International Research Journal of Multidisciplinary Scope (IRJMS), 2025; 6(3):1450-1465

Original Article | ISSN (O): 2582-631X DOI: 10.47857/irjms.2025.v06i03.04494

AI Base Inpainting for Hex-Art Restoration

Gargi Phadke*, Siuli Das, Yogita Mistry, Shivangi Agrawal, Aditi Chhabria, Ritika Thusoo, Reshma Gulwani

Ramrao Adik Institute of Technology, DY Patil Deemed to be University Nerul Navi Mumbai, India. *Corresponding Author's Email: gargi.phadke@rait.ac.in

Abstract

In light of the growing popularity of digital photography and the decreasing dependence on hard-copy photographic formats, this paper provides a thorough analysis of the comprehension and application of image restoration software. As a result, many photographic memories now sit in home storage for long periods of time, eventually giving way to the deterioration of image quality. However, the purpose of this research is to shed light on the revolutionary possibilities of advanced algorithms and technology in the restoration of these. However, the suggested approach recognizes the significant difficulties brought about by the complexity of the algorithm and the resource-intensive nature of the restoration procedure, highlighting the necessity of continued research and development. In conclusion, picture restoration software proves to be a vital instrument for preserving our digital legacy, revitalizing visual stories, even though there is a constant need for creative ways to overcome the challenges it poses. Utilizing technologies such as Variational Autoencoders (VAEs), open CV, Generative Adversarial Network (GAN), deep learning, and their applications, this proposed system has the potential to fundamentally alter photography and the field of archaeological surveys, assisting in historical discoveries and revealing world truths and times that have been concealed for centuries in the form of images. Proposed system are validated on standard facial dataset.

Keywords: Convolutional Neural Network (CNN), Face Detection, Image Processing, Mean Square Error (MSE), Open CV, Peak to Signal to Noise Ratio PSNR.

Introduction

This section of the paper aims to discuss and present cutting-edge technology that accurately and easily restores historical images that have multiple issues such as discoloration, scratches, and low-resolution pictures due to such a long time and an improper way of storing images for a long time. These images, scarred by the passage of time, hold various stories that are waiting to be revealed, which can be done using this software. Images of old times hold not only memories, but also information that reveals truths about the world and culture at that particular time, which helps us to learn about the people of that time along with their beliefs. Those images are those with low resolution because of the hardware limitations at that time. They are also mostly scratched and damaged due to being buried in a place. By restoring these photos, the proposed system can open a door to the past and learn things that were once hidden, and most likely even destroyed, in those pictures (1). Images, these windows to our shared human experience, transcend time, capturing ephemeral moments and preserving the emotions that once enveloped them. A rich tapestry of visual history, dating back to the earliest days of photography in 1826, they offer profound insights into the epochs that have shaped our world. Whether ancient daguerreotypes or contemporary snapshots, these images serve as invaluable reservoirs of information that call us to unravel the mysteries of bygone eras. Figure 1 shows the low-resolution image that seems to be old and not readable and requires a heavy restoration to convert it to a readable format (2). The image fuzziness and removal of the same, using mean filter, median filter, wavelet and even Neural network are shown in Figure 2.

This system was proposed with many motivations and many problems to solve. They include points like historical images that suffer from an enormous amount of degradation need to be corrected to make them understandable. These historical images contain the snapshot of that particular time and are hence required to be unearthed which cannot be done unless the images are clean and scratch-free.

(Received 17th March 2025; Accepted 07th July 2025; Published 30th July 2025)

This can also be used to restore images from old times, such as ancestors and relics, or old documents that have pictures in them to learn and know about them. Historical cultures i.e. the culture followed by people in historical time can also be available from these snapshots.

Figure 1: Distorted First Image having Unreadable Damage and Low Resolution (1)

The pictures in historical times have completely low resolution because there is no presence of a camera with high resolution. Not just old pictures, sometimes today's pictures also have low-resolution pictures, and blurred images, and all of these images can be corrected by using this proposed software, face detection technique is used. This system can also be used for study purposes because this system uses complex systems to correct these images (3).

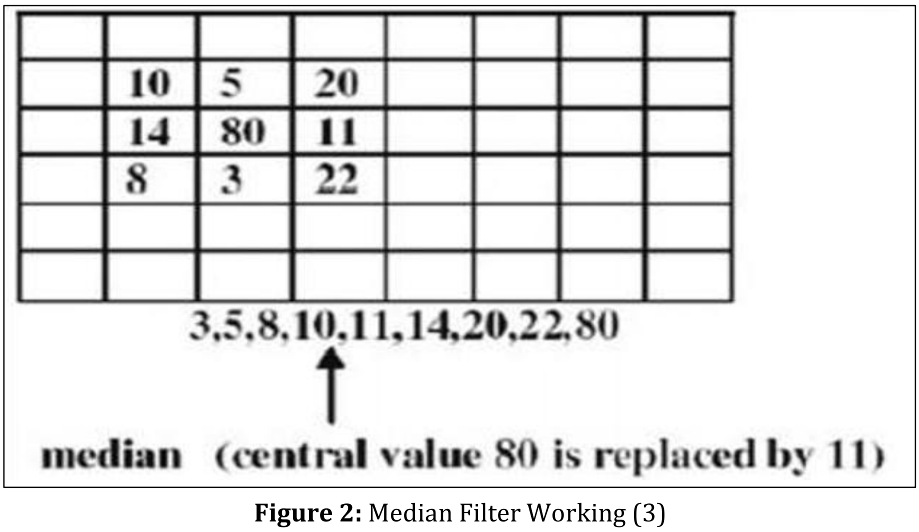


Figure 2: Median Filter Working (3)

In essence, this proposed system is a time-traveling artist, breathing life back into these fragments of the past. It takes blurry and blemished images, marred by the wear and tear of years, and imbues them with renewed vitality. With meticulous precision, it repairs scratches, bridges gaps, and enhances overall quality, aiming to offer a vivid and tangible connection to history itself. Through our efforts, these images, once elusive and incomplete, now stand as resplendent windows into bygone eras, ensuring that the legacy of 1826 and beyond is not merely remembered but vividly experienced, serving as a testament to the enduring power of human ingenuity (4, 5).

To comprehend the power of image restoration software, it is crucial to grasp its inner workings. At its core, this technology relies on sophisticated algorithms and advanced image-processing techniques. These algorithms meticulously analyze and manipulate digital images, identifying imperfections and artefacts caused by ageing, data corruption, or storage issues. By employing techniques like image inpainting and noise reduction, the software works its magic, bringing clarity and vibrancy to deteriorating images. It is a testament to human ingenuity in harnessing the potential of evolving technology to safeguard our digital heritage.

The objectives of the proposed system are:

Recreate Images

This is done for images that have scratches in them and have lost their resolution with time, this system aims to recreate these images by removing scratches and colourizing the whole image.

Historical Documents

Historical documents are normally easily readable but when images are included in these documents, it becomes a bit difficult to understand those images manually if they are completely destroyed.

Resolution

The proposed system can increase picture quality or increase the resolution of them every time the system is run. This feature is not only required for historical pictures but also new generation images which may suffer from low resolution.

This section provides all the salient features provided by the existing available methods and the limitations. The survey of the technologies and methodologies used in the existing solutions are highlighted. Denoising and cleaning the image which is degraded as given in Figure 3. In this method, Spatial [1] and Frequency [2] domain of are used (5). A comprehensive overview of the current state of Generative Adversarial Networks (GAN) and offers insights into their future prospects is explained in detail (6).

$$g(x, y) = h(x, y) \cdot f(x, y) + \eta(x, y) \qquad [1]$$

$$G(u, v) = H(u, v) \cdot F(u, v) + N(u, v) \qquad [2]$$

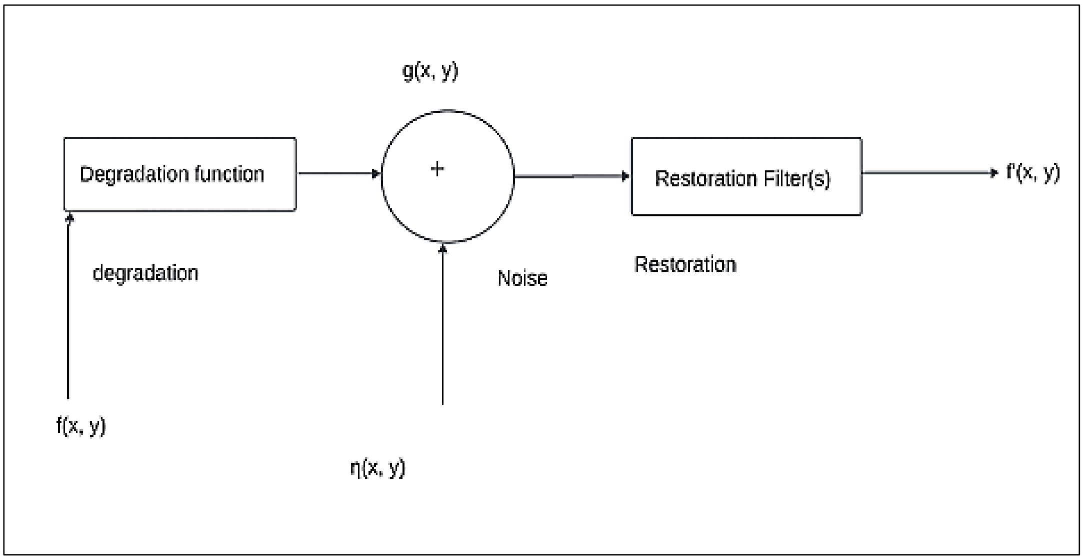


Figure 3: Image Restoration Model Filter for HexArt (5)

GAN, consist of a generator and a discriminator, both trained using the adversarial learning framework as shown in Figure 4. The primary objective of GANs is to approximate the underlying distribution of real data samples and generate new samples that resemble this distribution. Since their inception, GANs have garnered significant attention and exploration due to their vast potential across various applications, such as image and vision computing, speech and language processing, among others.

The system delves into the background of GANs, covering their proposal, theoretical foundations, implementation models, and diverse application domains. Subsequently, the strengths and weaknesses of GANs are explored and analysized. The loss function as given in equation [3].

$$Obj^{D}(\theta_D,\theta_G) = -\frac{1}{2}E_x \approx p_{data}(x)[\log D(x)] - -\frac{1}{2}E_z \approx p_z(z)[\log(1-D(g(z)))] \quad [3]$$

Notably, intricate relationship between GANs and parallel intelligence has been investigated. it can be concluded that GANs hold immense promise in the realm of parallel systems research, particularly in the context of virtual-real interaction and integration. GANs can also provide substantial algorithmic support for parallel intelligence (7). There is a very important type of GAN that this proposed system uses CycleGAN (7). The CycleGAN method helps train models to change images into different ones without needing pairs of examples. These models learn on their own using sets of pictures from one kind of image and another. Equation [4] shows the Cycle Consistency Loss and Adversarial training. These mappings aim to generate outputs with distributions identical to the target domains Y and X, respectively. It's worth to mention that, this G and F need to be stochastic functions (7).

$$loss_{cyc}(G,F,X,Y) = \frac{1}{m}[(F(G(x_i)) - x_i) + G(F(y_i) - y_i)] \quad [4]$$

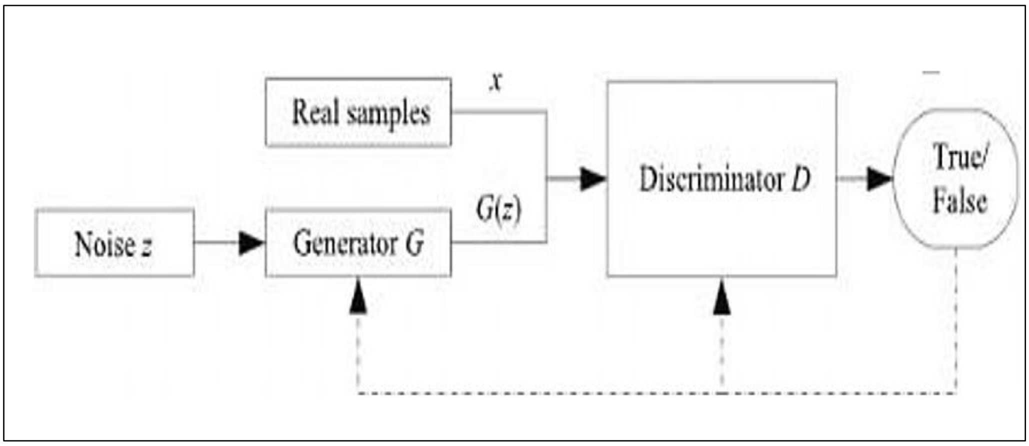


Figure 4: Structure of GAN (7)

However, when the network has ample capacity, it can map a set of input images to various random permutations of images in the target domain. Essentially, with a large enough network, different learned mappings can lead to output distributions that align with the target distribution, allowing for flexibility in the generated results. Image transformation method using cycle consistent adversarial networks (8). Improved Generative Adversarial Network(GAN) using u-nets method is proposed. This makes the image fuzziness removal process easier and better as shown in Figure 5 (9).

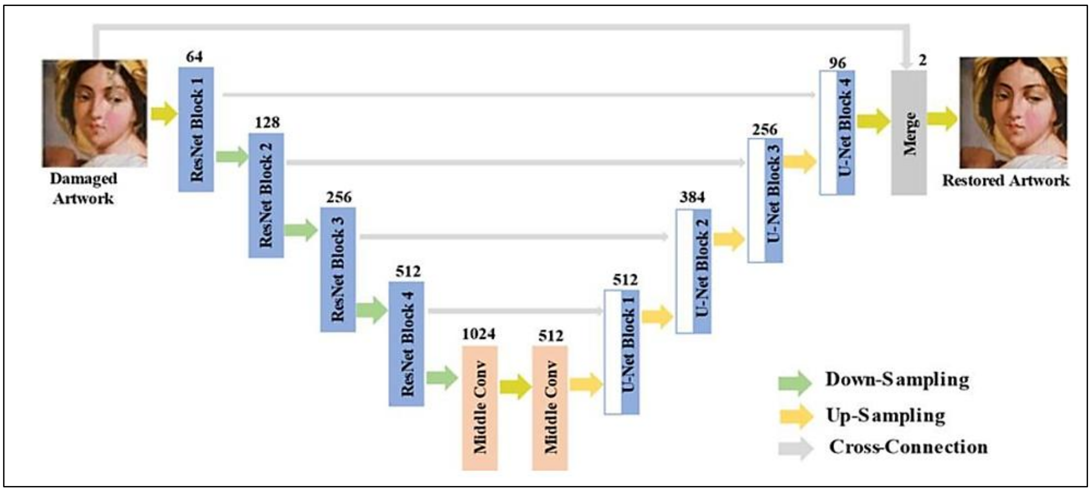


Figure 5: Architecture of Generator Suggested by Authors (7)

This technique is decision based filtering which is a combination of K means and PCA (10). To enhance image using deep learning and transformer based contextual optimization algorithm is proposed (11). Image Enhancement for visually degraded scenes proposed for object detection (12). Some researchers talk about different methods like GAN, Autoencoder networks, CNN etc. these methods are used to restore images in today's time with precision (13).
In previous studies, authors discussed about Mask R-CNN and image inpainting on Artwork Restoration through deep neural networks (14). Hybrid model was introduced that incorporates automatic mask generation relying on Mask R-CNN. Additionally, image inpainting is carried out using the U-Net architecture, featuring partial convolutions and automatic mask updates CNN based image restoration using implicit frequency selection is proposed for image enhancement (15, 16). The effectiveness of the proposed method is assessed through both qualitative and quantitative evaluations. Qualitative evaluation involves the input of expertise from three domain-specific art professionals (17). Extracted from the comprehensive document on the Research Status and Development Trend of Image Restoration Technology and Digital Image Restoration, the discourse delves into addressing the issue of discomfort through the utilization of the least square method an approach considered as the most straightforward means to resolve the problem at hand (18). Face detection methods for mask detection is proposed for detection of occluded face (19). For scene images improvement CNN based Inpainting method is explained (20). Auto-encoding variational bayes (21) and the learning a dynamic blur kernel methods are proposed for image restoration (22). The Comparing the Quality Assessment of Image Dehazing Using Different Techniques similar to inpanting. Deep image harmonization is proposed for image improvement (23).
Prompt-to-prompt image editing with cross attention control is proposed (24). Lie Detection System using Multimodal Biometric Analysis proposed for face detection (25). GAN-based image inpainting explained in detail (26). A novel Hankel singular values-based order truncation approach applied to SG filter.is proposed (27). Blind image denoising with generative adversarial networks is proposed (28, 29). Image enhancement method explained (30). For image classification CNN based method is used (31). Next Section gave details of proposed method.

Methodology

In this paper, the proposed methodology explained in different stages. The Method designed is a user-friendly GUI to breathe new life into blurry or scratched images. The system is powered by cutting-edge technologies such as GAN, deep learning, and open CV. Figure 6 shows the different stages flow for proposed method. Proposed system is tested on facial images dataset. The method unfolds in four meticulously orchestrated stages, each contributing its unique capabilities to enhance and complete the image. These four stages are:

Stage 1: Degrade Detection

The stage is done for the overall Quality improvement of the image. It is an important step to start because the historical picture which is extremely degraded and difficult to read and analyze. In this stage scratches and degradation of the image are identified to increase the quality.

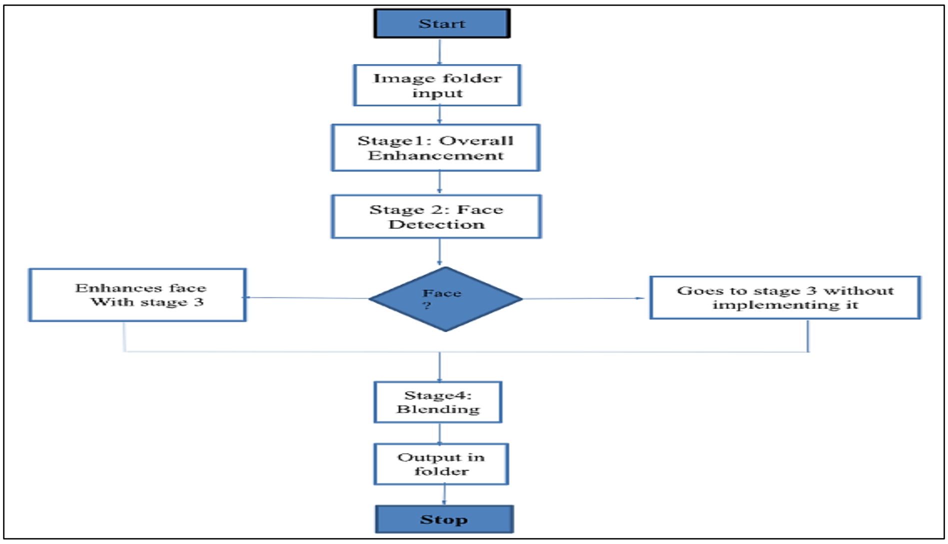


Figure 6: Flowchart of HexArt AI Image Restoration

Stage 2: Face Detection

dlib's pre-trained model is used for The face detection which is well-established component within Python libraries. This model employs 68 facial landmarks for accurate detection. The significance of facial detection lies in its crucial role in Stage 3, wherein objective is to enhance facial features. The paramount importance Dlib's detection method operates as follows:

All the features help in detecting the face first for future enhancement. Face detection is accurate because of the detection of this many points in the face.

Stage 3: Face Enhancement

This stage encompasses the refinement of facial characteristics, a process contingent upon the inputs derived from Stage 2, wherein facial detection is conducted. The rationale for positioning this enhancement stage here is predicated on the necessity of utilizing prior stage inputs for the refinement process.

Key components of Stage 3 include:

- Implementation of a specialized face enhancement model.
- The enhancement of specific facial features within the detected visages.
- The generation of superior-quality, enhanced facial images.

After this stage, the Image Shows Remarkable Enhancement Effort which is facilitated by Dlib. The successful identification of the facial components and a comprehensive enhancement of the entire image becomes evident. Notably, this enhancement is seamlessly achieved through the proficient utilization of Dlib.

Stage 4: Image Blending

The final stage of the image restoration process involves the harmonious blending of inputs from all preceding stages, culminating in the desired final output. This blending process entails aligning and warping the enhanced facial images back into the overall restored images, ensuring a coherent and visually appealing result.

Key elements of Stage 4 encompass:

The alignment and warping of the enhanced facial components into the comprehensive image. The blending of enhanced facial features with the entire image creates the ultimate output.

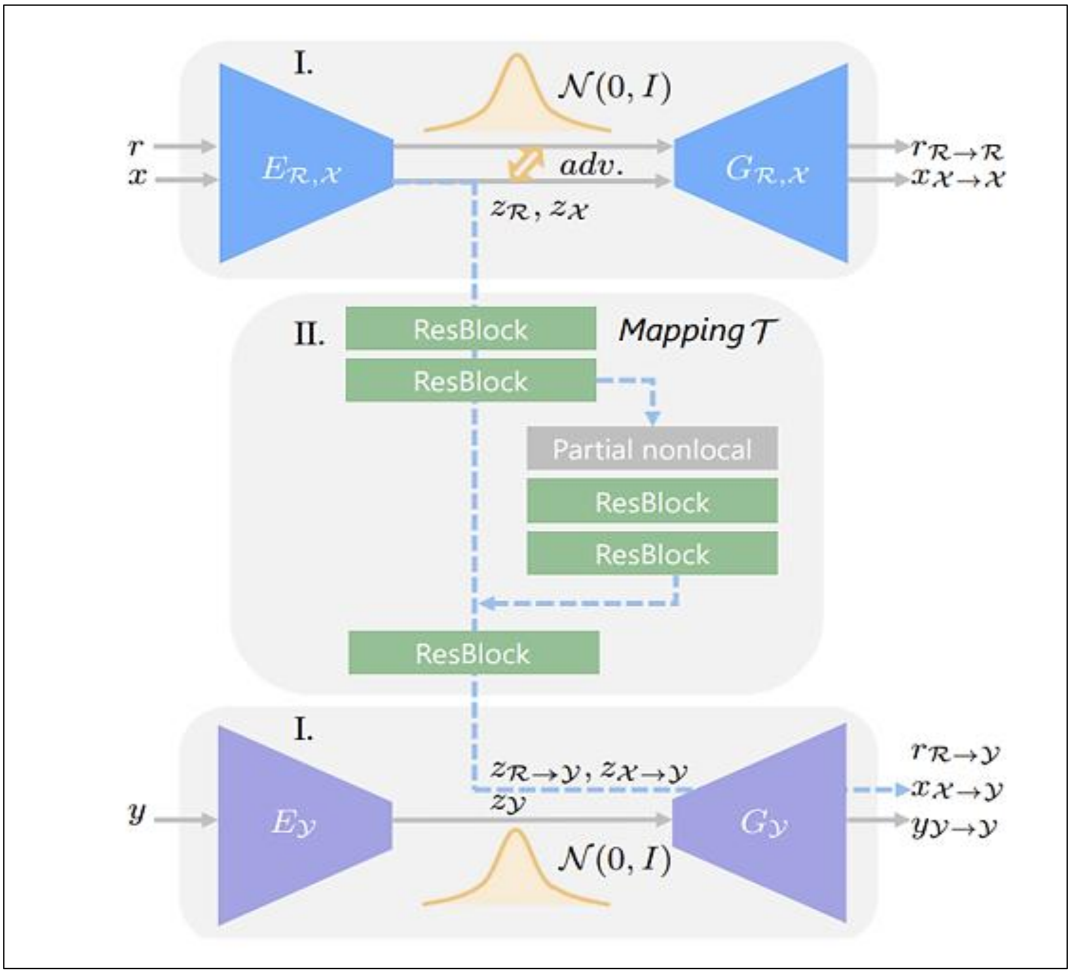


Figure 7: HexArt AI Image Restoration Architecture

Process: AI base Inpainting for Hex-Art Restoration

In the Proposed System HexArt, a two-step process is employed as shown in Figure 7. First, two VAEs (Variational Autoencoders) are trained (32). VAE1 is used for both real photos (denoted as ‘r’ in the ‘R’ category) and synthetic images (denoted as ‘x’ in the ‘X’ category). The aim is to make these images similar, and this is achieved by utilizing an adversarial discriminator. VAE2, on the other hand, is designed for clean images (‘y’ in the ‘Y’ category). These VAEs are instrumental in transforming images into a compact form in a space known as the “latent space.”

Table 1: HexArt AI Image Restoration Network Structure

Module	Layer	Kernel Size / Stride	Output Size
	Conv	7 x 7/1	256 x 256 x 64
Encoder E	Conv Conv	4 x 4/2	128 x 128 x 64
	ResBlock x 4	4 x 4/2	64 x 64 x 64
		3 x 3/1	64 x 64 x 64
	ResBlock x 4	3 x 3/1	64 x 64 x 64
Generator G	Deconv Deconv Conv	4 x 4/2	128 x 128 x 64
		4 x 4/2	256 x 256 x 64
		7 x 7/1	256 x 256 x 3
	Conv	3 x 3/1	64 x 64 x 128
	Conv Conv	3 x 3/1	64 x 64 x 256
		3 x 3/1	64 x 64 x 512
	Partial	1 x 1/1 [HTML]C0C0C03 x	64 x 64 x 512
	[HTML]C0C0C0 nonlocal	3/1	[HTML]C0C0C064 x 64 x 512
	Resblock x 2		
Mapping T	Resblock x 6	3 x 3/1	64 x 64 x 512
	Conv Conv Conv	3 x 3/1	64 x 64 x 256
		3 x 3/1	64 x 64 x 128
		3 x 3/1	64 x 64 x 64

Subsequently, the focus shifts to the restoration of damaged or old photos within this latent space. A series of opera-tions is performed to restore the old photos (’r’) to clean ones (’y’) within this latent space.

$$r_R \rightarrow y = G_y \circ T_z \circ E_R(r). \quad [5]$$

To ensure that ’R’ and ’X’ images are understood in the same way within the latent space, a technique called Variational Autoencoder (VAE) is employed. This technique compresses the images into compact forms, and an adversarial discriminator is used to further align and enhance their similarity. This alignment of ’R’ and ’X’ is a crucial aspect of the Proposed System HexArt’s approach.

In the first step, two VAEs are trained. VAE1 is used for ’r’ and ’x’ images and is responsible for mapping them into a shared latent space. VAE2 is responsible for ’y’ images. VAEs are designed to create compact representations of the images. Mathematical equation [5] are utilized to measure how well VAE1 can make the ’r’ and ’x’ images resemble the clean ones and encourage it to generate realistic images. The same process is repeated for VAE2 with ’y’ images.

The choice of VAEs over regular autoencoders is based on the fact that VAEs create denser representations within the latent space, making it easier to align ’R’ and ’X’ images and reduce the gap between them.

In addition to VAEs, an adversarial network is employed to further narrow the gap between ’R’ and ’X’ within the reduced latent space. This network ensures that ’R’ and ’X’ are mapped to the same space.

In the second step, the focus is placed on image restoration. Synthetic image pairs 'x' and 'y,' which share a compact representation in the latent space, are used. A mapping network (denoted as 'M') is used to restore these images. The advantage of this approach is that it is easier to restore images in a compact, low-dimensional latent space compared to the original high-dimensional image space. This mapping network is trained to carry out this restoration task. Mathematical terms are employed to measure the efficacy of this process. The outcome is that 'r' images can be restored to 'y' within the latent space, resulting in clean, undamaged images. This approach proves to be more effective than attempting to fix these images at the pixel level.

Different types of losses are applied to ensure the effectiveness of the restoration process and to guarantee that the images appear realistic. These loss functions in equation [6] are utilized to assess and enhance the quality of the restored images which include LSGAN denoted as LVAE1, GAN (33, 34).

$$L_{VAE}1\ (r) = KL(E_{R,X}\ (z_r/r)//\aleph(0, 1)) + \alpha E_{zr} \sim E_{R,X}((z_r\,|r)) \qquad [6]$$
$$//G_{R,X}\ (r_{R\to R}/z_r - r//_1] + L_{VAE}1,GAN\ (r)$$

Table 1 shows the major modules that used Encoder, Generator, and Mapping and explains the layers. Like the Encoder works for the convolution layer and ResBlock x 4 whereas the Generator works for the deconvolution layer. The table also talks about the Kernel size that each layer uses and its output size.

System Design for Hex-Art Restoration

A Detailed Insight into the Image Restoration Process Conducted in Stage 1. The software employed in Stage 1 distinguishes between three distinct entities: the original image, the image exhibiting issues and scratches, and the corrected image.

The Real Image: This constitutes the initial input image added to the graphical user interface (GUI) for correction and serves as the basis for the correction process. The real image is the primary focus of correction, necessitating alignment with the corrected image to remove imperfections and enhance overall quality.

The Image with Issues/Scratches: These synthetic images are computer-generated representations simulating the degradation process experienced by the photograph. Their creation serves to facilitate a more comprehensive understanding of the image's pre-degradation state.

The Corrected Image: This represents the anticipated output produced through the synthetic image. As the syn- thetic image replicates the degradation process, it offers a valuable reference for determining the desired corrected state. The proposed trained model has been fine-tuned with these features, achieving a high degree of accuracy.

After all these it can make out that the real image and the synthetic picture are close to similar, so entities' space for processing are align. To align their spaces, use VAE which encodes the images into a shared latent space. This aligned space provides major information on images and minimises the domain gap.

VAE1 for real photos and synthetic images:

- Encoder ER, X and generator GR, X.
- Objective function LVAE1(r) for real photos and LVAE1(x) for synthetic images.
- Adversarial discriminator DR, X to differentiate ZR and ZX.
- Latent adversarial loss LlatentVAE1, GAN to minimize the latent gap.

This is how the VAE by using mathematical operations combines or encodes both the images in the same shared latent space. The key innovation behind ResNet is the use of residual blocks. In traditional deep neural networks, as the network becomes deeper, it becomes increasingly challenging to train because of issues like vanishing gradients. Residual blocks help alleviate this problem by introducing "skip connections" or "shortcut connections" that allow the model to learn residual functions. These shortcut connections enable the gradient to flow more easily through the network, making it possible to train very deep neural networks effectively.

A residual block architecture as shown in Figure 8 consists of two main paths:

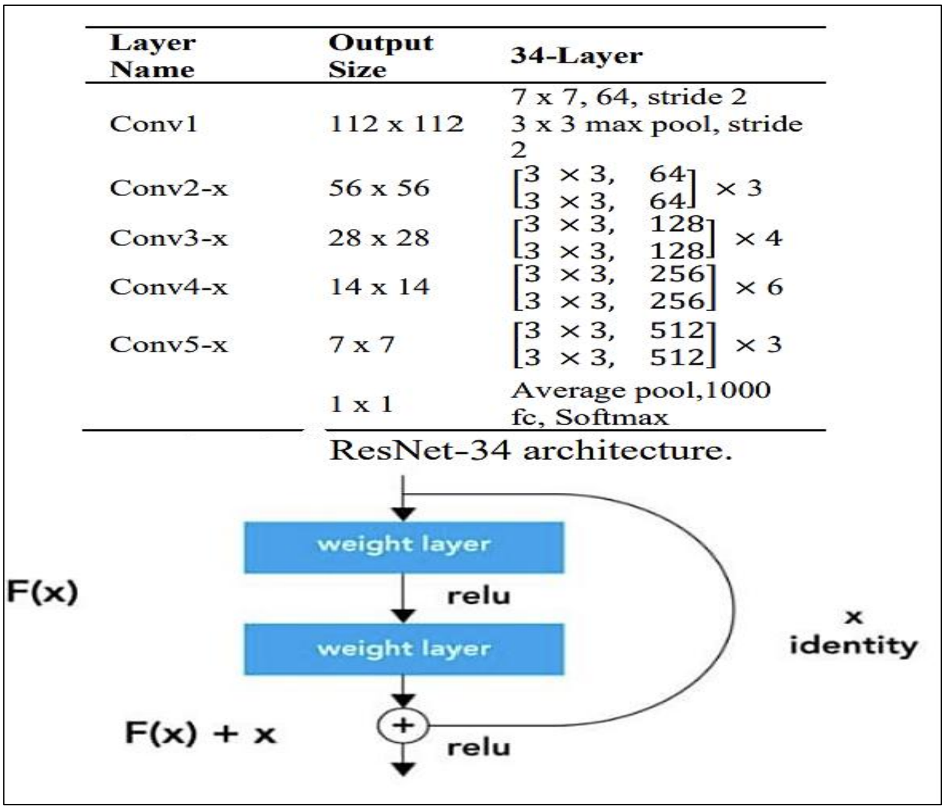
The "identity path" directly passes the input to the output without any transformation. The "residual path" contains a series of convolutional layers and other non-linear operations, which learn the residual mapping (the difference between the desired output and the input). The idea is that if the identity path is the optimal mapping, the residual path can learn to represent the difference between the input and the desired output, making it easier for the network to approximate the target function.

ResNet architecture in Figure 8, come in various depths, such as ResNet-18, ResNet-34, ResNet-101, and ResNet- 152 etc. The number in the architecture's name indicates the number of layers, and deeper networks are capable of achieving better performance on a variety of computer vision tasks. ResNet has been widely adopted and has set the state of the art in many image classification and object detection challenges as

$$y = F(x) + x \qquad [7]$$

Where x is input, y is output, and F is the residual function. Each basic block consists of 2 convolution layers and a pooling layer (3x3 size), followed by a (ReLU) activation function and batch normalization (BN). Using ResNet has greatly improved the performance of neural networks, where such networks given in equation [7].

are stacked with more layers for the creation of a deeper architecture, and hence, deeper learning, in contrast with shallower learning. ResNet-34 (ResNet with 34 layers) consists of 33 convolution layers a max-pooling layer (3x3 size) and an average pooling layer, followed by a fully connected layer.



Layer Name	Output Size	34-Layer
Conv1	112 x 112	7 x 7, 64, stride 2 3 x 3 max pool, stride 2
Conv2-x	56 x 56	$\begin{bmatrix} 3 \times 3, & 64 \\ 3 \times 3, & 64 \end{bmatrix} \times 3$
Conv3-x	28 x 28	$\begin{bmatrix} 3 \times 3, & 128 \\ 3 \times 3, & 128 \end{bmatrix} \times 4$
Conv4-x	14 x 14	$\begin{bmatrix} 3 \times 3, & 256 \\ 3 \times 3, & 256 \end{bmatrix} \times 6$
Conv5-x	7 x 7	$\begin{bmatrix} 3 \times 3, & 512 \\ 3 \times 3, & 512 \end{bmatrix} \times 3$
	1 x 1	Average pool,1000 fc, Softmax

Figure 8: ResNet-34 Learning Architecture

Figure 8 shows how the transition from the correct image to the corrupted image is made by mapping them together. Proceed with the image restoration in the latent space by pairing the images (The real image, and the correct image). This is a method to learn image restoration in the latent space. Use a mapping network to translate the corrupted image to the space of the correct image. In other words, the mapping network maps the corrupted image to the correct image for the overall correction. Latent space loss LT, '1 to penalize the 1' distance of corresponding latent codes. Adversarial loss LT, GAN to improve the realism of the restored images. Latent feature matching loss LFM to ensure structural consistency.

Results and Discussion

This section of the paper is aimed to highlight the results considering facial images. Figure 9 shows the GUI homepage of HexArt AI Image Restoration which has options like Input folder, Output folder, GPU, Checkpoint Name and check box options for images with scratches and high resolution (HR). The input image serves as the target for our correction process, which will involve multiple stages. The image exhibits noticeable degradation, likely due to its age, with significant scratches and imperfections. Proposed system designed to rectify these issues. Within the designated folder, there are two types' images, and our objective is to apply the correction process to all images.

Figure 10 provides a comprehensive overview of the sequential application of stages to the input image. Located on the left-hand side, the graphical user interface (GUI) empowers users to specify the directories for input and output image storage and flexibility to set the number of processing epochs. One

can choose 100 epochs to deliver highly accurate outcomes.
The interface further offers check boxes for "With Scratch" and "HR" options, each tailored to specific use cases. "With Scratch" caters to images afflicted with scratches, employing a model explicitly trained for this purpose to achieve remarkably precise results. As elucidated earlier, a diverse array of techniques and stages collaborates to produce the final restored image. The Global Restoration component, complemented by the "With Scratch" option, initially identifies and subsequently eliminates these scratches. On the other hand, the "HR" option is tailored for delivering high-resolution images that offer enhanced clarity and superior processing. This feature enables users to input high-quality images while expecting a commensurate quality in the output.

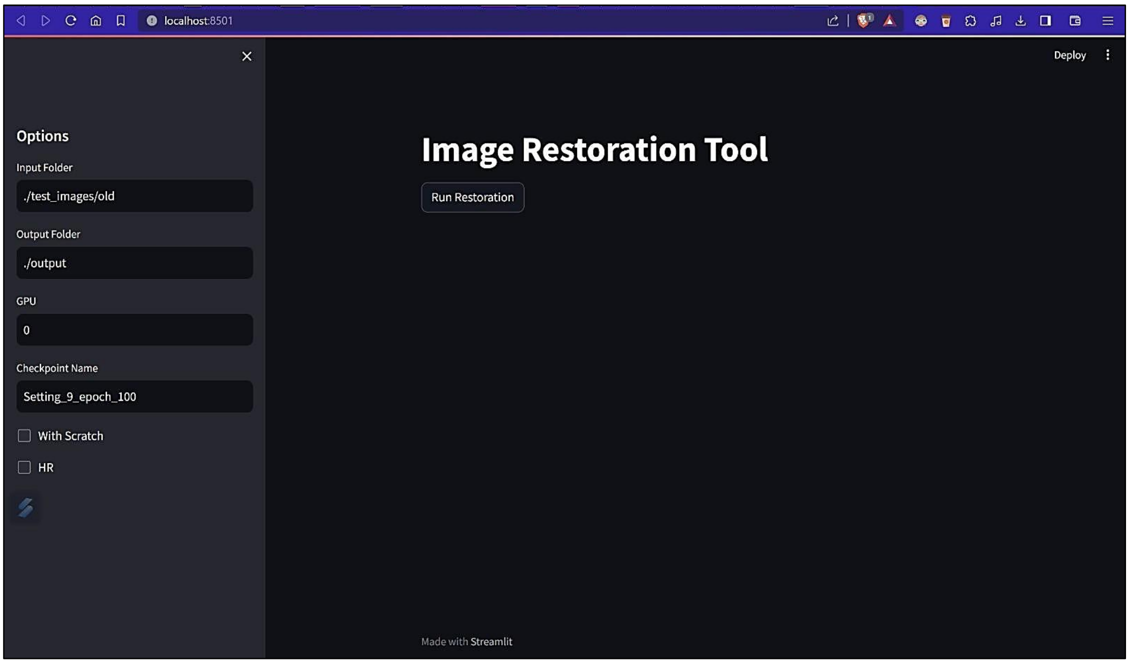


Figure 9: GUI Homepage of HexArt AI Image Restoration

The only limitation lies in its heightened GPU computational demands, occasionally leading to skips in processing due to insufficient memory. Figure 11 and Figure 12 shows the basic execution of HexArt with various stages in real-time and the final output images after computations respectively.
The outcomes of the proposed system is achieved through GPU-powered system. The proposed software adeptly erases a wide array of blemishes, fills in missing colours and objects as deemed suitable. Furthermore, it is unparalleled in its ability to rejuvenate historical images, rendering them comprehensible through cutting-edge technology, boasting an impressive accuracy rate. It consistently excels in the complete removal of imperfections, yielding images that are both easily legible and aesthetically gratifying. Each of the Stages further enhances the image, removing various issues from the pictures. The main enhancement stages are shown in Figure 12 as stage 1 Global Restoration and Face enhancement. The ROC curve illustrates the performance of scratch detection in various data settings as shown in Figure 15. By combining synthetic structured degradations and a small amount of labelled data, the scratch detection network in the Proposed System HexArt achieves excellent results. The face enhancement network is trained alongside the restoration network to enhance its generalization capabilities. The output of the triplet domain translation network denoted as 'rf,' ensures better generalization. This training scheme effectively suppresses any generated artifacts.
During the inference stage, the system first identifies the face parts in arbitrary photos and then refines this region using the proposed enhancement network. Sometimes, due to the generative nature of the model, there may be colour discrepancies between the reconstructed faces and the input degraded faces. This issue is resolved using histogram matching. Ultimately, the reconstructed facial representation undergoes integration with the initial input photograph, achieving the conclusive outcomes via a process of linear blending.

In terms of implementation, the training dataset involves the synthesis of old photos using images from the Pascal VOC dataset. Additionally, a dataset comprising 2,500 old photos is collected to form the old photo dataset. The training of the face enhancement network utilizes 10,000 aligned high-resolution face images. Figure 13 shows sample of dataset of old images. Figure 14 shows clear and enhanced output images generated by proposed model.

Training details include the use of the Adam solver with parameters $\beta 1 = 0.5$ and $\beta 2 = 0.99$, a learning rate set to 0.05 for the first 50 epochs, with a linear decrease to zero. Random cropping of images to 256x256 is performed during training. The parameters in Equations are set to $\alpha = 10$, $\lambda 1 = 50$, and $\lambda 2 = 10$.

Data generation involves two types of degradation: unstructured and structured. Unstructured degradation consists of operations like Gaussian white noise, Gaussian blur, JPEG compression, colour adjustments, and box blur, applied with varying parameters in a random manner. These operations are randomly omitted with a 30% probability to add diversity. Structured degradation includes the use of synthetic scratch and paper textures, combined with natural images. Film grain noise and random blur are also introduced to simulate overall defects.

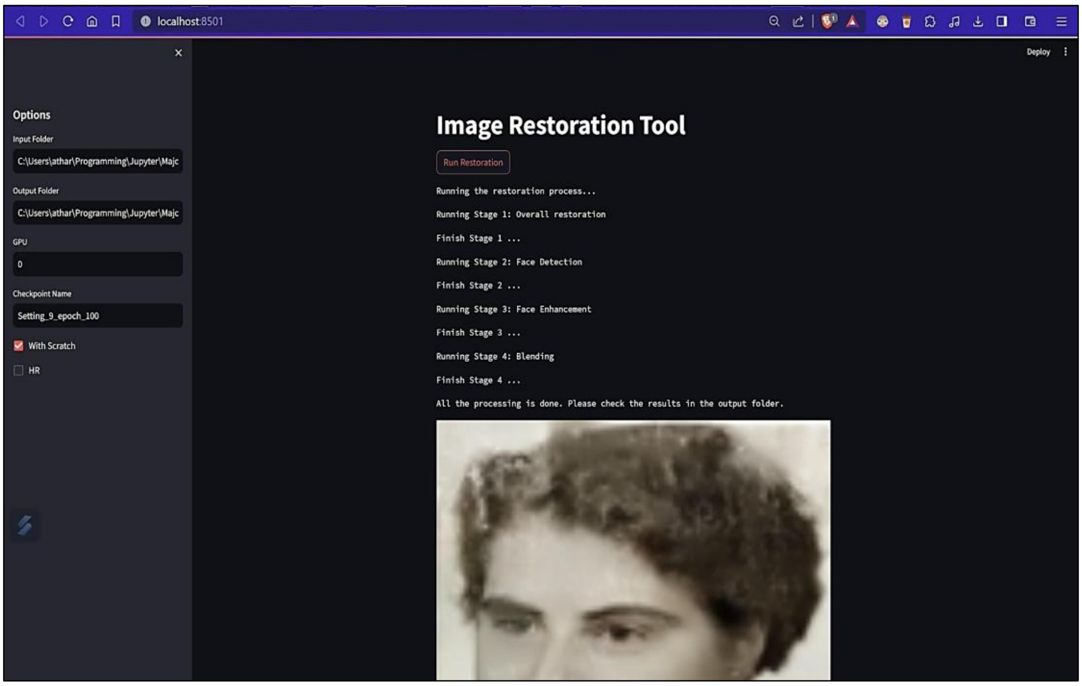


Figure 10: GUI Page of Hexart AI Image Restoration: Execution of Proposed System

Table 2: Study Results for Hexart AI Model on Dataset

Method	PSNR	SSIM	LPIPS	FID
Input	12.82	0.50	0.58	306.81
Attention[10]	**24.1**	**0.71**	0.30	208.2
DIP[15]	22.6	0.44	0.54	194.56
Pix2pix[14]	22.18	0.66	**0.24**	**135.16**
Sequential[16]	22.70	0.51	0.48	180.76
HexArt with partial non-local	**23.11**	**0.68**	**0.26**	**130.32**

To improve scratch detection on actual old photos, a dataset of 300 real old photos is collected and manually annotated for local defects, 200 of these images are used for training, and the remaining ones are for testing. Incorporating real data significantly enhances scratch detection performance on genuine old photos, achieving an AUC of 0.81. Digital Image Prior, often referred to as DIP, is employed to enhance an image without the need for additional training data beyond the image itself. This approach is effective in tasks such as noise reduction, super-resolution, and image inpainting (33). Pix2pix is a method for Image-to-Image translation utilizing Generative Adversarial Networks (GANs). In the Pix2Pix GAN, the loss function is modified to ensure that the generated image is not only coherent with the content of the target domain but also serves as a credible translation of the input image (33). Sequential is an approach employed in the application of the algorithm, where actions are executed in a specific sequence. When the technique is run sequentially in techniques like BM3D (34) and Edge Connect, it produces certain metrics, which are detailed in the accompanying Table 2 (35, 36). In a quantitative evaluation of various models using synthetic images from the DIV2K dataset, four metrics were applied, as shown in Table 2.

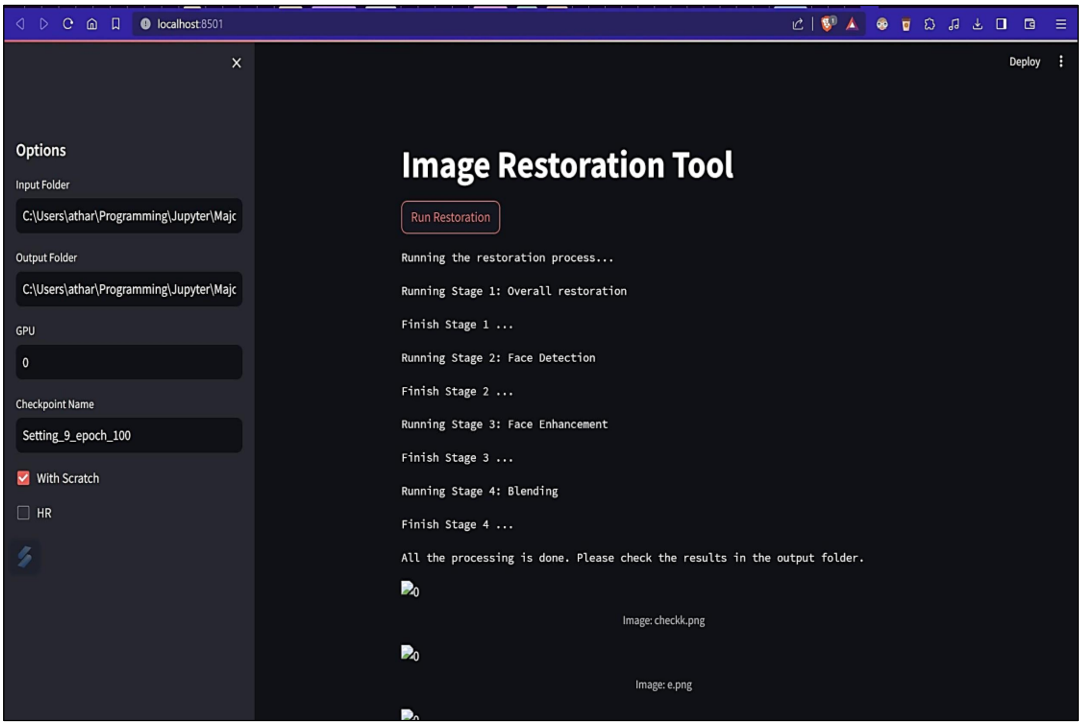


Figure 11: GUI of Hexart with Various Stages while Computing

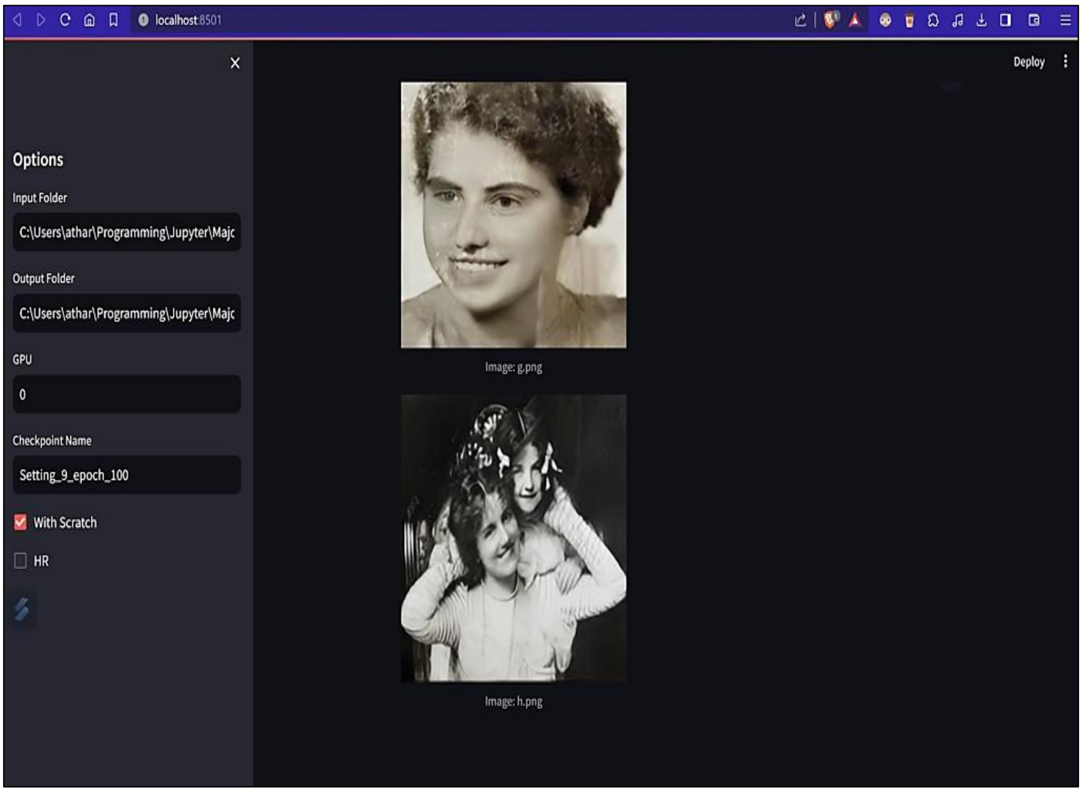


Figure 12: GUI of Hexart Final Output Images after Computation

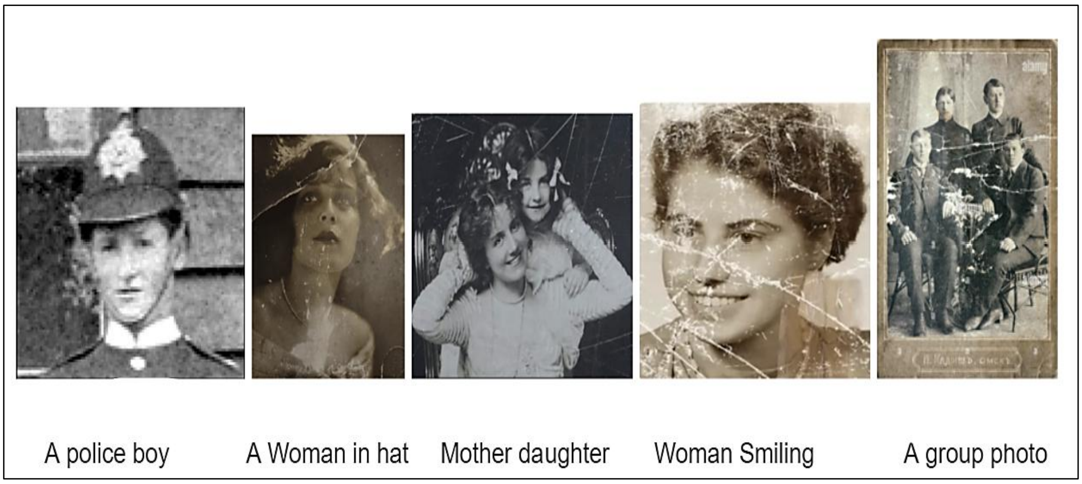


Figure 13: Input Images to Feed into the HexArt Model

Figure 14: Output Images Generated by HexArt Model

Conventional metrics such as peak signal-to-noise ratio (PSNR) equation [8] and the structural similarity index (SSIM) equation [9] were utilized for evaluating subtle distinctions at the low level between the restored output and the ground truth. Among these metrics, the operational-wise attention method demonstrated the most favorable PSNR/SSIM scores, with the Proposed System HexArt securing the second position. PSNR, succinctly expressed through mean squared error (MSE) equation [10], is calculated based on a noise-free m×n image I and its noisy counterpart K.

$$PSNR = 10 \cdot \log_{10} \frac{MAX^2}{MSE} \quad [8]$$

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c1)(\sigma_x^2 + \sigma_y^2 + c2)} \quad [9]$$

In equation [9] μ_x is the average of x, μ_y is the average of y, σ^2 is the variance of x, σ^2 is the variance of y, σ_{xy} is the *x, y* covariance of x and y, c_1 is $(k_1L)^2$ and c_2 is $(k_2L)^2$;

$$MSE = \frac{1}{mn\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}|I(i,j) - K(i,j)|^2} \quad [10]$$

To better align with human perception, the perceptual image patch similarity (LPIPS) metric was adopted. Here, both Pix2pix and the Proposed System HexArt scored similarly, outperforming the operational-wise attention method. Fre´chet Inception Distance (FID) was used to gauge the quality of generative models. In this case, the Proposed System HexArt and Pix2pix excelled, with the Proposed System HexArt having a slight quantitative advantage. In summary, the Proposed System HexArt proves comparable to leading techniques in synthetic data evaluation.

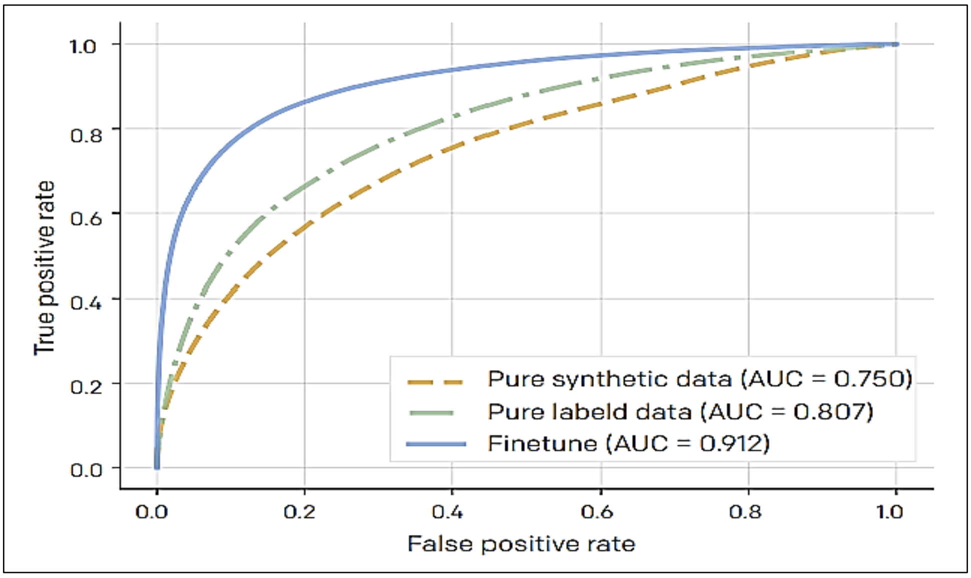


Figure 15: ROC of Hexart for Scratch Detection of Different Data Settings

Conclusion

In conclusion, the presentation of the proposed image restoration software and the sequence of stages applied to the input image reveal a comprehensive and effective solution for image enhancement. The software's intuitive GUI, offering options for input and output directories and processing epochs, provides users with a user-friendly and versatile tool for image restoration. The software consistently produces outstanding results, successfully eliminating various types of imperfections, intelligently reconstructing colors and objects, and excelling in the restoration of historical images with an impressive accuracy rate of up to 81 percent. The proposed system underscores the power of cutting-edge technology in breathing new life into visual relics and bringing the past to vivid and visually pleasing life. Looking ahead, this proposed system recognizes the potential for further expansion of its scope. Video enhancement stands as a compelling frontier that demands the development of a distinct model capable of reconstructing missing components and providing comprehensive restoration to moving images. This ambitious endeavor offers a promising avenue for future research, one that aligns with the ever-growing demand for advanced image and video restoration technologies.

Abbreviations

CNN: Convolutional Neural Network, FID: Fre´chet Inception Distance, GAN Generative Adversarial Network, GUI: Graphical user interface, HR: High resolution, MSE: Mean Square Error, PSNR: Peak to Signal to noise Ratio, VAEs: Variational Autoencoders.

Acknowledgement

We wholeheartedly thank all our colleagues for their support in the timely completion of this research work.

Author Contributions

All authors are contributed equally.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the content of this article.

Ethics Approval

Not Applicable.

Funding

None.

References

1. Gonzalez RC, Woods R. Digital Image Processing. 3rd ed. USA: Prentice-Hall. 2006. https://books.google.com/books?hl=en&lr=&id=a62xQ2r_f8wC&oi=fnd&pg=PA19&dq=1.%09Gonzalez+RC,+Woods+R.+Digital+Image+Processing.+3rd+ed.+USA:+Prentice-Hall%3B+2006.&ots=3B4uS3nG-E&sig=CN2zZ8bdIcatRBscrCTfawNkVYE
2. Abbas RF. Review on some methods used in image restoration. Int Multidiscip Res J. 2020;10:13–6.
3. Kumar MK, Kumari VR, Kishore, Bolla S, Sai Ram, Ramakrishna, Aravinda. A hybrid model for face detection using HAAR cascade classifier and single shot multi-box detectors based on OpenCV. Int Res J Multidiscip Scope. 2024; 5(01):650–60.
4. Kanrar S, Maji S. A study on image restoration and analysis. In: Advance Concepts of Image Processing and Pattern Recognition. Transactions on Computer Systems and Networks (TCSN). Springer. 2022: 35–61. https://link.springer.com/chapter/10.1007/978-981-16-9324-3_3
5. Madhan K, Shanmugapriya N. Real time object detection in low light environment using adaptive image information enhancement algorithm. Int Res J Multidiscip Scope. 2024;5(1):62–73.
6. Wang K, Gou C, Duan Y, Yilun L, Zheng X, Wang F-Y. Generative adversarial networks: Introduction and outlook. IEEE/CAA J Autom Sinica. 2017;4:588–98.
7. Goodfellow I. NIPS 2016 tutorial: Generative adversarial networks. arXiv Preprint. 2016. arXiv: 1701.00160. https://arxiv.org/abs/1701.00160
8. Zhu J-Y, Park T, Isola P, Efros A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proc IEEE Int Conf Comput Vis. 2017. 2223–32. http://openaccess.thecvf.com/content_iccv_2017/html/Zhu_Unpaired_Image-To-Image_Translation_ICCV_2017_paper.html
9. Kumar P, Gupta V. Restoration of damaged artworks based on a generative adversarial network. Multimedia Tools and Applications. 2023 Nov; 82(26):40967-85.
10. Ankita L. Research paper on image restoration using decision based filtering techniques. Int J Eng Dev Res. 2016;4:477–81.
11. Anandhi S, Jaiganesh M. An enhanced image restoration using deep learning and transformer based contextual optimization algorithm. Sci Rep. 2025;15:10324.
12. Liu H, Jin F, Zeng H, Pu H, Fan B. Image enhancement guided object detection in visually degraded scenes. IEEE Trans Neural Netw Learn Syst. 2024;35(10): 14164–77.
13. Liu Z. Literature review on image restoration. J Phys Conf Ser. 2022;2386:012041.
14. Gupta V, Sambyal N, Sharma A, Kumar P. Restoration of artwork using deep neural networks. Evolving Systems. 2021 Jun;12(2):439-46.
15. Huang ZJ, Lu W-H, Patel B, Chiu P-Y, Yang T-Y, Tong HJ. Convolutional neural network-based image restoration (CNNIR). In: Proc IEEE/ASME Int Conf Mechatron Embed Syst Appl (MESA). 2022: 1–6. https://ieeexplore.ieee.org/abstract/document/10004461/
16. Hu J, Wang Z. A novel CNN architecture for image restoration with implicit frequency selection. Connect Sci. 2025;37(1):2465448.
17. Suganuma M, Liu X, Okatani T. Attention-based adaptive selection of operations for image restoration in the presence of unknown combined distortions. In: Proc IEEE/CVF Conf Computer Vision and Pattern Recognition (CVPR). 2019: 9039–48. http://openaccess.thecvf.com/content_CVPR_2019/html/Suganuma_Attention-Based_Adaptive_Selection_of_Operations_for_Image_Restoration_in_the_CVPR_2019_paper.html
18. Yu L, Guo J, Yu C. Research status and development trend of image restoration technology. J Phys Conf Ser. 2022;2303:012081.
19. Bamankar S, Bhoir P, Pednekar S, Phadke G. Face mask and body temperature scanning system for COVID-19. In: Proc IEEE ICONAT. 2023:1-6. https://ieeexplore.ieee.org/abstract/document/10080641/
20. Mohite T, Phadke G. Image inpainting with contextual attention and partial convolution. In: Proc IEEE AISP. 2020:1-6. https://ieeexplore.ieee.org/abstract/document/9073008/
21. Kingma D, Welling M. Auto-encoding variational Bayes. arXiv Preprint. arXiv:1312.6114.2013. http://web2.cs.columbia.edu/~blei/fogm/2018F/materials/KingmaWelling2013.pdf
22. Yang Z, Xia J, Li S, Huang X, Zhang S, Liu Z, Fu Y, Liu Y. A Dynamic Kernel Prior Model for Unsupervised Blind Image Super-Resolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2024. p.26046–56. doi:10.1109/CVPR52733.2024.02461
23. Sigh S, Phadke G. Comparing the quality assessment of image dehazing using different techniques. In: Proc Int Conf Commun Intell Syst (ICCIS). 2021. https://www.researchgate.net/publication/360544684_Comparing_the_Quality_Assessment_of_Image_Dehazing_Using_Different_Techniques
24. Tsai YH, Shen X, Lin Z, Sunkavalli K, Lu X, Yang MH. Deep Image Harmonization. CoRR. 2017, http://arxiv.org/abs/1703.00069
25. Hertz A, Andonian A, Pfister H, Freeman WT. Prompt-to-prompt image editing with cross attention control. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2022. https://arxiv.org/abs/2208.01626
26. Phadke G, Chhabria A, Gulwani R, Mistry Y, Agarwal S, Thusoo R, Das S. Lie detection system using multimodal biometric analysis. J Inf Syst Eng Manag. 2025; 10(27). https://jisem-journal.com/index.php/journal/article/view/4399

27. Elharrouss O, Almaadeed N, Al-Maadeed S, Akbari Y. Image inpainting: A review. CoRR [Internet]. 2019 http://arxiv.org/abs/1909.06399
28. Agarwal S, Phadke G, Mistry Y, Mohindroo P. A novel Hankel singular values-based order truncation approach applied to SG filter. Lett High Energy Phys. 2024. https://www.lettersinhighenergyphysics.com/index.php/LHEP/article/view/694
29. Chen J, Chen J, Chao H, Yang M. Image Blind Denoising with Generative Adversarial Network Based Noise Modeling. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2018. doi:10.1109/CVPR.2018.00333
30. Mao X, Li Q, Xie H, Lau RYK, Wang Z, Smolley SP. Least squares generative adversarial networks. In: Proc IEEE Int Conf Comput Vis. 2017: 2794–802. DOI: 10.1109/ICCV.2017.304.
31. Varma A, Maurya Y, Kamath V, Chavan I, Phadke G, Das S. Comparative analysis of CNN models for insect detection system. In: Proc Int Conf Intell Comput Big Data Anal (ICICBDA). 2024. p. 223–36. https://link.springer.com/chapter/10.1007/978-3-031-74682-6_17
32. Wang TC, Liu MY, Zhu JY, Tao A, Kautz J, Catanzaro B. High-resolution image synthesis and semantic manipulation with conditional GANs. In: Proc IEEE Conf Comput Vis Pattern Recognit (CVPR). 2018: 8798–807. DOI: 10.1109/CVPR.2018.00917
33. Ulyanov D, Vedaldi A, Lempitsky V. Deep image prior. In: Proc IEEE Conf Comput Vis Pattern Recognit (CVPR). 2018: 9446–53. https://openaccess.thecvf.com/content_cvpr_2018/papers/Ulyanov_Deep_Image_Prior_CVPR_2018_paper.pdf
34. Dabov K, Foi A, Katkovnik V, Egiazarian K. BM3D image denoising with shape-adaptive principal component analysis. In: SPARS'09-Signal Process Adapt Sparse Struct Represent. 2009. Saint-Malo, France. Rennes: Inria Rennes - Bretagne Atlantique. https://inria.hal.science/inria-00369582
35. Brownlee J. A gentle introduction to pix2pix generative adversarial network. 2019. https://machinelearningmastery.com/pix2pix-for-image-to-image-translation-with-tensorflow/
36. Nazeri K, Ng E, Joseph T, Qureshi FZ, Ebrahimi M. EdgeConnect: Generative image inpainting with adversarial edge learning. arXiv Preprint. arXiv:1901.00212.2019.https://arxiv.org/abs/1901.00212