International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 17 (2025) pp. 1-18

https://doi.org/10.70917/ijcisim-2025-0293
Article

Analyzing the Relationship between Students' Emotional Changes and Ideological and Political Quality Enhancement in the Process of Civic Education Using LSTM Network Modeling

Lan Jiang *

Zhengzhou Academy of Fine Arts, Zhengzhou, Henan, 451450, China; 15736709619@163.com

Abstract: In this paper, we first preprocess the speech signal by methods such as pre-emphasis and windowing, and then extract the speech features by using the Mel Frequency Cepstrum Coefficients (MFCC) in the spectral features, and construct a speech emotion recognition model based on the bidirectional LSTM to capture the semantic correlation between long sequences. In order to synthesize the complementary information in the dataset and reduce feature redundancy, this paper constructs a speech emotion recognition model incorporating two-way CNN-LSTM and attention mechanism to analyze the emotional changes of students in the process of Civic Education. The results show that the method in this paper achieves the highest performance on the CASIA emotion corpus. The classroom as a whole presents a positive and active state (most of the emotion value P is distributed between [0,2]), the classroom atmosphere is good, and the teacher-student interaction has a positive effect on the improvement of ideological and political quality; in addition, the classroom emotion can also affect the classroom effect to a certain extent, which influences the effect of the students' quality improvement in the process of Civic and political education. Reasonably guiding the students' emotions such as "conscientiousness and empathy sublimation" in the classroom of Civic and Political Education can effectively improve the students' ideological and political quality.

Keywords: MFCC feature extraction; CNN-LSTM; Speech emotion recognition; Civic and political education

1. Introduction

Civic and political education is a long-term and arduous task, aiming at cultivating students' correct political theoretical concepts and ideological and moral qualities, so that they can become socialist builders and successors. In this process, the role of emotion can not be ignored, emotion and Civic and Political Education has a close connection, it can stimulate the enthusiasm and enthusiasm of students to help students establish correct political beliefs and attitudes to improve the ideological and political quality [1-3].

Emotions are the subjective experience of human beings they have an important influence on human cognition, emotion, will and action and other psychological processes [4]. In the process of Civic and political education, students' emotions play a vital role throughout [5]. In students' Civic and Political Education emotion is the internal power to motivate students to actively engage in learning, when students agree with a political theory, anger or excitement about a social phenomenon they will be more active to understand, think and accept the relevant knowledge [6-9]. The activation of emotion can make students have a strong emotional experience in the learning process, which can lead to the formation of strong political positions and beliefs [10-11].

Emotion as a part of the cognitive subject carries the content of cognition and plays an important role in the quality of cognition, in the Civic Education students' emotion is a concrete expression of in-depth understanding and perception of political theory is the inner feelings and experience of the political knowledge learned [12-15]. Through the communication and expression of emotions, teachers can better understand the students' political cognition and identity, which is conducive to adjusting teaching methods and teaching content, and improve the relevance and effectiveness of Civic and Political Education [16-18]. At the same time, emotion is also an important factor that affects the effect of civic education, students in the emotional appeal easily tend to believe in a certain political theory of social phenomena and historical events to produce a deep sense of understanding of the formation of the socialist cause of identity and support to establish a correct socialist political outlook, improve the quality of ideological and political [19-23].

In order to meet the needs of students' emotion analysis in the classroom of ideological and political education, this study uses six typical emotion types in the CASIA database, and constructs a speech emotion recognition model based on bidirectional LSTM after preprocessing, MFCC feature extraction, and long-short-term memory network (LSTM). In order to solve the problems of insufficient feature extraction and poor model recognition effect in the speech emotion recognition process in ideological and political classroom, this paper proposes a speech emotion recognition model that integrates CNN-LSTM and attention mechanism, and applies the model in the recognition of changes in teachers' and students' speech emotions in the classroom of ideological and political education. Finally, combined with multiple linear regression analysis, it explores the influence of four kinds of emotional changes on the improvement of the quality of Civic and Political Education.

2. LSTM based model for student emotion recognition

2.1. Speech emotion feature extraction

2.1.1. Emotional corpus

Affective speech libraries are the source of speech signal input and the basis for affective identification. In this study, an existing speech database was used. Considering cultural differences, pronunciation habits, and robustness, this study used the CASIA Chinese Emotion Database, which has a total of 1,200 speech sounds, including six emotions: fear, surprise, anger, sadness, astonishment, happiness, and neutrality.

2.1.2. Pre-processing

The preprocessing process consists of four main components: pre-emphasis, window addition, frame splitting, and endpoint detection.

(1) Pre-emphasis. Frequency domain analysis of the speech signal, the high-frequency part of the spectrum than the low-frequency part of the difficult to find, for this reason to be pre-emphasis processing in the pre-processing.

(2) Window addition. Adding a window means multiplying a certain window function $w(n)$ by $s(n)$, thus forming a windowed speech signal $sw(n) = s(n) * w(n)$. Commonly used window functions are rectangular windows and Hamming windows:

Rectangular windows:

$$w(n) = \begin{cases} 1, & 0 \le n \le N-1 \\ 0, & \text{Other} \end{cases} \tag{1}$$

Hamming windows:

$$w(n) = \begin{cases} 0.54 - 0.46\cos[2\pi n/(N-1)], & 0 \le n \le N-1 \\ 0, & \text{Other} \end{cases} \tag{2}$$

The rectangular window is essentially no windowing relative to the original signal, and the n in the window function, which refers to the window length (number of sample points), corresponds to one frame of the signal.

(3) Framing. In order to carry out short-time analysis and the analysis of the essential characteristic parameters of the speech signal, it is necessary to segment the signal first, in which each segment is called a frame, and each frame is usually taken as 10ms.

(4) Endpoint detection. The main purpose is to accurately find out the start and end points of the speech signal from a section of the speech signal, and its purpose is to separate the effective speech signal from the useless signal.

2.1.3. Speech emotion feature extraction

Speech signals carry a large amount of useful information which contains a variety of emotional feature parameters. The commonly used features for speech emotion recognition are mainly rhythmic features and spectral features, and spectral correlation based feature extraction is used in this study. The classification based on spectral correlation feature extraction is shown in Fig. 1.

Based on spectral correlation features
- Linear prediction cepstrum coefficient (LPCC)
- Mel Frequency Cepstrum Coefficient (MFCC)
- Logarithmic Frequency Power Coefficient (LFPC)
- Log-Mel inverse spectral coefficient (LOG-Mel)

Figure 1. Extract spectral features.

The process of MFCC extraction is as follows:

Speech signal → pre-emphasis → frame-splitting → windowing → fast Fourier transform → Meier filter → logarithmic operation → discrete cosine transform. Next, we mainly introduce the fast Fourier transform, Meier filter, logarithmic energy operation, and discrete cosine transform.

(1) Fast Fourier Transform (FFT)

This experiment is studied in the frequency domain, different energies are distributed in the frequency domain, which represent different characteristics of speech. After splitting frames and adding Hamming window, Fourier Fast Transform is performed on each frame of the signal to get the spectrum of each frame, and then the spectrum of the speech signal is squared by taking the mode of the speech signal, then the power spectrum of the speech signal can be obtained. Let the DFT of the speech signal be:

$$X_a(k) = \sum_{n=0}^{N-1} x(n) e^{-2\pi ik/N}, 0 \le k \le N \tag{3}$$

where $x(n)$ is the input speech signal and N denotes the number of points of Fourier transform.

(2) Mel filter

Pass the energy spectrum through a triangular filter bank. Define a filter bank with M filters, the triangular filter is used in this experiment:

$$H_m(k) = \begin{cases} 0, & k < f(m-1) \\ \dfrac{2(k - f(m-1))}{(f(m+1) - f(m-1))(f(m) - f(m-1))}, & f(m-1) < k < f(m) \\ \dfrac{2(f(m+1) - k)}{(f(m+1) - f(m-1))(f(m) - f(m-1))}, & f(m-1) < k < f(m) \\ 0, & k > f(m+1) \end{cases} \tag{4}$$

where $\sum_{m=0}^{M-1} H_m(k) = 1$.

After passing through the triangular bandpass filter, the signal spectrum is smoothed, eliminating harmonic effects and highlighting the resonance peaks of the original speech.

(3) The logarithmic energy of each filter

$$s(m) = \ln\left(\sum_{k=0}^{N-1} |X_a(k)|^2 H_m(k) \right), 0 \le m \le M \tag{5}$$

(4) Discrete Cosine Transform (DCT)

$$C(n) = \sum_{m=0}^{N-1} s(m) \cos\left(\frac{\pi n(m-0.5)}{M} \right), n = 1, 2, \ldots, L \tag{6}$$

2.2. *Speech emotion recognition model fusing two-way CNN-LSTM with attention mechanism*

2.2.1. Long and Short Term Memory Network Model (LSTM)

LSTM is improved on the basis of Recurrent Neural Network (RNN).The core structure of LSTM consists of four parts: forgetting gate, input gate, cell state update, and output gate.

(1) Oblivion Gate

It is mainly to selectively forget the incoming input from the previous node, and control how much information can be retained from the previous moment's unit state to the current moment's state by the output value of the Sigmoid function.

(2) Input Gate

It is mainly used to control how much information in the input of the network at the current moment can be preserved into the unit state at the current moment.

(3) Cell state update

Multiply the forgotten gate gate value just obtained with the $C(t-1)$ obtained at the previous time step, plus the result of multiplying the input gate gate value with the unupdated $C(t)$ obtained at the current time step. The updated $C(t)$ is finally obtained as part of the input for the next time step, and the whole process of cell state updating is the application of the forgetting gate and the input gate.

(4) Output Gate

It is mainly used to control how much information can be saved for output in the current unit state. The whole process of output gate is to generate the implied state $h(t)$.

2.2.2. Modeling

The speech emotion recognition model fusing two-way CNN-LSTM [24] with attention mechanism proposed in the paper is shown in Fig. 2. Its preprocessing and part includes a one-dimensional residual convolutional neural network and a two-dimensional multi-scale convolution; the feature fusion part based on the attention mechanism categorizes the original speech signals into emotions.

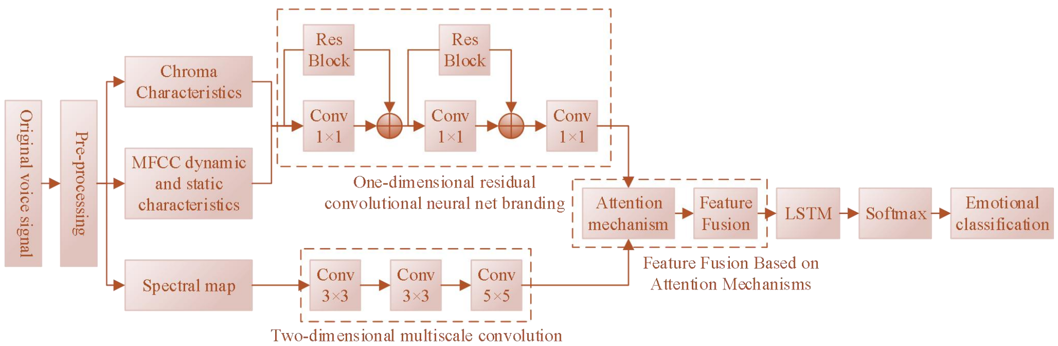


Figure 2. CNN-LSTM with attention mechanism for speech emotion recognition.

2.2.3. Two-way multi-dimensional multi-scale feature extraction

In order to make full use of feature diversity and complementarity and enhance the effective characterization of emotional features, a two-way multi-dimensional multi-scale feature extraction method is proposed, one way to extract Chroma features and MFCC features, and one way to extract speech signal spectrogram features.

(1) One-dimensional residual convolutional neural network feature extraction

Considering that the Chroma feature has a better resolution in the low-frequency part, which is in line with the human ear auditory perception, the Chroma feature is introduced as one of the emotional features in the paper, which is combined with the MFCC feature as the input to the one-dimensional residual convolutional neural network.

(2) Two-dimensional multi-scale convolutional feature extraction

The speech spectrogram is used as the input to the two-dimensional multi-scale convolutional neural network [25], and two different sizes of convolution kernels, 3×3 and 5×5, are used to perform convolution operations on the input data, and then the different scale features are learned to extract the local details and the overall contour information of the spectrogram, and the synthesized spectrogram and the time-domain waveform characteristics show the change process of the speech spectrogram over time in a better way.

2.2.4. Attention mechanism based feature fusion

Simple fusion of two-way features will cause problems such as feature overlap, information redundancy, and computational complexity, weakening the recognition rate of emotion classification. To address these problems, feature fusion based on the attention mechanism is proposed. $X_{\text{Chroma+MFCC}}$ refers to the one-dimensional residual convolutional neural network output features, X_{spc} refers to the two-dimensional multiscale convolutional neural network output features, K refers to the key, Q refers to the query, V refers to the value, W^Q, W^K, W^V refers to the matrix of the weight parameters, w and $1-w$ are self-attention weights and cross-attention weights, respectively.

The attention mechanism is divided into two parts, the first part calculates the self-attention Z_A and Z_B for $X_{\text{Chroma+MFCC}}$ and X_{spc}, respectively, and assigns attention to $X_{\text{Chroma+MFCC}}$ and X_{spc} by the attention weighting parameter, Z_A and Z_B denoted as follows:

$$Z_A = \text{Softmax}\left(\frac{Q_A\left(K_A\right)^{\mathrm{T}}}{\sqrt{d}}\right)V_A \tag{7}$$

$$Z_B = \text{Softmax}\left(\frac{Q_B\left(K_B\right)^{\mathrm{T}}}{\sqrt{d}}\right)V_B \tag{8}$$

Q_i, K_i, V_i is calculated as follows:

$$Q_i = X_i W^Q \tag{9}$$

$$K_i = X_i W^K \tag{10}$$

$$V_i = X_i W^V \tag{11}$$

where i is either $X_{\text{Chroma+MFCC}}$ or X_{spe}. The second part computes the cross-attention between $X_{\text{Chroma+MFCC}}$ and X_{spe}, fuses the important features of $X_{\text{Chroma+mFCC}}$ and X_{spe}, and synthesizes the complementary information of the two to reduce the feature redundancy effects. By assigning different weight coefficients to $X_{\text{Chroma+MFCC}}$ and X_{spe} with the final attention parameter Z, the final feature representation $X_{\text{Chroma+MFCC+spe}}, Z$ is obtained as follows:

$$Z = wZ_A + (1-w)Z_B \tag{12}$$

2.2.5. Extracting temporal characteristics of fused features

The fused two-way multidimensional multi-scale features XChroma+MFCC+spe are fed into the LSTM network for further extraction of temporal features and obtaining contextual semantic information.The LSTM network is computed by selectively updating and forgetting the information through the input gates, forgetting gates, and output gates, to remember the important information in a long data sequence and retain it:

$$\begin{aligned}
i_t &= \sigma\left(W_i \cdot \left[h_{t-1}, X_{\text{Chroma+MFCC+spe}}(t)\right] + b_i\right) \\
f_t &= \sigma\left(W_f \cdot \left[h_{t-1}, X_{\text{Chroma+MFCC+spe}}(t)\right] + b_f\right) \\
o_t &= \sigma\left(W_o \cdot \left[h_{t-1}, X_{\text{Chroma+MFCC+spe}}(t)\right] + b_o\right)
\end{aligned} \tag{13}$$

where $X_{\text{Chroma+MFCC+spe}}(t)$ denotes the input at the current moment, h_{t-1} denotes the hidden state at the previous moment, σ is the activation function that maps each element within $[0,1]$, W_i, W_o, W_f is the weight parameter matrix, and b is the bias.

$\tilde{C}_t$ is the information of the input element after scaling, which takes the value between 0 and 1. C_t is the cell state at the current moment, and the hidden state h_t at the current moment is calculated by Eq. (16):

$$\tilde{C}_t = \tanh\left(W_c \cdot \left[h_{t-1}, X_{\text{Chroma+MFCC+spe}}(t)\right] + b_c\right) \tag{14}$$

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{15}$$

$$h_t = o_t \times \tanh\left(C_t\right) \tag{16}$$

2.2.6. Categorization of emotions

The Softmax classifier is chosen for the sentiment four classification output, exponentiate each element in the LSTM extracted sequence $h_t = \{h_1, h_2, \cdots, h_N\}$ containing temporal information, N is the length of the signal and the exponentiation result is normalized, and finally output the probability of the The category corresponding to the largest element is used as the prediction result.The Softmax function is:

$$P_i = \frac{e^{h_{ti}}}{\sum_{j=1}^{N} e^{h_{tj}}} \tag{17}$$

where h_i is the LSTM output sequence containing N elements, and h_{ti}, h_{tj} denote the i, j elements in h_t, respectively.

2.3. Analysis of speech emotion recognition results

In this paper, weighted average recall (WAR) and unweighted average recall (UAR) are calculated as performance measures of the algorithm using CASIA database.

In this paper, four sentiment categories (ANGRY, SAD, NEUTRAL, HAPPY) plus a NULL category are set. Three different methods for setting the length of the label sequence are investigated: 1) the number of words in speech (WN); 2) the number of phonemes in speech (PN); and 3) twice the number of phonemes (PN*2).

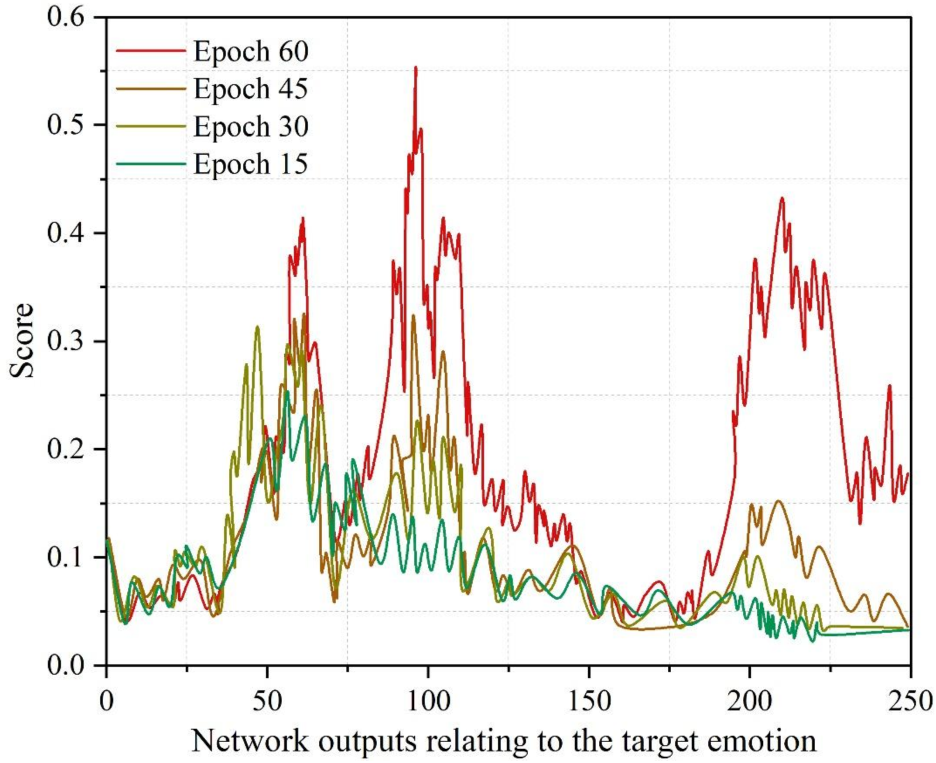
The experimental results of different methods on CASIA library are shown in Table 1. It can be seen that this paper's model achieves the best recognition results and is better than both $|y^{\text{seg}}|$= WN and $|y^{\text{seg}}|$= PN*2 when the label length $|y^{\text{seg}}|$= PN. The model in this paper has improved UAR and WAR values compared with the results of the DNN model, and significantly outperforms other LSTM models with sequence-to-label approach (t-test, $p<0.05$). That is, the number of speech emotion frames is comparable to the number of phonemes in speech. When the label length is set to the number of phonemes in speech, the network is just able to cover these emotion labels to most of the speech emotion frames; when the number of labels is the number of words, the number of labels is too small, and some emotion frames are not learned by the network; when the number of labels is twice the number of phonemes, the number of labels is too large, and some non-emotion frames are also learned by the network; therefore, the recognition effect of these two methods of setting the label length is worse Therefore, the recognition effect of these two label length setting methods will be poor.

Table 1. Experimental results of different methods on the CASIA library.

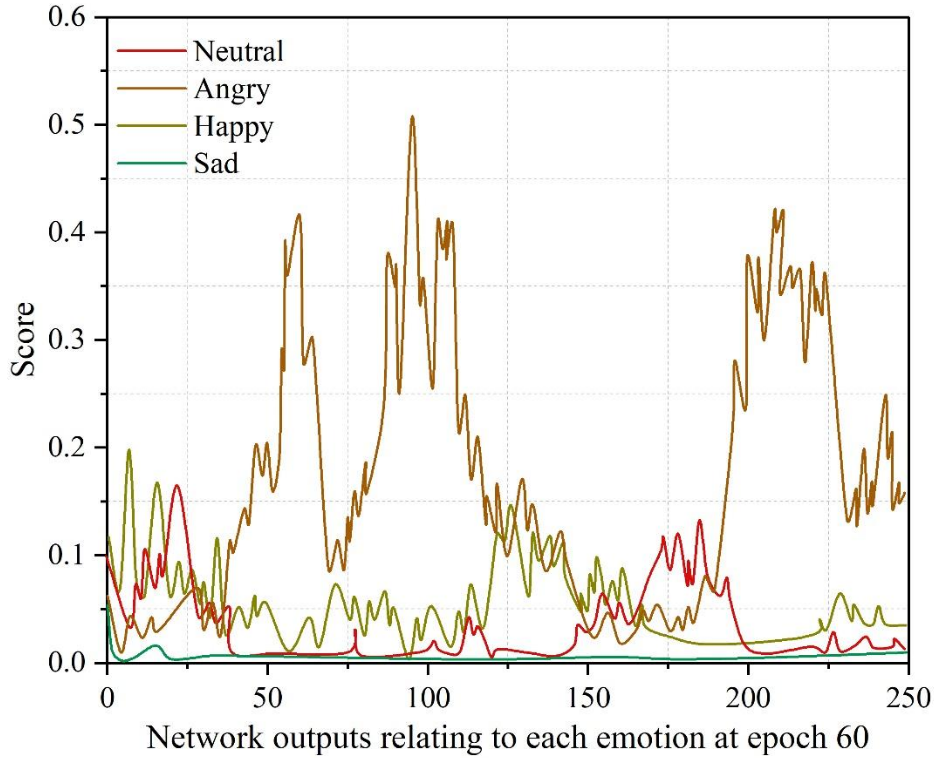
Method	Control methods	UAR(%)	WAR(%)
Control methods	DNN	62.43	61.67
	frame-wise LSTM	63.16	61.21
	final-pooling LSTM	53.01	52.64

	mean-pooling LSTM	63.78	62.66
	LSTM+CTC-like	59.73	57.64
	LSTM+ELM	63.48	62.27
This method	CNN-LSTM,$\lvert y^{seg}\rvert$=WN	63.75	63.35
	CNN-LSTM, $\lvert y^{seg}\rvert$=PN	65.67	64.38
	CNN-LSTM, $\lvert y^{seg}\rvert$=PN*2	65.04	63.04

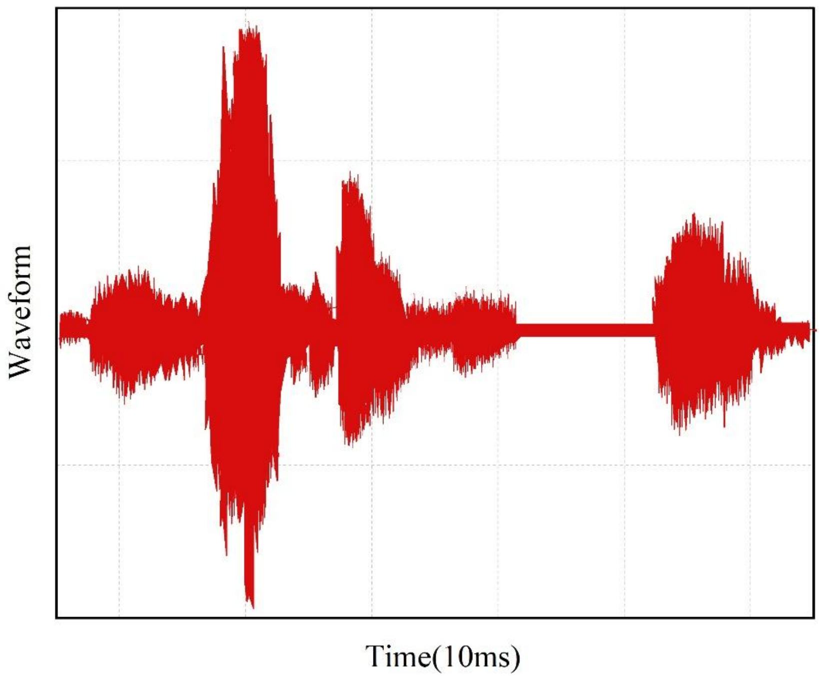
The network output results during the training process are shown in Fig. 3, where (a) the network output of the target emotion at different training rounds (b) the network output of the four types of emotions at the 60th training round (c) the original speech signal. If the activation value of the network is regarded as the probability of a speech frame as an emotion frame, it can be seen that the activation value of the network at the mute frames is very low and tends to zero, which indicates that they are non-emotion frames. Moreover, speech frames with high energy do not necessarily have the largest activation value, which indicates that the model does not just focus on the high energy places in speech, it is also able to consider the expression of emotional information elsewhere in speech.



(a) Target emotion network output under different training rounds



(b) Network output of four types of emotions



(c)Original audio

Figure 3. Network output during training.

The confusion matrix of the recognition results of the test set under the third cross-validation experiment is shown in Fig. 4. The experiment is aimed at optimizing the performance of UAR, so the “angry” and “sad” categories, which have fewer samples, also have higher recognition rates, both reaching more than 70%. Listening to the misclassified samples, the following misrecognition patterns are summarized:

1) The "happy" and "neutral" categories, as well as the "neutral" and "sad" categories, are easily confused.

2) In the "happy" category, emotions that are not particularly obvious are easily identified as "neutral".

3) In the "neutral" category, if the volume is too loud or the speaking speed is too fast, it is easy to be identified as "happy".

4) In the "neutral" category, when the tone is weak or there are many pauses, it is easy to be recognized as "sad".

Figure 4. Confusion matrix of test set recognition results.

3. Impact of Civic Education on Students' Emotional and Ideological and Political Qualities

3.1. Emotional Analysis of Teacher-Student Interaction Voice in the Process of Civic and Political Education

3.1.1. Experimental data processing

Based on designing the Civics classroom emotion recognition model, this study applies the emotion recognition technology to the evaluation of Civics education and teaching scenarios for detecting classroom learning emotion changes. The flow chart of classroom speech emotion analysis is shown in Figure 5, and the analysis steps are:

(1) Course speech extraction and recognition;

(2) Segmentation of utterances to obtain natural utterance audio.

(3) Recognize the voice pattern to complete the classification of teacher and student speech.

(4) Recognize the speech emotion and get the speech emotion P-value of each utterance.

(5) Classroom interaction emotion analysis: through the audio emotion P-value obtained from the classroom voice emotion model, draw the voice emotion change curve for classroom effect evaluation.

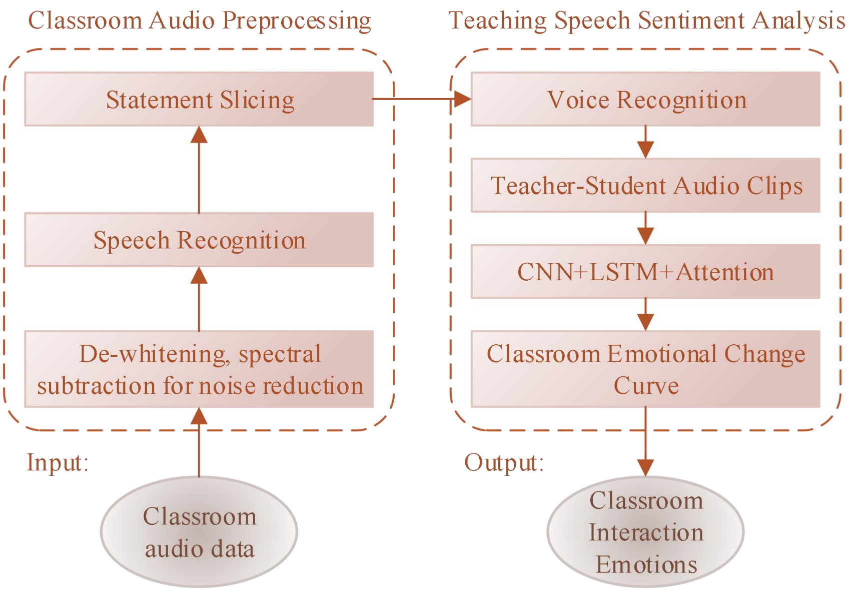


Figure 5. Classroom speech emotion analysis flow chart.

3.1.2. Trend Analysis of Emotional Changes

The classroom video selected for this experiment is a second grade Civic and Political Education class in a middle school, and this classroom video is a public video recorded in an offline teaching scenario. Non-voice emotion segments were removed and four effective interaction voice segments were retained, with durations of 310, 270, 260, and 220 seconds, respectively, totaling 1060 seconds of effective classroom interaction voice.

The voice emotion change curve of Civics classroom interaction segment I is shown in Figure 6. 108 seconds before the teacher lecture, the emotion value P in the range of [0.11,1.39], indicating that at the beginning of the Civics course the teacher emotion has not yet been fully activated, and is in a low degree of positive emotion; 108~226 seconds stage, the teacher based on the content of the previous class, the classroom questions to the students, so the overall emotion fluctuation is larger, and in the range of [-0.56,0.92] range, indicating that when carrying out the Civics classroom questions and answers, because the discussion content are with a questioning tone, and for the students to answer the questions together, the teacher and the students voice staggering, there is a large fluctuation in the value of the emotional assessment; 226~310 seconds stage, the teacher in the introduction of the wide application of political thought in life, the emotional value P is increasing, indicating that in the life of the common knowledge of knowledge expansion the teacher's emotion is in a positive state, and is conducive to activating the teacher's emotional state.

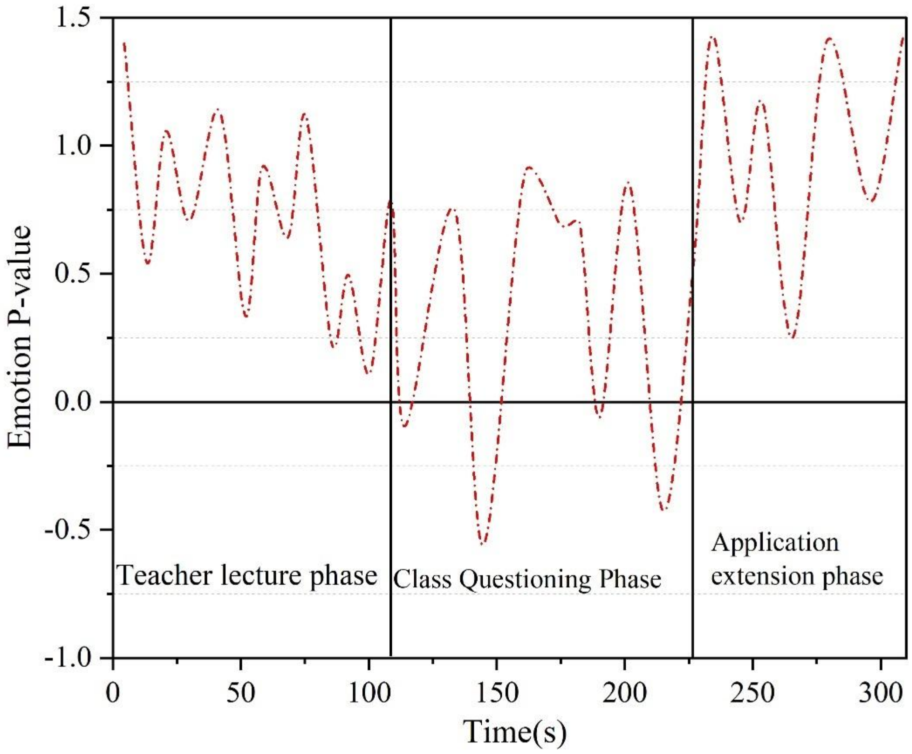


Figure 6. The voice emotion curve of the first classroom interaction segment.

The change curve of speech emotion of Civics classroom interaction fragment II is shown in Figure 7. It can be seen that the sentiment value P of clip two is mainly distributed between [1.31,2.03], before 75 seconds for the students to speak, the sentiment is positive and positive state, after 75 seconds, the teacher's sentiment shifts to about [1.31,2.03], which indicates that the active speech of the students in the process of Civics teaching has a positive effect on the teacher's sentiment, and there is a correlation between the teacher's emotion and the students' emotion. Compared with clip 1, clip 2 emotion value increases significantly, indicating that when the classroom proceeds to 580 seconds, the speech emotion value P reaches a high level, and the Civics classroom has been completely in the activation state.

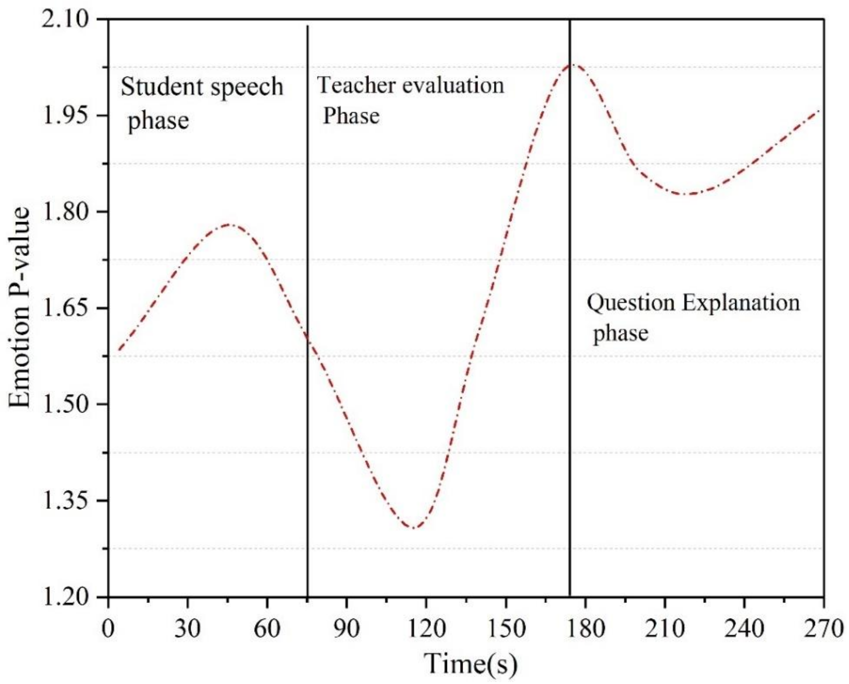


Figure 7. Voice emotion curve of Classroom interaction fragment 2.

The sentiment change curve of Interaction Segment III is shown in Figure 8. Interaction voice fragment three is after the student exercises and discussions, this fragment is a practice Q&A fragment, by the students for the blackboard questions to correct errors, the overall sentiment value P distribution between [1.24,1.78]. It is more similar to the emotional value of clip 2. During the teaching process of Civics, the teacher is asking questions continuously, the students' answers are always confident and correct, and the teacher gives encouragement and affirmation, indicating that the Civics classroom is always in an active state in the question-answer phase, and the students' emotions are positively mobilized.

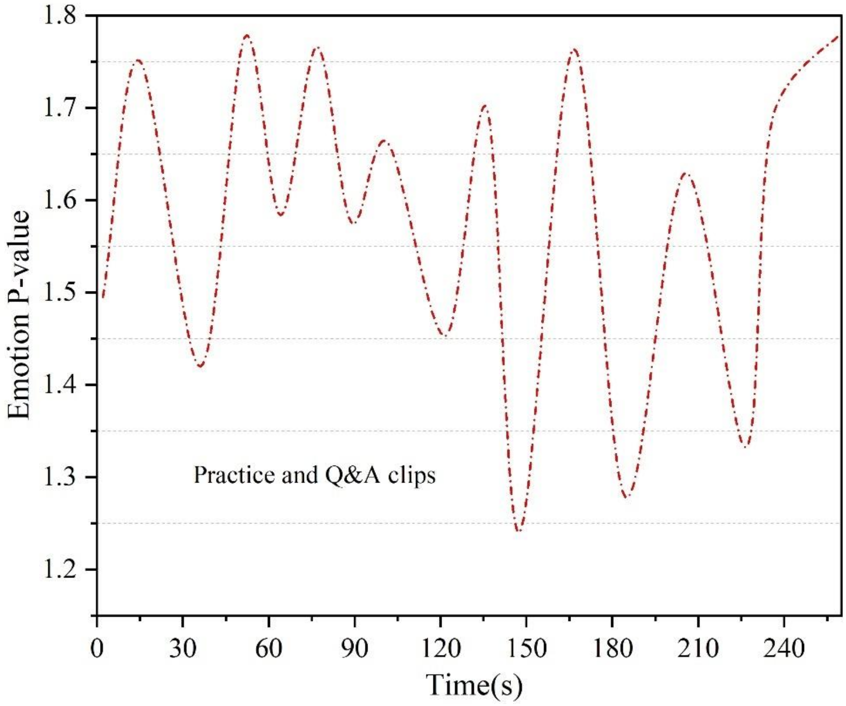


Figure 8. Interactive Segment 3 Emotional Change Curve.

In classroom interaction segment four, students developed explanations of why in response to the question, which was an open-ended question. The emotion change curve of interaction segment IV is shown in Figure 9. The first 60 seconds were spoken by two students on their own initiative, and later the teacher guided and communicated with the students according to the knowledge difficulties of the Civics teaching process, and the overall Civics classroom affective value P was distributed within the range of [0.83,1.37]. At 102 seconds, the teacher gives rewards to the students who perform well, and the voice affective value rises, while at 176 seconds and 198 seconds, the teacher asks the students in a questioning tone, and the voice pleasantness decreases to near 1.18, indicating a certain validity of the model.

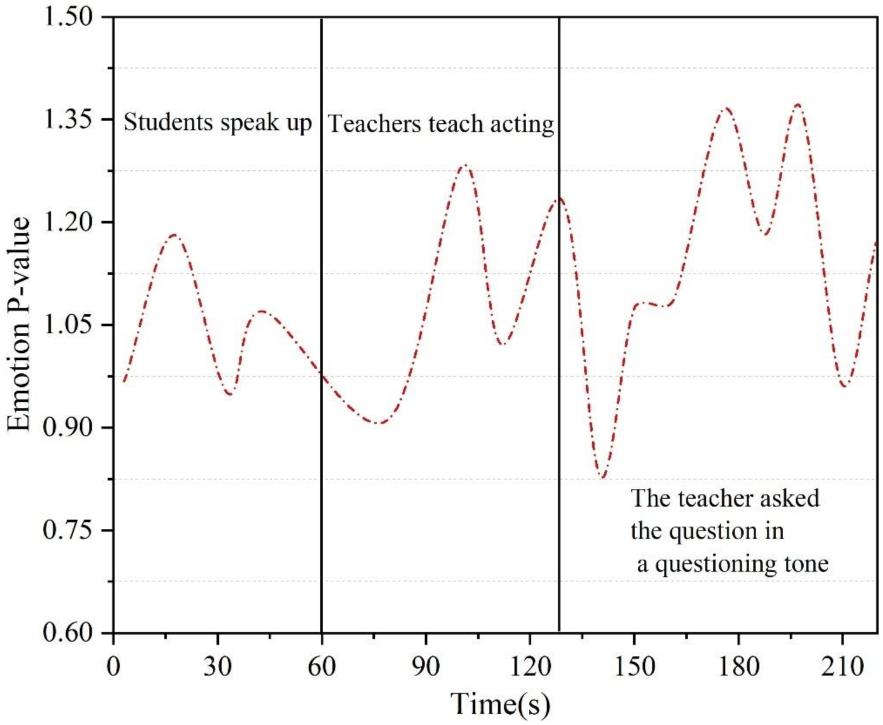


Figure 9. Interactive Segment 4: Emotional Change Curve.

Through the above experiments, it can be found that the overall emotional value P of the Civics classroom is mostly distributed between [0,2], indicating that the overall Civics teaching process presents a positive and active state, positive emotions have a positive impact on the effect of teacher-student interactions, the Civics classroom has a good atmosphere, and the overall effect of classroom teacher-student interactions forms a benign cycle.

3.1.3. Analysis of changes in classroom sentiment across ratings

In order to evaluate the correlation between classroom emotions in ideological and political education and the improvement of ideological and political quality, in the ideological and political education classrooms of the second year of junior high school, this paper selects the same course with more page views and comments from the provincial and municipal excellent teaching classrooms respectively, and conducts a visual analysis of classroom emotions in classrooms with different ratings. By comparing the teaching methods and teaching atmospheres of classrooms with different evaluation grades, Confirm the significance of emotional analysis in ideological and political classes for enhancing ideological and political qualities.

The curve of emotional change in the classroom of provincial excellence is shown in Figure 10. In the process of Civic Education, the overall classroom emotion fluctuates within the range of [0.2,1], first of all, the overall analysis, there are many points of emotional decline in the Civic Education classroom, based on the time of the decline point linked to the classroom activities, it is learned that these plummeting points for the students to read aloud in unison or speak in unison segments, which indicates that the explanation for the chaotic sound signal model is poor; through the combination of the Civic Education process of the interactive behavior of the classroom, there are About half of the time was spent on lectures, and half of the time was spent on teacher-student interactions and student-initiated problem solving for classroom problems, and it was observed that students had a high level of pleasurable emotions when they were performing exercises, which was conducive to the enhancement of classroom emotions in the process of Civic Education.

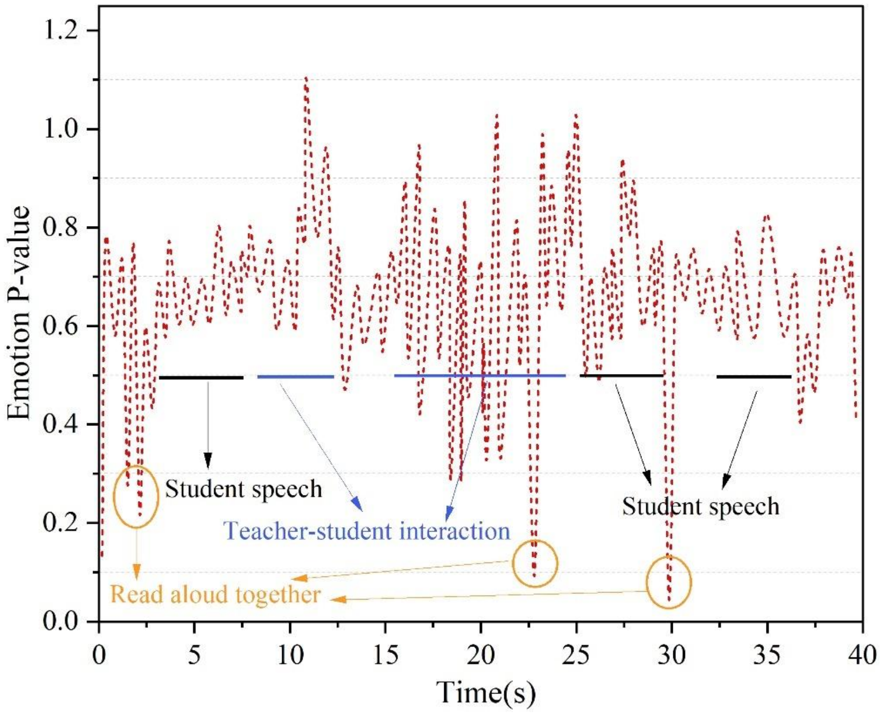


Figure 10. Emotional change curve of provincial excellent classroom.

The emotional change curve of the classroom of the city's best is shown in Figure 11. In the process of Civics education, the overall classroom emotion fluctuates within the range of [-0.5,0.8], and the overall classroom emotional state is lower, combined with the analysis of classroom voice behavior, the teacher-student interactions and students' active speeches in the process of Civics teaching are fewer and the P-value is in the higher state only when the students are speaking, which indicates that the students' voice emotion in the process of Civics teaching is slightly higher than that of the teachers.

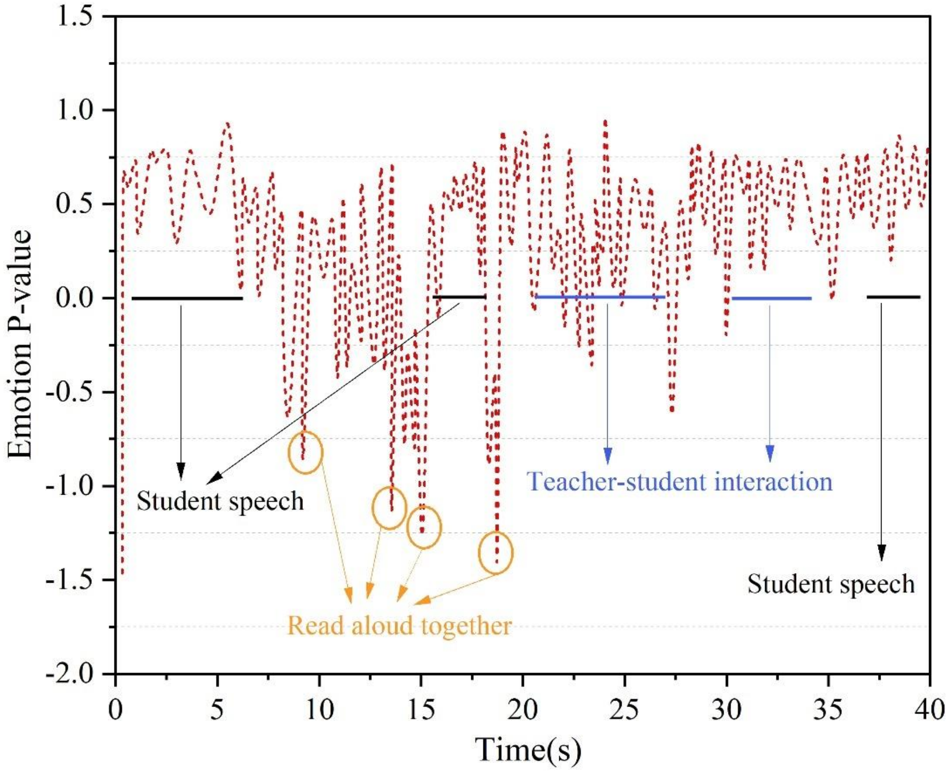


Figure 11. Emotion curve of municipal excellent classroom.

3.2. Analysis of the Relationship between Students' Emotional Changes and Ideological and Political Quality Enhancement

3.2.1. Regression Analysis of Students' Emotional Changes and Ideological and Political Qualities

In this paper, 156 students from three classes in the first year of X Middle School were used as a pilot class to investigate the relationship between the emotional changes in students' voice and the improvement of their ideological and political literacy in the process of Civic and Political Education. Here, four types of emotional changes are mainly selected, namely, "conscientiousness in taking up responsibilities, sublimation of empathy, internalization of identity, and firmness in refining".

Considering that multiple regression analysis can study the interdependence between various variables, this paper sets the improvement of students' ideological and political quality as Y as the dependent variable, and sets the four emotional changes as X as the independent variable, and carries out multiple linear regression analysis.

A total of 156 questionnaires were distributed in this research, and 100% of the questionnaires were effectively recovered. The reliability and validity of the questionnaire were 0.8872 and 0.8619 respectively, indicating that this questionnaire is more reliable and its results can be used as the basis for analyzing and arguing.

The descriptive statistics of the survey data showed that among the four independent variables, the mean value of take-charge self-awareness was the highest, followed by empathy sublimation and identity internalization, and lastly, quenching and firming emotional change. There are no outliers or missing values in the mean and standard deviation data of each variable. For the emotional change of quenching firmness, its mean value reaches 3.8501, indicating that most of the students very much recognize that quenching firmness and all the other three emotional changes can enhance the ideological and political education literacy.

The results of regression analysis between each variable and the improvement of ideological and political literacy are shown in Table 2. Through the significance test of regression contrast, it is found that its F value is 20.2058, $P < 0.05$, and there is a significant linear relationship between the dependent and independent variables, and the regression analysis is valid and statistically significant. The correlation coefficients of the four independent variables of "taking charge of self-consciousness, resonance sublimation, identity internalization and quenching firmness" for the dependent variable of the improvement of the quality of civic education are 0.0169, 0.0262, 0.044 and 0.0119, respectively, which are all $P < 0.05$ and statistically significant, indicating that the influence of these four independent variables on the dependent variable is significant. In addition, the four influencing factors of this model have a fit of 31.78% for the improvement of the quality of Civic Education, and the correlation between the respective variables and the dependent variable is good.

Table 2. Regression analysis results between variables and dependent variables.

Non-standardized coefficient	Standard coefficient	t	Sig.	Linear correlation statistic
Constant	1.629	0.2401		6.398
Take responsibility	0.2003	0.0908	0.0169	3.2008
Resonance and Sublimation	0.8983	0.05	0.0262	1.8022
Identify and internalize	0.1508	0.0427	0.044	4.1014
Temper your resolve	0.0732	0.0502	0.0119	1.7016

3.2.2. Regression Test of Emotional Change and Ideological and Political Quality Improvement

(1) Covariance diagnosis of the regression equation

The results of covariance statistics show that the eigenvalue is 0.1023 and the conditional index is 6.1577, which shows that there is no multicollinearity between the respective variables, and the regression coefficients are reliable, which proves that there is no covariance problem in this model.

(2) Residual analysis of regression equation

The histogram of the distribution of residuals with the normal distribution curve is shown in Figure 12; the normal P-P plot of the standardized residuals is shown in Figure 13. According to the distribution of residuals reflected in the calculation results of the questionnaire data, the histogram of residual distribution and the normal distribution curve attached to it can intuitively reflect that the residuals of the regression equation are normally distributed. From the observation of the normal P-P plot of the standardized residuals, it can be seen that the diagonal line corresponds to a normal distribution with a mean of 0. The scatter points in the plot are closely scattered near the diagonal line and almost cover it, which again verifies that the residuals obey a normal distribution.

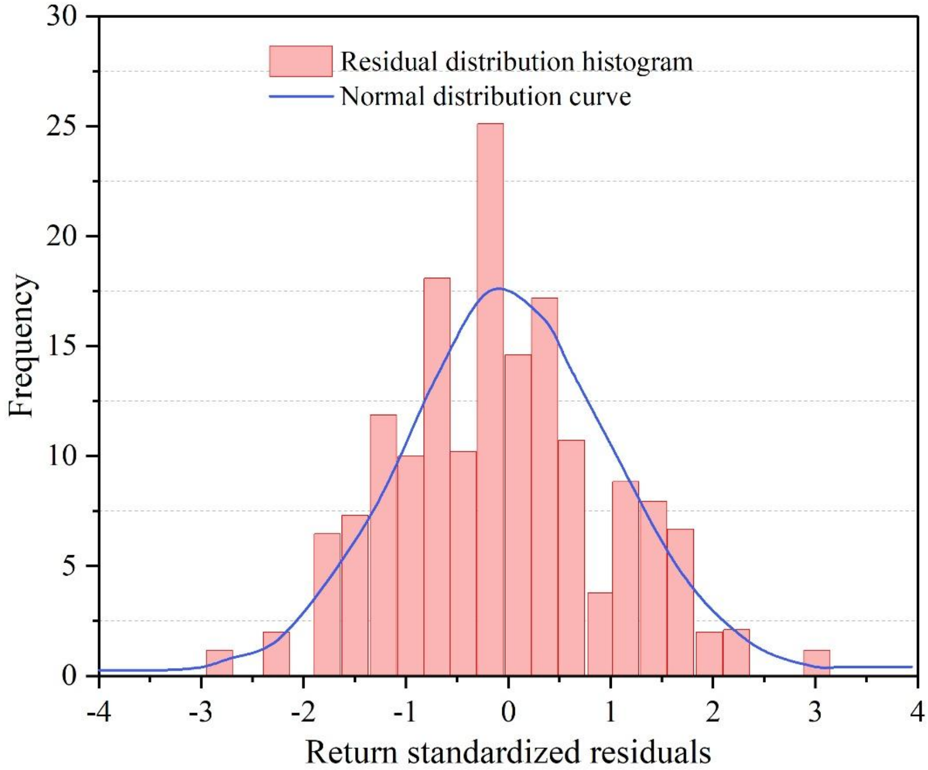


Figure 12. Residual distribution histogram and normal distribution curve.

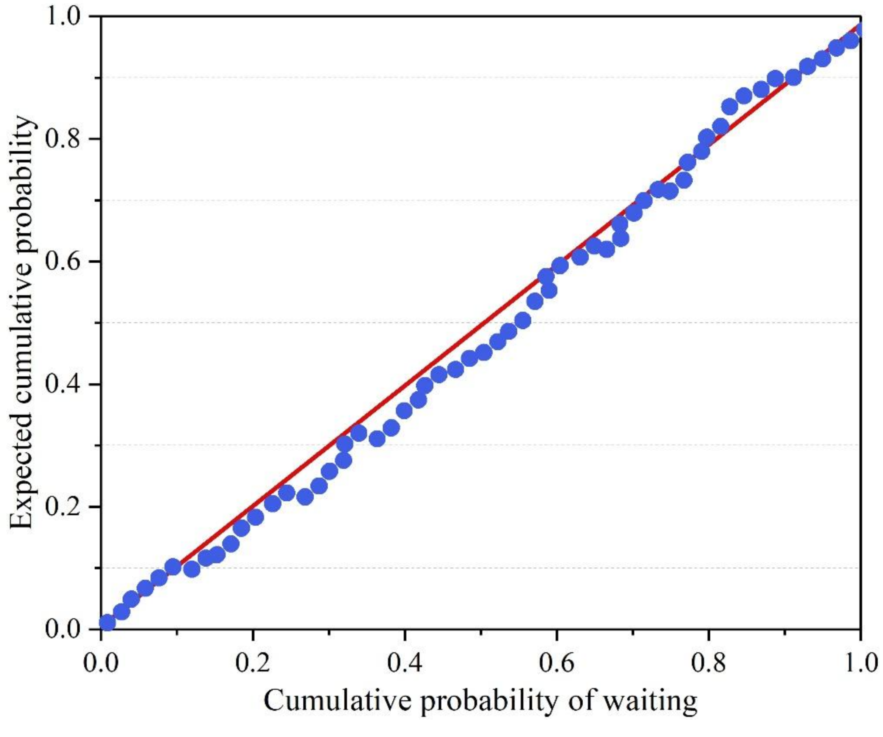


Figure 13. Standardized residual normal P-P plot.

4. Conclusion

This paper proposes a speech emotion recognition model that integrates two-way CNN-LSTM and attention mechanism to recognize the speech emotion changes of teacher-student interactions in the classroom of Civic and Political Education, and then combines with the multiple regression model to explore the correlation between the emotion changes of the students and the improvement of ideological and political quality in the process of Civic and Political Education.

The results show that: the accuracy of this paper's model for speech emotion recognition is better than the existing comparison model; the overall teaching interaction between teachers and students in the classroom presents a positive and active state, the students' emotional changes are rich, and the classroom as a whole forms a benign cycle with teacher-student interaction. Different ratings of classroom emotional changes corresponding to the overall classroom emotional atmosphere makes a difference, presenting the degree of emotional pleasure for the provincial excellent > city excellent, to a certain extent, affects the students' emotions, which in turn affects the enhancement of students' ideological and political quality. In addition, the emotional changes of taking responsibility consciously, resonance sublimation, identity internalization and quenching firmness will affect the students' ideological and political quality enhancement in the process of Civic Education to different degrees. For example, the effect of the emotional change of taking responsibility consciously on the enhancement of ideological and political quality is more obvious, while the effect of quenching firmness is smaller.

About the Author
Lan Jiang (1981.05-), female, Han ethnicity, Luohe, Henan Province, holds a postgraduate degree in Marxist Theory and Ideological-Political Education. She is an associate professor specializing in Marxist Theory and Ideological-Political Education.

References

1. Liu, J., & Lu, W. (2021). The Influence of Ideological and Political Education Ideas in College Courses on The Healthy Development of Students'Psychological Quality. *Psychiatria Danubina,* 33(suppl 6), 90-92.
2. Li, X., Gao, Y., & Jia, Y. (2022). Positive guidance effect of ideological and political education integrated with mental health education on the negative emotions of college students. Frontiers in Psychology, 12, 742129.
3. Zhang, X. (2025). Strategies to Improve Ideological and Political Education Quality in Higher Education Amid Comprehensive "Three-Way Holistic Education" Reforms. *Education and Social Work*, 1(1), 93-101.
4. Teunissen, P. W., & Bok, H. G. (2013). Believing is seeing: how people's beliefs influence goals, emotions and behaviour. *Medical education*, 47(11), 1064-1072.
5. Keegan, P. (2021). Critical affective civic literacy: A framework for attending to political emotion in the social studies classroom. *The Journal of Social Studies Research*, 45(1), 15-24.
6. Wang, Y., Jiang, G., Yao, Z., & Liu, L. (2024). The influence of teacher–student relationship on Chinese high school students' academic motivation for the ideological and political subject: the mediating role of academic emotions. Frontiers in Psychology, 14, 1329439.
7. Canova, R., & Ma'rof, A. A. (2023). Empowering Students With Ideological and Political Education and Bridging the Gap in Modern Democracies' Ideological Division. In Social Psychology Theories and Methods in Education (pp. 159-181). IGI Global.
8. Zha, W. (2022). Analysis of ideological and political education strategy based on college students'anxiety in the new era. Psychiatria Danubina, 34(suppl 4), 189-189.
9. Wu, J., & Zhang, H. (2022). The influence of the reform and innovation of college students'ideological and political education practice on students'spiritual and emotional anxiety. Psychiatria Danubina, 34(suppl 1), 631-632.
10. Su, L. (2021). New thoughts on ideological and political education in colleges and universities from the perspective of psychology. Psychiatria Danubina, 33(suppl 6), 272-274.
11. Zhang, Y., Yan, Y., Kumar, R. L., & Juneja, S. (2024). Improving college ideological and political education based on deep learning. International Journal of Information and Communication Technology, 24(4), 431-447.
12. Lin, K. (2025). Chinese ideology in the political education of students: how does ideology-based teaching impact students' consciousness?. Journal of the Knowledge Economy, 16(1), 4496-4515.
13. Zembylas, M. (2022). Affective ideology and education policy: Implications for critical policy research and practice. Journal of Education Policy, 37(4), 511-526.
14. Freeden, M. (2013). Emotions, ideology and politics. Journal of Political Ideologies, 18(1), 1-10.
15. Miao, X. (2025). Enhancing Ideological and Political Education Through Semantic Understanding and Sentiment Analysis: A Data-Driven Adaptive Learning Framework. International Journal of High Speed Electronics and Systems, 2540678.
16. Pingping, L. (2023). Application of Emotional Education in Ideological and Political Education of College Students. Application of Emotional Education in Ideological and Political Education of College Students, 5(3), 254-259.

17. Huang, H., & Liang, H. (2023). Ideological and political education combined with psychological care on patients with bipolar disorder. CNS Spectrums, 28(S2), S104-S105.
18. Robinson, M. D., Boyd, R. L., & Fetterman, A. K. (2014). An emotional signature of political ideology: Evidence from two linguistic content-coding studies. Personality and Individual Differences, 71, 98-102.
19. Qi, W., & Yu, L. (2024, December). Strategic Research on the Cultivation of Political Identity Literacy among College Students in University Ideological and Political Education. In 2024 7th International Conference on Humanities Education and Social Sciences (ICHESS 2024) (pp. 312-322). Atlantis Press.
20. Xuan, P. (2025). A Study on the Integration of Mental Health Education and Ideological and Political Education in Universities Under New Circumstances. *International Educational Research*, 8(2), p90-p90.
21. Zhou, Y. (2025). A Study on the Pathways for Cultivating College Students' Psychological Quality from the Perspective of Ideological and Political Education in Higher Education Institutions. *Advances in Curriculum Design&Education*, 1(2), 1-9.
22. Shen, Z. (2025). The Emotional Driving Effect of Role Model Narratives in Ideological and Political Education for College Students. Contemporary Education Frontiers, 3(3), 1-6.
23. Li, L., Hu, W., Tang, X., Wang, Z., & Cu, Z. (2019, April). Research on the Reform of Ideological and Political Education from the Perspective of" Core Literacy" and" Key Abilities". In The First International Symposium on Management and Social Sciences (ISMSS 2019) (pp. 259-263). Atlantis Press.
24. Dimple Tiwari, Bhoopesh Singh Bhati, Bharti Nagpal, Nazik Alturki & Lotta Bayisenge. (2025). Attention-augmented hybrid CNN-LSTM model for social media sentiment analysis in cryptocurrency investment decision-making. Scientific Reports,15(1),33201-33201. https://doi.org/10.1038/S41598-025-18245-X.
25. Hao Yu, Wan Xiaohua, Yan Rui, Liu Zhiyong, Li Jintao, Zhang Shihua... & Zhang Fa. (2022). VP-Detector: A 3D multi-scale dense convolutional neural network for macromolecule localization and classification in cryo-electron tomograms. *Computer Methods and Programs in Biomedicine*, 221,106871-106871. https://doi.org/10.1016/J.CMPB.2022.106871.