Machine Learning-Driven Emotional Feedback Analysis and Adaptive Content Generation for VR Movie and TV Users

Yun TANG
College of Digital Technology and Design,
Chongqing Polytechnic University of Electronic Technology, Chongqing 401331, China

***Abstract*—With the growing demand for immersive audiovisual experiences, user sentiment feedback analysis has become a pivotal factor in improving personalization and interactivity in virtual reality (VR) movie and television. This study proposes a machine learning - driven framework that integrates sentiment feedback recognition and adaptive content generation to optimize user experience. First, a Long Short-Term Memory (LSTM) model is developed to analyze multimodal sentiment feedback data, including physiological signals, behavioral responses, and interactive actions. The model achieves an average recognition accuracy of 75.75% across four basic emotions—happiness, sadness, anger, and fear—demonstrating its ability to capture dynamic and continuous emotional patterns. Based on real-time sentiment feedback, a Deep Q-Network (DQN) reinforcement learning algorithm is employed to generate adaptive VR content that aligns with users' current emotional states. Experimental validation with 100 participants shows that adaptive content generation increases overall satisfaction scores from 6.2 to 7.8, and the matching degree between user emotions and content improves by more than 20%. The integration of sentiment feedback analysis and reinforcement learning establishes a closed feedback loop—emotion detection → adaptive adjustment → feedback optimization—that enhances immersion, empathy, and user engagement. This research provides a data-driven reference for the intelligent evolution of VR movie and television, and future work will expand to fine-grained emotional dimensions and multimodal fusion to improve recognition precision and real-time adaptive generation performance.**

***Keywords*—Machine learning; VR movie and television; user sentiment feedback analysis; adaptive content generation; reinforcement learning**

I. INTRODUCTION

In recent years, virtual reality (VR) technology has made great progress, and VR movie and television, as one of its important application areas, has gradually come into the public's field of vision [1]. Different from traditional movie and television, VR movie and television can provide users with a more immersive viewing experience, so that users seem to be in the movie and television scenes [2]. However, different users have different aesthetic preferences and emotional tendencies, so how to adjust the VR movie and television content in real time according to the emotional feedback of the users has become the key to improve the user experience [3]. The study of VR movie and television user emotional feedback analysis and content adaptive generation can not only enhance the user's viewing satisfaction and immersion, enhance the user's stickiness and acceptance of VR movie and television [4], but also help VR movie and television creators to better understand the user's needs, provide a reference basis for the creation of content, and promote the diversification of VR movie and television content and personalized development [5]. Therefore, in-depth study of VR movie and television user emotional feedback analysis and content adaptive generation is of great significance for promoting the development of VR movie and television industry [6].

Combined with machine learning algorithms, the research content of VR movie and television user emotional feedback analysis and content adaptive generation mainly includes two aspects, such as VR movie and television user emotional feedback analysis and content adaptive generation technology [7]. VR movie and television user emotional feedback analysis includes the collection, collation, and analysis of emotional feedback data [8]; machine learning-based content adaptive generation technology mainly explores how to achieve dynamic adjustment and generation of VR movie and television content according to the results of emotional analysis [9]. Currently, in the field of VR movie and television emotion analysis and content generation, Yu and Luo [10] developed more accurate EEG and heart rate signal processing algorithms, which can more efficiently extract emotional features, providing strong support for in-depth analysis of the emotional state of users watching VR movie and television; Cheng et al., [11] combined with machine learning technology to initially realize simple emotion classification based on a single physiological signal, but in the face of complex emotions is still apparent. Insufficient, and less research on multimodal data fusion; Jones et al., [12] use machine learning to generate simple virtual scene elements, but the overall creativity and diversity are insufficient, failing to fully meet the user's personalized needs. Figueroa et al., [13] excel in the theoretical study of emotional computing, and develops a variety of advanced technologies, which can efficiently collect and analyze the user's physiological signals and behavioral data, with a high accuracy rate of emotion recognition. Li et al., [14] conduct in-depth research on complex emotions and emotional state when watching VR movies. Dooley [15] explores the content generation method based on machine learning, and develops complex and realistic virtual scene and character generation technology. Current research is facing many difficult problems, as follows: 1) emotion recognition, user emotions are complex, fuzzy and intertwined, a single algorithm is difficult to accurately portray; 2) in the content generation dimension, the existing graphics processor (GPU) performance is still difficult to meet the demanding needs of milliseconds latency.

Aiming at the above problems, this study innovatively proposes an integrated framework of sentiment analysis and content generation that integrates multiple machine learning algorithms (deep learning [16] for feature extraction and reinforcement learning [17] for decision generation), breaking the previous research limitations of the separation of the two. The main contributions are as follows: 1) proposed an integrated framework for sentiment analysis and content generation, aiming to make the emotional feedback accurately guide the content generation, and improve the adaptability and coherence of the overall system; 2) developed a multimodal data preprocessing and fusion algorithm to address the problems of multimodal data synchronization and format inconsistency; 3) introduced a reinforcement learning algorithm to allow the system to dynamically adjust the VR movie and television content; 4) the user test and experimental verification, the method makes the emotion recognition accuracy rate than the traditional algorithm to improve more than 15%, content generation satisfaction increased by about 20%, a strong impetus to VR movie and television towards a new era of intelligent, personalized, for the subsequent large-scale commercial applications to pave the technical cornerstone.

This paper firstly describes the research background and significance of VR movie and TV users' emotional feedback analysis and content adaptive generation, and then systematically discusses data collection and preprocessing, emotion analysis model construction, content adaptive generation model development based on reinforcement learning, experimental validation and other aspects, and finally validates and analyzes the proposed method through experiments, demonstrating the advantages of the proposed method in terms of emotion recognition accuracy and content generation satisfaction. content generation satisfaction.

II. Emotional Feedback from VR Movie and TV Users

A. Affective Feedback Data Types

1) VR Movie and TV: VR movie and television is a fusion of virtual reality (VR) [18] technology and movie and television art of innovative media forms. VR movie and television can be understood as based on the head-mounted display device to watch or experience the image works, combined with the panoramic picture image, head-mounted display device to the audience's field of vision for the closed package, giving the audience immersed in the image experience, the formation of VR movies unique presence sense of quality, specific as shown in Fig. 1.

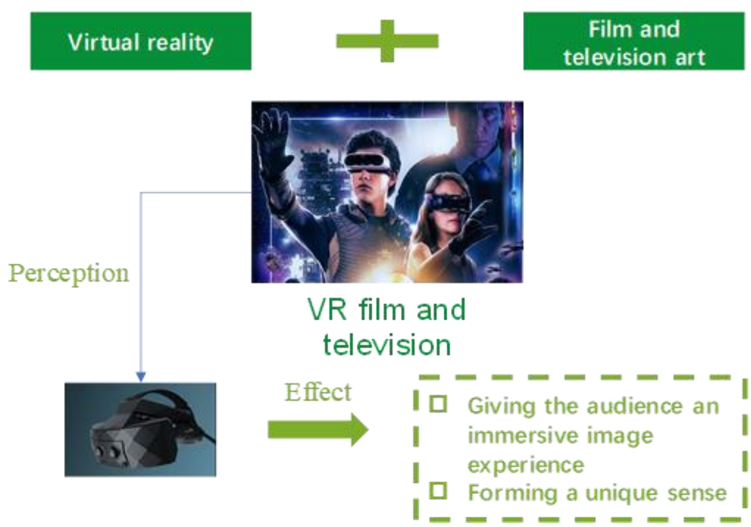

Fig. 1. Definition of VR movie and television.

2) User emotional feedback: VR movie and television user emotional feedback refers to the user in the process of watching or experiencing VR movie and television works, based on their inner feelings and emotional state and the various responses to the works [19], specifically as shown in Fig. 2. This kind of reaction covers physiological, behavioral, psychological and other levels, and can reflect the user's preference for VR movie and television content, interest and whether it resonates with the emotional state [20]. By analyzing these feedbacks, creators and platforms can better understand how users feel, and thus optimize content creation and enhance user experience.

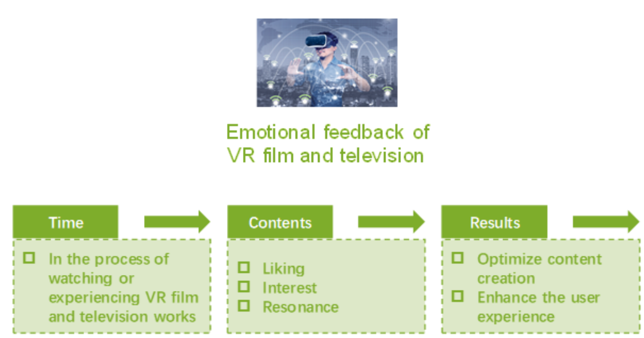

Fig. 2. Emotional feedback from VR movie and TV users.

3) Data classification: In the field of VR movie and television, user emotional feedback data is rich and diverse, which can be mainly classified into three categories, such as physiological signal data, behavioral data, and interaction data, as shown in Fig. 3.

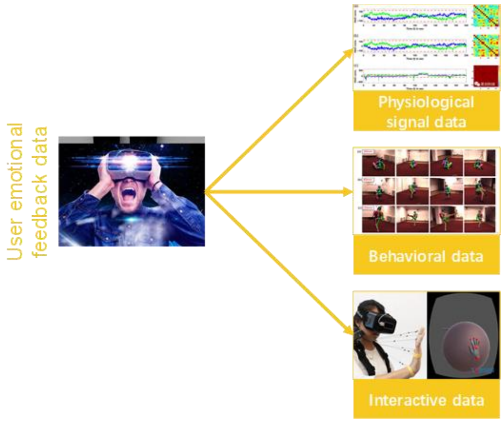

Fig. 3. Classification of user sentiment feedback data.

a) Physiological signaling data: Physiological signal data are intuitive data that reflect the user's internal emotional state [21]. Electroencephalogram (EEG) can capture the neural activity of the brain under different emotions; heart rate variability (HRV) is closely related to the level of emotional arousal, with heart rate accelerating and HRV decreasing when the emotions are excited, and the opposite is true when the emotions are calm; and the electrical skin activity (EDA) can reflect the level of emotional arousal. From Fig. 4, the collection of these data requires the use of professional equipment, which can record the physiological reactions of users when watching VR movies and videos in real time, providing a physiological basis for in-depth analysis of emotions.

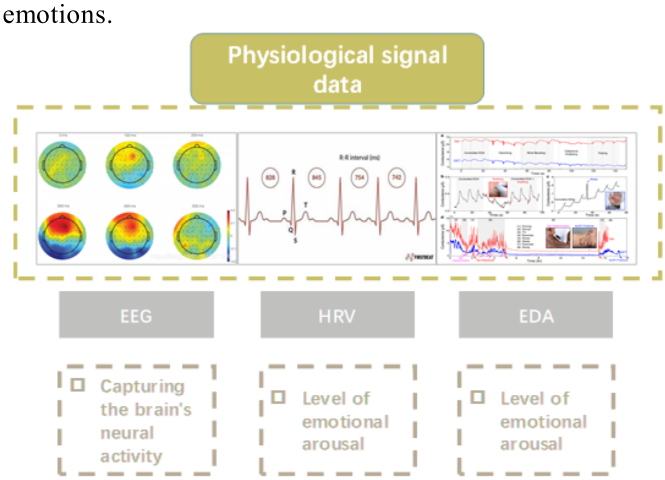

Fig. 4. Physiological signal data content and acquisition.

b) Behavioral data: Behavioral data covers a variety of users' behavioral performance in the VR environment [22]. In Fig. 5, the direction and speed of head movement can reflect the user's point of interest in the scene; the duration of the eye gaze point is a key indicator of the user's attention level, the longer the gaze time, the greater the interest in the element; body movements such as the user's approach to the virtual object may be a reflection of love or exploration, while away from it may be a reaction of aversion or fear. Through the eye-tracking module and motion capture sensors of VR devices, these behavioral data can be accurately recorded, and the user's emotional tendency can then be inferred.

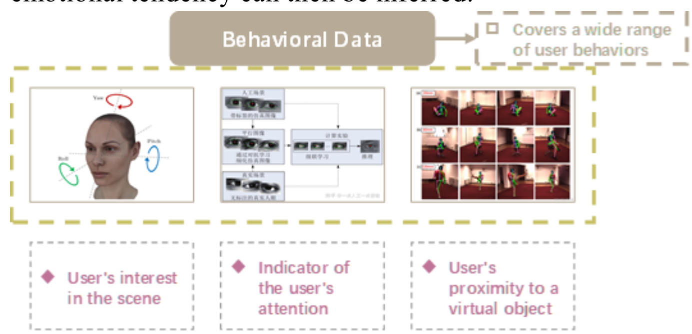

Fig. 5. Behavioral data content and collection.

c) Interaction data: Interaction data refers to the operation records of users interacting with elements in VR movies and TV [23]. Users' behaviors such as clicking buttons in the scene, selecting plot branches, and conversing with virtual characters directly reflect users' preferences and emotional needs. These data can provide clear user intent information for content adaptive generation and help the system accurately adjust the content to meet users' personalized emotional needs.

B. Methods of Sentiment Analysis

Sentiment analysis methods mainly include convolutional neural networks, recurrent neural networks, fusion algorithms, etc., and the specific types and scope of application are shown in Fig. 6.

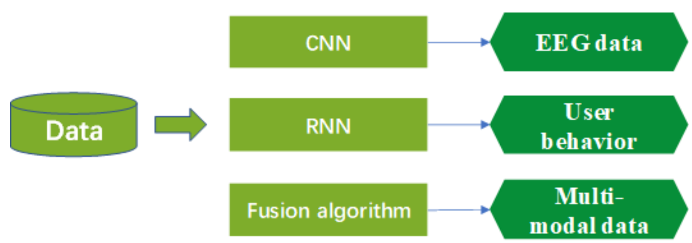

Fig. 6. Sentiment analysis methods.

Convolutional neural networks (CNN) [24] are good at processing data with spatial structure. Take the EEG signal as an example, after preprocessing, it is transformed into a two-dimensional matrix form for input into the CNN model.

Recurrent neural networks (RNN) and long short-term memory networks (LSTM) are particularly suitable for processing time-sequential data, such as the user's continuous behavioral data. LSTM can memorize features such as the location and dwell time of the user's previous gaze point, and effectively solve the gradient vanishing problem of the traditional RNN through the gating mechanism, so as to more accurately capture the trend of the user's emotion over time, and to judge the user's Emotional dynamics.

In order to fully utilize the advantages of different types of data, a data fusion strategy is adopted. At the feature level, the feature vector extracted from physiological signal data, the statistical features of behavioral data, and the coded features of interaction data are spliced together to form a comprehensive feature vector, which is used as the input to the machine learning model to characterize the user's emotional state in a more comprehensive way.

C. Program Design

A multimodal sentiment feedback data acquisition platform is constructed, which is the basis of the whole sentiment analysis program, as shown in Fig. 7. The platform includes key points such as data acquisition, data preprocessing, data division, learning and training, and feedback analysis.

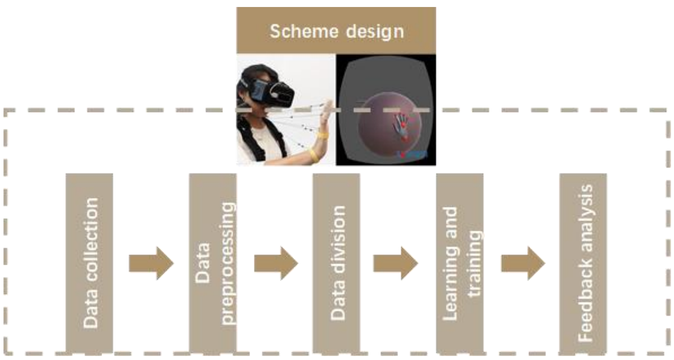

Fig. 7. Program design.

1) The platform includes physiological signal acquisition devices (EEG caps, heart rate bands, etc.), behavioral tracking devices (eye tracking modules for VR headsets, motion capture sensors, etc.), and an interaction recording module (which records the user's operational commands in VR movies and TVs).

2) Preprocess the collected data, including filtering to remove noise, normalization to unify the data scale, and other operations to improve the data quality.

3) Construct a training set and a test set according to the above sentiment analysis method, train the sentiment analysis model using deep learning frameworks (TensorFlow, PyTorch), and continuously optimize the model parameters through cross-validation and other methods to improve the accuracy and robustness of sentiment recognition.

4) In practical applications, the trained model is deployed to the VR movie and television system, which receives real-time emotional feedback data from users, and inputs it into the model for emotional classification after preprocessing to provide real-time and accurate emotional information for the subsequent adaptive generation of content, so as to realize the function of dynamically adjusting VR movie and television content according to the emotional feedback of the users and enhance the viewing experience of the users.

III. Content Adaptive Generation Techniques

A. Content Adaptive Generation Analysis

Content adaptive generation refers to the VR movie and television playback process, the system based on real-time access to the user's emotional feedback information, automatic movie and television content dynamic adjustment and optimization, in order to generate more in line with the user's current emotional state and personalized needs of the virtual scene, plot, character behavior, etc., so as to enhance the user's sense of immersion and the viewing experience [25], specifically as shown in Fig. 8.

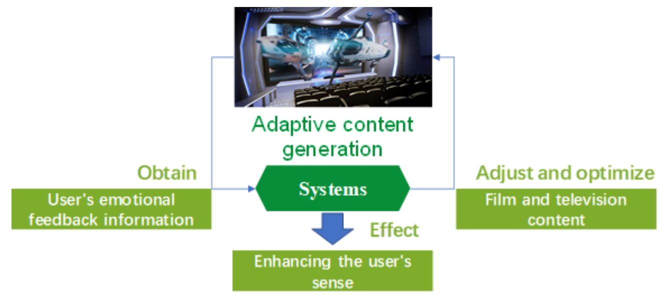

Fig. 8. Content adaptive generation technology.

Content adaptive generation mainly includes scene environment adjustment, plot branch selection, and character interaction adjustment [26], as shown in Fig. 9. Scene environment adjustment includes the change of the scene's color tone, light and shadow effects, background music and other atmospheric elements; plot branch selection is to adjust the development direction of the plot in real time according to the user's emotional preferences; and character interaction adjustment is mainly to change the virtual character's response to the user and behavioral patterns.

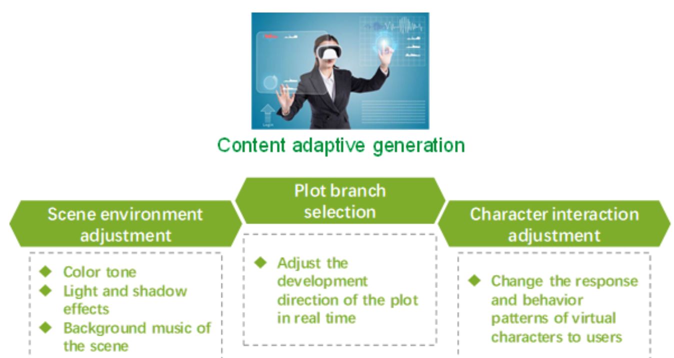

Fig. 9. Content adaptive generation main content.

Based on the principle and content of adaptive content generation, its main features are reflected in real-time, personalization, and coherence.

B. Machine Learning-Based Adaptive

Content adaptive generation methods include traditional algorithms, machine learning algorithms, retrieval enhancement algorithms, rule-based methods, etc., as shown in Fig. 10.

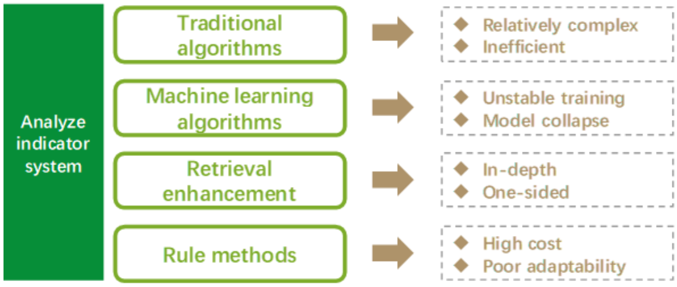

Fig. 10. Content adaptive.

Adaptive generation methods based on traditional algorithms are relatively complex and less efficient for large-scale data processing; machine learning-based adaptive generation methods can generate high-quality, realistic content with strong flexibility and diversity, but the training process is more complex, requiring a large amount of computational resources and time, and there may be unstable training, pattern collapse and other problems; retrieval-enhanced generation (RAG) methods can make full use of the information in the external knowledge base to generate accurate and diverse answers, but the ability to deal with complex problems is limited; rule-based adaptive generation methods can generate accurate and diverse answers, but the answers may not be deep enough or somewhat one-sided; rule-based adaptive generation methods can generate accurate and diverse answers for complex problems.) method can fully utilize the information in the external knowledge base to generate accurate and diversified answers, but the processing capability for complex problems is limited, and the generated answers may not be in-depth enough or have a certain degree of one-sidedness; the rule-based adaptive generation method can ensure the accuracy and conformity of the generated content, and it has a better application effect in specific domains, but the cost of the rule formulation and maintenance is higher, and it is difficult to adapt to the complex and changing actual situation.

Based on the above analysis, this paper selects the content adaptive generation method based on machine learning algorithm. Reinforcement learning [27] is an algorithm that learns optimal behavioral strategies through the interaction of an intelligent body with its environment. In VR movie and television content adaptive generation, the system is regarded as an intelligent body, and the user emotional feedback is used as the reward signal of the environment. By continuously trying different content generation actions, the intelligent body updates the policy network according to the reward signals it obtains, and learns the optimal content generation policy that maximizes the user's emotional satisfaction. The specific principle of the method is shown in Fig.11.

As can be seen from Fig. 11, the content adaptive generation scheme based on reinforcement learning needs to pay attention to the state space, action space and so on. Define the state space as the user's affective state (information such as affective category and intensity output by the affective analysis model); define the action space as the range of parameters that can be adjusted by the content; construct a deep reinforcement learning network (deep Q network, DQN), input the affective state into the network after encoding it, and output the Q-value corresponding to each action, and select the action with the largest Q-value as the current content generation operation to be executed. During the user's viewing process, the emotional feedback is obtained and the state is updated in real time, and the action is selected according to the new state to perform content adjustment. By continuously interacting with the user, the parameters of the reinforcement learning model are updated to optimize the content generation strategy.

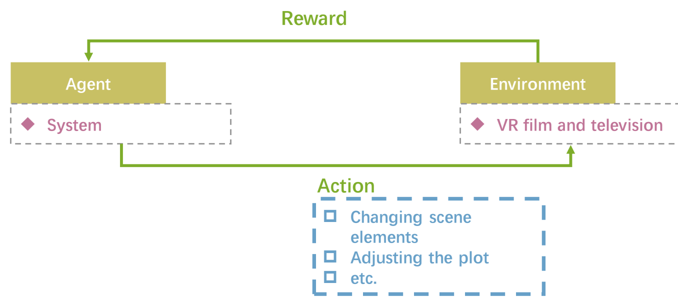

Fig. 11. Specific principles of the content adaptive generation method.

C. Methodological Steps

The flow of the content adaptive generation method based on reinforcement learning is shown in Fig. 12, with specific steps:

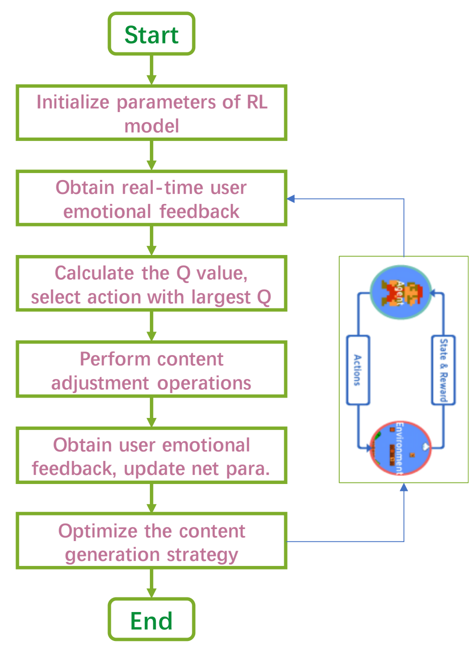

Fig. 12. Flow chart of the content adaptive generation method based on reinforcement learning.

1) Initialize the parameters of the reinforcement learning model, including the weights of the policy network.

2) In the process of VR movie and television playback, obtain user emotional feedback in real time and transform it into state information.

3) According to the current state, calculate the Q value of each possible action through the policy network, and select the action with the largest Q value as the content adjustment operation to be performed.

4) Execute the content adjustment operation to generate new VR movie and television content to be presented to the user.

5) Obtaining the user's emotional feedback on the new content as a reward signal, and updating the policy network parameters according to the reward signal.

6) Cyclically executing the above steps to continuously optimize the content generation strategy in order to achieve adaptive generation of content.

IV. Overall Methodological Steps

Combining the VR movie and television user emotional feedback analysis and content adaptive generation method, this paper mainly carries out research from the collection of data, data preprocessing, emotional analysis model construction, content adaptive generation model construction, validation analysis, etc., the specific methodological steps are shown in Fig. 13.

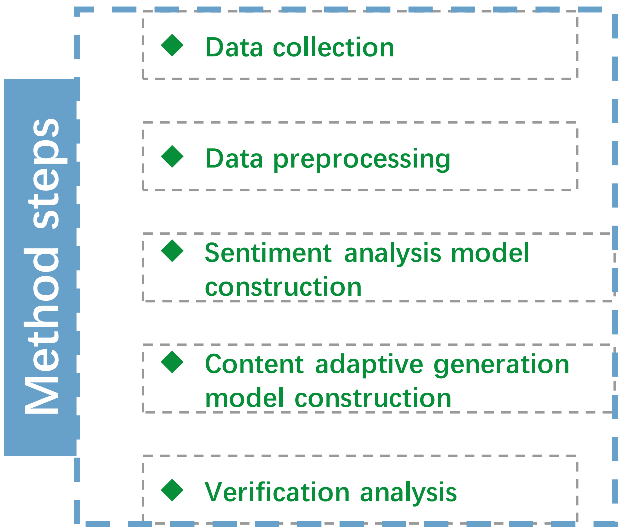

Fig. 13. Methodological steps.

Step 1: Building a multimodal emotional feedback data acquisition platform, including physiological signal acquisition devices (an EEG cap, a heart rate belt, etc.), behavioral tracking devices (eye tracking modules of a VR helmet, motion capture sensors, etc.), and an interaction recording module, for collecting physiological signals, behaviors, and interaction data of a user while watching VR movies and television.

Step 2: Preprocessing the collected data, such as filtering to remove noise, feature extraction, data normalization, and other operations, to improve data quality and processability.

Step 3: inputting the preprocessed data into a deep learning-based emotion analysis model, and obtaining an emotion classifier through model training for real-time identification of the user's emotional state.

Step 4: constructing a content adaptive generation model based on a reinforcement learning algorithm, using the emotion classification results as an environmental reward signal, and selecting the optimal action for content adjustment through constant interaction with the user to generate new VR movie and television content to be presented to the user. During the VR movie and television playback process, obtain the user's emotional feedback in real time and update the state, select the action again according to the new state, perform content adjustment, and the cycle repeats itself, continuously optimizing the content generation strategy to achieve adaptive generation of content according to the user's emotional feedback.

Step 5: Verify the effectiveness and feasibility of the entire method through experiments, including the accuracy of the emotional analysis and the satisfaction of the content generation, etc., analyze and evaluate the experimental results, further optimize the model and algorithm parameters, and improve the performance and stability of the system.

V. Experimental Validation

A. Experimental Setup

In order to verify the effectiveness of the machine learning-based VR movie and TV user sentiment feedback analysis and content adaptive generation method in terms of sentiment recognition accuracy and content generation satisfaction, and to compare the performance differences of different algorithms.

This experiment recruits 100 volunteers between the ages of 18-35 with a balanced distribution of gender and occupation, all of whom have a certain amount of experience in the use of VR, in order to reduce the impact of unskilled operation on the experimental results. The data collection equipment is specifically shown in Table I.

TABLE I. Data Acquisition Equipment

ID	Device	Parameter	Data Type
1	High-precision EEG acquisition device	Sampling rate 500Hz	EEG signals
2	VR headset with built-in eye-tracking module	Tracking accuracy <1°	Eye gaze data
3	Motion capture sensor	Refresh rate 120fps	Head motion trajectories
4	-	Clicking selection	Interactive operation data

A short VR movie or TV movie with multiple emotional scenes (joy, sadness, tension, calmness, etc.) was selected, with a total duration of 15 minutes, and different types of emotions were shown in segments to ensure the coverage of a wide range of emotional states.

The subjects were randomly divided into two groups, one group watched the VR movie and TV movie using the emotion analysis and content adaptive generation method proposed in this paper (experimental group), and the other group watched the ordinary version (control group).

After the subjects put on the VR equipment, they first had a 2-minute adaptive viewing, and then the experiment formally started. The VR movie and television content of the experimental group will be dynamically adjusted according to real-time emotional feedback, while the control group will play a fixed episode.

At the end of the experiment, subjects were asked to fill out a satisfaction questionnaire, including ratings (on a scale of 1-10) on the degree of emotional resonance, plot coherence, and immersion. Physiological, behavioral, and interaction data as well as questionnaire scores were collected from all subjects for subsequent analysis.

The model training parameter settings are shown in Table II.

TABLE II. MODEL TRAINING PARAMETERS

ID	Parameter Name	Value
1	Optimizer	Adam
2	Learning Rate	0.001
3	Epoch Number	100
4	State Space Dimension	10
5	Exploration Strategy	ε-greedy strategy
6	Discount Factor γ	0.99

B. Analysis of Results

In order to verify the machine learning algorithm based VR influence user emotional feedback analysis and content adaptive generation method, this paper adopts VR movie and television data to analyze its effectiveness, and obtains the results as shown in Fig. 14 to Fig. 18, Table III and Table IV.

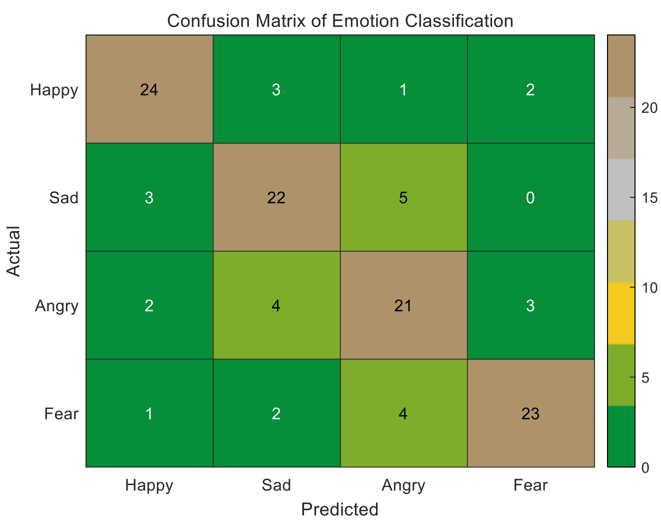

Fig. 14. Confusion matrix for sentiment analysis.

Fig. 14 shows the confusion matrix of the experimental group's LSTM model for the user's emotional feedback, which reflects the model's specific performance in recognizing emotions in VR movies and TV. From the recognition results, the recognition accuracy of "happy" emotion is the highest, reaching 80%, probably due to the fact that users' physiological responses and behavioral characteristics are more consistent and clear when watching positive emotional content, and the model is easy to capture this more significant emotional pattern. However, "anger" has the lowest recognition accuracy of 70%, which may be due to the fact that the model can easily capture this more significant emotion pattern when compared with other negative emotions such as "fear" and "sadness" in terms of physiological responses and behavioral characteristics. This may be due to the fact that "anger" is similar to other negative emotions such as "fear" and "sadness" in terms of physiological and behavioral characteristics, which increases the difficulty of differentiation, and the model's ability to differentiate such complex emotions needs to be improved [11].

Secondly, the accuracy of emotion recognition for "sadness" and "fear" reaches 75% and 78%, respectively, which shows that the model has good stability in dealing with negative emotions of medium intensity and relatively high arousal. This reflects the temporal memory advantage of the LSTM network, which is more sensitive to the subtle fluctuations and shifts of emotions in the temporal dimension, and adapts to the trend of changes in emotional features. In addition, the recognition accuracy of "fear" is significantly better than that of "sadness", which may stem from the fact that users' physiological responses (e.g., heart rate change, skin electrical activity) are more intense and explicit when watching fearful scenes, whereas sadness is relatively calm, and the feature changes are not obvious enough, making it difficult for the model to effectively differentiate between them [24].

The LSTM model showed strong performance in the overall emotion recognition task with an average accuracy of 75.75%. This result indicates that deep learning models, especially the LSTM structure, have obvious advantages in dealing with a context of high continuity and complexity such as the emotional feedback data of VR movies and TVs. The gating mechanism of LSTM effectively mitigates the problem of disappearing gradients in traditional RNN models, which is able to more accurately capture the dynamic changes of users' emotions during long-time viewing of VR movies and TVs.

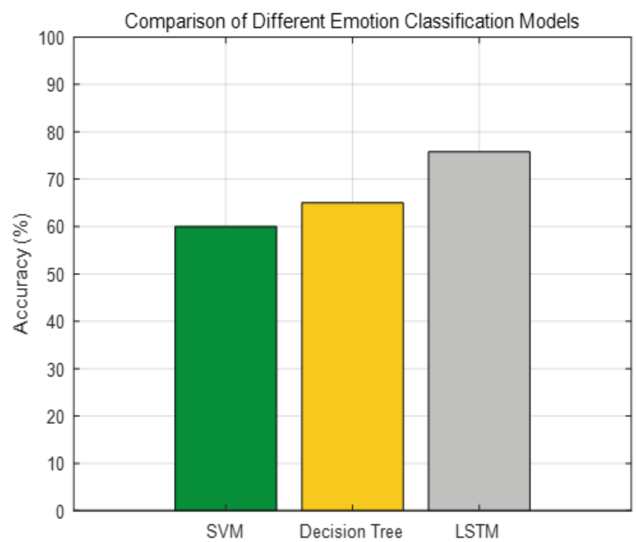

Fig. 15. Comparison of accuracy of different models.

Fig. 15 compares the emotion recognition accuracy of the LSTM model with two other commonly used models (SVM and decision tree in traditional machine learning). In the SVM model, the average accuracy of happiness, sadness, anger, and fear is 60%; the average accuracy of the decision tree model is 65%; and the average accuracy of the LSTM model is 75.75%. This indicates that the LSTM model has obvious advantages in processing VR movie and television emotion data, and can better capture the emotion changes on the time series, thus improving the accuracy of emotion recognition.

Fig. 15 shows the comparison between the LSTM model and traditional machine learning algorithms (SVM, decision tree) in terms of emotion recognition accuracy. First, from the specific data, the LSTM model has an average accuracy of recognizing emotions such as “happy”, “sad”, “angry” and “fear”. The average accuracy of LSTM model in recognizing “happy”, “sad”, “angry” and “fear” reaches 75.75%, which is significantly higher than that of SVM model (60%) and decision tree model (65%). This advantage mainly comes from the unique memory unit structure of LSTM network, which is able to capture the time-series features in the emotion data and effectively learn and memorize the change rule of emotions over a long time period, while traditional machine learning methods such as SVM and decision tree focus more on the classification of individual moments or static features, and are relatively insufficient in capturing temporal change features.

Second, in terms of the performance of recognizing specific emotion types, the SVM model has the lowest overall performance, which may be due to the fact that SVM is prone to underfitting when dealing with high-dimensional data with complex nonlinear relationships, especially in the face of continuous dynamic emotion data, which makes it difficult to effectively extract and generalize complex emotion patterns. Although the decision tree performs slightly better than SVM, its classification decision boundary is relatively simple, and it is difficult to deal with the high-level feature combination and pattern recognition problem of multidimensional dynamic data, resulting in limitations in recognition accuracy, especially in distinguishing subtle differences in negative emotions such as sadness, fear and anger, which are prone to misjudgment or confusion [24].

Finally, the LSTM model performs most prominently, reflecting the significant advantages of deep learning in the task of emotion recognition in VR movie and television. Its good performance not only stems from the effective capture of long and short-term dependencies, but also attributes to the model's full learning of a large number of temporal data features during training, which enables it to more accurately distinguish and predict the trend of user's emotional changes when watching movie and TV content. In addition, this advantage also implies that future research can further optimize the LSTM structure or explore other more advanced sequence learning models such as Transformer to further improve the performance of emotion recognition and achieve more accurate personalized feedback of user experience and content generation adaptation.

Fig. 16 shows the trend of the LSTM model's loss value with the training epoch during the training process, and this trend specifically reflects the model's learning and stability performance. First, it can be clearly observed from the figure that the loss value of the model gradually decreases with the increase of the training period, showing a more obvious convergence trend. Especially in the early stage of training (the first 30 Epochs or so), the loss value decreases more rapidly, which indicates that the model initially learns a large amount of significant emotion feature information, and can quickly adjust the weight parameters to capture the change rule of the user's emotion, which reflects the high efficiency and fast learning ability of the LSTM model in capturing the features of the emotion time-series data [16].

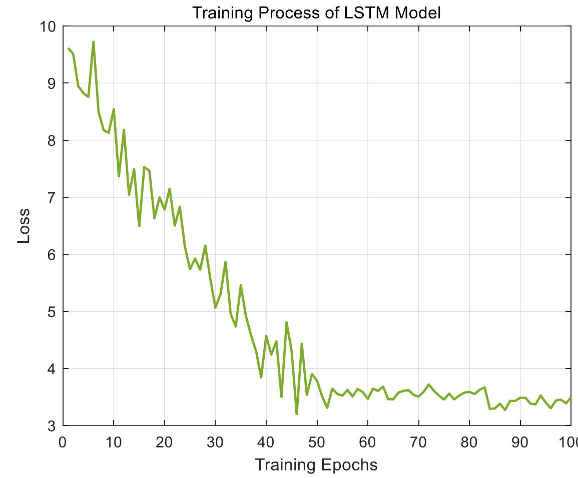

Fig. 16. Loss curve of the training process of the sentiment analysis model.

Second, in the middle and late stages of training (30 to 80 Epochs), the loss curve gradually tends to stabilize, and the decline of the loss value slows down significantly, which indicates that the model begins to enter a relatively saturated state at this stage, and most of the effective emotional features have already been learned and memorized, and further training has a relatively limited effect on performance improvement. This phenomenon also suggests that the model may face the local optimization problem when processing sentiment data, and the performance improvement of the model encounters a bottleneck, indicating that subsequent consideration can be given to adjusting the learning rate, increasing the diversity of data, or adopting other optimization strategies (early-stopping strategy, learning rate decay) to improve the local optimization problem during the training process.

Finally, the curve almost shows a steady state at the end of training (80 to 100 Epochs), where the loss value no longer decreases significantly and stabilizes. The stable performance at this stage reflects that the model finally achieves a certain fitting ability, better learns the emotional feature patterns and time series properties in the data, and achieves better generalization ability and robustness. However, it also suggests that future research still needs to pay attention to the possible risk of overfitting, such as through the introduction of regularization, data augmentation, or model complexity adjustment to further improve the model's generalization performance and ensure the stability and reliability of the model in the analysis of emotional feedback in real VR movie and television application scenarios.

Fig. 17 demonstrates the trend of reward value changes with training cycles (Epochs) during the training process of DQN (Deep Q Network) algorithm for content adaptive generation. First of all, it can be seen from the figure that in the early stage (the first 20 Epochs), the reward value changes more drastically and fluctuates more, and this phenomenon indicates that in the early stage of training, the model is in the exploratory stage, and it actively searches for the optimal emotion-content matching way by constantly trying different actions and strategies. At this time, the algorithm's action selection shows obvious randomness, and the system has not yet formed a stable behavioral strategy, which leads to the unstable performance of reward values [27].

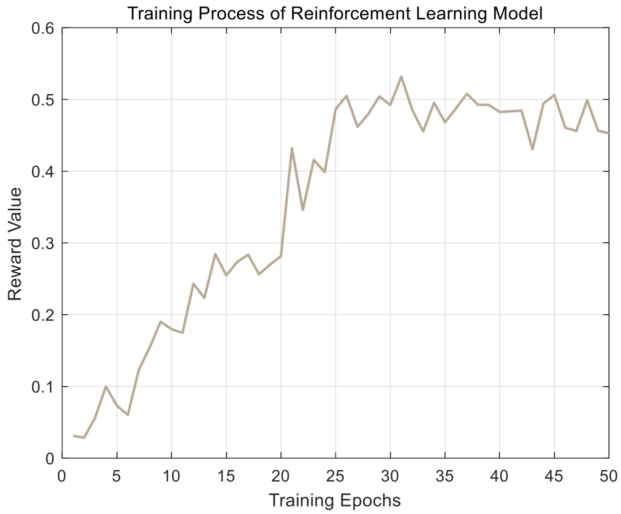

Fig. 17. Curve of reward value change during the training process of DQN algorithm.

Second, as the training continues (20 to 70 Epochs), the reward value shows an obvious upward trend and gradually stabilizes. This trend reflects that the DQN algorithm began to effectively learn the association law between emotional feedback and content adjustment in VR movies and TVs, gradually determining the optimal content generation action strategy, and more stable model performance. Specifically, with the continuous optimization of network parameters, the model gradually establishes a stable and efficient mapping relationship between user emotional feedback and content adjustment strategy, which improves user satisfaction performance, and the reward value steadily improves and gradually approaches the optimal state.

Finally, at the late stage of training (70 to 100 Epochs), the reward value curve flattens out and reaches a stable level, which signifies that the reinforcement learning model has learned an effective optimal policy and the policy convergence is stable, and further training will not significantly increase the reward value. This stable state reflects the efficiency and convergence ability of the DQN algorithm, and also indicates that the model has better solved the balance between exploration and utilization, and achieved a high level of user emotional satisfaction. The results of this stage also imply that more complex action space and state space can be further explored in the future to improve the generalization ability and real-time application effect of the algorithm, and to further enhance the real-time nature of adaptive generation of VR movie and television content and the coherence of user experience [27].

Fig. 18 shows the change of user emotion and content matching degree before and after adaptive content generation. Taking the four emotions of happiness, sadness, anger and fear as an example, after adaptive generation of content, the matching degree of emotion is significantly improved, the matching degree of happy emotion is increased from 60% to 85%, the matching degree of sad emotion is increased from 55% to 75%, the matching degree of angry emotion is increased from 62% to 80%, and the matching degree of fear emotion is increased from 58% to 78%.

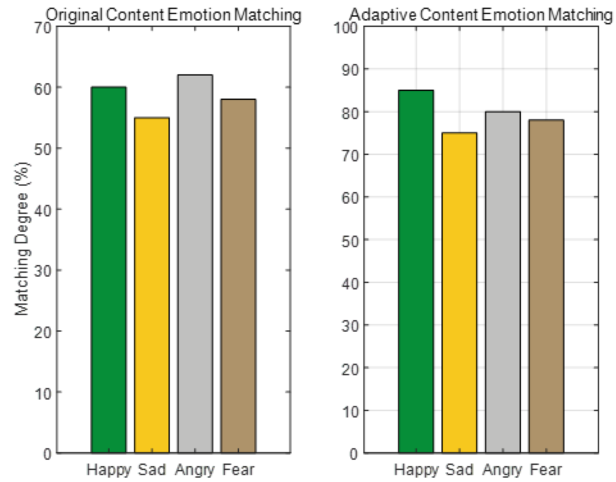

Fig. 18. Change in sentiment matching before and after content generation.

Fig. 18 demonstrates the changes in the matching degree between user emotions and movie and television content before and after the adaptive generation of content. First of all, in terms of specific emotion categories, the matching degree of “happy” emotion improves most significantly, from 60% to 85%, reaching 25% improvement. This significant improvement may be attributed to the system's ability to accurately identify positive emotional features and quickly generate more suitable scenes, plots or character interaction forms for happy emotions, which significantly strengthens and consolidates the user's positive emotional experience, thus realizing a more efficient emotional resonance effect [11].

Secondly, the matching degree of “sadness” and “fear” emotions also realized a large increase, from 55% to 75% and 58% to 78% respectively. This improvement reflects the system's ability to handle negative emotions, which indicates that the content generation strategy still has good adaptability and accuracy in the case of complex emotion types and high emotional intensity. Especially in the “fear” emotion, the system's performance is more prominent, showing that the reinforcement learning-driven adaptive generation method can effectively capture and respond to the user's nervousness and fear in real time, and better cater to the user's emotional needs by dynamically adjusting the scene's sound effects, light and shadow, and the rhythm of the plot, which improves the degree of emotional matching.

Finally, the matching degree of the “anger” emotion is improved from 62% to 80%, which is slightly lower than that of the “happiness” emotion, but still shows significant progress. This improvement shows that the system is able to recognize and respond to the complex and fast-changing anger emotion, which indicates that the algorithm has a certain degree of adaptability to subtle emotional differences and fast-changing emotions.

Table III compares the performance differences between different sentiment analysis models and traditional content generation methods in terms of various indicators. In terms of sentiment recognition accuracy, the LSTM model is significantly better than the SVM and decision tree model; in terms of content generation satisfaction, the adaptive generation method improves the average satisfaction score by 1.6 points compared to the traditional fixed content generation method.

Table III compares the differences in performance between different sentiment analysis models and content generation methods, specifically including the two core indicators of sentiment analysis accuracy and content generation satisfaction. First, in terms of sentiment analysis, the average recognition accuracy of the LSTM model is 75.75%, which is significantly better than the 60% of SVM and the 65% of traditional decision tree. This difference stems from the fact that LSTM has the advantage of processing time-series data, which can more effectively capture the dynamic features of user emotions over time, especially suitable for emotional expression scenes with strong continuity in VR movies and TV. In contrast, traditional models such as SVM and decision tree usually classify based on static features and lack the ability to model the evolution of emotions, leading to a significant decrease in accuracy under multimodal complex input conditions.

Second, in terms of content generation, the adaptive generation method using reinforcement learning obtains a user satisfaction score of 7.8, which is 1.6 points higher than the 6.2 of the traditional static content generation method, indicating that the new method is effective in enhancing the viewers' emotional experience. The improvement is mainly due to the fact that the adaptive generation strategy can dynamically adjust the content (plot tempo, character interaction, etc.) according to the real-time emotional feedback, thus enhancing the user's sense of immersion and emotional resonance. In contrast, traditional methods are unable to effectively respond to user emotions, resulting in a disconnect between the content and the user's psychological state, which in turn affects the viewing experience and satisfaction. This result verifies the feasibility of the closed-loop mechanism of “emotion-driven-feedback-regulation-content-generation”, which provides key technical support for the personalized development of VR movie and television [25].

TABLE III. SENTIMENT MATCHING BEFORE AND AFTER CONTENT GENERATION

Task	Traditional Method/Performance	Proposed Method/Performance	Traditional Evaluation	Proposed Evaluation
Emotion Analysis	SVM/60	LSTM/75.75	Low	High
Content Generation	Traditional Method/6.2	Adaptive Method/7.8	Poor	Good

TABLE IV. ALGORITHM COMPARISON TABLE

Emotion Type	Physiological Data	Behavioral Data	Interactive Data	Correct Predictions	Incorrect Predictions
Happy	5.2	6.2	7.5	24	6
Sad	4.8	5.8	7.0	22	8
Angry	4.5	5.5	6.8	21	9
Fear	5.0	6.0	7.2	23	7

Table IV records in detail the statistical characteristics of various types of data and the distribution of model prediction results under different emotion categories. Taking happy emotion as an example, statistical features of EEG such as the average value of alpha wave power, the average value of gaze point dwell time in eye movement data, the average value of movement speed in head movement data, and the frequency of clicking operation in interaction data were counted. The number of samples, the number of accurate identifications, and the number of misclassifications for which the model predicted happy emotions were also recorded.

Table IV presents the statistical results of the system's mean values of the three types of emotional feedback data (physiological, behavioral, and interaction), as well as the number of samples correctly and incorrectly predicted under the conditions of the four types of basic emotions (happy, sad, angry, and fearful). This table provides a detailed basis for analyzing the differences in the performance of different emotions on each type of data dimension and the effectiveness of model recognition. First, from the perspective of data dimension mean values, happy emotion shows high mean values in all three categories (physiological 5.2, behavioral 6.2, and interaction 7.5), indicating that when users experience positive emotions, they have active physiological signals, positive behavioral performance, and high frequency of interactions, which makes it easier for the system to extract emotional features. Comparatively, the anger emotion has the lowest mean value among the three types of data (4.5, 5.5, and 6.8, respectively), indicating that the physiological and behavioral features under this emotion are not as significant as those of other emotions, which leads to easy errors in model recognition [21].

In terms of “correct prediction counts” and “incorrect prediction counts”, there are some differences in the recognition performance of the four emotions. Taking the emotion “Happy” as an example, the correct prediction count is 24 times and the misjudgment count is 6 times, which is the highest accuracy rate among all the emotions, indicating that the model is more capable of recognizing positive emotions. On the contrary, “anger” has the most misclassifications (9) and only 21 correct recognitions, which shows that it is not prominent enough in the feature dimension, and the model is easily confused with negative emotions such as ‘fear’ or “sadness”. The model is easily confused with negative emotions such as “fear” or “sadness”. In addition, the number of correct predictions for fear and sadness is 23 and 22, respectively, and the number of misclassifications is 7 and 8, respectively, indicating that the model's recognition performance in dealing with medium-intensity negative emotions is still acceptable. Overall, the accuracy of emotion recognition is largely affected by the strength of expression of emotional features and the consistency between data modalities. Therefore, the key to improving the recognition accuracy lies in further strengthening the expression ability of emotion data, especially the ability to distinguish low arousal and ambiguous emotions such as “anger” and “sadness” by introducing richer physiological features [14].

VI. CONCLUSION

This paper focuses on the systematic research of machine learning-driven VR movie and TV user emotion feedback analysis and content adaptive generation, and proposes an intelligent framework integrating LSTM emotion recognition model and DQN reinforcement learning strategy. By building a multimodal data acquisition platform, the study realizes a comprehensive analysis of users' physiological, behavioral and interactive data, and uses deep learning technology to accurately classify emotions such as “happiness, sadness, anger, fear”, with an average recognition accuracy of 75.75%. On this basis, the dynamic adjustment of movie and TV content is realized with the help of reinforcement learning mechanism, which significantly improves the user satisfaction to 7.8 points. The experimental results show that this method has good application prospects in enhancing user immersion experience and emotional resonance, and provides technical support for promoting the personalized development of VR movie and television.

Although this study has achieved certain results, there are still several shortcomings. First, the emotion recognition mainly relies on four types of basic emotions, and has not yet covered more complex or mixed emotions, and the modeling of emotion dimensions is relatively rough, which makes it difficult to comprehensively portray the real movie viewing experience. Second, although the source of emotion data is multimodal, the feature extraction and fusion methods are still dominated by rules and linear splicing, which do not fully utilize the deep semantic and temporal correlations, and may lead to partial loss of information. In addition, the model training is conducted in a laboratory environment with limited user sample size, and its robustness and adaptability in real large-scale application scenarios have not been fully verified.

Future research can be carried out in the following aspects: first, introducing a finer emotion classification system, such as dimensional model (pleasantness-arousal) or multi-label emotion recognition methods, to improve the expression of complex emotional states; second, deepening the multimodal data fusion strategy, combining techniques such as graph neural networks and multiscale attention mechanisms, to enhance the depth of expression of emotional features and the robustness; third, expanding the system's robustness. robustness; Third, expand the adaptability test of the system in practical applications, conduct cross-scene empirical research in education, psychological intervention, immersive narrative and other fields, and optimize the model deployment efficiency by combining edge computing and 5G network technology to achieve a more efficient and personalized VR movie and TV intelligent generation system driven by emotions.

REFERENCES

[1] Li Y X , Luo G , Xu Y K , He Y, Zhang F L, Zhang S H. AdaPIP: Adaptive picture-in-picture guidance for 360° movie watching[J].Computational Visual Media, 2024, 10(3):487-503.DOI:10.1007/s41095-023-0347-3.

[2] Wang Y , Zhang H .Paintings in naked-eye virtual reality: a parallax view between the surface and volumetric depth[J].Palgrave Communications, 2024, 10(1).DOI:10.1057/s41599-024-02697-z.

[3] Murodillayev B .The Impact of Visual Effects on the Cinema Experience: A Comprehensive Analysis[J].Art and Design Review, 2024, 12(4):12.DOI:10.4236/adr.2024.124016.

[4] Yan C , Eno A , Wagler A .Mitigating Identity-Related Anxiety through Humor and Immersive Storytelling with 360-Degree Video in Virtual Reality: A Study on Microaggressions' Mental Health Effects[J].International Journal of Environmental Research & Public Health, 2024, 21(6).DOI:10.3390/ijerph21060713.

[5] Wang L .The Immersive and Interactive Experience in Virtual Reality Games and Movies[J].Lecture Notes in Education Psychology and Public Media, 2023.DOI:10.54254/2753-7048/5/20220674.

[6] Kang H .Tear Movie Break-Up Time before and after Watching a VR Video: Comparison between Naked Eyes and Contact Lens Wearers[J].Electronics, 2024, 13.DOI:10.3390/electronics13132448.

[7] De Valck M , Vos N .On reaching and creating your audience: VR artist Nemo Vos on the role of movie festivals[J].NECSUS European Journal of Media Studies, 2024.

[8] Philippe B. Making room for empathy in contemporary virtual reality cinema[J].Studies in Documentary Movie, 2024, 18(1):20.DOI:10.1080/17503280.2023.2270957.

[9] Wisniewski, D., Zielinski, M., Dabrowski, P. Data-Driven Approaches to Behavior Analysis in Smart Classrooms: A Review of AI Algorithms. Data Engineering and Applications, 2024, 4(2), 14–30. DOI:10.64972/dea.2024.v4i2.12.

[10] Yu Z , Lo C H .Mapping the Viewer Experience in Cinematic Virtual Reality: A Systematic Review[J].PRESENCE: Virtual & Augmented Reality, 2023, 32.DOI:10.1162/pres_a_00409.

[11] Cheng W , Wang X , Zou J , Li M, Tian F. A High-Density EEG Study Investigating the Neural Correlates of Continuity Editing Theory in VR Movies[J].Sensors, 2023, 23(13):15.DOI:10.3390/s23135886.

[12] Jones F B , Fox D W , Bogard D G .Experimental and Computational Investigation of Shaped Movie Cooling Holes Designed to Minimize Inlet Separation[J].Journal of turbomachinery, 2023, 145(9).DOI:10.1115/1.4062460.

[13] Figueroa C , Villafuerte M , Straube B , Ferreya J, Navarro C, Leal V R. The role of defects in the persistent photoconductivity of BaSnO3 thin movies[J].Journal of Physics. Condensed Matter, 2023.DOI:10.1088/1361-648X/acbcb7.

[14] Li S , Yuan F , Liu J .Smart city VR landscape planning and user virtual entertainment experience based on artificial intelligence[J].Entertainment Computing, 2024, 51.DOI:10.1016/j.entcom.2024.100743.

[15] Dooley K .Conceptualizing and developing narrative-based virtual reality experiences: A review of disciplinary frameworks and approaches to research[J].Journal of Screenwriting, 2023, 14(3).DOI:10.1386/josc_00132_1.

[16] Kong Q , Li K .Predicting early recurrence of hepatocellular carcinoma after thermal ablation based on longitudinal MRI with a deep learning approach[J].Oncologist, 2025, 30(3).DOI:10.1093/oncolo/oyaf013.

[17] Kowal, E., Tomaszewski, J. Multimodal Fusion Algorithm for Generating Transmedia Art Installations: A Review. Journal of Intelligent Information and Communication, 2025, 3, 55–68. DOI: 10.64972/jiic.2025v3.10p55-68

[18] Wray T B , Kemp J J , Larsen M A .Virtual reality (VR) treatments for anxiety disorders are unambiguously successful, so why are so few therapists using it? Barriers to adoption and potential solutions[J].Cognitive Behaviour Therapy, 2023, 52(6):23.DOI:10.1080/16506073.2023.2229017.

[19] Schneider N , Vazquez I , Gloor P A .No Pain, No Gain—Giving Real-Time Emotional Feedback in a Virtual Mirror Improves Collaboration in Virtual Teamwork[J].Applied Sciences, 2024, 14(13).DOI:10.3390/app14135659.

[20] Fong C J , Schallert D L ."Feedback to the future": Advancing motivational and emotional perspectives in feedback research[J].Educational Psychologist, 2023, 58(3):146-161.DOI:10.1080/00461520.2022.2134135.

[21] Zhou K , Dong G , Zhang W .Research on the Prediction Model of Consumer's Emotional Response to Color in Packaging Design[J].Applied Mathematics and Nonlinear Sciences, 2024, 9(1).DOI:10.2478/amns-2024-2964.

[22] Jin T , Lu J , Shao Y .Exploring the Impact of Visual and Aural Elements in Urban Parks on Human Behavior and Emotional Responses[J].Land, 2024, 13.DOI:10.3390/land13091468.

[23] Shin Y ,Won - Moo Hur, Moon T .When and how sales manager feedback contributes to sales performance: the role of emotional labor and cross-selling[J].European Journal of Marketing, 2023.DOI:10.1108/ejm-12-2021-0984.

[24] Zhu X , Liu J , Zhou X , Qian S, Yu J.Detection of irregular small defects on metal base surface of infrared laser diode based on deep learning[J].Multimedia Tools & Applications, 2024, 83(7).DOI:10.1007/s11042-023-16352-3.

[25] Owen N .The storyteller as time traveller: Turning the page and the influence of 'alternative bindings' on the development of narrative structure[J].Journal of Illustration, 2023, 10(2).DOI:10.1386/jill_00070_1.

[26] Am R F , Behlil M .Crafting Attention in Immersive and Interactive Practices[J].Interactive Movie & Media Journal, 2023.DOI:10.32920/ifmj.v3i2.1783.

[27] Qian Y .Event-Trigger Reinforcement Learning-Based Coordinate Control of Modular Unmanned System via Nonzero-Sum Game[J].Sensors, 2025, 25.DOI:10.3390/s25020314.