of 0.703 for the entire images and a PSNR of 18.23 and an SSIM of 0.616 inside facial image, further highlighting its capability to generate high-quality reflection-free images. Furthermore, even though the ERRNet and PRR models could not attain the same

Table 4: The Model Performance Comparison of the Facial Region after Reflection Removal

	Model	Metric				
		Average Confidence	Precision	Recall	mAP@.5	mAP@.5:.95
Pre-trained	ERRNet	0.628	0.986	0.737	0.742	0.517
	PRR	0.591	0.993	0.705	0.710	0.486
	YTMT	0.620	0.993	0.732	0.736	0.520
Fine-tune	ERRNet	0.685	0.987	0.795	0.800	0.561
	PRR	0.666	0.987	0.779	0.783	0.546
	YTMT	0.686	1.000	0.795	0.805	0.568
Freeze Discriminator	ERRNet	0.685	0.987	0.795	0.799	0.556
	PRR	0.689	0.981	0.795	0.799	0.552
	YTMT	0.684	0.980	0.784	0.789	0.551
Freeze Discriminator&	ERRNet	0.677	0.987	0.784	0.789	0.557
Reducing the Blurriness	PRR	0.658	0.987	0.795	0.800	0.561
	YTMT	0.666	1.000	0.774	0.779	0.540

PSNR and SSIM levels as the YTMT model, they still demonstrated significant improvements across all experiments. These results collectively demonstrate the effectiveness of our methods in improving reflection removal compared to the pretrained models, with YTMT consistently outperforming ERRNet and PRR across all experiments.

However, it is important to note that while our quantitative metrics exhibited significant improvements from the pretrained models, visual examination revealed that the generated images, particularly in terms of facial structure and color, still exhibited noticeable disparities from the ground truth, as depicted in Figure 3. This observation underscores the complexity of the reflection removal task, where quantitative metrics may not fully capture the nuances of visual quality. Further refinement and exploration in our methods may be required to achieve a closer resemblance to the ground truth in terms of facial attributes.

4.3.3 Performance Evaluation of Reflection Removal on a fine-tuned YOLOv7. To evaluate the performance of the images after reflection removal, we manually annotated the bounding boxes of all 50 sets of ground truth images, comprising a total of 191 faces. This annotation process involved using the positions of each face to create labels for each image. We then used these labels for the images resulting from each reflection removal model to measure their performance. The results are presented in Table 4. The performance measurement using a confidence threshold and an IoU threshold of 0.5 reveals that the average confidence scores for detecting faces using each method are relatively similar. This suggests that the models exhibit consistent confidence in recognizing faces. Notably, the Precision values stand out, with YTMT achieving a perfect Precision value of 1.0 for our test set. This signifies that the face detection model is exceptionally accurate, correctly identifying all accurate faces. Moreover, the average Recall value is approximately 0.77, indicating that the model can successfully detect approximately 77% of the ground truth faces. Furthermore, both the mAP@.5 and mAP@.5:.95 values exhibit significantly high levels of accuracy, highlighting the model's proficiency in face detection.

However, when annotating the ground truth images to generate labels for performance testing, one important consideration is that the YOLOv7 face detection model was trained with the WIDER FACE dataset and had already been annotated. In this research, we conducted additional annotation for the ground truth, which involved a different set of annotators. This could result in minor discrepancies in the face bounding box annotations due to the involvement of different individuals in the annotation process.

5 CONCLUSION

This paper addresses the challenging issue of reflection removal in facial images within the context of vehicles equipped with automotive glass and film. Our extensive experimentation with deep learning models, namely ERRNet, PRR, and YTMT, illustrated significant advancements in reflection reduction. Quantitative metrics, including PSNR and SSIM, consistently indicate substantial improvements, with YTMT emerging as the leading performer. Nevertheless, it is important to acknowledge that while our quantitative achievements are noteworthy, they do not always align perfectly with visual fidelity. Our research represents a significant stride forward in reflection removal for facial detection within the challenging environment of vehicles equipped with film on automotive glass. While our quantitative outcomes are promising, there remains room for enhancing visual fidelity, a critical aspect for practical applications. We hope our findings inspire further exploration and the development of more robust solutions for real-world scenarios.

REFERENCES

[1] Qiming Hu, and Xiaojie Guo. 2021. Trash or Treasure? An Interactive Dual-Stream Strategy for Single Image Reflection Separation. Advances in Neural Information Processing Systems, 30, 24683 – 24694
[2] Xuaner Zhang, Ren Ng, and Qifeng Chen. 2018. Single Image Reflection Separation with Perceptual Losses. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 4786-4794. https://doi.org/10.1109/CVPR.2018.00503
[3] Kaixuan Wei, Jiaolong Yang, Ying Fu, David Wipf, and Hua Huang. 2019. Single image reflection removal exploiting misaligned training data and network enhancements. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2019-June, 8170-8179. https://doi.org/10.1109/CVPR.2019.00837
[4] C. Wang, A. Bochkovskiy and H. Liao. 2023. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada,7464-7475. https://doi.org/10.1109/CVPR52729.2023.00721
[5] Sangwoo Mo, Minsu Cho, and Jinwoo Shin. 2020. Freeze Discriminator: A Simple Baseline for Fine-tuning GANs.

[6] Qingnan Fan, Jiaolong Yang, Gang Hua, Baoquan Chen, and David Wipf. 2017. A Generic Deep Architecture for Single Image Reflection Removal and Image Smoothing. Proceedings of the IEEE International Conference on Computer Vision, 2017-October,3258-3267, https://doi.org/10.1109/ICCV.2017.351

[7] Yu Li, and Michael S. Brown. 2014. Single image layer separation using relative smoothness. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2752-2759, https://doi.org/10.1109/CVPR.2014.346

[8] Mark Everingham, Luc Van Gool, Christopher K. I. Williams, John Winn, and Andrew Zisserman. 2010.The Pascal Visual Object Classes (VOC) Challenge.International Journal of Computer Vision, 303-338, https://doi.org/10.1007/s11263-009-0275-4

[9] S. Yang, P. Luo, C. C. Loy, and X. Tang. 2016. WIDER FACE: A Face Detection Benchmark. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 5525-5533, https://doi.org/10.1109/CVPR.2016.596