Deep Learning Inpainting Model on Digital and Medical Images-A Review

Jennyfer Susan
Department of Computer Science,
Avinashilingam Institute for Home Science and Higher Education for Women, India
19phcsf008@avinuty.ac.in

Parthasarathy Subashini
Department of Computer Science,
Avinashilingam Institute for Home Science and Higher Education for Women, India
Subashini_cs@avinuty.ac.in

Abstract: *Image inpainting is a method to restore the missing pixels on damaged images. Initially, the traditional inpainting method uses the statistics of the surrounding pixels to find the missing pixels. It sometimes fails to read the hidden information to attain plausible imagery. The deep learning inpainting methods are introduced to overcome these challenges. A deep neural network learns the semantic priors and hidden representation pixels in an end-to-end fashion in the digital and medical. This paper discusses the following: 1) The difference between the supervised and the unsupervised deep learning inpainting algorithm used in medical and digital images. 2) Discusses the merits and demerits of each deep learning inpainting model. 3) Discusses the challenges and solution for the deep learning inpainting model. 4) Discusses each model's quantitative and qualitative analysis in the digital and other medical images.*

Keywords: *Image inpainting, image restoration, deep learning, supervised method, unsupervised method.*

Received February 18, 2023; accepted May 9, 2023
https://doi.org/10.34028/iajit/20/6/9

1. Introduction

Image inpainiting is "*image completion*" or filling in the recovery of an image's missing or damaged pixel. It plays a vital role in computer vision and image processing research areas like restoring old or damaged images and documents, removing the desired object, and filling them with neighboring pixels [6]. It is usually modeled as in the Equation (1), where *x*, *y*, and *n* denotes the corrupted images, clear images and noise, respectively. The *i* denotes the image index *u* denotes the values of all the other image pixels corrupted by other factors. *M* represents the binary images where *0* represents the missing and *1* represents the digital images' valid pixels

$$x(i) = \begin{cases} y(i) + n(i), M(i) = 1 \\ u(i), \qquad M(i) = 0 \end{cases} \tag{1}$$

Initially, diffusion-based approaches depend on the partial differential equations to fill the hole or missing pixels in digital images. However, the diffusion-based method blurs the filled pixel when filling the large holes in the digital images [13]. To improvise the filling and reduce the blurring of large mask pixels in digital and medical images exemplar-based inpainting method has been introduced. The pixel-based inpainting works on filling the missing region pixel by pixel, which consumes more time, and some pixels are misplaced. The patch-based inpainting method has been introduced to overcome the issues. It searches for an undamaged part on the images, called candidate patches. It will search for the damaged portion and replace the suitable candidate patches in the masked region of the images [52, 5, 11]. The inpainting method is highly used in digital and medical images to remove unwanted information or to fill the missing regions. The cervical cancer digital images face the challenges of specular reflection, affecting the image quality. The reflection region is identified and removed from the cervical images. The removed region appears as the hole, and the exemplar-based inpainting methods are used to fill the missing region with the neighbouring pixel regions [37, 32].

Similarly, the exemplar inpainting method is applied to the object's digital images for quality purposes. Analyzing each traditional inpainting method and identifying similar patches or pixels is challenging. The search algorithm consumes more time in computation and involves the distance measure metrics in the filling process. To automate and fast process the missing region or pixel in global consistency and fine texture deep learning inpainting method is introduced [4, 35]. The deep learning inpainting algorithm is used to fill up the missing pixels, which is based on supervised and unsupervised algorithms. A supervised deep learning inpainting technique is the convolution neural network, and an unsupervised deep learning inpainting technique is the Generative Adversarial Network (GAN). The hierarchical representation of the image inpainting method is shown in Figure 1. This paper discusses the automated supervised and unsupervised deep learning inpainting methods in digital medical images. It plays a significant role in computer vision in improvising the quality of medical images and other object images. The paper is organized in the following ways: Section 2 discusses the searching strategies applied to find the related articles for inpainting digital images. Section 3 discusses the Convolutional Neural Network (CNN) for inpainting digital and medical images. The section 4 discusses the generative adversarial neural network for inpainting the missing region on medical and digital images. Section 5 is the discussion and finally, the paper's conclusion.

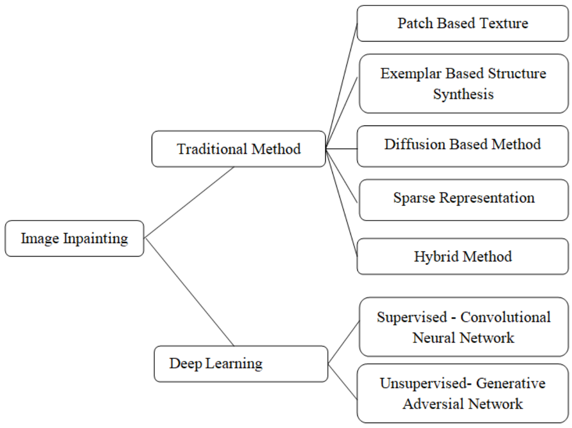


Figure 1. Hierarchical representation of the traditional and deep learning inpainting method used in digital and other medical images.

2. Searching Strategies

The purpose of the review paper is to include finding the challenges, advantages, and disadvantages of each deep learning inpainting model. From the whole data collected only deep learning inpainting method is included and other inpainting method are excluded for the review paper. From the different languages only, English paper is used in this article and other review paper are excluded from the review paper. The paper used for review is indexed in IEEE, Web of Science (WOS), Pub Med, and Scopus are filtered based on the Prisma diagram shown in Figure 2.

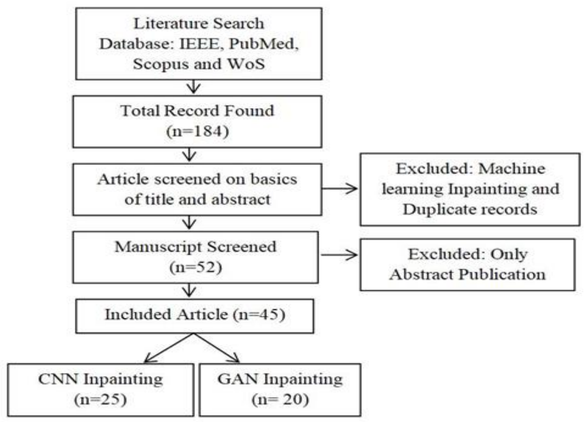


Figure 2. Prisma representation of the of study selection.

3. Convolutional Neural Network for Inpainting Digital Images

The supervised learning employs the convolutional filter on the images and replaces the removed or missing content with fixed pixel values. This section discusses in detail how CNNs and its modified version are used for inpainting digital and other medical images. The modified version of the CNN is shown in the Figure 3.

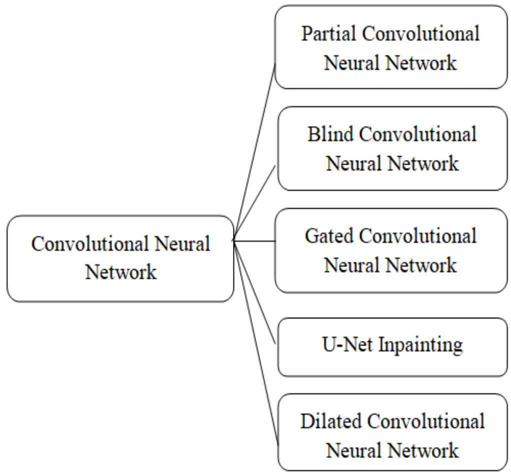


Figure 3. Hierarchical representation of the traditional and deep learning inpainting method used in digital and other medical images.

3.1. Partial Convolutional neural network (PConv)

The convolutional filter was applied to the valid pixels and predicted the missing value on the masked holes. Convolutional inpainting often leads to artifacts in colour discrepancy and blurriness of the images. Several filters and post-processing methods are applied to reduce the artifacts to overcome this problem. Nevertheless, it is time consuming and expensive and fails to achieve the desired result. Lie *et al*. [22] proposed a Partial Convolutional inpainting algorithm (PConv) for automated mask generation on the missing regions to overcome this problem. PConv inpainting is the reconstruction of image pixels based on the texture and structures of the digital images. It achieved higher-quality images in a single feed-forward pass along with the skip links to predict missing pixels. The proposed method combines partial convolutions and binary mask updates to perform inpainting on digital images. The PConv operation is expressed in Equation (2). Each pixel is unmasked, as shown in Equation (3).

$$x' = \begin{cases} W^T(X \odot M)\dfrac{sum(1)}{sum(M)} + b, & if sum(M) > 0 \\ 0, & otherwise \end{cases} \quad (2)$$

$$m' = \begin{cases} 1, & if sum(M) > 0 \\ 0, & otherwise \end{cases} \quad (3)$$

The convolutional filter W is applied to the binary masked M (image with missing pixels) images and feature value x with the corresponding bias b values. The element-wise multiplication is carried out to unmask the input images. After each partial operation, the binary mask is uploaded each time until all the pixels become the valid pixel in the digital images. The PConv inpainting is trained with the learning rate of 0.00005, along with the batch normalization and Adam optimizer.

Table 1. PConv inpainting on digital images.

Author	Objective	Dataset/Images	Masking and Metrics/Loss calculation	Merits	Demerits
Liu *et al*. [22].	Proposed the PConv inpainting method for the automatic updating of the binary mask with reduces artifacts.	PConv/ Image net, Places 2 and CelebA-HQ.	Masking: Free form mask. Metrics: L1 error, PSNR, SSIM, inception score. Loss calculation: Per-pixel loss, Visual Geometry Group (VGG) loss, Style loss, and Total variation loss.	Proposed for inpainting the irregular masked pixels. It reduces the artifacts like colour discrepancy and blurriness.	The image quality deteriorates catastrophically as the size of the hole increases. It fails to inpaint the sparsely structure images.
Patel *et al*. [28].	The PConv neural network approach along with the modified U-Net is used for the inpainting of digital images.	PConv/CelebA-HQ.	Masking: Free form mask. Metrics: L1 error, MSE, PSNR, SSIM. Loss calculation: Total variation loss, pixel loss, perceptual loss.	The modified U-Net along with the PConv neural network has reduced the loss rate than the existing P.Conv method.	The classical techniques are not much effective as the learning-based approaches for the image inpainting problem.
Mohite *et al*. [27].	Proposed PConv and contextual attention for the pixel wise filling the missing region.	PConv Contextual attention/Places, CelebA.	Masking: Square mask. Metrics: PSNR and MSE. Loss calculation: Perceptual loss and style loss.	The proposed method does not require the post processing for the removal of the noise in image inpainting.	The contextual attention deeply understands the available pixel and it consume more time for training.
Kaur *et al*. [19].	Proposed the U-Net and PConv for filling the mask hole in the manuscript images.	PConv/MPS dataset, Indian manuscript Digital images.	Masking: Free form mask and square mask. Metrics: PSNR, SSIM.	The inpainting model is perfect for small irregular holes and line and recovers the missing text in the perfect way.	The inpainting fails to generate the missing text when the holes are large.
Chen *et al*. [9].	Proposed the image inpainting with the sliding window for the digital Dunhuang mural images.	Permit Compliance System (PCS)/ Real-time dataset.	Masking: Random mask.	The proposed method restores the missing pixels with the promising accuracy and reduce time.	Some of missing regions are blurred with the visible traces.
Yan *et al*. [42].	Proposed the PConv inpainting with the attention mechanism to extract the long-distance pixel on images.	PCNet/CelebA-HQ, Place2, Paris-street view.	Masking: Free form mask Metrics: PSNR, SSIM, Loss calculation: Perceptual loss and style loss, smoothness loss, loss of holes, loss of non-hole effective pixels, total loss.	The convolutional kernel learns the feature distribution faster and helps in inpainting the large hole pixels on the scenery images or large area images.	It is not much suitable for the complex image reconstruction like human faces.
Kang *et al*. [18].	Proposed the WConv for inpainting the imbalanced pixels due to present of holes in the images.	WConv/Image net and CelebA, Digital images.	Masking: Free form mask and square masks. Metrics: PSNR, and SSIM. Loss calculation: L1 loss, L2 loss, and Total variation.	Resolves the instability and normalization problem due to present the invalid pixels.	The proposed method loss value is 7.755 where the loss of pixels is higher even it overcome the existing methods.

Patel *et al*. [28] used a robust PConv deep learning algorithm to fill up the empty portion of the digital images. For training, the modified U-Net model was proposed, built with seven encoders and a decoder. The proposed model consists of the ReLu activation function in each encoding layer and Leaky ReLu in the decoding except in the first layer. The image size is set with the i^{th} encoder, and the image size is automatically set the along with the kernel size. The proposed method is analyzed on two different applications, like automatic segmentation of the object removal and manual mask generation and inpainting of the removed pixels. Structural features like colors are highly considered for image inpainting, and pixel-per-information inpainting is more efficient than semantic inpainting for image reconstruction. The classical method is compared with the PConv method, PConv without style loss and PConv without perceptual loss. Based on the analysis, the classical method reduces the loss value in the inpainting process. Mohite *et al*. [27] proposed contextual attention with PConv for filling up the missing portion. Contextual attention is used pixel-wise inpainting by profoundly understanding the available content of the digital images. The PConv algorithm is replaced with the Stochastic Gradient Descent (SGD) to update the network's frequentative weight based on the training data. Initially, masked images are created, and original images are used for training the model Kaur *et al*. [19] proposed the deep learning approach to inpaint the manuscript images using PConv and U-Net architecture. The models are initially trained in two different steps; batch normalization is enabled with a learning rate of 0.0002 and 120 epochs. The batch normalization in the encoding stage is disabled with a learning rate of 0.00005 and 80 epochs. The Peak Signal to Noise Ratio (PSNR) and Structural Index Similarity (SSIM) give the maximum quality at the masking size of 16x16 and 4x512, respectively. Chen *et al*. [9] proposed the PConv along with the sliding window for inpainting the Dunhuang murals images. The sliding window is applied as the augmentation process for reducing the image patch size. It will reduce the computation time of the model. Yan *et al*. [42] proposed the Partial Convolution attention mechanism (PCNet) for the digital image inpainting method. The encoder layer is built with the convolutional and down-sampling layers. The generators and discriminators are applied to the encoder's stages. The decoders are built with deconvolutional, up-sampling layers to extract the digital images' features. The attention mechanism extracts the long-distance pixel information to fill the large holes in the digital images. It is replaced for batch normalization and helps normalize the pixel's feature in the digital images. Kang *et al*. [18] proposed a Weighted Convolution (WConv) to resolve the instability balances among pixels and the normalization issues present in the invalid pixels. The weighting scheme normalizes the layers and balances the invalid pixels caused by holes and zeroes padding. The WConv inpainting ratio of the mask is evaluated in the convolution operation using Equation (4) to update the binary mask. The normalization is applied on the updated mask, increasing the dynamic holes' stability with reduced complexity. The PConv neural network used for inpainting the digital images is shown in the Table 1.

- PConv *for medical images*: PConv inpainting is mainly used in medical images to remove unwanted artefacts. Susan et al. [36] applied the deep learning inpainting method to remove specular reflection on smart colposcopy images. The comparison analysis has been carried out on the three deep learning inpainting methods: Generative Multiple-column Convolutional Neural models (GMCNN), PConv model, and finally, the dilated convolutional model. The PConv completes the cervical images with plausible information in comparing the three deep learning methods. Yuan et al. [46] proposed a deep learning inpainting method to reconstruct the images that suffer from metal artifacts on the Computed Tomography (CT) image. The metal implants in the patient body affect the CT image quality. The U-Net with the PConv model is applied to reduce the metal artifact and improve image quality. For training purposes, the image size is set a 512x512 with a batch size of 2. The trained model generates the metal-free projection for all the 360-degree projections. Pimkin et al. [29] proposed a model for removing distorted metal trades in the sinogram. The PConv inpainting method has been modified for the quality of the improvised image. Compared with the initial PConv inpainting method, the modified version has improved the image quality. The sinogram inpainting U-Net with PConv was trained with the epoch value of 500 with the multiplication of the learning rate of 0.5. The learning rate is multiplied by 100, 200, and 300, respectively, with 0.1 on 400, 450, 475 and 490 epoch values. The proposed method outperforms the existing method with 64% of Mean Absolute Error (MAE). The PConv inpainting used for medical images is shown in the Table 2.

Table 2. PConv inpainting on medical images.

Author	Objective	CNN model /Dataset/Images	Masking and Metrics/Loss calculation	Merits	Demerits
Susan *et al*. [36]	Compared the different convolutional inpainting method for the identification of suitable method for smart colposcopy images.	GMCNN P.Conv D.Conv /Kaggle Dataset/ Digital cervical images.	Masking: Free form masking. Metrics: PSNR, SSIM. Loss Function: L1 and L2 loss.	-	The method is only implemented on the cervical medical images.
Yuan *et al*. [46]	The PConv method applied for inpainting the metal artifact reduction.	PConv / CT medical images.	Masking:Free form and square mask.	The PConv inpaints the boundary of the projection and well performs the existing method.	The method applied is analysed in the quantitative analysis.
Pimkin *et al*. [29]	Applied the PConv inpainting for the reduction of CT metal artifacts.	PConv /Real time dataset. CT metal artifacts.	Masking: Free form masking. Metrics: SSIM, MAE and MSE.	The proposed method reduces the overfitting to the metal shape set and achieve the-75% MSE value for the inpainting result.	The proposed method is specified for the multidomain. But it is analysed only on the brain CT images.

3.2. Bilateral Convolutional neural network (BConv)

Cai *et al*. [7] proposed a Blind Inpainting method based on a fully Convolutional Neural Network (BICNN) to remove invalid pixels. The SGD is built with the standard backpropagation to train the proposed BICNN to predict the missing pixel and produces the clean output with one pass of forwarding propagation. The BICNN is built with several convolutional layers with no pooling or fully connected layers. Liu *et al*. [24] proposed the effective blind inpainting method by using the deep CNN, which contains the encoder and decoder architecture to extract the digital images' features. The image gradient is computed to restore the missing information on the digital images. The robust loss function is used to fill the outline of the missing information using the equation. The proposed network is built with forty layers. It is tested on datasets like Set5, Set14, Urban, and Berkeley Segmentation Data Set (BSDS) 500 to attain the desired results. Wang *et al*. [38] proposed the new blind image inpainting method for predicting the unknown missing region pattern. The Visual Consistency Network (VCN) consists of two stages; first, it estimates the region to be filled, i.e., mask prediction network and what to fill in the missing region, i.e., Robust inpainting network. Schmalfuss *et al*. [33] proposed the sparse filter for improvising the blind inpainting method. The proposed method improvised the quality of the conventional CNN model and helped in the faster coverage of the network design. The Structured Receptive Filed Network (SRFN) filter is adopted in each CNN filter during the training process. It helps efficiently represent the salient feature of edges and corners of the digital images. The model is trained with 225,100 images with the 256x256 size collected from the dataset. The Blind convolutional inpainting on digital images is shown in the Table 3.

Table 3. Blind convolutional inpainting on digital images.

Author	Objective	CNN model /Dataset/ Images	Masking and Metrics/Loss calculation	Merits	Demerits
Cai *et al*. [7]	Proposed the blind inpainting method for the automatic inpainting the missing pixels.	BICNN/Digital images. Database: Training-image net, Testing-berkeley segmentation database.	Masking: Free form mask. Metrics: PSNR and SSIM are calculated.	The proposed method gives higher quality inpainting pixels in the small, dotted regions.	It the size of the missing pixels are larger than the sub images, then the inpainting pixels are blurred and large computation cost.
Liu *et al*. [24]	Presented a blind inpainting algorithm for restoration of the clear images.	3D-BICNN/Digital images. Database Set5, Set14, Urban100, BSDS500	Masking: Free form mask and square masking. Metrics: PSNR and SSIM are calculated. Loss calculation: Robust loss.	The average runtime time is reduced takes the 0.03 second for the computation.	It does not work on the important structure or details are corrupted. Example Nose and mouth of human face.
Wang *et al*. [38]	Proposed the new blind inpainting for training and predicting the missing patterns in the digital images.	VCNET/Digital images. Database: Flickr Face HQ dataset (FFHQ), CelebA-HQ image net, Places2.	Masking: Free form mask and square masking. Metrics: PSNR and SSIM are calculated. Loss calculation: Binary cross entropy loss.	Outperform the existing method and suitable for small holes and for automatic identification of the masking and refilling process.	For the large holes it cannot decide the dominant part and the performance dramatically decreases and fails in common occlusions problem.
Schmalfuss *et al*. [33]	Proposed the concepts of blind image inpainting method with the spares directional filter dictionaries approximation.	BI-Lightweight CNN/Digital images. Database: Places.	Masking: Free form mask and square masking. Metrics: PSNR and SSIM are calculated. Loss calculation: Binary cross entropy loss.	The proposed method gives the higher quality inpainting result with the minimum training data. The computation time is reduced to 2.453 msec.	The PSNR value is decreases for the pixel loss of 20-25% in the digital image. As the pixel loss increases the quality of the image's decreases.

- *Blind convolutional neural network inpainting on medical images*: Manjon *et al*. [25] proposed the 3D blind inpainting method to clean the lesion region to construct a healthy version of the Magnetic Resonance Imaging (MRI) brain images. It is the first 3D blind inpainting procedure in medical images for inpainting. The inpainting is applied to the lesion-affected images to generate a healthy version of the MRI images. It is obtained by masking the lesion region of the Multi Scale (MS) images and applying the inpainting process. It used of 3D-UNet along with the activation function of ReLu and Batch normalization to train the model. The blind convolutional inpainting on medical images is shown in the Table 4.

Table 4. Blind convolutional inpainting on medical images.

Author	Objective	CNN model/Dataset/Images	Masking and Metrics/Loss calculation	Merits	Demerits
Manjon *et al*. [23]	Proposed the deep network for blind inpainting the lesion in the brain images.	3D-BICNN/3D MRI images. Database: IXI dataset.	Masking: Free form masking. Metrics: PSNR and coefficient loss calculation MSE.	The proposed method computation process is 0.5 seconds in GPU and 8 seconds in CPU.	The correlation loss for the tested images is high.

3.3. Gated Convolution Neural Network (GCNN)

Yu *et al*. [45] proposed an image inpainting method using Gated Convolution (GConv) to learn the images without labelling. The vanilla convolutional treats all the empty pixels as valid pixels and affects the images during inpainting. The proposed method is highly meant for free-form masked images and learns the soft mask automatically. It is expressed in Equations (4), (5), and (6) where σ represents the sigmoid activation function and Ørepresent the other activation function.

$$Gating_{y,x} = \sum\sum W_g \,.\, I \tag{4}$$

$$Feature_{y,x} = \sum\sum W_f \,.\, Ib \tag{5}$$

$$O_{y,x} = \emptyset(Feature_{y,x}) \odot \sigma(Gating_{y,x}) \tag{6}$$

The W_g and W_f are the different convolution filter for inpainting the missing portion. The proposed methods learn the dynamic feature for each channel of the input images. The gated convolutional learned to highlight the masked region and sketched information of each channel to generate the inpainting results. Dai *et al*. [12] proposed the Spatial Temporal-based Gated Convolution Network (STGCN) for the cloud removal method using the deep learning algorithm and reconstructing the invalid pixel (i.e., removed portion) with the auxiliary information on the multi-temporal images. The gated convolution layer and sub-pixel convolution discriminate the cloud and non-cloud portion of the images to create the mask for inpainting and filling up the accurate pixels to create non-cloud images. The Atrous Spatial Pyramid Pooling (ASPP) module with different dilation rates is considered to extract the different features of receptive fields and concatenated to generate the missing regions. The model is trained with a learning rate of 10^{-4} and iterated with 500 epochs. Chang *et al*. [8] proposed the 3D-GCNN for inpainting the free-form mask and temporal patch GAN loss to improve temporal consistency. The GConv inpainting on digital images is shown in the Table 5.

Table 5. Gated convolutional inpainting on digital images.

Author	Objective	CNN model/Dataset/Images	Masking and Metrics/Loss calculation	Merits	Demerits
Yu *et al*. [45]	Proposed the gated convolutional to improvise the existing vanilla convolutional inpainting method.	GConv/Places2 and CelebA-HQ/Digital images.	Masking: Rectangle mask and free form mask. Loss calculation: L1 error and L2 error, L1 loss, and the GAN loss.	The proposed method outperforms the existing method and produces realistic results with seamless boundary transitions.	The proposed method fails in some of the human face images when the masking holes is large.
Dai *et al*. [12]	Proposed the cloud removal method using gated convolutional inpainting using the auxiliary information from the multi temporal images.	STGCN/WHU cloud dataset/Remote sensing images.	Masking: Free from masking, Loss calculation: PSNR, SSIM, spectral angle mapper, correlation coefficient Loss function: Joint loss, image-level loss, feature-level loss, and a total variation loss.	The proposed method outperforms the existing method by restoring the irregular clouds mask.	The proposed method is complicated and requires more computational time.
Chang *et al*. [8]	Proposed the 3D gated convolutional inpainting method and temporal Patch GAN loss for high quality videos.	3D-GCNN/Free form Video Inpainting (FVI)/Digital videos.	Masking: Free from masking. Loss calculation: FDI, MSE, LPI. Loss function: GAN loss.	The computation speed is faster with reduced number of parameters outperform the existing methods.	The method fails to generate the natural result for the thick mask.

3.4. U-Net Convolution Neural Network

The U-Net model is built with two main parts, one for the feature extraction and the other for the upsampling procedure to fuse the extracted feature on the digital images. Wei *et al*. [39] used the novel network with four channels with binary masked images where 1 represents the pixel to be repaired on the digital images. The RF_{i-1} represent the receptive field of the convolutional kernel of the upper layer and K_i represent the size of the convolutional kernel and S_n represent the steps involved in the convolutional of the n^{th} layers. The size of the convolutional kernel is set using the Equations (7) and (8).

$$RF = RF_{i-1} + (K_i - 1) * \prod_{n=1}^{i=1} S_n \tag{7}$$

$$K = k + (k-1)(d-1) \tag{8}$$

The k represents the convolutional kernel size with the dilation rate of d. The proposed method is evaluated with human eyes and objective data. The image inpainting is performed on the three groups, and four inpainting algorithms are implied to check the performance of the methods. The proposed method repairs the image with the known information of the images to defect the five-sense organ of the face's images. Yan *et al*. [43] proposed Shift-Net deep learning method to reduce blurry images and produce plausible results. The shift connection on the U-Net architecture fills the missing region of the sharp structure and fine texture detail of the digital images. The encoder extracts the feature from the known region of the pixel, and the decoder estimates the missing region using the features of the known regions on the images. From the dataset, 5000 are used for training, 900 for testing and 100 for the validation of the proposed model. The model is trained up to 30 epochs with a learning rate of $2x10^{-4}$. Zhou *et al*. [53] proposed a modified version of the Shift-Net architecture called Bishift-Net, which captures the encoder and decoder feature and rearranges the generated feature to produce the sharp texture on the digital images. The model produces fine details at the edges of each digital image. It is trained with 14,900 images with an image size of 256x256. The model is trained with 30 epochs with Adam optimizer and a learning rate of 2x0.0001 with a beta value of 1. The model is trained with the centre and the random masks and attains the PSNR values of 28.14db and 28.657, respectively. The similarity index for the centre and random masks is 0.937 and 0.919, respectively.

Hong *et al*. [17] used the fusion block to generate a flexible alpha composition map to combine the known and unknown regions to produce the structural and texture information. The fusion block initially extracts the raw completion image using the feature maps and predicts the alpha composition map to obtain the result by combining them. The masked images are done with the removal pixel of different ranges like [0-10%), [20-30%), [30-40%), and [40-50%). The model is trained with a batch size of 6 for each GPU processor and takes three days to complete the process. The learning rate is reduced from $2e^{-3}$ to $2e^{-6}$ in 20 epochs. The proposed method outperforms the existing method at each masking ratio, i.e., from 0-50%. Zeng *et al*. [47] proposed the Pyramid-context Encoder Network (PEN-Net) for inpainting digital images. It is built based on the U-Net architecture used to restore the image by encoding the input images' resolution and learning the semantic feature in the decoding stage. The images are resized and set as 256x256 for training and testing to training the model. The digital images are masked in the square size of 32x32, 64x64, and 128x128. The U-model used for inpainting the digital images is shown in the Table 6.

Table 6. U-model for Inpainting digital images.

Author	Objective	CNN model/Dataset/ Images	Masking and Metrics/Loss calculation	Merits	Merits	Demerits
Wei *et al*. [39]	Proposed a generator to repair the defect images using the improvised global and local discriminators.	U-Net/CelebA-HQ.	Masking: Square masking Metrics: PSNR, SSIM Loss calculation: Reconstruction loss (i.e., MSE).	The edges of the inpainted region are more detailed and it is same attributes as the entire image.	The edges of the inpainted region are more detailed and it is same attributes as the entire image.	Not applied on the free masking of the images.
Yan *et al*. [43]	Proposed to fill the missing portion of any shape with fine-detailed and sharp structure.	Shitft - Net/Places365, Paris Streetview, Real-world.	Masking: Rectangle, and free form Metrics: PSNR, SSIM. Loss calculation: Mean 2 loss, guidance loss, SN-Patch GAN loss, L1 loss.	The proposed method gives the fine details of the missing region with the fast speed computation.	The proposed method gives the fine details of the missing region with the fast speed computation.	The performance of the proposed is not specified for the large hole of the digital images.
Zhou *et al*. [53]	Proposed a modified shiftnet for the fine structure of inpainting using encoder and decoder.	BiShift-Net/CelebA-HQ and Paris Street-View, Digital images.	Masking: Center mask and random mask. Metrics: PSNR and SSIM, Loss calculation: Adversarial loss, Reconstruction loss, and Guidance loss.	Outperform the shiftnet model in inpainting the digital images.	Outperform the shiftnet model in inpainting the digital images.	The performance of the proposed is not specified for the large hole of the digital images.
Hong *et al*. [17]	Proposed a modified U-Net model with the fusion block to predict the unknown region.	DFNet/ Places2 and CelebA, Digtal images.	Masking: Free Form Masking Metrics: L1, Buffer Pool Extension (BPE), FID Loss calculation: Reconstruction loss, Perceptual loss, Style loss, Total variation loss, and Total loss.	The quality of the image completion at the boundary region are higher in the digital images.	The quality of the image completion at the boundary region are higher in the digital images.	The proposed model is not good for the larger hole because the information is not transmitted to the inner region of the images.
Zeng *et al*. [47]	Proopsed the PEN-Net to restore the missing region with high resolution.	PEN-Net/ Facade, Document Type Definition (DTD), CelebA-HQ, and places2, Digital images.	Masking: Square mask Metrics: Multi scale-SSIM, Inception score, and FID, Pyramid. Loss calculation: L1 loss, and adversarial training Loss.	The proposed method outperforms the existing method with the minimum computational time.	The proposed method outperforms the existing method with the minimum computational time.	The face detail's structure is missing in some of predicted images.

3.5. Dilated Convolution Neural Network (DCNN)

Salem *et al*. [31] deep learning uses the convolution filter to analyze each image and the mean value to replace the missing region. The dilated convolutional layer and multi-scale context are combined to fill the random missing region without affecting the resolution of the digital images. Typically, inpainting is a process that consumes more time to reduce the computational time, and the feature parameters dilated method is used to fill up the missing portion. The model is trained using 149000, selected from the Paris Street Review, where 36501 are selected for validation and 328501 are used for testing. The Adam optimiser is used for fast computation and trained with the NVIDIA GPU process. Prabhu *et al*. [30] adopt the structure of U-Shape architecture, which consist of feature encoding and feature decoder. The feature encoder is built with four convolutional layers, and the feature decoder helps to fuse the pixel information. The multiscale dilated convolution, up-sampling and down-sampling modules are used to improvise the synthesis images. The proposed model is built with the four convolutional layers with 128, 32, 32, 128, and the kernel size for the four layers is set as 3x3, 1x1, 3x3, and 1x1, respectively. It is trained with 1,500,000 images with SGD and a batch size of 8. The dilated CNN for inpainting the Digital Images is shown in the Table.7.

Table 7. Dilated convolutional model for inpainting digital images.

Author	Objective	CNN model/Dataset/Images	Masking and Metrics/Loss calculation	Merits	Demerits
Salem *et al*. [31]	Proposed for completing the random-shaped missing portion with variable size and arbitrary locations across the image.	DCNN/Places2, and Paris Street View/Digital images.	Masking: Random shaped mask. Metrics: PSNR, L1 loss, L2 loss. Loss function: Reconstruction loss.	Reduce the computational time and blurry region of the missing region.	The refilled region is marked with the fine lines and resolution of the pixel filled is low when compared with the original pixel.
Prabhu *et al*. [30]	Proposed the model to preserve the fine texture of the predicted pixel on the fingerprint images.	U-Finger: MS-DCN/Synthetic images.	Masking: Free form mask. Metrics: PSNR, MSE, SSIM.	Outperform the existing method and boosted the performance with less computation time.	Some of the large fingerprints are not slightly blurred and background regions are smudged.

4. Generative Adversarial Neural Network for Inpainting Digital Images

The unsupervised deep learning method used for inpainting the digital and other medical images is the Generative Adversarial Neural Network (GAN). The other modified version of the GAN model used for inpainting is shown in the Figure 4.

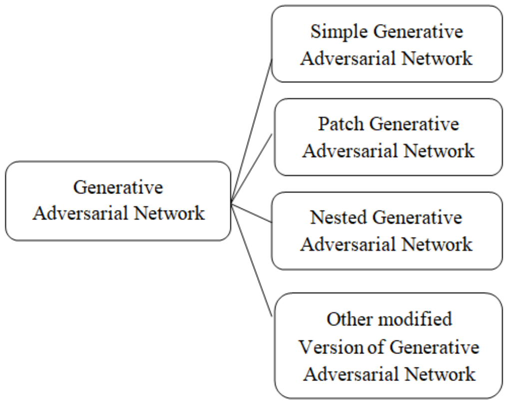


Figure 4. Modified version of the GAN.

4.1. Simple Generative Adversarial Network (GAN)

Aishwarya et al. [1] proposed the model to correct the noisy blurred and watery region of digital images. The model is built with the two adversarial networks to achieve the desired task and aims at the super-resolution of the images. The network efficiency is determined in terms of increased PSNR obtained, which is 28.19 dB with a less training period of approximately 14 hours compared with other network models that perform similar tasks. Zhang et al. [49], the model is constructed with the encoder and decoder, which act as the generative part along with the skip connection to fill the missing region of the face images. The discriminator is used to identify the real and fake pixels in the generated pixels. The model is trained with 13,233 face images with an image size of 128x128x3 with Adam optimizer and a learning rate of 2e-4. The discriminator network is trained with the RMSProp optimizer with a momentum of 0.5. Wei et al. [14] proposed that the model is constructed with the generator and discriminator for increasing the spatial resolution of the generated images in the missing region of the images. The discriminator uses the cross entropy loss to judge the missing pixel of the generated images as in Equation (9). The pi and qi represent the sample distribution, and qi is the generator generation distribution.

$$H(p,q) \coloneqq \sum_i p_i log q_i \tag{9}$$

The model is trained with the 10,177 face images from the CelebA dataset. Annotated Facial Landmarks in the Wild (AFLW) is a large-scale face database with multi-pose and multi view, where each face is marked with 21 feature points, which include images of different poses, expressions, lighting, and race. The model has been tested with the Super-Resolution (SR) direction. The model evaluation is used with datasets like Set14, BSD100, Urban100, and ManGA100 for evaluating model inpainting images. Based on the comparison analysis, the model outperforms the existing method, especially in the case of 8xSR. The model is faster and can produce a clear texture for the images under the conditions of 2×SR and 3×SR. Ghanem et al. [15] constructed the model with the generator based on the MobileNet V2 network to extract the landmarks prediction from the digital images. The landmarked and masked images are trained in the generator, which is based on the U-Net to produce the inpainted images. The generators GenP integrate missing pixels of the face images by taking the deteriorated images (C) along with the landmarks (L), which is taken as the input with θ P as network training parameters for the completion of the images (N) as in Equation (10).

$$N \coloneqq Gen_P(I^M, L;\ \theta_P) \tag{10}$$

The PSNR, SSIM and Fréchet Inception Distance (FID) are calculated for the proposed model for the random centre masks and centre random blocks masks. The random centre masks with the masking ratio of 10-20% attain a PSNR value of 34.97 with a similarity index of 0.989 and FID value of 1.82. The PSNR value is 29.12, the similarity index is 0.969, and the FID value is 2.91 for the masking ratio of 20–30%. The centre random blocks masks are taken for the digital images with a PSNR value of 27.12 and SSIM of 0.947. The FID for the proposed model is 3.506, higher than the existing methods. The simple GAN for inpainting digital images is shown in the Table 8.

Table 8. GAN for inpainting digital images.

Author	Objective	GAN model/Dataset/ Images	Masking and Metrics/Loss calculation	Merits	Demerits
Aishwarya *et al.* [1]	Proposed a method for removing the noise, blurry inpainting pixel in the digital images.	GAN/CelebA	Masking: Square mask. Metrics: PSNR. Loss function: L1 loss and MSE loss.	Our model automatically identify the locations that are distorted and abnormal and fill automatically with respect to the neighbouring pixels with less time.	The predicted pixel of the hole region resolution is very low for the digital images.
Zhang *et al.* [49]	Proposed the GAN model for repairing the color inconsistencies of the inpainting region on the digital images.	GAN/LFW dataset.	Masking: Square mask. Metrics: PSNR and SSIM. Loss function: MSE and structural loss.	The feature reconstruction loss makes the inpainting pixels to look more realistic.	But in some of the images the reconstructed region are blurred and resolution failed in the missing region.
Wei *et al.* [14]	Proposed a method for inpainting the missing pixels with higher resolution on the face images.	3D GAN/CelebA and AFLW.	Masking: Free form masking. Metrics: PSNR, and SSIM. Loss function: Cross entropy loss.	The proposed method result is satisfactory for the smaller missing region.	The model generate the pixels for the larger holes but not similar to original images.
Ghanem *et al.* [15]	Proposed a model for the improvement of facial completion on the digital images.	GAN/CelebA, CelebA-HQ, Novel landmarked face database.	Masking: Square mask. Metrics: PSNR, SSIM, FID.	The proposed method outperform the existing method in the PSNR and SSIM value.	The feature for real and generated images are very higher when compared to the existing method.

- *Simple GAN for medical images*: Armanious *et al*. [2] proposed the ipA-MedGAN framework to enable the arbitrarily shaped region for inpainting the medical field. The model is constructed with the new cascaded generator network based on the MultiRes-U-Net model. The model is trained using 3028 slices from 33 volunteers and 1072 slices from 11 volunteers for testing. The inpainting is applied on the 64x64 pixels randomly cropped using the pre-processed scans. Soganciolgu *et al*. [34] evaluated the performance of the recently published deep learning inpainting model using chest X-ray images. The generative model is trained with 1.2 million with 128x128 patches of healthy x-ray images and learned to predict the centre-masked region with the size of 64x64. Based on the comparative analysis, the semantic inpainting method attains a PSNR value of 33.85, context encoders attain a PSNR value of 26.31, and contextual attention attains a PSNR value of 31.80. Promising GAN Based Sinogram Inpainting Method for Ultra-Limited-Angle Computed Tomography Imaging proposed a method for inpainting medical images using the sinogram inpainting on the CT images. The model is constructed with reconstruction loss in the generator and discriminator to improve the medical images' quality for the experimental setup with the actual clinical dataset with the 2142 pleurable and cranial cavity 512x512 images for 12 patients. The model is trained with 1000 compelling images, and 200 images are selected for testing the model. The model is trained with the iteration of 50. The simple GAN for inpainting medical images is shown in the Table 9.

Table 9. GAN for inpainting medical images.

Author	Objective	Model/Dataset/ Images	Masking and Metrics/Loss calculation	Merits	Demerits
Karim *et al*. [2].	Proposed the pre-processing method for the correction in the PET/MRI images.	Real time dataset.	Masking: Square shaped mask. Metrics: PSNR. SSIM, UIQ. Loss function: MSE.	It act as the pre-processing method for further diagnosis of the medical images.	The pre-processing is not applicable for the square missing pixel.
Ecem *et al*. [34].	Proposed a model for the adoption of the medical images inpainting using generative model.	ChestX-ray14. X-ray image.	Masking: Square shaped region mask. Metrics: PSNR.	The proposed method helps in the enhancing the X-ray image which helps in identifying the abnormalities.	The model is only applied on the square mask of the X-ray images.

4.2. Patch Generative Adversarial Network (PGAN)

Liu *et al*. [23] proposed a method called Probabilistic Diverse (PD-GAN), which is built on the vanilla GAN and Spatially Probabilistic Diversity Normalization (SPD-Norm) to generate the image based on the random noise of the digital images. The Perceptual diversity loss empowered the loss of the diverse content generation of the missing coarse prediction, and mask images are initially sent to the SPD-Norm residual blocks to get the initial knowledge regarding the generation process. The SPD-Norm helps learn the scale and bias to generate the feature map. The model constructed is optimized using the Adam optimizer with β1and β2 as 0.0 and 0.99, respectively. The model is trained with 500K iterations with a batch size of 6. Chen *et al*. [10] proposed a method to reduce the edge loss function to enhance the inpainted effect of the posture and expression of the digital images. The model is constructed with multiple layers of discriminators along with the edge detector. The model is trained with 50,000 face images and tested with 12000 images collected from the CelebA dataset. The patch GAN for the digital images is shown in the Table 10.

Table 10. Patch GAN for inpainting digital images.

Author	Objective	CNN model/Dataset/Images	Masking and Metrics/Loss calculation	Merits	Demerits
Liu *et al*. [23]	Proposed a multiple inpainting to produce the visual realistic content on the missing portions.	PD-GAN/CelebA-HQ, Places2 and Paris Street.	Masking: Rectangle Mask Metrics: PSNR, SSIM, FID, and LIPIS. Loss Function: Perceptual diversity loss.	Outperform the existing method based on the qualitative analysis. The image with the masked ratio for 10-40% is giving higher performance.	The proposed method fails attains the quality for the masking ration of 40-50%.
Chen *et al*. [10]	Proposed the multi scale patch-based GAN model for inpainting the digital images.	MPGE/CelebA 256*256.	Masking: Square mask. Metrics: PSNR, SSIM, FID. Loss function: Perceptual diversity loss.	The model helps in reducing the edge loss of the missing region and outperform the existing method in qualitative analysis.	The method is performed only for the square masking region.

4.3. Nested Generative Adversarial Network (NGAN)

Li *et al*. [21] proposed a model built with the generative network to generate the inpainted images. The model is trained with 202,599 face images, the original face, and the missing pixel images. The original image size is set to 128x128, and the missing square region is masked with the size of 16x16 pixels. The proposed method is evaluated using the PSNR, MSE, SSIM, and FaceNet. Based on the quantitative analysis, the proposed method outperforms the existing method with the minimum loss errors. Wu *et al*. [40] proposed a novel generative framework to recover the missing region of the face images in quantitative and qualitative ways. The model is constructed with the generator, which helps synthesize missing regions and completes the face images. The model is trained with the Adam optimizer with the decay rate β1=0.5 and β1=0.99. The batch size is set as 16, and the generator updates the pixels for every five discriminator updates. The model training takes 15 hours for 200 thousand iterations on the NVIDIA TITAN X. The Nested GAN for inpainting digital images is shown in the Table 11.

Table 11. Nested GAN for inpainting digital images.

Author	Objective	CNN model/Dataset/ Images	Masking and Metrics/Loss calculation	Merits	Demerits
Li *et al*. [21]	Proposed the nested generative model for the construction of the non-blurred texture of the face images.	NGAN/CelebA	Masking: Square mask Metrics: PSNR, SSIM, MSE. Loss error: L1 loss error and L2 loss error.	The model generates the non-blurred and semantically non plausible and pleasing content in the missing region of the face images.	It fails to reconstruct the missing pixel on the profile images and it works only for the rectangle patch size of 16x16.
Wu *et al*. [40]	Proposed a method for the generating the realistic and high spatial resolution inpainting pixel in the missing portions.	Spatial Transformer Generative Adversarial Network (STN-GAN)/CelebA, 300-VW and face forensics	Masking: Centre free form mask. Metrics: PSNR, MS-SSIM, TV FID.	The proposed method outperform the existing method in completing the face images at different angles of faces.	The time consumption is higher for the proposed method than the existing method.

4.4. Other Generative Adversarial Neural Network

Mo *et al*. [26] proposed a model to generate realistic images in the face images. The generator consists of a convolutional layer, dilated convolution layers, and a deconvolutional layer. The Multi-Scale Generator (MS-GAN) was introduced after images to rebuild the missing portion of the images. For training, the reconstruction loss L_r is introduced to capture the structural information of the missing pixels on the digital image as represented in Equation (11), where L_2 is the distance between the pixel of the repair images and Z represents the noise mask of the digital images. The model is trained with 80,000 images and tested with 20,000 images. The batch size is set as 32 and goes under the iteration of 20,000, and each entire network will work 70,000 times.

$$L_{r(x,y)} = ||z\left(x - F\big((1-z)\odot x\big)\right)||_2^2 \tag{11}$$

Zeng *et al*. [48] proposed a model to capture the informative contexts and novel masking prediction using the generator and discriminator. The generator is built with the convolution, and Aggregated Contextual-Transformation (AOT) Block for enhancing the context information. AOT performs the splitting in the standard convolution into multiple sub-kernels. Each sub kernel performs a transformation of input feature x_1 with a different dilation rate and finally aggregates all the features of the standard convolution and learns the residual feature x_2 with the sum weighted as g as in Equation (12).

$$x_3 = x_1 \times g + x_2 \times (1-g) \tag{12}$$

The proposed model is trained with the 1.8 million and for testing 36,500 of data from the Place 2 dataset. From the CelebA-HQ dataset, 28,000 images for used for training and 2,000 for testing. Similarly, 15,975 images are used for training and 2,777 images are used for testing the model. The proposed method is trained with different masking ratios like 1-10%, 20-30%, 30-40%, 40-50% and 50-60%. Yeh *et al*. [44] constructed the model with the generator and discriminator. The model is trained with the 202,599 face images and tested with 2000 images collected for the CelebA dataset. The images are cropped in the centre region with 64x64 face images with various viewpoints and expressions. The Adam optimizer restricts z as [-1, 1] in each iteration and terminates the back propagation after 1500 iterations. The model is evaluated with different masking like centre, pattern, random and half image masking. In centre masking, the proposed model attained the PSNR value of 19.4 for the CelebFaces Attributes (CelebA) dataset, 19.0 for the Street View House Numbers (SVHN) dataset and 13.5 for the Stanford Cars data set. In pattern masking, the proposed model attained the PSNR value of 17.4 for the CelebA dataset, 19.8 for the SVHN dataset and 14.1 for the Stanford Cars data set. In random masking, the proposed model attained the PSNR value of 22.8 for the CelebA dataset, 33.0 for the SVHN dataset and 14.1 for the Stanford Cars data set. In half masking, the proposed model attained the PSNR value of 13.7 for the CelebA dataset, 14.6 for the SVHN dataset and 11.1 for the Stanford Cars data set. Xu *et al*. [41] proposed a model constructed with the generator and attention module to create the masking to differentiate the highlighted region from the greyscale images. The extracted mask images are trained in the auto-encoder to refill the highlighted region. The output images are sent to the discriminator to differentiate the original and fake pixels of the digital images. The difficulty is faced in the synthetic dataset using 3D modeling software to create the highlights on the grayscale images. The model is trained with 1000 images with a batch size of 1 and a learning rate of 0.0002. The decay rate is set as β1=0.6 and β1=0.9 along with the Adam optimizer. Zheng *et al*. [51] proposed a method to capture the digital images' long-range dependency and high-level semantics using Cascaded Modulation (CM-GAN). The model is constructed with the masked image encoder using the

Fourier convolution block and cascade spatial block module, which works as the decoder. The model is trained with the Adam optimizer with the learning rate and batch size of 0.001 and 32, respectively. The size of the images is set as 512x512, which is collected from the place2 dataset.

Xian *et al*. [41] proposed a generative model for predicting the missing pixel on the face images. The domain embedding net combines the convolutional and deconvolutional to locate the missing region on the images. From the CelebA dataset, 162770 images are selected for training, 19867 images are selected for validation, and 19960 images are selected for testing the models. The model is trained with the Adam optimizer with the learning rate of 0.0002, β1=0.5, β2=0.999. The model is trained 80 times with a batch size of 60. The model is evaluated by taking the individual parts of the face images like left eyes, right eyes, upper face, left face, right face and lower face. Normalization Cross-Correlation (NCC) has been commonly used to evaluate the similarity between two compared images. Hedjazi *et al*. [16] model is constructed with four GAN generators to fill the corrupted region of the digital images. The corrupted and masked images are concatenated and given as the input images to the constructed model. For the model training, the 1.8 million images with the 400 different categories of the scene are used from the place2 dataset and for evaluation, 30thousan cropped face images are used, which are collected from the CelebA dataset. The images are masked in the ratio of 10-20%, 20-30%, 30-40% and 40-50%. The Adam optimizer is used along with a decay rate of 0.5 and 0.99, respectively. The model is trained with a batch size of 32 with a learning rate of 10^{-4} for the generator and discriminators. The other GAN model used for inpainting the digital images is shown in the Table 12.

Table 12. Other GAN model for inpainting the digital images.

Author	Objective	CNN model/Dataset/ Images	Masking and Metrics/Loss calculation	Merits	Demerits
Mo *et al*. [36]	Proposed the multiscale discriminator method for improvising the visual consistency of the face images.	MS-GAN/LFW and celebrity face Image.	Masking: Square mask. Metrics: PSNR and SSIM. Loss function: Reconstruction loss, and distance loss.	The proposed method outperforms the existing method based on the quantitative analysis.	During the qualitative analysis the stage pixels appears in the filled region that cause artefacts and make images to look unnatural.
Zeng *et al*. [48]	Proposed the AOT GAN model for the large free masked missing pixels.	AOT GAN /Places2, CelebA-HQ, and QMUL-OpenLogo.	Masking: Free form mask. Metrics: SNR, SSIM, and FID.	The proposed method is gives the rich pattern in the extremely large missing region of the digital images.	-
Yeh *et al*. [44]	Proposed the high-level context generative inpainting method for filling the missing pixel on digital images.	Deep generative models/CelebA, Street view, and Stanford cars.	Masking: Square and free form. Metrics: PSNR Loss function: Weighted context loss, and prior loss.	The model is suitable for the small region of missing pixels.	The model fails in the refilling the complex data like face and in refilling the large holes in the scenery.
Xu *et al*. [41]	Proposed attentive GAN for the specular highlight removal from greyscale images.	Attentive GAN/Synthetic dataset using 3D modelling software.	Masking: Free form mask. Metrics: SSIM, PSNR, and MSE.	The proposed method outperform the existing method in the quantitative evaluation.	There is no specification of the reflection region detection to specify the filling quality of region.
Zheng *et al*. [51]	Proposed the inpainting method for long range dependency and high level semantic images.	CM-GAN/CelebA, CelebA-HQ.	Masking: Square mask, and random mask Metrics: PSNR, NCC. Loss function: L1 loss.	The high-resolution images produced by the model are best viewed by zooming in on the screen.	Our model still has limitations when it comes to synthesising big objects like people or animals.
Xian *et al*. [50]	Proposed a method for the generating the missing pixel with high resolution on the face images.	Domain Embedded Multi-model Generative Adversarial Networks (DE-GAN)/CelebA, CelebA-HQ.	Masking: Square mask, and random mask. Metrics: PSNR, SSIMand NCC. Loss function: L1 loss.	The proposed method outperform the existing methods with the resolution and also filling the missing pixels from the side images.	-
Hedjazi *et al*. [16]	Proposed multi GAN for performance improvisation and rendering efficiency.	Multi-GAN/Places2, 3D, CelebA-HQ.	Masking: Random mask. Metrics: PSNR, SSIM and MAE.	The global structure consistency and fine-grained textures is improvised and speeds up the computational time than the existing methods.	The proposed method is not focused in the super resolution of the digital images.

- *Other GAN for the Medical Images*: Armanious *et al.* [3] proposed a model for filling the missing pixels on the CT and MRI images. The model is trained with the weight of loss function like λ1=0.8, λ2=0.2 and λ3=λ4=0.0001was used. The model is trained for 200 epochs using NVIDIA Tital X GPU, and the training times consume an average of 24 hours with an approximate inference time of 100 milliseconds. The trained model with 50 volunteers was collected from the clinical CT scanners. The proposed model is applied to the CT and MRI images for quantitative analysis. The other GAN model for inpainting used for the medical images is shown in the Table 13.

Table 13. Other GAN model for inpainting the digital images.

Author	Objective	Model/Dataset Images/Image Type	Methodology	Merits	Demerits
Karim *et al.* [3]	Proposed a method for the removal of multitude of factors like metallic implants on the MRI images.	ipA-MedGAN/MR dataset (T2) Weighted MRI brain image.	Masking: Square shaped mask, free form mask. Metrics: PSNR, SSIM, UIQ, and MSE. Loss function: Non-adversarial loss.	It is enabled for the inpainting of arbitrary shaped regions with no a priori knowledge on the pixel locations of the regions of interest.	-
Li *et al.* [20]	Proposed a method for refilling the missing region of the CT images with the estimated sonograms.	Sinogram-Inpainting-GAN (SI-GAN)/Real clinical dataset.	Masking: Square shaped mask. Metrics: PSNR, RMSE, NMAD, and SSIM. Loss function: Sinogram loss.	The proposed method outperforms the existing method and helps in the refilling of the larger region of the images.	The errors of individual pixels in sinograms may lead to unknown errors in reconstructed image.

5. Discussion and Analysis

Based on the analysis, the convolution deep learning method is highly used to find plausible pixels on digital and medical images. Very few techniques are used and applied on the medical image inpainting to remove the noise and metallic part of the imaging devices. Based on above analysis the deep learning inpainting is the fast-growing research year from the year 2017-2022 as shown in the Figure 5.

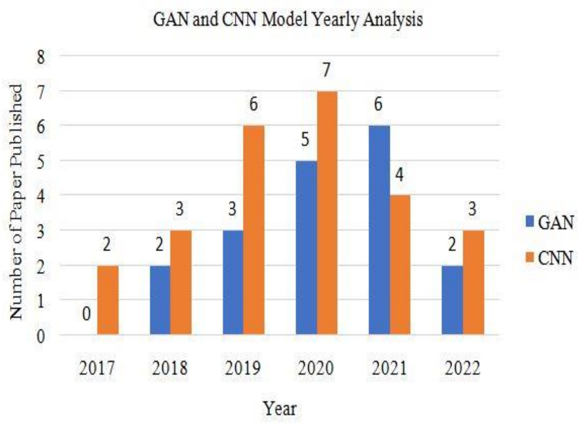


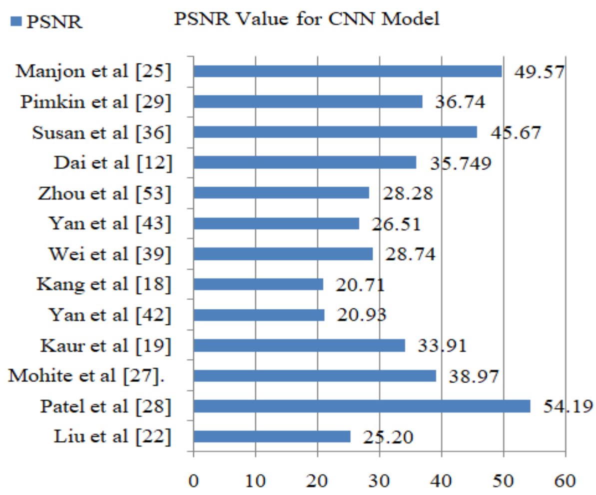
Figure 5. Yearly analysis of the CNN and GAN model.

5.1. Quantitative Analysis

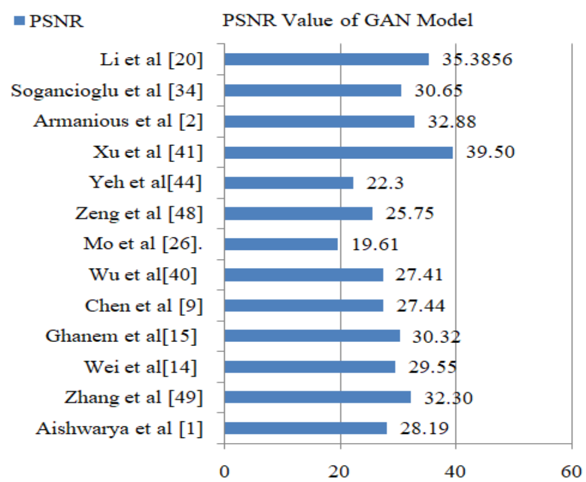
The quantitative analysis includes the performance calculation and the loss calculation in the deep learning inpainting method. In this section discusses the metrics used for the performance calculation and the loss calculation for the predicted pixel regions.

- *Performance Calculation*: there are various metrics used for the quantitative analysis of the inpainting method. The metrics are PSNR, SSIM [18, 19, 22, 27, 28, 39, 42, 43, 53], Universal Image Quality (UIQ), Root Mean Square Error (RMSE), MAE, and FID.

Dai *et al*. [12] used the FID and learned perceptual image patch similarity to measure the quality of the predicted output and value of 0.0276 and 0.472. Zeng *et al*. [47] use the Multiscale Structural Similarity Index (MS-SSIM) and inception distance for quality analysis and attained the value of 78.09% and 50.51% on the digital images. Sogancioglu *et al*. [34] used the average RMSE and Normalized Mean Absolute Distance (NMAD) and attained the value of 0.0097and 0.1467 respectively. The inpainting performance results in varies depending on the different mask sizes, and it also depends on the various resolution size of the digital images like 512x512,256x256,128x128. Xu *et al*. [41] used Learn Perceptual Image Patch Similarity distance (LPIPS) is calculated and attained 0.129 for the Co-Modulated (CoMODGAN) mask dataset. The Paired/Unpaired Frechet Inception Distance (P-IDS/U-IDS) is calculated and attained the value of 40.20 and 25.56 respectively.

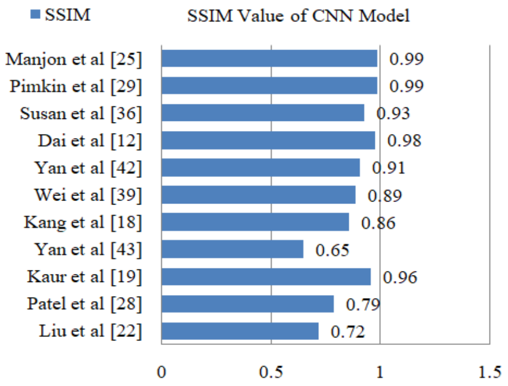


a) PSNR value of the CNN models.

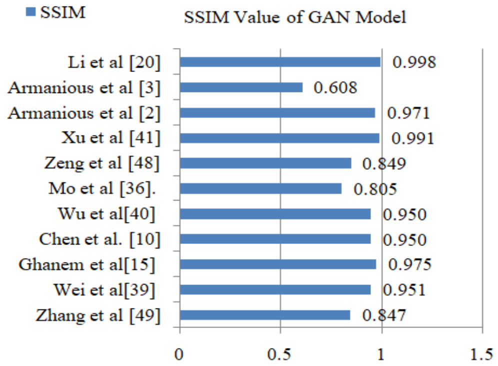


b) PSNR value of the GAN models.

Figure 6. Performance analysis of the deep learning inpainting model.



a) SSIM value of the CNN models.



b) SSIM value of the GAN models.

Figure 7. Performance analysis of the deep learning inpainting model.

The PSNR value calculated shown in the Figure 6-a). PSNR is a commonly used metric to assess the quality of image compression algorithms, with higher values indicating better preservation of image details. The authors like Liu *et al*. [22], Kaur *et al*. [19], and Dai *et al*. [12], achieved PSNR values in the range of 25 to 35. These results suggest that their algorithms provide a moderate level of image detail preservation. Interestingly, Susan *et al*. [36] achieved a PSNR value of 45.67, indicating a relatively high level of image quality preservation. Additionally, Manjon *et al*. [25] achieved a PSNR value of 49.57, which is the second-highest value. Similarly, the PSNR value of the GAN models is shown in the Figure 6-b). Based on the analysis, the Xu *et al*. [41] achieves the highest PSNR value of 39.508 indicating the strong image quality preservation. Mo *et al*. [26] obtains the lowest PSNR value of 19.61, suggesting potential image detail loss. Chen *et al*. [9] and Wu *et al*. [40] both achieved relatively low PSNR values, indicating suboptimal image quality. Finally, Li *et al*. [20] obtained a commendable PSNR value of 35.35 by effective preservation of image details. The SSIM value for the CNN model is shown in the Figure 7-a). Based on the observation, Pimkin *et al*. [29] and Manjon *et al*. [25] achieves the highest SSIM value of 0.99 indicating the strong similarity to the original images. Kaur *et al*. [19] and Dai *et al*. [12] also attain the higher SSIM value of 0.96 and 0.98 preserving the structural details. Similarly, the SSIM value of the GAN model is shown in the Figure 7-b). Li *et al*. [20] achieving the highest SISIM value, which indicates the generated images have the higher resembles to the original images. Zhang *et al*. [50] obtain the SSIM value of 0.99 by preserving the structural details.

- *Loss Calculation*: like the metrics, the loss calculation is an important analysis for measuring the missing pixel on the inpainting images. The lesser the loss value higher the accuracy of the inpainting images. There are various loss calculations for the inpainting method, like L1 loss, total loss, L2 loss, and GAN loss, Robust loss, Binary cross entropy loss, Mean Square Error loss (MSE), Joint loss, image level loss, feature loss, total variation loss etc., Wang *et al*. [38] calculated the binary cross entropy to know the loss of pixel on the predicted pixel of the digital images. Schmalfus *et al*. [33] proposed model has the loss percentage of 0-5% in the predicted pixels. It helps in identifying the minute loss happened in the predicted pixels. Manjon *et al*. [25] used the mean square loss to reduce the loss of pixel in the predicted pixels of the missing regions.

Yan *et al*. [42] introduced the guidance loss to minimize the loss between the predicted and the ground truth images. Salem *et al*. [31] used the L1 loss and L2 to predict the percentage of loss occurred in the predicted pixels i.e., the predicted output has the loss percentage of 6.45 and 1.6 in the predicted output. Very few methods are used for improvising the medical image quality, and loss calculation is not calculated to determine the inpainted pixel quality [30, 34, 45].

5.2. Research Challenges

Based on the above analysis, some of the most common challenges still faced in the field of deep learning inpainting are:

1) Training time: the number of images used for the training process is proportional to the quality of the results obtained. But the consumption time for the training process is randomly high in deep learning inpainting algorithms. To address this problem some of the models are trained using only one of the datasets, followed with the procedure called fine tuning. To test the model within different applications and its generalization capacity, the process is repeated for each one of the datasets. Aishwarya *et al*. [1] model, approximately takes 14 hours with K80 GPUs, to complete the entire training procedure. The model is trained with only one dataset with the images of 50,000 images. But to improvise the model it should be tested with different types of models. Different approaches have been tested. By combining spatially discounted reconstruction loss with a weighted mask, the Yu *et al*. [24] model was able to cut the training period to just one week. The length of training differs for each model, though. The

quality of the results and generalization abilities can be improved and multiple dataset training can be accomplished by shortening the training time, enabling models to be trained just once. Deep learning uses a lot of data to produce photographs with improved quality. Similar to how the inpainting used specific data for training.

2) Post processing: some models produce regions that have slight color differences from the nearby regions. They remedy this with a quick post-processing that involves mixing the finished region with the color of the pixels around it. The Yu *et al*. [24] model used the post processing approach to acquire the required pixel on the final results, however many results still fall short of producing results that are believable and also cause a greater loss value.

3) Computational resources: despite the positive outcomes, certain works need a lot of computer power and memory, which makes the model computationally intensive. More than 100M parameters are used in coarse-to-fine models [44]. About 33 million parameters are used by partial convolution-based techniques [49]. The model by Yu *et al*. [24] is one of the heaviest since it employs two stacked generative networks. It is still difficult to reduce processing resources without significantly lowering the quality of the reconstruction.

4) Arbitrary masks: a limitation of local discriminator models is that they can only handle a single rectangular hole region. Therefore, the local discriminator will not work if any hole emerges in real-world applications with random forms, sizes, and positions. Liu *et al*. [22] results are astounding, handling masks of any size, shape, placement, or distance from the image borders without performance degradation as the size of the holes increases. Another approach to deal with this is to include the mask as an additional input, however doing so requires the user to mark the area that needs to be filled, which slows down the automation process. Some of the dataset used for inpainting the medical images are shown in the Table 14.

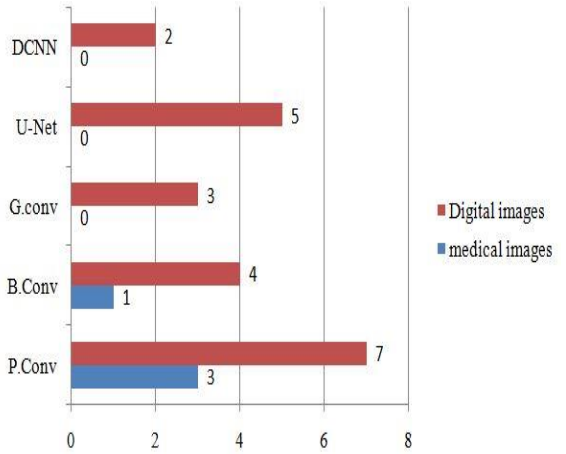
Table 14. Dataset used of inpainting the digital images.

Dataset	Related Papers
Image Net	[7, 18, 22, 38, 45]
Places2	[16, 17, 22, 23, 27, 33, 38, 42, 43, 45, 47, 48]
CelebA	[8, 11, 14, 15, 17, 18,21, 27, 40, 41, 45, 51]
CelebA-HQ	[16, 22, 23, 28, 38, 39, 43, 45, 48, 47, 53]
Medieval Paleographical Scale (MPS) dataset,	[19]
Paris-Street View	[23, 42, 23, 44, 53]
Berkeley segmentation database	[7]
Set5, Set 14, Urban100, BSDS500	[24]
FVI	[49]
LFW dataset	[14]
AFLW	[24]
LFW and Celebrity	[41]
Synthetic	[40]
WHU cloud dataset	[12]

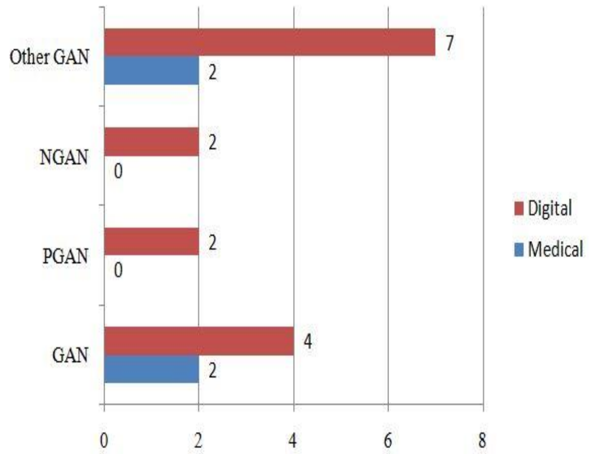
5) Medical image inpainting: to the medical image are considered for the further analysis so the quality of the medical images is highly important. But the quality of the images always fails during the images capturing due to the noise like metal artifacts or specular reflection covering the certain region [3, 25, 29]. So, this region should be identified and removed using the inpainting process. But very few inpainting methods are applied on the medical images as shown in the Figure 8. Similarly, for inpainting highly number of data are required but for medical images getting the higher number of data is very challenging. The only solution is to provide for medical images is to develop the model to give maximum quality with the lesser dataset. Some of the dataset used for inpainting the medical images are shown in the Table 15.

Table 15. Dataset used of inpainting the digital images.

Dataset	Related papers
Kaggle dataset	[22, 36].
Information eXtraction from Images (IXI) dataset	[25]
ChestXray14	[34].
MR dataset,	[3].
Real clinical dataset	[42].



a) Total CNN models used for inpainting digital and medical images.



b) Total GAN models used for inpainting digital and medical images.

Figure 8. Deep learning inpainting used for medical and other digital images.

6. Conclusions

Image inpainting is the one of the important tasks in the computer vision application. Due to the digitization the inpainting techniques is used in the field of Geographical, medical and other digital images. It is highly used for the quality enhancement and the restoration of the missing pixels. So, in this paper summarized the overall research challenges faced in the deep learning inpainting method. The merits and demerits of each model are summarized in this paper. The performance analysis of each model is analyzed and other quantitative analysis metrics are also discussed in this paper. The importance of the inpaining method in the medical field is also discussed and the challenge in the medical field is the lack of dataset. To conclude, there is no method that can inpaint all types of distortion in images but using learning techniques provides some promising results for each category of the analyzed cases.

7. Acknowledgement

The authors would like to express deep gratitude to the "*Centre for Machine Learning and Intelligence*" for supporting and sharing the resource for this research work. The Department of Science and Technology supports it under the scheme of DST CURIE (AI) sanctioned for the "Core Research grant for Artificial Intelligence (AI)" in the year 2021-2023.

References

[1] Aishwarya G. and Krishnan K., "Generative Adversarial Networks for Facial Image Inpainting and Super-Resolution," *Journal of Physics: Conference Series*, vol. 2070, no. 1, pp. 1-10, 2021. doi:10.1088/1742-6596/2070/1/012103

[2] Armanious K., Kumar V., Abdulatif S., Hepp T., Gatidis S., and Yang B., "ipA-MedGAN: Inpainting of Arbitrary Regions in Medical Imaging," *in Proceedings of the IEEE International Conference on Image Processing*, Abu Dhabi, pp. 3005-3009, 2020. doi: 10.1109/ICIP40778.2020.9191207

[3] Armanious K., Mecky Y., Gatidis S., and Yang B., "Adversarial Inpainting of Medical Image Modalities," *in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Brighton, pp. 3267-3271, 2019. doi: 10.1109/ICASSP.2019.8682677.

[4] Ayush T. and Sayak P., "Introduction to Image Inpainting with Deep Learning," WandB, https://wandb.ai/site/articles/introduction-to-image-inpainting-with-deep-learning, Last Visited, 2023.

[5] Borole R. and Bonde S., "Patch-Based Inpainting for Object Removal and Region Filling in Images," *Journal of Intelligent Systems*, vol. 22, no. 3, pp. 335-350, 2013. https://doi.org/10.1515/jisys-2013-0031

[6] Buyssens P., Daisy M., Tschumperle D., and Lezoray O., "Exemplar-based Inpainting: Technical Review and New Heuristics for Better Geometric Reconstructions," *IEEE Transactions on Image Processing*, vol. 24, no. 6, pp. 1809-1824, 2015. DOI:10.1109/TIP.2015.2411437

[7] Cai N., Su Z., Lin Z., Wang H., Yang Z., and Ling B., "Blind Inpainting Using the Fully Convolutional Neural Network," *The Visual Computer*, vol. 33, no. 2, pp. 249-261, 2017. DOI: https://doi.org/10.1007/s00371-015-1190-z

[8] Chang Y., Liu Z., Lee K., and Hsu W., "Free-Form Video Inpainting with 3D Gated Convolution and Temporal PatchGAN," *in Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, pp. 9066-9075, 2019. https://doi.org/10.48550/arXiv.1904.10247

[9] Chen M., Zhao X., and Xu D., "Image Inpainting for Digital Dunhuang Mural Using Partial Convolutions and Sliding Windiw Method," *Journal of Physics: Conference Series*, vol. 1302, no. 3, pp. 1-6, 2019. DOI:10.1088/1742-6596/1302/3/032040

[10] Chen G., Zhang G., Yang Z., and Liu W., "Multi-Scale Patch-GAN with Edge Detection for Image Inpainting," *Applied Intelligence*, vol. 53, no. 4, pp. 3917-3932, 2023. https://doi.org/10.1007/s10489-022-03577-2

[11] Criminisi A., Perez P., and Toyama K., "Region Filling and Object Removal by Exemplar-Based Image Inpainting," *IEEE Transactions on Image Processing*, vol. 13, no. 9, pp. 1200-1212, 2004. DOI:10.1109/TIP.2004.833105

[12] Dai P., Ji S., and Zhang Y., "Gated Convolutional Networks for Cloud Removal From Bi-Temporal Remote Sensing Images," *Remote Sensing*, vol. 12, no. 20, pp. 3427, 2020. https://doi.org/10.3390/rs12203427

[13] Elharrouss O., Almaadeed N., Al-Maadeed S., and Akbari Y., "Image Inpainting: A Review," *Neural Processing Letters*, vol. 51, pp. 2007-2028, 2020. https://doi.org/10.1007/s11063-019-10163-0

[14] Gao X., Nguyen M., and Yan W., "Face Image Inpainting Based on Generative Adversarial Network," *in Proceedings of the 36th International Conference on Image and Vision Computing*, Tauranga, pp. 1-6, 2021. DOI:10.1109/IVCNZ54163.2021.9653347

[15] Ghanem R. and Loey M., "Face Completion Using Generative Adversarial Network with Pretrained Face Landmark Generator," *International Journal of Intelligent Engineering and System*, vol. 14, no. 2, pp. 295-305, 2021. DOI:10.22266/ijies2021.0430.26

[16] Hedjazi M. and Genc Y., "Efficient Texture-

Aware Multi-GAN for Image Inpainting," *Knowledge-Based Systems*, vol. 217, no. 3, pp. 1-24, 2021. https://doi.org/10.1016/j.knosys.2021.106789

[17] Hong X., Xiong P., Ji R., and Fan H., "Deep Fusion Network for Image Completion," *in Proceedings of the 27th ACM International Conference on Multimedia*, Nice, pp. 2033-2042, 2019. https://doi.org/10.1145/3343031.3351002

[18] Kang J., Lee S., Heo S., and Lee S., "Image Inpainting Using Weighted Mask Convolution," *in Proceedings of the Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, Auckland, pp. 1262-1267, 2020. https://ieeexplore.ieee.org/document/9306435

[19] Kaur A., Raj A., Jayanthi N., and Indu S., "Inpainting of Irregular Holes in a Manuscript using UNet and Partial Convolution," *in Proceedings of the 2nd International Conference on Inventive Research in Computing Applications*, Coimbatore, pp. 778-784, 2020. DOI:10.1109/ICIRCA48905.2020.9182917

[20] Li Z., Cai A., Wang L., Zhang W., Tang C., Li L., Liang N., Yan B., "Promising Generative Adversarial Network Based Sinogram Inpainting Method for Ultra-Limited-Angle Computed Tomography Imaging," *Sensors*, vol. 19, no. 18, pp. 3941,2019. DOI:10.3390/s19183941

[21] Li Z., Zhu H., Cao L., Jiao L., Zhong Y., and Ma A., "Face Inpainting Via Nested Generative Adversarial Networks," *IEEE Access*, vol. 7, pp. 155462-155471, 2019. DOI:10.1109/ACCESS.2019.2949614

[22] Liu G., Reda F., Shih K., Wang T., Tao A., and Catanzaro B., "Image Inpainting for Irregular Holes Using Partial Convolutions," *in Proceedings of the 15th European Conference on Computer Vision*, Munich, pp. 89-105, 2018. https://doi.org/10.1007/978-3-030-01252-6_6

[23] Liu H., Wan Z., Huang W., Song Y., Han X., and Liao J., "PD-GAN: Probabilistic Diverse GAN for Image Inpainting," *in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Nashville, pp. 9371-9381, 2021. DOI:10.1109/CVPR46437.2021.00925

[24] Liu Y., Pan J., and Su Z., "Deep Blind Image Inpainting," *in Proceedings of the 19th International Conference Intelligence Science and Big Data Engineering. Visual Data Engineering, Lecture Notes in Computer Science*, Nanjing, pp. 128-141, 2019. https://doi.org/10.1007/978-3-030-36189-1_11

[25] Manjón J., Romero J., Vivo-Hernand R., Rubio G., Aparici F., De La Iglesia-Vaya M., Tourdias T., and Coupé P., "Blind MRI Brain Lesion Inpainting Using Deep Learning," *in Proceedings of the 5th International Workshop on Simulation and Synthesis in Medical Imaging*, Lima, pp. 41-49, 2020. https://doi.org/10.1007/978-3-030-59520-3_5

[26] Mo J. and Zhou Y., "The Image Inpainting Algorithm Used on Multi-Scale Generative Adversarial Networks and Neighbourhood," *Automatika: Journal for Control, Measurement, Electronics, Computing and Communications*, vol. 61, no. 4, pp. 704-713, 2020. https://doi.org/10.1080/00051144.2020.1821535

[27] Mohite T. and Phadke G., "Image Inpainting with Contextual Attention and Partial Convolution," *in Proceeding of the International Conference on Artificial Intelligence and Signal Processing*, Amaravati, pp. 1-6, 2020. DOI:10.1109/AISP48273.2020.9073008

[28] Patel H., Kulkarni A., Sahni S., and Vyas U., "Image Inpainting using Partial Convolution," *arXiv Preprint*, arXiv:2108.08791v1, 2021.

[29] Pimkin A., Samoylenko A., Antipina N., Ovechkina A., Golanov A., Dalechina A., and Belyaev M., "Multidomain CT Metal Artifacts Reduction Using Partial Convolution Based Inpainting," *in Proceedings of the International Joint Conference on Neural Networks*, Glasgow, pp. 1-6, 2020. DOI:10.1109/IJCNN48605.2020.9206625

[30] Prabhu R., Yu X., Wang Z., Liu D., and Jiang A., *Inpainting and Denoising Challeng*es, Springer, 2019. https://doi.org/10.1007/978-3-030-25614-2_3

[31] Salem N., Mahdi H., and Abbas H., "Random-Shaped Image Inpainting Using Dilated Convolution," *International Journal of Engineering and Advanced Technology*, vol. 8, no. 6, pp. 641-647, 2019. DOI:10.35940/ijeat.F8089.088619

[32] Sami M., Naiyer I., Khan E., and Uddin J., "Improved Semantic Inpainting Architecture Augmented with a Facial Landmark Detector," *The International Arab Journal of Information Technology*, vol. 19, no. 3, pp. 353-362, 2022. https://doi.org/10.34028/iajit/19/3/9

[33] Schmalfuss J., Scheurer E., Zhao H., Karantzas N., Bruhn A., and Labate D., "Blind Image Inpainting with Sparse Directional Filter Dictionaries for Lightweight CNNs," *Journal of Mathematical Imaging and Vision*, vol. 65, no. 2, pp. 323-339, 2023. https://doi.org/10.1007/s10851-022-01119-6

[34] Sogancioglu E., Hu S., Belli D., and Van Ginneken, B., "Chest X-ray Inpainting with Deep Generative Models," *arXiv Preprint*, arXiv:1809.01471v1, 2018.

[35] Sujaykhandekar., "Removing Objects from Pictures with Deep Learning," Medium, https://medium.com/analytics-vidhya/removing-objects-from-pictures-with-deep-learning, Last

Visited, 2023.

[36] Susan M., Subashini P., and Krishnaveni M., "Comparison of Various Deep Learning Inpainting Methods in Smart Colposcopy Images," *International Journal of Computational Intelligence Studies*, vol. 11, no. 1, pp. 53-72, 2022. https://doi.org/10.1504/ijcistudies.2022.123347

[37] Wang X., Li P., DU Y., Lv Y., and Chen Y., "Detection and Inpainting of Specular Reflection in Colposcopic Images with Exemplar-based Method," *in Proceedings of the IEEE 13th International Conference on Anti-counterfeiting, Security, and Identification (ASID)*, Xiamen, pp. 90-94, 2019. DOI:10.1109/ICASID.2019.8925202

[38] Wang Y., Chen Y., Tao X., and Jia J., "VCNet: A Robust Approach to Blind Image Inpainting," *in Proceedings of the 16th European Conference on Computer Vision ECC*, Glasgow, pp. 752-768, 2020. https://doi.org/10.1007/978-3-030-58595-2_45

[39] Wei R. and Wu Y., "Image Inpainting via Context Discriminator and U-Net," *Mathematical Problems in Engineering*, vol. 2022, no. 1, pp.1-12, 2022. DOI:10.1155/2022/7328045

[40] Wu Y., Singh V., and Kapoor A., "From Image to Video Face Inpainting: Spatial-Temporal Nested GAN (STN-GAN) for Usability Recovery," *in Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, Snowmass, pp. 2385-2394, 2020. DOI:10.1109/WACV45572.2020.9093473

[41] Xu H., Li Q., and Chen J., "Highlight Removal from A Single Grayscale Image Using Attentive GAN," *Applied Artificial Intelligence*, vol. 36, no. 1, pp. 1-19, 2022. DOI:10.1080/08839514.2021.1988441

[42] Yan S. and Zhang X., "PCNet: Partial Convolution Attention Mechanism for Image Inpainting," *International Journal of Computers and Applications*, vol. 44, no. 8, pp. 738-745, 2022. https://doi.org/10.1080/08839514.2021.1988441

[43] Yan Z., Li X., Li M., Zuo W., and Shan S., "Shift-Net: Image Inpainting via Deep Feature Rearrangement," *in Proceedings of the 15th European Conference on Computer Vision ECCV*, Munich, pp. 3-19, 2018. https://doi.org/10.1007/978-3-030-01264-91

[44] Yeh R., Chen C., Lim T., Schwing A., Hasegawa-Johnson M., and Do M., "Semantic Image Inpainting with Deep Generative Models," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Honolulu, pp. 6882-6890, 2017. DOI:10.1109/CVPR.2017.728

[45] Yu J., Lin Z., Yang J., Shen X., Lu X., and Huang T., "Free-Form Image Inpainting with Gated Convolution," *in Proceedings of the IEEE/CVF International Conference on Computer Vision*, Seoul, pp. 4470-4479, 2019. DOI:10.1109/ICCV.2019.00457

[46] Yuan L., Huang Y., and Maier A., "Projection Inpainting Using Partial Convolution for Metal Artifact Reduction," *arXiv Preprint*, arXiv,abs/2005.00762, 2020. https://api.semanticscholar.org/CorpusID:218486892

[47] Zeng Y., Fu J., Chao H., and Guo B., "Learning Pyramid-Context Encoder Network for High-Quality Image Inpainting," *in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, pp. 1486-1464. DOI Bookmark:10.1109/CVPR.2019.00158

[48] Zeng Y., Fu J., Chao H., and Guo B., "Aggregated Contextual Transformations for High-Resolution Image Inpainting," *IEEE Transaction on Visualization and Computer Graphics*, vol. 29, no. 7, pp. 3266-3280, 2023. DOI:10.1109/TVCG.2022.3156949

[49] Zhang C. and Wang T., "Image Inpainting Using Double Discriminator Generative Adversarial Networks," *in Proceedings of the 2nd International Seminar on Computer Science and Engineering Technology*, Shanghai, pp. 1-6, 2020. DOI:10.1088/1742-6596/1732/1/012052

[50] Zhang X., Wang X., Kong B., Shi C., Yin Y., Song Q., and Lyu S., "High Quality Image-based Face Inpainting using Domain Embedded Multi-model Generative Adversarial Networks," *Journal of Latex Class Files*, vol. 14, no. 8, pp. 1-11, 2015. https://arxiv.org/pdf/2002.02909.pdf

[51] Zheng H., Lin Z., Lu J., Cohen S., Shechtman E., Barnes C., Zhang J., Xu N., Amirghodsi S., and Luo J., "CM-GAN: Image Inpainting with Cascaded Modulation GAN and Object-Aware Training," *arXiv Preprint*, arXiv:2203.11947, 2022. https://doi.org/10.48550/arXiv.2203.11947

[52] Zhou T., Johnson B., and Li R., "Patch-based Texture Synthesis for Image Inpainting," *arXiv Preprint, arXiv, abs/1605.01576,* 2016. https://arxiv.org/pdf/1605.01576.pdf

[53] Zhou X., Dai T., Jiang Y., and Xia S., "Bishift-Net for Image Inpainting," *in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Toronto, pp. 2470-2474, 2021. DOI:10.1109/ICASSP39728.2021.9413380

Jennyfer Susan is a Research Scholar in Department of Computer Science at Centre for Machine Learning and Intelligence at Avinashilingam Institute for Home Science and Higher Education for Women, Tamil Nadu, and India. The author has published 2 journals, 1 national conference and 2 international conferences, 1 book chapter and 2 books publication. The research area of interest includes Vision AI, Medical Image Processing and Neural Networks.

Parthasarathy Subashini is the coordinator of the Centre for Machine Learning and Intelligence sanctioned by Department of Science and Technology, Avinashilingam University for Women, Tamil Nadu India. She has authored and co-authored 4 books, 4 book chapters, 1 monograph, 145 research papers both at international and national level. She has ten sponsored research projects of worth more than 2.54 crores from various Government funding agencies. She also extended her contribution towards various international collaborations with universities from USA, Germany and Morocco.