HexART :AI Art restoration

Shruti Bhattacharyay, Atharva Kapade, Tisha Agarwal, Lisshika Shetty,Gargi Phadke*, Siuli Das*
*Computer Engineering,Instrumentation Engineering**
Ramrao Adik Institute of Technology
D Y Patil Deemed to Be University
Navi Mumbai, Maharashtra
bhattacharyayshruti0808@gmail.com, atharva.kapade18@gmail.com
lisshika588@gmail.com, tishaagarwal2002@gmail.com

Abstract*—This Paper is a comprehensive examination of the understanding and implementation of image restoration software, a vital necessity in the context of the escalating use of digital photography and the declining reliance on hard-copy photographic formats. As a consequence, a multitude of photographic memories now languish in home storage for extended periods, inevitably succumbing to the erosion of image quality over time. However, this paper aim to illuminate the transformative potential of cutting-edge technology and complex algorithms in the restoration of these deteriorating images. The proposed system delves deeply into the intricate mechanisms that underpin image restoration technology, driven by sophisticated algorithms and advanced image processing techniques It emphasises the myriad benefits, including image preservation, historical documentation, enhanced accessibility, creative possibilities, and educational applications. At the same time, the proposed system acknowledges the formidable challenges posed by algorithmic intricacy and the resource-intensive nature of the restoration process, emphasizing the need for ongoing research and development efforts. In summation, image restoration software emerges as an indispensable tool in the safeguarding of our digital heritage, breathing new life into visual narratives, albeit with the continuing demand for innovative solutions to surmount the complexities it presents. This proposed system aims to use technologies like VAEs, open CV, GAN, deep learning, and their applications to propose this system, which might essentially change the course of photography and bring a change into the field of archaeological surveys, helping historical discoveries and uncovering the times and truths of the world that were hidden for centuries in the form of pictures.

I. 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 like discolouration, 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 from old times hold not only memories but also information that reveals truths about the world and culture at that particular time, which helps us learn about the people of that time along with their beliefs. Those images are the ones that have low resolution because of the hardware limitations of that time. They are also mostly scratched and damaged due to being buried in a place. By restoring these photos, the proposed system may open a door to the past and learn things that were once hidden—and most likely even destroyed—in those pictures.

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, beckoning us to unravel the mysteries of bygone eras. Like this picture here: Figure 1 shows

Fig. 1. Distorted First image having unreadable damage and low resolution

the low-resolution image which seems to be old and quite unreadable and requires some heavy restoration so that it might be converted into a format that is readable.

This system was proposed with a lot of motivations and a lot of 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.
- This can also be used to restore images from old times like 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.
- 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.
- This system can also be used for study purposes because this system uses complex systems to correct these images.

II. Problem Statement and Objectives

A. Problem Statement

In essence, this proposed system is a time-travelling 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.

This paper aims to discuss this software's current position and compare it to what can be done in the very near future there are also a few problems that remain due to the complexity of the algorithm which needs to be addressed. This complexity if solved then the software may even start restoring things like Video etc very smoothly and most importantly with speed.

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.

B. Objectives

This Proposed system has some objectives that it aims to reach:

- 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.

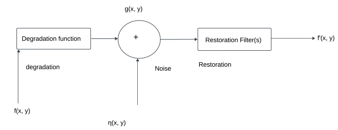

Fig. 2. Image Restoration Model Filter for HexArt[2]

III. Literature Survey

This part of the paper aims to introduce to the reader suggestive literature that led to the identification and subsequent selection of the task. This section should provide all the salient features provided by the existing available system. Also, the limitations of available solutions should be mentioned. The work can include a survey of the people who use these solutions/potential users of this solution w.r.t. their requirements and improvements expected. The survey of the technologies and methodologies used in the existing solutions can also be highlighted. This should end by introducing to the reader the general overall observations of the survey as the concluding paragraph[1].

A. Survey of existing System

Soumen Kanrar and Srabanti Maji's paper [2] is all about the noise that is present in the images that are degraded which also talks about how these noise and uncleaned pixels are removed and hence cleaned. They also talk about cleaning done in two domains: Frequency and Spatial which have different formulas and methods for application.

For Spatial domain:

$$g(x, y) = h(x, y) \cdot f(x, y) + \eta(x, y) \tag{1}$$

[2]

For Frequency domain:

$$G(u, v) = H(u, v) \cdot F(u, v) + N(u, v) \tag{2}$$

[2]

Rabab Farhan's paper is talking about the image fuzziness and removal of them. It says about mean filter, median filter, wavelet and even Neural network for the removal of these fuzziness with ease[1].

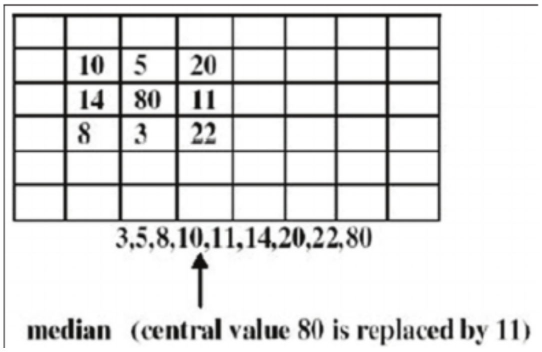

Fig. 3. Median filter working[1]

Praveen Kumar and Varun Gupta talks[3] about GAN and how they are not always usefull and hence they devised a paper with a new and improved GAN using u-nets to develop and that only makes the image fuzziness removal process easier and better[3].

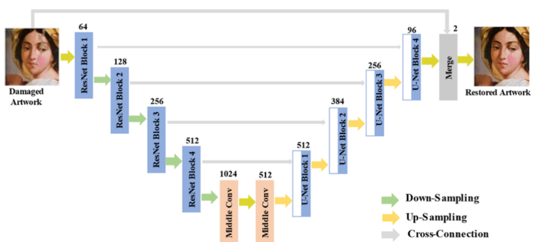

Fig. 4. Architecture of generator suggested by authors[3]

Paper from Ankit and Er. Lavina discusses about a technique which is used for filtering of the images they use a decision based filtering which is a combination of Kmeans and PCA. [4]

[4]

Fig. 5. PSNR comparison of Hybrid filter and decision based filtering

PSNR means peak signal-to-noise ratio and more the PSNR value, better the method is and the paper shows clearly the comparison of two methods using PSNR value[5].

The paper by Zihan Liu talks about different methods like GAN, Autoencoder networks, CNN etc. that are used to restore images in today's time and how each of these methods work along with their precision. This paper defined the parameters to check the accuracy of the proposed system as well. This paper defined the PSNR and its formula to calculate the proposed systems accuracy.

	Real data	CVAE	CGAN	FM-CGAN	CVAE-GA
Top-1 acc	99.61%	8.09%	61.97%	79.76%	97.78%

Fig. 6. Quantitative result of different methods proposed by author[5]

B. Methods survey

Mask R-CNN and image inpainting were explored in the research paper on Artwork Restoration through deep neural networks. The authors introduced a hybrid model 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. 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.[6].

Regularization: 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.[7].

Generative Adversarial Networks[3]: Generative Adversarial Networks, commonly known as GANs, consist of a generator and a discriminator, both trained using the adversarial learning framework. 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. This review paper provides a comprehensive overview of the current state of GANs and offers insights into their future prospects[8].

Initially, the system delves into the background of GANs, covering their proposal, theoretical foundations, implementation models, and diverse application domains. Subsequently, we explore the strengths and weaknesses of GANs and analyze the trends shaping their evolution. Notably, we investigate the intricate relationship between GANs and parallel intelligence, concluding that GANs hold immense promise in the realm of parallel systems research, particularly in the context of virtual-real interaction and integration. Clearly, GANs can provide substantial algorithmic support for parallel intelligence[8].

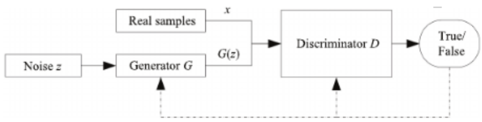

Fig. 7. Stucture of GAN

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

[8]

Equation 3 is the equation for the loss function. There is a very important type of GAN that this proposed system uses CycleGAN. 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, even if they aren't connected in any obvious way.

This simple yet powerful technique gives impressive results, especially in turning pictures of horses into zebras and vice versa. Image-to-image translation means controlled changes to a picture, and usually, you need lots of examples of pairs of pictures, which can be challenging to put together or may not even exist.

CycleGAN is a method for training models that can change images without needing paired examples. It does this using a type of architecture called GAN, and it works with sets of images from two different types that don't have to be related. People have used CycleGAN in different ways, like changing seasons in pictures, transforming objects, transferring styles between images, and even making pictures from paintings. Equation 4 mentions the equation for Cycle Consistency Loss and Adversarial training, it's theoretically possible to train mappings G and F. These mappings aim to generate outputs with distributions identical to the target domains Y and X, respectively. It's worth noting that for this to happen, G and F need to be stochastic functions[9].

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[10].

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

IV. Proposed Methodology

A. Stages

Introducing our innovative system: a user-friendly GUI designed to breathe new life into blurry or scratched images. Powered by cutting-edge technologies such as GAN, deep learning, and open CV, our platform has been meticulously crafted using Streamlit.

Our solution unfolds in four meticulously orchestrated stages, each contributing its unique capabilities to enhance and complete the image. These four stages are:

1. Stage 1: This stage is done for the overall Quality improvement of the image and this is an important step to start with because the historical picture that we are trying to improve has extreme degradation at the level that it is very difficult to read and hence we first are required to increase the quality for even the software to able to read the scratches and degradation before correcting them.

Significant progress is evident in the images, with all discernible scratches having been successfully eliminated through the methodology outlined below.

2. Stage 2: The face detection model employed in our study leverages dlib's pre-trained model, a 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 our objective is to enhance facial features, thus emphasizing its 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.

In the images presented, we observe a deliberate emphasis on facial elements; our software's primary focus is the accurate detection and enhancement of facial attributes.

3. 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.

Post-facial detection, the Image Shows Remarkable Enhancement Effort, Facilitated by Dlib Upon the successful identification of the facial components, a comprehensive enhancement of the entire image becomes evident. Notably, this enhancement is seamlessly achieved through the proficient utilization of Dlib.

4. Stage 4: Image Blending The final stage of our 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.

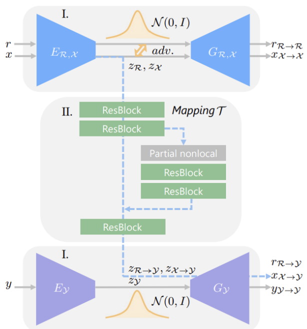

Fig. 8. HexArt AI Image Restoration Architecture

B. Process

In the Proposed System HexArt, a two-step process is employed as shown in fig 8.

First, two VAEs (Variational Autoencoders)[11] are trained. 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."

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

$$r_R \to y = G_y \circ T_z \circ E_R(r). \tag{5}$$

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 formulas 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.

TABLE I
HEXART AI IMAGE RESTORATION NETWORK STRUCTURE

Module	Layer	Kernel size / stride	Output size
Encoder E	Conv Conv Conv ResBlock x 4	7 x 7/1 4 x 4/2 4 x 4/2 3 x 3/1	256 x 256 x 64 128 x 128 x 64 64 x 64 x 64 64 x 64 x 64
Generator G	ResBlock x 4 Deconv Deconv Conv	3 x 3/1 4 x 4/2 4 x 4/2 7 x 7/1	64 x 64 x 64 128 x 128 x 64 256 x 256 x 64 256 x 256 x 3
	Conv Conv Conv	3 x 3/1 3 x 3/1 3 x 3/1	64 x 64 x 128 64 x 64 x 256 64 x 64 x 512
	Partial nonlocal Resblock x 2	1 x 1/1 3 x 3/1	64 x 64 x 512 64 x 64 x 512
-3*Mapping T	Resblock x 6 Conv Conv Conv	3 x 3/1 3 x 3/1 3 x 3/1 3 x 3/1	64 x 64 x 512 64 x 64 x 256 64 x 64 x 128 64 x 64 x 64

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 are utilized to assess and enhance the quality of the restored images which include LSGAN[12] denoted as LVAE1, GAN.

$$\begin{aligned} L_{VAE_1}(r) = KL(E_{R,X}(z_r|r)||\aleph(0,1)) + \alpha E_{z_r \sim E_{R,X}((z_r|r))} \\ [||G_{R,X}(r_{R\to R}|z_r - r||_1] + L_{VAE_1,GAN}(r) \end{aligned} \tag{6}$$

Table I shows the major modules that we use i.e. Encoder, Generator, and Mapping and explains the layers that they work at. 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.

V. System Design

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.

1) 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.
2) 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.
3) The Corrected Image: This represents the anticipated output produced through the synthetic image. As the synthetic image replicates the degradation process, it offers a valuable reference for determining the desired corrected state. Our trained model has been fine-tuned with these features, achieving a high degree of accuracy.

After all these we already can make out that the real image and the synthetic picture are close to being similar, so we align their entities' space for processing. To align their spaces, we use VAE better known as Variational Autoencoder 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 consists of two main paths:
  The "identity path" directly passes the input to the output without any transformation. The "residual path"

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

Table 2: ResNet-34 architecture.

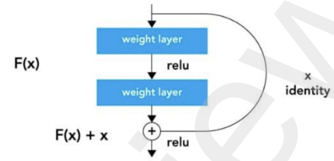

Fig. 9. ResNet-34 Learning Architecture

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 architectures come in various depths, such as ResNet-18, ResNet-34, ResNet-101, and ResNet-152, among others. 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.

$$(y = F(x) + x) \tag{7}$$

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 are stacked with more layers for the creation of a deeper architecture, and hence, deeper learning, in contrast with shallower learning.

ResNet-34 [23] (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.

- We then start to figure out how the transition from the correct image to the corrupted image by mapping them together. We 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.
  - We use a mapping network to translate the corrupted image to the space of the correct image. In other

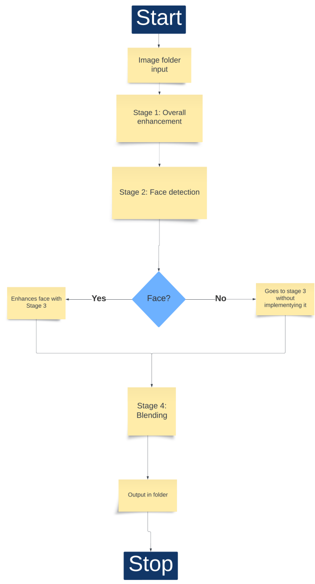

Fig. 10. Flowchart of HexArt AI Image Restoration which shows all stages in execution and decision flow

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.

VI. Results

A. *Implementation*

This section of the paper is aimed to highlight the results that was achieved by applying the algorithm after studying multiple papers and multiple algorithms. Before the proposed system was implemented, all of the algorithms were not only applied but also compared to get the best algorithm and best results.

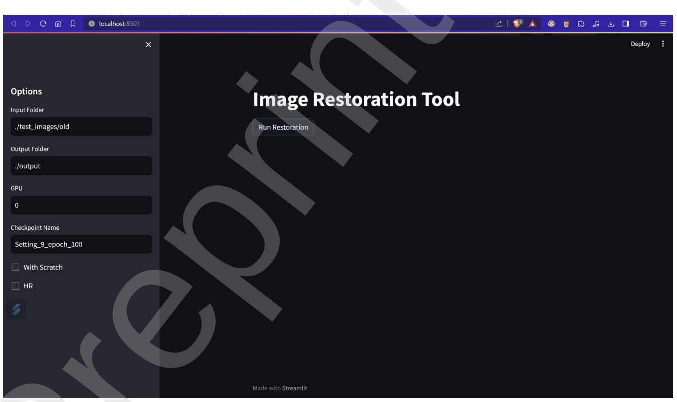

Fig. 11. This is the GUI homepage of HexArt AI Image Restoration which has options like Input folder, Output folder, GPU, Checkpoint Name and check box options for images with scratches and HR

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. Our software is designed to rectify these issues. Within the designated folder, there are two images, and our objective is to apply the correction process to all images contained therein.

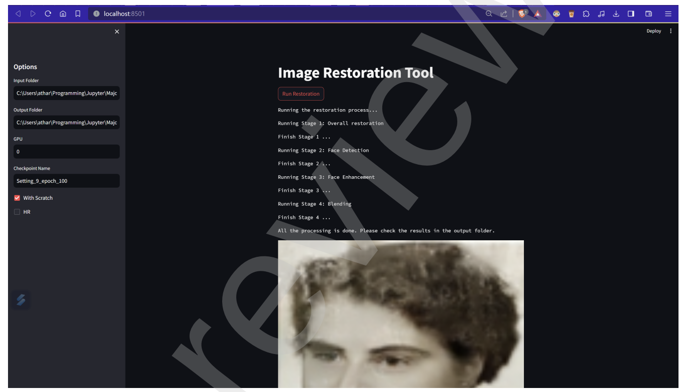

Fig. 12. This is the GUI page of HexArt AI Image Restoration which shows the execution of our system. It shows Stage 1, Stage 2, and Stage 3 execution. It also has the final output image below

Figure 12 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. Additionally, users enjoy the flexibility to set the number of processing epochs, a choice carefully made in favour of 100 epochs known for their capacity 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. The only limitation lies in its heightened GPU computational demands, occasionally leading to skips in processing due to insufficient memory. Figure 12 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. Additionally, users enjoy the flexibility to set the number of processing epochs, a choice carefully made in favour of 100 epochs known for their capacity to deliver highly accurate outcomes.

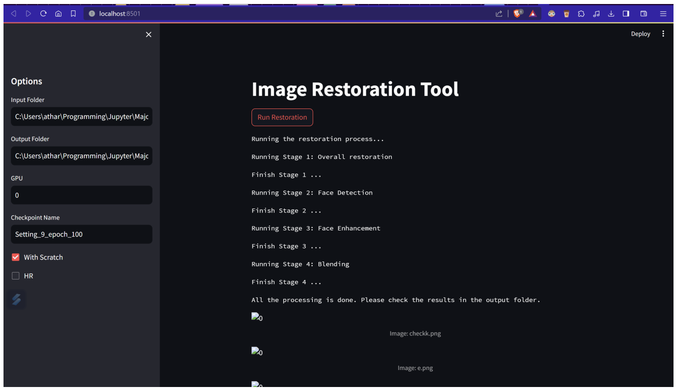

Fig. 13. This image shows the basic execution of HexArt which shows various stages in real-time while computing. In the end, the output images can be seen

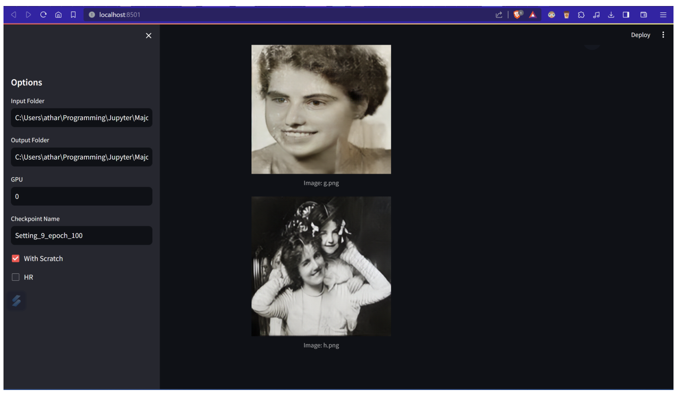

Fig. 14. The figure shows the final output images after computations using various algorithms and passing through various stages

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. The only limitation lies in its heightened GPU computational demands, occasionally leading to skips in processing due to insufficient memory.

B. Result Analysis

The outcomes achieved through our GPU-powered system are truly remarkable. The proposed software adeptly erases a wide array of blemishes, skillfully filling 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

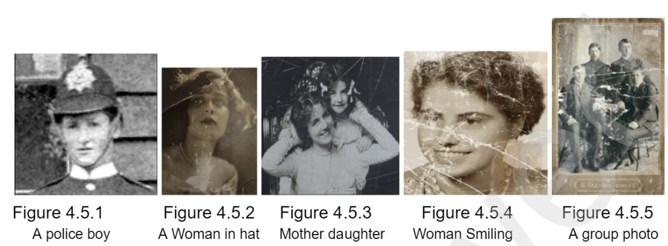

Fig. 15. Input Images to feed into the HexArt Model

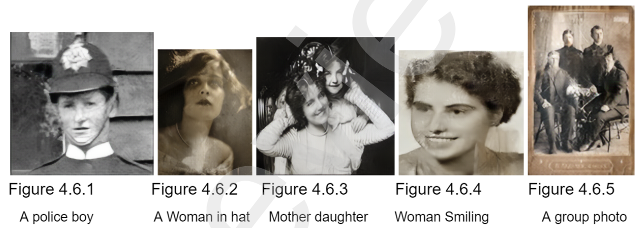

Fig. 16. Output Images generated by HexArt Model

impressive accuracy rate of up to 81%. 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 below as stage 1 Global Restoration and Face enhancement. The ROC curve illustrates the performance of scratch detection in various data settings. 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 artefacts. 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,

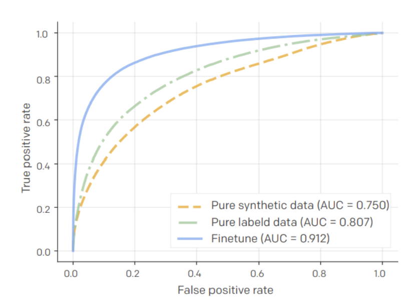

Fig. 17. ROC of HexArt for scratch detection of different data settings

TABLE II
STUDY RESULTS FOR HEXART AI MODEL ON DATASET

Method	PSNR	SSIM	LPIPS	FID
Input	12.82	0.50	0.58	306.81
Attention[13]	**24.1**	**0.71**	0.30	208.2
DIP[14]	22.6	0.44	0.54	194.56
Pix2pix[15]	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**

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.

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 6 are set to α = 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.

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[14].

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[17].

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 and EdgeConnect, it produces certain metrics, which are detailed in the accompanying table II[18].

In a quantitative evaluation of various models using synthetic images from the DIV2K dataset, four metrics were applied, as shown in table II

Conventional metrics such as peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) 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), is calculated based on a noise-free m×n image I and its noisy counterpart K.

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}[I(i,j)-K(i,j)]^2 \tag{8}$$

$$PSNR = 10 \cdot \log_{10}\left(\frac{MAX_I^2}{MSE}\right) \tag{9}$$

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{10}$$

In equation 10 μ_x is the average of x, μ_y is the average of y, σ_x^2 is the variance of x, σ_y^2 is the variance of y, σ_{xy} is the covariance of x and y, c_1 is $(k_1L)^2$ and c_2 is $(k_2L)^2$;

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.

Fréchet 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.

VII. 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 per cent. The proposed system underscores the power of cutting-edge technology in breathing new life into visual relics and bringing the past to vivid, visually pleasing life.

Looking ahead, this proposed system recognizes the potential for further expanding its scope. Video enhancement stands as a compelling frontier, demanding 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.

Our ongoing efforts will focus on further refining the software, addressing algorithmic complexities, and exploring opportunities for interdisciplinary applications in art, history, and media production. As we continue to enhance the capabilities of this system, we anticipate contributing to the broader landscape of digital restoration and expanding its impact across various domains.

References

[1] R. F. Abbas, "Review on some methods used in image restoration," *International multidisciplinary research journal*, vol. 10, pp. 13–16, 2020.

[2] S. Kanrar and S. Maji, "A study on image restoration and analysis," pp. 35–61, 2022.

[3] P. Kumar and V. Gupta, "Restoration of damaged artworks based on a generative adversarial network," *Multimedia Tools and Applications*, pp. 1–19, 2023.

[4] L. Ankita, "Research paper on image restoration using decision based filtering techniques," *International Journal of Engineering Development and Research (IJEDR), ISSN:2321-9939, Volume*, vol. 4, pp. 477–481, 2016.

[5] Z. Liu, "Literature review on image restoration," in *Journal of Physics: Conference Series*, vol. 2386, no. 1. IOP Publishing, 2022, p. 012041.

[6] V. Gupta, N. Sambyal, A. Sharma, and P. Kumar, "Restoration of artwork using deep neural networks," *Evolving Systems*, vol. 12, 06 2021.

[7] L. Yu, J. Guo, and C. Yu, "Research status and development trend of image restoration technology," in *Journal of Physics: Conference Series*, vol. 2303, no. 1. IOP Publishing, 2022, p. 012081.

[8] K. Wang, C. Gou, Y. Duan, L. Yilun, X. Zheng, and F.-Y. Wang, "Generative adversarial networks: Introduction and outlook," vol. 4, pp. 588–598, 09 2017.

[9] I. Goodfellow, "Nips 2016 tutorial: Generative adversarial networks," *arXiv preprint arXiv:1701.00160*, 2016.

[10] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.

[11] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.

[12] X. Mao, Q. Li, H. Xie, R. Y. Lau, Z. Wang, and S. Paul Smolley, "Least squares generative adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2794–2802.

[13] M. Suganuma, X. Liu, and T. Okatani, "Attention-based adaptive selection of operations for image restoration in the presence of unknown combined distortions," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 9039–9048.

[14] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 9446–9454.

[15] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional gans," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8798–8807.

[16] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Bm3d image denoising with shape-adaptive principal component analysis," in *SPARS'09-Signal Processing with Adaptive Sparse Structured Representations*, 2009.

[17] J. Brownlee, "A gentle introduction to pix2pix generative ad-versarial network," *(Last checked on*, 2019.

[18] K. Nazeri, E. Ng, T. Joseph, F. Z. Qureshi, and M. Ebrahimi, "Edgeconnect: Generative image inpainting with adversarial edge learning," *arXiv preprint arXiv:1901.00212*, 2019.