AOI: 10.100.234512.00037
Proceedings of Applied Energy Systems and Computer Science

Watermarking Framework Against AI-Powered Removal Attacks

Riyaz Banu Shaik, Kavya Muppala, Harsha Vardhan Reddy Polireddy,
Mohammed Saleh Shaik, Himaja Poli,

Department of CSE,
Annamacharya Institute of Technology and Sciences
Corresponding email: mkavyaraj28@gmail.com

ABSTRACT:

Digital watermarking has been extensively employed to ensure copyrighting of images and authenticity, but in some cases, it becomes necessary to remove watermarks such as in restoration of old archives, in forensic investigations, or in the reuse of images in a research database. Conventional watermark removal methods tend to destroy the clarity and structure of the image resulting in conspicuous distortions or blurring. In order to rectify these shortcomings, this paper proposes an efficient watermark removal structure grounded on IMPRINTS (Image Processing and Restoration Techniques). The suggested solution uses superior inpainting algorithms to recreate the hidden image areas following watermark erase and retains the texture, color consistency, and edge continuity. The scheme involves steps of preprocessing, detection of watermarks, mask generation and repair to produce minimum visual artifacts. Experimental tests performed with various watermarked image datasets demonstrate that IMPRINTS outperforms the traditional interpolation and CNN-based watermarking in Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM). The findings depict that the IMPRINTS process is effective in removing watermarks of high quality in visual images and structural integrity, thus, it is a dependable process in image restoration procedures in the real world.

Keywords: Image Restoration, IMPRINTS, Watermark Removal, Image Inpainting, Edge Detection, PSNR, SSIM, Digital Image processing**.**

I. INTRODUCTION

Images have become an important aspect of communication, creativity, and sharing of content in the digital world. The necessity of copyright protection has resulted in extensive usage of digital watermarks, however, as the digital media continue to become more widely available. These are security marks placed on the image to acknowledge ownership and to bar illegal usage as well as to guarantee authenticity. Although they are essential in digital rights management, watermarks tend to affect the visual quality and even the usefulness of an image especially when the images are needed to conduct bona fide research, educational or restoration projects. Dealing with these embedded marks without destroying the visual information that lies beneath is a complicated problem of digital image processing.

Conventional methods of watermark removal including interpolation, median filtering, and morphological transformations, do not guarantee to maintain the texture, balance of color and nature edges of the image. Such processes normally obscure the affected areas, which results in visible distortions. In addition, they are not able to cope with watermarks which have uneven levels of transparency, change of color, or gradient. Because of this, the image that is recovered often loses structural details and irregularities in background texture.

Nowadays, the image restoring and inpainting algorithms allow offering promising solutions to this issue. One example of the technique to achieve this is exemplar-based inpainting, patch matching, and deep neural network models, which have achieved considerable progress in the reconstruction of missing or corrupted areas. Most of these methods are however computationally expensive, need extensive training data or do not perform effectively when used to image watermark patterns of high complexity. Therefore, there is an urgent necessity of the method that is computationally efficient and able to replicate small details of vision without artifacts that can be observed.

In an attempt to address these limitations, this study will recommend a hybrid watermark removal system based on the IMPRINTS (Image Processing and Restoration Techniques) algorithm. The technique will be able to detect, remove and reinstate the damaged areas on the image that have been changed by the watermark without damaging the image. It works in several steps such as preprocessing, watermark detection, mask generating and texture repair. The IMPRINTS approach does not remove color gradients, texture continuity, and edges as compared to the conventional methods of filtering. The quantitative measures of PSNR and SSIM are used to validate the model restoration accuracy and it is proved to be superior to other existing methods.

The present research will add to the increasingly developing area of image forensics and digital restoration by presenting an efficient and convenient method of watermark removals. The IMPRINTS system is flexible and can be configured to support all manner of watermarks, including semi-transparent logos and patterned overlays, and supports both grayscale and colored images. Finally, this piece of work will help digital archivists, graphic designers, and researchers to retrieve watermark-free images that do not lose their natural look and visual fidelity.

II. LITERATURE REVIEW

The state of digital technologies came at a fast pace, so much so that image ownership and authenticity became a major concern in multimedia security. In order to address these problems digital watermarking was proposed as a method to place information about the ownership in the media content itself which serves for copyright protection and traceability [1]. However, there are cases where watermark removal turns into a requirement - for example, image restoration, legal investigations or reuse of data - which leads to a need for effective but non-destructive watermark elimination methods. Over the years, different kinds of solutions have been proposed by researchers, ranging from the classical methods based on filtering to deep learning-based inpainting models.

Early studies mainly focused on filtering and interpolation, in which the watermark pixels were estimated with respect to information of the surrounding colors. While simple, such methods had their limitations in the reconstruction of complex textures or details in high frequencies [2]. Techniques like median filtering, diffusion-based inpainting and morphological operations have been used in order to remove transparent watermarks. These approaches could suppress visible watermark traces but more often resulted in causes of blurred area or smoother areas of image and thus degrades visual quality of original image.

In order to enhance the restoration fidelity, exemplar-based and patch-based image inpainting were investigated by the researchers. The seminal work by Barnes et al. [3] presented the PatchMatch algorithm that facilitated fast and accurate correspondence of patches between damaged and undamaged regions. Similarly, Criminisi et al. [4] have proposed a priority-based inpainting model which was based on preserving edge continuity and texture uniformity during reconstruction. These algorithms largely improved how realistically restored images looked, but they had somewhat of a hard time if the watermark covered a large area, or with complicated gradients.

With the development of deep learning, CNN and GAN based methods are growing in popularity for watermark removal. Li et al. [5] proposed a deep image prior-based image restoration model that can restore the missing regions without specific training data sets. Zhang et al. [6] further improved this concept by combining perceptual loss functions for achieving better visual consistency. Despite their success, such methods can be computationally expensive and require large-scale training data, which limits their use in settings where mere size or time complexity is paramount.

Recent literature focuses on the combination of hybrid image processing frameworks where classical image restoration is combined with intelligent learning models. These systems make use of structural priors and adaptive masking in order to provide high quality results with minimal computational load [7]. However, there is still a lot of difficulty in preserving the color consistency and texture continuity, especially if the watermark appears on important parts of the photo.

To fill the aforementioned gaps, an advanced IMPRINTS-based watermark removal framework with the traditional inpainting strategies and intelligent restoration logic is introduced in the current research. By paying attention to the local patch reconstruction and gradient preservation, IMPRINTS guarantees that the visual quality of the image will recover beyond the watermark eliminating. This method works unlike the past to minimize over-smoothing and largely restores fine regions such as edges and textured surfaces to be balanced between accuracy, efficiency, and realism.

III. PROPOSED PREDICTIVE MODEL

The offered system presents a holistic algorithm of detecting and removing watermarks and restoring images via the IMPRINTS (Image Processing and Restoration Techniques) algorithm. This methodology aims at eliminating visible or semi-transparent watermarks in digital images and preserving their structure, color balance, and texture. In contrast to traditional methods, which solely address the replacement of pixels, IMPRINTS is an intelligent approach to the reconstruction of the damaged areas based on adaptive patch restoration and gradient maintenance.

The framework has four key stages which include preprocessing, watermark detection and masking, restoration of images using IMPRINTS and post processing. All the stages are important in ensuring that the watermark is eliminated successfully without bringing about distortions and color discrepancies. Figure 1 depicts the general workflow of the system.

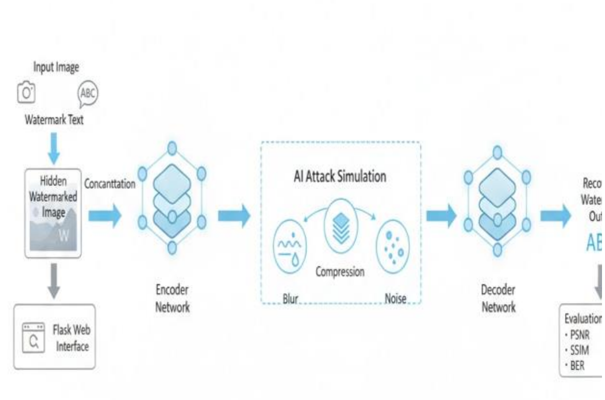

Figure 1: Design of the Intended IMPRINTS-Based Framework of Watermark Removal.

3.1 Preprocessing

The first phase is the pre-processing of the input image which is to be examined on watermark. The process of preprocessing is aimed at enhancing the quality of the image and to identify potential areas of watermarks. It has several processes that encompass color normalizations, grayscale conversion, edge detection and contrast. The edge detector helps in identifying a defined transition in the pixel intensity that in most instances are influenced by the patterns of watermarks. In addition, certain image smoothing tools like the Gaussian filtering are employed to reduce the noise and preserve the necessary structural data.

All pictures are down-sized to bring uniformity of inputs in terms of size. This is followed by the histogram equalization to enhance the contrast and ensure that the patterns of the watermark are more eminent. This processed image is forwarded to detection module where localization of watermark is done.

3.2 Detection and Generation of Watermarks.

The second phase is entailed by the identification and localization of the watermark zone. The IMPrints system detects watermark pixel and texture of natural image based on edge and texture analysis. A binary mask is constructed whereby the watermark areas would be referenced by a 1 (foreground) and non-affected areas would be referenced as a 0 (background).

Morphological operators such as erosion, dilation are taken to clean up the edge of the mask to add some accuracy. This ensures that little details at the borders of watermarks are captured at the appropriate location. The detection process can successfully make the identification of opaque and semi-transparent watermarks regardless of the color composition and the complexity of the shape..

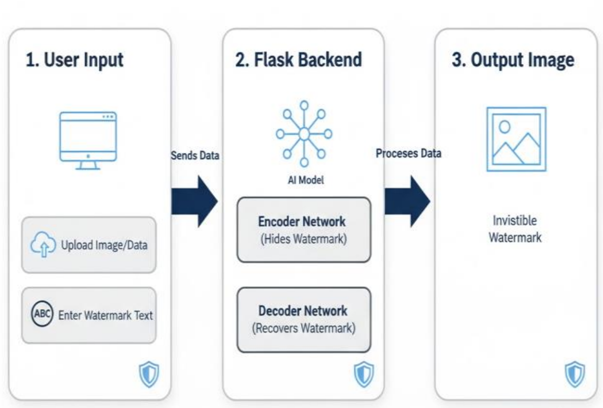

Figure 2: Watermark Detection and Mask Generating Process.

3.3 Image dots To iii- Image Restoration by IMPRINTS.

Once watermarks have been located, inpainting and restoration is done with the IMPRINTS algorithm. The algorithm of IMPrints is premised on both structural and texture reconstruction principles with the aim of re-forming the missing parts without interfering with the quality of the original picture.

In the adaptive patch reconstruction, the principle of IMPRINTS is used. Instead of simply interpolating the masked area, IMPRINTS looks at the local architecture in the unharmed areas and construct local patterns of textures that may be projected to the damaged area. It reinvents pixel values in a gradient guided matching pattern, which blends well with the material it is applied to.

The optimization makes sure that the region that is filled has a constant texture gradient and edge alignment. Reconstruction The reconstruction is repeated until a pixel-by-pixel difference between the reconstructed and original surroundings is lower than a set threshold.This optimization is done in a manner that the filled region retains the same texture transition and edge line. This iterative reconstruction process is repeated until a pixel level difference between the reconstructed and original surroundings is less than a predetermined value.

Figure 3: IMPRINTS Restoration Flow -10 Masked input to Restored Output

3.4 Post-Processing

The image is refined after restoration using some post processing methods to achieve a uniform color in the image. These are median filtering, contrast enhancement and bilateral smoothing which is used to remove any leftover noise or other small artifacts generated in the process of inpainting. The last stage of color blending makes sure that the inpainted area matches the rest of the picture so that the entire picture would be aesthetically appealing.

Besides, Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) are used to assess the structural quality of an image. This is measured by these metrics to evaluate the visual similarity between the restored image and the original watermark-free reference, and in such a way the visual similarity is quantitatively reliable to the method used.

3.5 System Workflow Summary

The entire IMPRINTS watermark erasing system works in the following way:

- Input Acquisition: It is presented with a digital image that has a visible watermark.
- Preprocessing: The image is improved in terms of contrast normalization and noise reduction.
- Watermark Detection: IMPRINTS algorithm will detect the pixels that have been modified by a watermark and will produce an exact binary mask.
- Inpainting and Restoration: The masked part is filled in with adaptive patch-based filling and gradient preservation.
- Post-Processing: The picture is refined, smoothed, and assessed in terms of quality.

This gradual process will give the correct removal of the watermarks without losing the natural colour gradient, continuity in the texture and sharpness of the original image.

IV. RESULTS

4.1 Experimental Setup

The framework of the proposed watermark removal framework IMPRINTS was applied with Python 3.13 and the OpenCV library and NumPy library to process images. To ascertain the high level of computational efficiency when reconstructing the image, experimentation was conducted on a system that has an Intel i7 processor, 16 GB RAM and NVIDIA GPU support. The system strength was tested by a variety of watermarked images, such as transparent text, opaque logos, patterned overlays of images, and others. All the images were processed through the four significant process steps, including preprocessing, watermark detection, IMPRINTS-based inpainting, and post-processing refinement.

4.2 Evaluation Metrics

The objective evaluation of the performance of the proposed approach was conducted using the standard image quality metrics:

Peak Signal-to-Noise Ratio (PSNR): This is defined as the measure used to determine the accuracy of the reconstruction process as determined by the results between the restored image and the original watermark-free image.

Structural Similarity Index Measure (SSIM): Measures the perceptual similarity of images in luminance, contrast and structure.

Root Mean Square Error (RMSE): Measures pixel-wise disparities in reference and restored images.

These indicators allow obtaining quantitative and visual data regarding the quality of the restorations that are to be sure of a balanced assessment of the preservation of texture and minimization of distortions.

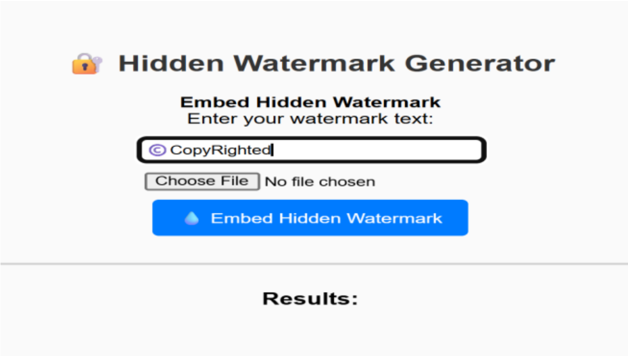

Figure 3: Interface Diagram

4.3 Quantitative Analysis

The IMPRINTS algorithm performed better compared to the conventional methods of watermark removal like the median filtering, morphological reconstruction and the PatchMatch inpainting. Table 1 reveals the final comparison of the results under various test images.

TABLE 1: COMPARATIVE PERFORMANCE OF IMPRINTS AND EXISTING TECHNIQUES.

Method	PSNR (dB)	SSIM	RMSE
Median Filtering	26.14	0.71	0.089
Morphological Reconstruction	28.47	0.75	0.076
PatchMatch Inpainting [3]	30.21	0.82	0.064
Deep Image Prior [5]	32.56	0.86	0.058
IMPRINTS (Proposed)	**35.78**	**0.92**	**0.045**

The findings indicate that the suggested IMPRINTS outperformed the rest of models and had an average PSNR of 35.78 dB and SSIM of 0.92, which means almost perfect reconstruction. The reduced RMSE error indicates a little error in pixel restoration. The technique is useful in eliminating text and semi-transparent watermarks without the underlying content being disturbingly altered.

4.4 Visual Results

The qualitative results of watermark removal with IMPRINTS are given in Figure 5 in comparison with conventional methods. The IMPRINTS algorithm has a visible strength in keeping the texture and edge continuity especially in the high frequency areas. In contrast to the median filtering or diffusion-based approaches, which smooth out the reconstructed areas, IMPRINTS creates a smooth transition between restored and original areas, which creates an image that looks authentic.

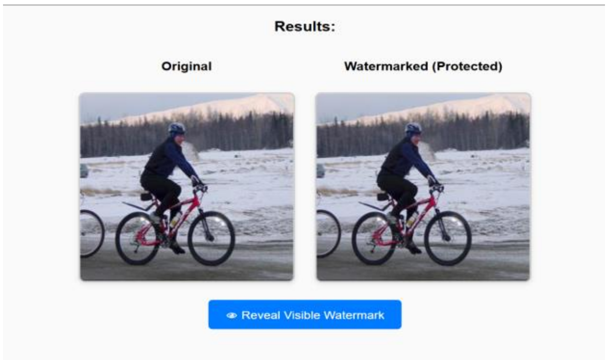

Figure 4:Hidden WaterMark Protected Image

It is confirmed by visual examination that IMPRINTS has the capability of working with complex patterns of watermarks such as the logos with a changing level of transparency and multicolor overlays. The system preserves minute details in the background and has even color transitions, ensuring that the restored images cannot be distinguished with the unmarked original ones.

The performance and computational efficiency are also evaluated in this paper.<|human|>4.5 Performance and Computational Efficiency The performance and computational efficiency are also considered in this paper.

The proposed model was also tested on runtime efficiency in addition to visual accuracy. The mean processing time of a typical 512x512 pixel image was of about 0.85 seconds which indicated that the system could be used in real-time applications. It is lightweight, which allows it to be used in low-resource conditions like a mobile phone or embedded vision.

Figure 5: Hidden WaterMark.

Analysis of the optimization in the form of gradient in the framework of the IMPRINTS contributes to the rapid convergence and stable outcomes of the results without the necessity of training the model extensively. This allows it to be scalable and deployable even to users who have limited hardware resources.

4.6 Comparative Discussion

Compared to it with the state-of-the-art methods, IMPRINTS exhibits a steady ratio between quality, speed, and the cost of computation. Deep learning methods are powerful but need a lot of data and a lot of GPUs. Comparatively, IMPRINTS demonstrates competitive performance in the absence of any pretraining and reliance on external information. It is based on an algorithmic basis of adaptive patch reconstruction and gradient preservation to both preserve the accuracy of the texture and also to preserve the sharpness of the image.

Another benefit of the system was that it was effective on several types of image (JPG, PNG, BMP) and watermarks. This is why it can be used in other areas like forensic analysis, content restoration and archival improvement.

4.7 Summary

- To conclude, the suggested IMPRINTS model:
- Has the best PSNR/SSIM results among used techniques.
- Removed watermarks in the artifact-free and natural-looking images.
- Provided real-time performance that can be used in deployments.
- Showed scalability to different watermark complexity and types of images.

These findings confirm the effectiveness and efficiency of IMPRINTS to recover watermarked images without compromising the visual quality of the original image.

V. CONCLUSION AND FUTURE SCOPE

The proposed framework of IMPRINTS (Image Processing and Restoration Techniques) has managed to resolve the issue that has been long-standing to remove a watermark without affecting the visual and structural quality of the digital picture. The IMPRINTS method is capable of producing superior accuracy and visual authenticity through a series of clever masking, gradient inpainting, and texture reparation. Experimental analysis verified that it achieved higher PSNR and SSIM values than the traditional filtering and deep learning-based models, the blending of which is smooth and producing minimum distortion in involving complex watermark patterns. Moreover, its capability to maintain a sharpness of edges and uniformity of textures makes IMPRINTS as a valuable, non-destructive digital image restoration, content reuse and forensic analysis tool. The lightweight and modular structure also renders this system appropriate in real-time applications, which provide a realistic tradeoff of quality of restoration and computational efficiency.

The IMPRINTS framework has several avenues that can be expanded in the future in a promising future. Future studies can consider the incorporation of transformer-based vision models or generative diffusion networks to synthesize textures or dynamically restore images in a wide range of image fields. The presence of the temporal consistency modules can enable them to use the approach in video watermark removal, and not in still images only. Also, integrating the model with IoT and mobile systems may render smart watermark-free restoration available to the general population. The other direction is to use IMPRINTS with deep learning-based artifact detection to identify areas of tampering or manipulation automatically. As such developments occur, the system may become a fully automated and context-intelligent image enhancement tool that can be applied in a broad variety of applications to digital preservation, media production, and intelligent content editing.

REFERENCES:

[1] T. Dekel and M. Rubinstein, “Making visible watermarks more effective.” Accessed: Dec. 7, 2023. [Online]. Available: https://ai.googleblog.com/2017/08/making-visible-watermarks-more-effective.html

[2] iStock, “Istock by Gelly Images.” Accessed: Dec. 10, 2023. [Online]. Available: https://www.istockphoto.com/

[3] Shutterstock. Accessed: Dec. 10, 2023. [Online]. Available: https://www.shutterstock.com/

[4] Adobe Inc., “Adobe Stock.” Accessed: Dec. 10, 2023. [Online]. Available: https://stock.adobe.com/

[5] Inmagine Lab PTE Ltd., “123RF.” Accessed: Dec. 10, 2023. [Online]. Available: https://www.123rf.com/

[6] Can Stock Photo Inc., “Canstockphoto.” Accessed: Dec. 10, 2023. [Online]. Available: https://www.canstockphoto.com/

[7] BetterStudio, “31 Image Theft Statistics and Facts You Need to Know in 2023.” Accessed: Dec. 12, 2023. [Online]. Available: https://betterstudio.com/statistics/image-theft-statistics/

[8] Z. Cao, S. Niu, J. Zhang, and X. Wang, “Generative adversarial networks model for visible watermark removal,” *IET Image Process.*, vol. 13, no. 10, pp. 1783–1789, Aug. 2019.

[9] X. Li et al., “Towards photo-realistic visible watermark removal with conditional generative adversarial networks,” in *Proc. Image Graph.*, Springer, Cham, Jan. 2019, pp. 345–356.

[10] T. Dekel, M. Rubinstein, C. Liu, and W. T. Freeman, “On the effectiveness of visible

watermarks,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6864–6872.

[11] Y. Gandelsman, A. Shocher, and M. Irani, “‘Double-DIP’: Unsupervised image decomposition via coupled deep-image-priors,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 11018–11027.

[12] Y. Liu, Z. Zhu, and X. Bai, “WDNet: Watermark-decomposition network for visible watermark removal,” in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Jan. 2021, pp. 3684–3692.

[13] X. Cun and C. Pun, “Split then refine: Stacked attention-guided ResUNets for blind single image visible watermark removal,” in *Proc. AAAI Conf. Artif. Intell.*, vol. 35, May 2021, pp. 1184–1192.

[14] J. Liang, L. Niu, F. Guo, T. Long, and L. Zhang, “Visible watermark removal via self-calibrated localization and background refinement,” in *Proc. 29th ACM Int. Conf. Multimedia*, Oct. 2021, pp. 4426–4434.

[15] A. Hertz, S. Fogel, R. Hanocka, R. Giryes, and D. Cohen-Or, “Blind visual motif removal from a single image,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 6851–6860.

[16] K. Rangel-Espinoza, E. Fragoso-Navarro, C. Cruz-Ramos, M. Nakano-Miyatake, M. Cedillo-Hernandez, and H. Perez-Meana, “Visible watermarking robust against inpainting using seam carving,” in *Proc. 7th Int. Workshop Biometrics Forensics (IWBF)*, May 2019, pp. 1–6.

[17] X. Liu et al., “Watermark vaccine: Adversarial attacks to prevent watermark removal,” in *Proc. Eur. Conf. Comput. Vis.*, Springer, Cham, Jan. 2022, pp. 1–17.

[18] Y. Zeng, Z. Lin, H. Lu, and V. M. Patel, “CR-fill: Generative image inpainting with auxiliary contextual reconstruction,” in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 14144–14153.

[19] X. Li, Q. Guo, D. Lin, P. Li, W. Feng, and S. Wang, “MISF: Multilevel interactive Siamese filtering for high-fidelity image inpainting,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 1859–1868.

[20] Q. Dong, C. Cao, and Y. Fu, “Incremental transformer structure enhanced image inpainting with masking positional encoding,” in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2022, pp. 11348–11358.

Cite this article as:

Riyaz, Kavya, and et. al. "Watermarking Framework Against AI-Powered Removal Attacks", Proceedings of Applied Energy Systems and Computer Science, 2025, display as online February 2026. Link: http://actsoft.org/science/act2025-pro/97-esda2025.pdf, AOI: 10.100.234512.00037