Research Article

Virtual Reality Video Image Classification Based on Texture Features

Guofang Qin¹ and Guoliang Qin²

¹Art College, Anyang Normal University, Anyang 455000, China

²Hubei Guowang Huazhong Science and Technology Development Co., Ltd., Wuhan 430077, China

Correspondence should be addressed to Guofang Qin; qgf@aynu.edu.cn

Received 19 February 2021; Revised 20 April 2021; Accepted 7 May 2021; Published 18 May 2021

Academic Editor: Wei Wang

Copyright © 2021 Guofang Qin and Guoliang Qin. 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.

As one of the most widely used methods in deep learning technology, convolutional neural networks have powerful feature extraction capabilities and nonlinear data fitting capabilities. However, the convolutional neural network method still has disadvantages such as complex network model, too long training time and excessive consumption of computing resources, slow convergence speed, network overfitting, and classification accuracy that needs to be improved. Therefore, this article proposes a dense convolutional neural network classification algorithm based on texture features for images in virtual reality videos. First, the texture feature of the image is introduced as a priori information to reflect the spatial relationship between pixels and the unique characteristics of different types of ground features. Second, the grey level cooccurrence matrix (GLCM) is used to extract the grey level correlation features of the image in space. Then, Gauss Markov Random Field (GMRF) is used to establish the statistical correlation characteristics between neighbouring pixels, and the extracted GLCM-GMRF texture feature and image intensity vector are combined. Finally, based on DenseNet, an improved shallow layer dense convolutional neural network (L-DenseNet) is proposed, which can compress network parameters and improve the feature extraction ability of the network. The experimental results show that compared with the current classification method, this method can effectively suppress the influence of coherent speckle noise and obtain better classification results.

1. Introduction

The rapid development of computer technology has promoted the output of smart devices, and smart products have been widely used in all aspects of people's lives and learning, bringing them great convenience [1, 2]. Virtual reality (VR) technology is an emerging practical technology developed in the 20th century [3–5]. This technology has its own unique advantages in the early stages of development. Based on the continuous advancement of technology, the simulation system and simulation environment created by virtual reality have opened new breakthroughs for the transformation and upgrading of many industries. At present, good applications have been achieved in many fields such as my country's manufacturing, education, medical care, culture and art, entertainment and film, business, and tourism [6].

Specifically, virtual reality technology refers to the simulation of a virtual world with the help of computer technology, allowing participants to interact with the virtual world with the help of VR headband display, data gloves, headphones, and other devices and allowing participants to complete certain aspects of training in "real" experience. Image classification refers to the image processing technology that excavates and obtains the different features presented in the image information and distinguishes different types of image data [7]. The core content of image classification is the extraction of image features. The extraction of high-level semantic features can correctly understand the content or essence of the image, thereby improving the accuracy of image classification. With the advent of the era of big data and the rapid development of Internet technology, deep learning methods have become a research hotspot in image classification tasks [8]. As one of the

most widely used methods in deep learning technology, convolutional neural networks have powerful feature extraction capabilities and nonlinear data fitting capabilities and are one of the most effective algorithms in today's image classification tasks [9]. Many researchers have carried out a lot of research on convolutional neural network and image classification technology and made excellent progress, but the convolutional neural network method still has complex network models, long training time, excessive consumption of computing resources, and convergence speed. Slowness, network overfitting, and classification accuracy need to be improved.

The deep convolutional network model has great advantages in solving complex problems. In order to further improve the performance of some recognition tasks, deeper convolutional network models have been proposed one after another. However, it also makes the deep convolutional network model more complex and the number of parameters increases sharply. Under certain conditions, increasing the depth of the network can properly improve the recognition performance of the network, but it will also bring the problem of model redundancy. In general, the excessive number of parameters in the convolutional neural network will cause the following problems: (1) the network being easy to fall into overfitting; (2) gradient vanishing phenomenon; (3) too many training parameters making it difficult to optimize the network, and being difficult to fit the data in the process of network training.

In order to effectively solve the problems of network degradation caused by too many network parameters, excessive occupation of computing resources, and storage resources and balance the relationship between network parameters and model accuracy, in this paper, a dense convolutional neural network classification algorithm based on texture features is proposed for images in virtual reality video. The validity of the model is proved by some experiments.

The remainder of this paper is organized as follows: Section 2 reviews the related work; Section 3 introduces the proposed methods; Section 4 reports the experiments and results; and Section 5 concludes our work.

2. Related Works

As an emerging technology, virtual reality technology involves multiple application fields [10, 11]. Its characteristic lies in constructing their ideal virtual environment according to the needs of users, allowing their brains to deeply immerse themselves in the real scene through visual, tactile, auditory, and other perception behaviours. Innovative applications of VR have developed rapidly in recent years. Many innovative products have been applied to people's work and life, improving their work efficiency and happiness in life. Researchers at NASA's Ames Laboratory are working on a project called "Virtual Planet Exploration" [12]. Multiple VR training systems have been established in the project, including a VR training system, station VR training system, and VR education system. The scientific research team at Loma Linda University Medical Centre applied computer graphics and VR technology together to

discuss problems related to neurological diseases and successfully pioneered the VR paediatric treatment method [13]. Bortone et al. [14] used VR to simulate a virtual game scene of manual rehabilitation training. Patients can grab different objects and move them to a fixed container. Repeatedly playing this game can improve the patient's motor and cognitive abilities. Lin et al. [15] studied methods based on virtual reality technology to reduce nerve pressure. Participants received VR-based stress therapy training. The content elements included island environment, forest environment, and calm instrumental music. Their research also confirmed that the integration of virtual reality technology could reduce nerves.

With the development of technology, it has become a trend to combine virtual reality technology with image classification. Image classification refers to the image processing technology that extracts the features of the input original image by related algorithms to obtain the key features of the target and classifies it into a known category. In image classification technology, the feature extraction of image data is the most important stage. To ensure that the extracted features can effectively interpret and express the content information of the image, whether the key features can be extracted directly determines the accuracy of the entire classification task. The design of the classifier model is also the main task in image classification technology. The performance of the classifier directly determines the quality of the entire classification task. The choice of classifier model mainly includes decision tree [16], support vector machine [17], main component analysis [18], K-nearest neighbour [19], and AdaBoost [20]. These methods extract the global features of the image. The process of extracting the global features is more cumbersome. Moreover, the image recognition technology that relies on the global features of the image is susceptible to illumination, occlusion, missing, size change, and other factors, and the image data processed is more complex. It did not have the desired effect.

Texture analysis and texture feature extraction have always been an active field of image processing research. Using texture features to classify texture images is the focus of attention of researchers. How to extract representative texture features is an important part of successful texture image description, which directly affects the correct rate of subsequent classification [21]. Di and Gao [22] proposed an improved grey level cooccurrence matrix (GLCM). This method has strong adaptability and good robustness. After that, researchers began to use various random field models [23] and fractal theory [24] to describe the texture characteristics of images. Yang and Yang [25] introduced wavelet transform to texture feature extraction. Wavelet transform can simultaneously obtain frequency domain and spatial domain information and extract multiscale texture features. Xia et al. [26] proposed a local binary pattern (LBP) for describing image texture features. This method is easy to calculate and has the advantages of rotation invariance and grey invariance. Wan and Chang [27] used WorldView-2 images to map the bamboo forests in India. The first and second principal components of the images and the corresponding GLCM mean values obtained by the recursive

feature elimination are important variables for classification. Muthukumar and Kavipriya [28] used SVM classifier for information extraction, and after adding the multiscale texture features extracted by the Gabor filter, the total classification accuracy increased from 93.82% to 97.40%.

The advantage of the convolutional neural network is that it can process the input multidimensional image directly. The target image is operated by convolution operation and pool processing, and the feature of the target image is extracted layer by layer, and the key image features are obtained continuously. Obtain high-level semantic information that can reflect the image category, thereby avoiding the tedious preprocessing and complex feature extraction of the original image. Yildiz [29] used deep learning methods to learn the texture features of video images and classify them. Rajagopal et al. [30] used a convolutional neural network training method to improve the accuracy of image recognition. Patrini et al. [31] combined convolutional neural networks and transfer learning methods, which performed well in the ImageNet dataset.

The artificially designed low-level feature method is cumbersome in feature extraction tasks, and the extracted features do not describe the target image well, which makes the classification accuracy of the image not high. The use of deep learning methods to extract features has a strong ability to describe target data and can achieve better classification performance. Therefore, this paper takes the texture feature of the image as input and sends it to the network for training, to realize the classification of video images.

3. Abnormal Behaviour Detection Algorithm

3.1. Virtual Reality Video Acquisition. The complete virtual reality video transmission architecture includes five parts: panoramic acquisition, stitching, mapping, encoding, and transmission, as shown in Figure 1.

- (1) *Panoramic Acquisition.* A multicamera combination panoramic camera takes the VR video. Using a panoramic camera to collect natural images can greatly simplify the production of virtual reality content.
- (2) *Video Splicing.* Video splicing is the postprocessing process of panoramic video collection. It combines the videos of several cameras into a panoramic video to form a complete virtual reality video and provide it to users.
- (3) *The Mapping Process.* In order to facilitate storage and compression coding, the spherical video content needs to be geometrically mapped to the plane. Mapping affects the number of pixels before VR video encoding and to a certain extent determines the amount of information contained in the video content.
- (4) *Video Coding.* VR video uses compression coding to reduce redundant information in the video.
- (5) *Network Transmission.* VR videos need to be distributed to users via the Internet.

The processing flow of the virtual reality video proposed in this paper is shown in Figure 2. In the stage of virtual reality content production, the sound-visual scene of the real physical world is recorded by a set of cameras or a camera device with multiple cameras and sensors and audio sensors. The output of the device is a set of digital videos and audio signals.

Compared with traditional video, the unique feature of virtual reality video is that usually only part of the video in the panoramic video corresponding to the current viewpoint is displayed. This feature can be used to use viewpoint-related virtual reality video processing technology to reduce video transmission bandwidth requirements or reduce video decoding complexity while ensuring the same video resolution/quality for users, thereby improving the overall performance of the virtual reality video system.

3.2. Texture Feature Extraction. Feature extraction is the most critical step in texture image classification and detection. Short extraction time, strong discrimination, and good robustness are indicators for extracting texture features [32–36]. In order to describe the characteristics of the distorted texture image, this paper extracts the GLCM feature and GMRF feature of the texture image.

3.2.1. Texture Feature Extraction Based on GLCM. The GLCM texture feature is obtained based on the GLCM statistics. The GLCM is obtained by calculating the joint probability density of two pixels in the Q direction separated by s. The calculation formula is

$$p_{ab}(Q, s) = \frac{F_{ab}}{\sum_{a=1, b=1}^n F_{ab}}. \quad (1)$$

Among them, F_{ab} is the number of occurrences of grey level a and grey level b , n is the grey level of the image, and Q generally chooses 0° , 45° , 90° , and 135° . There are 14 types of texture feature statistics commonly used in GLCM. Due to the correlation between various feature statistics, there is no need to calculate all features. This paper uses the following four unrelated feature statistics.

- (1) *Energy.* The calculation formula is

$$Eg = \sum_{a=1}^n \sum_{b=1}^n \{p_{ab}(Q, s)\}^2. \quad (2)$$

- (2) *Entropy.* The calculation formula is

$$Ep = \sum_{a=1}^n \sum_{b=1}^n p_{ab}(Q, s) \log_{10}(p_{ab}(Q, s)). \quad (3)$$

- (3) *Contrast.* The calculation formula is

$$Ct = \sum_{a=1}^n \sum_{b=1}^n (a - b)^2 p(a, b). \quad (4)$$

- (4) *Correlation.* The calculation formula is

FIGURE 1: Virtual reality video transmission architecture.

$$Cr = \frac{\sum_{a=1}^n \sum_{b=1}^n (a, b) p(a, b) - \sum_{a=1}^n \sum_{b=1}^n a p(a, b) \sum_{a=1}^n \sum_{b=1}^n b p(a, b)}{\sum_{a=1}^n \sum_{b=1}^n (a - \sum_{a=1}^n \sum_{b=1}^n a p(a, b))^2 p(a, b) \sum_{a=1}^n \sum_{b=1}^n (b - \sum_{a=1}^n \sum_{b=1}^n b p(a, b))^2 p(a, b)}. \quad (5)$$

Among them, $p(a, b)$ is the value of the element of GLCM at (a, b) .

3.2.2. Texture Feature Extraction Based on the GMRF Model. In the Markov random field, the intensity value $G(a)$ of any pixel a of the image is related to all surrounding neighbouring pixels, which can be expressed in the form of conditional probability as $p(G(a)|G(a+t), t \in n)$.

Let U be the point set on the image block, $U = \{(k, h), 1 \leq k \leq n, 1 \leq h \leq n\}$; the GMRF process can be expressed by a symmetric difference equation as

$$G(a) = \sum_{t=1}^n \alpha_t (G(a+t)). \quad (6)$$

Among them, n is the GMRF neighbourhood of point a ; $G(a+t)$ is the point in U , and α is the weight of the symmetric neighbourhood pixel. Substituting each pixel in area U into equation (6), the difference equation is obtained:

$$G(a) = G(U)(\alpha + \beta). \quad (7)$$

Among them, β is the feature vector to be estimated of the model. Estimating and solving with the least square error criterion can be obtained:

$$\beta = (\sum G(U)(G(U))^T)^{-1} (\sum G(U)(G(a))). \quad (8)$$

The obtained model parameter β is the GMRF texture feature describing the image block.

3.3. Classification Algorithm. The virtual reality video images have rich texture features, which are input into the network by introducing GLCM-GMRF texture features as a priori information. Using the powerful learning ability of the network for features, through layer-by-layer feature learning, more representative and discriminative advanced features can be obtained, thereby improving the classification accuracy.

FIGURE 2: Virtual reality video processing flow.

The spatial information extracted by GLCM can effectively suppress the influence of noise on the classification results and can distinguish different targets at the same time. The GMRF feature can also express the spatial information of the image, and it is two kinds of texture features that are not related to the GLCM feature. Based on the GLCM feature, the further introduction of the GMRF feature can increase the discrimination of different texture regions. The network is highly sensitive to features. It combines the intensity value of the image with the GLCM-GMRF texture feature and inputs it into the classifier. Through training, it can learn more abstract and high-level features to achieve high-precision image classification.

Affected by the structural superiority of the DenseNet model, the L-DenseNet network also uses dense connection technology. However, the traditional DenseNet increases the number of layers to as many as 100 layers. Experiments show that too many network layers will lead to network degradation and increase in model error rate, thereby affecting the training of the entire model. In order to avoid these problems, the network designed in this paper is a shallow structure, and the network is set to three levels of structure: 16 layers, 22 layers, and 28 layers.

The overall framework of the L-DenseNet model is shown in Figure 3. The three models designed include three L-Dense modules. Among them, the convolution operation and pooling operation between the input layer and the L-Dense module and the adjacent L-Dense module is called the transition layer, which aims to adjust the size of the feature map to facilitate the effective connection between the two modules. Therefore, the three designed models L-DenseNet-16, L-DenseNet-22, and L-DenseNet-28 all contain four transition layers. Among them, the convolution kernel size of the convolution operation in the transition layer is 3×3 , and the pooling kernel size of the pooling operation is set to 2×2 . The difference between them lies in the number of Dense-Conv modules. L-DenseNet-16 contains four dense convolution modules. In addition, L-DenseNet-22 and L-DenseNet-28 contain six and eight dense convolution modules, respectively.

The main function of the Dense-Conv module designed in this paper is to extract key features rich in high-level semantics from the target data. The Dense-Conv module designed in this paper includes BN, ReLU, compression layer, and expansion layer. Among them, the expansion layer is divided into two parts, expansion layer one and expansion layer two. The expansion layer 1 contains multiple 1×3 and

FIGURE 3: The overall framework of the L-DenseNet model.

3×1 one-dimensional convolution kernels, and the expansion layer 2 contains multiple 1×5 and 5×1 one-dimensional convolution kernels. This method can reduce network calculations, reduce network parameters, obtain different types of target features at multiple scales, increase the diversity of features, and improve network recognition accuracy. Finally, all the obtained feature vectors are feature stitched and then input into the subsequent layers.

4. Results and Discussion

In order to carry out the effectiveness and comparability of the research on image classification technology, the experiment in this chapter selects high-quality, standard, open, and universal image datasets. The image classification dataset used in this chapter is CIFAR-10 dataset. CIFAR-10 dataset is a common classical object image dataset. The dataset is divided into 10 categories and contains 60,000 colour images, of which 50,000 colour images are identified as the training set and 10,000 colour images are defined as the test set.

The hardware environment of the experiment is Intel i7-7700 CPU, 1T solid-state hard disk, 32 G memory, NVIDIA GTX1080TI GPU. The software environment includes Ubuntu 16.04 and TensorFlow for the deep model framework. The stochastic gradient descent algorithm is used in the experiment. Batch size is set as 64, network iteration algebra is set as 300, initial learning rate of network training is set as 0.1, weight attenuation is set as 0.001, and momentum factor is set as 0.9.

4.1. The Influence of Parameters on the Network. The number of hidden layers in the network structure affects the complexity of the network structure. In theory, studying the increase in the number of hidden layers can improve the ability of network feature extraction, but blindly pursuing the number of layers will increase the difficulty of learning. In this experiment, the numbers of hidden layer nodes are all

the same, 200 iterations are performed for different hidden layer numbers, and the average value of each iteration is 10 times. The experimental results are shown in Figure 4.

Observing Figure 4, we can see that the difference in the number of hidden layers in the network model will bring about different experimental results. When the number of hidden layers is two, the recognition accuracy is the lowest. When the number of hidden layers is six, the accuracy of recognition does not increase, but the effect is worse than other models. When the number of hidden layers is four, the overall recognition accuracy of the network model is the highest. With the continuous increase in the number of hidden layers, the recognition accuracy of the L-DenseNet network model in the process of feature training, learning first increases and then decreases, which does not continuously improve the recognition accuracy. This means that when the number of hidden layers is too large, the complexity of the network will increase and the extraction of feature information will be weakened, resulting in a downward trend in the accuracy and efficiency of recognition.

The number of hidden layer neuron nodes will also have a great impact on the training and learning ability of network feature information. Based on a fixed number of hidden layers of four, explore the influence of the number of hidden layer neurons on the generalization ability of the entire network model.

Observing and studying Figure 5, the accuracy of image recognition increases with the increase of the number of hidden layer nodes, which means that with the increase of hidden layer nodes, the network can improve the full absorption of feature information and the parameter optimization of the network structure is more effective. The stronger the expression ability of the data distribution, the more the meaning of the data can be fully expressed. It can also be seen that the number of hidden layer nodes cannot be too many or unlimited. If the number of nodes is too large, for example, the feature learning ability of the number of ascending nodes will be weakened.

FIGURE 4: Experimental results with different numbers of hidden layers.

4.2. L-DenseNet and Original DenseNet Experiment Comparison. In order to verify the performance of the proposed model, this paper compares the L-DenseNet model with DenseNet. There are three models of L-DenseNet: L-DenseNet-16, L-DenseNet-22, and L-DenseNet-28. DenseNet also has three models: DenseNet-16, DenseNet-22, and DenseNet-28. From the 10th generation to the 100th generation of the network algebra, the network training results are output every 5 generations. Figure 6(a) shows the relationship between the cross-entropy loss of each model and the iteration algebra during the training process. Figure 6(b) shows the relationship between the classification accuracy of each model and the number of iterations during the training process.

The network parameters of models DenseNet-16, DenseNet-22, DenseNet-28, L-DenseNet-16, L-DenseNet-22, and L-DenseNet-28 are 0.1 M, 0.3 M, 0.5 M, 0.5 M, 0.8 M, and 1.0 M. In order to visualize the experimental results more intuitively, this paper outputs the results after 100 iterations of each model; that is, the cross-entropy loss and recognition accuracy values obtained after each network are fully trained, as shown in Figure 7.

It can be obtained from Figure 7 that the cross-entropy loss of the networks DenseNet-16, DenseNet-22, DenseNet-28, L-DenseNet-16, L-DenseNet-22, and L-DenseNet-28 after full training is 0.389, 0.287, 0.167, 0.141, 0.122, and 0.098, respectively, the classification accuracy rates are 0.768, 0.833, 0.855, 0.892, 0.902, and 0.952 respectively. Combined

FIGURE 5: Experimental results of the number of nodes in different hidden layers.

FIGURE 6: The relationship between the cross-entropy loss, classification accuracy of each model, and the number of iterations during the training process. (a) The relationship between cross-entropy loss. (b) The relationship between classification accuracy.

with Figure 6, we can conclude that the most superior performance among the six models is the L-DenseNet-28 network. Of course, L-DenseNet-16 and L-DenseNet-22 also perform well, and the recognition accuracy is higher than that of the original DenseNet networks. In terms of network parameters, the L-DenseNet model has more parameters than the original DenseNet. Nevertheless, for DenseNet-28 and L-DenseNet-16 with the same number of parameters, the L-DenseNet-16 model has a higher classification accuracy than the DenseNet-28 network. From the experimental results and experimental analysis, we can conclude that the L-DenseNet model has a stronger feature extraction ability than the original DenseNet network. The proposed L-DenseNet can extract the multiscale and diverse features of the target data. It helps to reduce cross-entropy loss and improve recognition accuracy in network training.

4.3. Experimental Comparison of Multitexture Features. In the research process, the data information of training and learning multifeature fusion is used to complete the classification and recognition results. In order to verify the superiority of multifeature fusion, a single feature comparison experiment is carried out.

The comprehensive recognition effect of the network model reaches the expected value, and the impact of different feature learning is tested under the same dataset. Figure 8 shows that, for single feature input, multifeature input has the most obvious effect. Its sample image training learning accuracy rate and test sample recognition accuracy rate are both the highest under the same conditions, which are 0.953 and 0.952, respectively. The training of a single feature can achieve the recognition effect of the image, but the result is not ideal. Therefore, the multifeature input network model

FIGURE 7: Experimental comparison of cross-entropy and classification accuracy of L-DenseNet and DenseNet.

FIGURE 8: Classification accuracy of different feature inputs.

can fully absorb feature information and has a high feature representation ability, and the accuracy of classification and recognition has been improved.

4.4. Experimental Comparison of L-DenseNet Model and Other Deep Learning Methods. In order to demonstrate the performance of the L-DenseNet model in image classification tasks, we also compare it with different classic deep networks. The experimental results are shown in Figure 9. Figure 9 shows the performance comparison between the L-DenseNet model and different deep learning methods, and the comparison of the method parameters is given.

Through the network parameters and classification accuracy in Figure 9, it can be found that the performance of

the L-DenseNet model is better than most classic deep learning methods. The following conclusions can be drawn:

- (1) The L-DenseNet model has a shallow network structure and a small number of parameters. The design of the L-DenseNet model structure is helpful for the dissemination of feature information in network training. LBP have not effectively improved the feature extraction technology.
- (2) The L-DenseNet model can effectively alleviate the problem of gradient disappearance and overfitting, enhance feature propagation, and encourage feature reuse. Literature [29] and literature [30] aim to improve the network architecture to obtain better classification performance but ignore the feature

FIGURE 9: Comparison of classification accuracy and parameter values of different models.

reuse technology, and the extracted features are lost too much when propagating between networks, which is not conducive to network training and image classification.

- (3) The L-DenseNet model ensures the ability to extract network features while reducing parameters. Although the parameters in [30] are effectively compressed, the feature extraction ability of the model has not been effectively improved, resulting in poor classification performance.
- (4) The L-DenseNet model can extract and obtain the diversified features of the target image. These more detailed target features can effectively improve the recognition performance. MRF and SVM do not consider multiscale feature extraction technology.

5. Conclusion

Aiming at virtual reality video images, this paper proposes an image classification method based on GLCM-GMRF texture features and L-DenseNet. First, this paper introduces the GLCM-GMRF texture feature to suppress the influence of noise, while increasing the discrimination between different targets. Secondly, in view of the large number of parameters of the network and the weak feature extraction ability brought by the compressed network, a network called L-DenseNet is proposed. This model can effectively reduce the number of network parameters and strengthen the feature extraction capabilities of the model. Experimental results show that the algorithm in this paper effectively improves the classification accuracy of images and has good application value.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] L. Lin and T. Johnson, "Shifting to digital: informing the rapid development, deployment, and future of teaching and learning," *Educational Technology Research and Development*, vol. 69, no. 1, pp. 1–5, 2021.
- [2] D. G. A. Smith, L. A. Burns, D. A. Sirianni et al., "Psi4NumPy: an interactive quantum chemistry programming environment for reference implementations and rapid development," *Journal of Chemical Theory and Computation*, vol. 14, no. 7, pp. 3504–3511, 2018.
- [3] S. GÜDÜM, "Jekyll & Hyde online: a critical analysis on consumption of the self through advertisements promoting transhumanism," *Winter 2021*, vol. 12, no. 44, pp. 11–21, 2021.
- [4] A. Valentine, T. Van Der Veen, P. Kenworthy et al., "Using head mounted display virtual reality simulations in large engineering classes: operating vs. observing," *Australasian Journal of Educational Technology*, vol. 37, no. 3, pp. 119–136, 2021.
- [5] S. Kavanagh, A. Luxton-Reilly, B. Wuensche, and B. Plimmer, "A systematic review of virtual reality in education," *Themes in Science and Technology Education*, vol. 10, no. 2, pp. 85–119, 2017.
- [6] X. Pan and A. F. D. C. Hamilton, "Why and how to use virtual reality to study human social interaction: the challenges of exploring a new research landscape," *British Journal of Psychology*, vol. 109, no. 3, pp. 395–417, 2018.
- [7] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: an overview," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6690–6709, 2019.
- [8] J. Zhang, Y. Xie, Q. Wu, and Y. Xia, "Medical image classification using synergic deep learning," *Medical Image Analysis*, vol. 54, pp. 10–19, 2019.

- [9] W. Rawat and Z. Wang, "Deep convolutional neural networks for image classification: a comprehensive review," *Neural Computation*, vol. 29, no. 9, pp. 2352–2449, 2017.
- [10] G. Riva, B. K. Wiederhold, and F. Mantovani, "Neuroscience of virtual reality: from virtual exposure to embodied medicine," *Cyberpsychology, Behavior, and Social Networking*, vol. 22, no. 1, pp. 82–96, 2019.
- [11] M. Slater, "Immersion and the illusion of presence in virtual reality," *British Journal of Psychology*, vol. 109, no. 3, pp. 431–433, 2018.
- [12] K. S. Bitting, M. J. McCartney, K. R. Denning, and J. A. Roberts, "Conceptual learning outcomes of virtual experiential learning: results of google earth exploration in introductory geoscience courses," *Research in Science Education*, vol. 48, no. 3, pp. 533–548, 2018.
- [13] M. Moresco, F. Pizza, E. Antelmi, and G. Plazzi, "Sodium oxybate treatment in pediatric type 1 narcolepsy," *Current Drug Metabolism*, vol. 19, no. 13, pp. 1073–1079, 2018.
- [14] I. Bortone, D. Leonardis, N. Mastronicola et al., "Wearable haptics and immersive virtual reality rehabilitation training in children with neuromotor impairments," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 7, pp. 1469–1478, 2018.
- [15] J. Lin, L. Cao, and N. Li, "Assessing the influence of repeated exposures and mental stress on human wayfinding performance in indoor environments using virtual reality technology," *Advanced Engineering Informatics*, vol. 39, pp. 53–61, 2019.
- [16] G. Ke, Q. Meng, T. Finley et al., "Lightgbm: a highly efficient gradient boosting decision tree," *Advances in Neural Information Processing Systems*, vol. 30, pp. 3146–3154, 2017.
- [17] J. S. Raj and J. V. Ananthi, "Recurrent neural networks and nonlinear prediction in support vector machines," *Journal of Soft Computing Paradigm (JSCP)*, vol. 1, no. 1, pp. 33–40, 2019.
- [18] Z. Huang, Z. Wang, and H. Zhang, "A diagnosis algorithm for multiple open-circuited faults of microgrid inverters based on main fault component analysis," *IEEE Transactions on Energy Conversion*, vol. 33, no. 3, pp. 925–937, 2018.
- [19] J. Gou, H. Ma, W. Ou, S. Zeng, Y. Rao, and H. Yang, "A generalized mean distance-based k-nearest neighbor classifier," *Expert Systems with Applications*, vol. 115, pp. 356–372, 2019.
- [20] Y. Wu, Y. Ke, Z. Chen, S. Liang, H. Zhao, and H. Hong, "Application of alternating decision tree with AdaBoost and bagging ensembles for landslide susceptibility mapping," *Catena*, vol. 187, p. 104396, 2020.
- [21] L. Armi and S. Fekri-Ershad, "Texture image Classification based on improved local Quinary patterns," *Multimedia Tools and Applications*, vol. 78, no. 14, pp. 18995–19018, 2019.
- [22] H. Di and D. Gao, "Nonlinear gray-level co-occurrence matrix texture analysis for improved seismic facies interpretation," *Interpretation*, vol. 5, no. 3, pp. SJ31–SJ40, 2017.
- [23] R. Touati, M. Mignotte, and M. Dahmane, "Multimodal change detection in remote sensing images using an unsupervised pixel pairwise-based Markov random field model," *IEEE Transactions on Image Processing*, vol. 29, pp. 757–767, 2019.
- [24] Y. Aimin, L. Shanshan, L. Honglei, and J. Donghao, "Edge extraction of mineralogical phase based on fractal theory," *Chaos, Solitons & Fractals*, vol. 117, pp. 215–221, 2018.
- [25] P. Yang and G. Yang, "Statistical model and local binary pattern based texture feature extraction in dual-tree complex wavelet transform domain," *Multidimensional Systems and Signal Processing*, vol. 29, no. 3, pp. 851–865, 2018.
- [26] Z. Xia, X. Ma, Z. Shen, X. Sun, N. N. Xiong, and B. Jeon, "Secure image LBP feature extraction in cloud-based smart campus," *IEEE Access*, vol. 6, pp. 30392–30401, 2018.
- [27] S. Wan and S.-H. Chang, "Crop classification with WorldView-2 imagery using Support Vector Machine comparing texture analysis approaches and grey relational analysis in Jianan Plain, Taiwan," *International Journal of Remote Sensing*, vol. 40, no. 21, pp. 8076–8092, 2019.
- [28] A. Muthukumar and A. Kavipriya, "A biometric system based on Gabor feature extraction with SVM classifier for Finger-Knuckle-Print," *Pattern Recognition Letters*, vol. 125, no. 15, pp. 150–156, 2019.
- [29] K. Yildiz, "Identification of wool and mohair fibres with texture feature extraction and deep learning," *IET Image Processing*, vol. 14, no. 2, pp. 348–353, 2019.
- [30] A. Rajagopal, G. P. Joshi, A. Ramachandran et al., "A deep learning model based on multi-objective particle swarm optimization for scene classification in unmanned aerial vehicles," *IEEE Access*, vol. 8, pp. 135383–135393, 2020.
- [31] I. Patrini, M. Ruperti, S. Moccia et al., "Transfer learning for informative-frame selection in laryngoscopic videos through learned features," *Medical & Biological Engineering & Computing*, vol. 58, no. 9, pp. 1–25, 2020.
- [32] Q. Ke, S. Zeng-guo, Y. Liu, W. Wei, M. Woźniak, and R. Scherer, "High-resolution SAR image despeckling based on nonlocal means filter and modified AA model," *Security and Communication Networks*, vol. 2020, Article ID 8889317, 2020.
- [33] F. Orujov, R. Maskeliūnas, R. Damaševičius, and W. Wei, "Fuzzy based image edge detection algorithm for blood vessel detection in retinal images," *Applied Soft Computing*, vol. 94, p. 106452, 2020.
- [34] M. Zhang, W. Jing, J. Lin et al., "NAS-HRIS: automatic design and architecture search of neural network for semantic segmentation in remote sensing images," *Sensors*, vol. 20, no. 18, p. 5292, 2020.
- [35] H. Li and R. Mukundan, "Robust texture features for emphysema classification in CT images," in *Proceedings of the 2020 28th European Signal Processing Conference (EUSIPCO)*, pp. 1220–1224, Amsterdam, Netherlands, January 2021.
- [36] S. Dhingra and P. Bansal, "Relative examination of texture feature extraction techniques in image retrieval systems by employing neural network: an experimental review," in *Proceedings of International Conference on Artificial Intelligence and Applications*, pp. 337–349, Zurich, Switzerland, November 2020.