2023 6th International Conference on Information Systems and Computer Networks (ISCON)
GLA University, Mathura, India. Mar 3-4, 2023

A Comprehensive Study of Image Inpainting Techniques with Algorithmic approach

Tammineni Shanmukhaprasanthi
AUCE, Andhra University
Visakhapatnam
prashanthitammineni.rs@andhrauniversity.edu.in

Swarajya Madhuri Rayavarapu
AUCE, Andhra University
Visakhapatnam, India
madhurirayavarapu.rs@andhrauniversity.edu.in

Yenneti Laxmi Lavanya
AUCE, Andhra University
Visakhapatnam
lavanya9234@gmail.com

Gottapu Sasibhusana Rao
AUCE, Andhra University
Visakhapatnam
sasigps@gmail.com

2023 6th International Conference on Information Systems and Computer Networks (ISCON) | 979-8-3503-4696-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISCON57294.2023.10112205

***Abstract*— The process of image inpainting is indeed a method that allows you to fill in particular regions of an image using novel content by estimating from surrounding pixels, from other images, or from various sources. The term "image inpainting" refers to a wide range of techniques used to repair visual media, from restoring old photographs and films to erasing undesirable text or objects. When only individual components of a picture are missing, traditional methods can produce high-quality solutions, but they are unable to detect new elements that were not included in the picture. Improvements in the quality of image inpainting have been achieved by the use of Deep Learning techniques, which have made it possible to generate appropriate hole filling and separate objects that aren't included in a real image. However, there is a great deal of room for development in this regard, especially in the areas of adapting to different image sizes, using free-form masks, making high-resolution textures, using fewer computational resources, and shortening the training process. The comprehensive study carried on various image inpainting techniques and their applications will be useful to the new researchers to explore the possibility of developing new efficient algorithms for further improved image inpainting techniques truer to life.**

Keywords—Image inpainting; GAN(Generative Adversarial Network); Diffusion based inpainting; Convolution based inpainting; Exemplar based inpainting

I. Introduction

As early as the Renaissance, inpainting was used to restore damaged painted pictures caused by ageing, scratches. For the purpose of restoring priceless paintings and protecting their cultural heritage, physical inpainting is indeed a time-consuming process that is performed by hand by trained artists using whatever methods end up being the most successful in restoring the paintings to as close to their original state as possible. Following the introduction of the photography and also the videos, the requirement to reconstruct media developed, and digital inpainting had been formed. There are various image inpainting applications i.e., restoration of old photographs, removing the scratch and object removal in image editing, text removal, subtitles, logos, lost block recovery in transmission of images as shown in Fig.1. The method of digital inpainting focuses on the use of advanced algorithms to recreate digital picture data. development of special effects for cinema and television, robot vision, etc. The basic block diagram of inpainting is shwn in fig.2.

(a) Image restoration

(b)Text removal

(c)Scratch removal

(d) Object removal

Fig. 1. Various image inpainting applications (a) Image Restoration [15] (b)Text Removal [14] (c)Scratch Removal [15] (d) Object Removal [13].

This study describes and analyses several inpainting methods. Section 2 provides a Literature review of image inpainting, organized by category and chronology. In Sections 3 and 4, some of the inpainting techniques and quantitative measures from the examined papers are presented. Section 5 concludes with the findings.

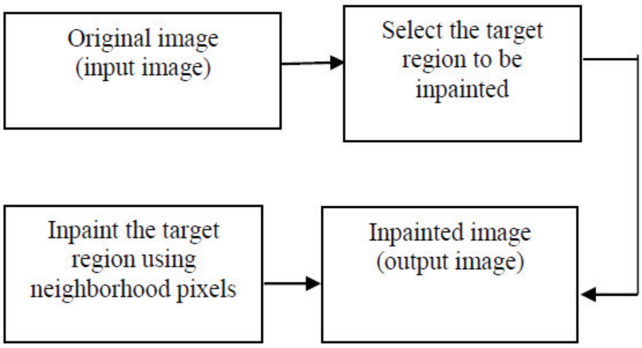

Fig. 2. Basic building blocks of image inpainting

II. LITERATURE REVIEW

Bertalmio et al. [1] used partial differential equations to automatically fill up picture gaps in the year 2000. Using Isophotes constant light intensity curves are employed. Total Variational (TV) Inpainting had been invented by the Chan and Shen [2] in 2001. Criminisi et al. suggested an effective patch-based and diffusion-based method later that year.

Cheng introduced a generalized priority function to improve picture reconstruction [3]. In 2009, Barnes et al. [4] improved Wexler and Simakov methods using fast-approximate nearest neighbour patch search. Patch Match determines patch correspondences using random sampling and nearest-neighbor field (NNF).

Oliveira et al. [5] developed a fast convolution-based image inpainting method in 2001. This approach repeatedly makes convolutions with the target region and the diffusion kernel. Hadhoud et al. reduced inpainting time and picture quality by changing the convolution step.

Deep Learning (DL) by LeCun [6] enabled supervised image classification inpainting. Each image is labelled, and a convolutional neural network (CNN) trains to match labels to pictures. Xie et al. presented a new Denoising Auto-encoder (DA) training method in 2012 [7] it denoises and automatically inpaints images on same frame. Two-layer neural network DA recovers input from noise.

Eigen's 2013 three-layer CNN model removed raindrops and soil. CNN's ability to blindly inpaint graphics when the missing area's shape is uncertain, it is beneficial in the context of everyday life situations. A robust deconvolution model [8] was developed by Xu et al. in 2014. Deconvolution restores convolution damaged images. This deconvolution convolutional neural network (DCNN) uses separable kernels for robust artefact free deconvolution and successfully works in the non-blind deconvolution of images.

The author Kohler had valuated mask shape and performance. It was observed that inpainting can be improved by inclusion of the mask shape in the input layer. Both sparse auto-encoders (SARs) and convolutional neural networks (CNNs) were found by Ren et al. (2015) to have inherent translation invariant operators. When DL requires interpolation with a translation variable, this drastically slows down the process (TVI). Many DL methods still have blurry graphics and cannot inpaint complicated scenes because of a lack of semantic picture comprehension.

With Goodfellow's work as basis, Pathak et al. [9] created Context Encoders (CE) in 2016 that work uses the conditional (GAN) Generative Adversarial Networks. The encoder-decoder networks which are CE's that can predict missing data and exploit adversarial loss during training to improve accuracy.

Zeng et al. presented U-Net-based (PEN-Net) Pyramid-context Encoder Network in 2019. Picture inpainting may achieve visual and semantic consistency by pyramidally transferring attention from deep to shallow regions to replace missing content Model for semantic inpainting presented by Sagong et al. [10] is called PEPSI (acronym termed as the Parallel Extended Decoder Path for Semantic Inpainting). The coarse and inpainting routes of the single shared encoding network are used to enhance Yu's original model, which previously relied on a two-stage feature encoding approach.

In 2021, Liu et al. [11] introduced PD-GAN for image inpainting using vanilla GAN. PD-GAN decodes a random noise vector instead of sending input pictures to CNN for content synthesis. Suin et al. developed a distillation-directed training approach to lead different layers of a network to improved optimums. Their concept included the Inpainting Network (IN) and the Auxiliary Network (AN). Masked and unmasked photos feed IN and AN networks. Jam et al. developed reverse masking image inpainting.

Yu et al. [12] came up with a model-based interactive picture-painting method in 2022. The authors present ESAP - External Spatial Attention Module that uses a lightweight external attention approach to enhance inpainting by integrating encoded and contextual picture components with actual spatial information.

A properly built stack of context reasoning layers makes up the generator's Aggregate Contextual Transformations (AOT) block. Split-transform-merge is used in AOT blocks. By integrating the results of several changes, rich pattern of interest is recorded, improving reasoning about the missing sections of the picture.

III. IMAGE INPAINTING TECHNIQUES

There are two distinct kinds of image inpainting approaches: conventional and deep learning-based approaches.

A. Conventional approaches

There are three conventional approaches for image inpainting based on diffusion, exemplar, and convolution.

1) Diffusion-based techniques:

Bertalmio et al. created this algorithm first. PDE necessitates the mathematical representation of pictures, such as a two-dimensional matrix containing integers encoding the grayscale value. Using this user-provided mask and the algorithm, the picture will be taken as a mixture of the three channels RED, GREEN, and BLUE. For each of the three channels, the required inpainting regions are filled in by transmitting information from outside the masked zone and along the border lines [16]. This method is effective as a result of its Isophote-driven method; grayscale values in the matrix that provides the most of the information and utilise it to fill the picture in less time. PDE-based picture inpainting has the problem of working on every pixel, resulting in a fuzzy output. This technique is not ideal for curved edges because edges are extended in a linear way. The fig.3 gives a detailed architecture of diffusion based inpainting technique.

Algorithm for Diffusion based inpainting method

Step1: Select the destroyed part in the image
Step2: Select the width of inpainting region
Step3: Select the range of the inpainting region with respect to rows and columns
Step4: Calculate R, G and B values
Step5: Apply filtering to the R, G and B components independently
Step6: Diffuse image pixel with R, G and B
Step7: In the inpainting region, repeat the above steps for all pixels

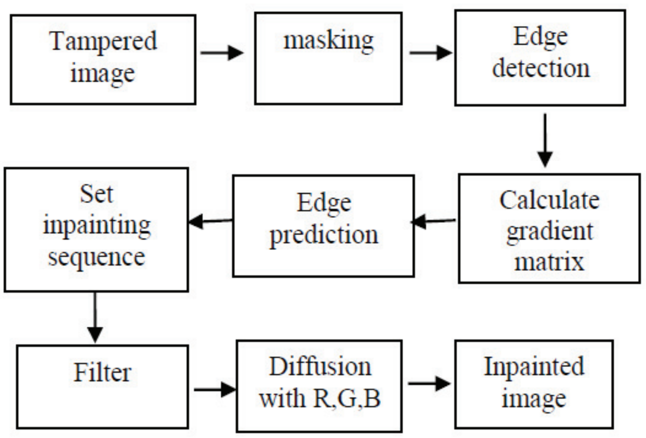

Fig. 3. Architecture of diffusion based image inpainting technique

2) Exemplar-based techniques:

There is a term in Exemplar Inpainting that is referred to as "text redness." The pixel to be filled in the target area is determined by the neighborhood "text redness" level [17]. Strong linear structures are weakened by noise therefore the utility of the additional computation is decreased. Exemplar-based texture synthesis is also used to propagate extended linear picture structures; hence, a separate synthesis process for isophotes is not required. This method is sufficient for spreading texture and structure into the patched area. In this algorithm, the best-matching patches from known places are picked using certain metrics and then put into the missing area. Using this strategy, large objective areas are reconstructed. The detailed description of exemplar based technique is given in fig.4.

Algorithm for exemplar based inpainting method

Step1: Read the image
Step2: Target image is obtained by excluding the object of interest from the given image
Step3: Create the target region of the excluded object and source region as the rest
Step4: Create patches with border pixels as their centres on the outside of the desired area.
Step5: Compute the priority function which is the product of confidence term and data term for identifying highest priority of target region.
Step6: Swap out the fix for an example and revise the reliability numbers.
Step7: To ensure that the target area has no residual pixels, till then repeat the step 4.
Step8: Inpainted image is the final output.

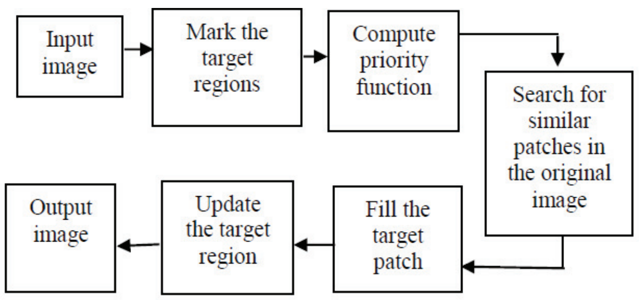

Fig. 4. Architecture of exemplar based image inpainting technique

TABLE I. A QUALITATIVE REPORT ON DIFFERENT IMAGE INPAINTING TECHNIQUES

Features	Diffusion based model	Exemplar based model	Convolution based model	Deep learning based model
Advantages	It uses isophote driven approach and preserves all the structure information	More efficient for larger target regions	It requires less time for inpainting and produces the output without blur because no repetition in convolution	Very efficient model and produces constant output
Disadvantages	It gives a blur effect for larger regions and take a long time. It works for pixel level only.	It requires high calculation time	When large objects are removed in natural images, poor results are produced	It requires complex training and non-convergence for complex and large data
Size of missing region	Small	Large	Large	Large
Data model	Structure	Structure and texture	Texture and structure	Structure and texture
No. of images required	1	1	1	Database of the image
Computational time	High	Less	Less	Very high

3) Convolution based techniques:

Image inpainting based on convolution algorithms are well known for fast results, however in some conditions, adequate results are not obtained in sharp details like edges. This method works as follows the gradient of the image is used to calculate the coefficient of the mask. With the help of a suitable kernel, we convolve the area around the broken pixels in the picture we're trying to inpaint. The algorithm benefits from being quick, iterative, and easy to implement. Because the central frequency of the scattering masks is assumed to be 0 at the beginning of this procedure because the relevant pixels in the original photograph is unknown [18]. While the oliveira approach requires 100 iterations to remove big items on symmetric backgrounds, the fast convolution-based (CN) digital inpainting algorithm only needs 1. However, it is difficult or impossible to fix huge items when working with a real image. The convolution based method architecture is shown in fig.5.

Algorithm for Convolution based inpainting method

Step1: Input the distorted image
Step2: Select the inpainting region
Step3: Initialize the inpainting region by clear its color
Step4: Create the masked region used image segmentation
Step5: Convolve masked region with kernel
Step6: Output is restored image

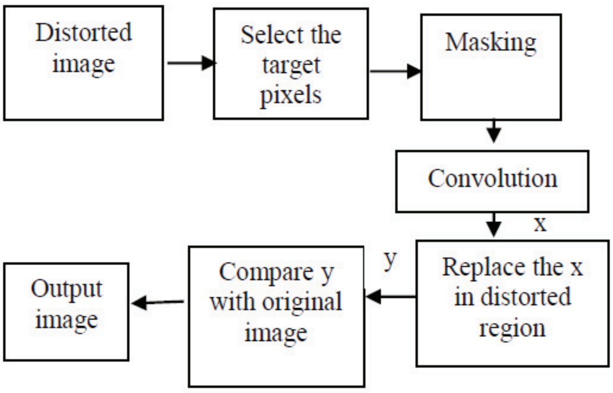

Fig. 5. Architecture of convolution based image inpainting technique

B. Deep Learning based approaches

Recent developments in deep learning, in particular the use of GAN Networks for object recognition, have enabled the creation of a tool that had previously been unavailable for picture inpainting. The Generative Adversarial Network, sometimes known as GAN, is a method that is frequently used in image inpainting. Within the area of GANs, the Deep Convolutional Generative Adversarial Network (DCGAN) in particular is very popular. A DCGAN is made up of two components: a generator and a discriminator. While it is the job of the generator to produce an image that has been fixed, the discriminator's job is to try to tell the difference between the restored image and the reference image. Up sample the input image by the generator, which uses several decoders. The majority of decoders include a batch normalization layer, a fractional-strided convolutional layer, and a ReLU activation in their construction. However, the final decoder is consisting of a Tanh activation a single transposed convolutional layer. On the other hand, the discriminator down samples both the created and the original images. The first down sampler is made up of a strided convolutional layer and a Leaky ReLU activation, whereas the other down samplers are composed of strided convolutional layers, Leaky ReLU activations and batch normalisation layers.

DCGANs place a greater emphasis on the utilization of convolutional layers than traditional GANs do. Because of this, DCGANs make it possible for the discriminator and the generator to learn their own spatial up sampling and down sampling. The block diagram of deep learning based inpainting is shown in the fig.6

Algorithm for GAN based inpainting method

Step1: Read the vector random values of image as an input
Step2: Generator is connected to the input noise
Step3: Convert the random vectors size using a project and reshape operation
Step4: Upscale the resulting arrays to required size array using a series of convolution layers with batch normalization and Rectified Linear unit layer
Step5: Transposed convolution vector is given to tanh layer
Step6: Real image and generated images given to discriminator network
Step7: Discriminator predicts whether it is real image or not which is taken as output

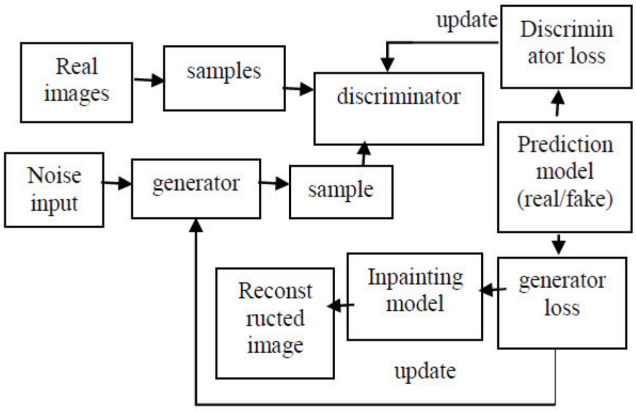

Fig. 6. Architecture of GAN based image inpainting technique

The comparative analysis of various image inpainting techniques is summarized in Table.1

IV. EVALUATION METRICS

The performance measures used are PSNR and MSE.

a) The term "mean square error" (MSE) refers to the average squared deviation that exists between the values that are estimated and the values that are really present. MSE may be determined by the use of a formula,

$$\text{MSE} = \frac{\sum_{M,N} [I_1(m,n) - I_2(m,n)]^2}{M*N} \quad (1)$$

Where n and m represent the total number of columns and rows respectively.

b) The peak signal-to-noise ratio (PSNR) is the ratio of the greatest possible signal power to the maximum possible noise power that degrades the accuracy of the signal representation. The formula for determining PSNR is as follows.

$$\text{PSNR} = 10\log_{10}\left(\frac{R^2}{MSE}\right) \quad (2)$$

R represents the amount of variation that may be found in the supplied picture data type.

Using these formulae performance can be measured, if the PSNR value is high then the errors in that is very less and vice versa.

V. CONCLUSION

This paper presents a comprehensive survey of image inpainting algorithms. Image inpainting is a necessary task for computer vision applications due to the requirement for modification of image data such as image repair and restoration. Both conventional and deep learning approaches have been presented as different categories of approaches. It is observed that no single technique can inpaint all sorts of image distortion, but that deep learning produces very promising image inpainting outcomes for each of the analyzed cases. Nevertheless, the issue with deep learning is its computational complexity and memory requirements. The "SSIM" and "PSNR" are the fundamental quantitative measures that can be used in both global and regional domains to determine the picture quality. These metrics are used in quality testing of an image. Future research must focus on presenting algorithms that are less complex and capable of being implemented with limited memory resources. In addition, extensive research is being conducted to use image inpainting technique as a key tool for video restoration.

REFERENCES

[1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in Proceedings of the 27th annual conference on Computer graphics and interactive techniques, 2000, pp. 417–424.

[2] T. Chan, "Local inpainting models and tv inpainting," SIAM J. Appl. Math., vol. 62, no. 3, pp. 1019–1043, 2001.

[3] W.-H. Cheng, C.-W. Hsieh, S.-K. Lin, C.-W. Wang, and J.-L. Wu, "Robust algorithm for exemplar-based image inpainting," in Proceedings of International Conference on Computer Graphics, Imaging and Visualization, 2005, pp. 64–69.

[4] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman, "Patchmatch: A randomized correspondence algorithm for structural image editing," ACM Trans. Graph., vol. 28, no. 3, p. 24, 2009.

[5] M. M. O. B. B. Richard and M. Y.-S. Chang, "Fast digital image inpainting," in Appeared in the Proceedings of the International Conference on Visualization, Imaging and Image Processing (VIIP 2001), Marbella, Spain, 2001, pp. 106–107.

[6] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," nature, vol. 521, no. 7553, pp. 436–444, 2015.

[7] J. Xie, L. Xu, and E. Chen, "Image denoising and inpainting with deep neural networks," in Advances in neural information processing systems, 2012, pp. 341–349.

[8] L. Xu, J. S. Ren, C. Liu, and J. Jia, "Deep convolutional neural network for image deconvolution," in Advances in neural information processing systems, 2014, pp. 1790–1798.

[9] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in Advances in neural information processing systems, 2014, pp. 2672–2680.

[10] Y. Zeng, J. Fu, H. Chao, and B. Guo, "Learning pyramid-context encoder network for high-quality image inpainting," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2019, pp. 1486–1494.

[11] H. Liu, Z. Wan, W. Huang, Y. Song, X. Han, and J. Liao, "Pd-gan: Probabilistic diverse gan for image inpainting," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 9371–9381.

[12] W. Yu, J. Du, R. Liu, Y. Li et al., "Interactive image inpainting using semantic guidance," arXiv preprint arXiv:2201.10753, 2022.

[13] https://www.cs.toronto.edu/~mangas/teaching/320/assignments/a2

[14] https://www.researchgate.net/figure/Inpainting-for-text-removal-a-Image-with-text-b-Inpainting-using-ICA-learned-basis_fig12_267074143

[15] https://en.wikipedia.org/wiki/Inpainting

[16] Patel, Hiral Nileshbhai. "A Survey on Different techniques for Image Inpainting." International Research Journal of Engineering and Technology (IRJET) (2016).

[17] Criminisi, Antonio, Patrick Pérez, and Kentaro Toyama. "Region filling and object removal by exemplar-based image inpainting." IEEE Transactions on image processing 13, no. 9 (2004): 1200-1212.

[18] Vadhel, Bhavini, and Bhailal Limbasiya. "Survey on Different Techniques for Image Inpainting." International Research Journal of Engineering and Technology (2016).