

Research Article

Artificial Intelligence-Assisted Fresco Restoration with Multiscale Line Drawing Generation

Guanghui Song and Hai Wang 

School of Architectural & Artistic Design, Henan Polytechnic University, Jiaozuo, Henan 454000, China

Correspondence should be addressed to Hai Wang; wanghai@hpu.edu.cn

Received 21 January 2021; Revised 3 February 2021; Accepted 3 March 2021; Published 12 March 2021

Academic Editor: Wei Wang

Copyright © 2021 Guanghui Song and Hai Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this article, we study the mural restoration work based on artificial intelligence-assisted multiscale trace generation. Firstly, we convert the fresco images to colour space to obtain the luminance and chromaticity component images; then we process each component image to enhance the edges of the exfoliated region using high and low hat operations; then we construct a multistructure morphological filter to smooth the noise of the image. Finally, the fused mask image is fused with the original mural to obtain the final calibration result. The fresco is converted to HSV colour space, and chromaticity, saturation, and luminance features are introduced; then the confidence term and data term are used to determine the priority of shedding boundary points; then a new block matching criterion is defined, and the best matching block is obtained to replace the block to be repaired based on the structural similarity between the block to be repaired and the matching block by global search; finally, the restoration result is converted to RGB colour space to obtain the final restoration result. An improved generative adversarial network structure is proposed to address the shortcomings of the existing network structure in mural defect restoration, and the effectiveness of the improved modules of the network is verified. Compared with the existing mural restoration algorithms on the test data experimentally verified, the peak signal-to-noise ratio (PSNR) score is improved by 4% and the structural similarity (SSIM) score is improved by 2%.

1. Introduction

Computer graphics and computer vision have gradually come into the limelight as their performance has developed rapidly. Nowadays, computer graphics and computer vision have become an integral part of computer development, and image segmentation, image enhancement, image recognition, and image restoration have always played an important role in these developments [1]. Among them, image restoration, an ancient problem dating back to the European Renaissance, is the study of repairing information in damaged areas of an image using manual or related algorithms based on information about the undamaged areas of the image, so that the damaged areas can be restored or their original information can be restored to some extent, and so that the restored image appears continuous and complete to the viewer [2]. Digital image restoration techniques are

currently used in many fields, such as photo restoration, target removal from images and videos, medical imaging, and conservation and restoration of cultural objects [3]. The application of image restoration technology in the field of heritage conservation and restoration has not only improved the efficiency of restoration but also avoided to a certain extent the damage caused by human misuse to the heritage [4]. Therefore, digital image restoration technology in the conservation and restoration of cultural relics is increasingly and widely used, and it has very important research significance [5]. Frescoes in the record of history at the same time also by time in its body carved traces of history, frescoes in the millenniums of years through weathering, rain, and snow, and other natural factors such as erosion, earthquakes, man-made, and other external damage, this witness to the history of mankind's bright pearl becomes very fragile, and some have been a considerable degree of defective [6].

Therefore, the conservation and restoration of frescoes have become increasingly urgent [7]. At present, the restoration of frescoes is mostly done manually, and professionals with expertise in history and art can use their professional skills to restore the textures to the overall structural information of the frescoes [8]. However, the premise is that restorers need to have solid and comprehensive skills or knowledge of history, humanities, fine arts, and archaeology, and these diverse professional needs have led to a lack of professional conservators [9]. Moreover, even with professional expertise, the conservation and restoration of cultural objects is still a time-consuming and complex process. It is these factors that severely restrict the development of fresco conservation and restoration and even cultural heritage protection.

In order to speed up the efficiency of mural restoration and save costs, it is necessary to adopt more advanced methods to study mural restoration.

After Gunning et al. first proposed computer image restoration techniques, the field of image restoration immediately received a lot of attention, followed by a large number of related or extended research works, which can be broadly classified into partial differential equation-based image restoration, texture synthesis-based image restoration, and a combination of both [10]. Checa et al. first proposed a computerized image restoration algorithm based on PDE (partial differential equation), which is called BSCB and which simulates the process of restoration by professionals by diffusing the restoration from the outside of the defective region inward along iso-illumination lines, with the help of the complete pixel information around the defective region. The missing region is repaired with the help of the complete pixel information around the defective region [11]. Later, Malik proposed an image restoration algorithm based on total variation (TV) based on BSCB, in which the image restoration problem is transformed into solving a constrained optimization problem [12]. This algorithmic model is easy to implement and maintains edge continuity, but it destroys the connectivity principle in vision theory [13]. Later, to solve the problems of the TV algorithm, Turner introduced curvature-driven diffusions (CDD) algorithm model based on curvature term and achieved good results. There are also Mumford–Shah based image restoration algorithms [14]. Papanastasiou et al. proposed an improved image restoration algorithm based on a selective adaptive interpolation algorithm combining diffusivity functions for the slow BSCB [15], and Zheng et al. proposed an image restoration algorithm based on the p-Laplace operator [16]. However, in general, most of the partial differential equation-based restoration algorithms use similar ideas to interpolation algorithms to repair defective areas, and, therefore, they can achieve good results when repairing small-scale and simple textured areas, and the repair time is positively correlated with the area of the area to be repaired [17]. When using partial differential equation-based algorithms for large-scale restoration of images with rich textures in the damaged areas, a more blurred restoration result is often obtained. The purpose of fresco image colour restoration is to restore the original colour of faded

fresco images. Although the convolutional neural network colour restoration method based on image feature similarity does not fall into an unreasonable local optimum solution due to the gradient disappearance problem, the colour content is often mismatched because of the complex structure of mural images [18]. Neural feature matching and optimal linear local model colour reduction methods are mostly used to process images with the same semantic structure; however, the semantic similarity of mural images is not strong [19]. The convolutional neural network-based image colour restoration method is to separate the content information and colour texture information in the image and then perform colour restoration and texture synthesis, but because the drawing process of mural images is extremely tedious, their content information and colour information will not be easily separated.

In this process of mural image colour restoration, the feature extraction of mural images is especially important. The colour restoration method based on global image statistical information has a good restoration effect only for images with a single global colour and is less effective for mural images with a wide variety of colours. The slope and kurtosis of global image data are adjusted by using higher-order moments, and the continuous mapping with the same probability density function makes the source image can be converted to the target image, but the colour restoration requirement cannot be achieved for the mural images with low similarity. The adaptive colour restoration method can extract the local texture information better so that the local features of the restored image can be preserved, but when the mural image has a large variety of colours or luminance, the method cannot accurately discern the number of sample blocks to be extracted, which leads to inefficient execution. In this paper, we use the maximum mean difference constraint to extract the global colour features of mural images and Markov random field constraint to extract the local features of mural images and propose a multiple-constrained convolutional neural network-based mural image colour virtual restoration method, which aims to extract the global colour information of mural images while preserving the local colour texture information of mural images. In this paper, we introduce semantic segmentation based on an expanded convolutional neural network to guide the colour restoration of mural images to avoid the problems of mismatching colour contents of mural images and poor similarity of mural images and then use a convolutional neural network to extract features of mural images and propose a mural image colour restoration method combining semantic segmentation and convolutional neural network to improve the accuracy of mural image colour restoration.

2. Analysis of Artificial Intelligence Restoration of Multiscale Line Drawings

2.1. Mural Multiscale Line Drawing Design. The observation shows that there is both texture and structure information in the detached area. This chapter proposes an improved restoration method to address the problem that the existing

sample-based image restoration algorithm produces poor structural continuity of restoration results when restoring ancient burial murals. First, linear weighting and redefinition of priority calculation are used to give higher weights to data items so that structural information is prioritized for restoration; then a new block matching criterion is defined to measure the degree of difference between sample blocks according to the structural similarity metric, and the structural information of the image is fully utilized in selecting the best matching blocks to improve the matching accuracy and achieve effective restoration of structural and texture information in the detached area [20]. The continuity of image texture and structural information is often used as one of the criteria to measure the quality of mural image restoration. Although, compared with the pixel-based image restoration algorithm, the sample-based restoration algorithm can better preserve the texture information of the image during the restoration process, the structural continuity of the restoration results is poor.

The priority is determined by both the repair block confidence term and the data term, which determine the repair order of the sample blocks. The conventional algorithm priority calculation is defined as the product of the confidence term and data term of a single-pixel point on the defect boundary, which is vulnerable to the influence of extreme pixel points. Moreover, when the confidence term gradually converges to zero in the late stage of repair, the repair priority of this pixel point also converges to zero, which in turn affects the repair order and produces wrong repair results. Therefore, to avoid the influence of extreme pixel points on the priority calculation and to prevent the error of priority calculation when the confidence level tends to zero even if the repair block data item value is large and the structure information is rich in the later stage of repair, as shown in Figure 1.

To compensate for the limitation of selecting the best matching sample block for image restoration solely from the sum of squares of grayscale differences, this paper attempts to transform the best sample block selection problem into a structural similarity metric problem by defining a new block matching criterion to select the best matching block, introducing more structural information to increase the matching accuracy. In computer vision, the application scenarios of neural networks today mainly contain target detection, image recognition, and semantic segmentation. The recognition and detection of objects in an image is target detection. Image recognition is the identification of objects in an image. Semantic segmentation is the interpretation of the above two problems in terms of pixels. Semantic segmentation is segmentation by the semantics of the image using the computer [21]. In semantic image segmentation, semantics refers to the content of the image, that is, the understanding of the content of the image, and segmentation is the separation of different objects in the image at the pixel level and the labelling of each pixel in the image. Semantic segmentation can be applied in many fields: unmanned vehicle driving, robot design, global positioning systems, and biomedical analysis. In the field of unmanned vehicle driving, semantic segmentation is the core

methodological technique. By putting images obtained from on-board cameras or LIDAR into a neural network, the computer can actively segment and classify the images, thus enabling the vehicle to avoid various obstacles such as pedestrians and vehicles.

Semantic segmentation is also widely used in robot design technology. Firstly, the camera is used to acquire images captured during robot motion; secondly, semantic segmentation of image information is performed to distinguish planar and nonplanar regions in the image and identify feature points in nonplanar regions; then, feature points in nonplanar regions and external depictions of planar regions are constructed, to build matching correspondence between feature points and planar regions in each. Based on the above correspondence, a likelihood function is established; finally, the described likelihood function is optimized to obtain the hybrid 3D map and camera motion parameters. The neural network can be trained in GPS to allow the computer to recognize and detect the ground, buildings, river roads, mountains, and so on and label each pixel in the image from the input satellite remote sensing images obtained. With the strong development of artificial intelligence, combining neural networks with biomedicine in the field of biomedical analysis is also a hot research topic. In intelligent medicine, semantic segmentation can be applied to tumour image recognition, lesion diagnosis, and so on.

There are two important concepts in transfer learning, one of which is the domain. The goal of domain adaptation is to apply what is learned in the source region to different regions, which are interrelated target regions. The actual goal is to construct a transformation function that, when applied, minimizes the difference between the values in the source and target regions, which solves the problem of how to measure the difference between the data distribution in the source and target domains, hence the use of maximum mean difference. In the maximum mean difference, if the samples of two distributions are equal in some function corresponding to the mean value of the imaging, then the two distributions can be considered as the same distribution, which is generally used to measure the similarity between the two distributions. The maximum mean difference can be calculated in an arbitrary space or the regenerative kernel Hilbert space. In the arbitrary space, by constructing a continuous set of functions in the sample space to obtain the average of the sample values of different distributions on the continuous set of functions and then make a difference between the sample averages of different distributions to get the mean difference corresponding to different distributions, from which the maximum value is then found, and this maximum value is called the maximum mean difference, which can be used as a test calculation quantity. In regenerative kernel Hilbert space, the unit ball is taken as a function set, and the maximum mean difference can be estimated by a limited number of observations. Each function set corresponds to a feature mapping and based on the feature mapping, and embedding means with constraints are defined for distribution, and the square of the maximum mean difference is obtained for each distribution under the

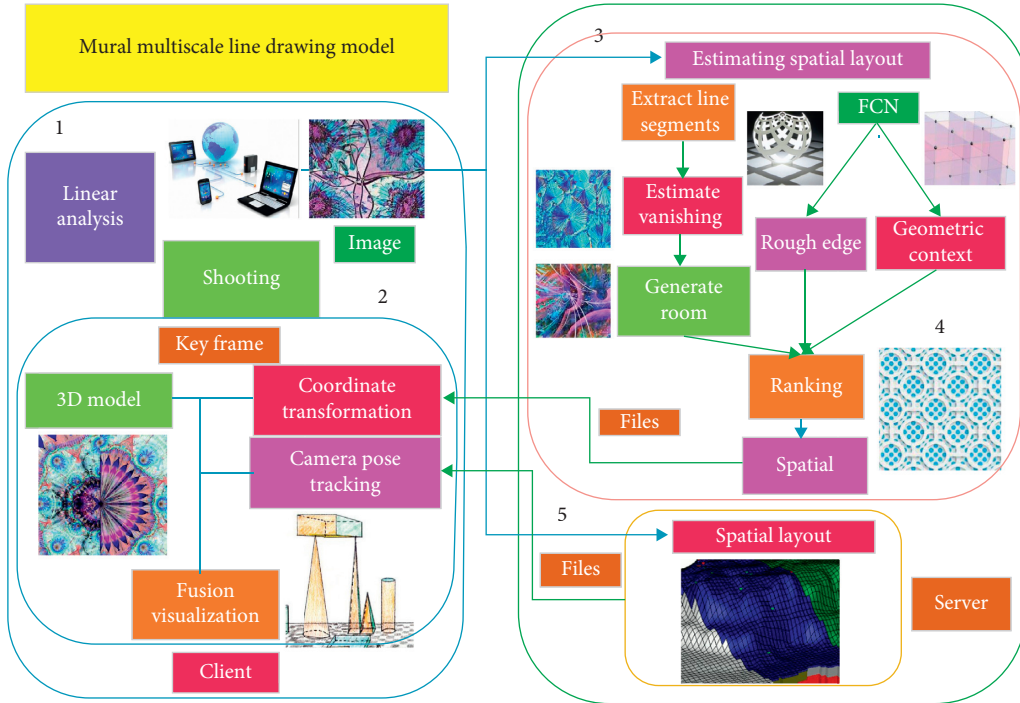


FIGURE 1: Mural multiscale line drawing model.

condition of the existence of the embedding mean. In domain adaptation, the maximum mean difference is usually used in feature learning to establish a regularization to constrain the representation so that the features of the distributions are approximately similar over the domain, as shown in Table 1 [22].

Through interactive display, the audience can enter the visual platform of the exhibition hall as a participant, and the electronic flipbook has become popular in increasingly special exhibition halls due to its interesting interactive mode. The electronic flipbook is based on various intelligent technologies and image sensing and capturing devices in one human-computer interaction, which can present the audience with the most complete state of the content to be presented in front of people, perfect visual experience, and interactive features to mobilize the audience's active sense of participation, making it easier for the audience to accept new things. The scene experience is not limited to physical interaction with users. Generally speaking, if the space of certain theme pavilions can be reasonably used with the support of technology, the scene experience can often be maximized, for example, the water curtain projection created by the scenic spot, which projects the history and culture on the water curtain wall, can vividly show several three-dimensional images, which can bring a certain degree of freshness and visual impact to the audience [23]. At present, the designers of many pavilions have introduced a variety of entertaining science and technology into the design concept of the pavilion, and the result is that many seemingly monotonous works can become more flexible. At the same time, it is possible to design interesting exhibition halls according to the preferences of different people (such as children, the elderly, college students, office workers, and

other people). At the same time, the knowledge, interaction, and information seep into the display content together. Combining the visual changes of fashion and technology in the space brings emotional resonance to the spiritual level of the audience and is a more new and effective way of interaction, as shown in Figure 2.

The open operation of the image can eliminate small protrusions, smooth the boundary, remove the convex corners, and separate two targets that are finely connected, while the closed operation can fill the small voids within the target and connect two neighbouring target regions. The crack information after threshold segmentation presents a discontinuous segmentation phenomenon, which is difficult to be distinguished from the noise of the image. Therefore, the characteristics of the closure operation can be used to make the binary image of the cracks more coherent by performing a closure operation on the segmented image. To eliminate the regions in the binary image that have a similar structure to the target information, a connected domain marking operation can be performed on them. For an image, all pixels in the target region can be uniquely labelled using this method, making the binary image convert into a labelled marked image with markers. Based on the market value, feature values such as area and circularity can be selected as market conditions according to the requirements. In this paper, the area and pixel block aspect ratio are used as market conditions to mark the connected domain for the crack binary image.

2.2. Artificial Intelligence-Aided Algorithm Design. To solve the current problems faced in mural restoration, this chapter extracts the texture, colour, and high-level structural features

TABLE 1: Sensory field size and step length [22].

Semantic attributes	Number of subclasses	Sample distribution of each subclass
Leaf	8	5
Roundness	7	4
glitch	10	3
Detail	10	6
Texture	1	4
edge	9	10
Malignancy	10	1
Internal structure	2	3
Calcification	7	5

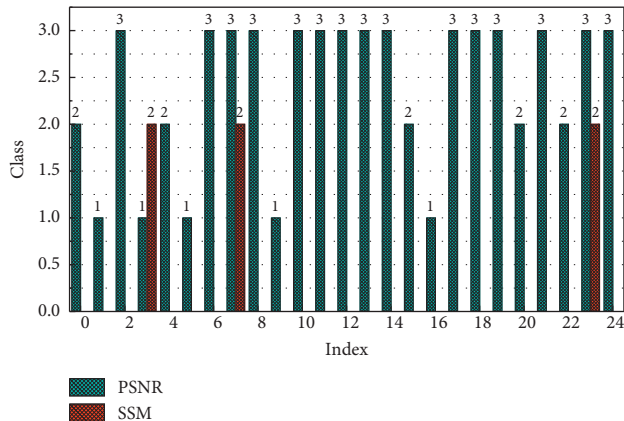


FIGURE 2: Graph of parameter versus average value.

in the mural by using the encoder part of the generator in the generative adversarial network and then generates mural defective regions based on the extracted features and hands the results of generating mural defective regions to the discriminator for discrimination; that is, the discriminator is used to supervise the generator, which makes the generator continuously optimize the generated results, and the result of generating the defective areas of the frescoes is passed to the discriminator; that is, the discriminator supervises the generator, so that the generator continuously optimizes the generated results, thus achieving the purpose of repairing the whole structure and texture of the defective areas of frescoes [24]. The network structure of this algorithm consists of a generator and a discriminator, as shown in Figure 3. The generator consists of an encoder composed of a convolutional layer and a decoder composed of a deconvolutional layer. The convolutional layer is responsible for extracting the colour, texture, and high-level structural features of the mural image, and then the decoder composed of a deconvolutional layer decodes the extracted features and generates the restoration results of the defective area.

The discriminator is responsible for judging the authenticity of the generator repair results, that is, whether the repaired results achieve the effect of falsity. True means that the discriminator considers that the result is not generated by the generator and the similarity between the generated data and the real data is maximized. False means that the discriminator considers that the result is generated by the generator and the similarity between the generated data and

the real data is much different. Through the adversarial training of the generator and the discriminator, we can finally achieve the result generated by the generator that the discriminator cannot determine the true or false; that is, the discriminator thinks that none of the results are generated by the generator, thus indicating that the generator generates data to the optimum. The discriminator consists of convolutional blocks, and the output result of the discriminator represents the judgment of the truth or falsity of the input data [25–27]. Finally, the images of the defective areas generated by the generator are filled into the missing areas of the original mural to obtain the final restored results. The network itself learns the feature extraction and adversarial learning process so that the final restoration result can recover the original colour, texture, and structure information of the mural as much as possible.

Among them, the convolution layer consists of convolution kernels of specified channels, through which convolution operations are performed on the input data, and then the gradient back-propagation algorithm (BP) is used to optimally adjust the convolution layer weights. Different convolution kernel sizes can acquire the image features in the image under different perceptual fields. Generally, a large convolution kernel can acquire the large-scale features in the image, but, at the same time, it brings more parameters and increases the computational effort, and the convolution parameters are calculated by the following:

$$ICC(A, 1) = \frac{MS_R - MS_E}{MS_R + (k - 1)MS_E}. \quad (1)$$

The input of each layer of the neural network before the introduction of the activation function is a linear function of the output of the upper layer, so it will result in a neural network whose output is a linear combination of inputs. By introducing a nonlinear activation function, the nonlinear fitting capability of the neural network can be greatly increased, allowing the neural network to fit any function:

$$ICC(C, 1) = \frac{MS_R - MS_W}{MS_R + (k - 1)MS_W}. \quad (2)$$

The role of the decoder is to decode the features mapped by the encoder into the hidden space and then restore the true data distribution. In this mural restoration task, the decoder decodes the features mapped by the encoder to the hidden space and then generates the information of the

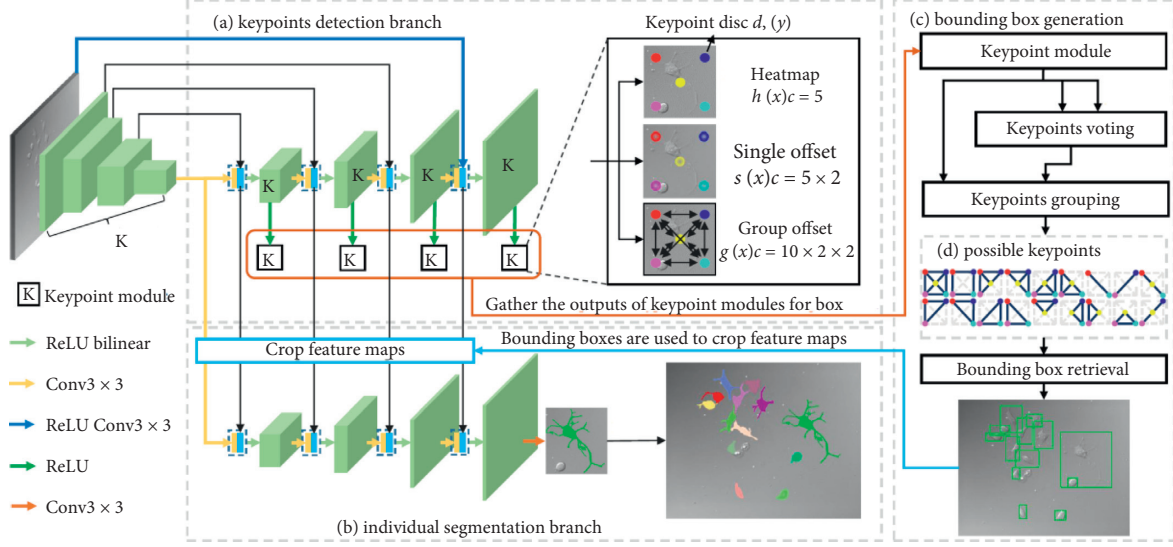


FIGURE 3: Schematic diagram of the overall architecture.

missing regions of the mural. The following formula for the deconvolution output feature size is as follows:

$$GW_{l,x,y}^k = f \left(\sum_{p=1}^{n-1} M_l^{k,p} \otimes L_{l=1}^p \right)^{-\beta}. \quad (3)$$

The entire network is trained; that is, the encoder maps the probability distribution of the input data into the hidden space and converts it into another probability distribution, a process that compresses the features of the input data. The decoder restores the data distribution of the region to be repaired based on the compressed feature information in the hidden space. During the training process, the encoder and decoder adjust their parameters according to the loss function, so that the encoder can encode and compress the features common to the data, and the decoder can restore the data distribution of the region to be repaired based on these encoded features. Therefore, the neural network structure with encoder and decoder is well suited for image restoration tasks:

$$HU_l^k = f \left(\sum_{p=1}^{n-1} J_l^{k,p} \otimes L_{l=1}^p \right). \quad (4)$$

A mapping transformation relationship is established based on the feature information in the input mural image and the feature information in the reference sample image. This process can be thought of as making the second-order statistically distributed feature representation of the input mural image as close as possible to the second-order statistically distributed feature representation of the reference sample image:

$$GW_{l,x,y}^k = f \left(\sum_{p=1}^n M_l^{k,p-1} \otimes L_{l=1}^p \right)^{-\beta}. \quad (5)$$

The core idea of Markov random field constraint is to first divide the reference mural image and the input mural image into several blocks and then find and approximate the closest colour block for each image block. This block matching-based approach has a significant highlight compared with the statistical distribution-based approach; that is, the block matching-based approach can well preserve the local structural information in the mural images:

$$H(Z)_j = \frac{e^{Z_j}}{\sum_{k=1}^k e^{Z_k}}, \quad (6)$$

$$K(t_n, t_m) = -(1 - t_m) \log(1 - t_n).$$

For the problem of retaining the structural texture information of the input mural image, this section fully considers the depth of the features extracted by the convolutional neural network and applies the image constraints to be restored to the 6th and 10th layers of the convolutional neural network model, which can effectively make the computational effort reduced and also enhance the structural texture information in the final colour restored mural image according to the input mural image:

$$M = \{M_1, M_2, \dots, M_k\}. \quad (7)$$

Regularization is often used to reduce testing errors. When building a machine learning model, the goal is to have the model perform well even when new data is introduced into the model. However, when using complex models, overfitting is more likely to occur, which can reduce the generalization ability of the network model. Using regularization can reduce this situation and make the model less complex. Common regularization methods include L1 parametrization and L2 parametrization:

$$N = \{N_1, N_2, \dots, N_k\},$$

$$H(M)_j = \frac{e^{M_i}}{\sum_{k=1}^k e^{M_k}}. \quad (8)$$

L1 regularization is achieved by adding the sum of absolute values of all feature coefficients to the objective function, which is suitable for feature selection. L2 regularization is achieved by adding the sum of squares of all feature coefficients to the objective function, which can avoid overfitting of the network model. However, it is not suitable for noise elimination, so squared gradient regularization is chosen. In the process of colour virtual restoration of mural images, some noises are easily enhanced, leading to false colours or artifacts in the restored images. So, this section adds constrained square gradient norm to the convolution layer with Markov random field constraint applied for noise suppression in the image restoration process, as shown in Figure 4.

This module is mainly responsible for the calibration of the peeling areas in the mural. It mainly includes the colour space conversion of the mural, shedding area generation, and shedding area calibration functions. Among them, the colour space conversion is to process the mural image with a colour space closer to the visual pattern of human eyes; each component image obtained after the colour space conversion is contrast stretched using morphological high and low hat operations respectively, which is used to enhance the detached edges, and then a multistructured morphological filter is constructed to smooth the noise. After morphological processing, there are obvious differences between the detached and intact regions of the mural image, and then a modified region growing algorithm is used to segment the detached regions of the mural, and the mask images of each component obtained after segmentation are fused to obtain the final mask image. The final calibration image is obtained by fusing the mask image with the original image of the mural.

It mainly includes shedding restoration and restoration result in saving functions. The mask image generated from the original mural image is input at the same time, the original mural image is converted to HSV colour space to obtain each colour feature component, followed by restoration using the structural similarity block matching algorithm mentioned, and, finally, the mural image restoration results are obtained by RGB transformation, and then the mural restoration results are saved.

3. Results and Analysis

3.1. Multiscale Line Drawing Generation Results. Data enhancement refers to the processing of the original data by image rotation, flipping, scaling, panning, cropping, contrast transformation, and so on, so that the original one image becomes multiple images and expands the sample size. Data enhancement has a great role in the training of neural networks, which adds multiple copies to a single image,

improves the utilization of the image, and effectively prevents the network from overfitting the learning of the structure of an image. Images have more redundant information and data augmentation can create different noises, and if the neural network can overcome these noises, its generalization performance is bound to be good. By doing the same operation on the edge images, the corresponding edge images can be obtained. After the above image enhancement, an enhanced training dataset can be obtained, which is 32 times larger than the original dataset. Figure 5 gives an example of an enhanced training dataset. From left to right, the original image, the image after the original image is flipped, the image after the original image is rotated by 90 degrees, the image after the original image is rotated by 90 degrees and flipped, and the corresponding edge images are shown in the bottom row.

The TV model, due to its lack of curvature constraint in diffusion, leads to severe distortion of the image edges and obvious disintegration; while the CDD model does not differ much from the restoration effect of the two models and also has the problem of broken edges; in contrast, the restoration effect of the algorithm in this paper is coherent and smooth, and the restoration results are better than the above algorithms. Its ratio of the signal represents the maximum possible power and the damage noise power affects the accuracy of its representation. Structural similarity, on the other hand, is an index used to measure the similarity between the output image and the original image after the colour restoration of the mural image. From Figure 5, it can be found that the PSNR and SSIM values obtained by the method in this paper for the virtual restoration results of the colour of the images of Buddha, jewellery, costume, and floral frescoes have a large improvement compared to the values obtained by other methods in the literature.

The defective area is first framed, and the proposed method uses a rectangular frame to frame the defective mural as in Figure 6, and then the rectangular frame is the area to be restored by inputting it into the network as in Figure 6 after data preprocessing can make the network achieve the best effect on the defective area. The comparison method of this experiment uses objective evaluation and subjective evaluation, respectively, where the objective evaluation of image quality can be divided into full-reference image quality evaluation, that is, the original mural image, which is completely known, and the restored mural image that can be evaluated based on the original mural image; semireferenced image quality evaluation, that is, having partial information of the original fresco images, comparing the restored frescoes based on this partial information of the original fresco images, and then evaluating the quality; and unreferenced image quality evaluation, that is, the original fresco images that are completely unknown and the quality of the restored frescoes which is evaluated only based on some image features of the restored frescoes.

From Figure 7, it can be seen that the TV algorithm based on the partial differential equation is unable to repair such a large area of defect, and the repair result not only cannot see the original texture but also cannot obtain even the rough structure; Figure 7 shows the results of the MCA

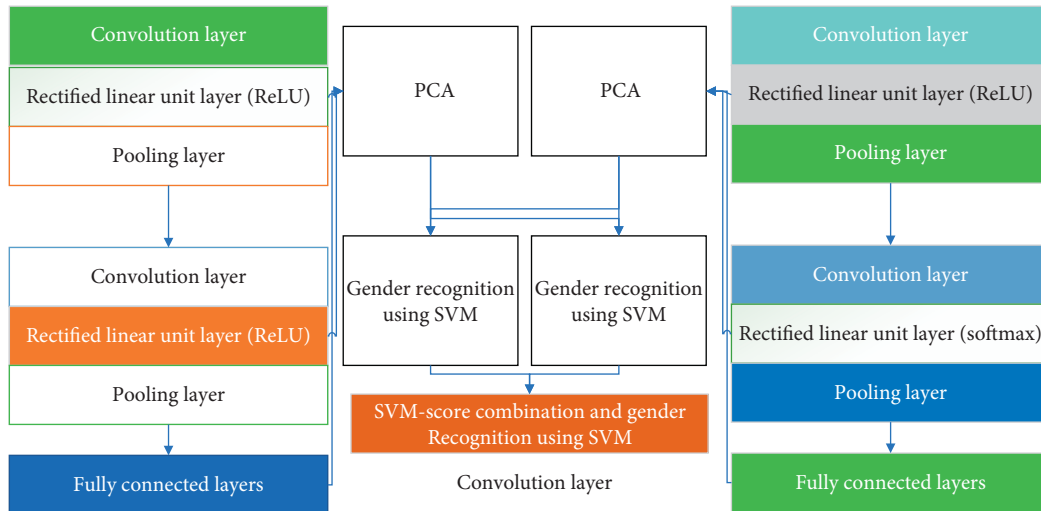


FIGURE 4: Mural off area calibration module diagram.

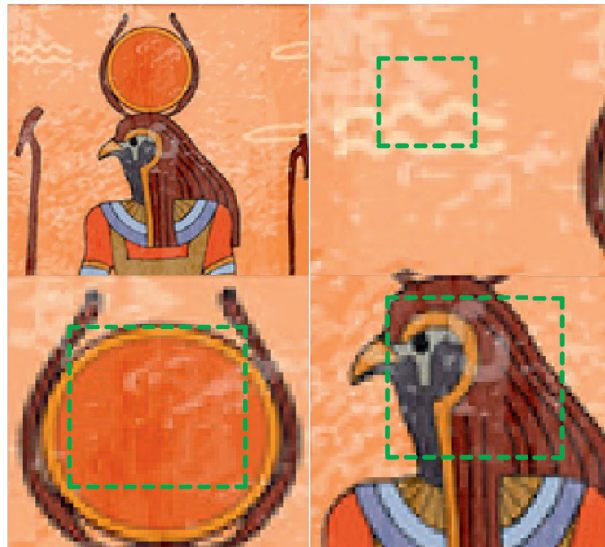


FIGURE 5: Test image.

repair algorithm based on RGB colour space that incorporates texture and partial differential equation to repair a large area of defect in the mural. From the results, the algorithm in Figure 6 is somewhat improved over Figure 7, but the original structure and texture information still cannot be seen; 7 switches the colour space from RGB to LAB and then uses the MCA algorithm for restoration. It can be seen from the figure that the restoration effect has been improved, and the texture effect of the repaired defective area is more in line with the texture characteristics of the original area, but the restoration effect in terms of structure is still not satisfactory, and no structural information can be seen at all from the restoration result, so although the algorithm can repair the texture information better, it still seems to be unable to repair the defects in the large area of the mural structure; the repair algorithm proposed in this chapter has good repair results in both structure and texture repair, and it can be seen from the figure that the repaired area has more

reasonable structure information compared with the previous algorithm. The restoration results in terms of the texture of the defective area are also not worse than those of the comparison algorithm.

From Figure 7, it can be demonstrated that the scores of the previously stated objective evaluation and the perception of the observer's human eye are not entirely consistent because the human eye is influenced by various factors when observing an image; for example, the observer's observation of one area is influenced by the adjacent areas. Therefore, the introduction of subjective evaluation in image restoration is completely necessary to better illustrate the restoration effect of fresco image restoration algorithms. The subjective evaluation in this thesis is carried out using relative evaluation, where the images restored by different algorithms are evaluated directly by the participants, and then the order of merit of the restoration quality of each image is given, which leads to the corresponding evaluation scores.

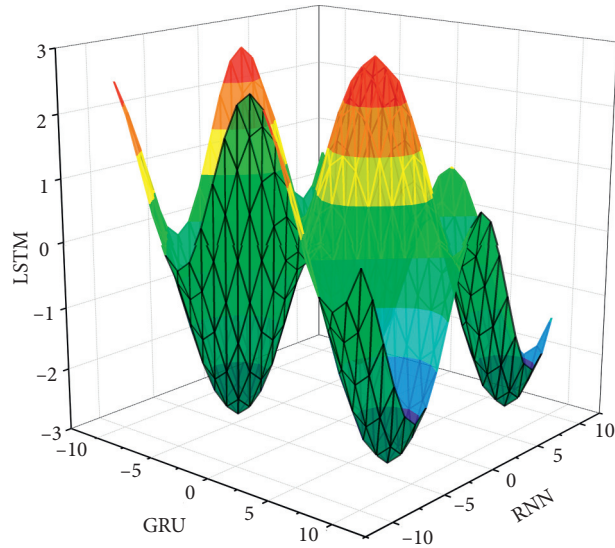


FIGURE 6: Graph of loss value change.

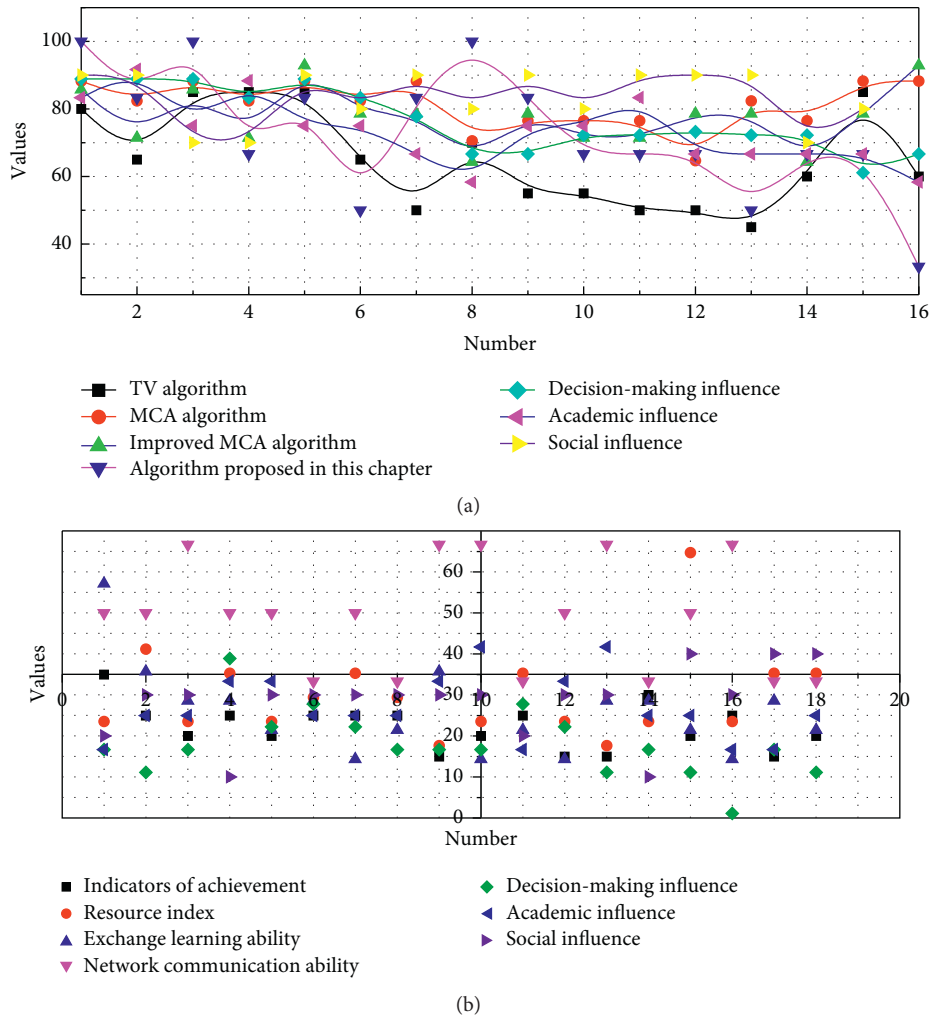


FIGURE 7: Experimental restoration of large defective areas of frescoes 1 objective evaluation score.

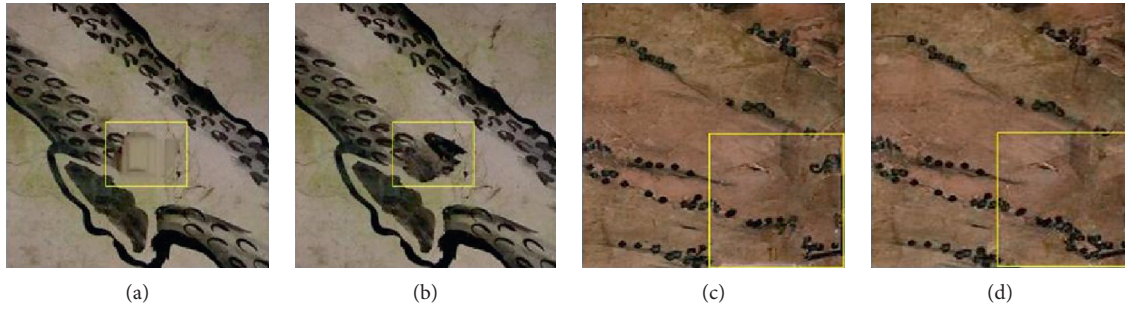


FIGURE 8: Results.

3.2. Mural Restoration Results Analysis. The effect of a different number of iterations on the colour restoration results of the mural images is investigated. In this chapter, the restoration results of this model are tested for 100, 500, 1000, 1500, 2000, 2500, and 3000 iterations. The comparative results are shown in Figure 8. From Figure 8, the mural image obtained from 100 iterations is blurred and the colour is not restored. At 500 iterations, the overall colour of the mural image obtained by the network is lighter and the overall colour is still blurred. At 1000 iterations, compared with figures (c) and (d), the blurring phenomenon is alleviated and the colour is restored, but the colour of the face is not uniformly restored, and there is a faint layer of Gray. At 1500 iterations, the colour of the face is more uniform compared with figure (a), but there is still unevenness in the colour alternation. At 2000 iterations, the overall skin fading colour of the statue is mostly reduced. At 2500 iterations, the colour of the mural image is restored, the colour is uniform, and the blurring phenomenon disappears, but the overall colour is dark. At 3000 iterations, the overall colour is brightened based on figure (b), so that the mural image is almost the same as the original image, and the colour reproduction is more realistic.

The improved joint loss function is proposed, and to demonstrate the effect of adding the local reconstruction loss function and not adding the local reconstruction loss function on the stable training of the network, this section trains the network by using the joint with and without adding the local reconstruction loss function, respectively, and then uses the trained network model to restore the mural to verify the effect of the improved loss function. Selected results from the comparison experiments are shown in Figure 9.

By comparing the experimental results and analysing them, it is concluded that the L1 distance used in the reconstruction loss function is more inclined to repair structural information, so using only reconstruction loss to train the network will make the repaired area blurred. By adding a discriminator to the network training to act as a dynamic adversarial loss, it can effectively compensate for the detailed texture part, thus solving the problem of blurring the repaired defective area by using only reconstruction loss. However, the use of adversarial loss without the restriction of reconstruction loss leads to over-complementation in terms of details, which usually results in

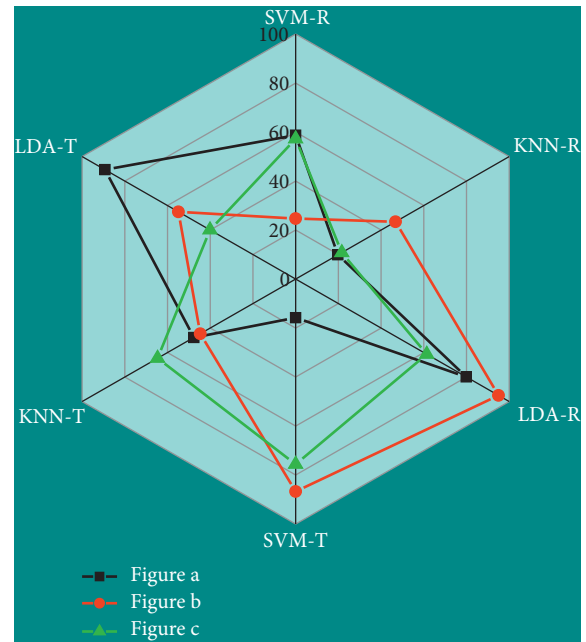


FIGURE 9: Comparison of PSNR values of restoration results by algorithms.

the loss of structural information in the repaired region. Therefore, it is necessary to add the reconstruction loss to constrain it when adding the corresponding antagonistic loss to avoid the loss of structural information due to excessive supplementation of the detailed texture part, as shown in Figure 10.

As the key part of the proposed algorithm to extract and encode the fresco features and repair the defective area based on the extracted features, the quality of the result of the whole restoration task is determined by the good or bad design of the generator structure. Therefore, the proposed algorithm in this chapter improves the whole generator from two parts, encoder, and decoder, and the following experiments are conducted to verify the effectiveness of the improvement. First, all the parts except the generator part are guaranteed to be the same, and then the models at different stages are selected to perform the repair task for the defective region separately to verify whether the improved generator has improved performance compared with the previous unimproved generator. Row 1 of the figure shows the results

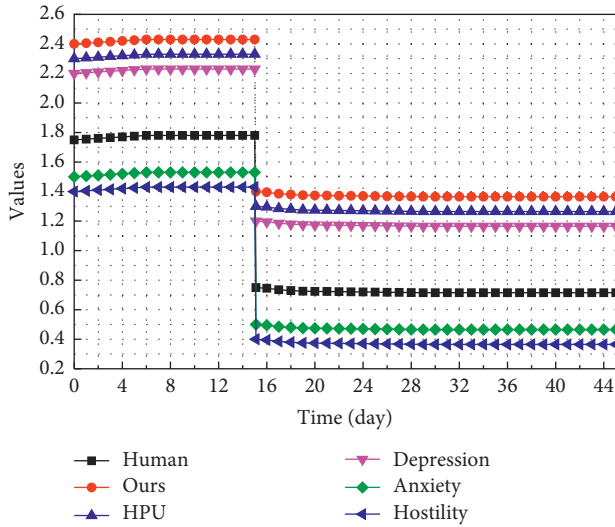


FIGURE 10: Comparison of the results of the algorithms.

of training with the unimproved generator model and performing the repair, row 2 of the figure shows the results of training with the improved generator and performing the repair using the obtained network model, and the area marked by the red box in the middle is the repaired defective area. From the results in Figure 10, it can be seen that the repaired defective area of the unimproved model lacks meaningful structural and textural features, and no meaningful information can be observed from the observer's point of view, while the repaired results using the improved generator model can show the general structural and textural features, and the repaired results using the improved generator are more natural from the observer's point of view.

When the network extracts the relevant features of the mural and restores the missing areas, the global weight of each pixel in the missing area is larger in the smaller resolution mural images, so if the current pixel is restored incorrectly, it will have a larger impact on the restoration of the global missing area information. On the other hand, the global weight of each pixel in the missing area is relatively small in the larger resolution images, so that the restoration of the global missing area information will not be affected if a pixel is restored incorrectly, so the high-resolution images will get better restoration results when restoring the same proportion of the missing area information.

4. Conclusion

The emergence of computer-aided fresco restoration methods, by virtually repairing fresco defects on a computer, has given the restoration work the ability to be retracted and to be rolled back without causing damage to the real fresco in the event of an error. The advent of computer-aided fresco restoration algorithms further frees up manpower, but, unfortunately, current fresco restoration algorithms are only suitable for repairing small defects that are scratched or contain simple textures. A dataset of 855 images for Dunhuang mural restoration was created by collecting image data of Dunhuang murals and manually selecting, adjusting,

and simulating the defects. The proposed improved algorithm addresses the problems of generators, discriminators, and joint loss functions in the current generative adversarial networks for mural restoration and validates the effectiveness of the improvements through experiments. The peak signal-to-noise ratio (PSNR) score is improved by 4% and the structural similarity (SSIM) score is improved by 2% on the test data compared to the current mural restoration algorithm. By adopting the structure of the proposed network for defect restoration, it can be applied to mural image colour restoration, providing a fully automated algorithm for mural colour restoration. The proposed generative adversarial network-based fresco restoration algorithm achieves better results in repairing defective areas of frescoes and can complete the colour restoration of faded frescoes, which solves the problem that the current traditional algorithm cannot reasonably repair large defective areas and restore the colour of faded frescoes as a whole.

Data Availability

No datasets were generated or analysed during the current study.

Consent

All authors approved the publication of the paper.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] L. Ge, H. Liang, J. Yuan et al., "Real-time 3D hand pose estimation with 3D convolutional neural networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 41, no. 4, pp. 956–970, 2018.
- [2] W. Xu, A. Chatterjee, M. Zollhofer et al., "Mo2Cap2: real-time mobile 3D motion capture with a cap-mounted fisheye camera," *IEEE Transactions on Visualization and Computer Graphics*, vol. 25, no. 5, pp. 2093–2101, 2019.
- [3] R. Thilagar and R. Sivaramakrishnan, "Fuzzy neuro-genetic approach for feature selection and image classification in augmented reality systems," *International Journal of Robotics and Automation (IJRA)*, vol. 8, no. 3, pp. 194–204, 2019.
- [4] D. Mehta, S. Sridhar, O. Sotnychenko et al., "VNect: real-time 3D human pose estimation with a single RGB camera," *ACM Transactions on Graphics*, vol. 36, no. 4, pp. 1–14, 2017.
- [5] Z. Yao, Y. Liu, Z. Ji, Q. Sun, P. Lasang, and S. Shen, "3D driver pose estimation based on joint 2D-3D network," *IET Computer Vision*, vol. 14, no. 3, pp. 84–91, 2020.
- [6] H. He, G. Liu, X. Zhu, L. He, and G. Tian, "Interacting multiple model-based human pose estimation using a distributed 3D camera network," *IEEE Sensors Journal*, vol. 19, no. 22, pp. 10584–10590, 2019.
- [7] F. Guo, Z. He, S. Zhang, and X. Zhao, "Estimation of 3D human hand poses with structured pose prior," *IET Computer Vision*, vol. 13, no. 8, pp. 683–690, 2019.
- [8] C. Li, X. Sun, X. Sun, and Y. Li, "Information hiding based on augmented reality," *Mathematical Biosciences and Engineering*, vol. 16, no. 5, pp. 4777–4787, 2019.

- [9] H. A. Alhajja, S. K. Mustikovela, L. Mescheder et al., "Augmented reality meets computer vision: efficient data generation for urban driving scenes," *International Journal of Computer Vision*, vol. 126, no. 9, pp. 961–972, 2018.
- [10] Q. H. Gao, T. R. Wan, W. Tang, and L. Chen, "Object registration in semi-cluttered and partial-occluded scenes for augmented reality," *Multimedia Tools and Applications*, vol. 78, no. 11, pp. 15079–15099, 2019.
- [11] S. Kim, S. Kim, D. Lee, and B. Ko, "Depth-map estimation using combination of global deep network and local deep random forest," *Electronic Imaging*, vol. 2019, no. 16, pp. 4-1-4-5, 2019.
- [12] J. Hetherington, V. Lessoway, V. Gunka, P. Abolmaesumi, and R. Rohling, "SLIDE: automatic spine level identification system using a deep convolutional neural network," *International Journal of Computer Assisted Radiology and Surgery*, vol. 12, no. 7, pp. 1189–1198, 2017.
- [13] J. Kim, H. Jung, M. Kang, and K. Chung, "3D human-gesture interface for fighting games using motion recognition sensor," *Wireless Personal Communications*, vol. 89, no. 3, pp. 927–940, 2016.
- [14] T. Nishio, H. Okamoto, K. Nakashima et al., "Proactive received power prediction using machine learning and depth images for mmwave networks," *IEEE Journal on Selected Areas in Communications*, vol. 37, no. 11, pp. 2413–2427, 2019.
- [15] Z. Hu, Y. Hu, B. Wu, J. Liu, D. Han, and T. Kurfess, "Hand pose estimation with multi-scale network," *Applied Intelligence*, vol. 48, no. 8, pp. 2501–2515, 2018.
- [16] R. Davies, I. Wilson, and A. Ware, "Stereoscopic human detection in a natural environment," *Annals of Emerging Technologies in Computing*, vol. 2, no. 2, pp. 15–23, 2018.
- [17] H. Tong, Q. Wan, A. Kaszowska, K. Panetta, H. A. Taylor, and S. Agaian, "ARFurniture: augmented reality interior decoration style colorization," *Electronic Imaging*, vol. 2019, no. 2, pp. 175-1-175-9, 2019.
- [18] E. Togootogtokh, T. K. Shih, W. G. C. W. Kumara, S.-J. Wu, S.-W. Sun, and H.-H. Chang, "3D finger tracking and recognition image processing for real-time music playing with depth sensors," *Multimedia Tools and Applications*, vol. 77, no. 8, pp. 9233–9248, 2018.
- [19] W. G. C. W. Kumara, S.-H. Yen, H.-H. Hsu, T. K. Shih, W.-C. Chang, and E. Togootogtokh, "Real-time 3D human objects rendering based on multiple camera details," *Multimedia Tools and Applications*, vol. 76, no. 9, pp. 11687–11713, 2017.
- [20] L. Ge, H. Liang, J. Yuan, and D. Thalmann, "Robust 3D hand pose estimation from single depth images using multi-view CNNs," *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4422–4436, 2018.
- [21] V. T. Hoang and K. H. Jo, "3-D human pose estimation using cascade of multiple neural networks," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 4, pp. 2064–2072, 2018.
- [22] H.-J. Kim and B.-H. Kim, "Implementation of young children English education system by AR type based on P2P network service model," *Peer-to-Peer Networking and Applications*, vol. 11, no. 6, pp. 1252–1264, 2018.
- [23] A. O. Alghabri, F. H. Osman, and N. Y. Ahmed, "FPGA-based real time hand gesture and AR marker recognition and tracking for multi augmented reality applications," *Arab Journal of Nuclear Sciences and Applications*, vol. 50, no. 3, pp. 66–76, 2017.
- [24] V. Albu, "Measuring customer behavior with deep convolutional neural networks," *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, vol. 7, no. 1, pp. 74–79, 2016.
- [25] F. Rameau, H. Ha, K. Joo, J. Choi, K. Park, and I. S. Kweon, "A real-time augmented reality system to see-through cars," *IEEE Transactions on Visualization and Computer Graphics*, vol. 22, no. 11, pp. 2395–2404, 2016.
- [26] J. Wang and Y. Zhang, "Median filtering forensics scheme for color images based on quaternion magnitude-phase CNN," *Computers, Materials & Continua*, vol. 62, no. 1, pp. 99–112, 2020.
- [27] M. B. Nejad and M. E. Shiri, "A new enhanced learning approach to automatic image classification based on Salp Swarm algorithm," *Computer Systems Science and Engineering*, vol. 34, no. 2, pp. 91–100, 2019.