High-Speed Pixel Enhancement in Military Imagery:
A Novel Approach Utilizing MobileNet_v3_small and Canny Edge Integration

Members:

Aviral Srivastava 22BAI1187
Kshitij Kumar 22BAI1161
Garv Bhaskar 22BAI1371
Shilpi Anand 22BAI1477

Paper	Summary/ Important Points
Enhancing UAV Aerial Image Analysis: Integrating Advanced SAHI Techniques With Real-Time Detection Models on the VisDrone Dataset	This research integrates the RT-DETR-X model with the Slicing Aided Hyper Inference (SAHI) methodology to enhance real-time object detection in aerial images, using the VisDrone-DET dataset. Key Points: Integration of the RT-DETR-X model with the Slicing Aided Hyper Inference (SAHI) technique. RT-DETR-X selected for its flexibility and efficiency in handling UAV aerial imagery complexities. RT-DETR-X is compared with other models (e.g., RT-DETR-R18, RT-DETR-R34) based on architecture, AP values, FPS, and computational efficiency. ONNX format for compatibility and integration; Gradio interface for user-friendly demonstration. Development of a Gradio interface for interactive, user-friendly demonstration of drone detection capabilities. Dataset Information: VisDrone-DET includes 6,471 images with 343,204 labeled boxes, predominantly small-sized targets. Conducted on a four-GPU setup with a learning rate of 0.0001 and a batch size of 4. RT-DETR-X achieves 54.8% Average Precision (AP) The model runs at 74 frames per second (FPS), surpassing similar models in both speed and accuracy. The RT-DETR-X model was finely tuned to handle the complexities of aerial images, particularly in detecting varied object sizes, enhancing its real-world applicability.
Using super-resolution generative adversarial network models and transfer learning to obtain high resolution digital periapical radiographs Maira B.H. Moran, Marcelo D.B. Faria, Gilson A. Giraldi, Luciana F. Bastos, Aura Conci, Using super-resolution generative adversarial network models and	1. Approaches to Super-Resolution: Qiu et al. classify super-resolution techniques into four categories: interpolation, reconstruction, enhancement, and learning-based methods. Recently, deep learning-based techniques have showed tremendous promise in enhancing picture resolution in a variety of applications, including medical imaging and defense-related imagery. 2. Limitations of Traditional Interpolation Methods: Commercial software frequently use interpolation methods to improve image resolution, but these techniques have severe limitations. According to Kositbowornchai et al.,

transfer learning to obtain high resolution digital periapical radiographs, Computers in Biology and Medicine,	these interpolation-based zoom tools can only improve resolution up to a point; beyond that, magnification may damage accuracy by blurring boundaries and distorting features. 3. Deep Learning-Based Super-Resolution: To address the shortcomings of previous approaches, deep learning-based super-resolution has emerged as a promising option. These algorithms, like SRGAN, have the ability to produce higher-quality images by learning complicated patterns in data, avoiding the drawbacks of simple interpolation techniques.
Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi, W. (2016). Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network. ArXiv. /abs/1609.04802	1. Super-Resolution Challenge: Creating a high-resolution (HR) image from a low-resolution (LR) counterpart is difficult, particularly at large upscaling factors, where fine texture details are frequently lost. 2. Objective Function in SR Methods: Traditional super-resolution methods aim to minimise the mean squared error (MSE) between the reconstructed HR image and the ground truth, which yields high peak signal-to-noise ratios (PSNR) but frequently lacks perceptually significant high-frequency features. 3. Limitations of MSE and PSNR: While the MSE and PSNR metrics are widely used, they are insufficient for capturing perceptual quality and high texture detail in super-resolved images because they rely on pixel-wise differences that ignore perceptual similarities. 4. Introduction to SRGAN: SRGAN is the first framework capable of creating photorealistic images for 4× upscaling factors. It uses a generative adversarial network (GAN) to bring super-resolved images closer to the natural image manifold. 5. Perceptual Loss Function: The SRGAN uses a novel perceptual loss function that combines an adversarial loss that promotes realism in images with a content loss motivated by perceptual similarity rather than pixel similarity. 6. GAN-Based Approach: The adversarial loss employs a discriminator network to distinguish between super-resolved and original images, directing the network to generate more natural and realistic images.
HCL Whitepaper on HCL_IMVE_WP-ImageProcessing_vs_DL	The HCL whitepaper comprehensively addresses the transition from traditional image processing techniques to the innovative realm of deep learning approaches. It offers insights into how automation via deep learning surpasses conventional methods by improving accuracy in image quality enhancement while managing complexity in datasets. This paper elaborates on the capacities of various deep learning architectures, illustrating their capability to

	handle intricate tasks that often result in significantly enhanced resolution outputs. The whitepaper also emphasizes the efficiency of these modern methods in processing vast amounts of data, which is a critical requirement in fields such as military surveillance and autonomous systems. By critically evaluating existing literature, the paper not only benchmarks the performance of deep learning against traditional techniques but also outlines potential future directions, making it essential reading for researchers and practitioners pursuing advancements in image processing technologies.
MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications	The paper detailing MobileNets showcases the development of highly efficient convolutional neural networks specifically crafted for mobile and embedded vision applications. By employing a streamlined architecture using depthwise separable convolutions, MobileNets significantly reduce computation loads while maintaining competitive performances. The introduction of two global hyper-parameters, tailored to optimize the trade-offs between latency and accuracy, allows model builders to customize their networks based on application-specific constraints. Extensive experimental results highlight MobileNets' impressive scalability and adaptability across various tasks, such as object detection and face attribute classification, demonstrating robustness comparable to more traditional, heavier models like ResNet. This research not only expands the applicability of deep learning in resource-constrained environments but also provides valuable insights into model design that can be leveraged in the pursuit of lightweight solutions for real-time processes in areas that require immediate feedback, such as autonomous systems and real-time surveillance.
Research on Improved Canny Edge Detection Algorithm	In the research presented in the MATEC Web of Conferences, an enhanced Canny edge detection algorithm is proposed, addressing common shortcomings of the original Canny method, particularly concerning noise robustness and the prevalence of false edges. The study innovatively integrates statistical filtering for denoising, coupled with a genetic algorithm that determines adaptive thresholds for edge detection. This dual approach significantly elevates the accuracy of edge localization and the algorithm’s overall robustness against noise interference. The MATLAB simulations conducted reveal the effectiveness of this improved algorithm across various applications requiring precise edge detection, particularly in challenging environments afflicted by noise. By overcoming traditional limitations, this research marks a notable advancement in edge detection techniques, positioning the proposed model as a vital tool in the performance enhancement of image processing tasks, thus expanding its potential applications in sectors such as security and medical imaging.
Aerial Image Object Detection Based on Improved YOLOv5	Objective: Enhances YOLOv5 for better aerial image object detection.

Authors: Jian Zhao, Jie Li, and Xiaoyong Du Published: 2022	Model Improvements: The architecture is modified to improve small object detection and integrates a spatial pyramid pooling (SPP) layer for multi-scale feature detection. Additionally, anchor box settings are optimized to better suit aerial data. Datasets & Evaluation: Shows improved accuracy on DOTA and VisDrone datasets, with higher precision/recall rates. The model also maintains real-time processing capabilities, making it suitable for UAV surveillance. Challenges Addressed: Effectively handles complex backgrounds and variability in object appearance, which are common in aerial imagery. Conclusion: YOLOv5 is made more robust for aerial image detection, with suggestions for further optimizations, such as adapting the model for varying altitudes and different environmental conditions.
Super-Resolution of Multispectral Satellite Images Using Convolutional Neural Networks Authors: M. U. Müller, N. Ekhtiari, R. M. Almeida, C. Rieke Published: 2023	Objective: Improves the resolution of multispectral satellite images using CNNs. Approach: Combines high-resolution panchromatic bands with low-resolution multispectral data, comparing four CNN architectures: SRCNN, RedNet30, AESR, and SRResNet. Results: RedNet30 performs best, balancing spatial detail and spectral integrity. The results show that CNNs with autoencoder structures and residual blocks are particularly effective. Evaluation Metrics: Advanced metrics like FSIM and ISSM confirm the superior visual quality, despite PSNR values suggesting that simpler interpolation methods might perform better. Conclusion: Highlights the effectiveness of CNNs in enhancing image resolution and suggests further refinement of pansharpening techniques to improve the overall quality and utility of satellite images for remote sensing applications.
A Residual Dense U-Net Neural Network for Image Denoising Authors: Javier Gurrola-Ramos, Oscar Dalmau, Teresa E. Alarcón	Objective: Introduces RDUNet for image denoising, focusing on noise reduction while preserving detail. Architecture: Based on U-Net with residual dense blocks for better feature reuse. The model also incorporates global residual learning, where it predicts the residual noise rather than the clean image, improving performance, especially with Gaussian noise. Methodology: The network was trained on images with additive white Gaussian noise at various intensities, making it robust against different noise levels. Strided convolutions are used for downsampling instead of max-pooling to better preserve local contrast information. Results: RDUNet outperforms traditional methods with higher PSNR and SSIM on standard datasets. Significance: RDUNet is suitable for high-quality image

	restoration, balancing complexity and performance. The model's versatility in handling different noise levels without prior knowledge of the noise characteristics adds to its practical application in fields like medical imaging and remote sensing.
LSwinSR: UAV Imagery Super-Resolution based on Linear Swin Transformer Authors: Rui Li, Xiaowei Zhao Published: 2023	Objective: Proposes LSwinSR for enhancing UAV imagery using the Linear Swin Transformer. Key Features: The model utilizes Swin Transformer with kernel attention to reduce computational complexity while maintaining high performance. It focuses on overcoming challenges related to limited resolution due to UAV flight constraints. Results: LSwinSR shows competitive accuracy with faster inference speeds compared to existing methods like SwinIR, making it ideal for real-time UAV applications. Applications: It is particularly effective in land monitoring and environmental surveillance, where high-resolution imagery is crucial. Significance: Demonstrates advancements in UAV image processing, offering practical solutions for enhancing UAV-captured images, which are vital for accurate analysis in environmental monitoring and urban planning.
A Residual Dense U-Net Neural Network for Image Denoising (Second Paper) Authors: Javier Gurrola-Ramos, Oscar Dalmau, Teresa E. Alarcón Published: 2023	Objective: This paper also presents the RDUNet model, focusing on image denoising through densely connected convolutional layers combined with residual learning. Methodology: Built on the U-Net architecture, the model improves denoising by predicting residual noise instead of the clean image. It was tested on various datasets corrupted by different types of noise, proving its robustness across scenarios. Results: Achieves superior results in denoising images, outperforming conventional methods like BM3D and DnCNN. Significance: RDUNet balances performance and complexity, making it a robust solution for various image denoising tasks, including those in medical imaging and photography, where maintaining image detail is critical. The paper also suggests future work to reduce the computational burden for real-time applications.
A Review on Image Enhancement Techniques April 2019Southeast Europe Journal of Soft Computing 8(1) 8(1) DOI:10.21533/scjournal.v8i1.175 LicenseCC BY 4.0	1.An Overview of Image Enhancement Techniques Throughout History: Several techniques have been established over the years, leading to a substantial evolution in image improvement. The early techniques were mainly limited to simple procedures such as contrast modification using histogram equalization. With the advancement of research, more sophisticated methods were created to address particular image types and applications, including satellite imagery, surveillance, and medical imaging. These

	methods included adaptive histogram equalization, unsharp masking, and edge enhancement. 2. Present-day Methods and Trends: Sophisticated algorithms that make use of artificial intelligence, deep learning, and machine learning are at the forefront of recent developments in picture enhancement. Several other picture formats, such as grayscale, color, infrared, and even films, can be handled using these techniques. Convolutional neural networks (CNNs) and generative adversarial networks (GANs) are two methods that are being used more and more to enhance image quality, especially in difficult situations like dimly lit or noisy surroundings. 3. Limitations and Challenges: In spite of the advancements, there are still a number of difficulties in the field of picture improvement. The trade-off between improving image quality and maintaining crucial information without adding artifacts is one of the main problems. Furthermore, different approaches could be effective in certain situations but ineffective in others, thus it's important to select the right approach depending on the application sector. The study also underlines the flaws of the state-of-the-art techniques as they are today and the necessity of more research to solve them.
	1.The history of single image super-resolution (SISR) In computer vision, Single Image Super-Resolution (SISR) has been an important field of study with the goal of recovering high-resolution images from their low-resolution counterparts. Conventional approaches include bilinear and bicubic interpolation algorithms, which are straightforward but frequently fall short in recovering fine details. By learning intricate mappings from low-resolution to high-resolution images, more sophisticated techniques have been developed to increase the quality of the upscaled images, including patch-based methods, deep learning approaches, and sparse coding. 2. Present-day Methods and Trends: Sophisticated algorithms that make use of artificial intelligence, deep learning, and machine learning are at the forefront of recent developments in picture enhancement. Several other picture

E. Agustsson and R. Timofte, "NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 2017, pp. 1122-1131, doi: 10.1109/CVPRW.2017.150. keywords: {Image resolution;Atmospheric measurements;Particle measurements;Image quality;Image restoration;Degradation;Agriculture},	formats, such as grayscale, color, infrared, and even films, can be handled using these techniques. Convolutional neural networks (CNNs) and generative adversarial networks (GANs) are two methods that are being used more and more to enhance image quality, especially in difficult situations like dimly lit or noisy surroundings. 3. The DIV2K Dataset's Introduction and Effects: A sizable and varied dataset created especially for SISR, the DIV2K dataset was first made available as part of the NTIRE 2017 competition. This dataset, which offers a more thorough evaluation of model performance over a range of image kinds and resolutions, has emerged as the new benchmark for assessing SISR approaches. By contrasting cutting-edge solutions from the challenge with conventional techniques, the research highlights how leveraging the DIV2K dataset can advance SISR. As a whole, the challenge and the dataset have improved performance on both long-standing and recently introduced evaluation metrics, setting new standards for SISR research in the future.
NTIRE 2017 Challenge on Single Image Super-Resolution: Dataset and Study Eirikur Agustsson CVL, ETH Zurich, Switzerland aeirikur@vision.ee.ethz.ch Radu Timofte CVL, ETH Zurich & Merantix GmbH radu.timofte@vision.ee.ethz.ch	1.Historical Context of Single Image Super-Resolution (SISR): Research on Single Image Super-Resolution (SISR) has been conducted over several decades, with notable advancements made from more sophisticated techniques to traditional methods such as bicubic interpolation. The primary goal of the field's fundamental studies is to overcome SISR's ill-posedness, which allows a low-resolution image to correspond to several high-resolution options. Numerous algorithmic techniques have been used to tackle this problem, from patch-based techniques to the first deep learning models such as SRCNN and its offspring. The creation of universal benchmarks, including Set5, Set14, B100, and Urban100, has made it easier to compare techniques under consistent settings. 2.Benchmarking and Recent Developments: The area of SISR has advanced quickly in recent years, especially with the introduction of deep learning-based techniques that have increased SISR's performance to unprecedented levels. On widely used datasets, techniques like VDSR, DRCN, and SRResNet have produced state-of-the-art results. In order to ensure that progress can be monitored objectively, increasingly complex models and approaches have made ongoing benchmarking on existing datasets necessary. The

	need for more hard and diversified standards has been brought to light by the reliance on a small number of datasets. 3. Overview of the NTIRE 2017 Challenge and the DIV2K Dataset: A more thorough benchmarking was required, which is why the DIV2K dataset was created. With 1000 high-resolution photos, this dataset offers a more varied and difficult benchmark for SISR techniques. The DIV2K dataset was used for the NTIRE 2017 Challenge, which encouraged participants to push the limits of current techniques and marked a noteworthy milestone in the field. Beyond the conventional bicubic downscaling, the challenge offered new degradation operators to test the resilience of SISR algorithms. The challenge's outcomes have established new standards and brought to light the advantages and disadvantages of the state-of-the-art SISR methods.
Sensitivity Increase Through a Neural Network Method for LOR Recovery of ICS Triple Coincidences in High-Resolution PixelatedDetectors PET Scanners	The proposed methodology introduces a real-time capable technique designed to enhance sensitivity in Positron Emission Tomography (PET) by accurately including Inter-Crystal Scatter (ICS) triple coincidences in the imaging process. This is achieved through careful computation of the Line-of-Response (LOR) using an Artificial Neural Network (ANN) that is trained with preprocessed raw data. The focus is on improving sensitivity while minimizing any degradation in image quality, particularly by analyzing scenarios where one photoelectric 511-keV event coincides with two additional events whose energy sums to 511 keV
Algorithm for Single Image Enhancement Based on Semantic Segmentation Assistance	The proposed methodology integrates image semantic segmentation with image enhancement to improve the accuracy and quality of images that have been degraded by noise. By utilizing an alternating boosting network that combines multiple segmentation and enhancement modules, this method is designed to leverage the synergistic relationship between image enhancement and semantic segmentation. The experimental results indicate that this approach not only enhances image quality but also significantly improves segmentation accuracy, thereby achieving outcomes close to those obtained with clean images.

	The methodology proposed involves the development of a segmentation-enhancement alternating boosting network (SEABN) that consists of multiple segmentation and enhancement modules1. The SEABN operates on the principle that enhancing the image prior to segmentation can significantly improve segmentation accuracy, particularly in unclear images1. Conversely, the segmentation results can be utilized to enhance the performance of the image enhancement processes, creating a feedback loop that benefits both tasks
Reduce The Effect Of Vignetting On Mimo Optical Wireless System By Reducing Ber Of Channel	The proposed methodology involves addressing the issue of vignetting in pixelated Multiple-Input Multiple-Output (MIMO) optical wireless communication (OWC) systems using spatial Orthogonal Frequency Division Multiplexing (OFDM). By implementing a technique aimed at reducing the Bit Error Rate (BER) through simulation in MATLAB, the methodology significantly improves the accuracy of signal transmission in the presence of vignetting. This process has critical implications for enhancing the quality of pixelated images received in these communication systems. The proposed methodology specifically examines the adverse effects of vignetting on pixelated systems and aims to mitigate these effects.
Computational investigation of CdZnTe pixelated detector performance for x-ray based medical imaging applications	The proposed methodology involves the computational modeling of pixelated Cadmium Zinc Telluride (CdZnTe) detectors, utilizing an integrated framework combining Monte Carlo and Finite Element numerical methods. This approach aims to enhance the performance of medical imaging applications by accurately simulating various detector configurations, including different pixel sizes and thicknesses. The outcomes can significantly contribute to the optimization of pixelated image processing in medical imaging systems, ultimately leading to improved image quality and diagnostic capabilities. This modeling incorporates charge transport effects, allowing for the investigation of parameters like pixel size and detector
Exploring the Effect of Image Enhancement Techniques on COVID-19 Detection using Chest X-ray Images Authors: Muhammad E. H. Chowdhury	This study emphasizes the critical role of image enhancement techniques in improving the detection accuracy of COVID-19 using chest X-ray (CXR) images. A comprehensive dataset, COVQU-20, containing 18,479 images, was compiled to evaluate detection performance effectively. It includes 3,616 COVID-19 images, 6,012 non-COVID lung opacity images, and 8,851 normal CXR images. Five distinct image enhancement methods were utilized: histogram equalization, contrast limited adaptive histogram

	equalization, image complement, gamma correction, and Balance Contrast Enhancement Technique. The gamma correction method proved to be the most effective, achieving an accuracy of 96.29% in detecting COVID-19 from standard and segmented lung CXR images. This research provides significant insights into leveraging image enhancement methods to improve the efficiency of COVID-19 detection systems in medical imaging.
Image Enhancement Method Based on Deep Learning Authors: Zhang	This paper introduces a deep learning-based image enhancement network capable of transforming original images into high-quality color images. It discusses various methodologies such as Convolutional Neural Networks (CNN), Residual Neural Networks, and Generative Adversarial Networks (GANs) for addressing common imaging issues like noise and low illumination. The proposed framework emphasizes multitasking through adaptive feature learning to significantly enhance image quality, improving visual experience and performance in computer vision applications. By utilizing advanced loss functions, including periodic consistency loss, the model effectively addresses issues of image quality degradation due to environmental variables. The research results indicate a marked improvement in the overall visual quality and detail retention of images, affirming the efficacy of deep learning in image enhancement tasks.