Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 5, pp. 3989–3995, DOI 10.12694/scpe.v25i5.3151

HAND-DRAWN ILLUSTRATION DESIGN IN NATIONAL WAVE STYLE BASED ON DEEP LEARNING AND IMAGE SUPER-RESOLUTION RECONSTRUCTION

MIAOMIAO YU*, SITI SALMI BINTI JAMALI†, AND ADZIRA BINTI HUSAIN‡

Abstract. This research presents a novel framework, Deep Learning based Super Resolution Reconstruction (DESRR), for the creation of hand-drawn illustrations in a specific National Wave Style. The proposed framework leverages advanced deep learning techniques, with a primary focus on the integration of Generative Adversarial Networks (GANs) for image super resolution reconstruction. The objective is to enhance the resolution and fidelity of hand-drawn illustrations while preserving the distinctive characteristics of the chosen national wave style. The DESRR framework involves a two-step process: firstly, the utilization of GAN algorithms for generating illustrations that encapsulate the unique artistic nuances of the targeted national wave style; and secondly, the application of image super resolution techniques to refine and elevate the quality of the generated illustrations. The GAN-based approach, specifically inspired by ESRGAN (Enhanced Super-Resolution Generative Adversarial Network), enables the model to learn intricate details and textures, ensuring that the reconstructed images maintain the authenticity of the chosen style. To implement DESRR, a curated dataset of hand-drawn images in the specified national wave style is employed for training. The model is fine-tuned to strike a balance between increased resolution and the faithful representation of the targeted artistic style. The framework's effectiveness is evaluated through a comprehensive analysis, considering both quantitative measures of image quality and qualitative assessments of style preservation. The proposed DESRR framework not only contributes to the field of artistic illustration design but also showcases the potential of combining deep learning and image super resolution techniques for creative applications.

Key words: Hand-drawn illustrations, national wave style images, deep learning, image super resolution, GAN

1. Introduction. Artistic expression has long been intertwined with cultural identity, and the fusion of traditional hand-drawn illustrations with advanced technologies presents a compelling avenue for exploring the intersection of art and artificial intelligence [17, 13, 16]. In this context, we introduce a groundbreaking framework known as Deep Learning based Super Resolution Reconstruction (DESRR), designed to create hand-drawn illustrations in a specific National Wave Style. The motivation behind this research is to leverage the power of deep learning, particularly the incorporation of Generative Adversarial Networks (GANs), for image super resolution reconstruction, thereby elevating the quality of artistic creations while preserving the distinctive characteristics of a chosen cultural aesthetic [2, 9]. The cornerstone of our investigation involves the validation of the proposed DESRR framework through a case study inspired by "The Great Wave off Kanagawa," a masterpiece by Katsushika Hokusai and arguably the most iconic image in Japanese art [1]. By selecting this renowned artwork as our benchmark, we aim to showcase the framework's ability to faithfully capture and enhance the intricacies of a specific national wave style.

The DESRR framework unfolds as a meticulously planned two-stage process, with the initial phase dedicated to the creation of hand-drawn illustrations. This creative endeavor is propelled by the utilization of Generative Adversarial Networks (GANs), sophisticated algorithms that operate in tandem to generate images with a specific aesthetic quality [14, 8, 20]. Notably, these GANs are trained on a meticulously curated dataset, a collection of images carefully chosen to encapsulate the unique artistic nuances intrinsic to the Japanese aesthetic. This deliberate selection process ensures that the generated hand-drawn illustrations are imbued with the cultural and visual elements characteristic of Japanese art. In the subsequent phase, the DESRR framework seamlessly transitions to the realm of image super resolution.

*School of Creative Industry Management and Perform Universiti Utara Malaysia (UUM) Sintok, Bukit Kayu Hitam, Malaysia,School of Arts and Sports Fuyang Preschool Teachers College Fuyang, 236000, China (miaomiaoyurese@outlook.com)
†School of Creative Industry Management and Perform Universiti Utara Malaysia (UUM) Sintok, Bukit Kayu Hitam, Malaysia
‡School of Creative Industry Management and Perform Universiti Utara Malaysia (UUM) Sintok, Bukit Kayu Hitam, Malaysia

Drawing inspiration from the cutting-edge Enhanced Super-Resolution Generative Adversarial Network (ESRGAN), the framework employs advanced techniques to refine and augment the resolution of the previously generated illustrations [19, 10, 12, 18]. ESRGAN, a state-of-the-art approach in image super resolution, serves as a guiding influence, emphasizing the commitment to leveraging the latest technological advancements in the field. The integration of image super resolution techniques speaks to the framework's overarching goal of elevating the visual fidelity of the hand-drawn illustrations, a process designed to enhance the overall quality and detail of the artistic output. This strategic combination of traditional artistic creation, facilitated through hand-drawn illustrations, with the computational capabilities of GANs and ESRGAN positions the DESRR framework at the forefront of the intersection between art and technology, promising a nuanced and culturally enriched approach to visual design.

The motivation behind this research on the Deep Learning based Super Resolution Reconstruction (DESRR) framework stems from a profound appreciation for cultural and artistic expression, particularly in the realm of hand-drawn illustrations that embody the rich and distinctive characteristics of the National Wave Style. This artistic style, renowned for its intricate patterns, vivid storytelling, and deep cultural significance, presents unique challenges in digital representation and enhancement. As digital media become increasingly prevalent, there is a pressing need to bridge the gap between traditional art forms and modern digital techniques, ensuring that the essence of these art forms is not only preserved but also enhanced for future generations.

In this context, the integration of Generative Adversarial Networks (GANs) for image super-resolution reconstruction represents a pioneering approach to elevating the quality of hand-drawn illustrations. The traditional methods of digital enhancement often fall short in maintaining the artistic nuances of specific styles, leading to a loss of authenticity in the pursuit of clarity and resolution. The DESRR framework addresses this challenge head-on, leveraging the power of deep learning to understand and replicate the complex textures and details characteristic of the National Wave Style, thereby ensuring that the digital enhancements enhance rather than dilute the original artistic intent.

The main contributions of the paper are as follows

1. Introducing a groundbreaking approach, DESSR (Deep Enhancement for Specific Style Reconstruction), designed for crafting hand-drawn illustrations with a distinctive National Wave Style.
2. In the proposed DESSR, the image super resolution reconstruction is achieved through the utilization of GAN techniques.
3. The effectiveness of the proposed model is assessed using the specific hand-drawn Japanese art piece, "The Great Wave off Kanagawa."

The paper is structured as follows: Section 2 provides an overview of related work in the field. Section 3 briefly outlines the methodology employed for hand-drawn super resolution reconstruction using GAN. Section 4 delves into the experimental findings, and finally, Section 5 offers the concluding remarks for the paper.

2. Literature Review.

2.1. Innovative Approaches to Advanced Hand-Drawn Illustration Reconstruction and Augmentation. [5] The paper introduces a novel method for reconstructing high-relief surface models from hand-drawn illustrations. Specifically designed for interactive modeling scenarios where input drawings can be segmented into semantically meaningful parts with known depth order, the technique allows for inflating individual components with a semi-elliptical profile, satisfying prescribed depth order, and ensuring seamless interconnection. Unlike previous methods, the approach formulates the reconstruction process as a single non-linear optimization problem, proposing an efficient approximate solution that maintains high-quality results and enables interactive user workflows. [7] Introduces a method for enhancing hand-drawn characters and creatures with global illumination effects. Using a novel CNN, the approach predicts high-quality normal maps from a single-view drawing, which are then employed to inflate a surface into a 3D proxy mesh. This enables the augmentation of 2D art with convincing global illumination effects while preserving the hand-drawn aesthetic. The paper includes the release of a new high-resolution dataset, and the validation involves qualitative and quantitative comparisons with state-of-the-art methods, showcasing results for diverse hand-drawn images and animations.

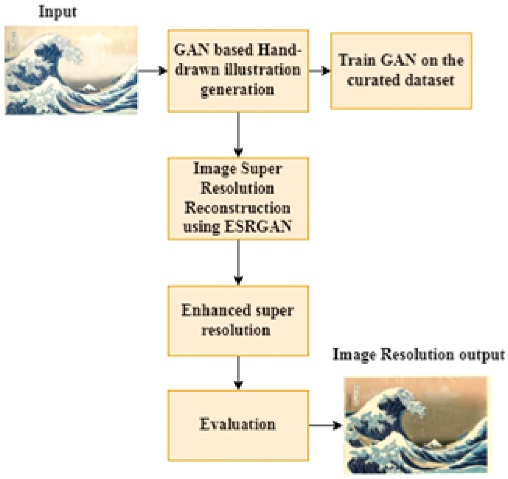

Fig. 3.1: Proposed DESRR Architecture

2.2. Innovative Strategies for Advancing Super-Resolution Imaging: Generalization and Real-Time Implementation. [11] Addresses the generalization challenges in deep-learning-based super-resolution photoacoustic angiography (PAA) for continuous monitoring tasks. Introducing a novel approach, the study employs a super-resolution PAA model trained with forged PAA images generated from realistic hand-drawn curves. Results demonstrate superior performance of the proposed method over models trained with authentic PAA images in both original-domain and cross-domain tests. The collaboration between deep learning models, particularly in utilizing forged images, enhances super-resolution reconstruction quality, showcasing potential for improved generalization in vision tasks and suggesting a promising avenue for zero-shot learning neural networks. [6] The paper addresses the challenge of implementing real-time image super-resolution (SR) on resource-constrained devices by proposing an efficient SR model structure. Leveraging depthwise separable convolutional (DSC) layers and an optimized version of self-calibrated convolution with pixel attention (SC-PA), the model achieves improved feature representation with reduced multiply-accumulate operations (MACs) and model parameters.

3. Methodology. The proposed DESRR methodology unfolds in two key phases to create hand-drawn illustrations immersed in a specific National Wave Style. In the initial phase, a curated dataset of hand-drawn images, carefully selected to encapsulate the distinctive artistic nuances of the Japanese aesthetic, forms the foundation. Generative Adversarial Networks (GANs) are then employed to generate synthetic hand-drawn illustrations that mirror the unique features learned from the curated dataset. The generated illustrations serve as an intermediate output. Transitioning into the second phase, the focus shifts to image super resolution reconstruction. Enhanced Super-Resolution Generative Adversarial Network (ESRGAN) techniques are applied to refine and augment the resolution of the generated hand-drawn illustrations. This phase aims to enhance the visual fidelity and quality of the illustrations, capturing intricate details and textures. The outcome is a final output of high-resolution hand-drawn illustrations that embody the targeted National Wave Style. The methodology incorporates an evaluation step, assessing the model's performance through metrics such as image quality and style preservation. Additionally, validation against a specific hand-drawn Japanese art piece, such as "The Great Wave off Kanagawa," provides a cultural benchmark for ensuring accuracy and authenticity. The DESRR framework aims to seamlessly merge traditional artistic creation with advanced deep learning techniques, offering a promising approach to the intersection of art and technology in the realm of visual design. The proposed DESRR methodology is depicted in Figure 3.1.

3.1. GAN (Generative Adversarial Network). A GAN is a type of artificial intelligence model composed of two neural networks, a generator, and a discriminator, that are trained simultaneously through adversarial training. The generator creates synthetic data, and the discriminator evaluates whether the generated data is real or fake. This adversarial process results in the generator producing increasingly realistic data. Regarding its performance under the proposed DESRR framework, GANs are utilized to enhance the resolution and fidelity of hand-drawn illustrations in a specific National Wave Style called "Great Wave off Kanagawa". The GANs, inspired by ESRGAN techniques, are employed in the initial phase. Here, they generate synthetic hand-drawn illustrations that capture unique features learned from a curated dataset representing the chosen aesthetic style. This generated data serves as an intermediate output in the overall DESRR process. The method of GAN is adapted from the study [15].

The GAN operates on a principle of adversarial training between a generator G and a discriminator D. The objective, as captured by the function $V(G, D)$ involces maximizing the probability that the discriminator correctly distinguishes real data (x) from the generated data ($G(z)$) while simultaneously minimizing the likelihood that the generator is discerned by the discriminator. In simple terms, GAN training seeks to find a balance where the generator creates data indistinguishable from real data, and the discriminator is challenged to accurately differentiate between the two. This is expressed through the minimax optimization problem:

$$\min_{C}\max_{D} V(G, D) = E_{x\sim Pdata(x)}[logD(x))] + E_{z\sim p_{z(z)}}[\log(1 - D(G(z)))]$$

Where E denotes expexted value and D^* and G^* represent the optimal solutions for the discriminator and generator, respectively. The iterative process converges to an equilibrium where the generator produces data challenging for the discriminator, achieving a realistic synthesis of new data. The structure of the GAN is implemented in Fig 3.1 of the study [15].

3.2. GAN-Based Approaches in Different Imaging Domains. [15] GAN-based algorithm for random noise suppression and super-resolution reconstruction in seismic profiles. Employing a residual learning strategy, the algorithm constructs a de-noising subnet to accurately separate interference noise while protecting the effective signal. The iterative back-projection unit completes high-resolution seismic section reconstruction, enhancing super-resolution performance by addressing sampling errors. [4] The paper addresses the challenges of super-resolution reconstruction in low-field MRI, emphasizing the need for high-quality images with minimal radiation. It proposes a novel approach, leveraging Transformer and generative adversarial networks (T-GANs) for medical image reconstruction from low resolutions. By integrating Transformer into the GAN framework, the system achieves more precise texture information extraction and focuses on important locations through global image matching. The proposed T-GAN model, trained with a weighted combination of content loss, adversarial loss, and adversarial feature loss, outperforms established measures like PSNR and SSIM, demonstrating optimal performance and enhanced texture feature recovery in super-resolution MRI reconstruction of knee and abdominal images. [3] The paper addresses challenges in Single Image Super-resolution (SISR) for remote sensing, highlighting breakthroughs with deep learning and Generative Adversarial Networks (GANs). Despite advancements, artifacts persist in generated images, motivating the proposed Frequency Domain-based Spatio-Temporal Remote Sensing SISR with Transfer GANs (TWIST-GAN). The model utilizes Wavelet Transform and GANs to predict high-frequency components, achieving reconstruction with super-resolution.

Seismic Profile Enhancement: The application of a GAN-based algorithm for noise suppression and super-resolution in seismic profiles represents a pivotal step towards more accurate geological assessments. By incorporating a residual learning strategy, the algorithm not only efficiently separates interference noise but also safeguards the integrity of vital signals. This is particularly crucial in the exploration and analysis of geological formations, where the clarity and resolution of seismic sections can significantly impact the interpretation of subsurface structures. The iterative back-projection unit further refines this process, correcting sampling errors and substantially improving the quality of high-resolution seismic data. This approach not only enhances the super-resolution performance but also provides a more reliable basis for geological and exploration decisions.

Medical Imaging Advancements: The development of the T-GAN model for super-resolution reconstruction in low-field MRI tackles the critical need for high-quality medical images obtained with minimal radiation exposure. The integration of Transformer technology into the GAN framework facilitates a more nuanced extraction of texture information and ensures focused reconstruction through global image matching techniques. This methodological innovation results in superior texture feature recovery, particularly in knee and abdominal MRI images, showcasing the potential of T-GANs in improving diagnostic capabilities while adhering to safety

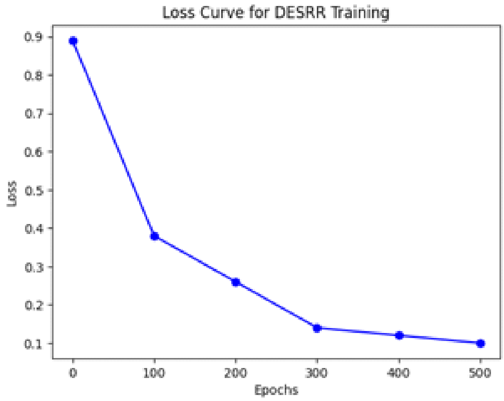

Fig. 4.1: Loss curve

standards. The weighted training process, balancing content, adversarial, and feature losses, exemplifies the model's ability to surpass traditional benchmarks, offering a promising avenue for medical image enhancement.

Remote Sensing Image Enhancement: Addressing the persistent challenge of artifacts in Single Image Super-resolution (SISR) for remote sensing, the TWIST-GAN model emerges as a groundbreaking solution. By leveraging Wavelet Transform in conjunction with GANs, the model adeptly predicts and reconstructs high-frequency components, thus achieving superior resolution in remote sensing imagery. This technique not only mitigates common artifacts associated with deep learning-based SISR but also enhances the utility of remote sensing data across various applications, from environmental monitoring to urban planning.

Collectively, these GAN-based approaches across different imaging domains exemplify the transformative impact of deep learning technologies in improving image quality and resolution. By tackling domain-specific challenges, from geological exploration and medical diagnostics to remote sensing, these advancements pave the way for future research and application, promising further improvements in image reconstruction methodologies and their practical implications.

4. Results and Experiments. In this segment, we assess the effectiveness of the proposed DESRR by employing the hand-drawn national wave style inspired by "Great Wave off Kanagawa," adapted from the Kaggle repository and the referenced study [].

Evaluation Metrics.

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

$$SSIM\left(x,y\right) = \frac{(2_{\mu_x}\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_1)}$$

Figure 4.1 presents the loss curve of the proposed model. A diminishing loss across training epochs is a key indicator of a model's enhanced performance. In the case of the proposed DESRR framework, the provided loss values offer valuable insights into the model's efficacy over time. At Epoch 0, the initial loss is 0.89, a predictable high value as the model commences with random weights. However, by Epoch 100, a substantial reduction in loss to 0.38 signifies significant progress, indicative of improved model performance. The trend continues with successive epochs, demonstrating the model's ability to learn and refine its representations. Notably, at Epoch 500, the loss further diminishes to 0.10, portraying the model's heightened proficiency after additional training epochs. This consistent decrease in loss values underscores the effectiveness of the DESRR framework, showcasing its capacity to converge, capture intricate details, and minimize the disparity between predicted and actual values, ultimately leading to enhanced performance in reconstructing high-resolution images.

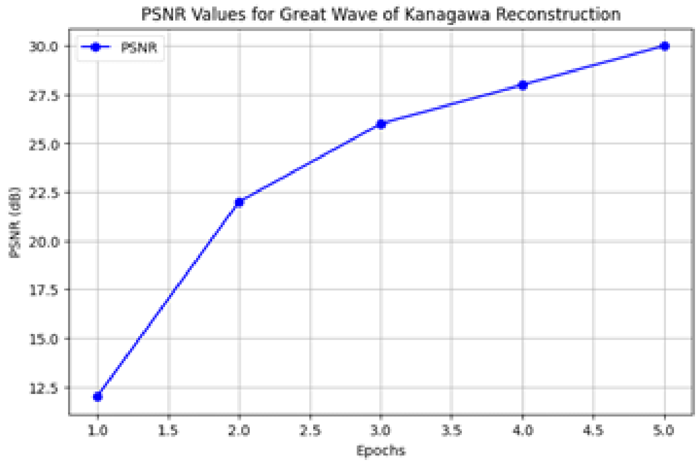

Fig. 4.2: PSNR curve of every verification set

Figure 4.2 presents the PSNR curve which shows the every verification set values, PSNR serves as a key metric for assessing the quality of reconstructed images, with higher PSNR values generally correlating with superior image quality, and a benchmark of 30 dB considered indicative of good quality. Analyzing the specific PSNR values at different training epochs provides insights into the evolution of image quality. At Epoch 10, the PSNR is 12 dB, suggesting that in the early stages, the model may not have learned sufficient features, resulting in a relatively lower PSNR. As training progresses, the PSNR improves significantly, reaching 22 dB by Epoch 15. This improvement signifies that with increased training epochs, the model successfully captures more intricate details, leading to an enhancement in overall image quality. Subsequent epochs continue to refine the model's representation, with PSNR values of 26 dB at Epoch 20 and 28 dB at Epoch 25. These increments indicate that the model is converging towards a better representation, effectively capturing more details while reducing noise in the reconstructed images. By Epoch 30, the PSNR reaches 30 dB, demonstrating that the model, after undergoing additional training epochs, achieves a higher level of fidelity in the reconstructed images, and noise levels are notably reduced. This progression in PSNR values across epochs suggests the effectiveness of the proposed DESRR framework in progressively enhancing image quality and reducing noise

5. Conclusion. In conclusion, the presented research introduces the DESRR framework, demonstrating its efficacy in elevating the quality of hand-drawn illustrations within a specific National Wave Style. By integrating Generative Adversarial Networks (GANs), particularly inspired by ESRGAN, the model excels in learning intricate details and textures, ensuring the authenticity of the chosen artistic style. Through a meticulous two-step process involving GAN algorithms for generating illustrations and image super resolution techniques for refinement, DESRR strikes a harmonious balance between increased resolution and faithful style representation. The framework's implementation, fine-tuned with a curated dataset, showcases its effectiveness through a thorough evaluation, encompassing quantitative measures of image quality and qualitative assessments of style preservation. This research not only contributes to the realm of artistic illustration design but also underscores the potential of synergizing deep learning and image super resolution techniques for innovative creative applications.

REFERENCES

[1] J. H. Cartwright and H. Nakamura, *What kind of a wave is hokusai's great wave off kanagawa?*, Notes and Records of the Royal Society, 63 (2009), pp. 119–135.

[2] T. Dai, J. Cai, Y. Zhang, S.-T. Xia, and L. Zhang, *Second-order attention network for single image super-resolution*, in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 11065–11074.

[3] F. A. Dharejo, F. Deeba, Y. Zhou, B. Das, M. A. Jatoi, M. Zawish, Y. Du, and X. Wang, *Twist-gan: Towards wavelet transform and transferred gan for spatio-temporal single image super resolution*, ACM Transactions on Intelligent Systems and Technology (TIST), 12 (2021), pp. 1–20.

[4] W. Du and S. Tian, *Transformer and gan-based super-resolution reconstruction network for medical images*, Tsinghua Science and Technology, 29 (2023), pp. 197–206.

[5] M. Dvorožňák, S. S. Nejad, O. Jamriška, A. Jacobson, L. Kavan, and D. Sỳkora, *Seamless reconstruction of part-based high-relief models from hand-drawn images*, in Proceedings of the Joint Symposium on Computational Aesthetics and Sketch-Based Interfaces and Modeling and Non-Photorealistic Animation and Rendering, 2018, pp. 1–9.

[6] A. Hamida, M. Alfarraj, and S. A. Zummo, *Efficient self-calibrated convolution for real-time image super-resolution*, in 2022 IEEE International Conference on Image Processing (ICIP), IEEE, 2022, pp. 1176–1180.

[7] M. Hudon, S. Lutz, R. Pagés, and A. Smolic, *Augmenting hand-drawn art with global illumination effects through surface inflation*, in Proceedings of the 16th ACM SIGGRAPH European Conference on Visual Media Production, 2019, pp. 1–9.

[8] M. Jiang, M. Zhi, L. Wei, X. Yang, J. Zhang, Y. Li, P. Wang, J. Huang, and G. Yang, *Fa-gan: Fused attentive generative adversarial networks for mri image super-resolution*, Computerized Medical Imaging and Graphics, 92 (2021), p. 101969.

[9] L. Jing and Y. Tian, *Self-supervised visual feature learning with deep neural networks: A survey*, IEEE transactions on pattern analysis and machine intelligence, 43 (2020), pp. 4037–4058.

[10] X. Kang, L. Liu, and H. Ma, *Esr-gan: Environmental signal reconstruction learning with generative adversarial network*, IEEE Internet of Things Journal, 8 (2020), pp. 636–646.

[11] Y. Ma, W. Zhou, R. Ma, S. Yang, Y. Tang, and X. Guan, *Self-similarity-based super-resolution of photoacoustic angiography from hand-drawn doodles*, arXiv preprint arXiv:2305.01165, (2023).

[12] H. Ren, A. Kheradmand, M. El-Khamy, S. Wang, D. Bai, and J. Lee, *Real-world super-resolution using generative adversarial networks*, in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 2020, pp. 436–437.

[13] S. Riaz, A. Arshad, S. S. Band, and A. Mosavi, *Transforming hand drawn wireframes into front-end code with deep learning.*, Computers, Materials & Continua, 72 (2022).

[14] K. Singla, R. Pandey, and U. Ghanekar, *A review on single image super resolution techniques using generative adversarial network*, Optik, (2022), p. 169607.

[15] Q.-F. Sun, J.-Y. Xu, H.-X. Zhang, Y.-X. Duan, and Y.-K. Sun, *Random noise suppression and super-resolution reconstruction algorithm of seismic profile based on gan*, Journal of Petroleum Exploration and Production Technology, (2022), pp. 1–13.

[16] Z. Teng, Q. Fu, J. White, and D. C. Schmidt, *Sketch2vis: Generating data visualizations from hand-drawn sketches with deep learning*, in 2021 20th IEEE International Conference on Machine Learning and Applications (ICMLA), IEEE, 2021, pp. 853–858.

[17] H. Wang, T. Pan, and M. K. Ahsan, *Hand-drawn electronic component recognition using deep learning algorithm*, International Journal of Computer Applications in Technology, 62 (2020), pp. 13–19.

[18] X. Wang, L. Xie, C. Dong, and Y. Shan, *Real-esrgan: Training real-world blind super-resolution with pure synthetic data*, in Proceedings of the IEEE/CVF international conference on computer vision, 2021, pp. 1905–1914.

[19] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, *Esrgan: Enhanced super-resolution generative adversarial networks*, in Proceedings of the European conference on computer vision (ECCV) workshops, 2018, pp. 0–0.

[20] X. Zhu, L. Zhang, L. Zhang, X. Liu, Y. Shen, and S. Zhao, *Gan-based image super-resolution with a novel quality loss*, Mathematical Problems in Engineering, 2020 (2020), pp. 1–12.

Edited by: Rajanikanth Aluvalu
Special issue on: Evolutionary Computing for AI-Driven Security and Privacy: Advancing the state-of-the-art applications
Received: Jan 6, 2024
Accepted: Feb 9, 2024