Image Regeneration for Old Damaged Real Picture

Aswathy V[1]
Department of Computer Science and Engineering,
Panimalar Institute of Technology,
Chennai, India

Irene Jerusha P[2]
Department of Computer Science and Engineering
Panimalar Institute Of Technology,
Chennai, India

Nuvvuru Divya[3]
Department of Computer Science and Engineering
Panimalar Institute Of Technology,
Chennai, India

***Abstract*—** Old photos are little things that keep memories alive. Each photograph is worth a thousand memories. However, several photographs have permanent deterioration as a result of unsuitable environmental conditions that lead to photodegradation. Motion blur, noise, and camera focus issues are some forms of degradation. Photo restoration is a technology to repair old photos for those who wish to bring old photos back to life. The two primary techniques utilized in conventional picture restoration to replace the destroyed areas of the image are thermal diffusion and mathematical calculations. It is challenging for this technology to be incorporated into people's daily lives since it can only restore photographs with simple structures and minimal deterioration. Nonetheless, there are still many of improvements being made in this time-honored method of picture restoration. A triplet domain translation network may now be used to stop the combined degradation of old photographs. When the translation to clean images is learned in latent space, the domain gap between historical photos and synthetic images is reduced. When compared to previous methods, this has a generalization problem, which is its main disadvantage. The development of deep learning technologies has boosted the scope of image restoration research. This project's primary goal is to offer an image restoration method based on the Resnet-50, GFP-GAN, and Parsenet algorithms in order to enhance the results of picture restoration of old images and provide additional possibilities for the procedure. The structure, concept, and loss function of an image restoration model built on Resnet-50, GFP-GAN, and Parsenet are shown. In order to create deblurred images and repair old, damaged photographs, GFP-GAN and Parsenet are employed, while Resnet-50 is used to train exceptionally deep neural networks. The historical significance of image restoration techniques is also discussed in this paper. The project concluded that, in the experiment for blur correction, its algorithm was superior to other algorithms. This was done by conducting a comparative experiment to compare this model to other models. The study's model, then, has the greatest effect on picture restoration.

***Keywords*— Generative adversarial network, Resnet-50, Parsenet, Convolutional Neural Network, Image Preprocessing**

I. INTRODUCTION

The world is rapidly evolving in the sphere of technology, to the point that the entire globe is in our hands, with various types of technologies and networks. In this current period of technological advancement, the most recent and fascinating topic of artificial intelligence, which includes a wide range of subsections such as Machine Learning, Deep Learning, and so on. Deep learning and other machine learning approaches encourage computers to acquire knowledge by modeling human behavior.

Several useful applications, including fraud prevention and detection, relationship management systems for clients, computer vision, voice AI, natural language processing, data augmentation, autonomous vehicles, supercomputers, and many more, fall under the broad field of deep learning.

Happy times come and go, but the memories stay forever. A photograph can keep a moment from running away. Image restoration, sometimes referred to as "picture restoration," is the process of restoring damaged photos to their original, undamaged form. Motion blur, low resolution, and noise are just a few of the several ways that the corruption can be observed. Image noise is defined as differences in color and brightness of a representation in comparison to an ideal image of the real scene. Image noise is caused by air disturbances, heat in semiconductor devices, or the random process of incoming photons. Visually, noise adds "dirty" grains of varying intensities to pictures, which in certain situations substantially affects visual enjoyment and image features like edges. Image noise is common as a result of a lack of light or defective camera sensors. When snapping photos, especially at night, the resulting photographs are generally contaminated with unclean pixels, which is known as image noise.When snapping photos, especially at night, the resulting photographs are generally contaminated with unclean pixels, which leads to higher image noise. As they preserve important memories and capture important events in people's life, photos are a good method to relay information. A few decades ago, barely a few paper-based images could fit in picture books and frames.

Many old photographs have oxidised over time and turned yellow; they are simple to treat improperly. Image restoration technology has improved to the point that it is now feasible to digitally restore on computer images that have faded or even been destroyed in order to preserve people's memories of the past. The primary objective of this project is to use new deep learning techniques to recover damaged old photographs that contain a number of important memories. The image processing method that we follow here is both supervised as well as unsupervised since it involves Convolutional Neural Networks [CNN] and Generative Adversarial Networks [GAN]. The specific techniques that are involved are ResNet-50, ParseNet and GFP-GAN. The deep convolutional neural network ResNet-50 architecture was trained using the ImageNet dataset. Semantic segmentation is accomplished using a deep convolutional neural network architecture called ParseNet.

It is designed to perform dense pixel-wise labeling of an image, where each pixel is assigned a class label based on its appearance and context within the image. GFP-GAN creates high-resolution pictures by reducing noise and deterioration from the input image.In the instance of picture regeneration employing a CNN and GFP GAN, the GAN would most likely be utilized to produce high-resolution images that match the content of low-resolution input images. The CNN may be used to analyze the input photos and extract important characteristics that could then be given into the GAN as input, assisting it in producing more accurate results. Overall, combining a CNN and a GFP GAN for picture regeneration would certainly provide high-quality results, but it would also need a substantial quantity of training data and processing resources.

II. REVIEW OF RELATED LITERATURE AND STUDIES

1. *Unpaired Image Super-Resolution using Pseudo-Supervision*

According to Shunta Maeda and Navier, the research produced a generative adversarial network-based unpaired SR method that does not require paired or matched training data. A pseudo-paired SR network and an unpaired kernel/noise-correction network make up the network. The rectification network, which also alters the kernel, cleans up the input LR image before the SR network upscales it. The SR network learns a paired mapping from the pseudo-clean LR image to the inputted HR image after the rectification network creates a pseudo-clean LR picture from the inputted HR image during the training phase. The paper further stated that the proposed method can be used with a variety of datasets, hyperparameter adjustment is required in each instance to achieve the best performance. Future work will involve making the network more resistant to hyperparameters.

2. *Automatic Damage Recovery of Old Photos Based on Convolutional Neural Network*

Every photograph has value in life. Photographs represent most lost memories. In this paper, Tien-Ying Kuo, Yu-Jen Wei, Ming-Jui Lee, Tzu-Hao Lin claimed to have used manual procedures and to have replaced the feature extraction and reconstruction layers with 3*3 CNN layers. They also included an additional three-layer CNN block to check whether the natural features had been extracted from the pictures. To build a model they have considered 190 photos, of which 100 are unbroken and 90 are damaged. The model employs various 100 different damaged photos that generate 5000 different pieces of training data to determine where and what type of damages are present in a specific image so that a model can forecast damages in future input photos. However, the execution of the images' opacity is flawed. The major disadvantage is that it is not justified to use PSNR (Peak Signal-to-Noise Ratio) or SSIM (structural index similarity) to determine the image difference after the damaged area has been fully repaired because the repair result may differ significantly from the original image.

3. *High-Fidelity Pluralistic Image Completion with Transformers*

According to Ziyu Wan, Jingbo Zhang, Dongdong Chen, and Jing Liao, this work offered the two best methods for picture completion: texture replenishment using CNN and earlier restoration with transformer. The latter type of CNN enhances the local texture features of coarse priors produced by high-resolution masked photos; the former restores pluralistic coherent structures in addition to certain coarse textures. The proposed strategy significantly outperforms state-of-the-art approaches. Future issues that might be studied are sampling strategy, inference efficiency, and architectural design.

4. *Face Video Deblurring using 3D Facial Priors*

Wenqi Ren, Jiaolong Yang, Senyou Deng, David Wipf, Xiaochun Cao, and Xin Tong introduced a face video deblurring network based on 3D facial priors in this paper, which outperformed current face deblurring approaches that only analyzed single frames and did not account for facial structure and identification information. The model was divided into two parts: a face video deblurring subnetwork and a 3D face reconstruction and rendering branch for predicting 3D priors of significant facial features and identification information. To deblur the input face footage, they used both picture intensity and high-level identification information generated from the rebuilt 3D faces.

5. *Multiple Cycle-in-Cycle Generative Adversarial Networks for Unsupervised Image Super-Resolution*

In a more systematic way where the down sampling process is unknown and the LR input is impaired by noises and blurring, Yongbing Zhang, Siyuan Liu, Chao Dong, Xinfeng Zhang, and Yuan Yuan proposed a multiple Cycle-in-Cycle network structure to address the single image super-resolution problem. The framework consists of a number of CycleGAN-like cycle networks, with the latter cycle encircling the former. Once the initial cycle returns the noisy and blurry LR input to a clean bicubic-downsampled LR space, a new cycle is formed by inserting a well-trained *2 EDSR model to further denoise, deblur, and super-resolve the recovered LR picture.

6. *Face Hallucination Using Cascaded Super-Resolution and Identity Priors*

This paper Klemen Grm, Walter J. Scheirer and Vitomir Štruc tackled the problem of how to deal with the hallucinating problem that is faced in high-definition pictures. This system was evaluated on the Labeled Faces in the Wild (LFW), Helen, and CelebA datasets and reported better performance. In terms of future work, they saw the possibility of adapting their model to other modalities, e.g., video sequences, via recurrent attention models.

III. METHODOLOGIES

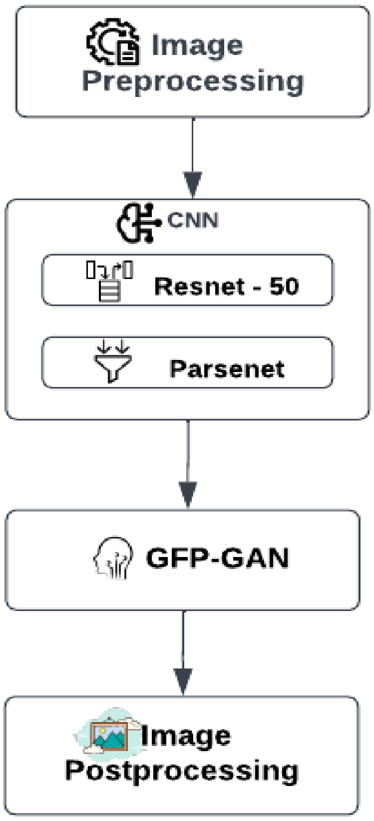

Fig.1. Development Stages

3.1 IMAGE PREPROCESSING:

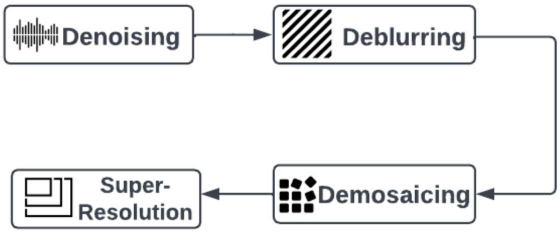

Fig.2. Image Pre-Processing

"Image pre-processing" describes actions taken on images at the most fundamental abstraction level. By minimizing undesirable distortions or boosting particular visual features that are crucial for later processing and analysis activities, pre-processing seeks to enhance the picture data.

Image Preprocessing can be done by the following techniques such as Image enhancement, image restoration, color image processing, object detection, and Image segmentation. For image restoration, the various techniques in image preprocessing are used to obtain correct input to train the model or to proceed with further steps of regeneration of an image. Image preprocessing also includes noise reduction and impulsive noise removal occurred by using large filters.

3.2 FEATURE EXTRACTION:

This module is in charge of extracting the important features and material from the damaged image, which will then be fed into the generator network. The following steps must be taken in order to extract features

- ❖ AutoEncoders
  Convolutional autoencoders learn to break down picture data into latent and spatial components.
- ❖ Transformers
  Transformers record global interactions across contexts and visual challenges through a self-attention technique.
- ❖ Deconvolutional Layers
  By using end-to-end mapping, deconvolutional layers sample the feature maps and restore picture information.
- ❖ Convolutional Layers
  Convolutional layers retain image content inference while removing corruptions.

3.3 CONVOLUTION NEURAL NETWORKS [CNN]:

Convolutional Neural Networks (CNN) are a key component of one particular technique that has seen significant breakthroughs in Deep Learning throughout time. Through the use of pertinent filters, a convolution network may effectively capture the spatial and temporal dependencies in a picture. The filters will shift to the right with a set Stride Value after a photo has been analyzed throughout its entire width to find its natural features.

The main objective of convolution is to extract the high-level properties of the input picture, such as its edges. Convolutional Nets don't have to have just one convolutional layer. The initial Convolutional Layer usually captures low-level features such as edges, colour, gradient direction, and others. The design adapts to the High-Level features as the number of layers increases, giving us a network that interprets the dataset's images holistically in a way that is comparable to how we do.

3.4 GFP-GAN:

A generator and a discriminator are the two primary parts of a GAN. As the discriminator works to separate the generated samples from actual samples, the generator creates new samples, such as new images or new texts.

GFP-GANs have the advantage of allowing quality control over the generated samples, which makes them useful for producing high-resolution photos, text, and other types of data.

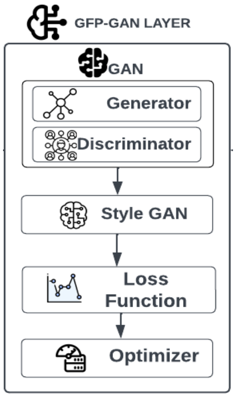

Fig.3. Generative Facial Prior

GFP-GAN operates using a variety of modules, including Degradation elimination, Generative facial prior, Channel-Split Feature Transform, and Model Goals. The suggested architecture begins with a U-Net degradation removal module that seeks to eliminate degradations and also helps in extracting latent features for mapping the picture to the closest StyleGAN-2 latent code,as well as a collection of spatial features with multiple resolutions for adjusting StyleGAN-2 intermediate feature mappings. The bottleneck layer provides the latent properties of an image, while the decoder portion of the U-Net provides the spatial features of an image.

The bottleneck layer of the U-NET latent features are converted to style vectors by a number of MLPs, which creates intermediate convolutional features that can be further manipulated by the spatial information. The affine transform parameters are used to scale and shift the feature maps in the generator, and the spatial features are utilized to forecast these parameters. The numerous losses obtained from a number of small local discriminators trained on patches of the eyes, lips, and other body parts are finally eliminated.

3.5 RESNET-50:

Convolutional neural network ResNet-50 has a total of 50 layers. A network that has previously been trained on more than a million pictures may be loaded from the ImageNet database. The trained network can categorise images into several object types depending on the user's demands.

ResNet was the first to introduce the skip connection. The problem of vanishing gradient is lessened by skip connections. Each of the five stages in the ResNet-50 model has a convolution and an identity block. Each identity block and convolution block includes three convolution layers. Millions of parameters may be taught with the ResNet-50.A ResNet, on the other hand, is simpler and contains fewer filters than a VGGNet[visual geometry group].

3.6 PARSENET:

ParseNet is a fully convolutional neural network that takes an input image and outputs a pixel-wise segmentation map. It is trained end-to-end using a combination of supervised and unsupervised learning, where the supervised part involves minimizing a cross-entropy loss between the predicted segmentation map and the ground truth, and the unsupervised part involves minimizing a boundary loss that encourages the predicted boundaries between different segments to match the ground truth.

"ParseNet" [parametric surface fitting network architecture] is trained on a vast dataset of artificially generated 3D shapes, ParSeNet collects high-level semantic priors for shape decomposition. The end-to-end trainable ParSeNet technique decomposes a 3D point cloud into parametric surface patches, including B-spline patches and basic geometric primitives. Using many patches instead of a single continuous patch has the advantage of allowing for the generation of a far greater variety of geometric properties and surface topologies.Overall, ParseNet achieved state-of-the-art performance on several benchmark datasets for image segmentation, demonstrating its effectiveness for this task.

3.7 IMAGE POSTPROCESSING:

The image postprocessing step is an important procedure to perform an enhanced visualization of the results from the previously processed image. Most image enhancement techniques are used to improve the image's aesthetic appeal. This can be achieved by boosting contrast, improving brightness, and lowering noise and lack of sharpness. This step mainly focuses on increasing the resolution of the image. Super-Resolution is a technique that is used to increase the resolution of an image. It involves generating high-resolution images from low-resolution ones by using algorithms such as deep learning. The post-processing procedure is repeated until the image's resolution is satisfactory for proper development.

3.7.1 LIBRARIES USED:

- Open CV
- OS
- NumPy
- Shutil

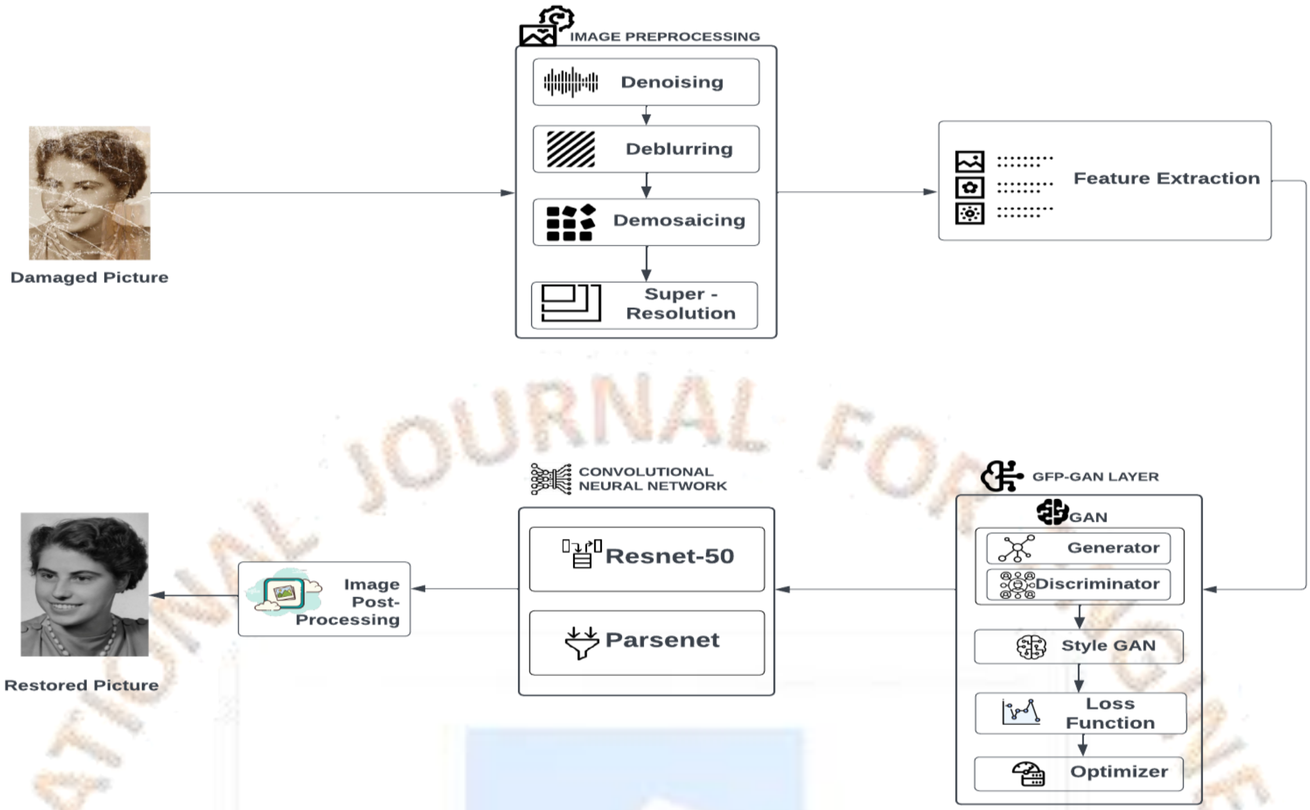

Fig.4. Overall Flow of the Project

IV. RESULT

The main technology for repairing ancient photographs and images is image restoration using GFP-GAN. Because damaged and unclear historical photographs are frequently available, it is required to provide an adequate restoration technique to recreate the scene of the old pictures. The introduction and explanation of the generic convolutional neural network and the generative confrontation network follows. This project develops a novel picture restoration technique based on the Resnet-50, GFP-GAN, and ParseNet algorithms, as well as a loss function. The area of photo restoration would benefit more from the model utilized in this study.

V. CONCLUSION & FUTURE WORK

The image that was recreated using the above methodologies was able to reclaim its distinct quality and structure as well as preserve earlier, more valued memories. This method of image regeneration produced the greatest results since it improved the image's quality and color to produce a photo with higher resolution. This strategic approach delivers better predicted outcomes and is more accurate when compared to classic picture regeneration approaches. Although the restoration effect is satisfactory from an experimental standpoint, there is still room for development. Further work on this might result in a complete software application with a high-level graphical user interface that could be more easily accessed by end users and offer greater services to society.

REFERENCES

[1] Liu, G., Li, X. and Wei, J., 2021. Large-area damage image restoration algorithm based on generative adversarial network. Neural Computing and Applications, 33, pp.4651-4661.

[2] Lin, K. and Cai, Q., 2022, August. Extended StyleGAN Encoder for Image Restoration. In 2022 26th International Conference on Pattern Recognition (ICPR) (pp. 2157-2164). IEEE.

[3] Demoment, G., 1989. Image reconstruction and restoration: Overview of common estimation structures and problems. IEEE Transactions on Acoustics, Speech, and Signal Processing, 37(12), pp.2024-2036.

[4] Rani, S., Jindal, S. and Kaur, B., 2016. A brief review on image restoration techniques. International Journal of Computer Applications, 150(12), pp.30-33.

[5] Maurya, A. and Tiwari, R., 2014. A novel method of image restoration by using different types of filtering techniques. International Journal of Engineering Science and Innovative Technology (IJESIT), 3(4).

[6] Maeda, S., 2020. Unpaired image super-resolution using pseudo-supervision. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 291-300).

[7] Wan, Z., Zhang, J., Chen, D. and Liao, J., 2021. High-fidelity pluralistic image completion with transformers. In Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 4692-4701).

[8] Ren, W., Yang, J., Deng, S., Wipf, D., Cao, X. and Tong, X., 2019. Face video deblurring using 3D facial priors. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 9388-9397).

[9] Zhang, Y., Liu, S., Dong, C., Zhang, X. and Yuan, Y., 2019. Multiple cycle-in-cycle generative adversarial networks for unsupervised image super-resolution. *IEEE transactions on Image Processing*, *29*, pp.1101-1112.

[10] Bulat, A. and Tzimiropoulos, G., 2018. Super-fan: Integrated facial landmark localization and super-resolution of real-world low-resolution faces in arbitrary poses with gans. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 109-117).

[11] Zhang, Y., Liu, S., Dong, C., Zhang, X. and Yuan, Y., 2019. Multiple cycle-in-cycle generative adversarial networks for unsupervised image super-resolution. *IEEE transactions on Image Processing*, *29*, pp.1101-1112.

[12] Computing, W.C.A.M., 2023. Retracted: An Old Photo Image Restoration Processing Based on Deep Neural Network Structure.

[13] Sendik, O., Cohen-Or, D. and Lischinski, D., 2020. Crossnet: Latent cross-consistency for unpaired image translation. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 3043-3051).

[14] Richardson, E., Alaluf, Y., Patashnik, O., Nitzan, Y., Azar, Y., Shapiro, S. and Cohen-Or, D., 2021. Encoding in style: a stylegan encoder for image-to-image translation. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 2287-2296).

[15] Tov, O., Alaluf, Y., Nitzan, Y., Patashnik, O. and Cohen-Or, D., 2021. Designing an encoder for stylegan image manipulation. *ACM Transactions on Graphics (TOG)*, *40*(4), pp.1-14.

[16] Wan, Z., Zhang, B., Chen, D., Zhang, P., Wen, F. and Liao, J., 2022. Old photo restoration via deep latent space translation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *45*(2), pp.2071-2087.

[17] Safin, A., Kan, M., Drobyshev, N., Voynov, O., Artemov, A., Filippov, A., Zorin, D. and Burnaev, E., 2021. Towards unpaired depth enhancement and super-resolution in the wild. *arXiv preprint arXiv:2105.12038*.

[18] https://www.hindawi.com/journals/wcmc/2022/7415342/

[19]https://www.researchgate.net/publication/359419452_An_Old_Photo_Image_Restoration_Processing_Based_on_Deep_Neural_NetworkStructure