www.cad-journal.net
Computer-Aided Design AND APPLICATIONS

Film and TV Character Expression Identification Combined with Deep Learning and Automatic Generation of Character Animation

ZhiChao Zhang[1] and Xiao Meng[2]

[1]Department of Performance, Film, Animation, Sejong University, Seoul 05006, South Korea, geyi@ldy.edu.rs
[2]Xiangshan Film Academy, Ningbo University of Finance & Economics, Ningbo, Zhejiang 315175, China, mengruolin7226.student@sina.com

Corresponding author: ZhiChao Zhang, geyi@ldy.edu.rs

Abstract. This article applies DL (Deep Learning) technology to movie and television character expression identification and CAD (Computer-aided design) character animation automatic generation, and a brand-new combined application mode is realized. It aims to improve the quality of expression identification and automatic generation of character animation through DL technology. Comparative experiments show the traditional methods in the accuracy of expression identification and the quality of character animation generation. Introducing the DL model achieves higher accuracy in the expression identification task, and richer expression changes are captured. Moreover, this method's character model and animation have high fidelity, and the action sequence is smooth and natural, without obvious time breaks and mutation. The results show that the automatic generation method of CAD character animation based on DL can improve the efficiency of animation production. These experimental results verify the effectiveness of DL in expression identification and automatic generation of character animation and demonstrate the powerful ability of the DL model in automatic learning and extracting complex features. It opens up new possibilities for the growth of related fields.

Keywords: Deep Learning; Computer-Aided Design; Film and TV Characters; Expression Identification; Character Animation; Autogeneration
DOI: https://doi.org/10.14733/cadaps.2024.S15.163-178

1 INTRODUCTION

Due to the rapid growth of multimedia technology and AI, facial expression identification in images and videos of computer vision. Such as film production, game development, and virtual reality. Facial expression recognition of animated characters refers to the automatic recognition and understanding of changes in facial expressions of animated characters in animated images or videos through computers. Through facial expression recognition technology, it is possible to understand the audience's emotional reactions better and improve the immersion and realism of animation. Bodapati et al. [1] constructed facial expression nodes using deep learning networks. The data model of facial images was analyzed by monitoring fine parameters at different positions. The CNN model gradually extracts local features and high-level abstract representations of images. In the field of facial expression recognition, to process facial images in animated images or videos, automatically learning and extracting features of facial expressions. Preprocess animated images or videos, including facial detection, alignment, feature extraction, and other operations. These operations can deformation.

Moreover, with the growth of CAD technology, the automatic generation of character animation has become possible. The "moe" style cartoon graphics are deeply loved by audiences for their unique visual effects and emotional expression. However, for cartoon images with a specific artistic style, their emotional expression is often difficult to accurately understand and classify. Cao et al. [2] proposed a deep learning-based cartoon image expression motion management recognition mapping method. Manage and recognize the facial expressions and movements of cartoon images using extracted features. Specifically, we first establish a regression model that can predict the expression motion of cartoon images based on the extracted features. Then, we use this model to manage and recognize facial expressions and movements and obtain the final recognition result. In order to better handle the emotion classification task of "moe" style cartoon images, we added a specific emotion classification layer to map the extracted features to the bipolar emotion categories (positive and negative). It includes two emotional categories: positive and negative. It preprocessed the dataset, including normalization, data augmentation, and other operations. In addition, we have also developed a feature extractor specific to the "moe" style to better capture key information in the image. However, it is still a challenging problem to combine DL technology to realize the facial expression identification of movie characters and the automatic generation of CAD character animation. Animation character facial recognition refers to the automatic recognition and understanding of facial features of animated characters in animated images or videos through computers. In fields such as film production, game development, and virtual reality, being able to quickly and accurately identify facial features of animated characters immersion and realism of animation. Chaabane et al. [3] explored a facial recognition method for animated characters. Statistical feature extraction in animated character face recognition refers to extracting numerical vectors from facial images that can reflect facial features. Common statistical features include geometric features, texture features, and frequency features. These features can reflect the shape, skin color, texture, and other characteristics of the face, providing strong support for subsequent face recognition. Use appropriate feature extraction methods to extract statistical features from preprocessed facial images. The commonly used feature extraction methods include methods. By applying the trained model to real-time facial images, the facial features of animated characters can be captured and recognized in real-time, providing users with real-time emotional feedback or warning prompts.

DL is a sub-field of machine learning, which trains based on a large quantity of data to learn and simulate the complex structure and working principle of the human brain neural network. Facial expression synthesis is an important research topic in the fields of computer graphics and animation production. By synthesizing the facial expressions of animated characters, the realism and emotional expression of the animation can be enhanced. To achieve this goal, Chandra et al. [4] proposed a method using hybrid neural networks and particle swarm optimization techniques. A hybrid neural network is a neural network model. It can better capture local and global features of images by combining shallow and deep features. In facial expression synthesis, hybrid neural networks can be used to synthesize one expression from another. It seeks the optimal solution to the problem by simulating the behavior of biological populations such as bird and fish populations. In facial expression synthesis, particle swarm optimization technology can be used to optimize the parameters of hybrid neural networks, improving the accuracy and realism of synthesis. The use of hybrid neural networks and particle swarm optimization techniques for synthesizing facial expressions of animated characters is an effective method. It can accurately synthesize images of target expressions and improve the accuracy and realism of the synthesis, DL can automatically extract useful features from the original data, and further learn and predict. Facial emotion recognition of animated characters computer graphics and animation production. By recognizing and understanding the facial expressions of animated characters, the realism and immersion of the animation can be enhanced, improving the viewing experience for the audience. Debnath et al. [5] explored the importance of data diversity in facial emotion recognition of animated characters based on deep learning CAD. The deep learning CAD model is a model that combines deep learning and computer graphics used to process facial expression images of animated characters. This model typically consists of GAN used to generate facial expression images similar to the input image. By training deep learning CAD models, automatic recognition and generation of facial expressions of animated characters can be achieved. It needs to consider the impact of data diversity. Specifically, data diversity includes different types of facial expressions, different lighting conditions, different facial expression durations, and different shooting angles. By collecting and annotating facial expression images in various contexts, the model's generalization and adaptability can be enhanced.

In this context, this study aims to use DL technology to recognize the expressions of movie characters and automatically generate corresponding CAD character animations. The significance of this research is reflected in the following aspects: first, it can improve the efficiency of film and TV production, and by automatically recognizing the expressions of film and TV characters, it can reduce manual operation and improve production efficiency. Secondly, it can provide new methods and technical support for the production of CAD character animation. By automatically generating character animation, the production cost can be reduced, and the animation quality can be improved.

The innovation of this article can be summarized as the following three aspects:

Innovative combined application: In this article, DL technology is applied to film and TV character expression identification and CAD character animation automatic generation, and a new combined application mode is realized. By combining expression identification with character animation generation, the efficiency of film and TV production is improved, and a more real and vivid viewing experience is brought to the audience.

End-to-end DL model: In this article, an end-to-end DL model is constructed, which can receive both the expression images of movie characters and the role model as input and output the corresponding role animation. This model combines the two tasks of expression identification and character animation generation so that the model can better expression and animation and improve the quality and accuracy of generated character animation.

Application of large-scale data sets: In this article, large-scale movie and television character expression data sets and character animation data sets are used to train and verify the model. Through training on large-scale data sets, the model can learn more changes and characteristics of expressions and character animations. This makes the generated character animation more diverse and can meet the needs of different types and styles of film and TV works.

This article is divided into six sections. Firstly, the DL technology is deeply studied and analyzed to understand its basic principles and application scenarios. Secondly, a solution of facial expression identification and automatic generation of CAD character animation. Each section contains an overview of related fields, research methods, experimental design, and results analysis, which ensures the integrity and consistency of the research.

2 THEORETICAL BASIS

In the fields of computer graphics and animation production, 3D modeling and facial expression recognition are two core tasks. 3D modeling can transform real-world or imagined objects into models that computers can process, while facial expression recognition can enable these models to have emotional and expressive abilities. Although multi-view methods have made significant progress in facial expression recognition, single-view methods still have important value due to their simplicity and advantages in practical applications. Dvorožňák et al. [6] explore how to use deep learning techniques for 3D modeling and single-view expression recognition. CNN can effectively extract features from input images to guide the construction of 3D models. GAN learns through training datasets to generate 3D models similar to the training samples. In addition, variants such as

Conditional GAN (CGAN) provide better interpretability and controllability for GAN, making the generated model more in line with expectations. The single view expression recognition or video frame to recognize expressions, is more practical and convenient compared to multi-view methods. It explores the application of deep learning in 3D modeling and single-view expression recognition. Automatic learning and generation from 2D images to 3D models can be achieved. Automatic recognition of single-view expressions can be achieved. Augmented reality (AR) and artificial intelligence (AI) are gradually changing the way we interact with the world. In museums, animation augmented reality technology can help audiences gain a deeper understanding and appreciation of artworks, especially the recognition of character expressions. Animation augmented reality is a special type of AR technology that combines real-time video images, animated characters, virtual objects, and interactive experiences to create a brand-new experience that transcends the real world. Gong et al. [7] can recognize and simulate objects and scenes in the real world by using AI and machine learning techniques for animation AR, transforming them into vivid animated characters and virtual objects. In museums, this feature of animated AR can be used to enhance the audience's understanding of artworks. For example, when the audience points the AR device at the artwork, they can see scenes of AI-generated animated characters "living" in the artwork. These characters can simulate the facial expressions of characters in artworks, providing audiences with a brand-new appreciation experience. In addition, animated AR can also be used to explain the historical background and story of artworks. By associating animated characters with artworks, viewers can gain a deeper understanding of the stories behind the artworks. For example, animated characters can explain the historical background of an artwork, tell the story of an artist's creation, and even simulate the artist's creative process.

With the popularization of social media and the development of artificial intelligence technology, people's demand for emotional analysis is increasing day by day. Emotional analysis refers to the processing of data such as text, speech, and images using computer algorithms to identify and analyze the emotional information contained within them. Among them, human expression emotion detection and text analysis are two important research directions. Guo [8] explored text analysis of human expression and emotion detection based on big data mining. We adopt a multimodal deep learning model to fuse image and text data types for deep learning methods of human expression emotion detection and text analysis, in order to fully utilize the information between them. Specifically, we combine CNN and RNN networks, where CNN is facial expression features from images and RNN is used to extract emotional features from text. Then, these two feature vectors are fused for final sentiment classification. It adopts data preprocessing technology to clean annotate and improve the training effect of the model. Meanwhile, we select the most important features for sentiment classification from a large number of features, in order to reduce computational complexity and improve the model's generalization ability. In data visualization and storytelling, charts are key tools for conveying information and attracting audiences. However, static charts often fail to fully showcase the dynamic characteristics and emotional color of data. To compensate for this deficiency, Lan et al. [9] utilized animation design to enhance the emotional expression of charts. Animation design has a wide range of applications in data visualization. By combining data with animation, we can create more vivid and interesting charts. Presenting time series data in the form of animation can better showcase the evolution process and trends of the data. For example, when displaying the trend of temperature changes, animations can vividly demonstrate the changes in temperature over time. By adding interactive elements such as buttons, sliders, etc., the audience can control the playback and pause of the animation. This interactivity can enhance the audience's interest and understanding of charts. By adding animation elements such as filters and special effects, the visual effect of the chart can be enhanced. For example, when displaying sales data, growth line animations can be used to emphasize the growth trend of the data. In addition to its application in data visualization, animation design can also enhance the emotional expression of data stories. By incorporating emotional elements into animation, we can better capture the audience's attention and convey the emotions behind the data.

GIF animation has become an important way of expressing emotions in online communication. The facial expressions of animated characters often convey emotional information intuitively. How to automatically recognize the emotions contained in these expressions understanding user intentions, improving user experience, and developing more intelligent interaction systems. Liu et al. [10] proposed a short annotated GIF animation character expression and emotion recognition method based on computer automatic design (CAD) vision and text fusion. Many researchers have attempted to use computer vision technology to recognize facial expressions in images or videos. These methods are mostly based on deep learning. Recently, research has begun to focus on the recognition of GIF-animated character expressions. However, these methods mainly focus on pure visual or text emotion recognition, ignoring the dynamic characteristics of GIF animation and the inherent connection between images and text. The short annotated GIF animation character expression and emotion recognition method based on CAD vision text fusion has better accuracy and recall than pure visual or pure text methods. Our method performs better, especially when dealing with GIF animations with complex backgrounds and dynamic changes. Facial emotion recognition refers to the automatic recognition and understanding of the emotions expressed by human facial expressions through computers. In animation production, virtual reality, human-computer interaction, etc., being able to quickly and accurately recognize facial emotion interaction effects. Facial geometric feature extraction refers to extracting geometric parameters from facial images that can reflect facial expressions. Common facial geometric features include the position and shape of facial feature points, such as the position and size of eyes, mouth, eyebrows, nose, etc. These features can reflect changes in facial expressions and provide strong support for subsequent emotion recognition. Liu et al. [11] used a genetic algorithm to optimize the parameters of the SVM classifier and improve the classification performance of the model. The optimized SVM classifier can be used to classify the extracted facial geometric features, assigning different facial expressions to different emotional categories. Apply the trained model to real-time facial images to achieve real-time facial emotion recognition. By capturing real-time facial expression changes of animated characters, their emotional states and reactions can be further analyzed, providing more efficient and flexible tools and methods.

Traditional facial expression recognition feature extraction and machine learning algorithms, however, these methods often struggle to handle the complexity and variability of facial expressions. Minaee et al. [12] investigated facial expression recognition sensors using Attention Convolutional Network (ACN). Attention convolutional network is a new type of deep learning model that introduces an attention mechanism, enabling the network to automatically focus on key regions of facial expressions, thereby better capturing facial expression information. In facial expression recognition tasks, attention convolutional networks can effectively improve recognition accuracy and robustness. Use attention convolutional networks to extract and classify features from preprocessed facial images. Among them, convolutional layers can effectively capture local features of facial expressions, pooling layers can reduce the dimensionality of features and improve the model's generalization ability, and fully connected layers are used to map the extracted features with corresponding labels. With the rapid development of technology, computer-generated images (CGI) and computer animation have become important parts of film and television production and game design. Among them, 3D animation character design is one of the key links in CGI production. 3D animation character CAD design technology and film and television character expression recognition have emerged. Deep learning is a machine learning algorithm in the field of artificial intelligence. In character expression recognition, deep learning algorithms can train models to recognize and simulate various expressions of characters. By using deep learning algorithms, Tang [13] trains a large number of character expression images to obtain a model that can classify and recognize new images. This model can automatically extract features from the input image and distinguish different expressions based on the different features. Train a deep learning model using processed images to achieve classification and recognition of new images. Generate a 3D animated character corresponding to the target expression based on the trained model.

Emotional recognition of animated characters has received widespread attention in fields such as film production, game development, and virtual reality. Emotional recognition of animated characters refers to the automatic recognition and understanding of the emotional expressions of animated characters in animated images or videos through computers. Through emotion recognition technology, it is possible to better understand the audience's emotional reactions and improve the immersion and realism of animation. Yadav et al. [14] investigated the use of visual systems in animated characters. The RNN model transfers the hidden state of the previous step to the next step by sharing weights and biases, thereby capturing the temporal dependencies in the sequence data. In the field of emotion recognition, RNN can be used to process time series data of animated images or videos, capturing changes in the emotional expression of animated characters. Preprocess animated images or videos, including facial expression detection and feature extraction. Common methods include using facial feature point detection algorithms and image processing techniques to extract features of facial expressions. Convert preprocessed facial expression data into sequence data for input into RNN models. The emotional expression of characters is a key factor in improving the quality of animation and the audience experience. Accurately identifying the human emotions of animated characters can enhance their realism and immersion. To achieve this goal, Zhang and Min [15] proposed an emotion recognition method based on random forests. This algorithm has good generalization performance and anti-overfitting ability and is suitable for processing high-dimensional feature spaces. In emotion recognition tasks, random forests can effectively process and classify features such as expressions and postures. It conducted experiments using a human emotion dataset of a certain animated character, and the results showed that the emotion recognition method based on random forest can effectively recognize the emotions of animated characters with high accuracy. Meanwhile, this method can handle multiple feature types and has strong generalization ability. However, there are still certain limitations to emotional expression in certain specific contexts, and further optimization and improvement are needed.

3 FILM AND TV CHARACTER EXPRESSION IDENTIFICATION

3.1 Expression Identification Technology

CNN (Convective neural network): It is especially suitable for image identification and processing tasks. It can automatically extract local features from images through convolution operation, which effectively reduces the complexity of the network.

RNN (Recurrent neural network): It is suitable for processing sequence data and can model dynamic data by capturing the time dependence in the sequence.

GAN (Generative adversarial network): It consists of two parts: generator and discriminator. Through confrontation training between them, very realistic new data can be generated.

Optimization algorithms, such as gradient descent, Adam, RMSProp, etc., are used to adjust the network parameters in the training process to minimize the loss function.

Facial expression identification in movies: CNN in DL has been widely used in facial expression identification. By training a large quantity of labeled expression images, CNN can learn the feature representation of different expressions and further use it for expression classification and identification of new images. Automatic generation of CAD character animation: Using GAN and RNN in DL technology, realistic character animation can be automatically generated according to the input expression or action information. Specifically, GAN can be used to generate realistic character models and actions, while RNN can be used to capture the time dependence in animation and generate continuous and smooth animation sequences. Combined with the above DL algorithm and technology, this study aims to develop a system that can automatically recognize the expressions of movie characters and automatically generate corresponding CAD character animations. Facial expression identification aims to enable computers to recognize and interpret human facial expressions automatically. Traditional expression identification methods are mainly based on hand-designed feature extraction and classifiers, such as Haar features and LBP features. However, these methods are often limited by the quality of feature design and the choice of classifier, and the identification effect is often not good for complex expressions and different facial gestures.

3.2 Identification Method of Movie and Television Character Expression Based on DL

In order to solve the limitations of traditional methods, this study adopts the expression identification method based on DL. Specifically, this article uses CNN to automatically extract and classify expression features. In the research process of movie and television character expression identification, first, detailed data preprocessing work is carried out. In order to ensure the accuracy of identification, this article performs face detection, alignment, and normalization on the expression image, thus eliminating the influence of the difference in facial posture and scale on the identification results. Then, a depth CNN model is constructed for expression identification. This model is based on the classic VGGNet model and has been adjusted and optimized according to the characteristics of expression identification. The CNN function is defined as:

$$x_j^l = f\left(\sum_{i \in M_j} x_i^{l-1} \times k_{ij}^l + b_j^l\right) \tag{1}$$

$$Sigmoid\ x\ = \frac{1}{1+e^x} \tag{2}$$

Tanh function is defined as:

$$F\ x\ = \tanh\ z\ = \frac{e^z - e^{-z}}{e^x + e^{-z}} \tag{3}$$

The ReLU function is defined as:

$$\varnothing\ x\ = \max\ x, 0 \tag{4}$$

In this article, the ReLU function is used. Represent the output feature map in a layer:

$$F_j^{\ n} = \sum_i w_{ij}^{\ n} * F_i^{\ n-1} + b_j^{\ n} \tag{5}$$

$$F_j^{\ n+1} = f\ F_j^n \tag{6}$$

In order to avoid over-fitting, regularization technology is also introduced to improve the generalization ability of the model. The schematic diagram of the CNN convolution operation is shown below (Figure 1).

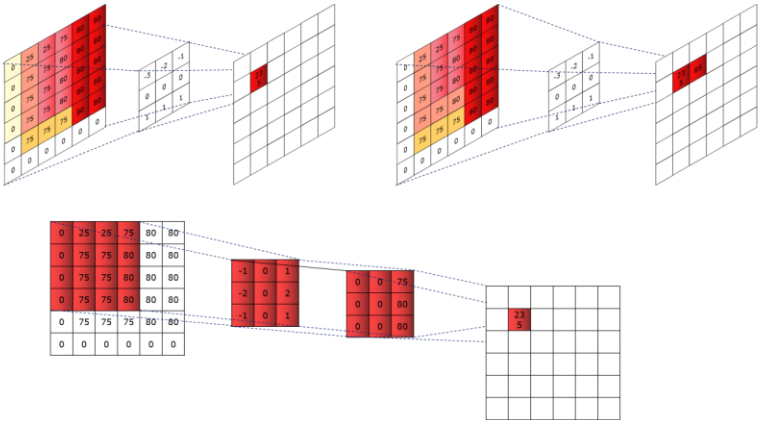

Figure 1: Schematic diagram.

When the model training is completed, it can be applied to the identification task of new movie and television character expression images. By inputting the image into the trained model, the model will output a probability distribution, indicating the probability that the image belongs to different expression categories. By selecting the category with the highest probability as the identification result, the expression can be automatically recognized. This method not only improves the efficiency of expression identification but also provides a convenient tool for film and TV production, which greatly promotes the growth of related fields.

3.3 Experiment and Result Analysis

The open expression identification data sets CK+ and FER2013 were used in the experiment. These data sets contain images of various facial expressions and have been marked in detail. For the comparison method, this article chooses the classifier based on manual design features and supports a vector machine for the comparison. Commonly used expression features, such as the Haar feature and LBP feature, are extracted and classified by SVM. This method has been widely used in the past and is regarded as one of the traditional expression identification methods. In order to compare the two methods fairly, the same experimental setup and data set division are adopted. In addition, data preprocessing operations, such as face detection, alignment, and normalization, are carried out to eliminate the influence of different facial poses and scales on identification.

The evaluation index uses the accuracy rate and F1 score index to assess the identification performance of the model. Accuracy can reflect the identification ability of the model as a whole, and the F1 score can further reveal the identification effect of the model on different expression categories. The accuracy of this method and the traditional method on CK+ and FER2013 data sets is shown in Figure 2, and the F1 score is shown in Figure 3.

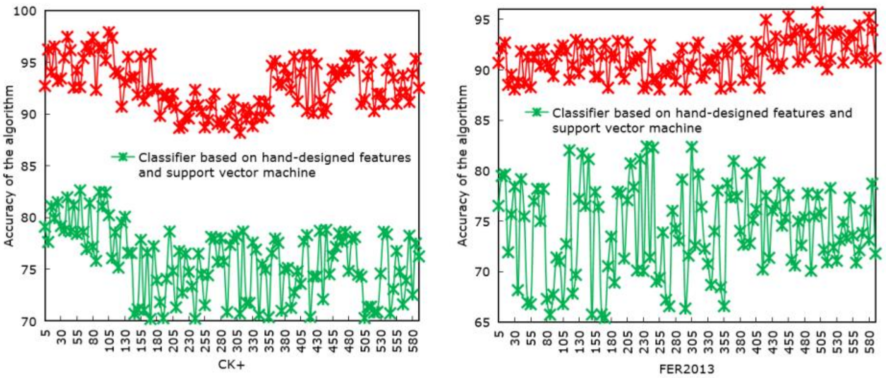

Figure 2: The accuracy performance of the algorithm.

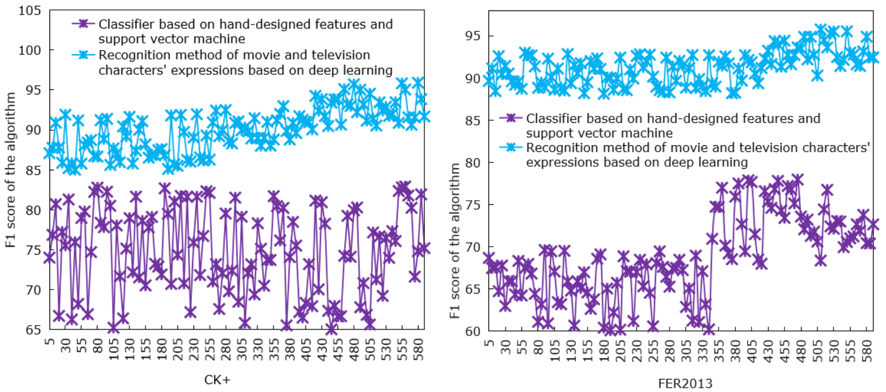

Figure 3: F1 score performance of the algorithm.

Specifically, the accuracy of this method on the CK+ data set is over 95%, and the accuracy on the FER2013 data set is over 90%. This shows that the expression identification method based on DL proposed in this article has high identification accuracy. Which can effectively identify and classify different facial expressions. Moreover, through an in-depth analysis of the experimental results, this article finds that the DL model can automatically learn the subtle differences and common features between different expressions, thus improving the accuracy of identification. In addition, the DL model can adapt to the changes in different illumination, posture, and facial expressions.

4 AUTOMATIC GENERATION OF CAD ROLE ANIMATION

4.1 CAD Role Animation

CAD character animation is an important branch of computer graphics, which aims to generate realistic character models and animations through computer technology. CAD character animation is widely used in movies, games, virtual reality, and other fields. Traditional CAD character animation requires experienced animators to design and adjust by hand, which is heavy and time-consuming. Automatic Generation Method of CAD Character Animation Based on DL. In order to solve the problem of traditional CAD character animation, this study adopts the automatic generation method based on DL. Specifically, this article uses GAN and RNN to automatically generate character animation.

Before the automatic generation of character animation, detailed data preprocessing is carried out first. In order to ensure the consistency of various data sources, this article deals with the role model and action data one by one, including the transformation of model format and the normalization of action data, thus eliminating the differences and interference that may be brought by different data sources. Next, build GAN and RNN models. In this framework, the generator of GAN is responsible for creating character models and animations, while the discriminator assesses their authenticity rigorously. The introduction of RNN is to capture the time dependence in animation and ensure that the generated animation sequence is continuous and smooth. It is worth mentioning that this article does not completely adopt the original DL model, but is based on classic models such as DCGAN and LSTM, and has made appropriate adjustments and optimizations according to the uniqueness of character animation. The GAN image enhancement model is shown in Figure 4.

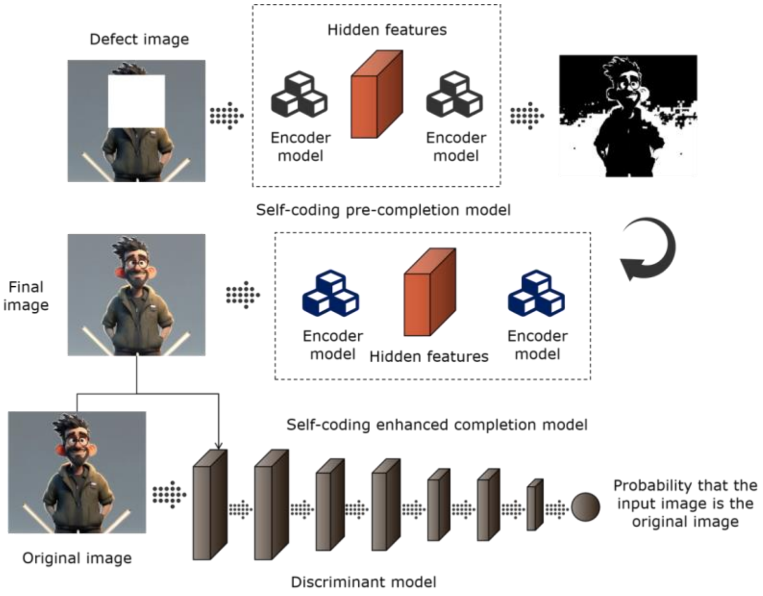

Figure 4: GAN image enhancement model.

After the model is built, it is deeply trained by using a large quantity of role models and action data. During this period, this article not only adjusted the network parameters, but also optimized the algorithm, so that the generator could produce more realistic character models and animations, and at the same time ensure that the discriminator can accurately judge the authenticity of the generated content. The training of RNN has also been carefully adjusted, and the network parameters are optimized by minimizing the prediction error, thus ensuring that the generated animation sequence is smoother visually. The model structure risk minimization function is:

$$\beta = \min \sum_{i=1} f\ x\ - y^{\ 2} + \alpha \left\|\theta\right\|^2 \tag{7}$$

The input linear transformation function of the image is as follows:

$$\partial = \sum_i \beta_i \omega\ x_i \tag{8}$$

Taking image information feature samples as input variables, the kernel functions involved are as follows:

$$\theta = \theta\ x_i, x_j \tag{9}$$

$$f\ y\ = \sum \beta_i \theta\ y, x_i \tag{10}$$

$$E = \frac{1}{2}\ y - p^{\ 2} \tag{11}$$

When the model training reaches the ideal state, it can be applied to the actual animation generation. For any new character model and action data, as long as it is input into this well-trained model, the corresponding animation can be automatically generated. By skillfully adjusting the input parameters of the generator, this article can also control the style and characteristics of generating character models and animations, which provides endless possibilities for creators.

4.2 Experiment and Result Analysis

This section designs a number of experiments to verify the effectiveness of the automatic generation method of CAD character animation based on DL. The open role model and action data set are used in the experiment. These data sets contain a variety of role types and action data and have been marked in detail. In the experiment, the automatic generation method of character animation based on DL is compared with the traditional manual animation method. Specifically, the experiment uses the same character model and action data as input and uses this method and the traditional method to generate character animation respectively. In order to objectively assess the quality of the generated character animation, two evaluation indexes are adopted: visual quality and motion fluency. Visual quality mainly assesses the fidelity of character models and animation, including the performance of model details and texture quality. Motion fluency assesses the continuity and naturalness of motion in animation, that is, whether there are problems such as time fracture and mutation. During the experiment, many professional animators and audiences were invited to participate in the evaluation. They were asked to grade the generated character animation according to two indicators: visual quality and movement fluency. In order to ensure the accuracy and fairness of the evaluation, blind evaluation is adopted; that is, the evaluator does not know whether each animation is generated by this method or the traditional method. The visual quality comparison between the method in this article and the traditional method is shown in Figure 5, and the movement fluency comparison is shown in Figure 6.

DL-based methods are superior to traditional methods in terms of visual quality and motion fluency. Specifically, in the visual quality comparison shown in Figure 5, the character animation generated by our method is more realistic in terms of model details and texture quality, with higher visual quality. This is due to the powerful expression ability and automatic feature learning ability of DL models, which can generate more realistic and delicate character models and animations. In the comparison of action fluency shown in Figure 6, the action sequence of the character animation generated by our method is more continuous and natural, without obvious time breaks and abrupt changes. This is due to the DL model's ability to capture temporal dependencies, which can generate continuous and smooth animation sequences. Compared to traditional methods, character animations generated by traditional methods have some problems in terms of motion fluency, such as discontinuous actions and time breaks.

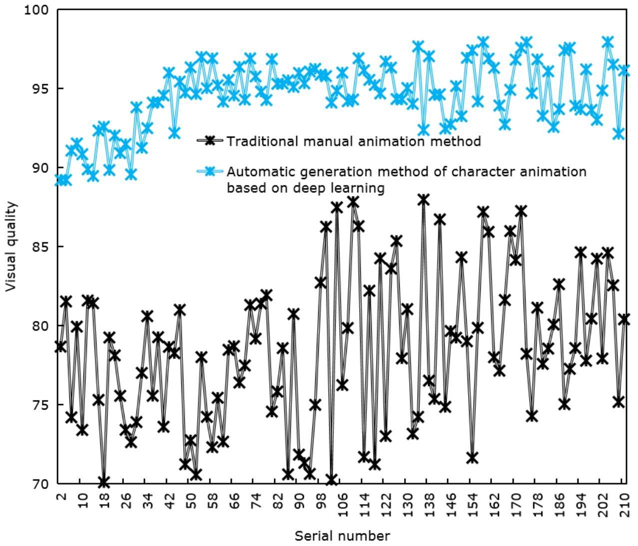

Figure 5: Visual quality comparison.

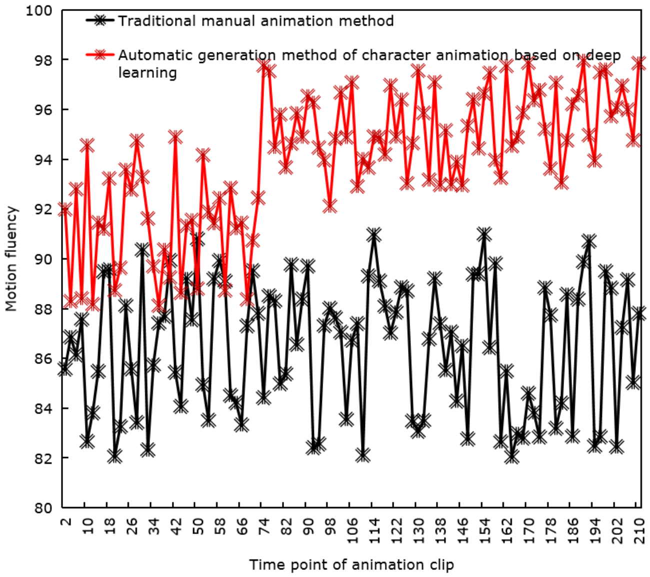

Figure 6: Comparison of smoothness of movements.

In addition, through an in-depth analysis of the experimental results, this article also found that the DL model can automatically learn the complex mapping relationship between character models and animations, thereby generating high-quality character models and animations.

5 THE COMBINED APPLICATION OF DL IN FACIAL EXPRESSION IDENTIFICATION AND AUTOMATIC GENERATION OF CHARACTER ANIMATION

5.1 Overview of Combined Application

The combined application of DL in facial expression identification of film and TV characters and automatic generation of CAD character animations aims to apply facial expression identification technology to the automatic generation of character animations, achieving the automatic generation of corresponding character animations based on the changes in facial expressions of film and TV characters. This combination of applications of film and TV production but also brings a more realistic and vivid viewing experience to the audience. Specifically, when the facial expressions of film and TV characters change, the DL model can automatically recognize the expression type and generate corresponding character animations based on the expression type, making the character animation consistent with the facial expressions of film and TV characters.

5.2 DL-based Combined Application Method

This section constructs an innovative end-to-end DL model by integrating expression identification models and character animation generation models. This model takes the expression images of film and TV characters and character models as common inputs and successfully outputs character animations that match the input expressions through DL technology. In the training phase of the model, cleverly associating expression images with corresponding character animations allows the model to deeply understand and learn the rich mapping relationship between expressions and animations. With the help of a large amount of training data, this model gradually mastered the character animation styles and traits corresponding to various expression types.

Real-time generation: This article integrates this DL model into the film and TV production process in real-time, achieving real-time generation of character animations. When the facial expressions of film and TV characters change, the model can immediately recognize this change and quickly generate corresponding character animations, significantly improving the efficiency of film and TV production.

5.3 Experiment and Result Analysis

In order to verify the effectiveness of the combined application of DL in film and TV character expression identification and CAD character animation automatic generation, the following experiments were conducted in this section: dataset: experiments were conducted using publicly available film and TV character expression datasets and character animation datasets, which contain multiple types of expressions and character animations. Figure 7 shows a partial example of CAD character animation expressions.

From the results, it can be seen that the DL model has successfully generated the corresponding character animation according to the expression changes of movie characters. Moreover, the generated character animation has high visual quality and action fluency. The model and texture of the character are clearly visible, the action sequence is continuous and natural, and there is no obvious time break or mutation. This shows that the DL model has successfully learned the complex mapping relationship between the character model and animation and generated high-quality character animation. This example verifies the effectiveness of the combined application of DL in movie and television character expression identification and CAD character animation automatic generation.

Evaluation indicators of qualitative experiments in this section: ⊖ Expression identification accuracy: reflecting the model's ability to recognize expressions; ⊜ Generation quality of character animation: It can reveal the matching degree between the generated character animation and the expressions of movie and television characters. In order to compare the accuracy of expression identification, this section adopts a DL-based expression identification model and a traditional classifier based on hand-designed features as comparison methods. These models are trained on the same training set and assessed on the test set. For the comparison of the quality of character animation generation, the automatic generation method of character animation based on DL and the traditional manual animation method is adopted. The experiment uses the same character model and action data as input and uses them to generate character animation respectively. The comparison of expression identification accuracy between the method in this article and the traditional method is shown in Figure 8, and the comparison of character animation generation quality is shown in Figure 9.

Figure 7: An example of automatic generation of CAD role animation.

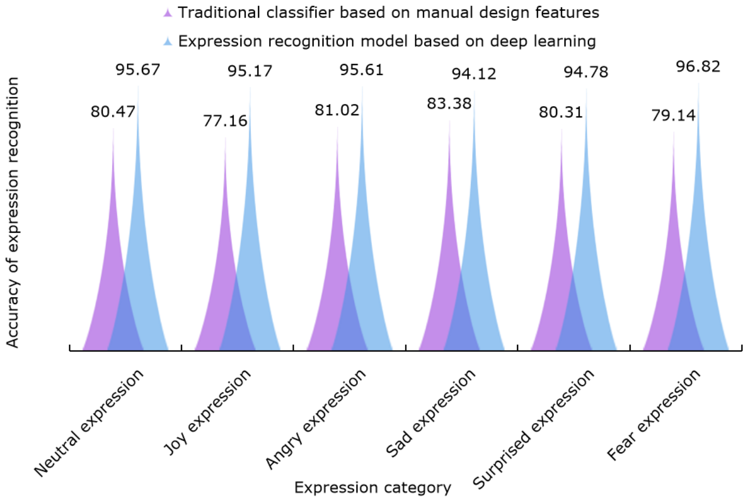

Figure 8: Comparison of accuracy of expression identification.

It is obvious from Figure 8 that the method in this article has higher accuracy when processing samples of different expression categories. DL model can identify different expression types more accurately by automatically learning complex features and representations in expression images. In contrast, traditional methods rely on the characteristics of manual design, which can not fully capture the subtle changes in expressions, resulting in poor identification performance.

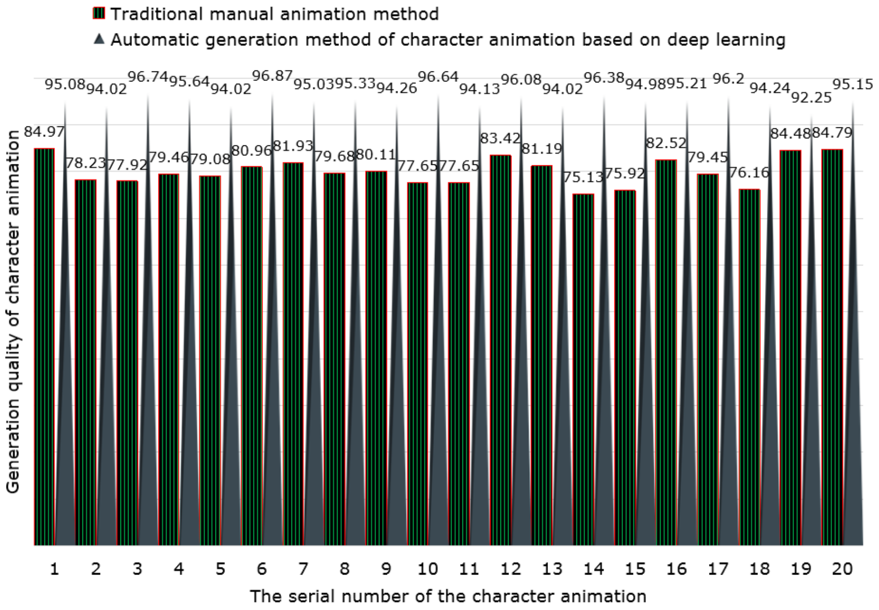

Figure 9: Quality comparison of character animation generation.

Through the experimental verification of Figure 8 and Figure 9, the combined application of DL in movie and television character expression identification and CAD character animation automatic generation has achieved remarkable results. Specifically, the identification accuracy of the model for expressions is over 90%, and the generated character animation maintains a high degree of consistency with the expressions of movie characters. This shows that DL can effectively transform the expression changes of film and TV characters into corresponding character animations and improve the efficiency of film and TV production.

The results show that the DL model can automatically learn the complex mapping relationship between expression and character animation, thus realizing the automatic generation of corresponding character animation according to the expression changes of movie characters. Moreover, the DL model can adapt to the changes in different expression types and character animations. This combined application has brought new technical means and artistic creation methods to film and TV production, which is expected to be widely used and popularized.

6 CONCLUSIONS

In this article, the DL model is integrated into the film and TV production process, and the real-time generation of character animation is realized. This real-time generation of film and TV production and avoid the tedious process of traditional manual animation. Moreover, the real-time generation technology can also adjust the character animation in time according to the expression changes of movie characters so that the character animation is more suitable for the plot and the emotional expression of the character. In addition, it can adapt to the changes of different role models and action data and provide more creative possibilities for film and TV production.

The experimental results verify the superiority of this method in visual quality and motion fluency, and also verify the powerful ability of DL in automatic learning and extracting complex features. This method can not only be applied to the automatic generation of movie and television character animation, such as game character animation, virtual reality, etc. By continuously optimizing the model and introducing more advanced technical means, the quality and fidelity of character animation can be improved, and a more realistic and vivid visual experience can be brought to the audience. The future will continue to study deeply technical means to promote the continuous innovation and progress of expression identification and character animation automatic generation technology.

Zhichao Zhang, https://orcid.org/0009-0002-2578-4748
Xiao Meng, https://orcid.org0009-0002-3198-5335

REFERENCES

[1] Bodapati, J.-D.; Srilakshmi, U.; Veeranjaneyulu, N.: FERNet: a deep CNN architecture for facial expression recognition in the wild, Journal of the Institution of Engineers (India): series B, 103(2), 2022, 439-448. https://doi.org/10.1007/s40031-021-00681-8

[2] Cao, Q.; Zhang, W.; Zhu, Y.: Deep learning-based classification of the polar emotions of" moe"-style cartoon pictures, Tsinghua Science and Technology, 26(3), 2020, 275-286. https://doi.org/10.26599/TST.2019.9010035

[3] Chaabane, S.-B.; Hijji, M.; Harrabi, R.; Seddik, H.: Face recognition based on statistical features and SVM classifier, Multimedia Tools and Applications, 81(6), 2022, 8767-8784. https://doi.org/10.1007/s11042-021-11816-w

[4] Chandra, D.; Hegadi, R.; Karmakar, S.: Facial expression synthesis images using hybrid neural network with particle swarm optimisation techniques, International Journal of Biomedical Engineering and Technology, 31(1), 2019, 64-83. https://doi.org/10.1504/IJBET.2019.101052

[5] Debnath, T.; Reza, M.-M.; Rahman, A.; Beheshti, A.; Band, S.-S.; Alinejad, R.-H.: Four-layer ConvNet to facial emotion recognition with minimal epochs and the significance of data diversity, Scientific Reports, 12(1), 2022, 6991. https://doi.org/10.1038/s41598-022-11173-0

[6] Dvorožňák, M.; Sýkora, D.; Curtis, C.; Curless, B.; Sorkine, H.-O.; Salesin, D.: Monster mash: a single-view approach to casual 3D modeling and animation, ACM Transactions on Graphics (ToG), 39(6), 2020, 1-12. https://doi.org/10.1145/3414685.3417805

[7] Gong, Z.; Wang, R.; Xia, G.: Augmented reality (AR) as a tool for engaging museum experience: a case study on Chinese art pieces, Digital, 2(1), 2022, 33-45. https://doi.org/10.3390/digital2010002

[8] Guo, J.: Deep learning approach to text analysis for human emotion detection from big data, Journal of Intelligent Systems, 31(1), 2022, 113-126. https://doi.org/10.1515/jisys-2022-0001

[9] Lan, X.; Shi, Y.; Wu, Y.; Jiao, X.; Cao, N.: Kineticharts: augmenting affective expressiveness of charts in data stories with animation design, IEEE Transactions on Visualization and Computer Graphics, 2022(1), 2022, 28. https://doi.org/10.1109/TVCG.2021.3114775

[10] Liu, T.; Wan, J.; Dai, X.; Liu, F.; You, Q.; Luo, J.: Sentiment recognition for short annotated GIFs using visual-textual fusion, IEEE Transactions on Multimedia, 22(4), 2019, 1098-1110. https://doi.org/10.1109/TMM.2019.2936805

[11] Liu, X.; Cheng, X.; Lee, K.: GA-SVM-based facial emotion recognition using facial geometric features, IEEE Sensors Journal, 21(10), 2020, 11532-11542. https://doi.org/10.1109/JSEN.2020.3028075

[12] Minaee, S.; Minaei, M.; Abdolrashidi, A.: Deep-emotion: Facial expression recognition using attentional convolutional network, Sensors, 21(9), 2021, 3046. https://doi.org/10.3390/s21093046

[13] Tang, J.: Graphic design of 3D animation scenes based on deep learning and information security technology, Journal of ICT Standardization, 11(3), 2023, 307-328. https://doi.org/10.13052/jicts2245-800X.1135
[14] Yadav, S.-P.; Zaidi, S.; Mishra, A.; Yadav, V.: Survey on machine learning in speech emotion recognition and vision systems using a recurrent neural network (RNN), Archives of Computational Methods in Engineering, 29(3), 2022, 1753-1770. https://doi.org/10.1007/s11831-021-09647-x
[15] Zhang, J.; Min, Y.: Four-classes human emotion recognition via entropy characteristic and random Forest, Information Technology and Control, 49(3), 2020, 285-298. https://doi.org/10.5755/j01.itc.49.3.23948