International Journal of Intelligent Information Technologies
Volume 20 • Issue 1

RGBD Synergetic Model for Image Enhancement in Animation Advertisements

Xuechun Wang, Huanghe Science and Technology University, China*
Wei Jiang, Huanghe Science and Technology University, China

ABSTRACT

This paper proposes a depth image symbiosis model to solve the problem of insufficient depth image quality in animated advertising. The model uses image surface information and image edge cues as the main guidance information to obtain image symbiosis information. Research data show that the model designed in this paper performs well in convergence, and enters a stable convergence process when the number of iterations is less than 5. Its PSNR data curve has the highest position and best performance, while a composite model structure has been adopted. Compared with the unitary model, the PSNR of this model reaches 41 dB when the number of iterations reaches 5, and the convergence effect of the three-step training is also better. Finally, in practical applications, the average PSNR value of the model mentioned in this article is the highest, 37.1 dB. From this comprehensive perspective, the depth image enhancement model in this study has better comprehensive performance and can provide better image enhancement effects for animation advertising depth images.

KEYWORDS

Animated Advertising, Image Enhancement, RGBD, Symbiosis

With the continuous transformation of image technology, depth information in image scenes has become more easily available. Depth information has been effectively applied in various directions such as autonomous driving, human-computer interaction, virtual reality, and so on. Wide animation also requires depth information of images to improve the fidelity and quality of images (Islam et al., 2020; Jiang et al., 2021; Monga et al., 2021). However, at present, most depth information acquisition devices obtain relatively low resolution depth information. Information images contain much noise, and the phenomenon of information loss in edge regions is relatively serious. It is necessary to use depth image enhancement technology to enhance low quality depth images captured by existing devices. This can better apply the depth image information obtained by the depth information collection device to animation advertising, remove noise from the image, and repair missing information in the image (Benz et al., 2020; Berman et al., 2020; Cai et al., 2021).

As a deep learning algorithm, Convolutional Neural Networks (CNN) can capture information features from images through a mass of training and learning and achieve image enhancement of animated advertisements under appropriate guidance. At the same time, using software-based image enhancement algorithms such as CNN can also help image enhancement technology break through

DOI: 10.4018/IJIIT.342478
*Corresponding Author

hardware limitations and achieve higher application performance (Alhichri et al., 2021; Alhussein et al., 2020; Yu et al., 2020). Therefore, based on the image common property information, the single image enhancement network, structural edge feature extraction network, and guiding image enhancement network are proposed. The paper aims to explore the image enhancement algorithm based on the symbiosis model, realize the fine processing of animated advertising images by analyzing the interaction relationship between pixels, and provide more effective image enhancement techniques and methods for the advertising industry. The research contributes a new idea and method for enhancing animated advertising images and explores the potential application of the symbiosis model in advertising. This study aims to contribute to the development and progress of the field of image processing by expanding the possibilities of RGB + Depth Map (RGBD) image processing technology in practical applications. It also provides new ideas and directions for image processing technology in the digital era.

RELATED WORKS

Presently, research in the field of image enhancement is continuously expanding. A study by Xu et al. (2020) proposed new insights into image processing for 360 degree video and images. Due to the spherical viewing range of 360 degree videos and images themselves, they had a larger data scale compared to ordinary videos and images. Therefore, compression and high-quality restoration of image data has become a key technology. The research analyzed the visual attention model and the spherical features of 360 degree images and looked forward to the future development of this technology (Xu et al., 2020). Ngugi et al. (2021) applied image processing technology to plant disease detection. The research combined machine learning technology with image processing technology, mainly using RGB images. The quality of CNN structural images was enhanced, which is the basis for the automatic recognition and formation of disease images. Lv et al. (2021) mainly used a CNN model with multiple branches to perform image enhancement on low light level images. This model combined attention mechanisms to form an end-to-end attention guidance strategy. At the same time, the form of dual attention map was adopted to achieve image enhancement and denoising. The first attention map could distinguish between underexposed and well-lit portions of the image, and the second attention map could distinguish between real texture and noise information. This method had strong fidelity.

Song et al. (2020) proposed an image enhancement model for underwater images. The model mainly depended on two optical parameters: background light and the transmission diagram. The model first performed a priori analysis of the dark channel of the underwater image and then used scene depth and reverse saturation to achieve color and contrast improvement of the image. At the same time, to balance the color and contrast of the image, white balance was used as the main post processing method. Its design method had excellent overall performance. Zhou et al. (2020) analyzed an image processing system that combines differentiators with traditional imaging systems. This image processing method could perform low-power and high-efficiency optical simulation, laying the foundation for further adaptive applications of computer vision processing and image enhancement processing technology.

On the other hand, as a relatively mature deep learning algorithm, the application of CNN in various fields is gradually deepening. Blanchet et al. (2020) proposed a stock price change prediction model combining CNN and Long Short-Term Memory (LSTM). This model could predict the closing price of stocks on the next day through the extracted characteristic data of stock price changes and could capture the closing price of stocks at different time points. This model had the ability to capture information characteristics over time, making it very suitable for predicting stock prices (Blanchet et al., 2020). Polsinelli et al. (2020) designed an automated diagnostic model for chest tomography images of COVID-19 based on light CNN. This model could effectively distinguish the chest CT images of COVID-19 from the healthy chest CT images. The model designed in this study could perform functions with more advantageous efficiency on notebook computers without hardware acceleration and had performance advantages (Polsinelli et al., 2020). Wang et al. (2020) designed an automated learning aid tool for CNN. This tool worked closely with CNN for users to use it more easily. Meanwhile, smooth transitions across abstraction levels were utilized to enhance the relevance between low-level mathematical computation and high-level model construction. This model could effectively assist users in algorithm learning.

Sharma et al. (2020) proposed a deep learning model for plant disease detection. It used segmented image data for CNN model training and used tomato crops and spot diseases as the main model diagnosis and classification targets. The performance of this research model nearly doubled compared to similar models. Basiri et al. (2021) proposed a bidirectional convolutional cyclic neural network model in accordance with attention mechanism, combining depth CNN and LSTM. This model was applied to automated emotional analysis. It extracted past and future contextual features from temporal information in two directions and extracted local features from different parts of the information on this basis. This model could be effectively applied to both long text emotional information and short text emotional information. The above research shows that CNN can efficiently extract local information, and image enhancement requires this information extraction capability. Therefore, this study applies the improved CNN to image enhancement technology to provide a technical basis for advertising image enhancement.

In conclusion, while many technologies have been applied in the field of image enhancement, current research still has some limitations. Specifically, there is a lack of specific research on enhancing animation advertising images. The current research concentrates on the treatment of 360 degree video, plant disease detection, and low-light images, among other fields. However, research on the enhancement of animated advertising images remains relatively limited. Although some studies have proposed new methods and technologies, their practical application may be limited due to poor performance in specific scenarios or the need for additional data support. Although some studies use deep learning algorithms such as CNN, the special needs and challenges of the problem of image enhancement in animated advertisements may not fully be considered. Therefore, the study proposes building an RGBD image synthesis enhancement model based on CNN to process color, depth, and other information in animated advertisement images, improving their realism and visual effects. By analyzing the characteristics of RGBD images and constructing an appropriate synicity model, more targeted and effective solutions can be found for enhancing animated advertisement images. When combined with deep learning algorithms such as CNN, the application of the RGBD synergetic model can further improve the effectiveness of animated advertisement image enhancement, resulting in higher quality image processing and enhancement.

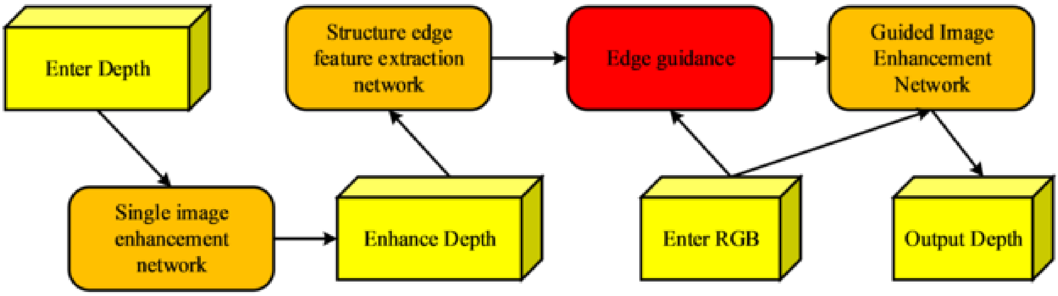
CNN-Based Design of RGBD Image Symbiosis Enhancement Model

A deep image enhancement algorithm is proposed to learn the genicity of RGBD images. The algorithm consists of three parts: the enhancement of a single depth image, the acquisition of a depth structure edge, and the enhancement of a guide depth image. First, low-quality depth images are input into a single depth image enhancement network for preprocessing. Next, the color image and the enhanced depth image are input into the previously proposed structure edge extraction network to complete the detection of the structural edge of the depth image. Then, the acquired depth image structure edge and color images are input into the guided depth image enhancement network to enhance the depth image while maintaining the syntactic relationship between RGBD images.

Architecture Design of a Symbiotic Image Enhancement Model

Research on an image symbiosis depth image enhancement algorithm based on CNN has been proposed. The reason CNN is used as the infrastructure is that it can extract image information features through numerous trainings, thereby achieving image enhancement functions under the guidance of appropriate information (Liu et al., 2020; Zhou et al., 2020). The algorithm designed in this study uses

Figure 1. Framework of Symbiotic Image Enhancement Model

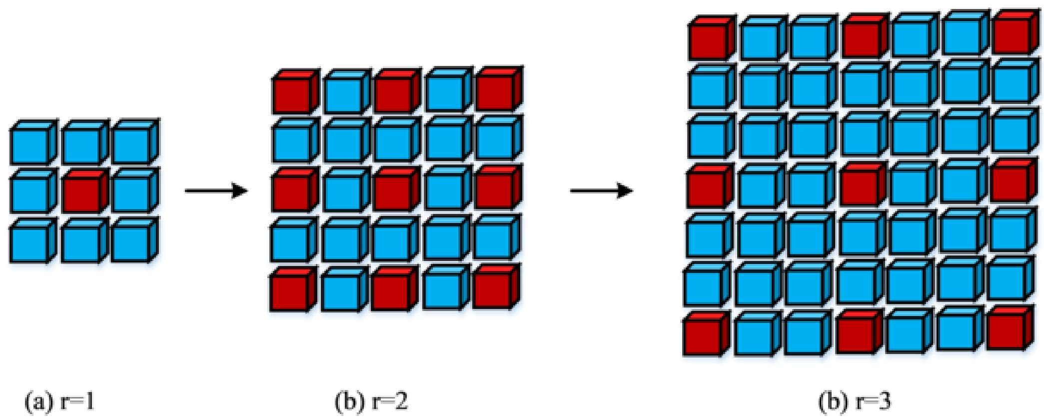


RGBD image surface information and image edge cues as the main guidance information to obtain RGBD image symbiosis information. The image edge clues themselves are the basis for symbiosis analysis set within the image gradient region. The symbiotic relationship between images is achieved by utilizing image edge clues to enhance depth in the image. RGBD image synicity information refers to the correlation and complementarity between RGB and depth images. In the field of image enhancement, RGB images provide visual information such as color and texture, while depth images provide information about the distance and depth of the objects in the scene. By combining both types of information, a better understanding of the image content can be achieved, resulting in more accurate image enhancement. Figure 1 is the specific symbiotic image enhancement model framework.

The overall framework of the image enhancement algorithm in Figure 1 shows the three main parts of the model. The first part is the Single Depth Image Enhancement Network (SDE), the second part is the Structural Edge Detection Network (SE), and the third part is the Guided Depth Enhancement Network (GDE). The SDE is mainly used for image preprocessing, which makes it easier to extract image features and provide high-quality input elements for image enhancement processing of the model. The SE is mainly used to locate the edge saliency of RGBD images, using the edge structural features of the image as the main information clues to obtain image symbiosis. GDE mainly uses the obtained image features as guidance to enhance the quality of images. Due to the relatively poor quality of depth images, the symbiotic relationship between color images and depth images is often difficult to accurately capture. This has brought some difficulties to model learning, so a SDE is proposed as a preprocessing model. During preprocessing, the image quality is processed first, and after structural edge extraction, a GDE is used for data analysis. Due to the need to preserve image details during the preprocessing process, a full convolutional network that does not cause loss of image details is studied as the main network structure. The full convolutional network structure is shown in Figure 2.

In this study, two types of network structure combination layers are used comprehensively. The first type is the convolution layer + batch standardization + modified linear element type. This type uses filters with a scale of 3*3 in the horizontal and vertical directions to generate 128 feature maps in the first four convolutional layers. The batch standardization processing layer is placed between the convolution layer and the correction linear unit, thereby effectively improving the convergence speed through normalization operations. The ability of CNN to perceive information content depends on its perceptual field, but traditional perceptual field expansion methods can easily lead to the loss of detailed features of image information. To solve this problem, this paper has selected void convolution to expand the receptive field without adding additional parameters while using jump connections to add input to output, making the network structure more able to learn residual mapping. Hole convolution can insert a convolution hole with a pixel value of 0 between individual pixels in the standard convolution. The convolution hole is shown in Eq. (1).

Figure 2. Full Convolution Network Structure



$$y(h,w)=\sum_{i=1}^{H}\sum_{j=1}^{W}x(h+r\cdot i,w+r\cdot j)\cdot f(i,j) \tag{1}$$

In Eq. (1), $x(h,w)$ represents the input signal and $y(h,w)$ represents the output signal formed after the convolution operation. $f(i,j)$ refers to the convolution filter, and r refers to the void convolution expansion rate. H represents the height of the filter, and W represents the width of the filter. In this paper, a filter with a convolutional kernel size of 3*3 is used at the last layer of the network structure, and it is more beneficial to improve the performance of the network structure without activating during operation. The network training strategy is separated into three main stages, as shown in Figure 3.

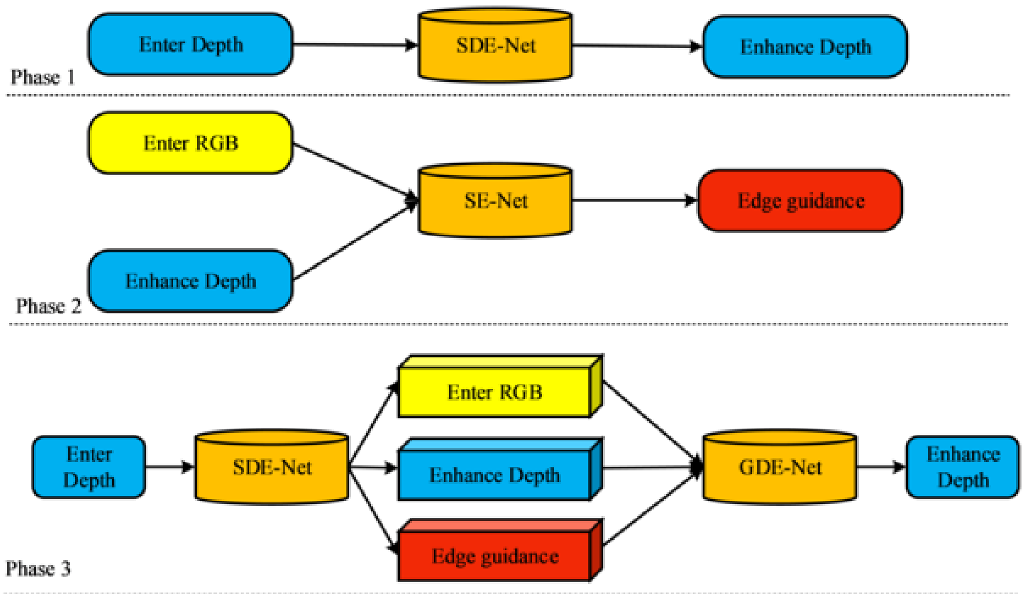
From Figure 3, targeted training is conducted for the three main parts of the network structure. By individually training a single image enhancement network, it is possible to enhance the preprocessing of low quality images during the input process. After that, the enhanced depth image and color image can be combined as training input for the structural edge feature extraction network, and this part can also be trained separately. Finally, the enhanced depth image, color image, and depth structure edge information are input into the guidance image enhancement network as training input information, and a joint training method is used to complete the final training steps. The training loss function of a single image enhancement network during training is defined as Eq. (2).

$$L_S(\Theta)=\frac{1}{2N}\sum_{i=1}^{N}\left\|F(D_i;\Theta)-D_i^{GT}\right\|^2 \tag{2}$$

In Eq. (2), F is a mapping function of the network structure. Θ is the set of parameters in the network structure. D_i represents the enhanced depth map. D_i^{GT} means a real image as a reference. N refers to the number of training samples for the model. Eq. (3) is the training loss function of the guided image enhancement network.

$$L_S(\Theta)=\frac{1}{2N}\sum_{i=1}^{N}\left\|F(C_i,S_i,D_i;\Theta)-D_i^{GT}\right\|^2 \tag{3}$$

Figure 3. Three-Stage Network Training Strategy



The G_i in Eq. (3) represents a color image of the input network structure. S_i is a structural edge information diagram.

Design of Edge Extraction Strategy for RGBD Image Structure

The symbiotic image enhancement model in this paper uses color images to guide depth images and then migrates the details of color images to depth images. However, in this process, because many methods can express the symbiosis of depth images only in a limited way, it is easy to form texture migration phenomena. Hence, this study adds a structural edge feature extraction network to the RGBD image symbiosis enhancement model. This is used to locate the edge saliency of RGBD images and provide certain edge clues for image enhancement. Supposing that a color image can be divided into two main parts: an illumination image and a reflection image. The overall image definition is Eq. (4).

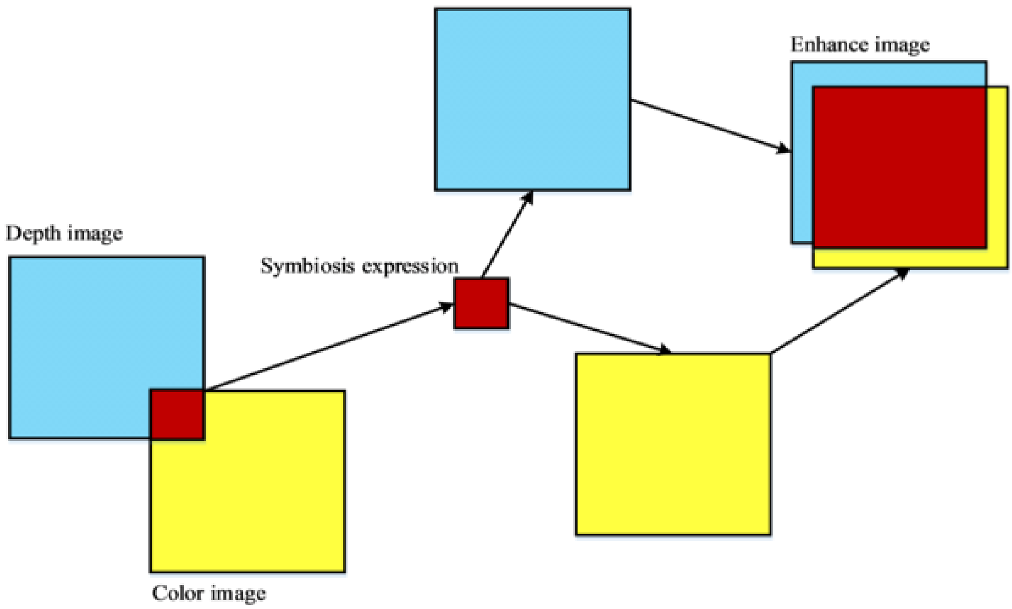
$$I\left(R,D,L\right)=S\left(N\left(D\right),L\right)\times R \tag{4}$$

In Eq. (4), I represents the RGB of a given color image, and R represents a reflection image. D is the depth image, and L refers to lighting. S is the illumination graph, and N means the surface normal graph. Generally, the lighting factor is often locally smooth or fixed in the image representation, so it can be regarded as a constant, which can form Eq. (5).

$$I_i\left(R,D,L\right)=S\left(N\left(D_i\right),L_\Omega\right)\times R_i, i\in\Omega \tag{5}$$

In Eq. (5), I_i is a color image and D_i is a depth image. There is a symbiosis between the color image and the depth image in Eq. (5), as shown in Figure 4.

Figure 4. Image Symbiosis



This symbiosis can also be referred to as the acquisition of structural symbiosis by shadow cues. Due to the uneven representation of information on the graphic surface and the presence of certain noise in the image information, it is difficult to extract the surface symbiosis between color images and depth images. Therefore, this study further conducts symbiosis analysis in the gradient domain and derives the color image factor I_i to obtain the Eq. (6).

$$\nabla I_i\left(R, D, L\right) = \nabla\left(S\left(N\left(D_i\right), L_\Omega\right) \cdot R_i\right) \tag{6}$$

Further, Eq. (7) can be derived.

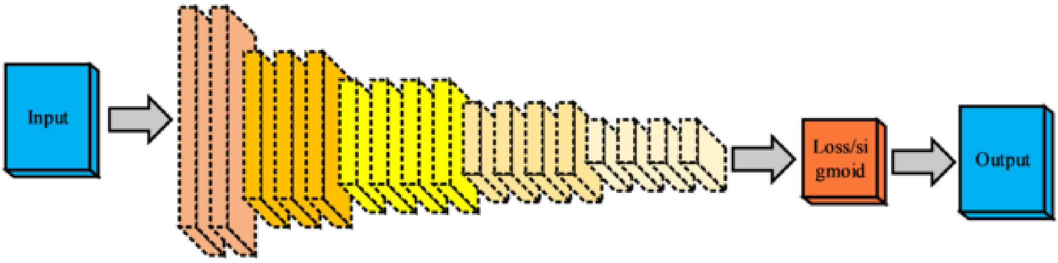
$$\nabla I_i\left(R, D, L\right) = S\left(N\left(D_i\right), L_\Omega\right) \cdot \nabla R_i + \nabla S\left(N\left(D_i\right), L_\Omega\right) \cdot R_i, i \in \Omega \tag{7}$$

In Eq. (7), I_i is a color image. ∇R_i refers to a reflected image. $S\left(N\left(D_i\right), L_\Omega\right)$ represents the gradient of the illumination map. Eq. (7) shows the determinants of gradient changes in color images. The variation is determined by the dual gradients of both reflected and irradiated images. If it is assumed that the illumination in local parts of the screen remains constant, a more simplified expression can be obtained, such as Eq. (8).

$$\nabla I_i\left(R, D, L\right) = S\left(N\left(D_i\right)\right) \cdot \nabla R_i + \nabla N\left(D_i\right) \cdot R_i, i \in \Omega \tag{8}$$

The $\nabla N\left(D_i\right)$ in Eq. (8) can be calculated as Eq. (9).

Figure 5. Edge Extraction Network Structure



$$\nabla N\left(D_i\right) = S\left(N\left(D_i\right), L_\Omega\right) \tag{9}$$

Generally, gradient changes in surface normal graphs correspond to gradient changes in depth image information. This gradient change is the cause of gradient changes in response to color images in most cases. That is, there is a certain degree of collinearity between color images and depth images. When there is no relatively large gradient change in the surface normal graph, it can be seen in the form of Eq. (10).

$$N\left(D_i\right) \approx 0 \tag{10}$$

At this time, the gradient change in the color image is mainly caused by the gradient change in the reflected image. In depth scenes with independence, this gradient change also corresponds to texture changes in color images. Therefore, color images can be used as guidance information to introduce texture noise into enhanced depth images. When detecting color structure edge clues in depth images, an edge extraction network structure grounded on the Visual Geometry Group (VGG) network framework is adopted, as shown in Figure 5.

The model performs edge prediction on the last layer of each stage in the network structure to achieve prediction and analysis of multi-scale image information. When training the model, the dataset can be recorded in the form of Eq. (11).

$$S = \left\{\left(X_n, Y_n\right), n = 1, \ldots, N\right\} \tag{11}$$

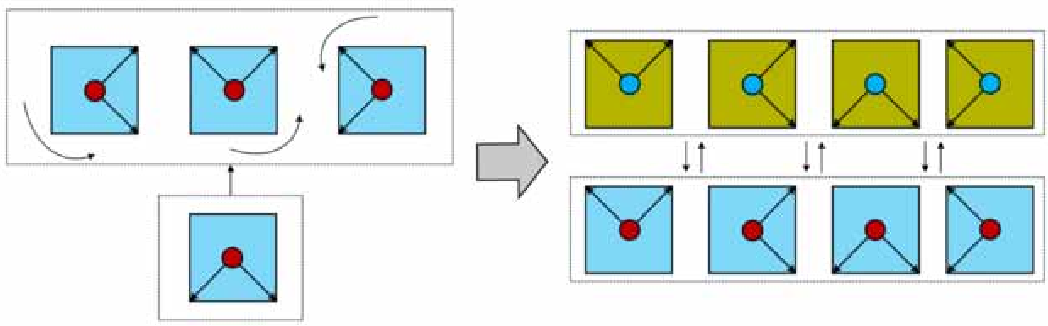
In Eq. (11), X_n is an RGBD image as input information and Y_n is a binary edge map that can be used as a true depth reference. The specific expression is Eq. (12).

$$Y_n = \left\{y_j^{(n)}, j = 1, \ldots, \left|X_n\right|\right\}, y_j^n \in \left\{0, 1\right\} \tag{12}$$

On this basis, the loss function Eq. (13) can be obtained.

$$\ell\left(W\right) = -\beta \sum_{j \in Y_+} \log P\left(y_j = 1 \middle| X; W\right) - \alpha \sum_{j \in Y_-} \log P\left(y_i = 0 \middle| X; W\right) \tag{13}$$

Figure 6. Model Training Input



W in Eq. (13) is the parameter set of the network. $|Y_-|$ represents the edge set of a binary edge graph. $|Y_+|$ refers to a non edge set of binary edge graphs. α and β are class equilibrium coefficients. $P(\ )$ represents a probability value. Marking the four class equilibrium loss functions used in the study separately and the set formed by them can be expressed as Eq. (14).

$$L_i\left(i \in \{1,2,3,4\}\right) \tag{14}$$

In Eq. (14), L_1, L_2, L_3, and L_4 represent four class equilibrium loss functions. The final total loss function is obtained by Eq. (15).

$$L_{final} = s_1 \cdot L_1 + s_2 \cdot L_2 + s_3 \cdot L_3 + s_4 \cdot L_4 \tag{15}$$

In addition, it should be noted that in general, when conducting deep neural network training, the model is required to have sufficient images for training. The more images trained, the stronger the generalization ability of the model itself. However, in actual model training, the information used for training may not be very sufficient. As a result, it is necessary to enhance the data of the training image to achieve the effect of changing the appearance of the image and forming a new image. The new image will also be used as training input for model training. Figure 6 shows the specific process.

Meanwhile, to prevent over fitting of the network structure during training, the original training image will be rotated by 90 degrees, 180 degrees, and 270 degrees, respectively, and then flipped left and right on this basis. With this assumption, the update of data is realized, the overall data volume of the training set is increased, and the generalization ability of the model is improved.

EFFECT VERIFICATION OF RGBD IMAGE SYMBIOSIS ENHANCEMENT MODEL

Symbiosis Test

The effectiveness of the RGBD image symbiosis enhancement model is tested from two perspectives: the symbiosis and enhancement of the model. The symbiosis test is mainly used to test the depth image edge detection effect of the model, while the enhancement test is mainly used to test the depth image enhancement effect of the model. The training of the model is divided into three steps, and the training settings for each step are shown in Table 1.

Table 1. Model Training Settings

SDE-NET	Step1	Training image specification	50*50
		Step	25
		Training batch specification	128
		Momentum size	0.9
		Weight attenuation coefficient	0.0005
		Gradient clipping parameters	0.1
		Iterations of learning rate change	10,000
	Step2	Number of training batches	128
		Momentum	0.9
		Weight attenuation coefficient	0.0002
		Weight coefficient of single loss function	1
		Iterations of learning rate change	8,000
		Data set	NYU v2
	Step3	Training image specification	50*50
		Step	25
		Training batch specification	60
		Momentum size	0.9
		Weight attenuation coefficient	0.0005
		Gradient clipping parameters	0.1
		Iterations of learning rate change	20,000

In Table 1, because the model is separated into three main parts, targeted training is required for the three parts when conducting model training. The first part of the training and the third part of the training have similarities in most training settings, with only some differences between the learning rate change iteration number and the weight attenuation coefficient. The second part of the training is aimed at structural edge feature extraction networks, so the training settings are slightly different from the first and third training settings. After training, this study will test the symbiosis and enhancement of the model. In the symbiosis test, NVUv2 is used as the main test data set, and the comparison of edge test results for different models is shown in Figure 7.

In Figure 7, the image edge detection model Holistically-Nested Edge Detection (HED) is mainly used for comparison with the SE model designed in the study. HHA represents the Horizontal Disparity, Height Above Ground, and Angle With Gravity of the three channels of depth image coding. At the same time, two indicators, Optimal Database Scale (ODS) and Optimal Image Scale (OIS), are used for analysis. The ODS indicator refers to the fixed P value that maximizes the F-measure in the dataset, while the OIS indicator refers to the fixed P value that maximizes the F-measure of a single image itself. In terms of ODS, the ODS values of the SE model for the three types of HHA, RGB, and HHA-RGB are 0.697, 0.726, and 0.755, respectively, which are higher than the three values of the HED model. In terms of OIS, the ODS values of the SE model for the three types of HHA, RGB, and HHA-RGB are 0.706, 0.738, and 0.767, respectively, which are higher than the three values of the HED model. The FPS values for the two models are consistent. The data proves that the edge detection effect of the model designed in the study is better. Figure 8 shows the model Precision-Recall (P-R) curve.

Figure 7. Edge Inspection Results

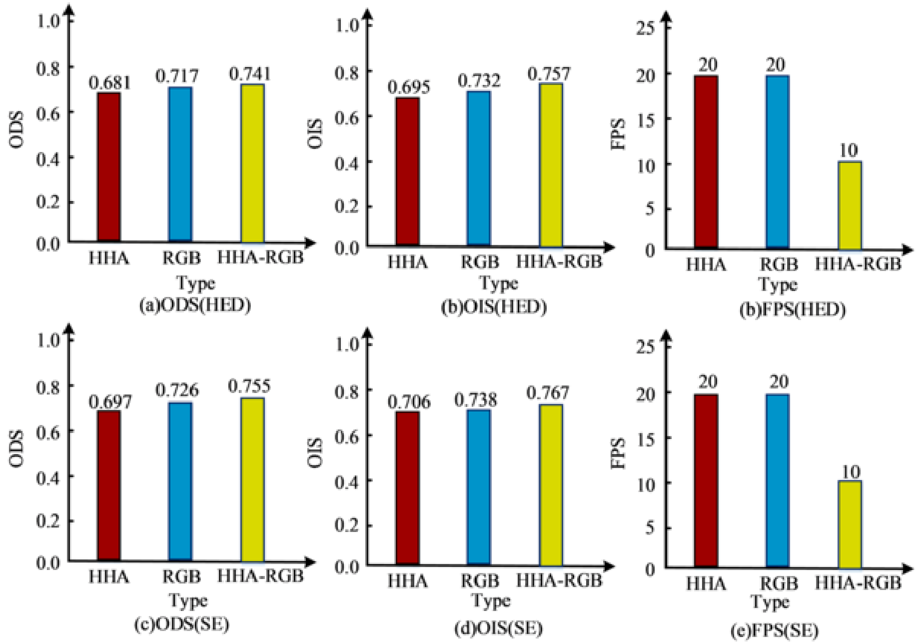
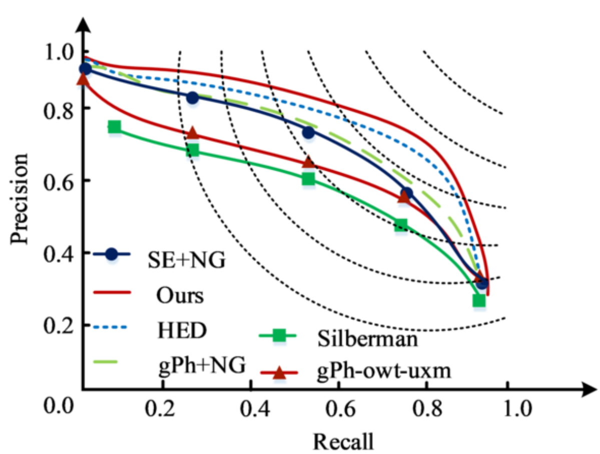


Figure 8 shows that the model under this study has the largest area under the PR curve, and its overall position is above the curves of other comparative models. Therefore, the model designed in the study is the best among the comparative models obtained in the dataset.

Enhancement Test

In the enhancement testing section, the effectiveness of the empty convolution Skip Connection (SK) and the jump connection Dilated Convolution (DC) in the designed network structure is verified.

Figure 8. Model PR Curve



In Figure 9, the model that does not use jump connections and void convolution has the worst convergence effect on the loss function variation curve. The curve enters a stable convergence process only after about 15 iterations, and the overall convergence value is high. The model using both jump join and hole convolution enters a stable convergence process when the number of iterations is fewer than five. At the same time, in the comparison of Peak Signal-to-Noise Ratio (PSNR) data, the model using both jump connection and hole convolution has the highest PSNR data curve position and best performance.

In Figure 10, compared to the single GDE model, the composite SDE+GDE model has a better convergence effect for the loss function. Meanwhile, from the perspective of PSNR numerical value, the composite SDE+GDE model also reflects a better numerical level. When the iteration number reaches five, it reaches 41 dB, significantly higher than the single GDE model. In addition, composite three-step training, also known as single training, has a better convergence effect, and the PSNR value is also better, between 39.5dB and 40.0dB. It shows that the model designed in the study can indeed achieve better performance effects under three-step training, and using a comprehensive model is better than using a single model. The study selects nine different types of images for image enhancement, and the specific enhancement effect comparison is shown in Figure 11.

Among the nine different types of images in Figure 11, the model in this study shows good image enhancement effects in all eight images. The enhancement effects shown in Figure 1, Figure 2, Figure 3, Figure 4, Figure 5, Figure 6, Figure 8, and Figure 9 are significantly stronger than those shown in the other four models. The enhancement effect in Figure 7 is slightly lower than that of other models, but this does not affect its significant advantage in the average PSNR value. The average PSNR value of the model is 37.1 dB, which is the highest average PSNR of all models. Overall, the model in this study has the highest comprehensive performance in actual image enhancement applications and can form better image enhancement effects. In actual animation advertising depth image application scenarios, it can provide more in-depth and three-dimensional advertising images, improve the overall quality and viewing level of animation advertising, and create a good viewing experience for users.

Effectiveness Verification

To verify the effectiveness of the RGBD image synogenesis enhancement model, the study first compares the performance of the model and then uses the PSNR and Structural Similarity (SSIM). The contrast models are the image enhancement methods (Lv et al., 2021; Ngugi et al., 2021; Song et al., 2020). The dataset used in the study is the TID2013 dataset, which includes 1,700 distortion images and corresponding subjective scores and is often used to evaluate the performance of the

Figure 9. Hole Convolution and Jump Connection Verification

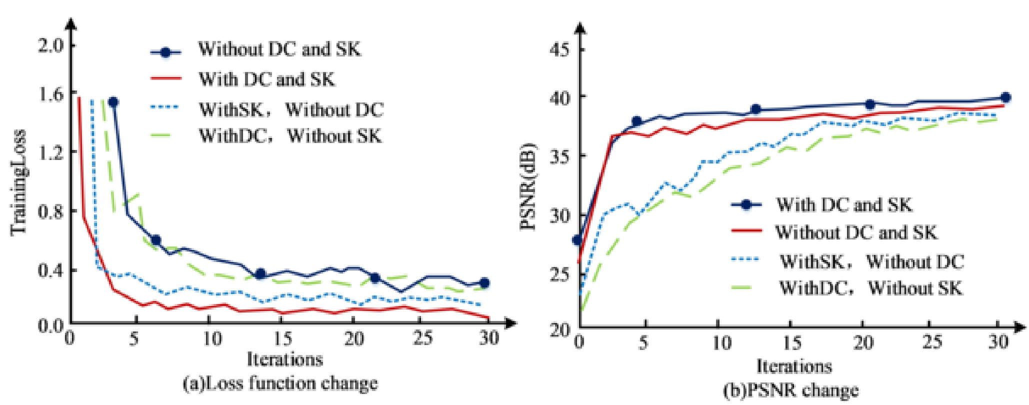


Figure 10. Composite Model Inspection

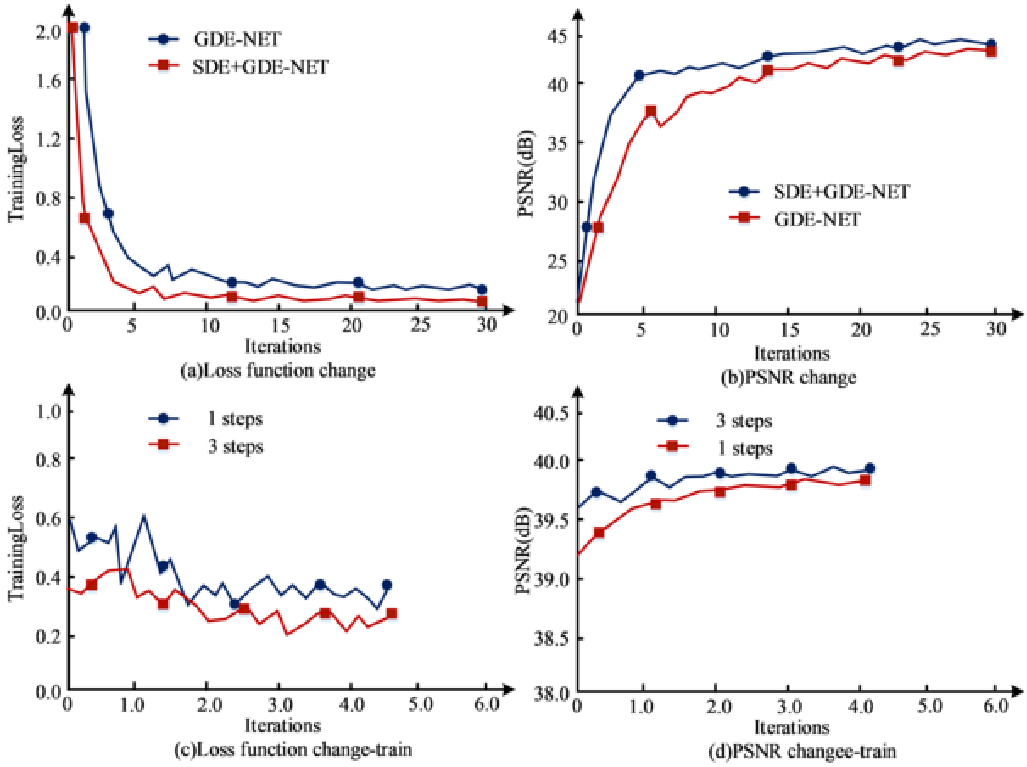


Figure 11. Comparison of Specific Enhancement Effects

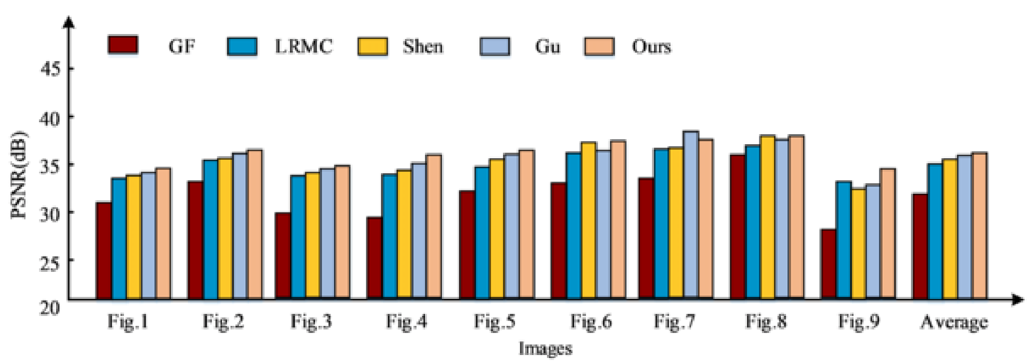
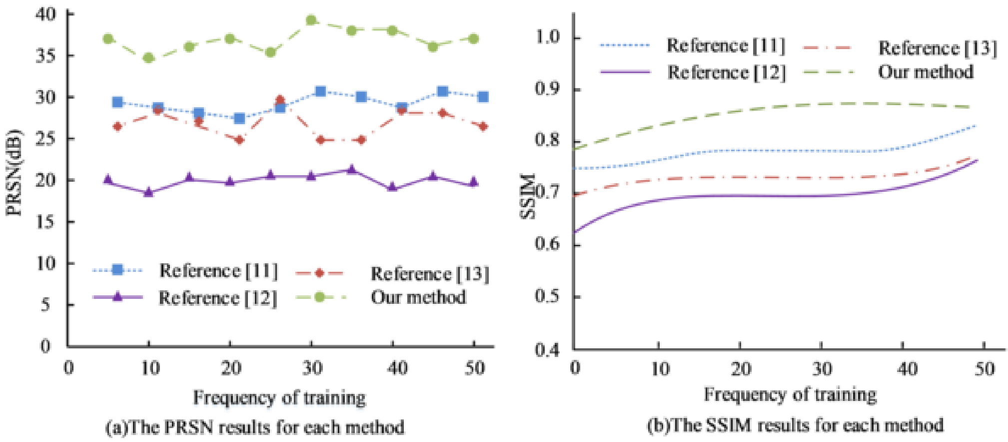


image quality assessment algorithms. The experimental environment is Intel Core i7-9700K (64GB), and the experimental platform is MATLAB. The comparison results of PSNR and SSIM for each image enhancement model are shown in Figure 12.

Figure 12(a) shows the PSNR value of each comparison method. The proposed RGBD image has the highest PSNR of 37.1dB, which is higher than other methods. Figure 12(b) shows the SSIM value of each comparison method. The output image of the RGBD image has the highest SSIM of 0.85,

Figure 12. Comparison Results of PSNR and SSIM Values of Image Enhancement



which is better than other methods. In conclusion, the results show that the proposed RGBD image synergetic enhancement model has achieved significant excellent performance in image enhancement. It can effectively improve the visual quality and realism of the image while maintaining the clarity and texture details of the image.

CONCLUSION

Animation advertising has the problem of insufficient depth image quality. This paper proposed a depth image symbiosis model consisting of three main network structures: SDE, SE, and GDE. The model used a three-step training and multi-layer progressive approach to improve the image enhancement performance. The research data showed that the ODS values of the SE were 0.706, 0.738, and 0.767 for the three types of HHA, RGB, and HHA-RGB, respectively, which were higher than the three values of the HED model. The FPS values of the two models were consistent, and the conclusion is that the edge detection effect of the research model is better. The model that used jump join and hole convolution simultaneously entered a stable convergence process when the iteration number was fewer than five.

The PSNR data curve had the highest position and best performance. Compared to the single model, the composite model had a PSNR of 41 dB when the iteration number reached five, which is significantly higher than the single model. Composite three-step training also had a better convergence effect than simplex training, and the PSNR value was also better, between 39.5dB and 40.0dB. In addition, from the practical application results, the average PSNR value of the designed model was the highest, at 37.1 dB. Eight of the nine images maintained the highest PSNR value. From a comprehensive perspective, the designed image enhancement model has more performance advantages. The proposed RGBD image synergetic enhancement model has practical significance in improving image quality and providing more attractive advertising content for the industry. This can lead to improved brand image and marketing effectiveness. However, the study has some limitations. The model's performance may be inadequate when dealing with certain scenes or complex images, and it may require further optimization and adjustment to accommodate a wider range of application scenarios. Future research directions could include optimizing the model's structure and parameters as well as improving its generalization ability and efficiency.

FUNDING STATEMENT

The research is supported by Private Education Brand Professional Construction Project of Henan Province Education Department: Computer Science and Technology (No.ZLG2019103); Industry-university Cooperative Education Project of Ministry of Education: Teaching Reform and Practice of Python Language Course Based on CDIO Mode (No.220601221223311); 2022 Zhengzhou Local Universities Urgent (special) Major Construction Project, Data Science and Big Data Technology, (grant number ZZLG202301).

CONFLICTS OF INTEREST

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

REFERENCES

Alhichri, H., Alswayed, A. S., Bazi, Y., Ammour, N., & Alajlan, N. A. (2021). Classification of remote sensing images using EfficientNet-B3 CNN model with attention. *IEEE Access : Practical Innovations, Open Solutions*, *9*, 14078–14094. doi:10.1109/ACCESS.2021.3051085

Alhussein, M., Aurangzeb, K., & Haider, S. I. (2020). Hybrid CNN-LSTM model for short-term individual household load forecasting. *IEEE Access : Practical Innovations, Open Solutions*, *8*, 180544–180557. doi:10.1109/ACCESS.2020.3028281

Basiri, M. E., Nemati, S., Abdar, M., Cambria, E., & Acharya, U. R. (2021). ABCDM: An attention-based bidirectional CNN-RNN deep model for sentiment analysis. *Future Generation Computer Systems*, *115*, 279–294. doi:10.1016/j.future.2020.08.005

Benz, D. C., Benetos, G., Rampidis, G., von Felten, E., Bakula, A., Sustar, A., Kudura, K., Messerli, M., Fuchs, T. A., Gebhard, C., Pazhenkottil, A. P., Kaufmann, P. A., & Buechel, R. R. (2020). Validation of deep-learning image reconstruction for coronary computed tomography angiography: Impact on noise, image quality and diagnostic accuracy. *Journal of Cardiovascular Computed Tomography*, *14*(5), 444–451. doi:10.1016/j.jcct.2020.01.002 PMID:31974008

Berman, D., Levy, D., Avidan, S., & Treibitz, T. (2020). Underwater single image color restoration using haze-lines and a new quantitative dataset. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *43*(8), 2822–2837. doi:10.1109/TPAMI.2020.2977624 PMID:32142424

Blanchet, F. G., Cazelles, K., & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, *23*(7), 1050–1063. doi:10.1111/ele.13525 PMID:32429003

Cai, Y., Hu, X., & Wang, H. Yulun, Zhang, Hanspeter, Pfister, Donglai, & Wei. (2021). Learning to generate realistic noisy images via pixel-level noise-aware adversarial training. Advances in Neural Information Processing Systems, 34, 3259-3270.

Islam, M. J., Xia, Y., & Sattar, J. (2020). Fast underwater image enhancement for improved visual perception. *IEEE Robotics and Automation Letters*, *5*(2), 3227–3234. doi:10.1109/LRA.2020.2974710

Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Zhou, P., & Wang, Z. (2021). Enlightengan: Deep light enhancement without paired supervision. *IEEE Transactions on Image Processing*, *30*, 2340–2349. doi:10.1109/TIP.2021.3051462 PMID:33481709

Liu, M., Jiao, L., Liu, X., Li, L., Liu, F., & Yang, S. (2020). C-CNN: Contourlet convolutional neural networks. *IEEE Transactions on Neural Networks and Learning Systems*, *32*(6), 2636–2649. doi:10.1109/TNNLS.2020.3007412 PMID:32692683

Lv, F., Li, Y., & Lu, F. (2021). Attention guided low-light image enhancement with a large scale low-light simulation dataset. *International Journal of Computer Vision*, *129*(7), 2175–2193. doi:10.1007/s11263-021-01466-8

Monga, V., Li, Y., & Eldar, Y. C. (2021). Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing. *IEEE Signal Processing Magazine*, *38*(2), 18–44. doi:10.1109/MSP.2020.3016905

Ngugi, L. C., Abelwahab, M., & Abo-Zahhad, M. (2021). Recent advances in image processing techniques for automated leaf pest and disease recognition–A review. *Information Processing in Agriculture*, *8*(1), 27–51. doi:10.1016/j.inpa.2020.04.004

Polsinelli, M., Cinque, L., & Placidi, G. (2020). A light CNN for detecting COVID-19 from CT scans of the chest. *Pattern Recognition Letters*, *140*, 95–100. doi:10.1016/j.patrec.2020.10.001 PMID:33041409

Sharma, P., Berwal, Y. P. S., & Ghai, W. (2020). Performance analysis of deep learning CNN models for disease detection in plants using image segmentation. *Information Processing in Agriculture*, *7*(4), 566–574. doi:10.1016/j.inpa.2019.11.001

Song, W., Wang, Y., Huang, D., Liotta, A., & Perra, C. (2020). Enhancement of underwater images with statistical model of background light and optimization of transmission map. *IEEE Transactions on Broadcasting*, *66*(1), 153–169. doi:10.1109/TBC.2019.2960942

Wang, Z. J., Turko, R., Shaikh, O., Park, H., Das, N., Hohman, F., Kahng, M., & Polo Chau, D. H. (2020). CNN explainer: Learning convolutional neural networks with interactive visualization. *IEEE Transactions on Visualization and Computer Graphics*, *27*(2), 1396–1406. doi:10.1109/TVCG.2020.3030418 PMID:33048723

Xu, M., Li, C., Zhang, S., & Patrick, L. (2020). State-of-the-art in 360 video/image processing: Perception, assessment and compression. *IEEE Journal of Selected Topics in Signal Processing*, *14*(1), 5–26. doi:10.1109/JSTSP.2020.2966864

Yu, C., Han, R., Song, M., Liu, C., & Chang, C.-I. (2020). A simplified 2D-3D CNN architecture for hyperspectral image classification based on spatial-spectral fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *13*, 2485–2501. doi:10.1109/JSTARS.2020.2983224

Zhou, X., Li, Y., & Liang, W. (2020). CNN-RNN based intelligent recommendation for online medical pre-diagnosis support. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, *18*(3), 912–921. doi:10.1109/TCBB.2020.2994780 PMID:32750846

Zhou, Y., Zheng, H., Kravchenko, I. I., & Valentine, J. (2020). Flat optics for image differentiation. *Nature Photonics*, *14*(5), 316–323. doi:10.1038/s41566-020-0591-3