ReflectionEraser: Implementing SoTA Reflection Removal Models

Anika Farzana
200041204
anikafarzana@iut-dhaka.edu

K. M. Abesh Ahsan
200041225
abeshahsan@iut-dhaka.edu

Namirah Rasul
200041239
namirahrasul@iut-dhaka.edu

Abstract

Reflections in images, such as those captured through glass, can obscure important details and degrade image quality, posing significant challenges in various applications including photography, surveillance, and computer vision. This project aims to develop a desktop application that automatically removes reflections from images, enhancing their clarity and usability. The application leverages an advanced reflection removal model inspired by the latest research in Single Image Reflection Separation (SIRS). Out code is publicly available at https://github.com/namirahrasul/ReflectionEraser.git.

1 Introduction

In various fields such as photography, surveillance, and computer vision, the presence of reflections in images captured through glass surfaces can significantly diminish their quality and usability. While existing methods often rely on multiple images or specialized equipment like polarizers to remove reflections, these approaches are impractical for scenarios where only a single image is available. This project addresses these limitations by developing a desktop application that autonomously removes reflections from single images.

Reflection removal from single images is a significant challenge in the computer vision field, and numerous methods have been explored to solve this problem. Zhang et el. [4] improved the process of removing reflections by incorporating adversarial and perceptual losses into their model and utilizing HyperColumn features [1] for improved feature extraction. They also employed a pre-trained VGG-19 network to enhance semantic awareness. Li et el. [9] demonstrated adversarial losses. Also, they introduced a two-stage network (RAGNet) that initially estimated reflection components and then used this estimation to predict transmission components, effectively separating reflection estimation from transmission. Additionally, Hu and Guo [8] introduced the YTMT approach, which emphasizes equal attention to both components. They also created a dual-stream interactive network to restoration both layers at the same time. However, their linear assumption often resulted in underwhelming predicted reflection components. Other approaches, such as BDN [3] and IBCLN [6], utilizes models that are assigned weights and iteratively estimates both components to avoid weak reflections.

However, the restoration process often faces challenges in completely separating the reflection from the transmission layer. Wen et al. [5] tackled the issue of nonlinear superimposition by using adversarial guidance to predict a three-channel alpha blending weight map from unpaired images they collected. Dong et al. [7] developed an iterative network and estimated a probabilistic reflection confidence map in each iteration. Recognizing limitations of previous reflection models, Hu and Guo [10] introduce a more versatile and powerful approach. They propose a general form of these models that incorporates a learnable residue term, resulting in improved effectiveness and flexibility. And among the models, the one by Hu and Guo [10] is state-of-the-art. The work of DSRNet [10] and IBCLN [6] has inspired our project. Our approach utilizes a combination of two state-of-the-art models: DSRNet and IBCLN. DSRNet employs a dual-stream network architecture with a semantic pyramid encoder, incorporating a novel learnable residue term to handle the complexities of reflection superposition. It features a mutually-gated interaction mechanism, facilitating more effective feature extraction and separation of reflection and transmission layers.

On the other hand, IBCLN introduces an Iterative Boost Convolutional LSTM Network for removing reflections. It uses a cascaded approach where transmission and reflection layers are refined iteratively, and employs Convolutional LSTM units to transfer information across cascade steps, addressing the vanishing gradient problem during training. Additionally, IBCLN features a novel residual reconstruction loss function, which enhances the accuracy of reflection removal by guiding the network through multiple cascade steps.

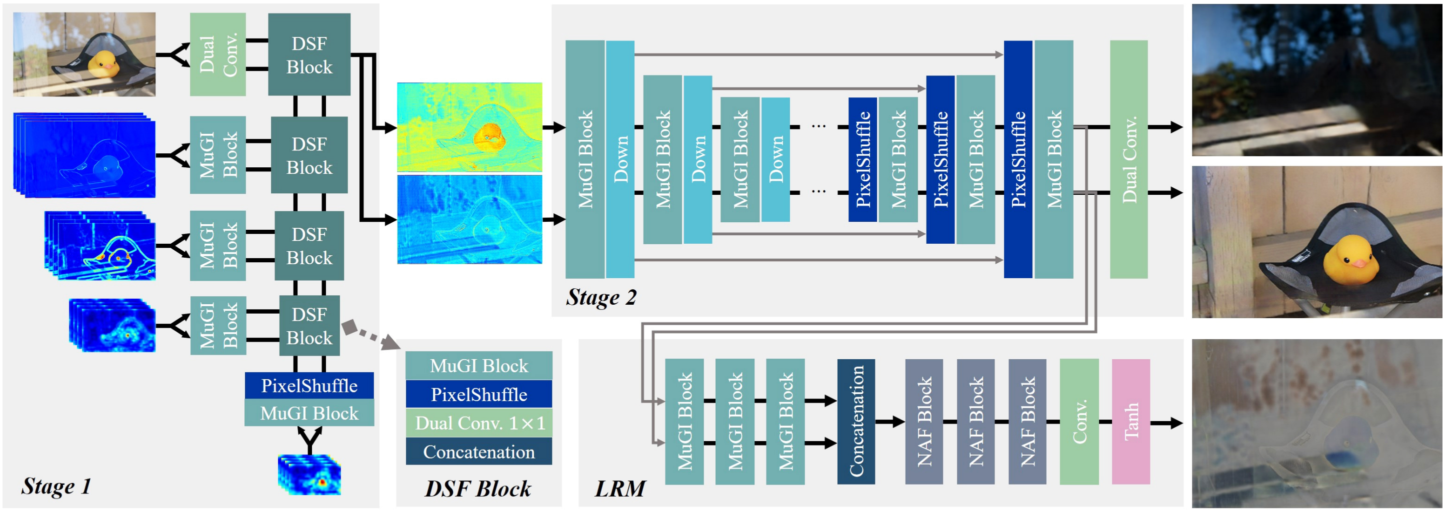

Figure 1: The DSRNet Architecture [10].

2 Methodology

Our study employs two state-of-the-art models for single image reflection removal: DSRNet [10] and IBCLN [6]. To ensure robust and equitable comparisons, both models were trained under identical experimental conditions.

For training, we used the datasets from DSRNet[10]:

- A subset of 7,643 images from the Pascal VOC dataset, center-cropped to 224x224 dimensions to generate training pairs.
- 90 real-world training pairs, specifically selected to challenge the models with complex real-world scenarios.
- An additional 200 real-world training pairs sourced from the IBCLN dataset, enriching the training corpus with varied reflection removal challenges.

This comprehensive dataset selection and rigorous experimental setup not only facilitate a thorough evaluation of model performance but also ensure the generalizability of our findings across diverse real-world scenarios.

2.1 Model Architectures

2.1.1 DSRNet

This model introduces a learnable residue term to the superimposition model, capturing residual information during layer separation. It proposes a dual-stream network with a novel interaction mechanism and a semantic pyramid encoder for effective decomposition.

Key Contributions:

- Provides a more flexible superposition model for handling various reflection scenarios.
- Enhances feature interaction through a mutually-gated interaction block.
- Demonstrates superior performance on real-world images for single image reflection separation (SIRS).

2.2 IBCLN

This model focuses on iteratively refining the transmission and reflection predictions,using LSTM for stable training, and employing a new loss function to better reconstruct the target layers. It uses a cascaded approach where transmission and reflection layers are refined iteratively to transfer information across cascade steps and address the vanishing gradient problem during training.

Key Contributions:

- Proposes s novel loss function called Reconstruction loss which helps in training over multiple cascade steps and guides more accurate estimates
- The authors created a new dataset with real-world images and ground-truth transmission layers to improve training data insufficiency.

3 Result Analysis

3.1 Experimental Setup

We conducted comprehensive evaluations using a diverse dataset comprising various real-world and benchmarking images. The dataset included 45

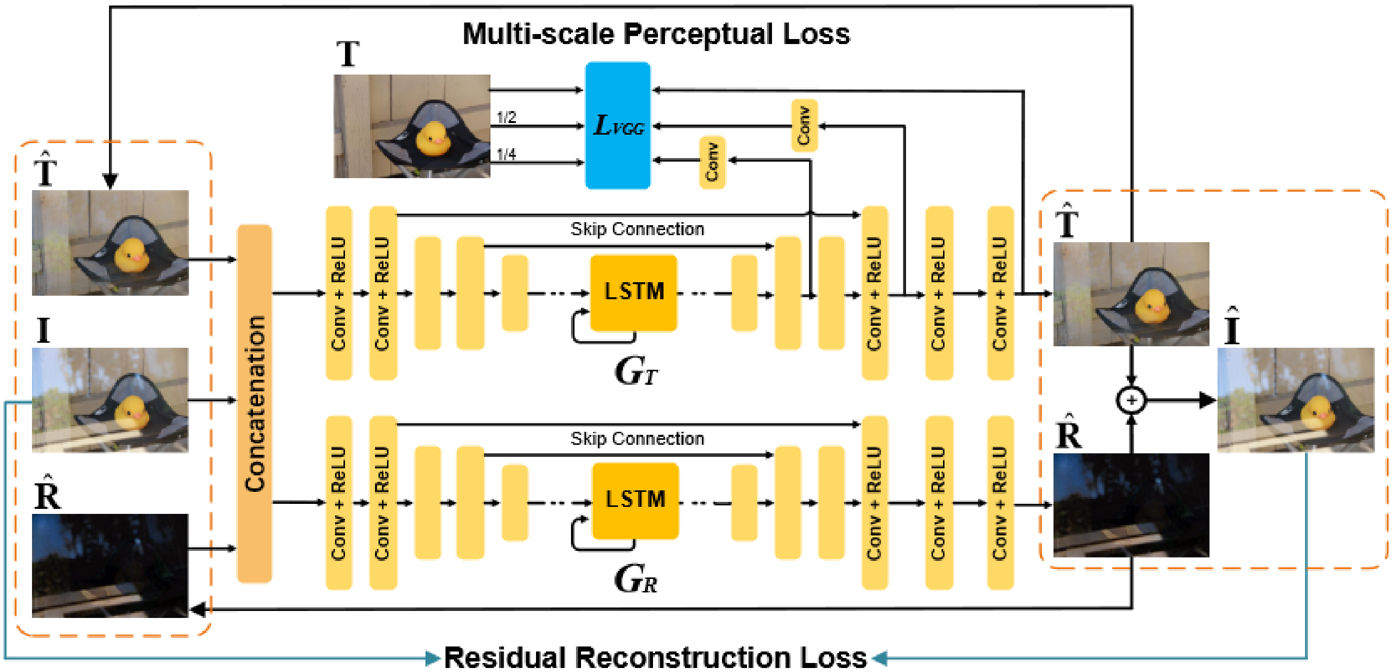

Figure 2: The IBLCN Architecture [6].

real-world testing images sourced from the CEILNet dataset, complemented by 20 pairs from Single Image Reflection Separation with Perceptual Losses [4] and an additional 20 pairs from IBCLN [6] for comparative analysis. Furthermore, we utilized 454 pairs from the SIR^2 dataset [2], categorized into three subsets: Objects (200 pairs), Postcard (199 pairs), and Wild (55 pairs).

Our experiments were performed on Kaggle's GPU P100, complemented by available CPU resources and 8GB of RAM. . The software environment utilized included Python 3.9 for scripting and PyTorch for implementing our reflection removal algorithms. We employed the Anaconda distribution for managing virtual environments, ensuring consistency and reproducibility in our experiments. Additionally, we utilized TensorBoardX for logging.

For testing on user data and implementing our application,we used a system equipped with a NVIDIA GTX 1050 Ti GPU featuring 4GB of VRAM, alongside a 7th generation Intel i3 CPU with 4 cores and 8GB of RAM.

3.2 Evaluation Metrics

For performance evaluation in reflection removal, we employed two key metrics: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM).

Peak Signal-to-Noise Ratio (PSNR) quantifies image quality by measuring the mean squared error (MSE) between the original and processed images, expressed in decibels (dB):

$$\text{PSNR} = 10 \cdot \log_{10}\left(\frac{\text{MAX}^2}{\text{MSE}}\right)$$

where MAX is typically 255 for 8-bit images. A higher PSNR indicates minimal distortion, crucial for assessing the fidelity of reflection removal algorithms.

Structural Similarity Index (SSIM) evaluates image similarity based on luminance, contrast, and structure:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

where:

- μ_x and μ_y are the average values of images x and y, respectively.
- σ_x^2 and σ_y^2 are the variances of images x and y, respectively.
- σ_{xy} is the covariance between x and y.
- C_1 and C_2 are small constants to avoid division by zero.

SSIM values range from -1 to 1, with 1 indicating perfect image similarity. SSIM assesses how well reflection removal algorithms preserve image features, crucial for maintaining visual fidelity.

3.3 Qualitative Analysis

The evaluation results in Table1 illustrate that DSRNet consistently outperforms IBCLN across various evaluation metrics for reflection removal. Visually, DSRNet shows superior performance, particularly evident in the higher average scores compared to IBCLN as shown in 3 . DSRNet achieves better preservation of image quality, as indicated by its higher average PSNR and SSIM values, which suggest less distortion and better structural similarity compared to IBCLN. These findings highlight DSRNet's efficacy in maintaining image fidelity while effectively removing unwanted reflections, making it a compelling choice for applications demanding high-quality image processing.

4 Prototype

As for our project, we have developed applications for both desktop and mobile platforms, using the trained models - DSRNet [10] and IBCLN [6].

4.1 Usage

To utilize the ReflectionEraser application, users begin by selecting an image. Upon uploading the selected image to the server, the application initiates rapid processing, typically completing within approximately 20 seconds. Users are then presented with the enhanced output, affording them options to save, share, or further refine the edited image to meet their specific requirements.

4.2 Desktop Application

The desktop counterpart of ReflectionEraser is developed using Python, leveraging the robust capabilities of PyTorch and additional machine learning libraries. The user interface is implemented using PyQt6, offering a sleek and minimalist design that prioritizes ease of use and functionality. The application includes dedicated folders for both DSRNet and IBCLN models, each showcasing distinct performance characteristics and output nuances.Depending on the command used to launch the application, either model can be selected for execution. The prototype is shown by 4

4.3 Mobile Application

The mobile application5 employs a client-server architecture designed to provide an intuitive interface, allowing end-users to appreciate the efficacy and practicality of the DSRNet model for reflection removal.The prototype is shown by 5

The server-side infrastructure is implemented using Flask, Python, PyTorch, and a comprehensive suite of machine learning libraries, creating a robust back-end that serves as a dedicated ReflectionEraser API. This API is designed to support a wide range of use cases and applications. On the client side, the application is developed with Flutter, using the Dart programming language to ensure seamless cross-platform compatibility. The application also incorporates functionality for capturing images directly from the phone's camera, providing users with the ability to easily upload images for processing.

5 Conclusion

This project successfully develops a desktop application that removes reflections from images, significantly enhancing their clarity and usability. Leveraging an advanced dual-stream network architecture, our model effectively separates reflection and transmission layers through innovative techniques such as the semantic pyramid encoder and a learnable residue term. The incorporation of the Mutually-Gated Interaction (MuGI) mechanism further improves feature extraction and separation, resulting in superior reflection-free images.

Despite these advancements, several limitations and challenges were encountered during the development and evaluation of our application. One notable limitation includes the computational intensity of the dual-stream network architecture, which may restrict real-time application on lower-end hardware. Furthermore, the model's performance can be sensitive to variations in lighting conditions and the complexity of reflection patterns, which may necessitate further refinements and adaptations for diverse real-world scenarios.

Looking forward, future work should focus on addressing these limitations and expanding the application's capabilities. Enhancements could involve optimizing the model for efficiency without compromising performance, exploring adaptive techniques to handle diverse reflection types more effectively, and integrating real-time processing capabilities to broaden its practical utility. Additionally, extending the dataset diversity and size used for training could further improve the model's generalization across different environments and scenarios.

In conclusion, this project underscores the transformative potential of advanced machine learning

Table 1: Comparison of evaluation metrics between DSRNet and IBCLN

Methods	Real20 (20)	Objects (200)	Postcard (199)	Wild (55)	Average
	PSNR / SSIM	PSNR / SSIM	PSNR / SSIM	PSNR / SSIM	PSNR / SSIM
IBCLN	21.86 / 0.762	24.87 / 0.893	23.39 / 0.875	24.71 / 0.886	24.10 / 0.879
DSRNet	24.23 / 0.820	26.28 / 0.914	24.56 / 0.908	25.68 / 0.896	25.40 / 0.905

(a) Original Image (b) ReflectionEraser using DSRNet (c) ReflectionEraser using IBCLN

Figure 3: Visual result comparison between DSRNet and IBCLN

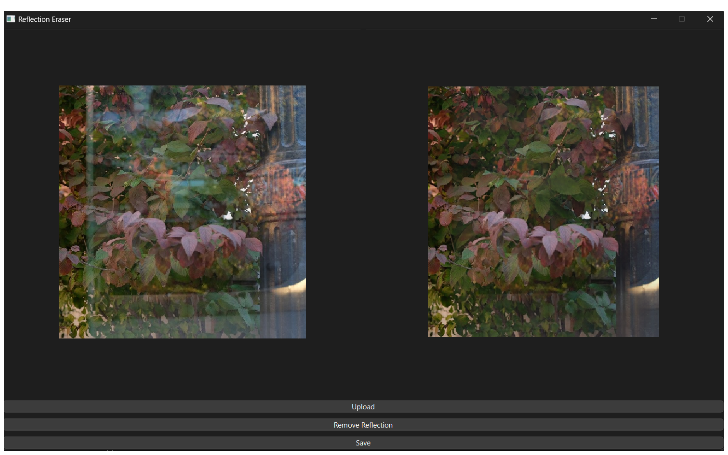

Figure 4: Prototype of the ReflectionEraser desktop application

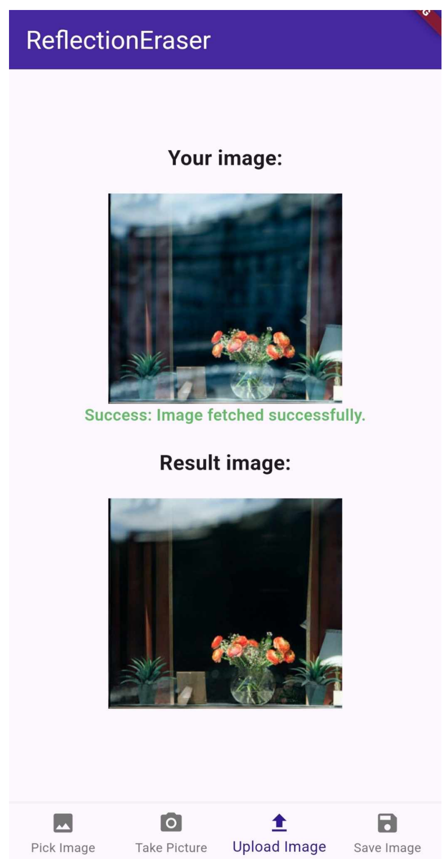

Figure 5: Prototype of the ReflectionEraser mobile application

techniques in practical image processing applications. By overcoming existing challenges in reflection removal from single images, our work contributes to advancing the state-of-the-art and opens new avenues for research and application in the field of computer vision.

References

[1] B. Hariharan, P. A. Arbeláez, R. B. Girshick, and J. Malik, "Hypercolumns for object segmentation and fine-grained localization," *CoRR*, vol. abs/1411.5752, 2014. arXiv: 1411.5752. [Online]. Available: http://arxiv.org/abs/1411.5752.

[2] R. Wan, B. Shi, L.-Y. Duan, A.-H. Tan, and A. C. Kot, "Benchmarking single-image reflection removal algorithms," in *International Conference on Computer Vision (ICCV)*, 2017.

[3] J. Yang, D. Gong, L. Liu, and Q. Shi, "Seeing deeply and bidirectionally: A deep learning approach for single image reflection removal: 15th european conference, munich, germany, september 8–14, 2018, proceedings, part iii," in Sep. 2018, pp. 675–691, ISBN: 978-3-030-01218-2. DOI: 10.1007/978-3-030-01219-9_40.

[4] X. Zhang, R. Ng, and Q. Chen, *Single image reflection separation with perceptual losses*, 2018. arXiv: 1806.05376 [cs.CV]. [Online]. Available: https://arxiv.org/abs/1806.05376.

[5] Q. Wen, Y. Tan, J. Qin, W. Liu, G. Han, and S. He, "Single image reflection removal beyond linearity," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 3766–3774. DOI: 10.1109/CVPR.2019.00389.

[6] C. Li, Y. Yang, K. He, S. Lin, and J. E. Hopcroft, *Single image reflection removal through cascaded refinement*, 2020. arXiv: 1911.06634 [cs.CV]. [Online]. Available: https://arxiv.org/abs/1911.06634.

[7] Z. Dong, K. Xu, Y. Yang, H. Bao, W. Xu, and R. W. Lau, "Location-aware single image reflection removal," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct. 2021, pp. 5017–5026.

[8] Q. Hu and X. Guo, *Trash or treasure? an interactive dual-stream strategy for single image reflection separation*, 2021. arXiv: 2110.10546 [cs.CV]. [Online]. Available: https://arxiv.org/abs/2110.10546.

[9] Y. Li, M. Liu, Y. Yi, Q. Li, D. Ren, and W. Zuo, *Two-stage single image reflection removal with reflection-aware guidance*, 2021. arXiv: 2012.00945 [cs.CV]. [Online]. Available: https://arxiv.org/abs/2012.00945.

[10] Q. Hu and X. Guo, *Single image reflection separation via component synergy*, 2023. arXiv: 2308.10027 [cs.CV]. [Online]. Available: https://arxiv.org/abs/2308.10027.