RESEARCH ARTICLE

The application of light and shadow visual technology in environmental landscape design

Ying Liu*

School of Humanities and Design, Open University of Henan, Zhengzhou, Henan, China.

Received: December 10, 2024; **accepted:** August 12, 2025.

Style and lighting are key elements in environmental landscape design. The development of modern intelligent technology enables style and lighting to assist in environmental landscape design. This study proposed a light and shadow vision technology based on light and shadow transfer and landscape image light and shadow detection using convolutional neural networks to achieve more intelligent environmental landscape design. The research employed semantic segmentation-guided spatial control and realism regularization strategies to mitigate style overflow and image distortion in light and shadow conversion. In addition, the proposed model improved the precision of image processing through the regression algorithm of local binary features. The performance validation of this study was also carried out. The results showed that the proposed model minimized its loss value to 123 iterations. Its accuracy for image feature extraction was up to 98.66%. The recognition accuracy of the proposed model was higher than 95% with a maximum of 98. 66%. The mean opinion score (MOS) of the proposed model for light and shadow transfer demonstrated that its overall transfer quality MOS was superior to that of ResNet, LeNet-5, and GoogleNet algorithms with a 6.2%, 9.4%, and 12.7% improvement over those three compared algorithms, respectively. The proposed method could effectively convey the style and lighting of images and could be applied to environmental landscape design.

Keywords: environmental design; image processing; optical transmission; brightness; spatial control.

***Corresponding author:** Ying Liu, School of Humanities and Design, Open University of Henan, Zhengzhou 450000, Henan, China. Email: lanhai0532@126.com.

Introduction

The aesthetics of the environmental landscape is a desirable goal for designers to achieve. Using light and image as key elements of environmental design not only enriches the environmental landscape but also perpetuates the environment [1]. Using the relationship between light and shadow, integrating light and shadow effects with the theme, style, color, texture, and atmosphere of the design can create an environmental landscape that is aesthetically pleasing and emotionally communicative [2]. The effective use of light and shadow vision can enhance people's sense of identity with the surrounding environment. Furthermore, it allows the space to communicate emotionally with people through light, shadow, and other such mediums, thereby satisfying people's psychological needs for the spatial environment. The vision of light and shadow has far-reaching significance in the design of environmental landscapes [3]. The continuous development of science and technology in modern society provides new ideas for the light transmission method in the environmental landscape. Among

them, light transfer is an algorithm that simulates the propagation and projection of light in a three-dimensional scene [4]. It can be utilized to calculate the interaction between light and objects, thereby facilitating the generation of realistic light and shadow effects. Furthermore, it serves as an invaluable auxiliary tool in environmental landscape design [5]. In recent years, with the rapid development of computer graphics and artificial intelligence technology, the application of light and shadow vision technology in environmental landscape design has received more and more attention. Researchers have gradually realized that the rational use of light and shadow can not only improve the beauty of the landscape but also affect people's psychological adaptation and identity of space. There have been studies using convolutional neural networks (CNN), deep learning, and other modern technologies to develop a variety of light and shadow detection and conversion algorithms. These algorithms have promoted the progress of intelligent environment design to a certain extent.

Scientists studied the visual elements of Chinese traditional culture Witchcraft including the cultural development, products creation, and the visual image using virtual reality (VR) based on 3D modeling technology. The study incorporated the relevant technical content of light and shadow vision to model the stage, props, and other objects and the results showed that the simulations based on the proposed technology had the necessary texture, color, light, and shadow variations, and were vivid and realistic [6]. Edensor and Hughes explored how the dynamics of light and shadow formed an integral part of daily emotional and sensory coherence and guided the movement of pedestrians by illustrating different light and shadow in both space and time, clarifying the role of light and shadow in different geographical environments. The results helped architects in the process of urban architectural design [7]. Lara applied the visual approach of light and shadow and combined them with a specific case of urban space design from the perspective of space and time to interpret people's inherent perceptions of the urban environment. The results of this project had important theoretical and practical significance for urban architects to create light and shadow imagery in cities and highlight urban architectural features [8]. Baslamisli *et al*. used the image decomposition algorithm to divide the color-generating components of the coloring component into two parts including direct color-generating (illumination) and indirect color-generating (color-generating) to achieve an effective intensity luminosity effect and reflectance change distinction and proposed an end-to-end deep CNN based on a fine structure with fine processing to extract specific spectral information from specific photometric effects to achieve a fine-grained portrayal of a real scene [9]. Eren and Yılmaz discussed the sketching of landscape architecture, interior design, or planning departments by combining computer technology and light and shadow visualization techniques. The study proposed the use of digital technology to assist students in landscape architecture design and found that the digital and traditional sketching skills acquired by the learners were directly related to their academic achievements and the work was ornamental [10]. Intelligent approaches to landscape design have also evolved at this stage with numerous researchers contributing to this effort. Eren *et al*. used a two-tiered approach to explore the influencing factors of urban landscape design, which included the review of literature on the impact of landscape color on individuals and translated them into a variety of color narratives, and then a questionnaire was used, and a factor analysis was conducted to reduce the statements in the questionnaire. The results showed that the strategy would guide landscape designers in their designs [11]. Kerr *et al*. described the prototype of main time augmented reality (AR) technology for those who were not familiar with the fundamentals of landscape architecture. The study explored the learning potential and benefits of AR, focused on digital narrative methods for different scenarios, and provided technical support for landscape architecture education [12]. Zhong *et al*. took "nature" as a

perspective and bionic design as a theoretical framework to explain the landscape architecture of “nature”. The study identified and compared the important structures in biological design and explained their main components and further explore the role of eco-friendly architecture in realizing sustainable development from the perspective of an ecological environment, which was quite meaningful for urban landscape design [13]. Yu *et al*. established a set of dynamic landscape models based on parameterization and proposed a complete set of parameterization schemes for environmental landscape design at multiple levels with good practical value [14]. Yang *et al*. developed a mathematical model of street microclimate through the measurement and simulation of neighborhoods and subsequently calculated a thermal comfort index to inform the landscape planning of streets by taking urban roads as the object and quantifying the factors of road pavement, greening, and shadow cover in urban roads and human thermal comfort to improve the quality of public space in old cities through landscape design to create resilient and humanized street space [15].

However, there are still many problems in existing methods such as image distortion and style overflow in the conversion of style and content. Therefore, how to achieve high-fidelity light and shadow transmission in landscape design and how to improve the accuracy and stability of the algorithm have become the key problems that need to be solved. Further, designers need effective tools to monitor and adjust the effects of light and shadow in different environmental conditions in practical applications. This study aimed to develop a smarter and more efficient tool to help designers accurately transmit and control light and shadow effects to create rich and deep environmental landscapes. An improved light and shadow transmission algorithm combined with CNN, semantic segmentation guided spatial control, and photorealistic regularization was proposed based on light transmission and light detection of landscape images to realize intelligent light transmission in environmental landscape design. Through deep learning technology, the style and content of landscape images were transformed to enhance image detail retention and visual effect. The spatial control method (SCSS) and photorealistic regularization strategy (PR) based on semantic segmentation were proposed to effectively reduce the problems of style overflow and image distortion in the process of light and shadow transmission. In addition, the color control idea (LC) based on brightness channel was innovatively introduced to maintain the consistency of image color. Meanwhile, the regression algorithm based on local binary feature (LBF) was used to improve the accuracy of image feature extraction. This research had important guiding significance for the future landscape design practice and helped designers to adjust the light and shadow effect more accurately to meet people's psychological needs for space environment. In addition, the proposed model provided a new perspective and foundation for subsequent research and promoted the further development of light and shadow vision technology in environmental design.

Materials and methods

Light and shadow visual transfer algorithm based on improved CNN

Light and shadow transfer can be considered as image style transfer. The difficulty of light and shadow transfer lies in the need to better transfer the artistic light and shadow on the reference image to the target landscape image while retaining the detailed information of the target landscape [16]. When using traditional methods for light and shadow transfer, there is usually an incomplete separation of light and shadow from the detailed features of the landscape [17]. CNN has excellent feature extraction ability and has been widely used in image target recognition, which can independently extract and process the style and content features of images and has practical applicability in the light and shadow transfer process of landscape images. Therefore, this

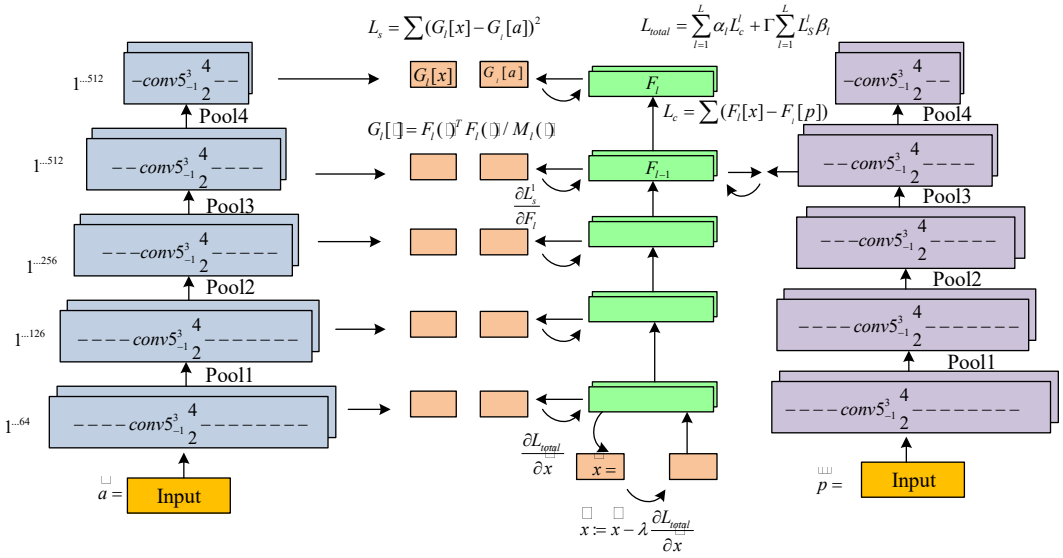

Figure 1. Neural network algorithm style transfer process.

study used CNN to facilitate the transfer of light and shadow, while simultaneously preserving the intricate details of the landscape. CNN was particularly adept at processing the content and style of images, rendering them an optimal choice for this task. The neural network style transfer algorithm utilized a trained CNN VGG19. For a given reference image *A* and a target image *P*, the reference image *A* passed through the network extracting a number of feature maps at each convolutional layer. The set of feature maps in each layer constituted the stylized representation of *A* in that layer $G_i[A]$. Similarly, the target image *P* passed through a layer of the network with a number of feature maps, and these feature maps formed the content representation of *P* at that layer $F_i[P]$. A randomly generated noisy image *X* was then generated and many feature maps were also extracted from this image as it passed through the network. The feature maps of each layer formed the style representation $G_i[X]$ and the content representation $F_i[X]$ of *X* at that layer, respectively. Then, the loss between *X* and *P* in a certain layer of content representation and the loss between *X* and *A* in multiple layers of style representation were calculated at the same time. *X* was adjusted by minimizing the loss function through continuous optimization, so that the random noise image *X* maintained the content of the target image *P* as well as has the style of reference image *A* (Figure 1). The style and content features were extracted from the reference image and target image, respectively, using the VGG19 network. By optimizing the random noise image, its content features were aligned with the target image, while incorporating the style of the reference image. The process entailed calculating and subsequently reducing the losses incurred in terms of style and content, which was achieved by adjusting the noisy image that facilitated style transfer and content preservation, thereby ensuring that the resulting image exhibited both the style of the reference image and the content of the target image. The content representation of the generated image *X* in the *l* layer of the network $F_i[X]$ was equal to the content of the image *P* in the *l* layer $F_i[P]$. Then, the mean square error loss function between the two types of feature representations was defined as follows.

$$L_c^l = \frac{1}{2N_l M_l}\sum_{ij}\left(F_l[X]-F_l[P]\right)_{ij}^2 \quad (1)$$

where N_i was the number of filters in the neural network layer, which produced N_i feature vectors

with a dimension of each as M_i. *ij* was the activation value of the *i*-th filter at position *j*. To obtain a stylized representation of the reference image *A*, this study used a feature space to capture the texture information [18], which consisted of the correlation between the feature responses of the different filters in any of the layers of the network and was shown in equation (2).

$$G_l[\cdot] = \frac{1}{M_l(\cdot)} F_l[\cdot]^T F_l[\cdot] \tag{2}$$

where $G_l[\cdot]$ was the correlation, whose main meaning was the Lagom matrix between the feature vectors of the image in the *l* layer. To generate a new image *X* that incorporated the style of the reference image *A* and the content of the target image *P*, it was necessary to synchronously minimize the loss between the images *X* and *P* in one layer of the content representation and the loss between *X* and *A* in multiple layers of the stylistic representation. The loss function was shown in equation (3).

$$L_{total} = \sum_{l=1}^{L} \alpha_l L_c^l + \Gamma \sum_{l=1}^{L} \beta_l L_s^l \tag{3}$$

where α_l was the weight factor of the content loss factor of the *l* layer over the total content loss. β_l was the weight factor of the style loss factor of the *l* layer over the total style loss. α_l and β_l were the inverse of their respective layers. β_l was the weight factor that balanced the content loss and style loss. The above methods could effectively transfer styles, but there were problems such as style overflow and detail distortion in the image conversion process [19, 20]. To improve the situation, this study proposed spatial SCSS and photorealistic regulation of photos (PR) deep photo stylization algorithms. A least squares constraint term was first introduced, and the vector of the output image *X* in the channel α was defined as $V_c[X]$. The loss of the photorealistic regularization term was then defined as shown in equation (4).

$$L_m = \sum_{C=1}^{3} V_C[X]^T M_p V_C[X] \tag{4}$$

where M_p was the symmetric matrix. When using this photo-realistic regular term in gradient processing, it was necessary to know the derivative of the output image as shown in equation (5).

$$\frac{dL_m}{dV_C[X]} = 2M_P V_C[X] \tag{5}$$

where L_m was the loss function of the image realism regularization term. When incorporating the light and shadow of the reference image into the target image, it was important to ensure that the scene matched. That is, the shaded areas should correspond to the regions passed to the target landscape, and the non-shaded areas passed onto the non-shaded regions. In neural algorithms, the Gram matrix of the style represented the overall distribution of the style of the entire image, which arranged the extracted feature vectors as equidistant vectors to implicitly encode the neural response of the exact distribution. This limited the ability to adapt to changes in the semantic content and could lead to an overflow of the style [21, 22]. To address this problem, the image was semantically segmented to generate a mask prior to delivery. Then a photomask was added to the input image as another channel to guide the photomask passing with the corresponding same semantic labels on the mask. The defined style loss function was shown in equation (6).

$$\begin{cases} L_{s+}^l = \sum_{c=1}^{C} \frac{1}{2N_{l,c}^2} \sum_{ij} \left(G_{l,c}[X] - G_{l,c}[A] \right)_{ij}^2 \\ F_{l,c}[X] = F_l[X] M_{l,c}[P] \\ F_{l,c}[A] = F_l[A] M_{l,c}[A] \end{cases} \tag{6}$$

where *C* was the number of light and shadow segmentation regions (number of channels).

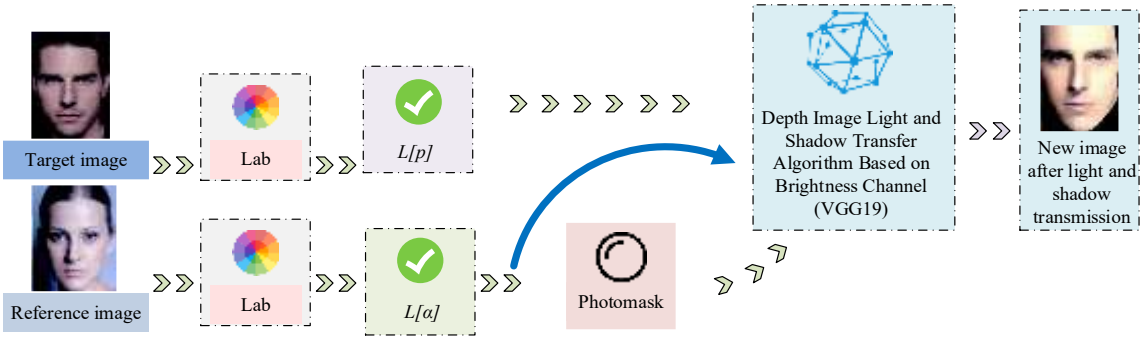

Figure 2. Color control method based on luminance channel.

$M_{l,c}[\cdot]$ was the c channel of the light and shadow mask of the l layer. $G_{l,c}[\cdot]$ was the Gram matrix corresponding to $F_{l,c}[\cdot]$. The maximum number of channels of semantic masks for this study was 3, i.e. black labels for the first channel (c = 1), gray for the second channel (c = 2), and white for the third channel (c = 3). The three components, the combined content loss, the improved style loss, and the photo-realistic regularity term constituted the objective function of style transfer as shown in equation (7).

$$L_{total} = \sum_{l=1}^{L} \alpha_l L_c^l + \Gamma \sum_{l=1}^{L} \beta_l L_s^l + \lambda L_m \quad (7)$$

where λ was the parameter of the image realism regularity term, which usually took the value of 103. To preserve the image color during the photo styling process, this study introduced the idea based on the color control of the LC. After extracting the LC $L[P]$ and $L[A]$ in the target and reference images, a deep photo style transfer algorithm was applied to these images to produce a new luminance map $L[X]$. This method preserved the color of the original image and effectively avoided the influence of the color of the reference image as shown in Figure 2 [23]. The light and shadow transfer algorithm proposed in this research was defined as SCSS-PR-LC. The target image and the reference image underwent Lab color space conversion to extract the brightness channel L. Then, the brightness channel separated from color information was input into the VGG19 network for style transfer, generating a new brightness map. The new brightness map retained the original color information of the target image while incorporating the light and shadow style of the reference image. In the final synthesis stage, the new image combined style features with the original color tone, achieving visually realistic color transfer with reference image lighting effects.

Preprocessing, feature extraction, and light detection model construction for landscape images

Landscape design involves various complex environmental factors and detailed information. To achieve better design results, a series of image processing is required before light and shadow transfer, aiming to facilitate smooth style and light and shadow transfer in the future. To accurately transfer the light and shadow of the reference image to the corresponding position of the target image, semantic segmentation of the target image and the reference image is required to obtain a light and shadow mask before the light and shadow transfer. The light-shadow mask is then used to guide the process of light-shadow transfer so that the light-shadow is transferred in the corresponding region. It is very difficult to extract the light-shadow mask for the target image because there is no light-shadow on it [24]. To solve this problem, by aligning the reference image with the target image, the distribution of the detail information of the two images is approximately the same, so that the target image can use the same light and shadow

mask as the reference image [25]. The landscape image alignment method in this research was mainly divided into two steps. First, an image feature point extraction algorithm based on LBF regression was used to extract feature points from the image. Then, using the image feature points as the control points, the reference image was deformed using a feature line-based image alignment algorithm to achieve image alignment. Finally, the MRF light and shadow detection method was used to extract the three-dimensional mask. Therefore, the feature extraction i.e. light and shadow detection model for landscape images could be written as LBF-MRF. LBF was a regression-based feature point extraction algorithm for LBFs, which was mainly divided into training and testing two parts. Among them, training was the most important step, which had two main core steps. The first used random forest to learn the LBF mapping function, which was used to extract LBF. The second used the obtained LBFs to train the global linear regression function, which was used to predict deformation. Assuming the feature mapping function was $\Phi^t = \left[\phi_1^t, \phi_2^t, \mathrm{K}, \phi_L^t\right]$ and the regression matrix was W^t. The LBF was mainly trained in two consecutive steps. First, each set of locally mapped features in Φ^t needed to be trained separately with the aim of returning to the true shape increment ΔS^t as shown in equation (8).

$$\min_{w^t, \phi_l^t} \sum_{i=1} \left\| \pi_l \circ \Delta S_i^t - w_l^t \phi_l^t \left(I_i, S_i^{t-1} \right) \right\|_2^2 \tag{8}$$

where i was the i-th data in the training dataset. ΔS^t was the increment of the current shape compared with the real one. $\pi_l \circ \Delta S_i$ was the actual offset of the data i at the l-th feature point. The first term of equation (8) represented the increment of the current shape compared to the real shape. The second term represented the increment of the current shape compared to the predicted trait. This model first extracted features from local regions of the original image, and a binary code was generated for each local region. Multiple LBFs were combined into a global feature vector through concatenation operations. By using this feature merging method, the feature distribution of each local region could be reflected, which could then be used for further image analysis and recognition. LBFs could be obtained after learning using random forest training. After this, the global linear regression mapping function W^t could be learned according to equation (9).

$$\min_{w^t} \sum_{i=1}^{N} \left\| s_i^t - w^t \Phi^t \left(I_i, S_i^{t-1} \right) \right\|_2^2 + \lambda \left\| W^t \right\|_2^2 \tag{9}$$

where the first term denoted the regression target and the second term denoted the *L2* norm regularization term on W^t. λ was used to regulate the degree of regularization. In the feature line-based image deformation algorithm, the feature correspondence between the original image and the target image was defined using some specific line segments. The correspondence of the pixel points of the line segments was determined by the distance from the corresponding pixel point to the line segment [26, 27]. The inverse mapping was used to compute the deformation of the image by processing each pixel in the target image one by one to find the corresponding pixel in the original image. The deformation method of inverse mapping was used (Figure 3a). For each pixel Y in the target image, two variables u and v could be found on each feature line, and then go to the corresponding feature line on the original image and bring in u and v. The corresponding pixel Y of the target image in the original image could be found, and the solution process was shown in equation (10).

$$X' = P' + u\left(Q' - P'\right) + \frac{v \cdot Perpendicular\left(Q' - P'\right)}{\left\|Q' - P'\right\|} \tag{10}$$

where W^t denoted the return of a vector perpendicular to the input vector with the same length as the input vector [28]. When deforming an image, the deformation of the whole image could not be accomplished using only the single line segment pair algorithm, so the constrained

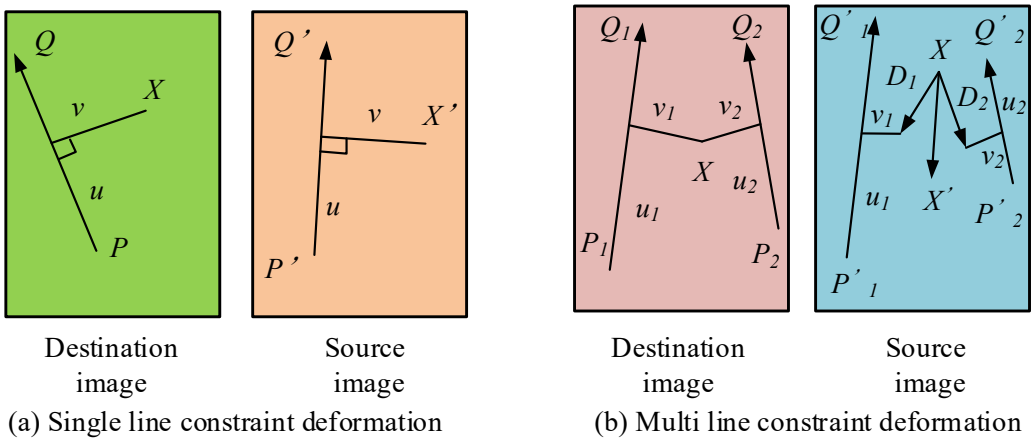

Figure 3. Single line segment constraints and multiple line segment constraints.

deformation of multiple line segments was derived from the single line segment pair algorithm Figure 3b). This study determined the position of the target image corresponding to the original image based on the weighted average method, and the formula for solving the weights was shown below.

$$weight[i]=\left(\frac{length[i]^{p}}{a+diat[i]}\right)^{b} \tag{11}$$

where *length* denoted the length of the current line segment. *dist* denoted the distance from the target image to the current line segment. *a*, *b*, *p* were the constants for controlling the distortion, respectively. The corresponding pixel point *Y'* could be solved by weighted average method as follows.

$$Y'=\frac{w_1\cdot Y_1^{'}+w_1\cdot Y_1^{'}+w_1\cdot Y_1^{'}+\mathrm{L}\ +w_1\cdot Y_1^{'}}{w_1+w_2+\mathrm{L}\ +w_i} \tag{12}$$

If the pixel point *Y'* obtained by equation (12) was not an integer, it could be further solved by bilinear interpolation. The light and shadow detection problem of an image could be considered as the problem of labeling the pixel points as light and non-light areas, i.e., the problem of obtaining the distribution of different classes of image pixels from the image. In this research, MRF light and shadow detection method was used to extract the three-part mask [29]. The label estimation of a pixel could be expressed by equation (13).

$$m=\arg\max_{M}\left\{P\left(n\left|m\right.\right)P\left(m\right)\right\} \tag{13}$$

where *m* was the realization of random field MRF, which represented the category label that each pixel point in the image was tagged with by maximum a posteriori probability (MAP). Assuming that the category label was $m = l_n$ and defining that the pixel gray value y_i was a Gaussian distribution of the parameter $\theta_i = \{\mu_t, \sigma_t\}$. μ_t denoted the mean value of the area labeled as *l* category, and σ_t denoted its standard variance. Given this, equation (14) is obtained.

$$P\left(y_i\left|m_i\right.\right)=\left(1/\sqrt{2\pi\sigma_t^2}\right)\cdot\exp\left(-\left(y_i-\mu_t\right)^2/2\sigma_t^2\right) \tag{14}$$

MAP-based image segmentation algorithms typically computed the optimal set of labels *M* that maximized the posterior probability distribution about *M*. The iterated conditional mode (ICM) algorithm was used to control the iterations, and the labels were updated by making the conditional probability approach the maximum value at each iteration as shown in equation (15).

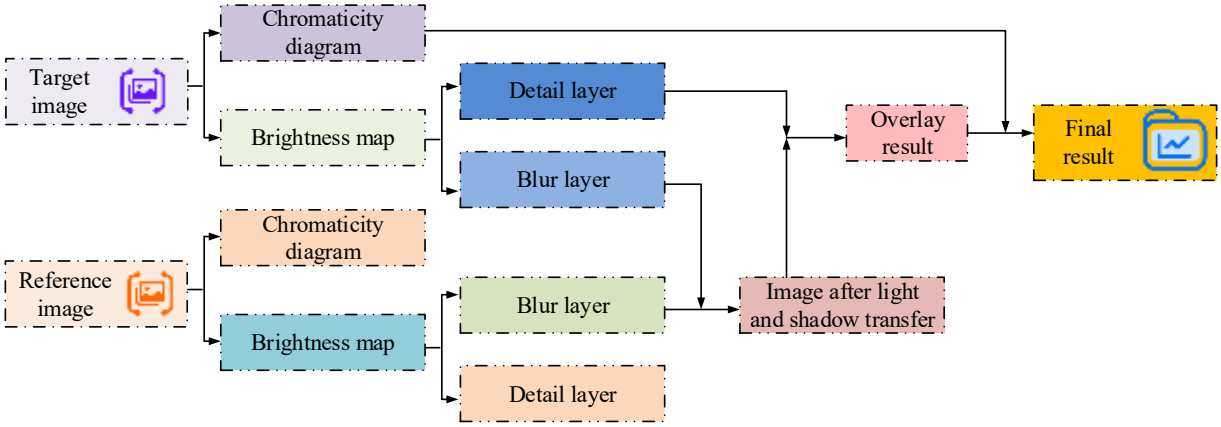

Figure 4. SCSS-PR-LC-LBF-MRF light and shadow transfer flowchart.

$$m_i = \arg^{\max}_{x_i \in \{1,2,\mathrm{L},C\}} \left\{ f\left(y_i \middle| m_i\right) \cdot f\left(m_i \middle| m_{Ni}\right) \right\}, \ i \in S \quad (15)$$

The final landscape light transfer model constructed in this research was named SCSS-PR-LC-LBF-MRF (Figure 4). The image alignment was performed to extract the light and shadow mask of the reference image before light and shadow transfer was performed using LC and depth photo light and shadow transfer algorithms based on LC and depth photo light and shadow transfer algorithms. The datasets used for image preprocessing and the information of the reference model for comparison were Places2 Challenge Dataset (Massachusetts Institute of Technology (MIT), Cambridge, MA, USA) with 8 million images covering 365 scenes dedicated to image restoration and scene recognition tasks, Cityscapes Dataset (Munich, Germany) covering street view structured annotated data, and COCO dataset jointly maintained by Microsoft, Caltech, and other institutions with annotated resources for object detection, segmentation, and other tasks. In addition, the study compared proposed SCSS-PR-LC-LBF-MRF model with LeNet-5 (New York, NY, USA), GoogleNet (Mountain View, California, USA), and ResNet (Redmond, Washington, USA) models [30-32].

Results and discussion

SCSS-PR-LC-LBF-MRF model training

To verify the feasibility and advancement of the proposed model, the study was analyzed through a variety of comparative experiments. The results of loss function and accuracy variation of the four algorithms demonstrated that the loss value of SCSS-PR-LC-LBF-MRF reached its lowest point at the 123rd iteration and quickly stabilized (Figure 5a), indicating its superior convergence performance. In comparison, LeNet-5, GoogleNet, and ResNet models were noticeably slower in convergence speed with more iterations than SCSS-PR-LC-LBF-MRF and had not achieved the same level of fast steady state. LeNet-5 and GoogleNet still showed significant fluctuations in their loss values after 150 iterations. The recognition accuracy of SCSS-PR-LC-LBF-MRF remained above 90% throughout the entire training period and reached the highest level of 97.23% after nearly 300 iterations, significantly better than other algorithms (Figure 5b). The recognition accuracies of LeNet-5, GoogleNet, and ResNet were relatively low, especially ResNet that had an accuracy of only about 80% after 300 iterations. This comparison further indicated that the application of SCSS-PR-LC-LBF-MRF in environmental landscape design had significant advantages, not only in convergence speed but also in recognition accuracy. To further highlight the robustness and generalization performance of the model, the study conducted further experiments on four datasets including seasonal and weather

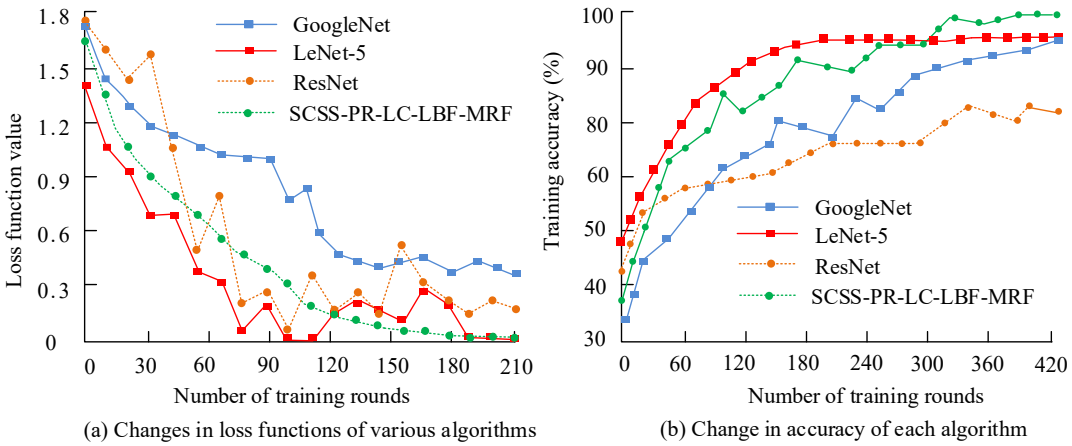

Figure 5. Loss function and accuracy variation results of four algorithms.

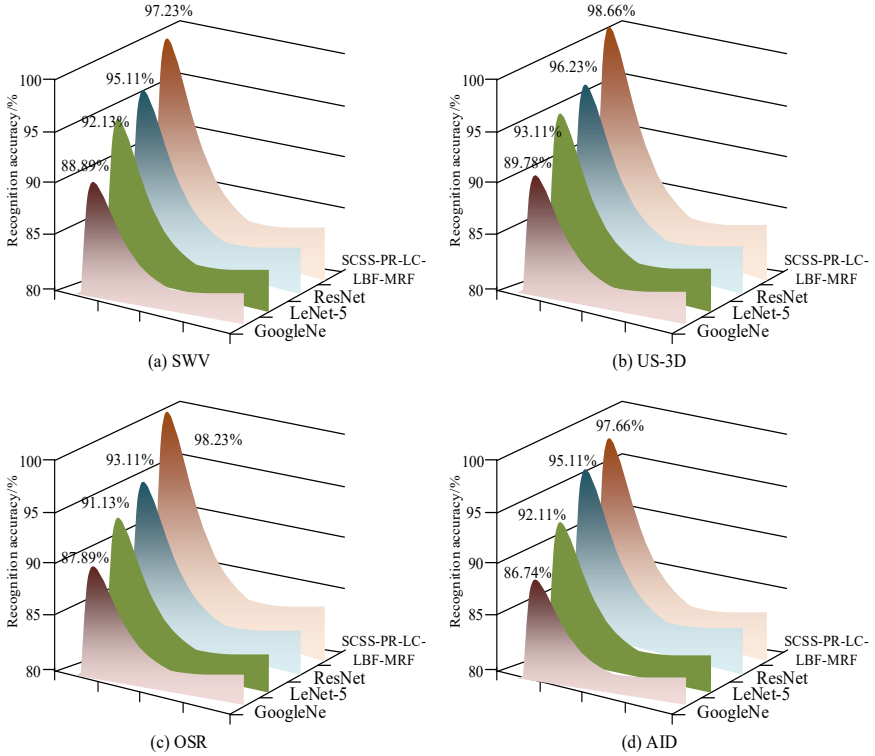

Figure 6. Recognition accuracies of four models for four datasets.

variations (SWV), UrbanScene3D (US-3D), CVPR outdoor scene recognition (OSR), and aerial image dataset (AID). Among them, SWV included landscape images of seasonal changes and different weather conditions. US-3D contained three-dimensional landscape images of urban environments, suitable for training with multiple light sources and complex backgrounds. OSR offered a diverse range of outdoor scenes including urban, natural, and mixed environments. AID contained high-resolution remote sensing images of multiple scene categories, suitable for landscape classification and recognition tasks. The results showed that

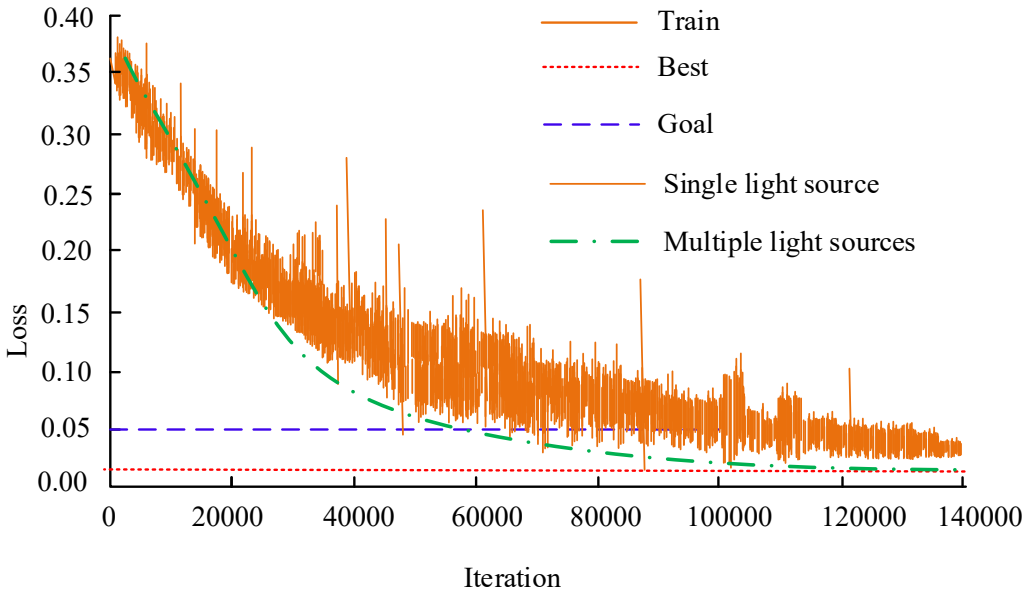

Figure 7. Convergence process of SCSS-PR-LC-LBF-MRF model under two light sources.

the accuracy of the SCSS-PR-LC-LBF-MRF model reached 98.66%, while the accuracies of LeNet-5, GoogleNet, and ResNet were 92.34%, 94.89%, and 96.15%, respectively (Figure 6a). The recognition accuracy of the SCSS-PR-LC-LBF-MRF model was 97.54%, while LeNet-5, GoogleNet, ResNet were 91.27%, 93.76%, and 95.08%, respectively (Figure 6b). The accuracies of the SCSS-PR-LC-LBF-MRF model, LeNet-5, GoogleNet, and ResNet were 97.81%, 90.23%, 93.14%, and 94.79%, respectively (Figure 6c). The recognition accuracies of the SCSS-PR-LC-LBF-MRF model, LeNet-5, GoogleNet, and ResNet were 98.32%, 90.89%, 92.68%, and 95.02%, respectively (Figure 6d). The data analysis showed that the SCSS-PR-LC-LBF-MRF model performed well on all datasets with an accuracy rate of over 97%, far higher than the other three models, demonstrating its superior performance and good generalization ability on different datasets. The impact of different image acquisition methods was explored on the model accuracy. The experiment continued to augment the data with images, and the final total number of images was 5,000. The process diagram of the network convergence situation during the training process showed that the loss value of the image obtained by a single light source showed significant fluctuations and converged slowly during the training process (Figure 7). It was not until about 135,000 iterations that the loss value gradually stabilized. This long-term unstable fluctuation indicated that images obtained from a single light source had adverse effects on model training with poor convergence and difficulty in quickly reaching the optimal state. In contrast, the model loss value of multi-light source image acquisition showed a smoother and more stable trend throughout the entire iteration process, which reached the plateau period at approximately 80,000 iterations, significantly earlier than the convergence time of single light source image acquisition. The results confirmed that multi-light source image acquisition could achieve faster convergence in model training, and the convergence process was more stable, thereby significantly optimizing the model. Therefore, the SCSS-PR-LC-LBF-MRF model had better convergence and stability under multiple light source conditions, thereby enhancing the application effect of light and shadow vision technology in environmental landscape design.

SCSS-PR-LC-LBF-MRF model based light and shadow transfer practical application effect

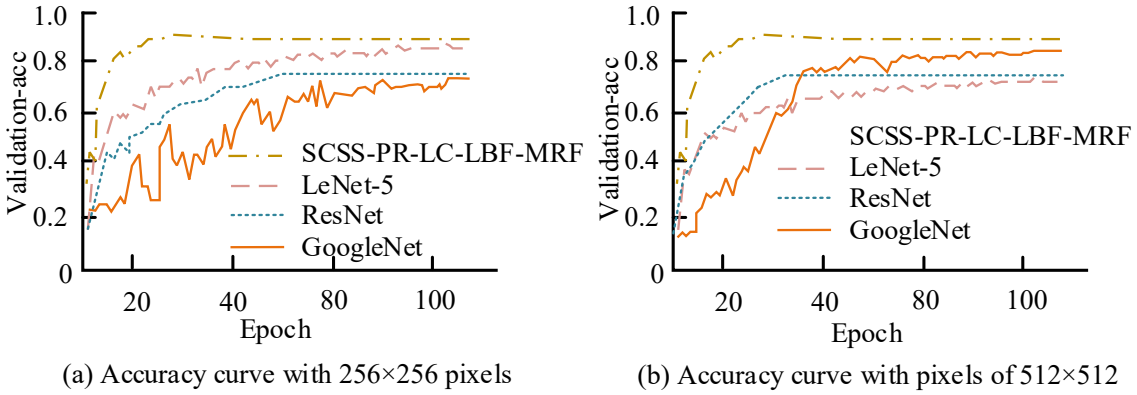

Figure 8. Performance comparison results of different models.

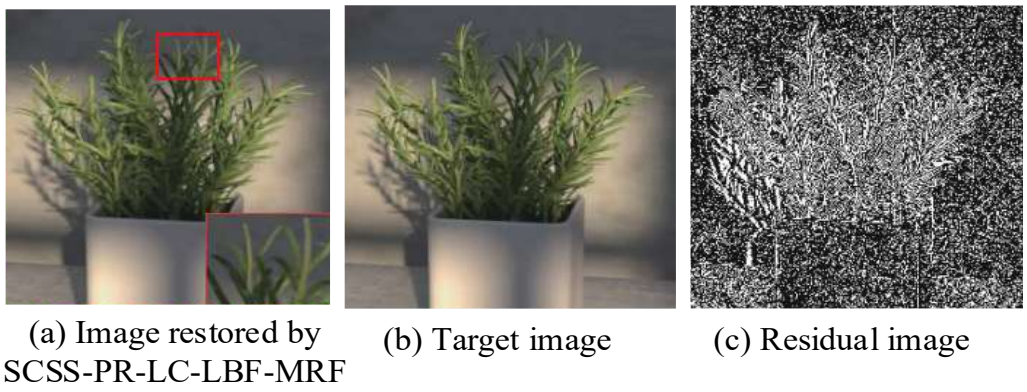

Figure 9. Effect of SCSS-PR-LC-LBF-MRF light and shadow transmission.

To verify the comparison between the optimized shallow and lightweight CNN algorithms for light and shadow transfer, the lightweight performance was analyzed and improved through a total of 200 iterations. The performance comparison results of different models showed that the SCSS-PR-LC-LBF-MRF algorithm demonstrated high accuracy and stability. The curve of SCSS-PR-LC-LBF-MRF maintained a high position during the iteration process with small fluctuations, which demonstrated its superior performance compared to LeNet-5, GoogleNet, and ResNet (Figure 8). Especially, compared to the original LeNet-5 structure, several performance improvements could be observed. The emergence of SCSS-PR-LC-LBF-MRF helped to reduce the use of computing resources while maintaining high accuracy, thus enabling efficient deep learning tasks to run on devices with limited computing power. Scenes in the dataset containing self-obscuring targets and rich in high and low frequency information were selected for the test of effect of light and shadow transfer using the images recovered by the proposed algorithm, the target image, and the residual image, respectively. The virtual angle image recovered by the proposed algorithm showed that the edge information recovered in the self-obscuring region of the tree leaves was more obvious and the low frequency information was more consistent (Figure 9). Therefore, SCSS-PR-LC-LBF-MRF was able to pass on the light shading in the scene and its residuals were at a low level. To better show the advantages of the proposed algorithm in this research, the MOS method was used to quantify the quality of the user's experience. The results showed that the overall

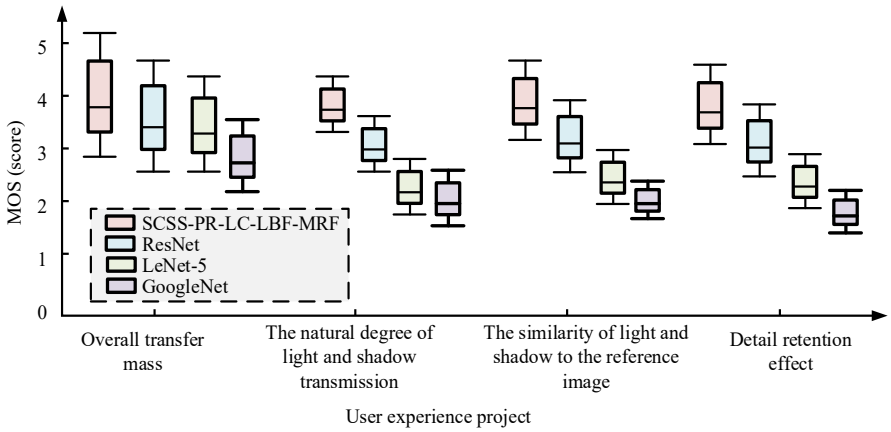

Figure 10. Comparison results of the MOS values of the four algorithms for the effect of landscape light and shadow transmission.

delivery quality MOS value of proposed model was higher relative to the other three algorithms with 6.2%, 9.4%, and 12.7% higher relative to ResNet, LeNet-5, and GoogleNet, respectively (Figure 10). The naturalness score of SCSS-PR-LC-LBF-MRF was 4.5, which was significantly better than that of ResNet's 4.2, LeNet-5's 4.1, and Google Net's 4.0. The SCSS-PR-LC-LBF-MRF method had a significant advantage in the naturalness of light and shadow transfer. In terms of similarity between the light and shadow of the reference image, SCSS-PR-LC-LBF-MRF scored 4.6 points, while ResNet, LeNet-5, and GoogleNet scored 4.2 points, 4.1 points, and 4.0 points, respectively, indicating that SCSS-PR-LC-LBF-MRF performed better in simulating real light and shadow. In terms of section retention effect, the SCSS-PR-LC-LBF-MRF method scored 4.7 points, higher than ResNet's 4.3 points, LeNet-5's 4.2 points, and Google Net's 4.1 points. The results validated the superiority of the SCSS-PR-LC-LBF-MRF method in various dimensions.

Conclusion

To realize the intelligent design of environmental landscapes, this study constructed a light and shadow SCSS-PR-LC-LBF-MRF model based on light and shadow transmission and landscape image light and shadow detection using CNN and validated the performance of the proposed model. The results showed that, when the iterations were carried out 123 times, the loss value of the proposed model dropped to the lowest point. As the number of iterations increased, the convergence of the proposed model tended to stabilize and outperformed LeNet-5, GoogleNet, and ResNet models. The results of image quality variation showed that the MSE of the images could reach the extreme value and the SSIM could reach a relatively large value at the number of training rounds around 34 for pixels of 256 × 256 and 512 × 512. The recognition accuracy of proposed model for the two datasets was higher than 95%, and the maximum was up to 98.66%. The practical application effect demonstrated that SCSS-PR-LC-LBF-MRF could recover the edge information of landscape images with consistent low-frequency information. In addition, SCSS-PR-LC-LBF-MRF could pass on the light shadows in the scene, and its residuals were also at a low level. The user MOS rating results of SCSS-PR-LC-LBF-MRF for the effectiveness of light and shadow transmission showed that its overall transmission quality MOS value was relatively high compared to the other three algorithms with 6.2%, 9.4%, and 12.7% higher than ResNet, LeNet-5, and GoogleNet, respectively. The naturalness of light and shadow transmission, similarity of light and shadow with reference images, and detail preservation effect of SCSS-PR-LC-LBF-MRF were all superior to the other three methods. The

limitation of this study was that the CNN was more time-consuming, which could be further optimized in the future.

References

1. Hong Y, Niu L, Zhang J. 2022. Shadow generation for composite image in real-world scenes. AAAI Conf Artif Intell. 36(1):914-922.
2. Gugliotti J, Landman MM. 2020. Flat dome lights tame reflections and shadows in machine vision. Vis Syst Des. 25(7):12-15.
3. Tsai TH, Chang CH, Chen S, Yao C. 2020. Design of vision-based indoor positioning based on embedded system. IET Image Process. 14(3):423-430.
4. Fenderson LE, Kovach AI, Llamas B. 2020. Spatiotemporal landscape genetics: Investigating ecology and evolution through space and time. Mol Ecol. 29(2):218-246.
5. Masnavi MR, Motedayen H, Saboonchi P, Hemmati M. 2021. Analyses of landscape concept and landscape approach from theoretical to operational levels: A review of literature. MANZAR. 13(57):22-37.
6. Wang Y, Hu X. 2020. Wuju opera cultural creative products and research on visual image under VR technology. IEEE Access. 8(1):161862-161871.
7. Edensor T, Hughes R. 2021. Moving through a dappled world: The aesthetics of shade and shadow in place. Soc Cult Geogr. 22(9):1307-1325.
8. Lara AL. 2021. Patagonia, lights and shadows: Territorial perceptions of a circumpolar space. Arctic Antarctic. 15(1):15-31.
9. Baslamisli S, Das P, Le HA, Karaoglu S, Gevers T. 2021. Shadingnet: Image intrinsics by fine-grained shading decomposition. Int J Comput Vis. 129(8):2445-2473.
10. Eren ET, Yılmaz S. 2022. The student attitudes towards digital and conventional drawing methods in environmental design studios and the impact of these techniques on academic achievement in the course. Int J Technol Des Educ. 32(1):617-644.
11. Eren ET, Alpak EM, Düzenli T. 2022. Color associations in landscape design and subscription levels to these associations. Environ Sci Pollut Res. 29(47):70842-70861.
12. Kerr J, Lawson G. 2020. Augmented reality in design education: Landscape architecture studies as AR experience. Int J Art Des Educ. 39(1):6-21.
13. Zhong W, Schröder T, Bekkering J. 2022. Biophilic design in architecture and its contributions to health, well-being, and sustainability: A critical review. Front Archit Res. 11(1):114-141.
14. Yu A, Altalbe N, Du N. 2022. Environmental landscape art design using dynamic nonlinear parameterization. Fractals. 30(2):2240077-2240089.
15. Yang X, He S, Zhang Q. 2020. Thermal comfort measurement and landscape reconstruction strategy of street space based on decomposition grid method: A case study of Yangmeizhu Street. J Landscape Res. 12(3):33-46.
16. Zhan F, Zhang C, Yu Y, Chang Y, Lu S, Ma F, *et al*. 2021. Emlight: Lighting estimation *via* spherical distribution approximation. AAAI Conf Artif Intell. 35(4):3287-3295.
17. Mussap M, Fanos V. 2021. Could metabolomics drive the fate of COVID-19 pandemic? A narrative review on lights and shadows. Clin Chem Lab Med. 59(12):1891-1905.
18. Sun H, Kuchenbecker KJ, Martius G. 2022. A soft thumb-sized vision-based sensor with accurate all-round force perception. Nat Mach Intell. 4(2):135-145.
19. He K, Gan C, Li Z, Rekikl YZ, Ji W, Gao Y, *et al*. 2023. Transformers in medical image analysis. Intell Med. 3(1):59-78.
20. Samanik S. 2021. Imagery analysis in Matsuoka's cloud of sparrows. Linguist Lit J. 2(1):17-24.
21. Tyagi S, Yadav D. 2023. A detailed analysis of image and video forgery detection techniques. Vis Comput. 39(3):813-833.
22. Nurani AM, Ozawa Y, Furuya T, Sakamoto Y, Ebine K, Matsunaga S, *et al*. 2020. Deep imaging analysis in VISUAL reveals the role of YABBY genes in vascular stem cell fate determination. Plant Cell Physiol. 61(2):255-264.
23. Golestani N, Khakzand M, Faizi M, Karimi B. 2020. Explaining the mutual relationship between landscape perception and participation in the process of participatory landscape development. The Monthly Sci J Bagh-e Nazar. 17(90):35-54.
24. Wang H, Li F. 2020. Application of landscape architecture design in the water conservancy project of the coastal city. J Coast Res. 112(SI):33-35.
25. Preethi P, Mamatha H. 2023. Region-based convolutional neural network for segmenting text in epigraphical images. Artif Intell Appl. 1(2):119-127.
26. Williams AE. 2023. Human-Centric functional computing as an approach to human-like computation. Artif Intell Appl. 1(2):118-137.
27. Wu L, He L, Chen W, Hao X. 2023. Recognition method of voltage sag sources based on RMT-CNN model. IAENG Int J Appl Math. 53(3):163-173.
28. Murthy M, Koteswararao A, Babu M. 2022. Adaptive fuzzy deformable fusion and optimized CNN with ensemble classification for automated brain tumor diagnosis. Biomed Eng Lett. 12(1):37-58.
29. Yasinska-Damri L, Babichev S, Durnyak B, Goncharenko T. 2022. Check for application of convolutional neural network for gene expression data classification. Lect Notes Data Eng Commun Technol. 149(3):3-24.
30. Yadav B, Indian A, Meena G. 2024. Recognizing off-line devanagari handwritten characters using modified lenet-5 deep neural network. Procedia Comput Sci. 235:799-809.
31. Williams WA, Arunprasad A, Aravamudhan S. 2024. Application of a modified set of GoogLeNet and ResNet-18 convolutional neural networks towards the identification of environmentally derived-MPLs in the Yadkin-Pee Dee River Basin. Environ Syst Res. 13(1):1-17.
32. Yadav A, Kumar E. 2024. Object detection on real-time video with FPN and modified mask R-CNN based on inception-ResNetV2. Wireless Personal Commun. 138(4):2065-2090.