Survey on Advanced Visibility Restoration and Object Segmentation Techniques in Adverse Weather Conditions

Saba Attar[1] Samiksha More[2] Snehal Bhujbal[3] Prof. Y. R. Khalate[4] Prof. J. H. Shaikh[5]
[1,2,3]Student [4,5]Professor
[1,2,3,4,5]Department of Computer Science & Engineering
[1,2,3,4,5]SVPM's College of Engineering, Malegaon (Bk), Maharashtra, India

Abstract — This paper provides a survey of three state-of-the-art techniques focusing on visibility restoration in diverse weather conditions, optical flow estimation in dense fog, and semi-supervised video object segmentation. The unified approach towards weather visibility restoration, the use of semi-supervised learning for optical flow in foggy environments, and the cyclic mechanism for video object segmentation are explored. These methods are compared in terms of system architecture, datasets, and performance metrics. Through this survey, we aim to provide insights into advancements in these domains, highlighting future trends and areas for improvement.

Keywords: Visibility Restoration, Optical Flow, Foggy Scenes, Semi-Supervised Learning, Video Object Segmentation, Adverse Weather, Deep Learning, Cyclic Mechanism

I. INTRODUCTION

Restoring visibility in images and videos captured in adverse weather conditions is crucial for numerous computer vision applications, including autonomous driving, surveillance, and object detection. Adverse climate conditions, similar as fog, rain, and haze, significantly degrade the quality of visual data. Additionally, dense fog conditions introduce complexities in optical flow estimation, affecting the accuracy of motion estimation and tracking systems.

Moreover, video object segmentation (VOS), where the goal is to separate foreground objects from the background, is an essential task in video analysis. The introduction of semi-supervised learning and cyclic mechanisms in video object segmentation shows potential in improving accuracy, particularly when large labeled datasets are unavailable.

This survey reviews three key papers, each addressing specific challenges in these domains:

1) Unified Multi-Weather Visibility Restoration.[1]
2) Optical Flow in Dense Hazy Scenes via Semi-Supervised Learning.[2]
3) Exploring the Cyclic Mechanism for Semi-Supervised Video Object Segmentation.[3]

II. LITERATURE SURVEY

A. Unified Multi-Weather Visibility Restoration.[1]

This paper presents a novel approach for visibility restoration under multiple weather conditions, including fog, haze, and rain. Unlike earlier approaches focused on specific weather conditions, this unified model is capable of handling a range of adverse weather scenarios. The technique uses deep learning to estimate visibility degradation and then applies a restoration model to improve the image quality.

The primary contributions of this paper include:

- A unified model for handling multiple weather conditions.
- Integration of both prior-based and learning-based approaches for improved accuracy.
- Use of multiple datasets representing diverse weather conditions.

B. Optical Flow in Dense Hazy Scenes via Semi-Supervised Learning.[2]

The paper addresses the challenge of estimating optical flow in dense foggy scenes using a semi-supervised learning approach. Optical flow plays a vital role in estimating motion within video sequences. However, dense fog introduces complex visibility issues that degrade the accuracy of traditional methods.

The proposed method leverages semi-supervised learning to reduce the need for large labeled datasets, which are difficult to obtain in foggy environments. Key contributions include:

- A semi-supervised learning framework to handle dense foggy scenes.
- Use of synthetic fog datasets for training, with real-world datasets for testing.
- Improved accuracy over existing methods for optical flow estimation under dense fog.

C. Exploring the Cyclic Mechanism for Semi-Supervised Video Object Segmentation.[3]

In this paper, the authors propose a cyclic mechanism for semi-supervised video object segmentation. The cyclic nature of the model ensures that the segmentation model can revisit frames to refine its output, thereby improving accuracy over time. The model excels at segmenting objects in videos where only sparse annotations are available.

The main contributions include:

- A novel cyclic mechanism to revisit frames for refining segmentation.
- Effective handling of semi-supervised scenarios with minimal annotated data.
- Significant improvement over baseline methods on standard benchmarks.

III. METHODOLOGY

A. Unified Multi-Weather Visibility Restoration.

The architecture is composed of two primary components: a weather classification module and a restoration module. The weather classification module predicts the type of adverse weather condition affecting the image, while the restoration module applies a learned mapping to improve visibility. Both modules are deep learning-based and trained using diverse weather datasets.

The proposed unified network has two key streams:

1) Degradation Removal Stream (DRS), which removes rain and snow streaks.
2) Feature Corrector Stream (FCS), which refines the coarse output from DRS.

The fusion of multi-scale features is captured through the Adaptive Multi-scale Feature Sharing Block (AMFSB):

$$AFB(f_j, f_k) = \lambda \times f_j + (1-\lambda) \times f_k$$

Where λ is a trainable weight optimized via the sigmoid function.

Performance: Achieved PSNR and SSIM improvements compared to state-of-the-art methods on various datasets (rain, snow, and veiling effects).[1]

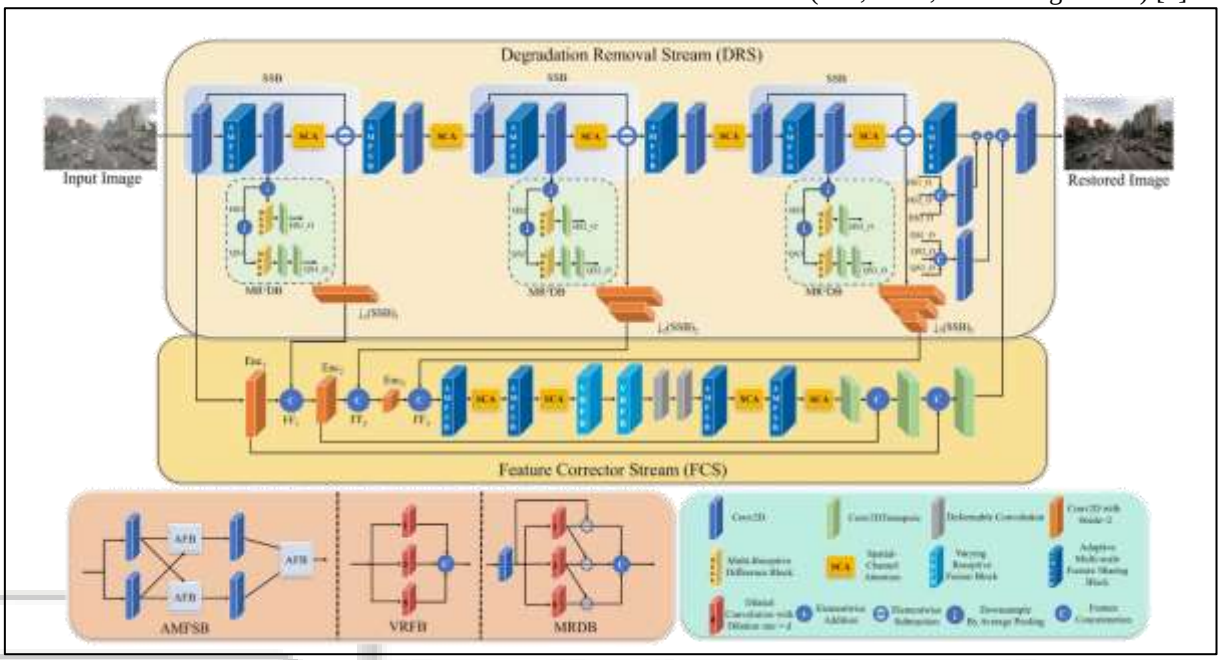

Fig. 1: System Architecture [1]

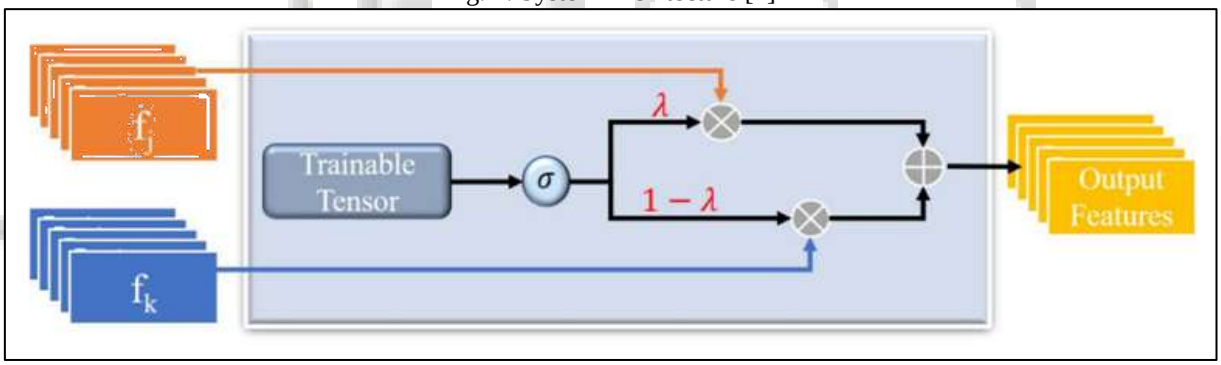

Fig. 2: Adaptive Feature Block

B. Optical Flow in Dense Hazy Scenes via Semi-Supervised Learning.

The architecture follows a semi-supervised learning approach, where an encoder-decoder network estimates the optical flow. The semi-supervised nature of the model allows it to learn from both labeled and unlabeled data. The training involves synthetic fog datasets, and testing is performed on real-world datasets to validate the model's generalizability.

Optical flow accuracy is evaluated using End-Point Error (EPE). The EPE loss quantifies the discrepancy between the predicted flow and the ground-truth flow:

$$L_{EPE} = \mathrm{E}(|| \widehat{o_f} - o_f ||^2)$$

Where $\widehat{o_f}$ is the predicted optical flow and o_f is the ground truth.

The domain transformation ensures robust learning from both clean and foggy images. Photometric and hazeline constraints further enhance performance in dense foggy conditions.

EPE Results: On synthetic fog datasets, the proposed method outperformed existing methods with significant reduction in EPE compared to PWCNet and other supervised models.[2]

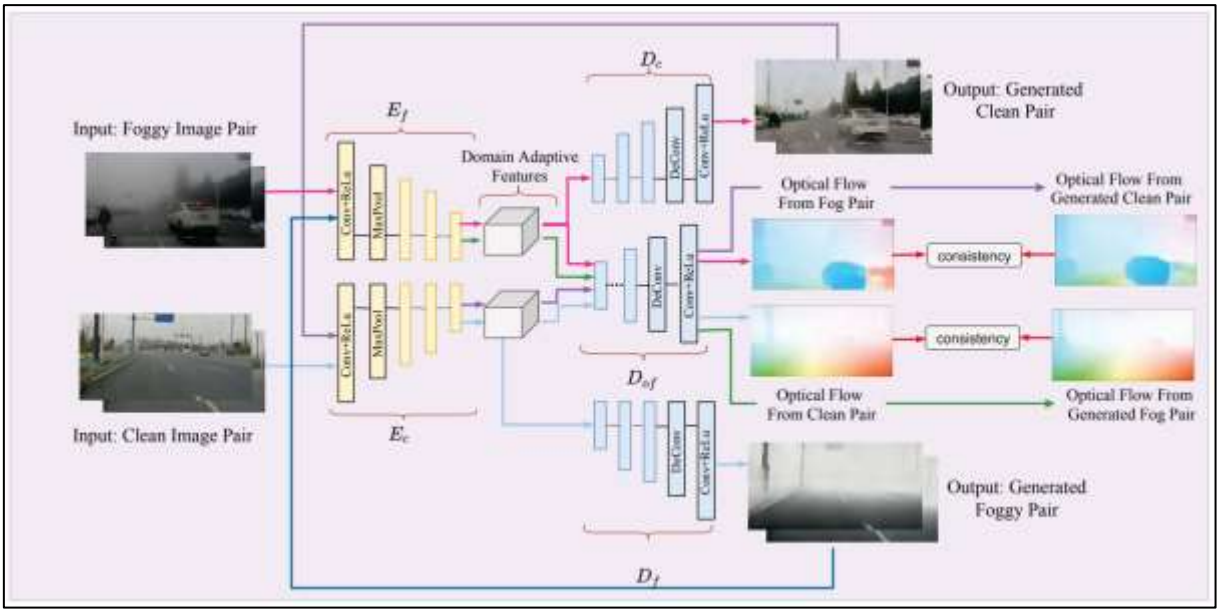

Fig. 3: System Architecture [2]

C. Exploring the Cyclic Mechanism for Semi-Supervised Video Object Segmentation.

The architecture consists of a cyclic feedback loop where the model revisits earlier frames to refine its segmentation. The model is based on a convolutional neural network (CNN) and relies on temporal information across frames to improve segmentation accuracy over time.

The cyclic mechanism is formulated using a forward-backward flow (cyclical structure) to bridge the gap between initial and intermediate reference masks, reducing error propagation.

Loss Function:

$$L_{cycle,t} = L(\hat{Y}_t, Y_t) + L(\hat{Y}_1, Y_1)$$

Where:

$L(\hat{Y}_t, Y_t)$ and $L(\hat{Y}_1, Y_1)$ are the cross-entropy losses for the predicted and reference masks. IOU (Intersection Over Union) and cross-entropy losses are combined for both forward and backward directions to minimize error propagation across video frames.[3]

Performance on DAVIS17 and YouTube-VOS:

Evaluation Metrics (Jaccard Index and Contour F-score - J&F):

- STM-cycle: J&F = 71.7, FPS = 38
- STM-cycle + Gradient Correction: J&F = 72.3, FPS = 9.3

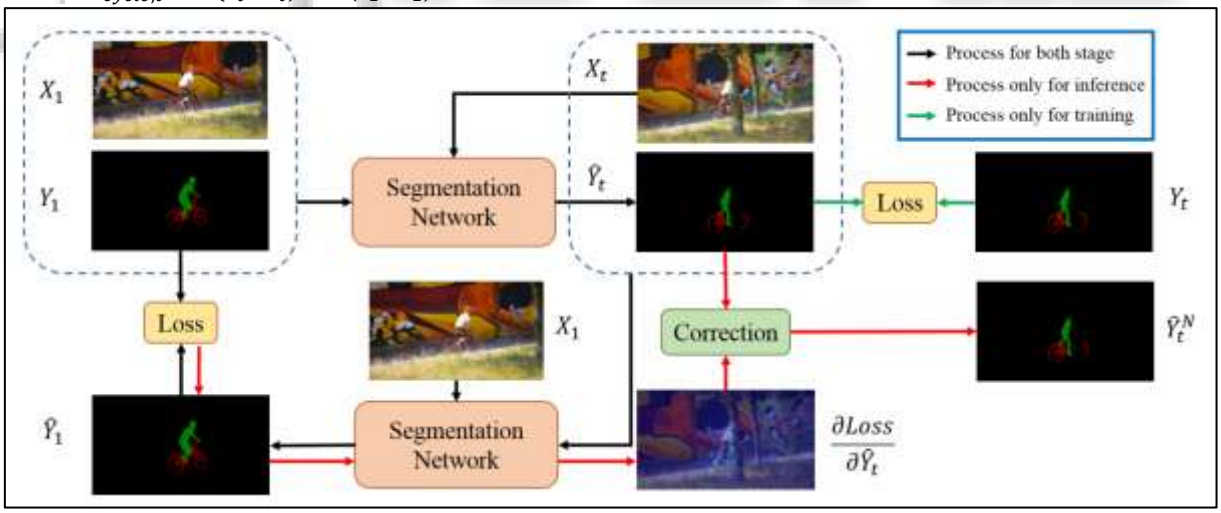

Fig. 4: System Architecture [3]

IV. DATASET

The datasets used in these papers are diverse and tailored to the specific challenges they address:

A. Unified Multi-Weather Visibility Restoration:

The model is trained on a mixture of datasets representing different weather conditions, including fog, haze, and rain datasets.[1]

B. Optical Flow in Dense Hazy Scenes via Semi-Supervised Learning.

The training is performed on synthetic fog datasets, and the model is tested on real-world foggy scenes.[2]

C. Cyclic Mechanism in Semi-Supervised Video Object Segmentation:

The model uses standard video segmentation datasets, such as DAVIS and YouTube-VOS, with semi-supervised annotations.[3]

V. RESULT ANALYSIS

Paper	Task	Metric	Result	Additional Notes
Unified Multi-Weather Visibility Restoration (Kulkarni et al.)	DIDMDN, Snow100K, SRRS, Outdoor-Rain. Multi-weather visibility restoration	1.PSNR (Peak Signal-to-Noise Ratio) 2.SSIM (Structural Similarity Index).	PSNR - Higher than state-of-the-art methods for rain, snow, haze. SSIM - Significant improvement over previous methods	Robust multi-weather image restoration with lower computational complexity. Single network handles multiple weather conditions with high efficiency
Optical Flow in Dense Foggy Scenes (Yan et al.)	Synthetic & Real Fog Datasets Optical Flow Estimation	1.EPE (End-Point Error) 2.Visual Comparison	EPE - Significant reduction compared to PWCNet. Improved optical flow accuracy in foggy scenes.	Robust performance in real foggy scenes with unsupervised learning. Outperforms baseline methods like PWCNet visually and quantitatively.
Cyclic Mechanism in VOS (Li et al.)	DAVIS17 / YouTube-VOS Video Object Segmentation	1.J&F (Jaccard & F-score) 2. FPS (Frames per Second)	71.7 (STM-cycle) 72.3 (STM-cycle+GC) 38 (STM-cycle) 9.3 (STM-cycle+GC)	FPS drops with Gradient Correction due to additional computation. FPS drops with Gradient Correction due to additional computation.

Fig. 5: Result Table

VI. FUTURE SCOPE

Future research in video de-hazing holds promising potential to enhance model performance and adaptability across diverse scenarios. One key area of improvement is in achieving smoother frame-to-frame transitions, particularly in scenes with fast motion. By focusing on temporal consistency, future models could reduce flickering and other visual inconsistencies, making the de-hazed video output more stable and natural.

Another important direction is the development of adaptive filtering techniques. Current models often apply uniform filtering across the frame, but future models could use intelligent, region-specific filters that automatically adjust based on haze intensity. This would allow for better visibility in videos with varied haze levels, ensuring clarity in both heavily and lightly obscured areas.

Additionally, expanding the models to handle multiple weather conditions—such as rain, snow, and fog—could make them more robust and versatile. A single, multi-

weather model would be more efficient for real-world applications, reducing the need for multiple specialized models.

Real-time processing is another promising goal. Optimizing these models for faster, lighter operation would allow deployment in critical areas such as surveillance, autonomous driving, and environmental monitoring, where timely information is crucial. Such advancements could enable consistent, reliable performance in adverse weather, ensuring that key applications remain operational in challenging conditions.

Finally, enhancing edge and detail clarity in de-hazed videos could significantly improve the realism and visual quality of the output. By refining feature extraction and preservation, future models could retain more fine details, making videos appear sharper and more accurate. This would be particularly beneficial in applications where precise visual information is essential, such as object detection and scene analysis in low-visibility conditions.

Together, these advancements could make video de-hazing models far more adaptable, efficient, and effective across a wide range of real-world scenarios.

VII. Conclusion

The survey of these papers shows progress in restoring visibility under challenging weather, estimating optical flow in tough conditions, and segmenting video objects accurately. A unified approach for visibility restoration works well across multiple weather types.

In optical flow estimation, using semi-supervised learning reduces the need for extensive labeled data, making it particularly effective in foggy scenes. This approach highlights how machine learning can handle complex tasks even with limited resources.

Finally, the cyclic mechanism in video object segmentation brings a new way to refine results frame by frame. Future research could work on making these models even faster and stronger to handle more challenging situations.

References

[1] A. Kulkarni, P. W. Patil, S. Murala, and S. Gupta, "Unified Multi-Weather Visibility Restoration," IEEE, 2023.
[2] W. Yan, A. Sharma, and R. T. Tan, "Optical Flow in Dense Foggy Scenes using Semi-Supervised Learning," National University of Singapore, 2020.
[3] Y. Li, N. Xu, J. Peng, J. See, and W. Lin, "Delving into the Cyclic Mechanism in Semi-Supervised Video Object Segmentation," Shanghai Jiao Tong University, 2020.
[4] J. Zhang, W. Ren, S. Zhang, H. Zhang, Y. Nie, Z. Xue, and X. Cao, "Hierarchical density-aware dehazing network," IEEE Transactions on Cybernetics, 2021.
[5] D. Liu, B. Wen, Y. Fan, C. C. Loy, and T. S. Huang, "Non-local recurrent network for image restora- tion," Advances in Neural Information Processing Systems, vol. 2018, pp. 1673–1682, 2018.
[6] H. Dong, J. Pan, L. Xiang, Z. Hu, X. Zhang, F. Wang, and M.-H. Yang, "Multi-scale boosted dehazing network with dense feature fusion," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 2157–2167, 2020.
[7] R. Li, R. T. Tan, and L.-F. Cheong, "All in one bad weather removal using architectural search," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 3175– 3185, 2020.
[8] W.-T. Chen, H.-Y. Fang, J.-J. Ding, and S.-Y. Kuo, "PMHLD: patch map-based hybrid learning de- hazenet for single image haze removal," IEEE Transactions on Image Processing, vol. 29, pp. 6773– 6788, 2020.
[9] W. Ren, J. Zhang, X. Xu, L. Ma, X. Cao, G. Meng, and W. Liu, "Deep video dehazing with semantic segmentation," IEEE Transactions on Image Processing, vol. 28, no. 4, pp. 1895–1908, 2018.
[10] H. Zhang and V. M. Patel, "Densely connected pyramid dehazing network," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3194–3203, 2018.

