www.nature.com/scientificreports

scientific reports

Check for updates

OPEN

The intelligent development and preservation of folk sports culture under artificial intelligence

Zhihui Li[1,2], Shamsulariffin Bin Samsudin[1]✉, Noor Hamzani Farizan[3], Zulkhairi Azizi Zainal Abidin[4] & Lei Zhang[2]

To promote the intelligent development and preservation of folk sports culture, this work proposes a model grounded in the Cycle-Consistent Generative Adversarial Network (CycleGAN) to produce high-quality human images that recreate traditional sports movements. In order to improve the performance of the model, a discriminative mechanism for pose consistency and identity consistency is innovatively designed, and an appearance consistency loss function is introduced. Finally, the effectiveness of the model in image generation is verified. Experiments conducted on the DeepFashion and Market-1501 datasets suggest that compared to other models, the proposed model achieves superior visual quality and realism in the generated images. In ablation experiments, the model incorporating the appearance consistency loss achieves improvements of 1.49%, 1.76%, and 2.2% in image inception score, structural similarity index, and diversity score, respectively, compared to the best-performing comparative models. This demonstrates the effectiveness of this loss function in improving image quality. Moreover, the proposed model excels across multiple evaluation metrics when compared to other models. In authenticity discrimination experiments, the generated images have a 58.25% probability of being judged as real, significantly surpassing other models. In addition, the results on the folk sports culture action dataset also show that the model proposed performs excellently in multiple indicators, and it particularly has an advantage in the balance between image diversity and quality. These results indicate that the CycleGAN model better reproduces the details and realism of folk sports movements. This finding provides strong technical support for the digital preservation and development of traditional sports culture.

Keywords Artificial intelligence, Deep learning, Folk sports culture, Human image generation, Cycle-consistent generative adversarial network

Research background and motivations

Folk sports culture, as a unique cultural heritage of a country or region, embodies rich historical, cultural, and social values. These cultural forms are often deeply rooted in specific geographical environments and social structures, reflecting the unique forms of sports activities and cultural expressions developed by local communities through long-term practices[1–3]. However, with the rapid advancement of modernization, many folk sports are gradually declining, and some are even at risk of disappearing entirely[4–6]. Traditional folk sports not only offer diverse forms of activity but also carry the cultural memory and social bonds of specific regions[7,8]. The challenge of protecting, preserving, and revitalizing these unique cultural forms in modern society has become an urgent issue that demands resolution. Against this backdrop, exploring innovative methods and technical means, especially the application of information technology and Artificial Intelligence (AI) technologies, appears particularly crucial and urgent.

In recent years, the rapid development of AI technologies, especially deep learning technologies, has provided new ideas and practical approaches for the digital protection and inheritance of traditional culture[9,10]. Among them, Generative Adversarial Network (GAN), as an innovative image generation technology, has demonstrated its powerful potential in multiple fields. In particular, in human body image generation and action simulation, GAN can accurately reproduce the key postures of the human body[11–13]. With the help of these technologies,

[1]Department of Sports Studies, Faculty of Educational Studies, Universiti Putra Malaysia, Serdang 43400, Selangor, Malaysia. [2]Department of Physical Education and Health, Lvliang University, Lvliang 033000, China. [3]Defense Fitness Academy, National Defense University of Malaysia, sungai besi camp, Kuala Lumpur 57000, Malaysia. [4]Department of Nature Parks and Recrearion, Universiti Putra Malaysia, Serdang 43400, Malaysia. ✉email: 20141031@llu.edu.cn

researchers can retain the essence of traditional culture and reproduce the important movements in folk sports through vivid and intuitive image representation, thus achieving their digital protection. Therefore, how to promote the intelligent development and inheritance of folk sports culture through AI technologies, especially the image generation technology based on GAN, has become an important motivation and goal of this work.

Research objectives

The key purpose of this work is to explore how AI technology, particularly Cycle-Consistent Generative Adversarial Network (CycleGAN), can be utilized to achieve the intelligent preservation and development of folk sports culture. The specific goals are as follows:

(1) Develop a human image generation model based on CycleGAN to recreate and simulate key action postures in folk sports, providing a digital preservation method for these traditional cultural forms.
(2) Enhance the quality of the images produced by designing and optimizing the model's loss function. It intends to ensure that the generated images not only retain the appearance features of the source images but also accurately reflect the realism and consistency of the target postures.
(3) Validate the model through experiments and performance evaluation to explore its effectiveness in practical applications and assess its potential value and prospects for the preservation of folk sports culture.

The innovations and contributions of this work are as follows:

(1) A human body image generation model based on CycleGAN. The proposed CycleGAN model can not only generate high-quality human body images without paired data, but also accurately retain the details and sense of reality of traditional sports movements under complex changes in actions and postures. Through the application of this model, traditional folk sports movements can be digitally archived and inherited.
(2) Discriminative mechanism for pose consistency and identity consistency: To ensure the consistency of the generated images in terms of pose and identity, two discriminators are innovatively designed in the CycleGAN model: a pose discriminator and an identity discriminator. The pose discriminator can effectively identify whether the generated image accurately reflects the target pose, while the identity discriminator ensures that the generated image is consistent with the person in the source image. This discriminative mechanism greatly enhances the authenticity and consistency of the generated images.
(3) The innovative introduction of an appearance consistency loss function. In order to improve the quality and sense of reality of the images, an appearance consistency loss function is innovatively introduced within the CycleGAN framework. This loss function can effectively reduce the feature differences between the source image and the generated image, thereby enhancing the appearance consistency and detail fidelity of the generated images.

In summary, this work innovates in the application of AI technology, promotes the digital inheritance of folk sports culture, and provides new technical paths and theoretical support for the future intelligent protection of cultural heritage.

Literature review

Recently, due to the swift advancement of AI, the methods for preserving and passing down traditional culture have gradually shifted from traditional oral transmission and physical preservation to more intelligent and digital approaches. Many scholars have begun to explore how modern AI technologies can be utilized for the digital development, protection, and inheritance of traditional culture. Sun et al. (2022) proposed a GAN-based design strategy. Then, they optimized the renovation design process of traditional historic districts by automatically generating stylized façades. The findings demonstrated the model's advantages in terms of accuracy, realism, and diversity, while also validating its feasibility and adaptability in practical applications[14]. Wang (2022) studied how AI technology could be used to ensure the protection and preservation of the cultural landscape heritage of traditional villages, using Wuyuan County in Jiangxi Province as a case study. The analysis revealed that the cultural landscape of traditional villages was damaged due to weak protection, and AI technologies like image restoration and intelligent monitoring were proposed to effectively protect these heritages[15]. Kasmahidayat and Hasanuddin (2022) explored how AI technology could help pass on and develop the cultural arts of the archipelago. They used innovative strategies to enhance the protection and inheritance of Kuda Renggong art and analyze the impact of technological advancements on these cultural arts[16]. These studies provide the technical application background in the field of folk culture protection for this work. Especially, the innovative applications of AI and GAN technologies in the protection of cultural heritage have inspired the application of AI to the generation and protection of the movements in folk sports culture. They offer references for the technical approaches of traditional culture protection, particularly regarding the role of AI image generation technology in the process of protection and inheritance.

Image generation technology in AI has provided a novel pathway for the transmission and development of traditional culture. Scholars have conducted extensive research on image generation technology. Ding et al. (2021) designed a text-to-image model based on Transformers, achieving state-of-the-art results in text-to-image tasks and demonstrating its potential in applications like style learning[17]. Wang et al. (2022) proposed a magnetic resonance imaging enhancement method combining deep transfer learning and digital twin technology. By introducing composite metasurfaces and an adaptive decomposition multi-modal image fusion algorithm, this method showed excellent performance in image realism, clarity, and color retention[18]. Saharia et al. (2022) developed a text-to-image generation model that combined large language models with diffusion models, significantly improving the realism of generated images and the consistency between text and image[19].

The research on these image generation technologies demonstrates the potential of modern AI technologies in cultural inheritance, especially their advantages in aspects such as style transfer and image quality improvement. They provide technical references for the CycleGAN model proposed here. In particular, the generation strategies when dealing with complex visual information and the technical methods for enhancing image quality are of great reference value for the design and implementation of the image generation model proposed.

In the protection and inheritance of culture, Bei (2023) utilized computer vision, deep learning, and GAN technologies to innovate the design of paper-cut patterns, achieving automated creation while maintaining the cultural value of paper-cutting[20]. Garozzo et al. (2021) used GANs to generate images of cultural heritage. By combining semantic ontology and data-driven strategies, they hierarchically generated independent objects and combined them into realistic scenes to make up for the shortage of training data. Moreover, this was applied to improve the image classification and retrieval system[21]. Kumar and Gupta (2023) proposed an art restoration method based on GANs. Using an improved U-Net architecture and a pre-trained residual network, it can achieve high-quality digital restoration of damaged artworks and assist in their actual restoration[22]. These studies reveal the extensive applications of GAN technologies in the restoration of traditional art and cultural heritage, especially in aspects such as generating realistic images, enhancing image quality, and restoring damaged artworks. This work draws on these methods to explore how to retain the unique elements and forms of traditional culture through precise image generation technologies when generating images of folk sports movements. These works provide technical inspiration for the innovative model here and help to understand how to combine GAN technologies with specific cultural backgrounds to achieve effective cultural protection.

The above-mentioned literatures offer profound insights into the application of AI technologies in different fields, especially in the protection of cultural heritage, and the generation and restoration of traditional art. They provide a solid theoretical foundation and technical support for the research direction of this work. In particular, the application of GAN technologies in image generation, cultural protection, and inheritance provides valuable references for the human body image generation model based on the CycleGAN proposed. However, there are still some deficiencies. First, most of the research focuses on general image generation tasks or applications in a single cultural field, lacking accurate modeling of specific cultural backgrounds, especially the unique movements and postures in folk sports culture. Then, although some studies have adopted GAN technologies, there has been no systematic and in-depth discussion on how to apply this technology in the digital preservation of folk sports culture. It is especially obvious in terms of the generation of dynamic movements and the retention of cultural elements. Therefore, based on the previous research, this work proposes a human body image generation model based on CycleGAN, aiming to achieve high-quality generation and inheritance of the classic movements in folk sports culture. By introducing a network structure optimized specifically for complex posture transformation, this work solves the technical problems in the generation of traditional culture images. Besides, it provides an innovative technical path for the digital protection and reproduction of folk sports culture.

Research model

Analysis of the principles of cyclegan

With the rapid progress of AI technology, GAN has been widely applied in many fields such as image generation, style transfer, and image enhancement. Compared with traditional deep learning methods, GAN can generate high-quality images from scratch through adversarial training, and it has extremely strong expressive and generative capabilities. Among them, as an unsupervised learning method, CycleGAN is particularly suitable for situations where paired data are lacking. It can achieve high-quality image conversion between two different domains without the need for labeled data[23–25]. The sports movements in folk sports culture are usually unlabeled and diverse, and it is very difficult to obtain paired data. The CycleGAN model, which can effectively convert source images into target images without the need for paired data, thus becomes an ideal choice. The core idea of CycleGAN is the introduction of cycle consistency loss. This confirms that an image moved from one domain to another is able to be effectively mapped back to the original domain, thereby enhancing the accuracy and fidelity of the transformation[26,27]. The cycle consistency loss of CycleGAN can ensure that the mapping from the source domain to the target domain has a high degree of accuracy. Besides, it can ensure that it maintains a high level of consistency when mapping back from the target domain to the source domain. This is crucial for retaining the details and sense of reality of traditional sports movements. CycleGAN can not only generate images with high visual effects but also handle the complex problems of movement and posture transformation in folk sports culture. Compared with traditional generation models, CycleGAN can reconstruct the details of traditional sports movements more flexibly and accurately, and it has a strong cross-domain mapping ability, which is particularly suitable for the unaligned data conditions here. Figure 1 illustrates the structure of CycleGAN and its key cycle consistency loss mechanism.

The CycleGAN structure is composed of two sets of generators and discriminators: Generator G_1 and G_2, and Discriminator D_A and D_B. The entire CycleGAN structure can be described as a combination of two adversarial networks, where the generators and discriminators continuously compete to enhance the generated images' quality[28–30]. The goal of the Generator G_1 is to map a sample *a* from domain *A* to domain *B*, and the Discriminator D_B verifies whether the generated sample $G_1(a)$ exhibits the characteristics of the value domain. Similarly, the Generator G_2 maps a sample *b* from domain *B* back to domain *A*, and the Discriminator D_A verifies the authenticity of the generated sample $G_2(b)$. This bidirectional generation process can be expressed as:

$$a \underrightarrow{G_1} G_1(a) \underrightarrow{G_2} G_2(G_1(a)) \approx a \tag{1}$$

$$b \underrightarrow{G_2} G_2(b) \underrightarrow{G_1} G_1(G_2(b)) \approx b \tag{2}$$

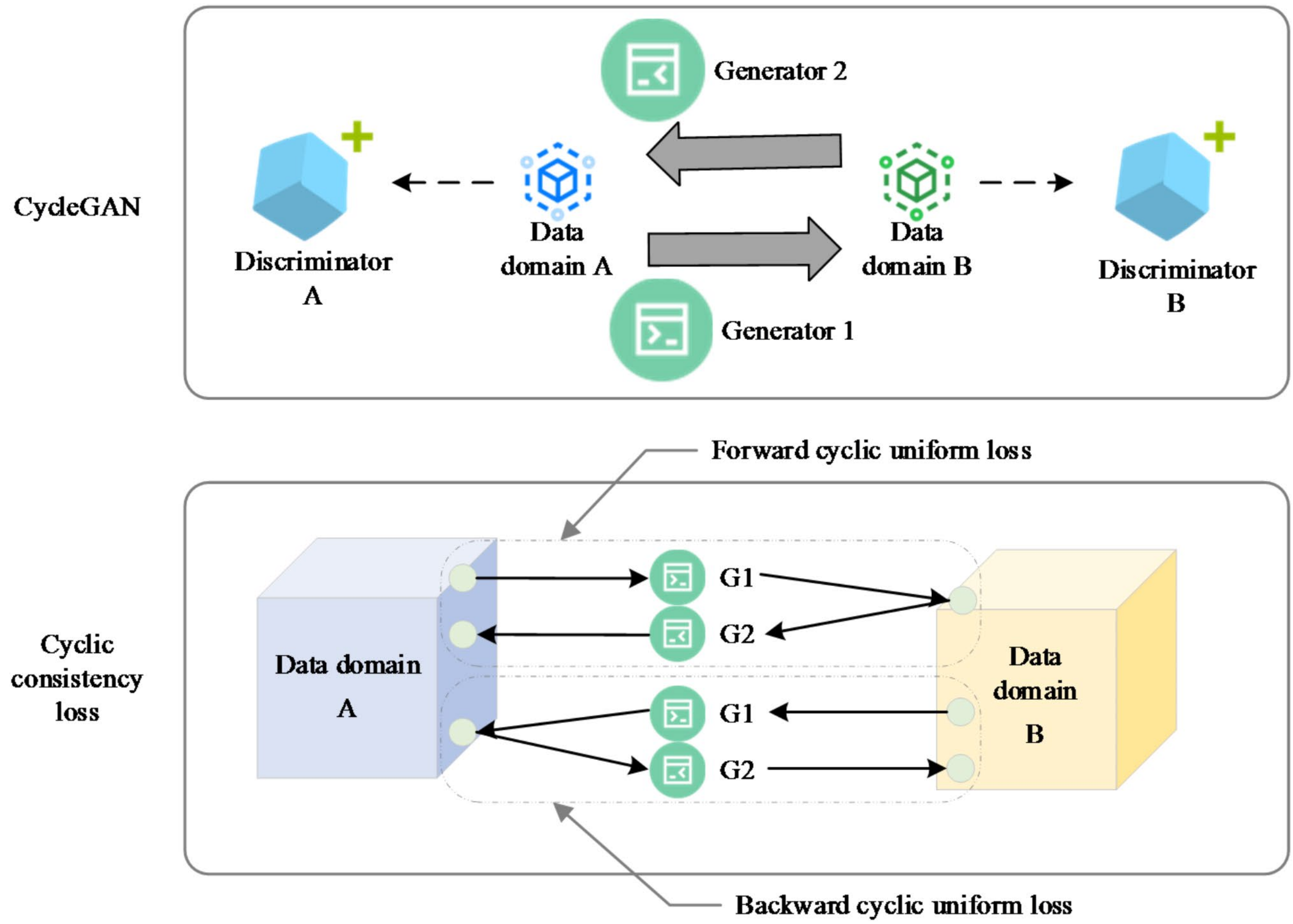

Fig. 1. CycleGAN Structure and Schematic of Cycle Consistency Loss.

$G_1(a)$ is the generated image obtained by mapping the sample a from domain A to domain B, while $G_2(G_1(a))$ represents the reconstructed image obtained by mapping the generated image back to domain A. This cyclical operation ensures that CycleGAN retains the primary features of the original sample while transforming it between different domains.

CycleGAN's core includes two types of losses: adversarial loss and cycle consistency loss. The former is adopted to optimize the competition between the generator and the discriminator, driving the generator to produce increasingly realistic images[31,32]. The adversarial loss between Generator G_1 and Discriminator D_B is expressed as:

$$\mathcal{L}_{GAN}(G_1, D_B, A, B) = \mathbb{E}_{b\sim P_{data(b)}}[\log D_B(b)] + \mathbb{E}_{a\sim P_{data(a)}}[\log(1 - D_B(G_1(a)))] \tag{3}$$

Similarly, the adversarial loss between Generator G_2 and Discriminator D_A is expressed as:

$$\mathcal{L}_{GAN}(G_2, D_A, B, A) = \mathbb{E}_{a\sim P_{data(a)}}[\log D_A(a)] + \mathbb{E}_{b\sim P_{data(b)}}[\log(1 - D_A(G_2(b)))] \tag{4}$$

The cycle consistency loss is a key innovation in CycleGAN. It constrains the variation between the generated image and the reconstructed image to guarantee that the generated image faithfully retains the features of the original image[33,34]. This loss comprises two elements: forward cycle consistency loss and backward cycle consistency loss, and is defined as:

$$\mathcal{L}_{cyc}(G_1, G_2) = \mathbb{E}_{a\sim P_{data(a)}}[\|G_2(G_1(a)) - a_1\|] + \mathbb{E}_{b\sim P_{data(b)}}[\|G_1(G_2(b)) - b_1\|] \tag{5}$$

$\|\cdot\|_1$ denotes the L_1 norm, which calculate the variation between the generated image and the original image. By minimizing this loss, CycleGAN ensures consistency in both content and style during the cross-domain transformation.

Overall, the total objective function for CycleGAN can be expressed as:

$$\mathcal{L}(G_1, G_2, D_A, D_B) = \mathcal{L}_{GAN}(G_1, D_B, A, B) + \mathcal{L}_{GAN}(G_2, D_A, B, A) + \delta\,\mathcal{L}_{cyc}(G_1, G_2) \tag{6}$$

δ represents a weight coefficient. By jointly optimizing these loss functions, CycleGAN is able to achieve high-quality cross-domain image generation under unsupervised conditions.

Human image generation model using cyclegan

In the aforementioned Sect. 3.1, the basic principles of CycleGAN and its mechanism for achieving high-quality image conversion in unsupervised learning are analyzed in detail. Based on the CycleGAN framework, this work proposes an improved model for human body image generation, aiming to accurately model complex pose changes and generate high-quality images with target pose features. To achieve this goal, it optimally designs the generator by integrating multiple Pose-Aware Texture Blending (PATB) modules. This is to better map the appearance features of the source image to the target pose. This improvement enables the model not only to perform conversions between different poses but also to effectively preserve the details and realism of the source image, thereby providing higher adaptability and accuracy for the human body image generation task. Next, the specific structure and implementation details of the human body image generation model based on CycleGAN are introduced.

The aim of the generator is to generate an image featuring the target pose based on the source image *x* and the target pose feature *y*, while preserving the appearance texture of the source image. This process incorporates multiple PATB modules. In the PATB module, the generator is split into two primary components: appearance texture encoding and pose encoding. The key to the appearance texture encoding path is the texture transformation of the source image using the pose attention mask M_t. The pose attention mask M_t contains both the original pose features and the target pose features, which can be represented as:

$$M_t = \sigma\ (conv(\ F^-\ )) \tag{7}$$

σ represents the activation function, $conv$ denotes the convolution operation, and F^- is the appearance feature of the source image. The mask M_t performs element-wise multiplication with the features of the source image and combines with residual connections to effectively transform the appearance texture. The appearance texture transformation is represented as:

$$F_t = econ\left(conv\left(F^-\right)\right) + F^- \tag{8}$$

F_t represents the feature image after the appearance texture transformation, and $econ$ denotes the appearance texture encoding module.

In the pose encoding path, the pose feature P_t is adjusted as the appearance texture features are updated. The pose encoding path is represented as:

$$P_t = conv\,(P)\,||\,F_t \tag{9}$$

P represents the initial pose feature. In the generator, each PATB module is responsible for transforming different regions of the image, and multiple PATB modules are concatenated to achieve the target pose transformation for the entire image.

The primary function of the discriminator is to distinguish whether the input image is a real image or a generated image, and to assist in improving the generator through adversarial learning[35–37]. To enable the discriminator to more effectively assess image quality and pose consistency, a pose discriminator D_p and an identity discriminator D_i are designed. The identity discriminator D_i primarily determines whether the generated image corresponds to the same person as the source image. In the absence of paired data, the identity discriminator operates only during the reverse mapping (from the target domain to the source domain) to guarantee that the generated image retains the appearance of the source image. The pose discriminator D_p is adopted to identify whether the pose of the generated image matches the target pose. D_{p1} and D_{p2} represent the pose discriminators in the two sets of adversarial networks in CycleGAN. The design of the pose discriminator is based on the ResNet architecture to enhance its performance in complex tasks. The objective function for the pose discriminator can be represented as:

$$\mathcal{L}_{D_p}\left(D_{p1}, D_{p2}\right) = \mathbb{E}_{y\sim P_{data(y)}}\left[\log D_{p1}\left(G(x,y)\right)\right] + \mathbb{E}_{x\sim P_{data(x)}}\left[\log(1 - D_{p1}\left(G(x,y)\right))\right] \tag{10}$$

D_{p1} and D_{p2} respectively assess the authenticity of the generated image in the target pose.

The model of Structure of the Human Image Generation Model Based on CycleGAN is mainly used to generate images of typical human body movements in folk sports culture, including the key postures of traditional sports events such as martial arts, dragon dance, lion dance, cuju (ancient Chinese football), and ethnic wrestling. By inputting the source image and the target posture, CycleGAN can accurately transform the posture while maintaining the appearance features of the person. As a result, the generated images not only conform to the real movement trajectories but also showcase the unique style of folk sports culture. In addition to enhancing the quality of the generated images, this design strengthens the model's adaptability to complex pose transformations, providing strong technical support for the digital preservation of traditional sports culture.

Model loss function design

With the purpose of improving the quality of the images generated by the model, three types of loss functions are designed: adversarial loss, reconstruction loss, and appearance consistency loss.

Adversarial loss
Adversarial loss is used to guide the adversarial process between the generator *G* and the discriminator *D*. In the CycleGAN model, there are two sets of adversarial losses. They correspond to the transformation from the source domain to the target domain ($X \to Y$) and the reverse transformation from the target domain to the source domain ($Y \to X$). Specifically, the adversarial loss $L_{X\to Y}$ for the transformation from the source domain to the target domain is defined as:

$$L_{X\to Y} = \mathbb{E}_{K\sim P_{data(K)}} [\log D_{p1} (K)] + \mathbb{E}_{K\sim P_{data(K)}} [\log(1 - D_{p1} (G (K| P_t)))] \tag{11}$$

D_{p1} is the pose discriminator, and $G (I| P_t)$ denotes the image generated by the generator at the target pose P_t. The adversarial loss $L_{Y\to X}$ for the transformation from the target domain to the source domain is defined as:

$$L_{Y\to X} = \mathbb{E}_{K\sim P_{data(K)}} [\log(D_{p2} (K) \cdot D_i(K))] + \mathbb{E}_{K\sim P_{data(K)}} [\log(1 - D_{p2} (K_t)) \cdot (1 - D_i(K_t))] \tag{12}$$

$\mathrm{D_{p2}}$ is another pose discriminator, and $\mathrm{D_i}$ is the identity discriminator. $\mathrm{I_t}$ denotes the image mapped back to the source domain from the target pose $\mathrm{P_t}$ through the generator G. It can be defined as:

$$K_t = G (G((K| P_t)| P_C) \tag{13}$$

P_C represents the pose condition or source domain condition.

Reconstruction loss
The reconstruction loss is used to ensure the stability and robustness of the generated images during the transformation process. This loss measures the pixel-wise difference between the generated image and the source image. The L_1 loss function is used to calculate this difference, and it is defined as:

$$L_{pixel} = \mathbb{E}_{K\sim P_{data(K)}} \left[\left\|\hat{K} - K_1\right\|\right] \tag{14}$$

$\widehat{K}$ is the image generated by the generator, and K is the real source image.

Appearance consistency loss
To ensure that the generated images are visually consistent with human perception, an appearance consistency loss is proposed. This loss includes both texture consistency and structural consistency losses[38,39]. The goal of appearance consistency loss is to make the generated image visually similar to the target image.

The texture consistency loss L_v is based on the perceptual loss using the Visual Geometry Group 19-layer (VGG19) Network, which involves extracting high-level feature maps from both the generated image and the target image. The equation is:

$$L_v = \sum_{j\in\varphi} \frac{1}{C_j H_j W_j} \left\|Q_j (K) - Q_j \left(\hat{K}\right)\right\|_2^2 \tag{15}$$

$Q_j (K)$ and $Q_j \left(\widehat{K}\right)$ denote the feature maps extracted from the VGG19 network at layer *j* for the target image and the generated image, respectively. The set ϕ represents the subset of layers in the network. C_j, H_j, and W_j represent the number of channels, height, and width of the feature maps, respectively.

The structural consistency loss L_s is designed to ensure that the generated image does not suffer from distortions during the pose transformation. It is based on the least squares function, and it is achieved by minimizing the standard linear system of the source image matrix expression:

$$L_s = \sum_x \left\|U_x \left(\hat{K}\right) - U_x (K)\right\|_2^2 \tag{16}$$

$U_x \left(\widehat{K}\right)$ and $U_x (K)$ represent the vectorized representations of the generated image and the source image in channel *x*, respectively.

Overall loss function
Combining the aforementioned loss functions, the overall loss function for the model is expressed as:

$$L_O = \alpha\, L_{X\to Y} + \alpha\, L_{Y\to X} + \beta\, L_{pixel} + L_{apper} \tag{17}$$

α is the weight coefficient for the generative adversarial loss. β is the weight coefficient for the reconstruction loss. L_{apper} represents the appearance consistency loss, which includes both texture consistency loss L_v and the structural consistency loss L_s.

By optimizing these loss functions, the model can achieve a comprehensive enhancement in image quality, pose consistency, and visual effects when generating human images related to folk sports culture. This improves the accurate reproduction of folk sports actions and cultural representation, thereby effectively supporting the intelligent development and preservation of folk sports culture.

Experimental design and performance evaluation
Datasets collection
To evaluate the performance of the constructed human image generation model, experiments are conducted on two representative public datasets: DeepFashion and Market-1501. The DeepFashion dataset provides a diverse range of fashion images, which can assist the model in training under complex circumstances of clothing and accessories. The Market-1501 dataset, on the other hand, offers the model a large number of pedestrian images and rich data on pose variations. This enables the model to handle different pose transformations more effectively when dealing with folk sports movements.

The DeepFashion dataset is a large-scale fashion image dataset containing over 800,000 annotated images. These images cover various fashion categories and scenes, including clothing, accessories, and their effects in practical wear. The dataset provides detailed image annotation information, including clothing categories, attribute labels, and key point locations. The pose labels and detailed annotations in DeepFashion help the model learn fine-grained pose transformations, improving the accuracy and realism of the generated images. In order to adapt to the human body image generation model, in the experiment, images from the In-Shop Clothes Retrieval part of the DeepFashion dataset are selected. Among them, there are 36,000 images for training and 12,000 images for testing. All the images have been standardized, and their sizes have been adjusted to 256×256 to facilitate input into the network for training. In the training data, all images have been normalized, and the pixel values have been scaled to the range of [0, 1] to ensure numerical stability during the training process.

In terms of data preprocessing, data augmentation techniques are adopted, including random rotation, flipping, translation, cropping, and brightness changes. These operations aim to simulate various environmental changes that may be encountered in reality, such as perspectives, lighting, and clothing combinations, to improve the generalization ability of the model and enhance its adaptability to different pose changes. In addition, in deep learning models, the image quality in the training dataset directly affects the fineness of the generated results. Therefore, low-quality or noisy images are removed during the data cleaning process.

The Market-1501 dataset focuses on person re-identification and includes images of 1,501 different identities from various camera viewpoints. This dataset includes a total of 32,668 pedestrian images and provides pedestrian identity labels and annotations for image regions. The Market-1501 dataset excels in the person image generation tasks involving pose transformation and pedestrian re-identification. It provides a rich variety of pose samples to support effective training of the model across different environments and angles. The Market-1501 dataset is randomly segmented into training and test sets in a 4:1 ratio. The size of each image is also uniformly adjusted to 256×256 to meet the input requirements of the model. Similar to the DeepFashion dataset, the images from Market-1501 are also normalized and undergo data augmentation processing. It is ensured that each image is subjected to operations such as random cropping, rotation, and flipping, to enhance the model's learning ability for different postures and scenarios.

In addition, a dataset of traditional movements in folk sports culture is collected to further verify the effect of the model in reproducing folk sports movements. This dataset includes classic movements of various folk sports events, such as dances, martial arts, and wrestling. Each movement is annotated in detail, including postures, key points of the movements, and cultural background information. These data are used as part of the training set to verify the performance and effect of the model when generating images of traditional sports movements. This dataset can help the model better understand and reproduce the specific details of folk sports movements, providing a more suitable application scenario for the digital protection of folk culture.

Experimental environment and parameters setting
Table 1 presents the experimental environment and parameter settings.

In order to ensure that the model does not overfit, several measures are taken during the training process to improve the model's generalization ability. First, through data augmentation techniques, the input data are processed by rotation, scaling, and flipping, which expands the diversity of the training set. Then, a multi-layer discriminator design is adopted to enhance the model's ability to judge the details of the generated images, thus preventing the generator from relying too much on a single feature. Finally, a relatively small learning rate is used, and an early stopping strategy is applied during the training process to prevent the overfitting problem caused by

Hardware/Parameter name	Parameter/Value
Operating system	Windows10
Processor	Intel(R)Core(TM)i9-9900k
Graphics processing unit	NVIDIA GeForce RTX 2080Ti
Deep learning framework	PyTorch
Language	python
Training epochs	120k
Learning Rate	$1\times e^{-5}$
Number of PATB modules	9
Optimizer	Rectified Adam
Batch size for deepfashion dataset	1
Batch size for market-1501 dataset	4

Table 1. Experimental environment and parameter settings.

excessive training. Through these means, it is ensured that CycleGAN maintains a strong generalization ability while generating images, preventing the model from overfitting to specific training samples.

To comprehensively evaluate the performance of the constructed human image generation model, four commonly used image quality assessment metrics are employed: Inception Score (IS), Structural Similarity Index (SSIM), Diversity Score (DS), and Fréchet Inception Distance (FID). These metrics offer a comprehensive assessment of the quality of the generated images from various perspectives.

Performance evaluation

Ablation study

In the constructed model, appearance consistency loss is introduced to constrain the generated images. To confirm the effectiveness of this loss function, an ablation study is designed. Under the same conditions, three models are tested on two datasets. Model 1 incorporates appearance consistency loss; Model 2 employs perceptual loss; and Model 3 does not use either appearance consistency loss or perceptual loss. Figure 2 presents the results.

Figure 2 suggests that the model incorporating appearance consistency loss (Model 1) exhibits significant advantages across multiple evaluation metrics. On the Market-1501 dataset, Model 1 outperforms the slightly better-performing Model 2 in IS, SSIM, and DS by 5.78%, 1.76%, and 2.78%, respectively. Similarly, on the DeepFashion dataset, Model 1 shows improvements in IS, SSIM, and DS by 1.49%, 12.67%, and 2.2% over Model 2. These results indicate that incorporating appearance consistency loss significantly advances the visual quality and diversity of the generated images, underscoring its crucial role in improving image quality and maintaining detail consistency.

(2) Comparison of Model-Generated Image Quality.

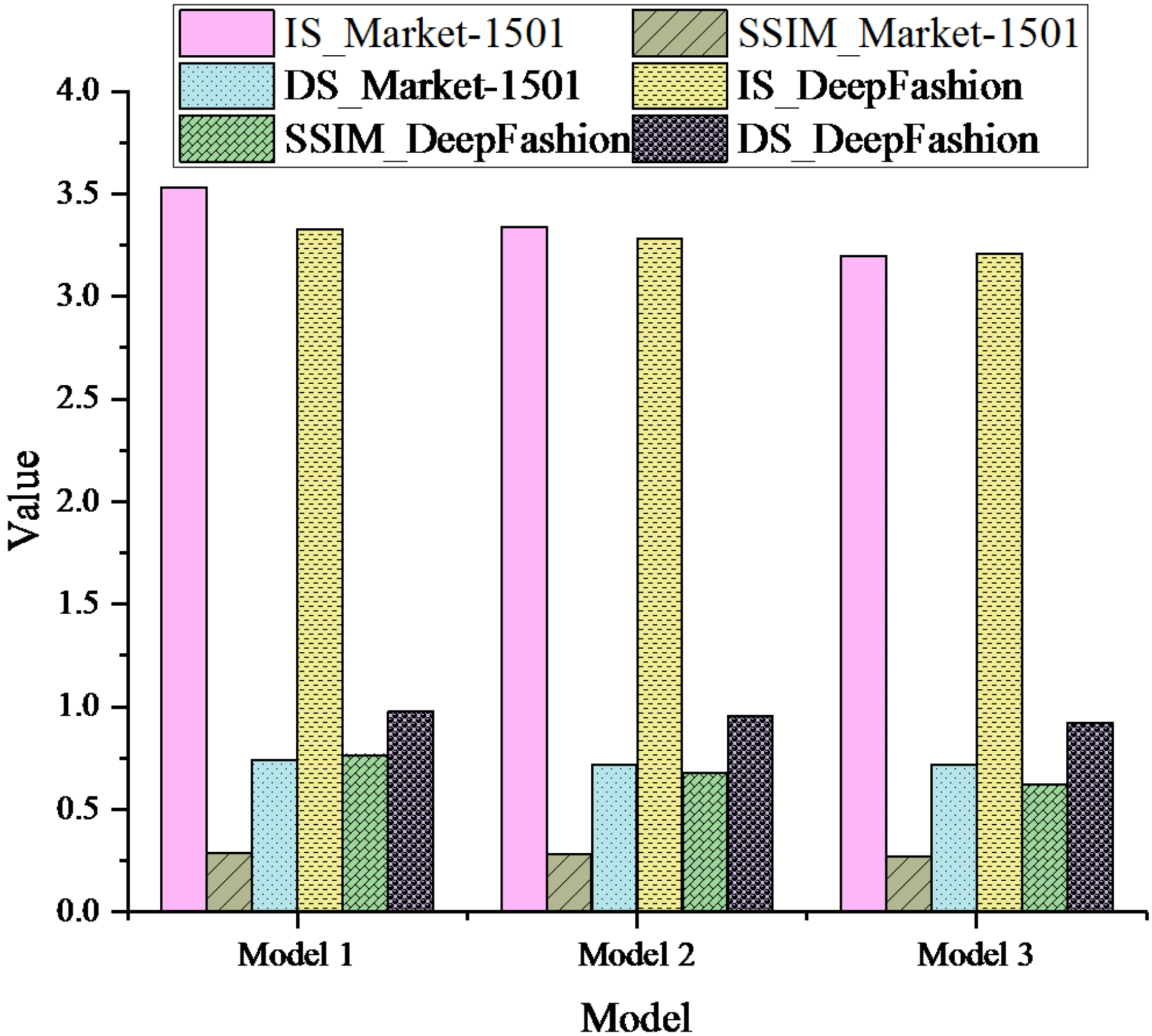

Fig. 2. Ablation study of appearance consistency loss.

Under the same conditions, the proposed model is compared with Video-to-Video Synthesis with Unsupervised Learning (VU-Net)[40], Pose-Associated Texture Network (PATN)[41], Pose Guided Generative Network (PG^2)[42], and Feature Space Perturbation GAN (FSP-GCN)[43]. Figure 3 illustrates the results on the DeepFashion dataset.

Figure 3 reveals that on the DeepFashion dataset, the established model achieves IS of 3.329, SSIM of 0.765, DS of 0.976, and FID of 15.649. Except for SSIM, which is slightly lower than PATN, all other metrics are higher than those of the other models. This result confirms the comprehensive advantages of the proposed model in image generation tasks, particularly regarding image quality, structural consistency, and diversity.

Figure 4 presents the results on the Market-1501 dataset.

Figure 4 shows that on the Market-1501 dataset, the established model achieves an IS of 3.533, SSIM of 0.289, and DS of 0.74, which are significantly higher than those of the other models. Although the PATN model has the highest SSIM, the proposed model performs better in other metrics, indicating that it not only maintains the structural details of the images but also presents superior visual effects. The results from both Figs. 3 and 4 reveal that the established model outperforms existing models in all aspects of image generation, enhancing both the quality and visual experience of the generated images.

To further validate the realism of the generated images and their similarity to real images, a questionnaire is designed containing multiple samples of image pairs, each consisting of one real image and one generated image. Volunteers are asked to distinguish between the generated and real images. Figure 5 presents the results of this experiment.

Figure 5 illustrates that the images generated by the proposed model are identified as real images with a probability of 58.25%, which is significantly higher than that of other models, such as VU-Net (14.9%), PG2 (5.5%), and PATN (31.78%). Additionally, the probability of real images being misclassified as generated images is 20.12%, which is also lower than that of most comparative models. This indicates that the images generated by the proposed model exhibit superior authenticity and visual quality compared to other models, effectively capturing the details and realism of the source image actions.

Finally, in order to verify the applicability of the model to folk sports culture, the model is also tested on the collected dataset of folk sports culture movements. Figure 6 shows the results.

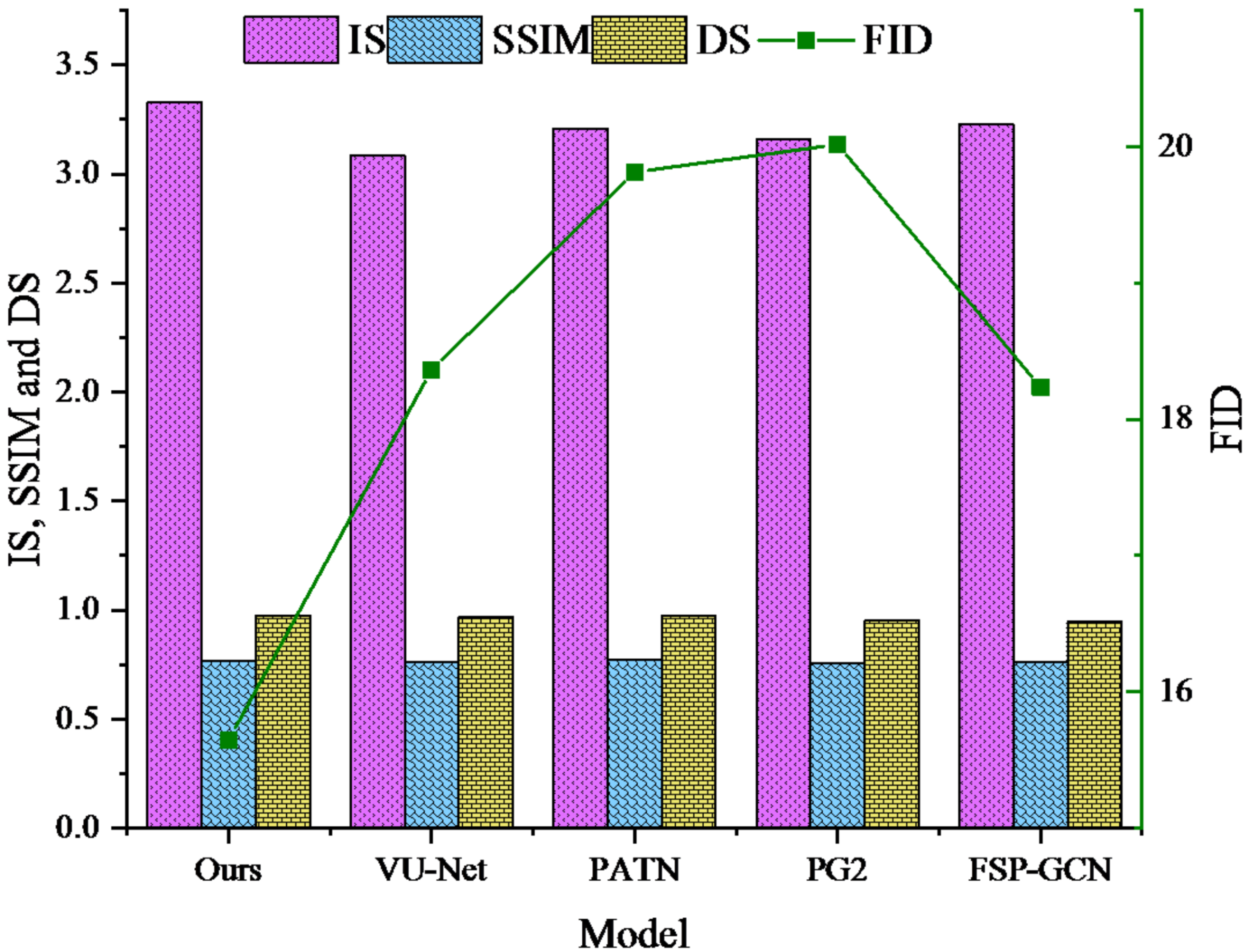

Fig. 3. Model comparison results on the deepfashion dataset.

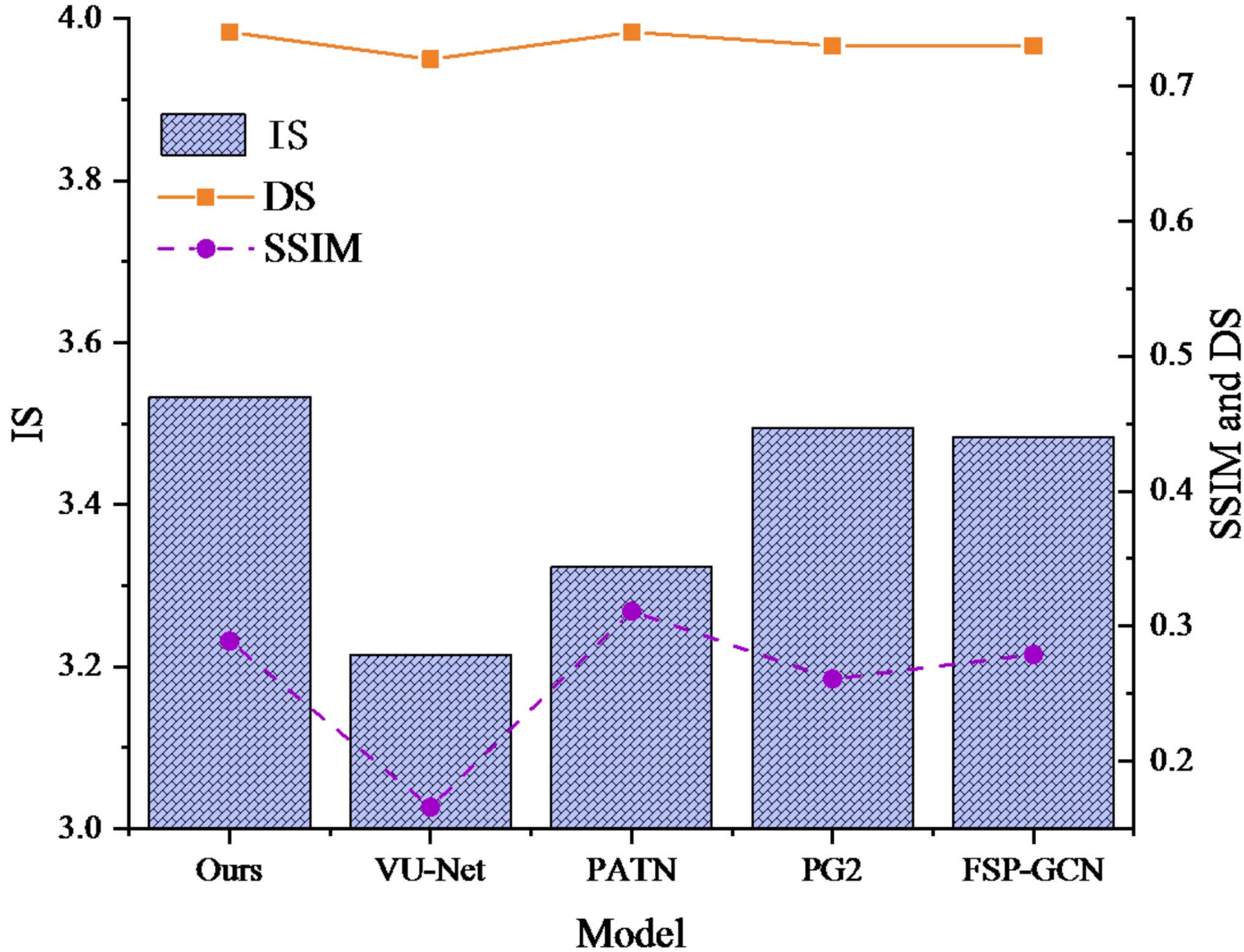

Fig. 4. Model comparison results on the market-1501 dataset.

Figure 6 shows that the model proposed achieves the highest value of 0.68 on DS, outperforming other models, which indicates that it has a greater advantage in the diversity of generated images. In addition, in terms of IS, the value of the model proposed is 3.105, slightly lower than that of PG2, but better than VU-Net and FSP-GCN, suggesting that the generated images have relatively high clarity. Regarding SSIM, PATN performs the best, and the model established comes second, indicating that it has a good effect in maintaining the appearance consistency with the source images. Overall, the model performs excellently in multiple indicators, especially having an advantage in the balance between image diversity and quality. This verifies its effectiveness in the generation of folk sports culture movements.

Figure 7 shows example generation results of the model on the DeepFashion and Market-1501 datasets. (a), (b), and (c) show DeepFashion data, while (d), (e), and (f) show Market-1501 data. From the figure, it can be seen that under different input conditions, the model successfully transforms the source image into the target pose image while maintaining high image quality and detail consistency. In (a), (b), and (c), the generated images from the DeepFashion dataset demonstrate smooth transitions from the source image to the target pose, while showcasing realistic changes in both clothing and posture. In (d), (e), and (f), the generated results from the Market-1501 dataset similarly highlight the model's ability to perform pose transformations across different individuals and backgrounds. Although some noise and detail loss are present, the model still manages to preserve the individual characteristics of the source image. These results indicate that CycleGAN performs exceptionally well in handling unlabelled data and complex pose transformations. Particularly, it can generate high visual quality and realistic images while maintaining consistency in both appearance and posture, which is of significant importance for the digital preservation of sports movements in traditional sports culture.

To explore the limitations of the model in complex scenarios, this work selects two typical challenging samples from the folk sports culture motion dataset. The experiment uses the same training parameters as in this work to compare the generation effects of the proposed model and the baseline model PATN. Table 2 presents the results of SSIM, FID, and human evaluation (the percentage of 50 volunteers judging generated images as real). The results in Table 2 show that the proposed model achieves an SSIM of 0.612 in the martial arts high-kick scenario, which is better than PATN's 0.598. However, the human realism assessment is only 42.1%, indicating that while the generated image has high structural similarity, the body blurring caused by fast motion still affects visual

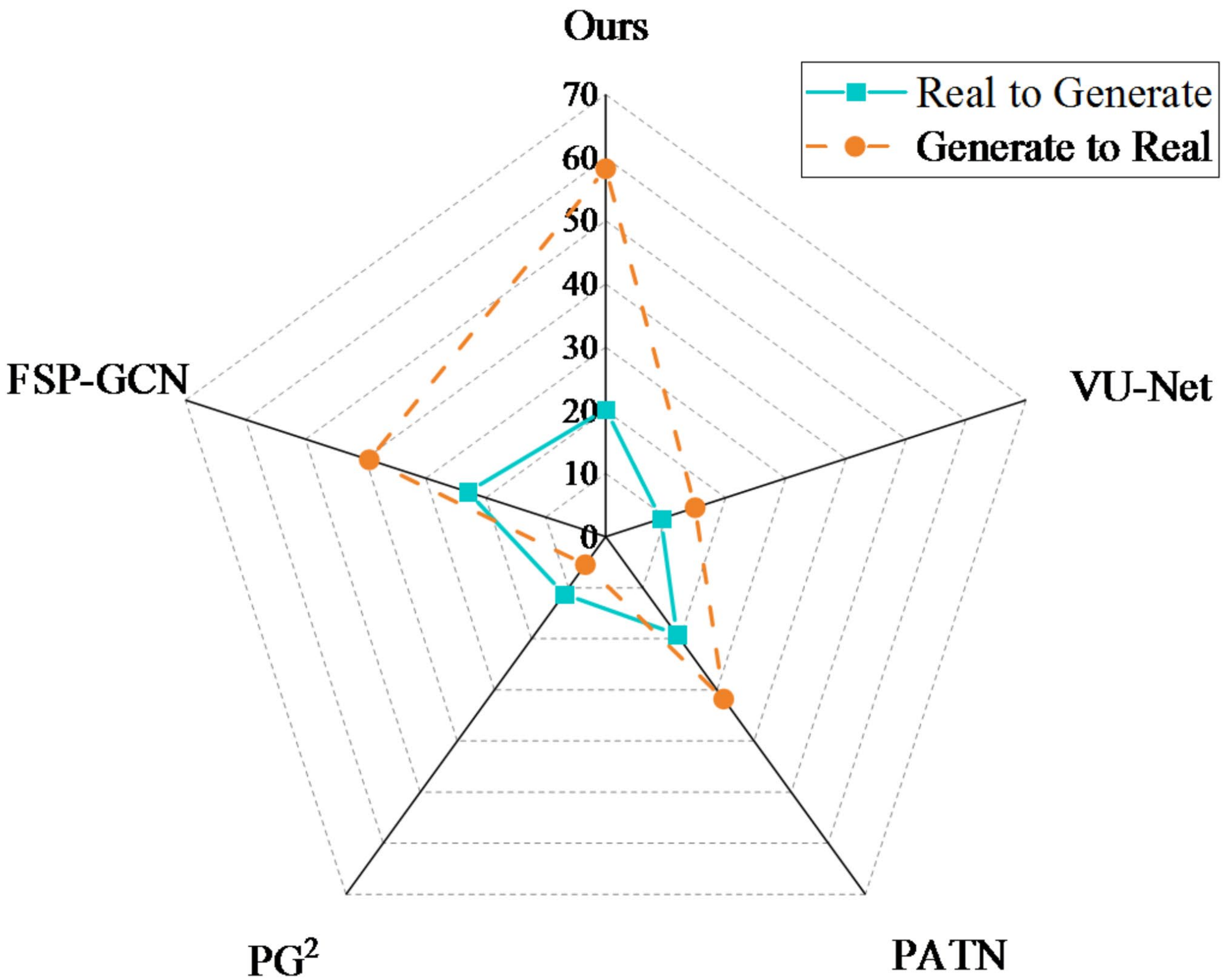

Fig. 5. Results of image authenticity discrimination experiment.

realism. In the dense crowd scenario, the proposed model has an FID of 31.02, which is significantly higher than the baseline dataset (FID = 15.649). This suggests that the complexity of the background increases the deviation between the generated image and the real distribution. The human evaluation shows that the misjudgment rate of volunteers for generated images with dense backgrounds reaches 66.8%, primarily because the model has not sufficiently learned the background-subject separation features. Under strong light conditions, the model-generated images exhibit local overexposure (SSIM = 0.553) and texture detail loss, reflecting the current model's insufficient robustness to lighting variations.

Discussion

This work proposes and trains a generative model based on CycleGAN, aiming to generate high-quality human body images to reproduce the key movement postures in folk sports. Specifically, through the CycleGAN architecture, it is possible to convert source images into generated images with target postures in an unsupervised manner, thus accurately simulating and restoring the movement performances in folk sports. Different from traditional motion capture technologies, this work uses the CycleGAN model to generate images that can accurately reflect specific movements and postures without the need for aligned data. The generated images not only showcase the dynamic characteristics of sports movements but also depict the unique costumes, scenes, and movement details in folk sports culture. This image generation process can provide high-quality visual records for these traditional sports movements and ensure the long-term preservation of these cultural manifestations through the digital storage and dissemination of the generated images. This approach helps to effectively protect and inherit folk sports culture in modern society. Especially when traditional inheritance is not feasible in practice, the generated images provide vivid learning and research materials for future generations.

In summary, by incorporating appearance consistency loss into CycleGAN, the model has achieved significant improvements in visual effects and detail preservation of generated images. Comparisons with related research further validate the effectiveness of the optimization. For instance, Maziarka et al. (2020) established a model combining CycleGAN and convolutional neural networks, focusing on retinal disease detection and localization,

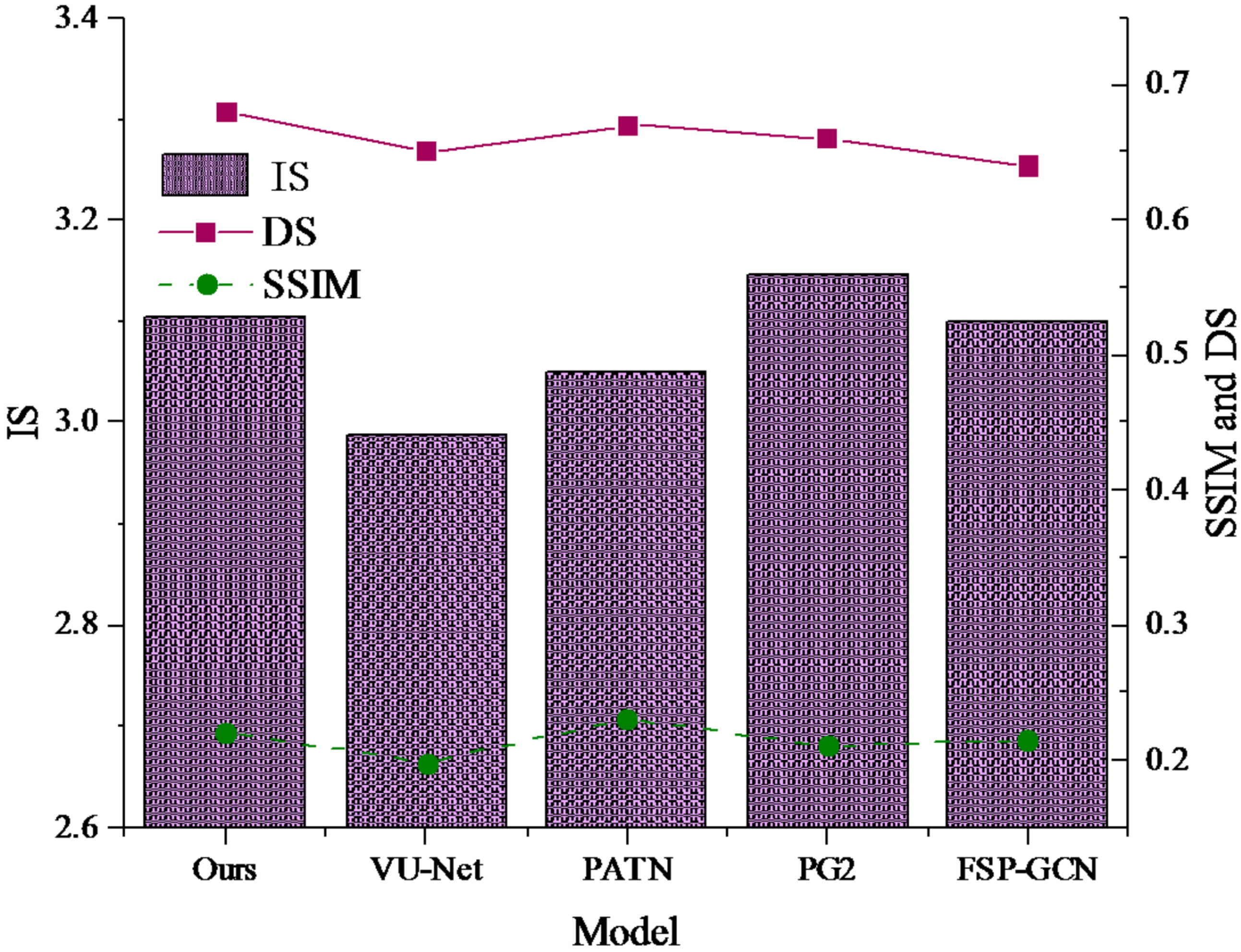

Fig. 6. Comparison results of the models on the dataset of folk sports culture movements.

which demonstrated advantages under limited training data conditions[44]. Zhang et al. (2021) introduced the Mol-CycleGAN model for generating structurally similar compounds optimized for specific attributes. This model excelled in the logP optimization task for drug molecules, surpassing previous results[45]. Zhang et al. (2024) developed a high-quality image-to-image translation model that improved the optimization scheme for image-to-image translation and addressed style bias issues in existing methods, significantly enhancing the visual quality of generated images[46]. These studies collectively indicate that the introduction of optimization strategies can significantly promote the performance of generative models in various applications, providing strong support and reference for the methods used in this research.

This work provides an intelligent and digital solution for the protection and inheritance of folk sports culture through AI technology, especially the human body image generation method based on CycleGAN. As an important part of intangible cultural heritage, folk sports culture contains rich regional characteristics and national spirit. However, due to the acceleration of the modernization process, many traditional sports events are at risk of extinction. This model can accurately capture and reproduce the key movement postures in folk sports, and achieve high-quality digital storage of traditional sports skills. This can provide data support for subsequent research and reproduction. Next, by using the unsupervised learning method, this model can perform posture transformation and image generation in the absence of paired training data. This improves its adaptability and makes it suitable for different types of folk sports events. The experimental results show that the images generated by the model established not only maintain high fidelity but also have strong diversity, which provides a technical guarantee for the visual reproduction and dissemination of traditional sports culture. In addition, by combining with other AI technologies, it is possible to achieve virtual display, digital teaching, and interactive experience of folk sports culture, enabling it to gain a wider audience in modern society and enhancing cultural identity. Through this work, traditional folk sports events can be reproduced in the digital space, and can be applied to fields such as cultural education and competitive simulation in combination with augmented reality or virtual reality technologies. This can break through the limitations of time and space and promote the modern dissemination and sustainable development of folk sports culture. Therefore, the model here not only improves the intelligent generation ability to folk sports movements but also provides a new research perspective and practical path for the digital protection and promotion of intangible cultural heritage.

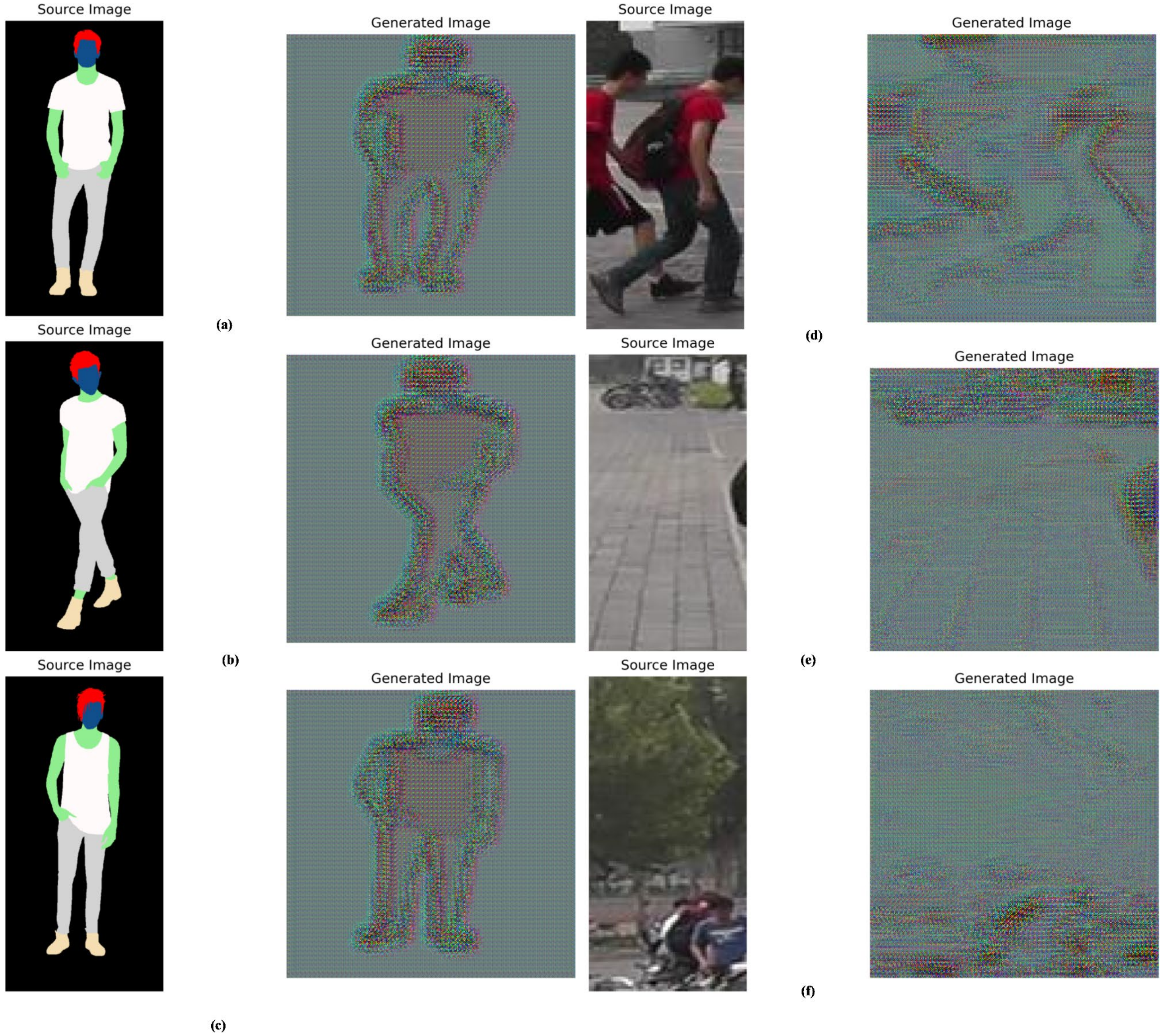

Fig. 7. The generation results of the model on the deepfashion and Market-1501 Datasets (**a**), (**b**), and (**c**) show DeepFashion data; (**d**), (**e**), and (**f**) show Market-1501 data).

Scene types	Model	SSIM	FID	Human authenticity (%)
Martial arts high-kick	The proposed model	0.612	24.73	42.1
	PATN	0.598	25.89	38.5
Dragon dance group collaboration	The proposed model	0.587	26.15	37.8
	PATN	0.562	28.44	34.2
Dynamic lighting changes	The proposed model	0.553	28.91	35.6
	PATN	0.541	29.73	32.4
Dense crowd background	The proposed model	0.521	31.02	33.2
	PATN	0.498	33.15	29.7

Table 2. Model performance comparison in typical failure scenarios.

Conclusion
Research contribution
This work analyzes the CycleGAN model and introduces appearance consistency loss to develop an optimized human image generation model, aiming to enhance the quality and authenticity of the reproduction of folk sports

movements. The performance of the proposed human image generation model is verified through experiments, leading to the following conclusions:

(1) Ablation experiment: In the ablation experiment, the model with appearance consistency loss achieves improvements of at least 1.49%, 1.76%, and 2.2% in IS, SSIM, and DS, respectively, compared to the slightly better-performing Model 2. This result strongly supports the effectiveness of introducing appearance consistency loss, indicating that this loss function significantly improves the visual effects and realism of the generated images.
(2) Model performance comparison: Compared to other models, the proposed model demonstrates superior performance across two different datasets. This indicates that the quality of the generated images in terms of structure, perceptual quality, and diversity is significantly better.
(3) Image authenticity assessment: In the image authenticity assessment experiment, the generated images from the proposed model are judged to be real with a probability of 58.25%, which is significantly higher than that of other models. This indicates that the images generated by the proposed model are markedly superior in authenticity and visual effect, effectively reproducing the details and realism of the source images.
(4) The results on the dataset of folk sports culture movements show that the model proposed performs excellently in multiple indicators. It particularly has an advantage in the balance between image diversity and quality, which verifies its effectiveness in the generation of folk sports culture movements.

Future works and research limitations

Although the proposed CycleGAN-based human image generation model has achieved significant improvements in image quality and realism, there are still some limitations. The current model's adaptability to lighting variations and background interference in different environments needs further enhancement. Additionally, the research has predominantly focused on specific datasets. Future work could extend to more diverse traditional sports datasets to validate the model's broader applicability and robustness. Future efforts may include further improvements to the model structure, incorporation of additional prior knowledge, and exploration of cross-domain applications, thereby advancing the digital development and preservation of traditional sports culture.

In addition, although the human body image generation model based on CycleGAN proposed has achieved good performance on the DeepFashion and Market-1501 datasets, the generalization ability of the model remains an issue that requires further exploration. The current experiments have verified the effectiveness of the model on these two datasets, but in practical applications, the inheritance of folk sports culture involves more complex and diverse environments and data characteristics. Therefore, future research should focus on expanding the scope of application of the model and addressing the problem of its generalization ability in different datasets and complex environments. First, further verifying the performance of the model on other datasets with different characteristics can help to evaluate its generalization ability. For example, in addition to clothing and pedestrian images, datasets containing different cultural backgrounds, sports movements, and scene settings can also be considered. By incorporating datasets such as sports event video datasets or datasets containing folk sports events from different countries and regions, the performance of the model in these environments can be explored. This can test whether the model can generate high-quality images in diverse cultural and environmental backgrounds and accurately reproduce folk sports movements. Next, the challenges of environmental factors, such as different lighting conditions, backgrounds, movement occlusions, and complex scenes, may affect the quality and accuracy of the generated images. To address these issues, future research can attempt to introduce more advanced data augmentation techniques, such as lighting change simulation and complex background modeling. By modeling and processing environmental variables, the adaptability of the model in different scenarios can be further improved, thereby enhancing its stability in practical applications. In addition, considering the high diversity of folk sports events in different countries and regions, future research can also develop customized generation models for the folk sports culture of specific regions. For example, region-specific cultural labels and movement characteristics can be introduced. This can enable the model to not only be accurate in terms of movement postures but also reflect the cultural background, costumes, and movement details, thus better achieving cultural inheritance. Therefore, the current model has already demonstrated good performance, especially in terms of the details and authenticity of traditional sports movements. However, improving the generalization ability of the model, expanding its application scope, and addressing environmental challenges remain important directions for future research.

Data availability

The datasets used and/or analyzed during the current study are available from the corresponding author Shamsulariffin Samsudin on reasonable request via e-mail 20141031@llu.edu.cn.

Received: 19 October 2024; Accepted: 15 April 2025
Published online: 23 April 2025

References

1. Li-Yun, Z., Cheng-Ke, W. & Qiang, Z. [Retracted] the construction of folk sports featured towns based on intelligent building technology based on the internet of things. *Appl. Bionics Biomech.* **2022** (1), 4541533 (2022).
2. Drozdek-Małolepsza, T. & Małolepszy, E. Sport in the activity of the folk sports teams province association in Katowice in the years 1952–1975. *Sport Tourism Cent. Eur. J.* **7** (2), 57–79 (2024).
3. Wang, Y. A probe into the mutual promotion between the inheritance of traditional culture and the development of National sports in private universities. *Adult High. Educ.* **5** (18), 14–19 (2023).
4. Ospankulov, Y., Zhumabayeva, A. & Nurgaliyeva, S. The impact of folk games on primary school Students[J]. *J. Educ. E-Learning Res.* **10** (2), 125–131 (2023).

5. Irfan, M. et al. Development of traditional sports-based through educational tourism model: edu Ortrad as a sports industry model. *Int. J. Hum. Mov. Sports Sci.* **11** (02), 411–417 (2023).
6. Damanik, S., Damanik, S. A. & Ritonga, D. A. Creating digital content for traditional children's games in North Sumatra to preserve culture and enhance sports participation. AL-ISHLAH: Jurnal Pendidikan **16** (2) 2211–2219. (2024).
7. Gao, J. Construction of digital protection path for intangible cultural heritage of ethnic minority traditional sports: take Tibetan traditional sports as an example. *J. Sociol. Ethnol.* **5** (9), 27–32 (2023).
8. He, P., Zheng, G. & Gong, Z. Survival of folk sports–related cultural heritage in China. *Int. J. Hist. Sport.* **37** (12), 1159–1171 (2021).
9. Zuo, Y. et al. How natural environments influence traditional sports and games: A mixed methods study from China. *Int. Rev. Sociol. Sport*. **58** (2), 328–348 (2023).
10. Sinaga, R. M. & Adha, M. M. Preservation of intangible cultural heritage: the role of Documentation in cultural conservation in the Semaka district, Tanggamus Regency. *Int. J. Adv. Technol. Social Sci. (IJATSS)*. **2** (3), 375–388 (2024).
11. Paananen, V., Oppenlaender, J. & Visuri, A. Using text-to-image generation for architectural design ideation. *Int. J. Architectural Comput.*, 14780771231222783. (2023).
12. Spennemann, D. H. R. ChatGPT and the generation of digitally born knowledge: how does a generative AI Language model interpret cultural heritage values?. *Knowledge* **3** (3), 480–512 (2023).
13. Steinfeld, K. Clever little tricks: a socio-technical history of text-to-image generative models. *Int. J. Architectural Comput.* **21** (2), 211–241 (2023).
14. Sun, C., Zhou, Y. & Han, Y. Automatic generation of architecture facade for historical urban renovation using generative adversarial network. *Build. Environ.* **212**, 108781 (2022).
15. Wang, X. Artificial intelligence in the protection and inheritance of cultural landscape heritage in traditional village. *Sci. Program.* **2022** (1), 9117981 (2022).
16. Kasmahidayat, Y. & Hasanuddin, H. Collaboration strategy in the development and inheritance of Archipelago's arts. *J. Indigenous Cult. Tourism Lang.* **1** (1), 1–20 (2022).
17. Ding, M. et al. Cogview: mastering text-to-image generation via transformers. *Adv. Neural. Inf. Process. Syst.* **34**, 19822–19835 (2021).
18. Wang, J. et al. Deep transfer learning-based multi-modal digital twins for enhancement and diagnostic analysis of brain mri image. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **20** (4), 2407–2419 (2022).
19. Saharia, C. et al. Photorealistic text-to-image diffusion models with deep language understanding. *Adv. Neural. Inf. Process. Syst.* **35**, 36479–36494 (2022).
20. Bei, Y. Research and implementation of innovative design of paper cuttings pattern based on artificial intelligence technology. *J. Artif. Intell. Pract.* **6**, 89–93 (2023).
21. Garozzo, R. et al. Knowledge-based generative adversarial networks for scene understanding in cultural heritage. *J. Archaeol. Science: Rep.* **35**, 102736 (2021).
22. Kumar, P. & Gupta, V. Restoration of damaged artworks based on a generative adversarial network. *Multimedia Tools Appl.* **82** (26), 40967–40985 (2023).
23. Huang, Z. et al. CaGAN: A cycle-consistent generative adversarial network with attention for low-dose CT imaging. *IEEE Trans. Comput. Imaging*. **6**, 1203–1218 (2020).
24. Xiang, Y. & Bao, C. A parallel-data-free speech enhancement method using multi-objective learning cycle-consistent generative adversarial network. IEEE/ACM Transactions on Audio, Speech, and Language Processing, **28** 1826–1838. (2020).
25. Zhou, H. et al. Unsupervised cycle-consistent generative adversarial networks for pan sharpening. *IEEE Trans. Geosci. Remote Sens.* **60**, 1–14 (2022).
26. Runz, M. et al. Normalization of HE-stained histological images using cycle consistent generative adversarial networks. *Diagn. Pathol.* **16**, 1–10 (2021).
27. Kearney, V. et al. *Attention-aware Discrimination for MR-to-CT Image Translation Using cycle-consistent Generative Adversarial networks*2e190027 (Artificial Intelligence, 2020). 2.
28. Banach, A. et al. Visually navigated bronchoscopy using three cycle-consistent generative adversarial network for depth estimation. *Med. Image. Anal.* **73**, 102164 (2021).
29. Li, P. et al. A deep multimodal adversarial cycle-consistent network for smart enterprise system. *IEEE Trans. Industr. Inf.* **19** (1), 693–702 (2022).
30. Kalantar, R. et al. CT-based pelvic T1-weighted MR image synthesis using UNet, UNet++and cycle-consistent generative adversarial network (Cycle-GAN). *Front. Oncol.* **11**, 665807 (2021).
31. Karabulut, D. et al. Cycle-consistent generative adversarial neural networks based low quality fingerprint enhancement. *Multimedia Tools Appl.* **79**, 18569–18589 (2020).
32. Wang, Y. Q. et al. Seismic impedance inversion based on cycle-consistent generative adversarial network. *Pet. Sci.* **19** (1), 147–161 (2022).
33. Wu, S., Dong, C. & Qiao, Y. Blind image restoration based on cycle-consistent network. *IEEE Trans. Multimedia*. **25**, 1111–1124 (2022).
34. Kwak, D. & Lee, S. A novel method for estimating monocular depth using cycle Gan and segmentation. *Sensors* **20** (9), 2567 (2020).
35. Bigioi, D. et al. Pose-aware speech driven facial landmark animation pipeline for automated dubbing. *IEEE Access.* **10**, 133357–133369 (2022).
36. Jung, E., Luna, M. & Park, S. H. Conditional GAN with 3D discriminator for MRI generation of Alzheimer's disease progression. *Pattern Recogn.* **133**, 109061 (2023).
37. Schwarz, K., Liao, Y. & Geiger, A. On the frequency bias of generative models. *Adv. Neural. Inf. Process. Syst.* **34**, 18126–18136 (2021).
38. Li, H. et al. Cross adversarial consistency self-prediction learning for unsupervised domain adaptation person re-identification. *Inf. Sci.* **559**, 46–60 (2021).
39. Yan, C. et al. Beyond triplet loss: person re-identification with fine-grained difference-aware pairwise loss. *IEEE Trans. Multimedia*. **24**, 1665–1677 (2021).
40. Zhuo, L. et al. *Fast-Vid2Vid++: Spatial-Temporal Distillation for Real-Time Video-to-Video Synthesis* (IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024).
41. Sharma, S. & D'Amico, S. Neural network-based pose Estimation for noncooperative spacecraft rendezvous. *IEEE Trans. Aerosp. Electron. Syst.* **56** (6), 4638–4658 (2020).
42. Khatun, A. et al. Pose-driven attention-guided image generation for person re-identification. *Pattern Recogn.* **137**, 109246 (2023).
43. Xue, H. et al. Face image de-identification by feature space adversarial perturbation. *Concurrency Computation: Pract. Experience*. **35** (5), e7554 (2023).
44. Maziarka, Ł. et al. Mol-CycleGAN: a generative model for molecular optimization. *J. Cheminform.* **12** (1), 2 (2020).
45. Zhang, Z. et al. Joint optimization of cyclegan and CNN classifier for detection and localization of retinal pathologies on color fundus photographs. *IEEE J. Biomedical Health Inf.* **26** (1), 115–126 (2021).
46. Zhang, Y. et al. HQ-I2IT: redesign the optimization scheme to improve image quality in CycleGAN-based image translation systems. *IET Image Proc.* **18** (2), 507–522 (2024).

Author contributions
Zhihui Li: Conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing—original draft preparation, Shamsulariffin Bin Samsudin: writing—review and editing, visualization, supervision, project administration, funding acquisition Noor Hamzani Farizan: methodology, software, validation, formal analysisZulkhairi Azizi Zainal Abidin: visualization, supervisionLei Zhang: formal analysis, investigation, resources.

Funding
This work was supported by Research on the Identification Path of Folklore Sports Culture in Rural Tourism along the Yellow River, oxShanxi Province Higher Education Teaching Reform and Innovation Project, (Grant No.J20231381).

Declarations

Competing interests
The authors declare no competing interests.

Ethics statement
The studies involving human participants were reviewed and approved by Department of Sports Studies, Faculty of Educational Studies, Universiti Putra Malaysia Ethics Committee (Approval Number: 2022.44556784). The participants provided their written informed consent to participate in this study. All methods were performed in accordance with relevant guidelines and regulations.

Additional information
Correspondence and requests for materials should be addressed to S.B.S.