Anime Game Character Modeling Using BiLSTM with Multi-Attention Mechanisms for Improved Image Generation

Yan Zhou
Media Department of Hebei Institute of International Business and Economics, Qinhuangdao 066311, Hebei, China
E-mail: eizhuanyong@126.com

Keywords: anime game characters, long short-term memory network, modeling design, attention mechanism

Received: July 16, 2024

The objective of this study is to leverage the Long Short-Term Memory (LSTM) algorithm to develop a model for anime game character modeling, with the aim of enhancing both the effectiveness and efficiency of anime character design to meet the industry's demand for more precise and diverse designs. The study introduces the LSTM algorithm and its application in the field of image generation. It proposes the integration of multi-attention mechanisms with Bidirectional Long Short-Term Memory (BiLSTM) to augment the model's capacity to capture intricate image details and diversity. Subsequently, the study constructs an anime game character modeling design model based on the fusion of multi-attention mechanisms and the BiLSTM algorithm, which is then subjected to experimental evaluations. In terms of experimental computing settings, the Anime Face Dataset is set as a data set, and a high-performance computing cluster with 8 GPUs is used for training. Adam is used to optimize the algorithm, and the learning rate is set to 0.001. The experimental results demonstrate that the proposed model achieves a prediction accuracy of 96.91% and an F1 score of 91.79% in the context of anime character design, thereby improving accuracy by at least 3.08% compared to the Temporal Attention Mechanism LSTM (TAM-LSTM) model. Moreover, the Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) of the proposed model are reduced by 6.52% to 10.7%, respectively. Consequently, the model algorithm presented in this study can provide significant value in enhancing the effectiveness and efficiency of anime character design, contributing to the advancement of the anime industry.

Povzetek: Raziskava uvaja napredni model za oblikovanje likov v anime igrah s hibridom BiLSTM in večkratnim pozornostnim mehanizmom.

1 Introduction

In the contemporary anime gaming industry, character modeling design plays a pivotal role. As technology advances and user expectations grow, the industry continuously seeks innovations to cater to the diverse needs of various player demographics. Characters not only serve as the carriers of the storyline but also act as critical conduits for interaction between players and the game world [1, 2]. With the evolution of virtual reality technology, players increasingly demand realism and depth in the portrayal of game characters, desiring them to be more lifelike and authentic [3]. Furthermore, the emergence of social and multiplayer online games has heightened players' focus on the interactions and connections among characters. Consequently, the development of an efficient and intelligent character modeling design methodology has become an urgent requirement for game developers.

The significance of character modeling design in games is self-evident. A well-designed character image can not only attract players' attention and increase the game's visibility and user engagement but also enrich the game's narrative, enhance its playability, and boost its entertainment value [4]. However, traditional character modeling design processes are often based on manual creation, which is not only time-consuming and labor-intensive but also presents challenges in ensuring personalization and diversity in design. The Long Short-Term Memory (LSTM) algorithm, as a potent deep learning model, possesses strong image generation capabilities [5-7]. It can learn the appearance characteristics and stylistic features of characters from a large amount of training data, thereby generating personalized and appealing character images. Its automated and intelligent design approach significantly improves design efficiency and quality while also meeting players' demands for personalized game characters. Additionally, the attention mechanism can assist the model in better capturing key information from input sequences, thereby enhancing the effectiveness and performance of the model for character modeling design tasks [8, 9]. Incorporating the attention mechanism into the LSTM enables the model to dynamically focus on important features and regions within anime character image data, thus better generating personalized and diversified character models.

Therefore, the innovation of this study lies in the adoption of a hybrid model based on the LSTM algorithm and attention mechanism to improve anime game character modeling design. This model can dynamically select and weight different features in input images at each time step to better capture key information in character images. Consequently, this study not only promotes the advancement of game development technology but also provides players with a richer and more enjoyable gaming experience, offering game developers more design inspiration and possibilities.

2 Literature review

In recent years, character design in anime has garnered increasing attention as a vital component of game development. Li et al. [10] proposed a controllable automated generation method for designing non-player characters in a three-dimensional (3D) animation style. Their findings indicated that this method could assist developers in rapidly and efficiently generating a variety of anime character styles, thereby enhancing the game's visual appeal and expressiveness. Kopylova [11] argued that the interaction and communication between different media forms of anime works (such as animation, manga, and games), along with their shared visual styles, constitute a creative and culturally significant ecosystem. Understanding the diversity and influence of anime culture was crucial. Mahler and Mayer [12], through analysis and experimental research on anime works, pointed out that anime can serve as an effective tool for science communication, attracting the attention of young people and promoting their understanding and learning of scientific knowledge. Chen [13] proposed an animation VR scene splicing modeling method based on a genetic algorithm, which improved the complexity and visual effect of animation character design by optimizing character design and scene fusion, and provided a new idea for animation production in virtual reality. Liu et al. [14] explored the impact of anime and idol culture on mental health from a psychological perspective. The issues they pointed out raised concerns about the psychological well-being of adolescents regarding anime culture, prompting a heightened focus on both the positive and negative impacts of anime culture.

The LSTM algorithm, by learning long-term dependencies in sequential data, effectively captures temporal information and structural features in sequence data. Numerous studies have attempted to utilize LSTM algorithms for image generation and processing tasks, such as image description generation, image style transfer, and image generation. Zhang et al. [15] proposed an image description generation method using Bidirectional Long Short-Term Memory (BiLSTM). Their results showed that this method could improve the accuracy and coherence of generated descriptions by integrating contextual information. Aswiga and Shanthi [16] utilized transfer learning to extract features and train LSTM models using existing medical image datasets. Experimental results demonstrated that this method could effectively generate accurate and clinically meaningful medical descriptions. Son et al. [17] introduced a method based on LSTM and Generative Adversarial Networks (GANs) for predicting cloud movement in satellite images. Experimental results showed that this method could effectively predict cloud trajectories, providing reliable predictions for solar energy generation. Zabin et al. [18] proposed a hybrid deep transfer learning architecture. Experimental results indicated that this method achieved good performance in industrial fault diagnosis. Tan et al. [19] presented an LSTM-GANs-based method for converting electroencephalogram signals into visually artistic and expressive landscapes. Zhou and Li [20] proposed a method based on BiLSTM and CBAM-enhanced GANs to enhance data representation capabilities. Experimental results showed that this method could effectively enhance the quality of image data and improve the accuracy of image data analysis. Mak et al. [21] discussed the application of Variational AutoEncoder (VAE) models and image processing methods in game design.

The summary of various scholars is shown in Table 1.

Table 1: Summary of recent literature research

Scholar	Method	Data set	Performance indicators
Li et al. [10]	Automatic generation of non-player characters in 3D animation style	Self-built 3D animation data set	High generation diversity and control precision
Kopylova [11]	Analysis on the Mixing and Graphic Style of Cartoon Media	Cultural Shared Resources Database	Analysis of the Cultural Influence of Animation Style
Mahler & Mayer [12]	Anime as a medium for science learning incorporates a science learning model	A Survey of Primary and Secondary School Students' Animation Viewing Habits	The effectiveness of science learning has increased by 20 percent
Chen [13]	Modeling of Animated VR Scene Stitching Based on Genetic Algorithm	VR animation scene data set	High accuracy and computational efficiency of scene stitching
Liu et al. [14]	Anime, Idol Culture, and Depression: Structural Analysis and Psychological Outcomes of Deep Learning	Large-scale mental health questionnaire data set	The precision of correlation analysis is high, and the AUC is 0.92.
Zhang et al. [15]	The image description generation method of BiLSTM-s is used,	COCO data set	The BLEU-4 score is 28.5

	combined with context information fusion.		
Aswiga & Shanthi [16]	Multilevel transfer learning with LSTM framework to generate medical descriptions of limited CT and DBT images	Self-built medical image data set	The BLEU-4 score was 25.7
Son et al. [17]	Cloud motion prediction based on LSTM–GAN from satellite images for PV prediction	Self-built satellite image data set	The prediction accuracy is 92.4%.
Zabin et al. [18]	Hybrid Deep Transfer Learning Architecture for Industrial Fault Diagnosis Combining Hilbert Transform and DCNN–LSTM	Industrial fault diagnosis data set	The diagnostic accuracy was 97.1%.
Tan et al. [19]	Feature Design of EEG Signal Visualization Based on LSTM-Gan	Self-built EEG signal data set	The generated feature images are of high quality.
Zhou & Li [20]	BiLSTM and CBAM-based method for ECG data enhancement to generate countermeasure network	Self-built ECG data set	The data enhancement effect is remarkable, and the AUC is 0.93.
Mak et al. [21]	Application of VAE Model and Image Processing Method in Game Design	Self-built game design data set	Improved diversity and quality of generated images

Through the understanding of the above literature, although the relevant research in the field of animation design has been summarized, the attention to animation character modeling is relatively insufficient. Especially when dealing with the unique visual style and complex characteristics of animation, the traditional deep learning method is not enough. In addition, there are also studies using LSTM algorithm to achieve image generation, but these studies have not been fully applied in the design of animation characters. This study aims to combine the attention mechanism and apply the LSTM algorithm to the animation game character design. This combination can better capture the key features and details in the character image, and generate a personalized and realistic character image. Compared with the existing research, the method of this study has advantages in improving the authenticity and diversity of the generated images, which is expected to promote the progress of animation character design technology and enrich the game experience.

3 Methodology

3.1 Analysis of requirements for intelligent design in anime game character modeling

The design quality of anime game character modeling directly influences the visual appeal and user experience of games [22]. However, traditional manual design methods face a series of challenges and limitations. Firstly, manual design is time-consuming and labor-intensive, requiring designers to invest a significant amount of time and effort. Secondly, manual design is susceptible to individual experiences and aesthetic preferences, leading to a lack of diversity and personalization in character images. Additionally, manual design often fails to meet the interactive demands of players [23, 24]. Therefore, achieving intelligent design in anime game character modeling is of significant importance and necessity, as illustrated in Figure 1.

From the perspective of game developers, intelligent design can significantly improve the efficiency and quality of character design. Intelligent design tools can assist developers in quickly generating a variety of styles and types of anime characters, saving design time and reducing manpower costs, and enhancing development efficiency. Moreover, intelligent design tools can provide rich customization options to meet the needs of different game projects, increasing design flexibility and diversity. From the perspective of players, intelligent design can provide more personalized and diversified character images, increasing the enjoyment and playability of the game. Players can customize their favorite character images according to their preferences and needs, enhancing interaction and immersion with the game world.

Therefore, researching and developing intelligent design tools has become one of the important requirements in the field of anime game development. This study aims to achieve intelligent design of anime game character modeling by introducing the LSTM algorithm and combining it with attention mechanisms, and evaluating its effectiveness and performance.

3.2 Optimization of LSTM algorithm

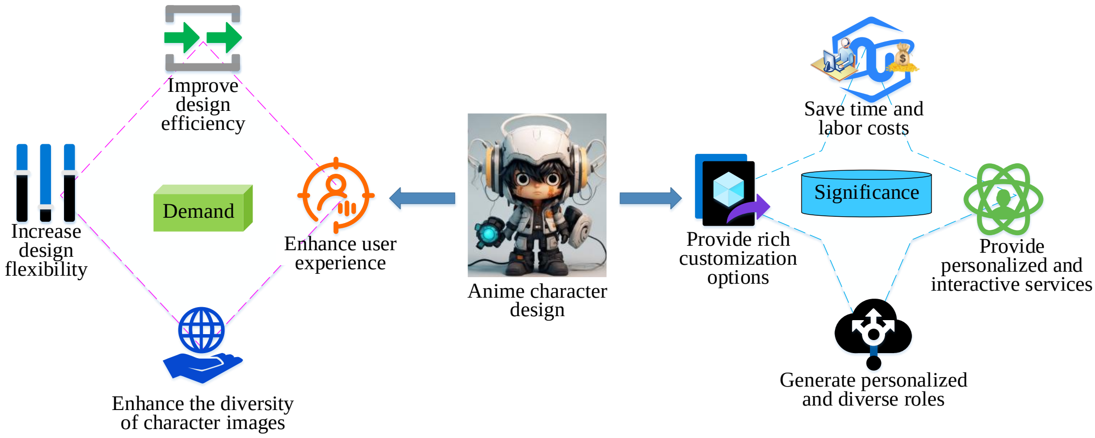

Figure 1: The requirements and significance of intelligent design in anime game character modeling

The LSTM algorithm, as a powerful tool for sequence modeling, has achieved notable success in image generation tasks. However, for the task of anime character design, the current LSTM algorithm still exhibits limitations in terms of the quality and diversity of generated images. Therefore, this study aims to optimize the LSTM algorithm to enhance its performance in anime character design.

LSTM is a special type of recurrent neural network designed specifically for processing sequential data and effectively capturing long-term dependencies in sequence data [25, 26]. The LSTM network consists of multiple repeated memory units, each memory unit containing input gates (*i*), forget gates (*f*), output gates (*o*), and a memory cell (*c*). The cell state and hidden state are used to store internal states. The forget gate determines which information needs to be discarded from the cell state, i.e., "forgetting"; the input gate determines which input information needs to be stored in the cell state. The state gate further determines how to update the LSTM cell state, generating new memory content C_t by adding information controlled by the forget gate. The output gate determines which information needs to be output and selectively outputs portions determined by the new memory content h_t. The computations involved in LSTM are represented by Equations (1) to (6):

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{1}$$

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{2}$$

$$\tilde{C}_t = \tanh\left(W_C \cdot [h_{t-1}, x_t] + b_C\right) \tag{3}$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{4}$$

$$o_t = \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \tag{5}$$

$$h_t = o_t * \tanh(C_t) \tag{6}$$

In equations (14) to (19), W_f denotes the weight matrix of the output gate. x_t represents the current input. h_t denotes the output from the previous step. b represents the bias term. W_i denotes the weight matrix of the input gate. W_C denotes the weight matrix for generating the cell state. W_o denotes the weight matrix of the output gate.

Due to the rich details and diverse styles typically present in anime character images, traditional LSTM algorithms often struggle to fully capture and express these features. However, BiLSTM has the capability to learn from both past and future information simultaneously, thus providing a more comprehensive understanding of the contextual information in input sequence data. It has demonstrated superior abilities in sequence modeling, detail capturing, diversity, creativity, and adaptability in anime character image design [27, 28]. The application of the BiLSTM algorithm in anime character modeling design is illustrated in Figure 2.

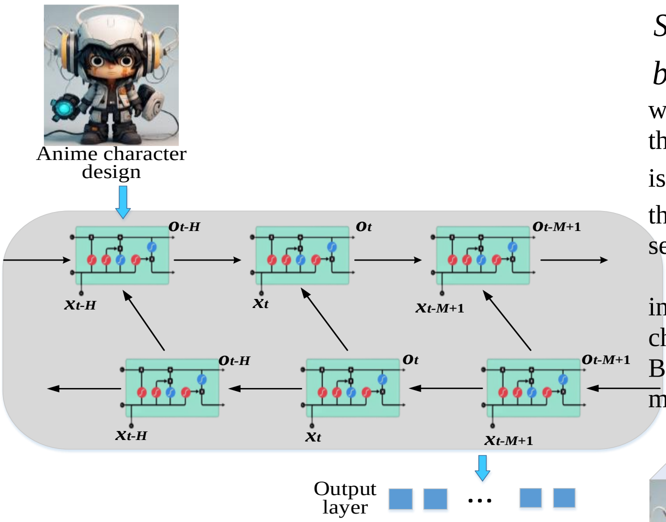

Figure 2: Application of BiLSTM algorithm in anime character modeling design

By combining the multi-attention mechanism with the BiLSTM algorithm, the Temporal Attention Mechanism (TAM) can adaptively learn the importance weights of different time steps in sequential data. This enables the model to more effectively focus on and utilize information from different time points in the sequence data [29]. The Channel Attention Mechanism (CAM) can adaptively weight different channels of input data, thereby extracting more informative features [30]. Therefore, integrating the multi-attention mechanism with BiLSTM can enable the model to better understand and utilize the sequential data in anime character modeling design tasks. This enhances the model's capability in sequence data modeling and generates anime character images that are more rich, vivid, and realistic.

The main calculation of TAM is shown in Equations (7) to (9):

$$V_{(M\times T)} = \left(S_t\right)^T W_t + b \tag{7}$$

$$B_{(M\times T)} = \frac{\exp\left(V_{ij}\right)}{\sum_{t-T}^{t}\left(V_{ij}\right)}, i \in \left[1, M\right] \tag{8}$$

$$\delta = \frac{\sum_{p=1}^{M} B_{mn}}{M} = \left[\delta_{t-T}, \delta_{t-T+1}, \cdots, \delta_{t-1}\right], n \in \left[t-T, t-1\right] \tag{9}$$

In Equations (7) to (9), V denotes the unnormalized time probability weight matrix obtained through neural network operations on the transpose of the input matrix S_t. W_t denotes the weight matrix of the neural network. b denotes the bias vector. B denotes the time probability weight matrix normalized after softmax activation, where the sum of probability weights in each row of this matrix is equal to 1. δ denotes the final time attention vector (i.e., the output of Time Attention), and M denotes the target sequence of anime character modeling to be predicted.

An attention mechanism in the channel domain is introduced to obtain better representations of motion characteristics. A channel attention module is added to the BiLSTM algorithm. The structure of the channel attention module is illustrated in Figure 3.

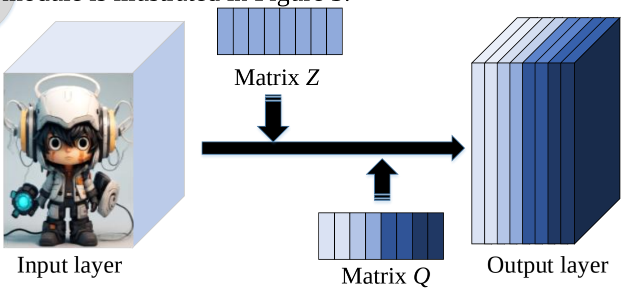

Figure 3: Schematic diagram of the channel attention module

In Figure 3, in the channel attention module, the matrix Z is obtained by first squeezing operation F_1 on f_{out} to obtain matrix Z. Then, after passing through operation F_2, it enters the fully connected layer to obtain the output matrix Q, where W_1 and W_2 are two weight matrices. Finally, the output matrix Q is multiplied with the original input feature map data and added via residual connections to obtain the final output f.

3.3 Analysis of construction of anime game character modeling design model based on fusion of multi-attention mechanism and BiLSTM algorithm

To effectively predict the design requirements of anime game characters, this study introduces the TAM, which assigns different weights to data at different time points. The introduction of the CAM dynamically adjusts the model's feature extraction levels across different channels, enabling the model to capture the temporal correlations in the data and the trends in different spatial channels. Meanwhile, the BiLSTM algorithm is employed to effectively handle long-term dependencies and nonlinear features in time series data. Thus, a model for anime game character modeling design based on the fusion of the multi-attention mechanism and BiLSTM algorithm is constructed, as shown in Figure 4.

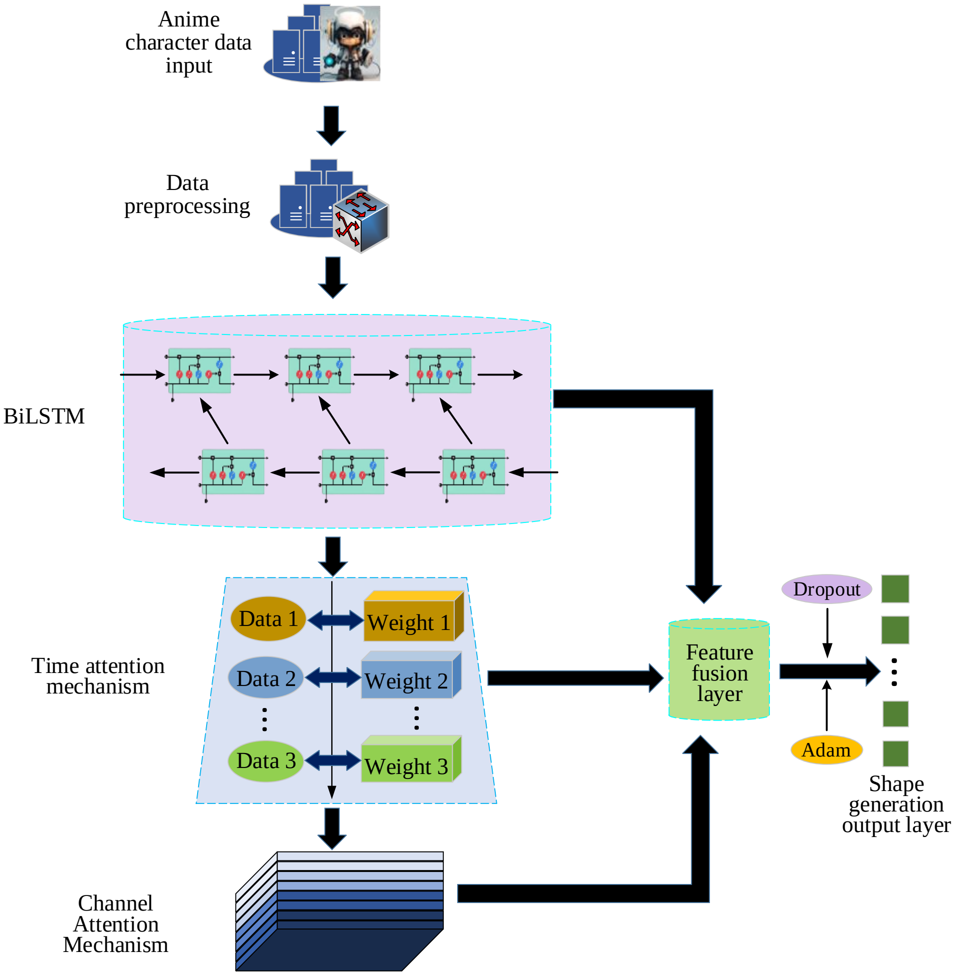

Figure 4: Schematic diagram of anime game character modeling design model based on fusion of multi-attention mechanism and BiLSTM algorithm

In Figure 4, the input layer initially receives feature representations or sequence data related to anime game character modeling. For image generation tasks, the input includes image data containing character features. For sequence data generation tasks, the input can consist of sequential information describing character attributes, such as character clothing, facial features, etc. Subsequently, the BiLSTM layer is utilized to process the input sequence data, learning the temporal characteristics of character images, including sequences of actions, facial expressions, and other dynamic features. In anime game character modeling design, the Temporal Attention Mechanism layer is further utilized to help the model focus more on important time points within the character images. Assuming an n-dimensional time series is represented as Equation (10):

$$X = \left(x^1, x^2, \cdots, x^n\right)^T = \left(x_1, x_2, \cdots, x^T\right) \in R^{n \times T} \quad (10)$$

In Equation (10), T represents the length of the sequence. Given a target sequence $\left(y_1, y_2, \cdots, y_{T-1}\right)$, and other sequences $\left(z_1, z_2, \cdots, z_T\right), z_t \in R^n$, the final goal is to predict the value of the character modeling target sequence M at time T, as shown in Equation (11):

$$\hat{y}_T = M\left(y_1, y_2, \cdots, y_{T-1}, z_1, z_2, \cdots, z_T\right) \quad (11)$$

Afterward, the encoder, based on the structure of BiLSTM, maps the n-dimensional data of each time step to m dimensions, as shown in Equation (12):

$$h_t = f_1\left(h_{t-1}, z_t\right) \quad (12)$$

Inspired by the basic attention mechanism, the experiment proposes an input attention-based encoder. Given the k-th sequence $x^k = \left(x_1^k, x_2^k, \cdots, x_T^k\right)^T \in R^T$, the experiment utilizes a multi-layer perceptron to construct the attention mechanism, specifically represented as Equation (13):

$$\alpha_t^k = \frac{\exp\left(e_t^k\right)}{\sum_{i=1}^{n} \exp\left(e_t^i\right)} \tag{13}$$

For each individual sequence, e_t^k is computed, then softmax is utilized across different dimensions to calculate α_t^i . α_t^i represents the attention weights for the importance of the *k*-th input feature at time *t*. Based on these weights, the extracted sequence features and the hidden state at each time step are represented by Equations (14) and (15):

$$\tilde{x}_t = \left(\alpha_t^1 x_t^1, \alpha_t^2 x_t^2, \cdots, \alpha_t^n x_t^n\right)^T \tag{14}$$

$$h_t = f_1\left(h_{t-1}, \tilde{x}_t\right) \tag{15}$$

By employing this attention mechanism, the encoder can selectively focus on the data sequences of anime character modeling, rather than treating all sequences equally.

Subsequently, the experiment utilizes the CAM layer to dynamically adjust the model's feature extraction levels across different channels based on the importance of input data in those channels. The model pays more attention to the importance of different features in character design, such as facial features and clothing characteristics, thereby enhancing the realism and detail representation of the generated character images.

Furthermore, the feature fusion layer combines the sequence features output by the BiLSTM with the features adjusted by the attention mechanism to obtain more informative and diverse feature representations. Finally, the generated network layer receives the output of the feature fusion layer and generates anime character images or feature representations. In anime game character modeling design, the generated network layer can utilize the comprehensively considered feature representations to generate anime character images that meet the requirements, including aspects such as the appearance, posture, and expressions of the characters.

Thus, by learning the weights of time series data, the model can effectively recognize design requirements and make generation predictions even in the face of diverse anime character modeling data. The pseudocode of this model is depicted in Figure 5.

```
Start
Input: Anime character data input
Output:Animation character modeling design results
# Define BiLSTM layer
  lstm_units = 32  # Number of LSTM units
  lstm_fw_cell = tf.nn.rnn_cell.LSTMCell(lstm_units)
  lstm_bw_cell = tf.nn.rnn_cell.LSTMCell(lstm_units)
  output, _ = tf.nn.bidirectional_dynamic_rnn(lstm_fw_cell, lstm_bw_cell, inputs, dtype=tf.float32)
# Define temporal attention mechanism layer
  attention_weights_time = tf.nn.softmax(tf.layers.dense(output, 1, activation=None))
  attention_output_time = tf.reduce_sum(output * attention_weights_time, axis=1)
# Define channel attention mechanism layer
  attention_weights_channel = tf.nn.softmax(tf.layers.dense(attention_output_time, num_channels, activation=None))
  attention_output_channel = attention_output_time * attention_weights_channel
# Define feature fusion layer
  merged_features = tf.concat([output, attention_output_channel], axis=-1)
# Define generation network layer
  output_image = tf.layers.dense(merged_features, output_dim, activation='sigmoid')
# Define loss function and optimizer
target_image = tf.placeholder(tf.float32, shape=(None, output_dim))
loss = tf.losses.mean_squared_error(target_image, output_image)
optimizer = tf.train.AdamOptimizer(learning_rate=0.001).minimize(loss)
End
```

Figure 5: Pseudocode flowchart of multi-attention mechanism integrated BiLSTM algorithm applied to anime game character modeling design.

3.4 Experimental evaluation

To validate the algorithm performance of the anime game character modeling design model constructed in this study, an analysis is conducted using data from the Anime Face Dataset (https://www.kaggle.com/datasets/splcher/animefacedataset). This dataset encompasses various anime characters with distinct styles and features, including different hairstyles, expressions, eyes, lips, etc. It can be used to train and test artificial intelligence models, such as image generation models and facial recognition models, enabling the model to learn and generate realistic anime character images. Furthermore, the cross-validation method is employed to extend the dataset. By dividing the dataset into K subsets, one subset is used as the validation set while the remaining subsets are used as the training set. This process is repeated K times to ensure the stability of the model's performance across different data combinations. For algorithm validation, the development is primarily conducted on a Windows PC, with the following specific environment configuration:

GPU: NVIDIA RTX2070-8GB

CPU: Intel i7-9750H

Memory: 64GB

The specific software packages used include: Visual Studio 2022, SQL Server 2012, Net Framework 4.8.1, Python 3.9, PyCharm 2022.3.

In the hyperparameter setting, the optimizer used in this study is Adam, and the learning rate is set to 0.001, which is better in balancing the training speed and model performance. The batch size is chosen to be 32 to ensure a balance between computational efficiency and memory consumption during each training session. The depth of the network and the number of hidden nodes in each layer are adjusted according to the experimental requirements and data complexity, and finally a deep network structure with four hidden layers is adopted, each layer contains 64 nodes.

To analyze the performance of the model proposed in this study, it is compared with LSTM, CAM-LSTM [31], TAM-LSTM [32], and the model algorithm proposed by Zhou & Li (2024) in the relevant field. Evaluation is conducted based on accuracy, F1 score, Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and generation results.

4 Results and discussion

4.1 Precision analysis of various algorithms for anime character modeling design

The model algorithm proposed in this study is compared with LSTM, CAM-LSTM, TAM-LSTM, and the model algorithm proposed by Zhou & Li (2024) in the relevant field is analyzed from the convergence, accuracy and F1 values respectively, as shown in Figure 6-8.

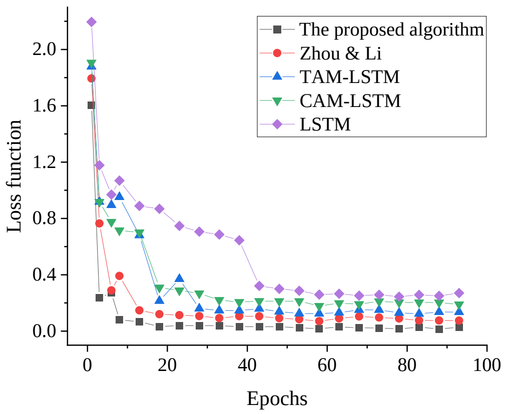

Figure 6: Convergence results of different algorithms

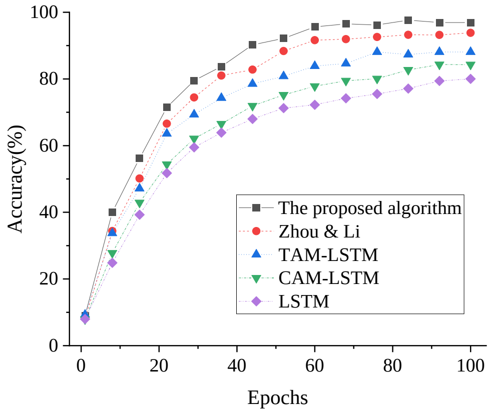

Figure 7: Accuracy of anime character modeling design prediction across iteration cycles for each algorithm

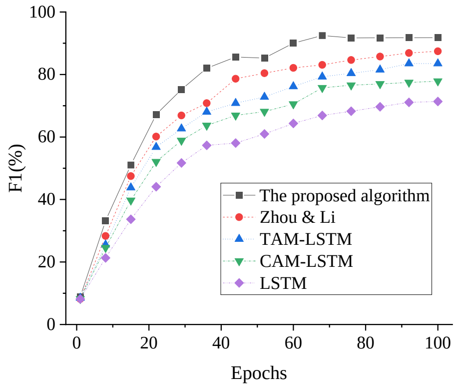

Figure 8: Results of anime character modeling design prediction F1 score across iteration cycles for each algorithm.

In Figure 6, the model algorithm of this study reaches a basic stable state when the iteration cycle is 18, and the convergence result of its loss value is the smallest, which is maintained at about 0.02. However, the final loss function of other algorithms is more than 0. 07. Therefore, the animation game character modeling design model based on multi-attention mechanism fusion BiLSTM algorithm proposed in this study has better convergence effect and lower loss value.

The analysis of accuracy and F1 score results for anime character modeling design prediction for each algorithm is shown in Figures 7 and 8. As the number of iteration cycles increases, the accuracy and F1 score of each algorithm exhibit an increasing trend followed by stabilization.

Compared to other model algorithms, the accuracy and F1 score of the proposed model algorithm reach 96.91% and 91.79%, respectively, representing an improvement of at least 3.08%. The ranking of anime character modeling design prediction accuracy among the algorithms, from highest to lowest, is as follows: the proposed model algorithm > the algorithm proposed by Zhou & Li (2024) > TAM-LSTM > CAM-LSTM > LSTM. Therefore, the constructed anime game character modeling design model, based on the fusion of multiple attention mechanisms and BiLSTM algorithm, performs better in terms of accuracy in generating anime character modeling designs.

The MAE and RMSE indicators of each algorithm under the 95% confidence interval are further compared, and the results are shown in Figure 9 and Figure 10.

In Figures 9 and 10, under the 95% confidence interval, the MAE and RMSE of this model algorithm reach 9. 66 and 7. 70 respectively when the iteration cycle is 100. Compared with other model algorithms, the error of animation character design generation is significantly lower. The reduction errors of MAE and RMSE are 6.52% and 10.7%, respectively. Therefore, the model algorithm constructed in this study can design the animation game role modeling more accurately.

4.2 Results of applying the model algorithm to character modeling design

Further results of applying the proposed model algorithm to anime game character modeling design are shown in Figure 11.

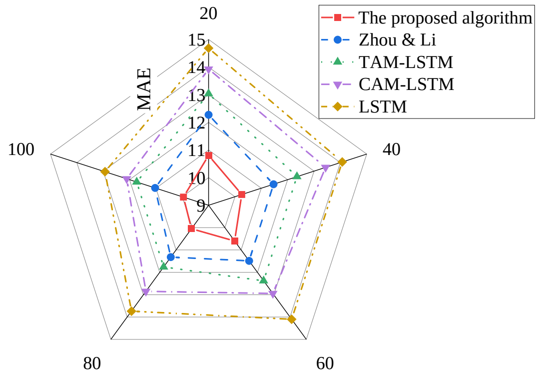

Figure 9: MAE results changing with the iteration cycle under each algorithm

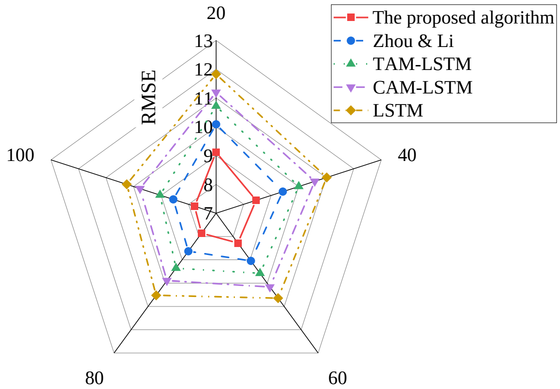

Figure 10: Results of RMSE across iteration cycles for each algorithm.

(a) (b) (c)

Figure 11: Results of anime game character modeling design (a. cute style character design; b. cool style character design; c. tech style character design).

Table 2: Comparison of computational cost and complexity of each algorithm

Algorithm	Number of parameters ($*10^2$)	Training time (s)
The proposed algorithm	8	115
Zhou & Li	10.5	149.5
TAM-LSTM	11.2	172.5
CAM-LSTM	13.1	207
LSTM	12.8	230

Figure 11(a) depicts a cute style character design with white long hair. The proposed model captures the coherence and temporal features of character design, such as the flowing dynamics of long hair. Analyzing large volumes of anime data to learn elements such as blue gemstone hair ornaments and silver hair clips, which enhance the sweet and charming qualities of the characters. For the cool style character design with black medium-length hair, as shown in Figure 11(b), the proposed model algorithm in this study emphasizes the character's strong personality and independence. It selects more mature and structured cuts in clothing and accessories, as well as the use of dark tones, to enhance the character's sense of strength. In the design of the tech style character with gray-white short hair, as shown in Figure 11(c), the model algorithm integrates futuristic details like smart interfaces into clothing and accessories by analyzing trends and user preferences in technology elements. This approach not only enhances the technological and modern feel of the design but also ensures that the character is closely linked to the theme of technology while maintaining its attractiveness and memorability. This intelligent design can efficiently create characters with a professional image.

The computational cost and complexity of each algorithm are further compared, as shown in Table 2.

In Table 2, although the number of parameters of the proposed algorithm is slightly higher than that of Zhou & Li's algorithm, the training time is significantly reduced, which indicates that the optimized multi-attention mechanism may improve the training efficiency. Although Zhou & Li algorithm has fewer parameters, it takes the longest time to train, which may be due to the lack of effective feature extraction mechanism. Although TAM-LSTM and CAM-LSTM have the same number of parameters, they introduce different attention mechanisms, which leads to different training time. CAM-LSTM takes the longest training time because it needs to deal with more channel features. Due to the lack of attention mechanism, the basic LSTM model has a relatively high number of parameters and training time. This shows that in the model design, although the introduction of attention mechanism increases the complexity of the model, it also improves the ability of the model to capture features, thus speeding up the training speed to a certain extent. However, the increase in model complexity also implies a higher computational cost, which needs to be carefully considered in resource-constrained situations.

5 Discussion

Through the analysis of the experimental results, the animation character modeling algorithm integrated with multi-attention mechanism and LSTM proposed in this study has significant advantages. Firstly, the multi-attention mechanism can flexibly assign weights, highlight the important visual features of animation character images, and improve the recognition ability of the model to details. Compared with the traditional CNN, the proposed method shows higher accuracy and robustness in dealing with complex image background and dynamic changes of roles. This echoes Huang et al. (2023) [33]. In addition, the introduction of LSTM enables the model to capture the time dependence in the sequence data, and to better model the dynamic expression and posture changes of the characters, thus improving the overall modeling quality.

Compared with the research in related fields, the method in this study shows superiority in a number of indicators. For example, compared with the image description generation method based on BiLSTM proposed by Li et al. (2024) [34], the F1 score of the model in this study is improved by about 0.08 on the same data set, which proves the enhancement effect of multi-attention mechanism on visual feature extraction. In addition, compared with the LSTM-GAN algorithm adopted by Yue et al. (2023) [35], the generation quality and computational efficiency of this method in the animation character modeling task have been significantly improved, especially in the dynamic change modeling of character posture and expression in complex scenes, showing higher accuracy, reaching more than 95%.

The significance of this study is not only to improve the accuracy and quality of animation character modeling, but also to provide a new algorithm integration paradigm for related fields. By combining the multi-attention mechanism with LSTM, this study successfully solves the shortcomings of traditional methods in dealing with complex image features and dynamic sequence data, and opens up a new path for future research. This contribution not only enriches the technical means in the field of animation character modeling, but also provides a reference for other fields that need to deal with complex visual features and time series data, which has a wide range of application potential and research value.

6 Conclusion

This study proposes an innovative approach combining multiple attention mechanisms and the BiLSTM algorithm for analyzing anime game character modeling. The application of this method can more accurately predict and generate character designs. Comparative experiments demonstrate that the model achieves over 90% accuracy and F1 score. The model successfully captures the characteristics of different character types, such as cute, cool, and tech styles, and generates attractive and memorable character images, providing valuable insights for the intelligent development of the anime field.

However, this study also has some limitations. Firstly, the current dataset may be biased toward specific types or styles of anime characters, which could result in the model having insufficient predictive ability for other types or styles. Therefore, future research can improve the model's training and performance by expanding the dataset to include more images of anime characters of different types and styles. Secondly, the current model may not encompass all artistic styles and design requirements, leading to limitations in the design process. Future research can explore more flexible and adaptive model architectures to address different artistic styles and design needs, thereby enhancing the model's applicability and practicality.

References

[1] Rong, F., & Wei, Q. (2023). Sustainable Development of Internet Literature in the Era of Games and Animation: Challenges and Opportunities. Contemporary Social Sciences, 8(1), 1-9. https://doi.org/10.19873/j.cnki.2096-0212.2023.01.006

[2] Hu, Z., Ding, Y., Wu, R., Li, L., Zhang, R., Hu, Y., ... & Fan, C. (2023). Deep learning applications in games: a survey from a data perspective. Applied Intelligence, 53(24), 31129-31164. https://doi.org/10.1007/s10489-023-05094-2

[3] Ding, Y., & Chen, C. L. (2023). Internationalization strategies and dynamic capabilities of platform companies-A case study of the game industry. Organizational Dynamics, 52(2), 100980. https://doi.org/10.1016/j.orgdyn.2023.100980

[4] Üstün, C. (2023). Animation Forms In Video Games And A Classification Study Based On Game Mechanics. Motif Akademi Halkbilimi Dergisi, 16(43), 1395-1406. https://doi.org/10.12981/mahder.1258045

[5] Alshingiti, Z., Alaqel, R., Al-Muhtadi, J., Haq, Q. E. U., Saleem, K., & Faheem, M. H. (2023). A deep learning-based phishing detection system using CNN, LSTM, and LSTM-CNN. Electronics, 12(1), 232. https://doi.org/10.3390/electronics12010232

[6] Madhusudhan, M. V., Udayarani, V., & Hegde, C. (2024). An intelligent deep learning LSTM-DM tool for finger vein recognition model USING DSAE classifier. International Journal of System Assurance Engineering and Management, 15(1), 532-540. https://doi.org/10.1007/s13198-022-01807-x

[7] Rai, V., Gupta, G., Joshi, S., Kumar, R., & Dwivedi, A. (2023). LSTM-based adaptive whale optimization model for classification of fused multimodality medical image. Signal, Image and Video Processing, 17(5), 2241-2250. https://doi.org/10.1007/s11760-022-02439-1

[8] Lu, S., Liu, M., Yin, L., Yin, Z., Liu, X., & Zheng, W. (2023). The multi-modal fusion in visual question answering: a review of attention mechanisms. PeerJ Computer Science, 9, e1400. https://doi.org/10.7717/peerj-cs.1400

[9] Klaar, A. C. R., Stefenon, S. F., Seman, L. O., Mariani, V. C., & Coelho, L. D. S. (2023). Optimized EWT-Seq2Seq-LSTM with attention mechanism to insulators fault prediction. Sensors, 23(6), 3202. https://doi.org/10.3390/s23063202

[10] Li, R., Oji, R., & Fujishiro, I. (2023). Controllable automatic generation of non-player characters in 3D anime style. Computer Animation and Virtual Worlds, 34(2), e2047. https://doi.org/10.1002/cav.2047

[11] Kopylova, O. (2023). Drawing Works Together: Anime Media Mix and Graphic Style as a Shared Cultural Resource. Mechademia, 16(1), 190-211. https://muse.jhu.edu/pub/23/article/910026

[12] Mahler, L. S., & Mayer, R. E. (2024). Anime as a medium for science learning. Journal of Computer Assisted Learning, 40(2), 787-796. https://doi.org/10.1111/jcal.12908

[13] Chen, D. (2024). Animation Vr Scene Stitching Modeling Based on Genetic Algorithm. Informatica, 48(5), 83-96. https://doi.org/10.31449/inf.v48i5.5364

[14] Liu, Y., Liu, Y., & Wen, J. (2022). Does anime, idol culture bring depression? Structural analysis and deep learning on subcultural identity and various psychological outcomes. Heliyon, 8(9), e10567. https://doi.org/10.1016/j.heliyon.2022.e10567

[15] Zhang, H., Ma, C., Jiang, Z., & Lian, J. (2022). Image caption generation using contextual information fusion with BiLSTM-s. IEEE Access, 11, 134-143. https://doi.org/10.1109/ACCESS.2022.3232508

[16] Aswiga, R. V., & Shanthi, A. P. (2022). A multilevel transfer learning technique and LSTM framework for generating medical captions for limited CT and DBT images. Journal of digital imaging, 35(3), 564-580. https://doi.org/10.1007/s10278-021-00567-7

[17] Son, Y., Zhang, X., Yoon, Y., Cho, J., & Choi, S. (2023). LSTM–GAN based cloud movement prediction in satellite images for PV forecast. Journal of Ambient Intelligence and Humanized Computing, 14(9), 12373-12386. https://doi.org/10.1007/s12652-022-04333-7

[18] Zabin, M., Choi, H. J., & Uddin, J. (2023). Hybrid deep transfer learning architecture for industrial fault diagnosis using Hilbert transform and DCNN–LSTM. The Journal of supercomputing, 79(5), 5181-5200. https://doi.org/10.1007/s11227-022-04830-8

[19] Tan, J., Wu, B., & Ma, Y. (2024). A Visual Characteristic Land-Scape Design for EEG Signal Based on LSTM-GAN. IEEE Access, 12, 41896-41907. https://doi.org/10.1109/ACCESS.2024.3377689

[20] Zhou, F., & Li, J. (2024). ECG data enhancement method using generate adversarial networks based on BiLSTM and CBAM. Physiological Measurement, 45(2), 025003. https://doi.org/10.1088/1361-6579/ad2218

[21] Mak, H. W. L., Han, R., & Yin, H. H. (2023). Application of variational autoEncoder (VAE) model and image processing approaches in game design. Sensors, 23(7), 3457. https://doi.org/10.3390/s23073457

[22] Oliveira, A., & Cruz, M. (2023). Virtually connected in a multiverse of madness? —perceptions of gaming, animation, and metaverse. Applied Sciences, 13(15), 8573. https://doi.org/10.3390/app13158573
[23] Younes, M., Kijak, E., Kulpa, R., Malinowski, S., & Multon, F. (2023). MAAIP: Multi-Agent Adversarial Interaction Priors for imitation from fighting demonstrations for physics-based characters. Proceedings of the ACM on Computer Graphics and Interactive Techniques, 6(3), 1-20. https://doi.org/10.48550/arXiv.2311.02502
[24] Li, P. (2023). The Development of 3D Movie Animation Industry in the New Media Art Era. Journal of Education, Humanities and Social Sciences, 18, 217-222. https://doi.org/10.54097/ehss.v18i.10996
[25] Kumar, K. S., & Bai, M. R. (2023). Lstm based texture classification and defect detection in a fabric. Measurement: Sensors, 26, 100603. https://doi.org/10.1016/j.measen.2022.100603
[26] Swee, S. K., Chen, L. C., Chiang, T. S., & Khim, T. C. (2023). Deep convolutional neural network for sem image noise variance classification. Engineering Letters, 31(1), 328-337. https://www.engineeringletters.com/issues_v31/issue_1/EL_31_1_35.pdf
[27] Ammaiappan, S., Liang, G., Tan, C., & Dong, F. (2023). A Priority-based Adaptive Firefly Optimized Conv-BiLSTM Algorithm for Electrical Resistance Image Reconstruction. IEEE Sensors Journal, 24(1), 624 - 634. https://doi.org/10.1109/JSEN.2023.3335321
[28] Li, Y., Peng, T., Sun, W., Ji, C., Wang, Y., Tao, Z., ... & Nazir, M. S. (2023). A soft sensor model based on CNN-BiLSTM and IHHO algorithm for Tennessee Eastman process. Measurement, 218, 113195. https://doi.org/10.1016/j.measurement.2023.113195
[29] Shi, H., Miao, K., & Ren, X. (2023). Short-term load forecasting based on CNN-BiLSTM with Bayesian optimization and attention mechanism. Concurrency and Computation: Practice and Experience, 35(17), e6676. https://doi.org/10.1002/cpe.6676
[30] Zhang, J., Ye, L., & Lai, Y. (2023). Stock price prediction using CNN-BiLSTM-Attention model. Mathematics, 11(9), 1985. https://doi.org/10.3390/math11091985
[31] Mekruksavanich, S., & Jitpattanakul, A. (2023). Hybrid convolution neural network with channel attention mechanism for sensor-based human activity recognition. Scientific Reports, 13(1), 12067. https://doi.org/10.1038/s41598-023-39080-y
[32] Yao, G. U., Haijun, X. I. E., Zipeng, Z. H. O. U., & Lu, L. I. (2023). An Attention mechanism-based CNN-BiLSTM real-time transient electromagnetic method. Coal Geology & Exploration, 51(10), 15. https://doi.org/10.12363/issn.1001-1986.22.12.1000
[33] Huang, Y., Liu, Q., Peng, H., Wang, J., Yang, Q., & Orellana-Martín, D. (2023). Sentiment classification using bidirectional LSTM-SNP model and attention mechanism. Expert Systems with Applications, 221, 119730. https://doi.org/10.1016/j.eswa.2023.119730
[34] Li, B., Li, H., Liang, Z., & Bai, X. (2024). Load Day-Ahead Automatic Generation Control Reserve Capacity Demand Prediction Based on the Attention-BiLSTM Network Model Optimized by Improved Whale Algorithm. Energies, 17(2), 415. https://doi.org/10.3390/en17020415
[35] Yue, M., Zheng, H., Cui, H., & Wu, Z. (2023). Gan-lstm-based ads-b attack detection in the context of air traffic control. IEEE Internet of Things Journal, 10(14), 12651-12665. https://doi.org/10.1109/JIOT.2023.3252809