International Journal of Engineering Research And Management (IJERM)
ISSN: 2349- 2058, Volume-08, Issue-10, October 2021

A Review of Image Inpainting

Gang Wei

Abstract— **As for actual content in the field of computer vision, image inpainting has increasingly become a research hotspot in recent years. Image inpainting is the task of restoring the lost image content according to the known image content. It has a broad application value in image editing, film and television stunt production, virtual reality, and digital cultural heritage protection. To make more researchers explore the theory and development of image restoration based on deep learning, this paper summarizes the research status in this field. Firstly, starting from the traditional image inpainting methods, this paper analyzes their existing problems, focuses on the overview of image inpainting methods based on deep learning, including image inpainting methods based on convolution neural network, generation countermeasure network, and cyclic neural network, introduces the principle and structure of various methods, and summarizes the application scope, advantages and disadvantages of various methods based on deep learning. Finally, the future research direction and focus have been prospected.**

Index Terms— **Image Inpainting, Computer vision, Deep Learning**

INTRODUCTION

Traditional image inpainting techniques can be divided into two categories: structure and texture-based image inpainting techniques. According to the similarity of image content and texture consistency, the traditional image inpainting adopts the method based on mathematical and physical theory to complete the small area damaged image restoration by establishing a geometric model or texture synthesis. At the beginning of the 21st century, betalmio [1] and others first proposed the image inpainting algorithm BSCB model, which realizes image inpainting by gradually diffusing the area's boundary to be repaired inside. The overall structure of the image in the BSCB model determines the repair result of the image. The missing image is divided by the edge line. The corresponding color is filled according to the boundary of the missing image area to generate the repair information. This method is suitable for repairing images without texture structure, and the structure is apparent. Still, the repaired areas tend to be blurred, so there is no way to meet the requirements of subjective vision for image connectivity. Texture-based image inpainting technology can realize the repair of the large damaged area of the image. Criminisi et al. [2] proposed block-based image inpainting technology, which achieves image inpainting by finding the optimal target block and copying the pixels to the area to be filled. Although this kind of image inpainting method effectively maintains the consistency of image texture, it ignores the image edge structure, can not repair it effectively and needs a lot of time. Due to the lack of image understanding and perception like human beings, there are often problems such as fuzzy content and semantic loss in restoring missing images in large areas.

In recent years, machine learning technology represented by deep learning technology has made a qualitative leap and has made a series of outstanding achievements in many research fields. With the extensive research of deep learning in academia and industry, its application advantages in image semantic extraction, feature representation, and image generation are becoming increasingly prominent, which makes the research of image inpainting method based on deep learning become a research hotspot at home and abroad and get more and more attention. Among them, convolutional neural networks (CNN), as a feedforward depth network, has strong ability in image feature learning and expression and excellent performance in large-scale image processing [3-6]. Researchers use convolutional neural networks to train data and efficiently predict the structure of images. However, the texture details of image inpainting are not satisfactory. On the other hand, Goodfellow [7] proposed generating a countermeasure network in 2014, composed of a generator and discriminator. The generator synthesizes the data from the given noise, and the discriminator distinguishes the similarity between the synthetic data and the actual data. If the calculated sample image is similar enough to the image of the area to be repaired, the purpose of image repair can be achieved. Because the generation of confrontation network has the great potential of ingenious game confrontation learning mechanisms and fitting data distribution, it has also been widely used in computer vision. These research results exceptionally make up for the shortcomings of traditional methods in image semantic understanding in image vision tasks, solve the semantic gap between image bottom features and high-level semantics to a certain extent, and make deep learning technology gradually occupy the forefront of the field of computer vision [8]. Image inpainting based on deep learning has set off a research upsurge and achieved remarkable results.

Image inpainting model based on convolutional neural network. Self-Encoder is a neural network composed of the input layer, hidden layer, and output layer. It belongs to the category of unsupervised learning. At first, it was mainly used for feature extraction and data dimensionality reduction. In 2006, Hinton et al. [9] proposed the Deep Auto-Encoder (DAE) network by improving the structure of the traditional self-encoder. They achieved effective data dimensionality reduction by training the multi-layer neural network to reconstruct the high-dimensional input vector and convert it into low-dimensional code. The Convolutional Auto Encoder (CAE) proposed by Masci et al. [10] in 2011 uses convolution layer and pooling layer to replace the entire connection layer in the traditional self encoder, uses convolution and pooling operation of convolution neural network to extract

Manuscript received Oct 18, 2021

Gang Wei, Master's degree in TianGong University. Research area: image processing

unsupervised features, and effectively solves the difficulty of supervised learning in image inpainting by CNN. Pathak et al. [11] proposed a network named context encoder for image inpainting, which is similar to self coding network and is also a network of the coding-decoding process. This network structure is also widely used in image style migration, image generation, image super-resolution reconstruction, image coloring, segmentation, video-based prediction, and other video image processing. The first half is a series of layer by layer down-sampling processing, and the second half is an inverse operation similar to the first half, that is, gradually reduce the image scale in the encoding process, gradually increase the image scale in the decoding process, and finally form an "hourglass" like network structure.

Image inpainting model based on generative adversarial networks. Since Goodfellow et al. [7] proposed a framework of generating a model through the confrontation process in October 2014, GAN has become one of the most promising methods in unsupervised learning complex distribution in recent years. Because the training process of the GAN image inpainting method is unstable, to improve the stability of training, Mirza proposed CGAN. By changing the unsupervised GAN into a semi-supervised or supervised model, increasing the constraint of network structure, introducing another conditional variable y, and merging the original input of GAN into a vector, the stability of the training process has been dramatically improved. Researchers are still improving the GAN model in recent years, and a series of derivative models such as DCGAN and WGAN has been born. The image inpainting method based on the generation model refers to the inpainting method that uses the powerful image generation ability of the trained generation model to infer the unknown distribution based on the known prior distribution of the damaged image. The representative generation models mainly include the autoregressive model [12], VAE, and GAN. The pixel RNN model based on the autoregressive model proposed by the Google team in 2016 [12] has achieved good results in image inpainting tasks. The model is essentially an improved two-dimensional cyclic neural network. By learning the discrete probability distribution of the image, the pixels in the two-dimensional image are predicted point by point for image inpainting. Bao et al. [13] proposed a generation network composed of the encoder, generator, classifier, and discriminator called CVAE-GAN. They tried to apply it to the image inpainting task and achieved some inpainting results. Zheng et al. [14] effectively combined VAE and GAN and proposed a diversity image inpainting model including inpainting network and generation network in parallel. Zhao et al. [15] regarded the diversity image inpainting task as a known edge probability distribution. Joint probability distribution solved the problem of the conditional probability distribution and proposed an unsupervised cross-space generation model. The GAN network is used to repair the image. In the training stage, the damage of the image is not considered. That is, the undamaged data is used for GAN network training in the training process. When the training is completed, the generation model G can generate a new image from the noise. Then the repair of a damaged image can be transformed into generating a new sample similar to the known part of the original image through the generation model G. To achieve this purpose, the newly generated image needs to be iteratively modified. Compared with the image inpainting method based on self-coding, the method based on the generation model can realize diversity in inpainting. However, due to the problems of unstable training and prone to mode collapse, such methods can only deal with low-resolution images.

With the advent of big data, the improvement of hardware computing power, and urgent application needs, image inpainting has attracted extensive attention from academia and industry at home and abroad and has become an important and challenging research topic in the field of computer vision. By summarizing the two types of image inpainting methods, this paper summarizes the characteristics of several image inpainting methods: the traditional image inpainting methods can achieve good inpainting results when the missing area of the image to be repaired is small, and the structure and texture are relatively simple. However, in the face of more complex image inpainting tasks, due to the lack of understanding and perception of the high-level semantics of the image, it is impossible to fill the missing areas with semantically consistent and reasonable content. CNN is widely studied, but there are deficiencies in texture repair; GAN can be applied to image inpainting without a large amount of data, but the instability in the training stage of GAN needs more in-depth research.

REFERENCES

[1] BERTALMIO M , SAPIRO G , CASELLES V , et al. Image inpainting [C]. International Conference on Computer Graphics and Interactive Techniques , 2000:417-424.

[2] CRIMINISI A , PEREZ P , TOYAMA K , et al. Region filling and object removal by exemplar-based image inpainting[J]. IEEE Transactions on Image Processing , 2004 , 13 (9): 1200-1212.

[3] KRIZHEVSKY A , SUTSKEVER I , HINTON G E. Imagenet classification with deep convolutional neural networks[J]. Communications of the ACM , 2017 , 60(6):84-90.

[4] SIMONYAN K , ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[J]. ar Xiv:1409.1556 , 2014.

[5] SZEGEDY C , LIU W , JIA Y , et al. Going deper with convolutions[C]‖Procedings of the IEEE Conference on Computer Vision and Patern Recognition.2015:1-9.

[6] HE K , ZHANG X , REN S , et al. Dep residual learning for image recognition[C]‖ Procedings of the IEEE Conference on Computer Vision and Patern Recognition.2016:770-778.

[7] GAN QIANG Z P , HE L B , CHEN X , et al. Survey on dep learning image inpainting methods[J]. Journal of Image and Graphics , 2019 , 24(3):447-463.

[8] HINTON G E , OSINDERO S , TEH Y W. A fast learning algorithm for dep belief nets[J]. Neural computation , 2006 , 18(7): 1527-1554.

[9] MASCI J , MEIER U , CIREAN D , et al. Stacked convolutional auto-encoders for hierarchical feature

extraction[C]‖International Conference on Artificial Neural Networks.Springer , Ber- lin , Heidelberg , 2011:52-59.

[10] Pathak D , Krahenbühl P , Donahue J , et al. Context encoders: feature learning by inpainting[C]/ / Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas , NV , USA: IEEE , 2016: 2536-2544. [DOI: 10. 1109 / CVPR.2016.278]

[11] OORD A , KALCHBRENNER N , KAVUKCUOGLU K.Pixel recurent neural networks[J].arXiv:1601.06759 , 2016.

[12] BAO J , CHEN D , WEN F , et al.CVAE-GAN:fine-grained im- age generation through asymmetric training[C]‖Procedings of the IEEE International Conference on Computer Vision.2017: 2745-2754.

[13] ZHENG C , CHAM T J , CAI J. Pluralistic image completion [C]‖Procedings of the IEEE Conference on Computer Vision and Patern Recognition.2019:1438-1447.

[14] ZHAO L , MO Q , LIN S , et al. UCTGAN: Diverse Image Inpainting Based on Unsupervised Cros-Space Translation[C]‖ Procedings of the IEEE/CVF Conference on Computer Vision and Patern Recognition.2020:5741-5750.