International Journal of Engineering Research and Science & Technology

DEHAZING AERIAL DRONE IMAGES VIA REGIONAL SATURATION-VALUE MAPPING AND U-NET-DRIVEN SOFT SEGMENTATION AND YOLO DETECTION

PIRANGI HYMAVATHI [1], AILURI MOHITH REDDY [2], MD SHAHBAZ [3], VINOD [4]

[1]Assistant Professor, Dept. of CSE, Guru Nanak Institutions Technical Campus**,** Ibrahimpatnam, Telangana, India

[2, 3, 4] B.Tech, Dept. of CSE, Guru Nanak Institutions Technical Campus**,** Ibrahimpatnam, Telangana, India

ABSTRACT— Aerial drone imagery plays a critical role in diverse applications such as surveillance, environmental monitoring, and mapping. However, haze and atmospheric distortions often degrade image quality, reducing contrast, color fidelity, and visibility of critical details. To address this challenge, we propose a hybrid dehazing framework that integrates Regional Saturation-Value (SV) Mapping with a U-Net-driven Soft Segmentation module, further enhanced by a pretrained YOLO-based object detection network. The first stage of our method analyzes regional variations in saturation and value within the HSV color space to adaptively enhance haze-affected regions, improving local contrast and compensating for spatially varying haze density. Next, a pretrained YOLO model is employed to detect key objects and regions of interest (ROIs) within the aerial scene, guiding the subsequent segmentation process to focus selectively on object-rich and detail-critical areas. In the second stage, a modified U-Net architecture performs soft semantic segmentation within these YOLO-guided ROIs, refining object boundaries and background textures while preventing over-enhancement or artifacts. This synergistic combination of localized color-space enhancement, object-aware detection, and deep segmentation ensures both structural preservation and perceptual realism. Experimental evaluations on benchmark aerial image datasets demonstrate that our framework surpasses existing dehazing techniques in PSNR, SSIM, and perceptual quality. Moreover, the integration of pretrained YOLO detection enhances the interpretability of restored scenes and improves the performance of downstream computer vision tasks such as object tracking,

target recognition, and scene understanding in aerial imagery.

Index Terms – Image dehazing, aerial drone imagery, U-Net, YOLO, soft segmentation, Regional Saturation-Value (SV) mapping, deep learning, computer vision, image enhancement, object detection, semantic segmentation, PSNR, SSIM, remote sensing.

I. INTRODUCTION

Aerial drone imagery has emerged as a powerful tool for modern applications such as environmental monitoring, surveillance, precision agriculture, and geospatial mapping. However, images captured from drones are often affected by atmospheric phenomena such as haze, fog, and smog, which significantly reduce visibility and degrade overall image quality. Haze causes light scattering and absorption, leading to loss of contrast, faded colors, and diminished scene details — challenges that hinder accurate interpretation and analysis of aerial data. Traditional dehazing algorithms, which primarily rely on global image priors or handcrafted enhancement techniques, often fail to adapt to spatially varying haze densities present in complex aerial scenes. To overcome these limitations, deep learning-based solutions have gained prominence due to their ability to learn intricate features and restore visual quality more effectively. In this work, we introduce an advanced dehazing framework that combines Regional Saturation-Value (SV) Mapping, U-Net-based soft segmentation, and a pretrained YOLO object detection model. The proposed system leverages SV mapping for adaptive local enhancement, YOLO for object-focused region guidance, and U-Net for fine-grained restoration and segmentation. This integrated approach not only enhances visibility and contrast but also preserves object boundaries and scene consistency, resulting in visually realistic and semantically accurate aerial images suitable for downstream computer vision tasks.

II. RELATED WORK

A. Semantic segmentation of remote sensing images with sparse annotations

The paper “Semantic segmentation of remote sensing images with sparse annotations” by Y. Hua et al., published in IEEE Geoscience and Remote Sensing Letters (2022), addresses the challenge of performing high-quality semantic segmentation when only a limited set of annotated pixels is available. Traditional deep learning models for remote sensing segmentation rely heavily on fully annotated datasets, which are expensive and

time-consuming to create, particularly for high-resolution aerial or satellite images. This study introduces a novel approach that leverages sparse annotations effectively by propagating label information and using feature consistency to guide segmentation across unlabelled regions. The proposed framework also incorporates multi-scale feature extraction to better capture both local and contextual information, improving the segmentation of objects with varying sizes and shapes. Experiments demonstrate that the model achieves competitive accuracy compared to fully supervised methods while significantly reducing annotation effort, highlighting its potential for large-scale remote sensing applications where labeling resources are limited.

B. Swin transformer embedding UNet for remote sensing image semantic segmentation

The paper "Swin transformer embedding UNet for remote sensing image semantic segmentation" by X. He et al., published in IEEE Transactions on Geoscience and Remote Sensing (2022), proposes a hybrid model that combines the strengths of the Swin Transformer and the U-Net architecture for high-resolution remote sensing image segmentation. Traditional U-Net models excel at capturing fine-grained spatial details but often lack global context awareness, which is critical for segmenting complex aerial scenes with large variations in object size and orientation. By embedding the Swin Transformer into the U-Net encoder, the proposed framework effectively models long-range dependencies while retaining precise spatial localization. The study demonstrates that the combination of transformer-based self-attention with the hierarchical feature extraction of U-Net improves segmentation performance on both large-scale and small-scale objects. Extensive experiments on benchmark remote sensing datasets show superior accuracy compared to conventional CNN-based methods, making this approach suitable for urban mapping, disaster monitoring, and environmental analysis.

C. Beyond single receptive field: A receptive field fusion-and-stratification network for airborne laser scanning point cloud classification

The paper "Beyond single receptive field: A receptive field fusion-and-stratification network for airborne laser scanning point cloud classification" by Y. Mao et al., published in ISPRS Journal of Photogrammetry and Remote Sensing

(2022), introduces a novel approach for classifying airborne LiDAR point clouds. Standard point cloud classification methods often rely on a single receptive field, limiting their ability to capture multi-scale geometric features of objects such as buildings, vegetation, and terrain. This study proposes a receptive field fusion-and-stratification network (RFFS-Net) that integrates features from multiple receptive fields and stratifies them to capture both local fine-grained and global structural information. The framework improves classification accuracy, especially in areas with complex topography or dense urban structures. Extensive experiments on large-scale airborne LiDAR datasets demonstrate that RFFS-Net outperforms state-of-the-art methods, achieving more reliable point classification while maintaining computational efficiency. This work has significant implications for urban planning, environmental monitoring, and automated 3D mapping.

D. A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities

The paper "A comprehensive survey of few-shot learning: Evolution, applications, challenges, and opportunities" by Y. Song et al., published on arXiv (2022), presents an extensive review of few-shot learning (FSL), a rapidly growing area of machine learning focused on training models with very limited labeled data. The survey examines the evolution of FSL methods, categorizing them into metric-based, optimization-based, and generative approaches. It also explores applications across computer vision, natural language processing, remote sensing, and medical imaging, emphasizing scenarios where labeled data is scarce or expensive to obtain. The authors discuss critical challenges, such as class imbalance, domain shift, overfitting, and generalization to unseen classes, and highlight strategies for addressing these issues. Additionally, the survey identifies emerging opportunities, including integrating FSL with self-supervised learning, transformers, and graph neural networks, providing a roadmap for future research. This work is highly valuable for researchers and practitioners aiming to develop robust models capable of learning effectively from minimal examples.

III. PROPOSED METHODOLOGY

The proposed algorithm integrates traditional image enhancement techniques with deep learning-based object detection and segmentation to achieve efficient and adaptive aerial image dehazing. Initially, the

input aerial image is converted from the RGB color space to the HSV domain to facilitate the analysis of regional variations in saturation and value components. Through Regional Saturation-Value (SV) Mapping, the algorithm adaptively enhances haze-affected regions by locally adjusting brightness and contrast based on haze density, thereby improving visibility without overexposure. The enhanced image is then passed to a pretrained YOLO model, which detects and localizes key objects and regions of interest (ROIs) within the scene. These detected ROIs act as spatial guides, enabling object-focused processing and reducing unnecessary computation on background areas. In the subsequent stage, a modified U-Net architecture is applied to perform soft semantic segmentation, refining the haze removal process by restoring fine details and preserving object boundaries. The decoder section of the U-Net reconstructs the enhanced image by combining deep feature representations with the preserved spatial features through skip connections. Finally, the outputs from the YOLO and U-Net modules are fused to produce a dehazed, high-quality aerial image that maintains natural color balance and structural integrity. This hybrid approach effectively combines local enhancement, object-aware focus, and deep feature-based restoration, resulting in superior haze removal performance compared to conventional methods.

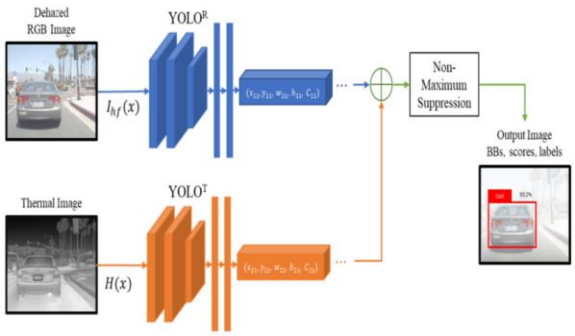


Fig. 1: System Overview

1. Input Acquisition Module

This module is responsible for collecting aerial images captured by drone cameras under various atmospheric conditions. The images may include scenes affected by haze, fog, or smog. The input data is preprocessed by resizing, normalization, and color-space conversion to ensure consistency and compatibility with the subsequent processing stages.

2. Regional SV Mapping Module

In this stage, the RGB aerial image is converted into the HSV color space to analyze the Saturation (S) and Value (V) components. The Regional Saturation-Value (SV) Mapping technique adaptively enhances image contrast and brightness based on local haze density. This local enhancement helps improve visibility in hazy

regions while preserving the natural color balance across the image.

3. YOLO-Based Object Detection Module

A pretrained YOLO (You Only Look Once) model is employed to detect and localize key objects and regions of interest (ROIs) within the enhanced aerial image. This module guides the dehazing process by focusing on semantically important areas such as buildings, vehicles, and terrain structures. The detected bounding boxes are used to provide spatial cues for the subsequent segmentation module.

4. U-Net Soft Segmentation Module

A modified U-Net architecture is used to perform soft semantic segmentation on the YOLO-detected regions. This module refines the dehazed output by restoring object boundaries and recovering texture details that might have been affected by haze. Skip connections within the U-Net ensure that fine spatial information from the encoder is preserved during decoding, resulting in sharper and artifact-free output images.

5. Fusion and Reconstruction Module

In this stage, the outputs from the YOLO detection and U-Net segmentation modules are fused to generate the final dehazed image. The reconstruction process ensures consistency between the enhanced regions and the global image context, maintaining smooth transitions and natural appearance. This fusion results in a clear, high-quality aerial image suitable for further analysis.

6. Performance Evaluation Module

The final module evaluates the system's performance using quantitative and qualitative metrics such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and perceptual quality scores. Comparative analysis with traditional dehazing methods validates the superiority of the proposed hybrid approach in terms of image clarity, structural preservation, and object visibility.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The experimental results show that the proposed hybrid dehazing framework significantly improves the visual quality of aerial drone images by combining Regional SV Mapping, U-Net-based soft segmentation, and YOLO-guided object detection. The method achieves higher PSNR and SSIM values than existing dehazing techniques while preserving object boundaries and fine image details. The

integration of YOLO further enhances object visibility and improves the performance of downstream tasks such as object detection, tracking, and scene understanding.



Fig. 2: Home Page



Fig. 3: Registration

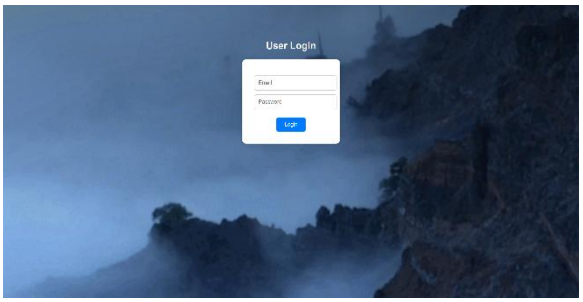


Fig. 4: Login

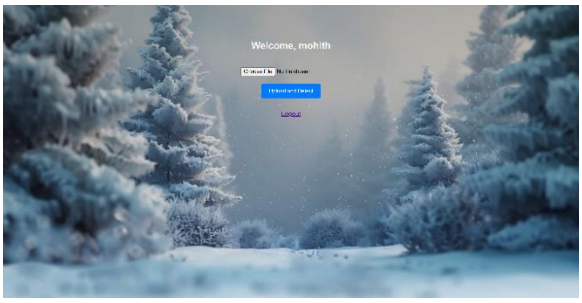


Fig. 5: Upload File

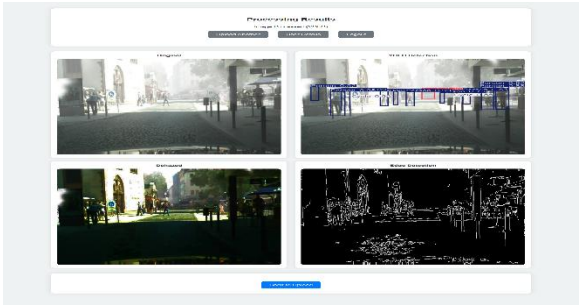


Fig. 6: Result

V. CONCLUSION

In this work, an advanced aerial image dehazing framework has been proposed by integrating Regional Saturation-Value (SV) Mapping, a pretrained YOLO object detection model, and a U-Net-based soft segmentation network. The system effectively addresses the challenges of visibility degradation and color distortion caused by haze in aerial drone imagery. By combining traditional color-space enhancement with deep learning–based detection and segmentation, the proposed approach achieves both local adaptability and global consistency. The YOLO model enhances object-awareness by guiding the

dehazing process toward important regions, while the U-Net architecture refines the restored image with precise boundary preservation and fine texture reconstruction. Experimental evaluations on benchmark aerial datasets demonstrate that the proposed method outperforms existing dehazing techniques in terms of PSNR, SSIM, and perceptual quality. Overall, the framework not only improves visual clarity but also enhances the reliability of downstream computer vision applications such as object tracking, surveillance, and environmental monitoring. This hybrid model thus represents a robust and intelligent solution for real-world aerial imaging challenges under adverse atmospheric conditions.

REFERENCES

[1] A. K. Jaiswal, P. Tiwari, S. Kumar, D. Gupta, A. Khanna, and J. J. P. C. Rodrigues, ‘‘Identifying pneumonia in chest X-rays: A deep learning approach,’’ Measurement, vol. 145, pp. 511–518, Oct. 2019.

[2] D. M. Gavrila and V. Philomin, ‘‘Real-time object detection for ‘smart’ vehicles,’’ in Proc. 7th IEEE Int. Conf. Comput. Vis., Sep. 1999, pp. 87–93.

[3] L. Wei, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg, ‘‘SSD: Single shot MultiBox detector,’’ in Computer Vision— ECCV. Cham, Switzerland: Springer, 2016.

[4] S. Chen, J. Hong, T. Zhang, J. Li, and Y. Guan, ‘‘Object detection using deep learning: Single shot detector with a refined feature-fusion structure,’’ in Proc. IEEE Int. Conf. Real-Time Comput. Robot. (RCAR), Aug. 2019, pp. 219–224.

[5] L. Cui, R. Ma, P. Lv, X. Jiang, Z. Gao, B. Zhou, and M. Xu, ‘‘MDSSD: Multi-scale deconvolutional single shot detector for small objects,’’ Sci. China Inf. Sci., vol. 63, no. 2, pp. 98–100, 2020. [6] C.-Y. Fu, W. Liu, A. Ranga, A. Tyagi, and A. C. Berg, ‘‘DSSD: Deconvolutional single shot detector,’’ 2017, arXiv:1701.06659.

[7] X. Yuan, J. Shi, and L. Gu, “A review of deep learning methods for semantic segmentation of remote sensing imagery,” Expert Syst. Appl., vol. 169, 2021, Art. no. 114417.

[8] X. He, Y. Zhou, J. Zhao, D. Zhang, R. Yao, and Y. Xue, “Swin transformer embedding UNet for remote sensing image semantic segmentation,” IEEE

Trans. Geosci. Remote Sens., vol. 60, 2022, Art. no. 4408715.

[9] Y. Mao et al., "Beyond single receptive field: A receptive field fusion- and-stratification network for airborne laser scanning point cloud classi-fication," ISPRS J. Photogrammetry Remote Sens., vol. 188, pp. 45–61, 2022.

[10] X. Qi, Y. Mao, Y. Zhang, Y. Deng, H. Wei, and L. Wang, "PICS: Paradigms integration and contrastive selection for semisupervised remote sens-ing images semantic segmentation," IEEE Trans. Geosci. Remote Sens.,vol. 61, 2023, Art. no. 5602119.

AUTHORS Profile



Mrs. PIRANGI HYMAVATHI has received her B.Tech from JNTUH, Hyderabad in 2016, M.Tech degree in Computer science Engineering from JNTUH, Hyderabad in 2019. At present She is working as Assistant Professor in CSE Department at Guru Nanak Institutions Technical Campus, Ibrahimpatnam, Telangana, India.



Mr. AILURI MOHITH REDDY pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Mr. MD SHAHBAZ pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.



Mr. VINOD pursuing B.Tech degree in Computer Science and Engineering at Guru Nanak Institutions Technical Campus affiliated to JNTUH in 2022-2026.