Jiang, Y. (2022) Emotion-Driven Training Using Emotion Recognition and Reinforcement Learning in a College Gymnastics Program. Revista Internacional de Medicina y Ciencias de la Actividad Física y el Deporte vol. 22 (85) pp. 303-317
DOI: https://doi.org/10.15366/rimcafd2022.85.019

ORIGINAL

Emotion-Driven Training Using Emotion Recognition and Reinforcement Learning in a College Gymnastics Program

Yu Jiang
[1] College of Sports and Leisure, Guangdong Ocean University, Zhanjiang, Guangdong, 524000, China
Email: jiangyu0529@126.com

UNESCO Code / UNESCO Code:
Council of Europe classification / Council of Europe classification:

Recibido 23 de enero de 2021 **Received** January 23, 2021
Aceptado 23 de enero de 2022 **Accepted** Januray 23, 2022

Abstract

In the process of college students' gymnastics learning, the phenomena of weakened motivation and emotional instability have seriously affected the learning effect of current gymnastics motor skill learning. The change of learning motivation is an internal factor for continuous learning of sports skills, and the generation and weakening of individual frustration can lead to the change of motivation, thus prompting college students to undergo a reversal of motivation, so that they can maintain the motivation for continuous training and learning. Currently, emotion-driven training is an effective method. In this paper, we utilize emotion recognition and reinforcement learning to drive the training of college students in gymnastics programs, with a view to reducing college students' frustration and emotional fluctuations in the training process. Specifically, we propose an emotion semantic recognition algorithm based on reinforcement learning, and the reinforcement learning model is able to recognize the complex emotional features in college students' speech and extract the emotion semantics. Among them, the multimodal semantic extraction model can fully extract the multimodal features among college students' speech, the emotion attention enhancement mechanism can guide the optimization of the semantic feature extraction algorithm of the multimodal semantic extraction model by effectively tracking and feeding back the results of the emotion analysis, and the general extended dictionary can pay more attention to the content that affects the emotion category. With the collaborative training method, the reinforcement learning-based emotion semantic recognition model effectively comprehends the emotion information of college athletes, which provides a positive reference for the emotion-driven training method.

Keywords. Emotion-Driven Training, Emotion Recognition, Reinforcement Learning, College Gymnastics Program

1. INTRODUCTION

In the process of learning and teaching gymnastics motor skills, college students are prone to diminished motivation to participate and inability to persist. This weakened level of motivation and withdrawal from learning seriously affects the learning effect of their gymnastic sport skills (Yüvrük, Kapucu, & Amado, 2020). In order to overcome the effects of weakened motivation and inability to persist on gymnastics motor skill learning, scholars have conducted a great deal of research from multiple perspectives. From the field of motor skill learning, the motor skill learning effect is largely influenced by intrinsic factors such as motivation and emotion (Azemi, Ozuem, & Howell, 2020; Chan & Sun, 2021). In order to improve the effect of motor skill learning, teachers should pay attention to the intrinsic motivation of students' learning in the teaching process (Chowdary, Nguyen, & Hemanth, 2021; Fathi, Greenier, & Derakhshan, 2021). The transformation of individual college students' motivation will put them in different emotional experiences, and when the individual's emotions change will react on the motivation, prompting a reversal of motivation (Ford & Troy, 2019). The motivational function of emotions suggests that emotions can be used as a motivation to continuously stimulate the production of individual behaviors. When an individual's motivational needs are not satisfied in objective things, negative emotions (frustration (Hajian, 2019)) that impede the efficiency of learning work will be generated, which acts as a condition for the reversal of motivation (Ishiwatari, Yasuda, Miyazaki, & Goto, 2020; Jeon, Yoon, & Yang, 2022). Studies have shown that changes in frustration affect the transformation of motivational states and the ensuing behavioral differences (Khalil et al., 2019; Korpershoek, Canrinus, Fokkens-Bruinsma, & de Boer, 2020) will affect the individual's motor skill learning outcomes (Yang et al., 2022).

In the teaching practice of gymnastics skill learning, people often experience frustration due to cognitive bias, which acts in their motivational changes. In cognition-related research, cognitive learning theory and cognitive evaluation theory both focus on the direct role of cognitive features on motivation, and both theories emphasize the intrinsic mechanisms by which learning occurs and a person's sense of interest, control, and initiative, respectively. In addition to its ability to act directly on individuals, cognition can also influence individual motivation through irrational perceptions in the learning process. Studies have shown that people's irrational perception of frustration is the real root cause of emotional deterioration and diminished motivation (Magarini et al., 2021; Maulud, Zeebaree, Jacksi, Sadeeq, & Sharif, 2021). Such irrational perceptions have been referred to by scholars as irrational beliefs (Miranda-Mendizabal et al., 2019; Moroń & Biolik-Moroń, 2021). In the field of psychology, exploring ways to ameliorate irrationality has been a more popular topic. Scholar Ellis believes that if rational thinking can replace irrational thinking with a maximum of irrational thinking in understanding things, it can greatly reduce people's negative emotions (Mukhopadhyay et al., 2020; Weiner, 1982).

The expressions of university students are complex and varied. "I like you" and "I like you very much" do not express the same emotional intensity, and "I like watching television" and "I like watching television, but I don't like watching cartoons" also express two different feelings. Another example is that "I like watching television" and "I like watching television, but I don't like watching cartoons" also express two different feelings. Another example is "I am very happy today!" which expresses a stronger emotion than "I am very happy today." Another example is "I am very happy today", which expresses a stronger emotion than "I am very happy today". These examples illustrate the problem of simplicity. They illustrate the problem that simple punctuation marks, negatives, and adverbs of degree can seriously affect and change the category of feelings. In past studies, researchers often labeled each word with a common lexicon of nouns, adverbs, etc., and then determined the emotional category of the text by counting the number of positive and negative labels. With the diversification of people's expressions and the emergence of more and more emerging words and phrases on the Internet, this kind of research method obviously loses its accuracy. Before the emergence of deep learning (Schoneveld, Othmani, & Abdelkawy, 2021; Smith et al., 2021), researchers usually used some special structures or symbols to express textual meaning. The emergence of deep learning has provided the field of natural language processing (Richey et al., 2019; Ryan, Deci, Vansteenkiste, & Soenens, 2021) with a new solution, which is to express the semantic structure of a text in terms of vectors.

However, in the field of reinforcement learning, researchers represent the semantics of a text by using a vector composed of feature values, where each dimension of the vector represents a feature. However, at present, there is no best feature representation scheme to express the semantics of text emotion, and not enough attention is paid to the semantics of emotion. Therefore, this paper proposes a reinforcement learning-based emotion semantic recognition algorithm, and the reinforcement learning model can recognize the complex emotional features in college students' speech and extract the emotion semantics. Among them, the multimodal semantic extraction model can fully extract the multimodal features of college students' speech, the affective attention enhancement mechanism guides the optimization of the semantic feature extraction algorithm of the multimodal semantic extraction model by effectively tracking and feeding back the results of the affective analysis, and the general extended dictionary can pay more attention to the content that affects the affective categories. With the collaborative training method, the reinforcement learning-based emotion semantic recognition model effectively comprehends the emotion information of college athletes, which provides a positive reference for the emotion-driven training method. The main contributions of this paper are as follows:

(1) The algorithm proposed in this paper is able to obtain a better expression of college students' emotional characteristics and generate accurate emotional polarity categories..

(2) In this paper, we propose an emotion semantic extraction algorithm based on reinforcement learning that can significantly improve the problem of

insufficient feature extraction and can provide a positive reference for emotion-driven training methods..

2. Related Works

2.1 Emotion Dictionary

The main purpose of constructing an emotional lexicon is to determine the emotional polarity of words in a text. Currently, there are three mainstream ways of expressing emotional polarity: emotional category, emotional score, and emotional distribution. Emotional category is to categorize words into positive and negative emotional polarity; emotional score is to convert each word into an emotional score, where a positive number between 0 and 1 indicates positive emotion and a negative number between 0 and -1 indicates negative emotion; and emotional distribution is to express the emotional category of each word as a real number.

The construction of these three emotion lexicons can be viewed as a process of learning mapping, i.e., mapping each word to an emotion category. In the existing research, the main emotion category lexicons are MPQA and GI, which are based on words and phrases. Among them, GI uses lexemes to differentiate the emotional category of a word under different lexemes on the basis of lexemes, which is used to accurately locate the emotional inclination of a word. The mapping function for constructing a lexicon of sentiment categories is as follows, and the range of sentiment for each word is {positive, neutral, negative}:

$$C = \begin{cases} positive & + \\ neutral & | \\ negative & - \end{cases} \qquad (1)$$

This type of lexicon can only represent positive and negative emotion labels, and it cannot represent the strength of emotional inclination of the words. The emotion scoring lexicon is based on the probability distribution of emotion categories (e.g., SentimentNet), which represents each word as the probability of different emotion states, the positive probability of positive emotion is larger than the negative probability, and vice versa. The mapping function for constructing the emotion scoring dictionary is:

$$C = \begin{cases} P(positive) & 0 \leq P(positive) \leq 1 \\ P(neutral) & 0 \leq P(neutral) \leq 1 \\ P(negative) & 0 \leq P(negative) \leq 1 \end{cases} \qquad (2)$$

where $P(negative)$ denotes the probability that a college student will speak with negative emotion, and by comparing the three probability sizes of words, the emotional label of each word is obtained. In addition to this, the emotional probability of a word can be used to determine the strength of the emotional inclination of the word. If $P(positive) - P(negative) > 0$, then the word is positively emotional, and the larger the absolute value of the difference, the stronger the emotional tendency. If $P(negative) - P(neutral) > 0$, the word is negatively emotional, and the absolute value of the difference is proportional to the strength of the emotional tendency.

An affective distribution lexicon uses a real number to represent the emotional inclination of a college student's spoken language, with positive and negative emotions indicated by symbols that judge the value of the affective category, and the intensity of the emotion indicated by absolute values. Unlike the second emotion score lexicon, the emotion distribution lexicon takes values in real space, while the emotion score lexicon can only take values in the range of -1 to 1 under the constraints of probability. The sentiment distribution lexicon is constructed by using the following mapping function:

$$C = \begin{cases} R^{+} & positive \\ R^{-} & negative \end{cases} \qquad (3)$$

where R^{+} is a positive real number, which indicates that the word is positively emotional. If R^{+} is larger, the greater the emotional intensity of the words; on the contrary, R^{-} is a negative real number, indicating that the words are negatively emotional. If R^{-} is larger, the greater the emotional strength of the words. If the size of the value of the emotion category is inclined to 0, then the word is neutral emotion.

2.2 Reinforcement Learning

Reinforcement learning is another type of machine learning. The idea of reinforcement learning is derived from behavioral psychology, in which an intelligent body learns through continuous experimentation and "trial and error". Reinforcement learning allows an intelligence to continuously interact with its environment and receive feedback signals from the environment, i.e., punishments or rewards. In a given situation, rewarded behaviors will be strengthened, while punished behaviors will be weakened, and the intelligent body continuously adjusts its strategy according to the feedback signals in order to achieve the optimal decision (as shown in Figure 1).

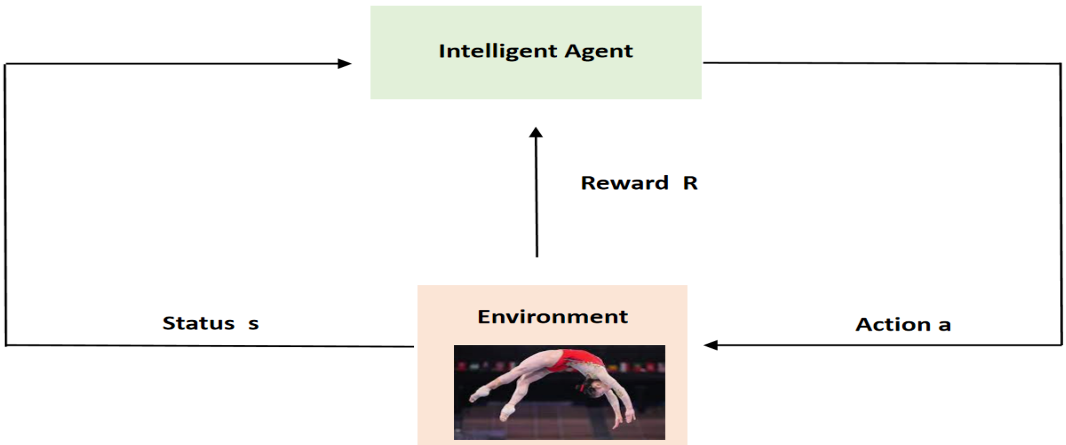

Figure 1. Schematic of the workflow of reinforcement learning

The intelligent agent perceives the current state s from the environment and then selects an action a from the action space, the environment generates the next state s' based on the perceived action and feeds a reward

R to the intelligent agent, which adjusts its own strategy to perform a new action through the reward R it receives, the optimal sequence of actions is determined by the rewards provided by the environment. The goal of the intelligent is to learn the strategy π^* that maximizes the cumulative reward, i.e., given a state, choose an action that maximizes the expected reward in the environment.Reinforcement learning can be described as a Markov Decision Process (MDP), which can be simply denoted as $M =< S, A, P_{s,a}, R >$, where S denotes the set of states, A denotes the set of actions, $P_{s,a}$ denotes the state transfer probability, i.e., the distribution of probabilities of performing an action $a \in A$ and transferring to the next state in the current state $s \in S$, and R denotes the reward function. Markov decision making is characterized by the fact that the future state is only related to the current state and not to the past state. The functional expression for the intelligence to find the optimal strategy π^* to maximize the return is shown below:

$$\pi^* = arg\max_{\pi} E[\sum_{t=0}^{T-1} \gamma^t r_{t+1} | \pi] \tag{4}$$

where π^* denotes the strategy acquired by the intelligent through learning, r_{t+1} denotes the reward obtained by the intelligent at time $t + 1$, γ denotes the discount factor with the value range of [0,1], and the smaller γ denotes the more importance attached to the immediate reward.

3. Methodology

Semantic extraction is a fundamental problem in the field of natural language processing and is often studied together with sentiment analysis algorithms. How to fully understand the emotional semantic information of college gymnasts, how to accurately extract the emotional semantic features and how to effectively mine the multimodal semantic content are three urgent problems that existing semantic extraction algorithms need to solve. The existing algorithms neglect the contextual relationship between historical and current texts when extracting textual features, resulting in insufficient semantic information; due to the lack of tracking and feedback on the analysis results, resulting in the inability to self-optimize the semantic extraction algorithms, so that some key emotional semantics have been neglected; due to the extraction of only lexical features, and the lack of other modal emotional features, resulting in the imperfection of semantic extraction, the results of semantic extraction are not satisfactory. The results of semantic extraction are not satisfactory. In order to solve the above problems, this paper proposes an emotion semantic extraction algorithm (MS-SAE) based on intensive learning. The algorithm consists of a multimodal semantic extraction model, an affective attention enhancement mechanism and an extended lexicon. The multimodal semantic extraction model can extract semantic features of the text from multiple perspectives during semantic extraction, and pays full attention to the extracted features to solve the problem of insufficient semantic extraction. In order to solve the problem of neglected emotional semantics, we constructed an emotional attention enhancement mechanism next to the multimodal semantic extraction model to provide real-time feedback on the analysis results. In order to solve the problem of a single perspective of semantic extraction, we constructed a generalized extended lexicon

(ExWordNet) to support the semantic extraction algorithm. In order to test the effectiveness of our proposed sentiment extraction algorithm, MS-SAE is combined with a sentiment analysis model based on LSTM to conduct a joint sentiment analysis experiment. The experimental results of large-scale sentiment analysis on three datasets, namely, SST-2, MR, and Subj, show that MS-SAE can solve the problems of poor semantic understanding of the text, errors in the analysis results, and significantly improve the accuracy of the students' sentiment recognition.

3.1 Reinforcement learning based emotion recognition algorithm

As shown in Figure 2, this paper proposes a reinforcement learning-based sentiment recognition algorithm (MS-SAE), which includes a multimodal semantic extraction model, an emotion attention enhancement mechanism and an extended lexicon (ExWordNet). In order to validate the semantic extraction effect of MS-SAE, we introduce a sentiment analysis model based on LSTM and conduct a joint sentiment analysis experiment. In the semantic extraction process, a multimodal semantic extraction model and a sentiment attention enhancement mechanism are proposed for extracting multimodal text sentiment feature vectors and improving the model's attention to the sentiment semantics, respectively. In addition, an extended lexicon is proposed to support the multimodal semantic extraction model to extract text features, so that the model can focus on multi-dimensional features such as lexical features, affective features, punctuation features, etc., and the semantic meaning of the extracted text will be more complete. In the process of sentiment analysis, we introduced the LSTM model to ensure the accuracy of sentiment analysis.

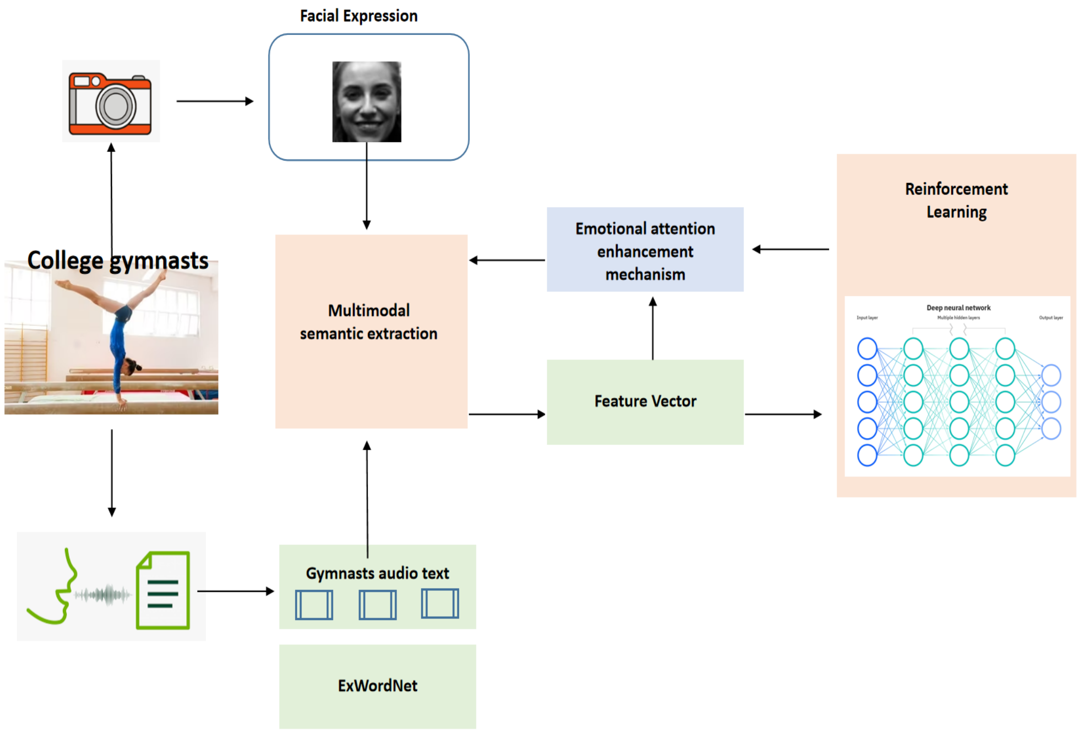

Figure 2. Schematic diagram of a reinforcement learning-based emotion recognition algorithm for college gymnasts

3.2 ExWordNet

We constructed a generic extended lexicon (ExWordNet) to support multimodal semantic extraction models, which consists of word and word embedding vectors. The traditional lexicon contains only common words such as nouns and adjectives, so only lexical features of the text can be obtained, which leads to insufficient attention to affective features .ExWordNet consists of five dimensions: common lexical, affective, adverbial of degree, negation and punctuation. The extended lexicon selects common sentiment words and sentiment scores (positive: 1, negative: -1) from SentiWordNet. Adverbs of degree strengthen/weaken the degree of emotion of emotion words, e.g. the emotion of "I quite agree with you" is obviously stronger than that of "I agree with you".

The extended lexicon selects 6 levels of adverbs of degree (defined as 0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3 depending on the intensity) from the web knowledge base. Negative words are able to change the category of emotions directly, so the lexicon collects common negative words (yes: 0.5, no: 0). The presence of punctuation can strengthen and weaken the affective categories of the text, e.g. the exclamation mark strengthens the tone of voice, so the introduction of punctuation enhances the accuracy of the extraction of semantic features (exclamation mark: 0.01, question mark: 0.01, no: 0).

3.3 Multimodal Semantic Extraction Model

The multimodal semantic extraction model is an important component of the proposed method, which can select semantic features based on their extraction, optimize and update the parameters of its own algorithm based on the feedback from the affective attention enhancement mechanism, and give the most reasonable semantic feature extraction results based on the information of the current text environment. At the same time, the extended lexicon provides support for the multistate semantic extraction model, so that it can give the complete word embedding vector and output it to the sentiment analysis model. The multimodal semantic extraction model adopts the logistic function $\pi_{\Theta}(o_i, e_i)$ to extract features, and defines it as the feature representation probability, $P_{\Theta}(o_i|e_i)$ denotes the probability of choosing o_i under the condition that the environmental information is e_i. In order to obtain accurate results, the model calculates the delayed feedback reward on a sentence-by-sentence basis. In order to obtain accurate probability results, the model calculates the delayed feedback reward on a sentence-by-sentence basis, i.e., once the word embedding vector of the whole sentence is obtained, $P_{\Theta}(o_i|e_i)$ is calculated.

$$\begin{aligned}\pi_{\Theta}(o_i, e_i) &= P_{\Theta}(o_i|e_i) \\ &= o_i\sigma(W * e_i + b) \\ &+(1-o_i)(1-\sigma(W * e_i + b))\end{aligned} \tag{5}$$

where σ denotes the $sigmoid$ function. It is worth noting that the following equations should be used to obtain the optimal choice of o_i' during testing.

$$o_i' = argmax\, \pi_\Theta\,(o_i, e_i) \qquad (6)$$

3.4 Emotional attention Strengthening mechanisms

Emotional Attention Enhancement Mechanism for College Gymnasts By constantly tracking and feeding back the results of emotional analysis, the multimodal semantic extraction model is guided to optimize the parameters of the algorithm, so as to make the algorithm more suitable for the current environment.

By paying more attention to the semantics that affect the sentiment of the text, the accuracy of sentiment analysis is greatly improved. Obviously, the affective attention enhancement mechanism is suitable for any sentiment analysis task.

The accuracy of categorization effectively reflects the strength of affective attention, as shown in the following equation, the affective attention enhancement mechanism calculates the feedback reward to the multimodal semantic extraction model based on the probability $P(C/S)$ that a sentence S is classified into category C (correctly categorized). Since the sentiment analysis task works on the whole text, $P(C/S)$ should be calculated for each sentence that has been analyzed by the sentiment analysis model.

$$R = \log P(C|S) \qquad (7)$$

The above mechanism for enhancing college gymnasts' affective attention evaluates the effectiveness of all the choices made by the multimodal semantic extraction model, which monitors the selection of the multimodal semantic extraction model, maximizes the average likelihood of the extracted features, and aligns the multimodal semantic extraction model's target function with that of the sentiment analysis model.

3.5 Emotion Recognition and Emotionally Driven Training

In the multimodal semantic extraction model, the model obtains a series of word embedding vectors of the same dimensions, which are exported to the sentiment analysis model. In order to ensure the accuracy of sentiment analysis, this paper uses the sentiment recognition model based on LSTM (as shown in Figure 3).

$$C = LSTM(V) \qquad (8)$$

where V is the word embedding vector of the sentence, C is the emotion category, and the result of the last layer is output to a $softmax$ classifier to predict the emotion type of college gymnasts.

$$P(C|S) = \mathrm{softmax}(W_s h_n^P + b_s) \qquad (9)$$

The proposed method uses cross entropy as a loss function.

$$Loss = -\sum_{S\in X}\sum_{C=1}^{K} p(C,S)\, log\, P\,(C|S) \qquad (10)$$

where $p(C,S)$ denotes the golden probability distribution and X denotes the training text converted from the college gymnast's voice.

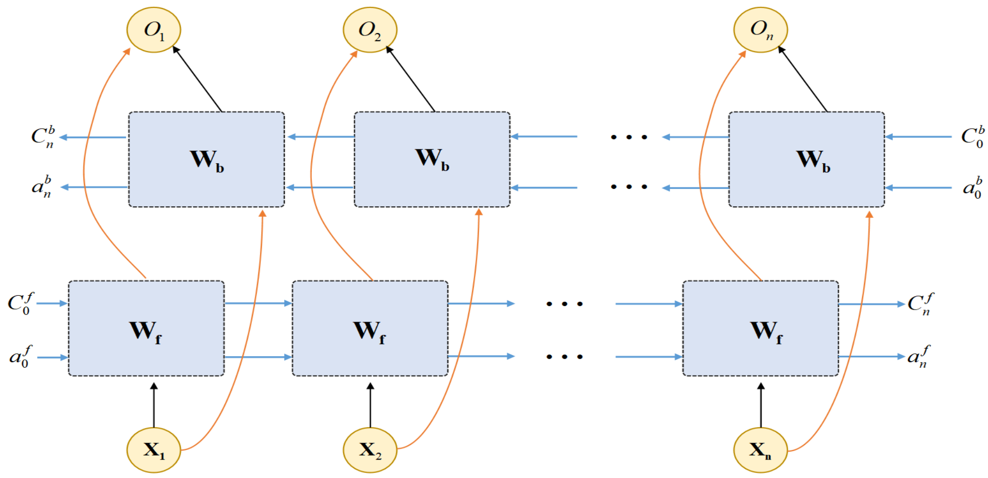

Figure 3. Schematic diagram of LSTM cell structure

After obtaining the results of emotional identification, it is suggested to adopt scientific methods and means to exert corresponding influence on the athlete's psychology, so that he or she can effectively control and regulate his or her own psychological state and behavior when training or competing, in order to achieve the best level of sports and the best sports performance. Gymnastics competition is more intense, the technical difficulty is getting bigger and bigger, especially in the case of equal technical level, instead of comparing the technique, it is more like comparing the psychological emotion, whoever has good psychological quality can give full play to the technique, and whoever can get the victory. On the contrary, if there is a lack of psychological training in general, it is very easy to make mistakes in the field, and it is bound to be a loss. It can be seen that athletes' excellent performance not only depends on their physical quality and technical and tactical training, but also has a close relationship with their psychological quality.

4. Experiment and Results

4.1 Dataset

Three common standard datasets (as shown in Table 1), SST-2, MR and Subj, are used in the experiments to verify the validity of the methodology in this paper.

(1) SST-2 is a common movie review dataset that categorizes reviews into positive and negative.

(2) MR is also a movie review dataset with the same classification objectives as SST-2.

(3) Subj is a dataset of subjective and objective statements, divided into two categories: subjective and objective.

Table 1. Three emotion recognition datasets

Data	c	l	N	$\|V\|$	$\|V_{pre}\|$	Test
SST-2	2	19	9613	16185	14838	1821
MR	2	20	10662	18765	16448	10-fold CV
Subj	2	23	10000	21323	17913	10-fold CV

4.2 Experimental setup

The word embedding vector dimension of this algorithm is 300 dimensions. Epoch is defined as 10 and the batch size is set to 128. In the short-term emotion classification model, three sizes of convolutional kernels of 3, 4 and 5 are used. In order to optimize the training of the emotion analysis model, the drupo ut is 0.5 and the learning rate is 0.5. The learning rate was 0.5. The learning rate was 0.001. The training model uses a batch gradient descent with a batch size of 32. The experimental environment is an Intel Core i5 C PU, and the comparative experiments use the same parameter settings as the corresponding paper.

4.3 Experimental results and analysis

Table 2. Experimental results on the validity of extended dictionaries.

Method	SST-2	MR	Subj
No-ExWordNet	83.2	79.2	91.6
ExWordNet	88.6	80.9	94.5
No-Multimodal	88.6	81.9	92.1
Multimodal	90.2	83.2	94.8
No-Attention	87.9	81.2	93.6
Attention	91.8	83.7	95.3

In order to verify the effect of the extended lexicon on the enhancement of the proposed method, this section compares the results of two experiments without and with the extended lexicon. In order to control the variables, the experiments use the LSTM-based sentiment analysis model, but not the multimodal semantic extraction model and the attention enhancement mechanism proposed in this chapter. From the results in Table 2, ExWordNet can enhance the semantic feature extraction. The model without extended lexicon extracts only lexical features when word2vec preprocesses the text, and the lexical features sometimes cannot accurately represent the emotional categories of the text. In contrast, the algorithm using ExWordNet can extract semantic features from a multimodal perspective, and the feature categories are more complex, especially the semantic information affecting the emotion categories of college gymnasts is given more attention.

In order to verify the effect of the multimodal semantic extraction model on the performance of the algorithm, this section compares the algorithm with and without the multimodal semantic extraction model. In order to control the variables, both groups of experiments use extended lexicon and LSTM-based sentiment analysis model, but do not use the emotional attention enhancement mechanism (Turner et al., 2022). Table 2 shows that in the comparison experiments, the algorithm without multimodal semantic extraction model simply extracts the lexical information of the text according to the extraction basis provided by the extended lexicon when extracting text features. The effective extraction and utilization of multimodal text features is neglected, resulting in incomplete text extraction results, which directly affects the accuracy of the sentiment analysis model. Compared with the algorithm without multimodal semantic extraction model, the accuracy of the algorithm with multimodal semantic extraction model is significantly improved. The multimodal semantic extraction model proposed in this chapter is able to focus on the word embedding vectors and the context information of the current text, and make the most appropriate semantic feature extraction choices every time. The model can also obtain accurate text features when the above text has an impact on the sentiment categories of the current text.

In order to verify the effect of the affective attention enhancement mechanism on the model results, this section compares the experimental results of the algorithm with and without the affective attention enhancement mechanism. In order to control the variables, the LSTM-based sentiment analysis model and the multimodal semantic extraction model are used, and the extended lexicon is not used. The experimental results in Table 2 show that the algorithm without LSTM cannot track and feedback the results of the sentiment analysis model in real time in the comparison experiments. For each text sentence, the trained model outputs the results directly, leaving the model with no way to better optimize it for the current textual sentiment analysis task. Compared with algorithms that do not use an affective attention enhancement mechanism, algorithms that use an affective attention enhancement mechanism are able to improve the accuracy of sentiment analysis. This is because the affective attention enhancement mechanism can provide feedback to the multimodal semantic extraction model based on the classification results of the sentiment analysis model, and constantly guide the optimization of the extraction model to fit the semantic feature extraction task in the current paper. The model is able to pay more attention to the semantics affecting the emotion categories and fully extract the emotion semantics to deliver the word embedding vectors that best fit the current text, which improves the accuracy of emotion recognition for college gymnasts.

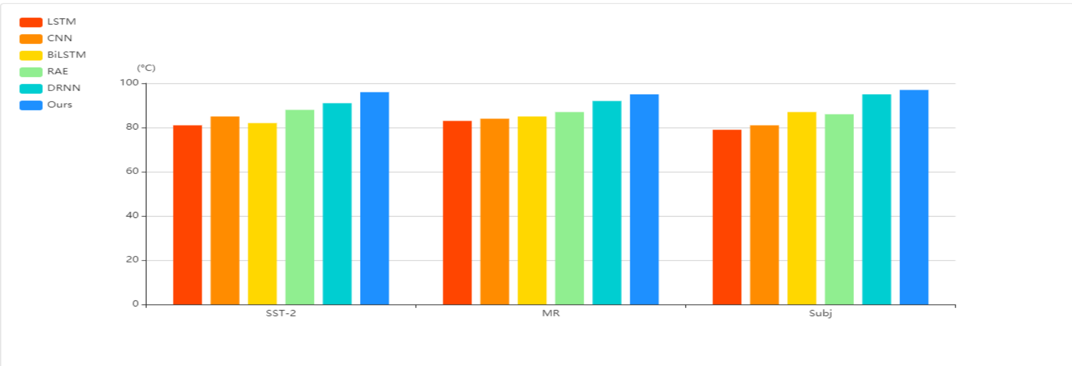

Figure 4. Histogram of the results of the comparison experiment

In terms of emotion semantic recognition and accuracy, the proposed method demonstrates better emotion recognition results compared to state-of-the-art algorithms. LSTM, CNN, BiLSTM, RAE, and DRNN generally do not have the ability to self-optimize, resulting in the algorithm not reaching the optimal state, especially the emotional semantics of college gymnasts are not considered enough. The emotional attention enhancement mechanism proposed in this article can track the classification results of the emotional analysis model, provide real-time feedback and guide the multi-modal semantic extraction model to optimize the extraction algorithm, and pay higher attention to the emotional semantics, effectively improving the accuracy of emotional analysis. In addition, under the influence of collaborative training methods, the emotional semantic recognition model based on reinforcement learning effectively understands the emotional information of college athletes, providing a positive reference for emotion-driven training methods (Teovanović et al., 2021).

5. Conclusion

In this paper, we use emotion recognition and reinforcement learning to drive the training of college students' gymnastics projects, in order to reduce the frustration and emotional fluctuations of college students during the training process. Specifically, we propose an emotional semantic recognition algorithm based on reinforcement learning. The reinforcement learning model can identify complex emotional features in college students' speech and extract emotional semantics. Among them, the multimodal semantic extraction model can fully extract the multimodal features of college students' speech, and the emotional attention enhancement mechanism guides the optimization of the semantic feature extraction algorithm of the multimodal semantic extraction model by effectively tracking and feedbacking the emotional analysis results. An expanded lexicon enables greater attention to content that affects emotional categories. Under the influence of collaborative training methods, the emotional semantic recognition model based on reinforcement learning effectively understands the emotional information of college athletes, providing a positive reference for emotion-driven training methods.

Conflicts of Interest

The authors do not have any possible conflicts of interest.

Funding Statement

This work was not supported by any foundation.

REFERENCES

Azemi, Y., Ozuem, W., & Howell, K. E. (2020). The effects of online negative word-of-mouth on dissatisfied customers: A frustration–aggression perspective. *Psychology & Marketing, 37*(4), 564-577.

Chan, H. W. Q., & Sun, C. F. R. (2021). Irrational beliefs, depression, anxiety, and stress among university students in Hong Kong. *Journal of American College Health, 69*(8), 827-841.

Chowdary, M. K., Nguyen, T. N., & Hemanth, D. J. (2021). Deep learning-based facial emotion recognition for human–computer interaction applications. *Neural Computing and Applications*, 1-18.

Fathi, J., Greenier, V., & Derakhshan, A. (2021). Self-efficacy, reflection, and burnout among Iranian EFL teachers: the mediating role of emotion regulation. *Iranian Journal of Language Teaching Research, 9*(2), 13-37.

Ford, B. Q., & Troy, A. S. (2019). Reappraisal reconsidered: A closer look at the costs of an acclaimed emotion-regulation strategy. *Current Directions in Psychological Science, 28*(2), 195-203.

Hajian, S. (2019). Transfer of learning and teaching: A review of transfer theories and effective instructional practices. *IAFOR Journal of education, 7*(1), 93-111.

Ishiwatari, T., Yasuda, Y., Miyazaki, T., & Goto, J. (2020). *Relation-aware graph attention networks with relational position encodings for emotion recognition in conversations.* Paper presented at the Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP).

Jeon, M.-K., Yoon, H., & Yang, Y. (2022). Emotional dissonance, job stress, and intrinsic motivation of married women working in call centers: The roles of work overload and work-family conflict. *Administrative Sciences, 12*(1), 27.

Khalil, R. A., Jones, E., Babar, M. I., Jan, T., Zafar, M. H., & Alhussain, T. (2019). Speech emotion recognition using deep learning techniques: A review. *IEEE Access, 7*, 117327-117345.

Korpershoek, H., Canrinus, E. T., Fokkens-Bruinsma, M., & de Boer, H. (2020). The relationships between school belonging and students' motivational, social-emotional, behavioural, and academic outcomes in secondary education: A meta-analytic review. *Research Papers in Education, 35*(6), 641-680.

Magarini, F. M., Pinelli, M., Sinisi, A., Ferrari, S., De Fazio, G. L., & Galeazzi, G. M. (2021). Irrational beliefs about COVID-19: A scoping review. *International journal of environmental research and public health, 18*(19), 9839.

Maulud, D., Zeebaree, S., Jacksi, K., Sadeeq, M., & Sharif, K. (2021). State of Art for Semantic Analysis of Natural Language Processing. Qubahan Academic Journal, 1 (2), 21–28. In.

Miranda-Mendizabal, A., Castellví, P., Parés-Badell, O., Alayo, I., Almenara, J., Alonso, I., . . . Gili, M. (2019). Gender differences in suicidal behavior in adolescents and young adults: systematic review and meta-analysis of longitudinal studies. *International journal of public health, 64*, 265-283.

Moroń, M., & Biolik-Moroń, M. (2021). Trait emotional intelligence and emotional experiences during the COVID-19 pandemic outbreak in Poland: A daily diary study. *Personality and individual differences, 168*, 110348.

Mukhopadhyay, M., Pal, S., Nayyar, A., Pramanik, P. K. D., Dasgupta, N., & Choudhury, P. (2020). *Facial emotion detection to assess Learner's State of mind in an online learning system.* Paper presented at the Proceedings of the 2020 5th international conference on intelligent information technology.

Richey, J. A., Brewer, J. A., Sullivan-Toole, H., Strege, M. V., Kim-Spoon, J., White, S. W., & Ollendick, T. H. (2019). Sensitivity shift theory: A developmental model of positive affect and motivational deficits in social anxiety disorder. *Clinical psychology review, 72*, 101756.

Ryan, R. M., Deci, E. L., Vansteenkiste, M., & Soenens, B. (2021). Building a science of motivated persons: Self-determination theory's empirical approach to human experience and the regulation of behavior. *Motivation Science, 7*(2), 97.
Schoneveld, L., Othmani, A., & Abdelkawy, H. (2021). Leveraging recent advances in deep learning for audio-visual emotion recognition. *Pattern Recognition Letters, 146*, 1-7.
Smith, J., Guimond, F.-A., Bergeron, J., St-Amand, J., Fitzpatrick, C., & Gagnon, M. (2021). Changes in students' achievement motivation in the context of the COVID-19 pandemic: A function of extraversion/introversion? *Education Sciences, 11*(1), 30.
Teovanović, P., Lukić, P., Zupan, Z., Lazić, A., Ninković, M., & Žeželj, I. (2021). Irrational beliefs differentially predict adherence to guidelines and pseudoscientific practices during the COVID-19 pandemic. *Applied Cognitive Psychology, 35*(2), 486-496.
Turner, M., Miller, A., Youngs, H., Barber, N., Brick, N., Chadha, N., . . . Evans, A. (2022). "I must do this!": A latent profile analysis approach to understanding the role of irrational beliefs and motivation regulation in mental and physical health. *Journal of sports sciences, 40*(8), 934-949.
Weiner, B. (1982). An attributionally based theory of motivation and emotion: Focus, range, and issues. *Expectations and actions: Expectancy-value models in psychology*, 163-204.
Yang, J., Yan, X., Chen, S., Liu, W., Zhang, X., & Yuan, J. (2022). Increased motivational intensity leads to preference for distraction over reappraisal during emotion regulation: Mediated by attentional breadth. *Emotion, 22*(7), 1595.
Yüvrük, E., Kapucu, A., & Amado, S. (2020). The effects of emotion on working memory: Valence versus motivation. *Acta Psychologica, 202*, 102983.