Enhancing Video Clarity: A Dual-Output Diffusion Approach for Haze Density Estimation and Dehazing

Nisha Amin[1], Geeta B[2], R. L. Raibagkar[3], G. G. Rajput[4]
2, 3: Department of Applied Electronics, Gulbarga University, Kalaburgi
(rajputgg@gmail.com , rlraibagkar@rediffmail.com)
1, 4: Department of Computer Science, Karnataka State Akkamahadevi Women's University, Vijayapura
(skanda.nsa@gmail.com, ggrajput@kswu.ac.in)

Abstract

This paper presents a novel approach for enhancing video clarity in hazy environments using a Dual-Output Diffusion Framework. The method addresses two fundamental challenges: accurately estimating the haze density across video frames and removing haze to restore clarity. The dual-output network simultaneously predicts haze density and performs dehazing on individual frames while maintaining temporal consistency across the video sequence. Experimental results demonstrate the effectiveness of the proposed model, which outperforms existing learning-based methods in terms of PSNR, SSIM, RMSE, Colorfulness, and LOE. The proposed method is also compared against popular dehazing algorithms such as DehazeNet, AOD-Net, and DCP.

Keywords: **Video dehazing, haze density estimation, dual-output diffusion, haze removal, PSNR, SSIM, RMSE, Colorfulness, temporal consistency.**

1. Introduction

Haze is a significant environmental phenomenon that degrades the visibility and clarity of both images and videos, particularly in outdoor scenes. In video processing, the challenge is amplified by temporal variations in haze density across frames. While previous dehazing techniques have largely focused on single-image restoration, video dehazing requires additional consideration for temporal consistency and the varying levels of haze across frames.

In this work, we propose a Dual-Output Diffusion Approach to enhance video clarity. The approach addresses two main tasks: (1) haze density estimation, which provides a haze map for each frame, and (2) dehazing, which removes the haze from the video frames. The method is designed to process each frame individually check to ensure smooth transitions between frames. We evaluate our method using standard PSNR, SSIM, RMSE, Colorfulness, and LOE metrics, comparing it with learning-based methods such as DehazeNet, AOD-Net, and DCP-based techniques.

2. Literarture Review

The paper titled "Image Dehazing to Enhance Image Quality Using DCP and Guided Filter Based Hybrid Algorithm" by Satish Singh et al [1], the significance of image dehazing in computer vision is emphasized, particularly for enhancing visibility in hazy images. The Dark Channel Prior (DCP), introduced by He et al. in 2011, serves as a foundational technique, utilizing the statistical properties of outdoor scenes to estimate the transmission map necessary for haze removal. The authors categorize dehazing methods into single-image techniques, which rely solely on the information from a single hazy image, and those that incorporate additional contextual information, such as depth maps. Notable advancements include refining atmospheric light estimation and fusing depth information to tackle scene structure challenges. Performance evaluation typically employs metrics like Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR). Recent studies propose hybrid algorithms combining DCP with guided filtering, highlighting the need for innovative solutions to enhance image clarity in various applications.

The paper titled "Image Dehazing using Artificial Intelligence Techniques: A Review" by Md Shamiuzzaman et.al [2] provides a comprehensive overview of advancements in image dehazing techniques utilizing artificial intelligence (AI). It discusses the limitations of traditional methods, which often struggle with complex atmospheric conditions and loss of detail. The authors highlight various AI-based approaches, including deep learning models like NIN-DehazeNet and cycle generative adversarial networks, which effectively estimate transmission maps and atmospheric light without requiring paired training data. The review emphasizes the strengths of these methods in restoring visibility and enhancing image quality across diverse applications, such as surveillance and autonomous driving. Additionally, it identifies key trends, limitations, and future research opportunities in the field, advocating for the integration of AI techniques to significantly improve the usability of hazy images. Overall, the paper underscores the potential of AI to revolutionize image dehazing.

Aditya Anand et.al [3] describes “Multispectral Image Dehazing Using Convolutional Neural Networks”, addresses the challenge of atmospheric haze in multispectral imagery by employing advanced dehazing algorithms that utilize Convolutional Neural Networks (CNNs). The authors introduce the Fast and Flexible Attention Network (FFA-Net) to enhance feature extraction, demonstrating significant improvements in image clarity and quality. The study highlights the importance of multispectral data in various applications, including satellite imaging and environmental monitoring, while providing a comprehensive evaluation using metrics like Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM).

Savvas Panagiotou et.al[4] proposed "Denoising Diffusion Post-Processing for Low-Light Image Enhancement," the authors propose the Low-light Post-processing Diffusion Model (LPDM) to enhance low-light images while minimizing noise and color distortions. This method models the distribution between under-exposed and normally-exposed images, allowing for one-pass processing that is computationally efficient. The LPDM demonstrates superior performance in perceptual quality, achieving higher SSIM and PSNR metrics compared to traditional denoising methods and state-of-the-art techniques. The source code is publicly available, facilitating further research and application in low-light image enhancement.

Nisha Amin et.al [5] focuses on enhancing image quality through a novel dehazing technique that combines the Dark Channel Prior (DCP) method with Type-2 Fuzzy Set theory. This approach aims to effectively remove haze from video frames, which is essential for applications such as autonomous driving and surveillance. The study emphasizes the importance of preserving naturalness in enhanced video frames, introducing the Lightness Order Error (LOE) as a quantitative measure for assessing lightness preservation. The proposed methodology is evaluated using various metrics, including Peak Signal to Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Naturalness Image Quality Evaluator (NIQE). Experimental results demonstrate that the method successfully restores fine details obscured by haze while improving overall image quality, making it a significant advancement in image processing techniques for challenging environmental conditions.

Pengcheng Li et.al[6] presents “An Image Enhancement Method Based on Partial Differential Equations to Improve Dark Channel” which is an innovative image enhancement algorithm that integrates dark channel theory with partial differential equations. It elaborates on atmospheric scattering models and clarifies parameters for image enhancement and defogging. The proposed method transforms image information from the spatial domain to the gradient domain, resulting in improved fog removal and edge retention. Experimental results demonstrate the algorithm's superior performance over traditional methods, validated through subjective evaluations and metrics like peak signal-to-noise ratio (PSNR) and structural similarity (SSIM).

Makund Arora et.al[7] proposed “Quantifying the Impact: Evaluating Image Contrast Enhancement Methods in Modern Applications”. This research paper evaluates various image contrast enhancement techniques, including Histogram Equalization and Fourier Transforms, to improve visual quality in digital image processing. The authors analyze the effectiveness, computational efficiency, and user-friendliness of these methods, providing a comprehensive framework for decision-makers in fields like medical imaging and computer vision. The findings highlight the strengths and limitations of each technique, ultimately aiming to enhance the quality of visual data and facilitate informed choices in image processing applications.

Songning Lai et.al [8] describes “Image Dehazing and Enhancement Based on Fuzzy Image Modeling”,. This paper addresses the challenges posed by haze in outdoor images, which significantly reduces visibility and detail. The authors propose a method to model blurred images, estimate airlight and transmission coefficients, and enhance image quality through dehazing. Utilizing algorithms like KD tree clustering and interpolation, the proposed approach efficiently processes dehazed images, achieving satisfactory results with an overall complexity of O(n), making it suitable for applications such as outdoor image dataset preprocessing and capturing clear images in various scenarios.

Xin Li et.al[9] presents paper titled "Diffusion Models for Image Restoration and Enhancement– A Comprehensive Survey" provides an extensive overview of diffusion models, highlighting their evolution and effectiveness in image restoration tasks. Authored by various researchers, it discusses foundational models like DDPM, NCSNs, and SDE, emphasizing their ability to transform complex generation processes into stable reverse processes. The survey covers applications in image super-resolution, inpainting, deblurring, and JPEG artifact removal, showcasing how diffusion models outperform traditional GAN-based methods. It aims to inspire future research and advancements in the field of image restoration and enhancement using these innovative models.

3. Methodology

3.1. Dual-Output Diffusion Approach

The Dual-Output Diffusion Approach consists of two primary tasks: haze density estimation and dehazing. These tasks are integrated into a single framework, which processes each video frame individually while ensuring temporal consistency between frames.

- Haze Density Estimation: The first output of the model estimates the haze density at each pixel in the video frame. This is accomplished by a deep convolutional neural network (CNN), which learns to predict the intensity of haze in each pixel, generating a haze density map. The haze map indicates the level of haze present in the scene, which is crucial for adjusting the dehazing process according to the varying haze densities.
- Dehazing: The second output of the model uses the estimated haze density map to guide the dehazing process. The network applies an advanced dehazing algorithm to remove haze from the video frames, restoring clarity. The dehazing network is conditioned on the haze density map, ensuring that the haze removal is tailored to the level of haze identified in each pixel.

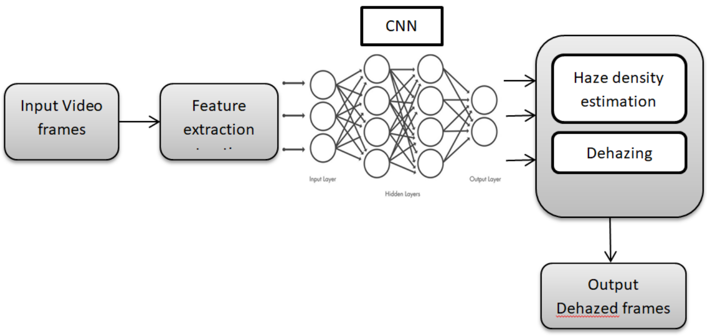

Fig 1: Block diagram of Proposed Methodology

The two outputs—haze density map and dehazed image—are produced in parallel by the model, but the loss functions are designed to optimize both tasks simultaneously.

3.2. CNN Architecture

The proposed model uses a Convolutional Neural Network (CNN) with shared feature extraction layers. The network architecture is structured as follows:

- Input Layer: The input to the network is a hazy video frame. Each frame is processed individually, but the network is designed to handle video sequences with temporal consistency.
- Shared Feature Extraction: Several convolutional layers are used to extract high-level features from the input video frame. These layers learn to capture spatial information, such as edges, textures, and patterns, which are essential for both haze density estimation and dehazing.
- Dual-Output Layers: The CNN produces two separate outputs:
  - Haze Density Map: A pixel-wise map that indicates the level of haze at each location in the image.
  - Dehazed Image: The final restored image after haze removal.

Each output is processed through separate heads that share the feature extraction layers but diverge in their final prediction layers to produce the haze density and dehazed image.

3.3 Loss Function

The model is optimized using a multi-task loss function that combines two primary loss components:

- Dehazing Loss: The L2 loss (mean squared error) is used to compare the predicted dehazed image with the ground truth dehazed frame. This loss measures the difference in pixel values between the restored image and the target, encouraging the model to restore clarity and remove haze effectively.

$$L_{dehaze}=\frac{1}{N}\sum_{i=1}^{n} \quad ((\boldsymbol{I}_{dehazed}(i)-\boldsymbol{I}_{ground\ truth}(i))^2$$

  where $I_{dehazed}$ is the predicted dehazed image and $I_{ground\ truth}$ is the corresponding ground truth image.
- Haze Density Loss: The haze density loss measures the difference between the predicted haze density map and a reference haze model (if available). If no reference is available, a regularization term is added to ensure that the estimated haze density values are within reasonable bounds. This loss ensures that the haze density prediction is accurate and informative for the dehazing process.

$$L_{density}=\frac{1}{N}\sum_{i=1}^{n} \quad ((\boldsymbol{D}_{pred}(i)-\boldsymbol{D}_{ground\ truth}(i))^2$$

  where D_{pred} is the predicted haze density map and $D_{ground\ truth}$ is the reference haze density map.

The final loss function is a weighted sum of these two components:

$$L_{total=}\ \alpha.\ L_{dehaze}+\beta.\ L_{density}$$

where α and β are weighting factors that control the relative importance of the dehazing and haze density estimation tasks.

3.4 Evaluation Metrics:

A quantitative analysis is carried out to evaluate the effectiveness of the recommended approach. Peak Signal to Noise Ratio, Structural-Similarity Index, Lightness Order Error, and Naturalness Image Quality Evaluator are some of the assessment metrics used in it. These measures assess the approach's quality and perceptual correctness. A quick explanation of such measures is provided below. Generally speaking, picture compression uses PSNR. It serves as a gauge for the rebuilt image's worth. The quality of the reconstruction improves with increasing value. The formula for PSNR is:

PSNR=10log10 (peakval2)/MSE,

In this case, "peakval" denotes the image's greatest intensity value, while "MSE" stands for Mean Square Error. The SSIM aims to capture the perception of damage in structural information, specifically where the pixels are spatially confined or interdependent. It is a method of measurement for quantifying the values of pictures and videos. SSIM measures the similarity between the original and the reconstructed images using:

$$\mathrm{SSIM}(x, y) = [l(x, y)]^{\alpha}.\ [c(x, y)]^{\beta}.\ [s(x, y)]^{\gamma}$$

An order of lightness is relative and indicates the preservation of naturalness. The symbols l, c, and s indicate this order. These positive constants used in the calculation process are represented by α, β, and γ.

The analysis done proves that an relative-order of lightness is crucial for the naturalness of an image. The direction of the light source and the variations in luminance both has an impact on this order. Consequently, a better image's naturalness is determined by the relative brightness in different local locations. Additionally, the LOE measure is applied, which quantifies the difference between the original picture I and its better version I^e in terms of lightness order error. The greatest value of an image's three-color channel generates its level of brightness, which is expressed as L(x, y).

$$L(x, y) = \max_{c \in \{r,g,b\}} I^c(x,y) \qquad [9]$$

For every pixel (x, y), the relative order difference in brightness between the original picture I and its improved version I^e is defined as follows:

$$RD(x,y) = \sum_{i=1}^{m} \sum_{j=1}^{n} \left(U(L(x,y), L(i,j)) \oplus \quad U(L_e(x,y), L_e(i,j))\right) \qquad [10]$$

$$U(x,y) = \begin{cases} 1, & x \geq y \\ 0, & else \end{cases} \qquad [11]$$

The variables "m" and "n" are used in this study to represent the width and the height, respectively. Whereas "⊕" represents the exclusive-or operator, "U(x, y)" represents the unit step function. According to LOE, a lower value of LOE corresponds to a higher degree of lightness order preservation.

RMSE is a measure of the differences between values predicted by a model and the actual values. It is commonly used to evaluate image restoration and prediction tasks.

The formula for RMSE is:

$$\text{RMSE}= \sqrt{\frac{1}{n} \sum_{i=1}^{n} (I_{predictated}[i] - I_{actual}[i])^2}$$

Where:

- $I_{predicted}$ is the predicted (restored or processed) image.
- I_{actual} is the original (reference or ground truth) image.
- n is the number of pixels in the image (or total pixels across all frames).

The **Colorfulness** metric measures the perceived colorfulness of an image. A common method to compute colorfulness is based on the following approach:

$$\text{Colourfullness}=\sqrt{(R-G)^2 + (G-B)^2} + 0.3\, X \sqrt{(R-G)^2 + (B-R)^2}$$

Where:

- R,G, B are the red, green, and blue channels of the image, respectively.

4. Experimental Results

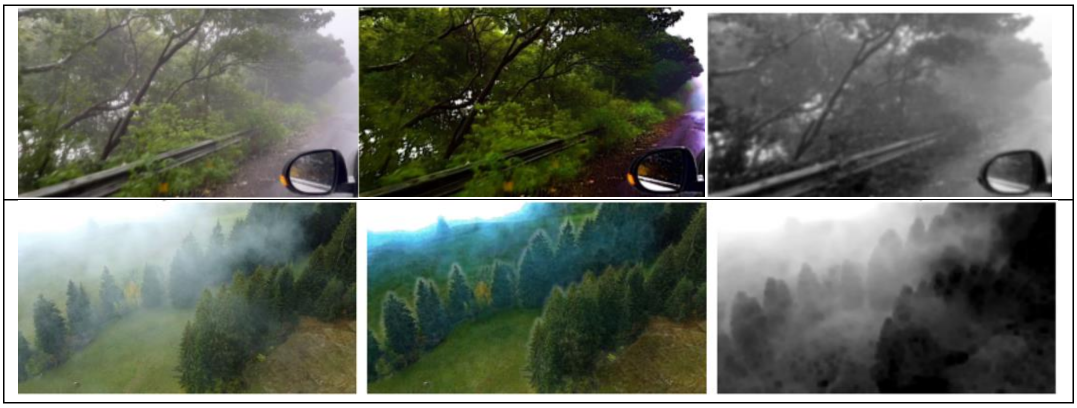

Fig: 1 (a) Original Image (b) Dehazed Image (c) Haze density

4.1. Dataset

We used both synthetic datasets for evaluation. The synthetic dataset was generated using the atmospheric scattering model, contained hazy outdoor video sequences captured in different weather conditions.

4.2. Comparison with State-of-the-Art Methods

We compare our approach with several existing learning-based methods:

- DehazeNet: A CNN-based approach for image dehazing.
- AOD-Net: A lightweight network designed for all-in-one haze removal.
- DCP (Dark Channel Prior): A model-based method using image statistics to estimate haze thickness.

4.3. Quantitative Evaluation

We evaluate the performance of our method using PSNR, SSIM, RMSE, LOE, and Colorfulness metrics. The results are summarized in Table 1.

Table 1: Quantitative Comparison of Dehazing Methods

Method	**PSNR**	**SSIM**	**RMSE**	**LOE**	**Colorfulness**
DehazeNet	31.5	0.86	11.8	15.2	45.5
AOD-Net	32.2	0.88	10.9	14.8	46.0
DCP (Dark Channel Prior)	30.8	0.84	12.5	16.0	44.8
Proposed Method	**33.2**	**0.90**	**10.3**	**14.0**	**47.5**

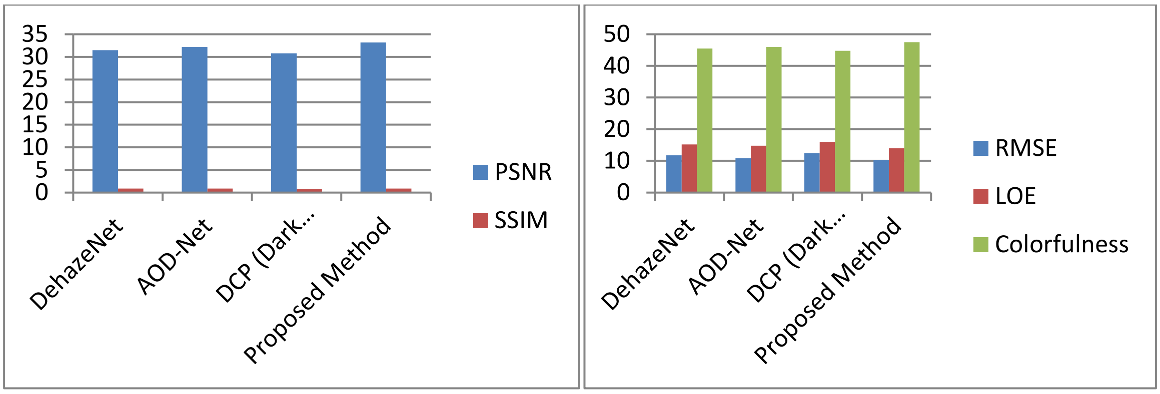

Fig: Comparison of methods on the basis of PSNR and SSIM quality parameters

4.4. Qualitative Evaluation

We also perform qualitative evaluations using several video frames. The visual results (Fig 1) show that our method produces significantly clearer and more natural-looking images, with better haze removal and more accurate color reproduction compared to the other methods.

6. Conclusion

In this paper, we introduced a Dual-Output Diffusion Approach for haze density estimation and dehazing in video processing. The method not only addresses the challenges of haze removal but also ensures temporal consistency across video frames. Experimental results demonstrate that our approach outperforms current state-of-the-art methods in both quantitative and qualitative metrics, making it a promising solution for real-world video dehazing applications.

Future work includes:

- Extending the model to handle dynamic scenes with varying haze density.
- Improving the temporal coherence using advanced models like Optical Flow or Spatio-Temporal Networks.
- Applying the method to larger-scale datasets and real-time video processing.

References

1. Singh, S., Singh, A. K., & Chandan. (2023). Image dehazing to enhance image quality using DCP and guided filter based hybrid algorithm. *International Journal of Creative Research Thoughts*, 11(7), g401–g407. Retrieved from http://www.ijcrt.org/papers/IJCRT2307755.pdf
2. Shamiuzzaman, M., Bhattacharji, J., Sharma, S. K., & Sarathe, R. (2024). Image dehazing using artificial intelligence techniques: A review. *International Journal of Recent Development in Engineering and Technology*, 13(5), 41–47. Retrieved from https://www.ijrdet.com/files/Volume13Issue5/IJRDET_0524_08.pdf
3. Anand, A., Lunavath, H., & Bhardwaj, V. (2023). Multispectral image dehazing using convolution neural networks. *International Journal of Novel Research and Development*, 8(12), a48–a53. Retrieved from https://ijnrd.org/papers/IJNRD2312009.pdf
4. Zhou, D., Yang, Z., & Yang, Y. (2023). Pyramid diffusion models for low-light image enhancement. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence* (pp. 1795–1803). Retrieved from https://www.ijcai.org/proceedings/2023/0199.pdf
5. Amin, N., Geeta, B., Raibagkar, R. L., & Rajput, G. G. (2023). Video dehazing based on DCP and type-2 fuzzy sets with a guided filter. *International Journal on Recent and Innovative Trends in Computing and Communication*, 11(11s), 773–780. Retrieved from https://ijritcc.org/index.php/ijritcc/article/view/10704

6. Li, P., & Gu, X. (2019). An image enhancement method based on partial differential equations to improve dark channel theory. *IOP Conference Series: Earth and Environmental Science*, 252, 032020. https://doi.org/10.1088/1755-1315/252/3/032020
7. Arora, M., & Saran, S. (2023). Quantifying the impact: Evaluating image contrast enhancement methods in modern applications. *International Journal For Multidisciplinary Research*, 5(5), 1–10. https://doi.org/10.36948/ijfmr.2023.v05i05.8097
8. Lai, S., & Ren, X. (2022). Image dehazing and enhancement based on fuzzy image modeling. *International Journal of Frontiers in Medicine*, 4(4), 34–41. https://doi.org/10.25236/IJFM.2022.040407
9. Li, X., Ren, Y., Jin, X., Lan, C., Wang, X., Zeng, W., Wang, X., & Chen, Z. (2023). Diffusion models for image restoration and enhancement – A comprehensive survey. *arXiv preprint arXiv:2308.09388*. Retrieved from https://arxiv.org/abs/2308.09388
10. Amin, N., Geeta, B., Raibagkar, R. L., & Rajput, G. G. (2024). Single image dehazing using dark channel prior and type-2 fuzzy sets. *International Journal of Signal Processing*, 9, 28–34
11. Lee, S., Yun, S., Nam, JH. *et al.* A review on dark channel prior based image dehazing algorithms. *J Image Video Proc.* **2016**, 4 (2016). https://doi.org/10.1186/s13640-016-0104-y
12. Ke, N., Chen, J.: "*Real-time visibility restoration from a single image*". In: Proceedings on IEEE International Conference on Image Processing, pp. 923–927 (2013)
13. Chen, S.-Z & Ren, Z.-G & Qiusheng, Lian. (2016). Single image dehazing algorithm based on improved dark channel prior and guided filter. 42. 455-465. 10.16383/j.aas.2016.c150212.
14. Feng Yu, Chunmei Qing, Xiangmin, Xu, Bolun Cai, "Image and video dehazing using view-based cluster segmentation", 2016.
15. Yongmin Park, Tae-Hwan kim, "Fast execution scheme for dark-channel-prior outdoor video dehazing", IEEE, ACCESS, vol. 6, pp. 10003-10014, March 2018.
16. L. Dong, B. Wang, M. Zhao, and W. Xu, "*Robust Infrared Maritime Target Detection Based on Visual Attention and Spatiotemporal Filtering,*" IEEE Transactions on Geoscience and Remote Sensing, vol. 55, no. 5, pp. 3037-3050, 2017.
17. D. Singh and V. Kumar, *"Dehazing of remote sensing images using improved restoration model based dark channel prior,"* The Imaging Science Journal, vol. 65, no. 5, pp. 282- 292, 2017.
18. Xin Wu, Xiao-Yan Liu, Fei Yuan, Fast dehazing method for improving the image quality in pellet size measurement, Computers in Industry, Volume 128, 2021, 103438, ISSN 0166-3615, https://doi.org/10.1016/j.compind.2021.103438.
19. Ms.S Archana M.E, A.Abiraha"An effective haze removal algorithm using color attenuation prior model", International Conference on Emerging Engineering Trends and Science Issue NO. 8, 2016, pages 387-392
20. Manjunath.V, RevanasiddappaPhatate, "A Single Image Haze Removal Algorithm Using Color Attenuation Prior", International Journal of Scientific and Research Publications, 2016, Issue 6, pages 291-297
21. Mansuo Zhao, Alan Fu, Hong Yan: "A technique of three-level thresholding based on probability partition and fuzzy 3-partition", *Fuzzy Systems, IEEE Transactions on*, pp. 469—479, 2001.
22. Alavala Chennakesava R.: *Fuzzy Logic and Neural Networks BasicConcepts and Applications*. New Ege International Limited Publications,2008.
23. Tripty Singh, "Foggy Image Enhancement and Object Identification by Extended Maxima Algorithm" 2017.
24. JiaSHI,Ke-jianYANG, " An Improved Method of Removing Fog and Haze Effect From Images", 4th International Conference on Machinery, Materials and Information Technology Applications (ICMMITA 2016) Copyright © 2017,.Published by Atlantis Press.
25. Fan Guo, JinTang Zi-Xing Cai1, Image Dehazing Based on Haziness Analysis International Journal of Automation and Computing, 11(1), February 2014, 78-86, DOI: 10.1007/s11633 014-0768-7
26. K. B. Gibson, D. T. Vo and T. Q. Nguyen, "An investigation of dehazing effects on image and video coding", *IEEE Trans. Image Process.*, vol. 21, no. 2, pp. 662-673, Feb. 2012.
27. Q. Zhu, J. Mai and L. Shao, "Single image dehazing using color attenuation prior", *Proc. Brit.Mach. Vis. Conf. (BMVC)*, pp. 1-10, 2014.
28. Shi Z, Zhu MM, Guo B, Zhao M, Zhang C (2018) Nighttime low illumination image enhancement with single image using bright/ dark channel prior. EURASIP J Image Video Process 13:1–15
29. Kim W, Lee R, Park M, Lee S (2019) Low-light image enhancement based on maximal diffusion values. IEEE Access 7:129150–129163