J. COMBIN. MATH. COMBIN. COMPUT. 127a (2025) 6909--6925

Journal of Combinatorial Mathematics and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Combinatorial Press

Computational Linguistic Interpretation and Modeling of Dress Metaphors in Ming Dynasty Novel Texts

Menghe Tian[1], Xiangyang Bian[1,✉]

[1] *Collage of Fashion and Design, Donghua University, Shanghai, 200050, China*

ABSTRACT

Dress metaphor is a very important way of expression in the novel text of Ming Dynasty, and the recognition and interpretation of the metaphor play a very important role in really understanding the novel text. This paper proposes a dress metaphor recognition model based on Transformer and graph convolutional neural network, and a dress metaphor interpretation method based on Seq2seq framework. The apparel metaphor recognition model performs feature extraction of global and local information of apparel metaphor sentences by Transformer. Graph Convolutional Neural Network is utilized to obtain syntactic structure information and sentence dependencies, in order to complete multi-word dress metaphor recognition. Then the obtained deep metaphor features and syntactic structure information of the sentence are input to the classification layer. The metaphor decoding method carries out costume metaphor understanding through the encoder-decoder, which chooses the LSTM network structure for both encoder and decoder to better obtain the semantic features of the novel text. The dress metaphor recognition model improved the recognition correctness on the dataset by 17.97% and 7.28%. The dress metaphor interpretation method based on the Seq2seq framework elaborates the interpretation content and can more accurately interpret the dress metaphors in Ming Dynasty novels. It verifies the practicality of the metaphor recognition and interpretation model in this paper in the task of interpreting dress metaphors in Ming Dynasty novel texts.

Keywords: Graph Convolutional Neural Network, Transformer, Seq2seq Framework, Dress Metaphor Recognition, Dress Metaphor Interpretation

1. Introduction

Dress is a manifestation of culture and a symbol of civilization. Roland Barthes is one of the earliest scholars to study dress aesthetically, and he emphasized that "only those dresses and dress phenomena that can represent something or a certain idea, and can be recognized and understood

✉ Corresponding author.
E-mail address: xiangyanbian@163.com (X. Bian).
Received 20 February 2024; Revised 15 May 2024; Accepted 20 November 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-384

belong to the category of symbols" [1]. Dress is part of the discourse of the times in the process of "humanization of nature" and "socialization of man". Clothing as a symbol of the wearer's social status, social identity, national identity, gender roles, interests and desires, as well as aesthetic habits of the outside of the continuous development and improvement of its function is no longer limited to the practical significance of covering the body, warmth, and prevention of various infringements, but became a symbol of the performance of a symbol, into in the manifestation [2-5]. And the depiction of clothing in literary works has also become the key for writers to shape the characters, and become a symbol of the character's personality and destiny [6].

In China, since the pre-Qin period, people have used clothes to send their thoughts and give them moral significance. For example, in the "Poetry - Zheng Feng - Dicky", the clothes of a man who studied in ancient times and the ribbon of a jade belt were used to convey the thoughts of a lover [7]. Generally speaking, the "Hundred Clothes" in the pre-Qin period injected rituals, ethical indoctrination and political connotations into the clothing that initially covered the body to avoid shame and regulate warmth and coldness [8]. Qu Yuan also gave the dress "clean and fragrant" aesthetic meaning, and linked the dress with the beauty of personality. Therefore, as an artificial object, the dress has a connotation beyond the material, with both material and spiritual attributes, and becomes a good "tool" for constructing philosophical thoughts and a "frequent guest" in literature, which has a profound influence on the literature of the later generations. Among the many ancient literary works, novels used costumes frequently and spread deeply. The Ming Dynasty was the rise of chapter and verse novels, and the classics include Romance of the Three Kingdoms, Water Margin, Journey to the West, etc. [9]. In the late Ming Dynasty, with the development of urban economy and the progress of the printing industry, in order to suit the needs of urban residents to read, reflecting the urban life of the popular novels are becoming more and more complicated, this period of time only long novels have more than one hundred kinds of out. In addition to "Three Words and Two Beats", the famous ones are "Plum in the Golden Vase", "Feng Shen Yan Yi", "East Zhou Li Guo Zhi", "Yang Jia Gong", etc. The social status and aesthetic value of these novels are remarkable [10-13]. The Ming Dynasty was a period of epoch-making development of Chinese classical literature, which reflected the new social and cultural life after the rise of the civic class, and these novels were representative and typical, and involved a rich variety of dresses, including male dresses, female dresses, hu clothes, and monastic and Taoist dresses, as well as the main body of dressmaking, craftsmanship, and dress texture [14]. Clothing in the Ming Dynasty has different application scenes, which is the witness of Ming social life and the characterization of aesthetic interests. The creators of novels used costumes partly with the symbolic meaning of the costumes themselves, and partly through artistic techniques to give new meanings to the costumes, which is essentially a kind of metaphorical thinking and expression [15-16]. Through the interpretation of clothing metaphors in novels, it can reveal the character, social status, cultural background and other information in the novels.

In this paper, a metaphor recognition model based on Transformer and GCN and a metaphor understanding model based on Seq2seq framework are constructed. The metaphor recognition model re-specifies the concept of local context to obtain fine-grained local context. The dependencies between words and syntactic structure information are obtained using graph convolutional neural network, and the two are fused and input to the classification layer. The LSTM network structure and attention mechanism are introduced in the metaphor comprehension model based on the Seq2seq framework to optimize the ability of the encoder and decoder, which solves the problem of information loss in the compression process. Experiments are conducted to explore the effectiveness and validity of the two models in the task of recognizing and understanding metaphors of Ming Dynasty novel costumes.

2. Current status of research on clothing metaphors

2.1. Clothing Metaphors

Dress, as an artificial cultural creation, is the result of interaction between the social and cultural background to which people belong and their own needs and ability to make choices, on the premise of objective prior natural material resources and ecological environment. One of the characteristics of modern dress style is that dress behavior has shifted from sameness to difference. However, the way of dressing will never become a purely private behavior, because it has a clear and continuous tendency of collective representation. Inside individual dress, there are always hidden social memories, values and cultural connotations that are shared by a certain community, group or class. Experiencing and representing something by means of something else is qualitatively compatible with the principle of metaphor. This characteristic of dress can therefore be called the metaphorical rule of dress.

Functionally, clothing is not only for human survival to protect the body from external injuries, but also to meet the needs of human life, providing comfort, convenience and aesthetics. Therefore, at the level of existentialism, clothing is the physical manifestation of human life state, and also a cultural statement of certain social restrictions and demands. The interpretation of such a complex social text as clothing, and the sorting out and interpretation of its metaphorical rules can only be carried out in a multi-dimensional research space, and cannot be cut and categorized in a simple way.

Clothing is an essential element of human daily life. Clothing is dualistic, with practical functions and at the same time an important symbol of human civilization, featuring social, aesthetic, ethical and local characteristics. The novels of the Ming Dynasty reached an unprecedented height in depicting the world's situation and fully demonstrated the dress life of the Ming Dynasty. The parasitic lyrics and songs in the Ming novels also became an important carrier of the costume life of the Ming Dynasty. There are two sources of parasitic words and songs in Ming novels: one is the author's own words and songs. The second one is the copying of other people's compositions. There are two types of paraphrase: one is the former dynasty's paraphrase. The second is the current dynasty lyrics and music. The author's own lyrics and music and the inherited words and music of the dynasty can become the “mirror” of the social life of the Ming Dynasty. Since the social life is very extensive, this paper intends to take the author's self-created and inherited parasitic lyrics and songs of the current dynasty in Ming novels as “mirrors”, and focus on the most common and characteristic dress life in the Ming Dynasty, so that we can get a glimpse of the whole “panther”, and reveal the characters' temperament under the metaphor of dress. This will reveal the characterization of the characters in Ming novels under the metaphor of clothing.

2.2. Application of Machine Learning in Metaphor Recognition

In the past decade, many scholars have conducted research on costume metaphor recognition using machine learning methods. Various studies usually start from one or several metaphor types, which can be categorized into dress noun metaphors, dress verb metaphors and dress adjective metaphors in terms of the lexical types of dress descriptions. Classified in terms of granularity, under the divisions of word level, phrase level, sentence level, paragraph level and article level, the studies of many scholars mainly focus on phrase level and sentence level. By calculating the similarity between nouns or utilizing machine learning methods to identify whether a noun is its literal or metaphorical usage. Those researching for dress verb-based metaphors determine whether it is a dress metaphorical usage by using indicators such as common verb collocation and whether it violates the preferred semantic choice of the verb. The metaphor recognition oriented to dress adjectives involves a cross-linguistic metaphor recognition method, which is of great significance in the directions of dress metaphor interpretation methods and machine translation in the application domain.

3. Transformer and graph convolutional neural based metaphor recognitio

This chapter first introduces the phenomenon of word metaphorical expression and multi-word metaphorical expression in metaphorical expression. Then a metaphor recognition model based on Transformer [17] and graph convolutional neural network is proposed, which is analyzed in depth from three aspects: contextual semantic computation based on Transformer, representation of sentence dependency information based on graph convolutional neural network, and metaphor recognition based on fused semantic vectors.

3.1. Analysis of Metaphor Recognition Problems and Solution Ideas

A costume metaphor is a non-literal expression. Costume metaphors have been widely used in all types of literature and writings, especially in fiction. Costume metaphors can effectively convey to readers the complex sensations, emotions, and visuals that appear in fictional texts through costume depictions.

In this paper, we propose a dress metaphor recognition model based on graph convolutional neural network and Transformer, which can effectively recognize dress metaphors as well as solve the metaphorical expressions jointly triggered by multiple words in order to recognize dress metaphors in novel texts of the Ming Dynasty. The framework of the model is shown in Fig. 1, firstly, the semantic representation of words is obtained by using the RoBERTa [18] model, and the lexical information of the words is added as supplementary information, which is inputted into the Transformer structure to extract the global contextual information and the local contextual information of the sentence, and then these two kinds of contextual semantic information are aggregated by using multi-head attention. Secondly, the dependency information graph is constructed using the dependency syntax tree, and the syntactic information representation of words is obtained by learning the dependency relationship between words through graph convolutional neural network to extract the potential metaphorical feature information. Finally, the contextual semantic representation and syntactic information representation of the target word are fused and inputted into the classification layer, and the probability distribution is used to determine whether the target word is a metaphorical expression or a non-metaphorical expression.

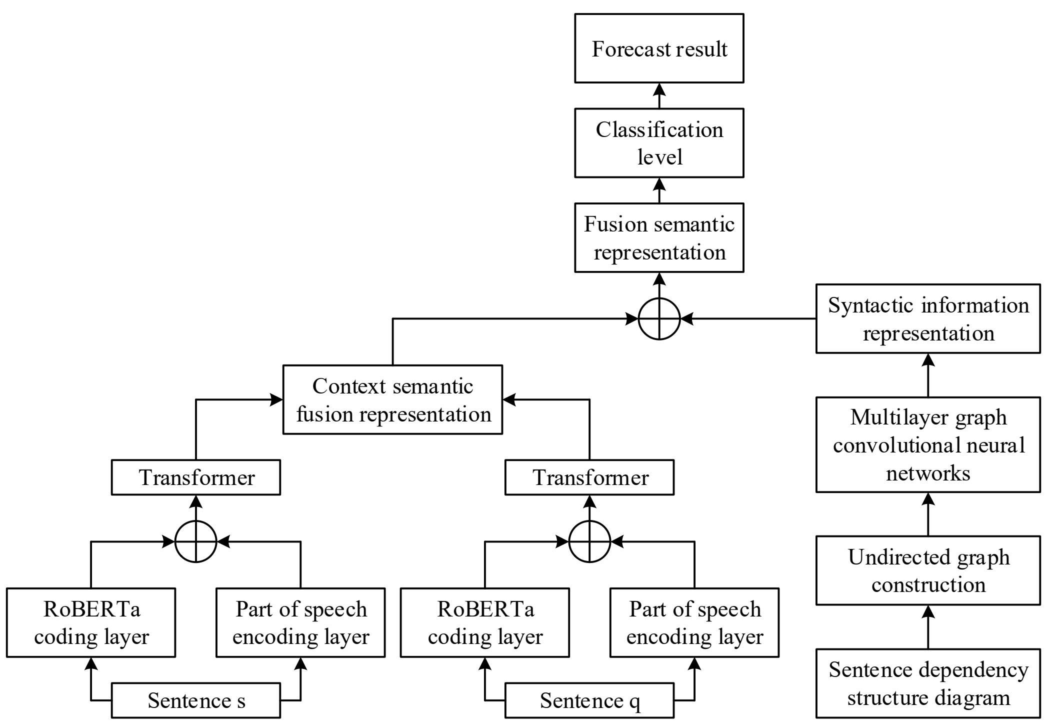

Fig. 1. Schematic diagram of metaphor recognition model

3.2. Transformer-based contextual semantic computation

Considering only the concept of local context has some shortcomings. If it is a long sentence without punctuation splitting, then the local context may become the global context, which obviously does not conform to the concept of local context. Therefore, the model obtains a fine-grained local context by limiting the window size k and thus standardizing the concept of local context.

Given sentence $S = \{w_1, w_2, \cdots, w_n\}$ and target word w_t, clause $q = \{w_t - k, \ldots, w_t, \ldots, w_{t+k}\}$ is first constructed within a window of size k centered on target word w_t.

The meaning of words may change subtly between different contexts, and RoBERTa is used to generate contextualized word representation vectors to capture the contextual semantic information features of words.RoBERTa model has shown very excellent results in various natural language understanding tasks. The robertalarge-mnli is used on the English dataset and the roberta_zh is used on the Chinese dataset.

First, word embeddings are encoded for each word in sentence s and clause q using RoBERTaEmbedding, which converts character representations into computable real-valued vectors. Second, the coarse-grained lexical labels and fine-grained lexical labels of the words are obtained by spaCy, and the two lexical labels are encoded in a unique hot code. The encoded lexical label vectors are then spliced with the word embedding representation vectors to obtain the final word vector representations x_i. Sentences s and clauses q after encoding are denoted by C_s and C_q, respectively.

$$x_i = RoBERTaEmb(w_i) \oplus e(POS) \oplus e(FGPOS) \tag{1}$$

where RoBERTaEmb(w_i) denotes the contextual word vector representation of word w_i encoded by the RoBERTaEmbedding layer, and $e(POS)$ and $e(FGPOS)$ denote the coarse-grained lexical labeling vectors and fine-grained lexical labeling vectors obtained by spacy.

In order to obtain richer contextual semantic information, two Transformer structures are used to extract the global contextual features of target word w_t in C_s and the local contextual features of target word w_t in C_q, respectively, to obtain the global contextual semantic representation and the local contextual semantic representation. The two Transformers share the weight parameters during training, so that the model not only learns the global context information and local context information of the target words from different perspectives, but also avoids the double storage of weight parameters.

3.3. Sentence Dependency Information Representation Based on Graph Convolutional Neural Networks

Numerous studies have shown that various natural language processing tasks have shown some improvement in results after using graph convolutional neural network models. While traditional natural language processing using serialization modeling does not tap into the nonlinear and complex relationships in a sentence, graph convolutional neural networks are able to do so effectively.

The graph convolutional neural network mines the potential metaphorical information of the target word in the sentence from the dependency structure of the sentence, and uses this information to assist in determining whether the target word is a metaphorical expression or a non-metaphorical expression. Conveniently using the graph structure can effectively mine the complex semantic relationships in the sentence. First, the dependency analysis of the sentence is performed by spacy to obtain the information of the dependency structure of the sentence. The dependency structure of the sentence is used to obtain the dependency relationship between words, and an undirected graph is constructed through this dependency relationship. The potential dress metaphorical feature information of the target word in the sentence is further mined by graph convolutional neural network.

For the undirected graph $G=(V,E)$, V is the set of vertices, i.e., each word in the sentence is a vertex in graph G, and E is the set of edges, including the dependencies between words in the sentence and the self-loops. The graph convolution process is shown below, except for the first layer $(i=1)$, the outputs of layer i are derived from the output of layer $i-1$ after the convolution operation.

$$\tilde{A}=D^{-\frac{1}{2}}AD^{-\frac{1}{2}} \tag{2}$$

$$H^1=\sigma(\tilde{A}XW^1) \tag{3}$$

$$H^i=\sigma(H^{i-1}W^i) \tag{4}$$

3.4. Metaphor Recognition Based on Fused Semantic Vectors

The attention mechanism [19] is able to efficiently select the most important elements in a given homogeneous representation while minimizing information loss. Using the multi-head attention mechanism, the model can obtain the feature representation of the dress metaphor semantic information from different perspectives.

The process of contextual semantic vector fusion is shown below:

$$Q^i,K^i,V^i=W_qh^{i-1},W_kh^{i-1},W_vh^{i-1} \tag{5}$$

$$S^i=softmax(\frac{Q^iK^i}{\sqrt{d_k}}) \tag{6}$$

$$Attention(Q,K,V)=S^iV^i \tag{7}$$

$$head_j=Attention(QW_q^j,KW_k^j,VW_v^j) \tag{8}$$

$$MultiHead(Q,K,V)=Concat_{j-1}^n(head_j)W_o \tag{9}$$

where Q^i,K^i,V^i is the query matrix, key matrix and value matrix of the ind block matrix, h^{i-1} is the input hidden state, W_q,W_k,W_v,W_o are the attention mechanism weight matrix, d_k is the scaling factor that counteracts the effect of overgrowth of the dot product, S^i is the scoring matrix of the i th block attention matrix, $head_j$ is the jth attention header i.e., the matrix after the fusion of the value matrix and attention scoring matrix, and the function Concat is the splicing operation.

$$\alpha=MultiHead(C_s,C_q) \tag{10}$$

$$m=Concat(\alpha,H) \tag{11}$$

where α represents the mixed semantic representation of the target word, H is the output result of the graph convolutional neural network i.e. the syntactic information representation of the target word, and m is the fusion vector representation of the mixed semantic representation and syntactic information representation of the target word.

The multi-head attention mechanism can obtain the deep dress metaphorical feature representations in C_s and C_q through different angles.The Transformer structure may neglect the sentence dependent structure features when extracting the sentence context information.By splicing α and H, the deep metaphorical feature representations of the target word are preserved and the sentence dependent structure features are supplemented.

Finally, the fused semantic vectors are fed into a fully connected neural network (MLP) to obtain the classification results of the target words. The model parameters are trained by the Adam optimizer with adaptive learning rate, and the objective function is the cross-entropy loss function L. The loss function is continuously minimized through a finite number of iterations so that the model achieves the best results. The classification layer function and cross entropy loss function L are shown below:

$$\overline{y} = softmax(MLP(m)) \tag{12}$$

$$L = -\sum_{i=1}^{N} \{y_i \log(\overline{y}_i) + (1 - y_i)\log(1 - \overline{y}_i)\} \tag{13}$$

where, N is the number of training samples, y_i is the true label of the target word, and $\overline{y}_i$ is the prediction result of the target word.

4. Metaphorical understanding based on the Seq2seq framework

On the basis of Chapter 3, this chapter proposes a Seq2seq-based metaphor comprehension method. Metaphor recognition is the prerequisite for metaphor comprehension, and the metaphor recognition model accurately identifies the dress metaphors in the text of the Ming Dynasty novels, and then further analyzes the meanings of the dress metaphors through the metaphor comprehension method. Metaphor understanding then provides deep contextual support for metaphor recognition. Both methods in this paper contribute to the interpretation of dress metaphors in Ming Dynasty novel texts.

Costume metaphor understanding is an important task in metaphor computing. Current research in metaphor computing mainly focuses on metaphor recognition, and most of the existing metaphor understanding methods are based on semantic selection. This chapter proposes a deep learning based approach for metaphor understanding. This chapter will specifically introduce the method of metaphor understanding based on the Seq2seq framework, which first summarizes and analyzes the structure of the Seq2seq model, then introduces the attention mechanism to improve the ability of decoding, and describes the structure of the model network and algorithms used in the experiments.

4.1. Seq2seq model

Seq2seq is a neural network framework for a class of sequence transformation problems, i.e., an encoder-decoder sequence transformation framework, where the encoder and decoder generally use a recurrent neural network framework. The main role of the encoder is to convert the input sentence sequence into a fixed-dimension text vector representation, and the main role of the decoder is to convert the fixed-dimension text vector converted by the encoder into the output sentence sequence.

The input to the decoder at each moment is the output vector of the previous moment and the vector of the encoder's output at the last moment. The basic principle is as follows: given an input sentence X, the decoder-encoder framework generates sentence Y. X and Y are represented as sequences of words, respectively. The encoder encodes the sentence and transforms the input sentence into a semantic vector C by a nonlinear transformation:

$$C = F(x_1, x_2, ..., x_n) \tag{14}$$

Then, the word to be generated at moment t is obtained by the decoder based on the semantic vector C of sentence X and the generated Y:

$$y_t = G(C, y_1, y_2, ..., y_{t-1}) \tag{15}$$

When the lengths of the input and output sequences are different, LSTM can better capture the contextual information. Since this chapter focuses on solving the costume metaphor understanding problem, in which the lengths of the input and output sequences are not necessarily the same. Therefore, this chapter uses the LSTM-based Seq2seq model for costume metaphor understanding.The encoder and decoder of the Seq2seq framework use the LSTM model, which enables a complete stream from input to output, i.e., the state obtained at the end of the encoding layer is inputted into the first cell memory unit of the decoding layer, so that from an overall point of view, there is a complete linear data stream from input to output. This way, from the overall point of view, there is a complete linear data flow from input to output.

When given an input sequence X and an output sequence Y, where the lengths of X and Y are $L, L^{'}$, the probability of final output Y is:

$$p(y_1,....,y_L \mid x_1,...,x_L) = \prod_{l=1}^{L} p(y_l \mid v, y_1,..., y_l) \tag{16}$$

When the text is generated and decoded, the output sequence with maximum probability needs to be found out, and the decoding layer applies a greedy search algorithm to find out the optimal solution.The Greedysearch algorithm finds out the output with maximum probability of decoding each word of the input until the sentence reaches the maximum length or a terminator appears.

In the Seq2seq framework, the input and output parts are the input metaphorical sentence and the result of understanding the metaphorical sentence, the input sequence is encoded by the neural network to get the final state, which is input to the decoder as the initial state, and the optimal result is searched by the Greedysearch algorithm, that is, the optimal result of the apparel metaphor understanding.

In the encoder-decoder framework, the input sentence sequence is compressed into a fixed-size vector by the encoder, and the compression process results in the loss of information, which is solved by the attention mechanism, which does not focus only on the semantic interrogation in the encoding stage, but also on the part of information related to the target vocabulary. Instead, the encoder encodes the input text into multiple semantic vectors to form a series of semantic vectors, and when decoding, a subset of this series of semantic vectors is selected for decoding according to the temporal order, which makes it possible to use more textual features in the decoding stage. The weights are first calculated based on the features of the encoder and decoder, and then a weighted summation value is calculated for the encoder's features, which is used as an input to the decoder.The Seq2seq model incorporates an attention mechanism, where the encoder calculates an attention probability distribution for each word.

With the addition of the attention mechanism to the Seq2seq framework, the coding layer is a bi-directional recurrent network (Bi-RNN), including a forward recurrent network and a reverse recurrent network. The input sequence is $X = \{x_1, x_2,..., x_T\}$. The forward recurrent network is fed sequentially in chronological order and computes the hidden layer states, and the reverse recurrent network passes through in back-to-front order and then computes the states of each hidden layer. In this case, the vector of each word in the input sentence associates the states of the previous and next hidden layers so that they contain the surrounding word features.

In this chapter, on the dress metaphor comprehension task, the encoder uses a bi-directional LSTM with better embedding. The decoding process uses the LSTM model and the local optimal solution algorithm is used to compute the point with the highest probability, as shown in equation (17). s_i is the hidden layer state of the neural network at the moment t. c_i is the semantic vector of the context, and the weighted sum is calculated by Eq.

$$P(y_t \mid y_1,..., y_{t-1}, x) = g(y_{t-1}, s_t, c_t) \tag{17}$$

$$s_t = f(s_{t-1}, y_{t-1}, c_t) \tag{18}$$

$$c_t = \sum_{j=1}^{r_i} \alpha_{tj} h_j \tag{19}$$

The encoder is a bidirectional long and short memory network. Its basic idea is that the forward and backward directions are two long and short memory neural networks respectively and both are connected to an output layer. Forward LSTM computes the hidden layer state with front-to-back sequential inputs. The reverse LSTM computes the hidden layer states with back-to-front sequential inputs, and each word connects the forward hidden layer state and the backward hidden layer state, and therefore, contains the contextual information of the word.

$$s_t = f(U * x_t + W * s_{t-1}) \tag{20}$$

$$s_t^{'} = f(U^{'} * x_t + W^{'} * s_{t-1}) \tag{21}$$

where U, V and W are weight matrices and f is the activation function of LSTM. The final output depends on s and $s^{'}$, where g is the activation function of the LSTM.

$$o_t = g(V * s_t + V * s_t^{'}) \tag{22}$$

4.2. A Framework for Metaphorical Understanding Systems

In this chapter, a LSTM based Seq2seq model for metaphor understanding is designed. This model can encode and decode the input metaphorical sentences, and metaphor understanding can be better understood by utilizing this model for metaphor understanding, and the system flowchart is shown in Fig. 2:

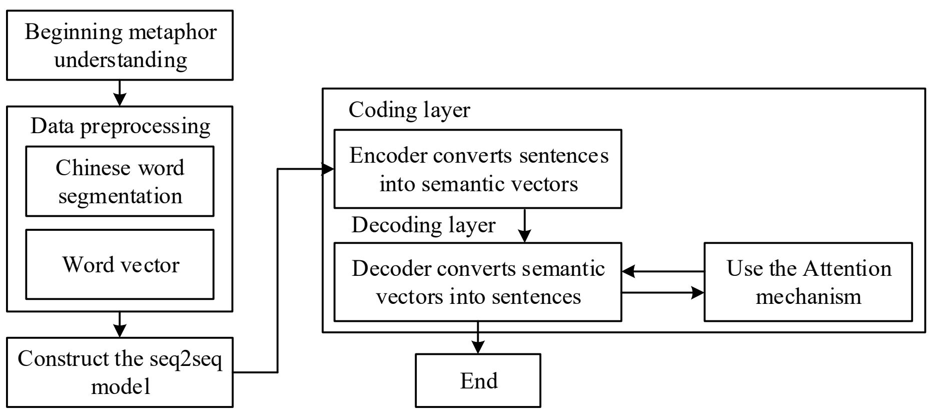

Fig. 2. Metaphor understanding flow

The metaphor comprehension corpus is firstly disambiguated, and then a dictionary of word-index mappings is constructed, while special characters are added.

When the model is trained, the encoder and decoder generate output vectors separately, where the decoder outputs word by word, and all the generated results of the decoder and the output of the encoder are added to a list, and together they are passed as the output of the regression training to update the parameters of the network, and finally the model is generated. In the prediction phase, the model gets the optimal result by the greedy algorithm when the test sentence is input.

Since in traditional recurrent neural networks, the state of the hidden layer of the encoder is passed from front to back. However, in practice, the output sequence of the current moment does not only contain the hidden layer state of the previous moment, but is also related to the semantics of the subsequent moment. Therefore, the encoder in this paper uses a bidirectional LSTM, which can better obtain contextual semantic features. In the decoder, we introduce the attention mechanism can make the semantic vector contain longer distance information and allow the decoder to assign different weights to the hidden layer state at different moments, and the decoder uses LSTM.

5. Clothing Metaphor Recognition Results and Analysis

5.1. Representation in different metaphors and genres of fiction

In this chapter, the Chinese metaphor corpus is selected as the test dataset for this paper, which is used for the study of the recognition of dress metaphors in Ming Dynasty novel texts. The performance of the proposed Transformer and graph convolutional neural-based metaphor recognition model is compared with the baseline on this test dataset, and the data with the largest number of dress metaphors present in the dataset is also used, and the experimental results of the model on the dress metaphors data are reported. The modeling results are reported on the clothing metaphor data. The choices in the paper include “square scarf”, “wide sleeve”, “long skirt”, “lapel shirt”, The data source is the corpus of the Chinese Linguistics Research Center of Peking University. Sentences containing these six verbs are randomly selected by searching and manually annotated, and the sentences are labeled as metaphorical expressions and regular expressions, in which metaphorical sentences were selected for experimentation only as single metaphorical expressions of words. The experimental data are shown in Table 1 (where Baseline refers to the proportion of metaphorical sentences in the sentences).

Table 1. Experimental results contrast

Costume metaphor	Sentence number	Metaphorical sentence	Baseline/%
Square towel	78	44	56.41
Wide sleeve	77	45	58.44
Long skirt	79	47	59.49
Cardigan	74	39	52.70
Uniform	72	41	56.94
Round collar	74	49	66.22
Total amount	454	265	58.37

In order to show the effectiveness of the experiments, two groups of comparative experiments are designed in the paper.The first group adopts the maximum entropy model (MEM) for the recognition of dress metaphors. The second group adopts the method based on Conditional Random Field Model (CRF), and the correct rate of adopting classification and the accuracy and recall of metaphor recognition are used as the evaluation criteria of the experimental results, and the experimental results are shown in Tables 2 and 3, respectively.

In terms of the correct rate of classification, comparing the results of this paper's method with those of the MEM model, it can be seen that the overall effect of the paper's method is obviously better than the experimental effect of the MEM model, except that the correct rate of the word "knitting" is slightly lower, which is 11.2% higher than that of the MEM model, and that the correct rate of the other five words of the dress metaphor category is obviously higher than that of its maximum entropy and conditional random field model results. Except for the word "knitting", which is slightly lower, with an improvement of 11.2% compared with the MEM model, the classification rates of the other five clothing metaphors are significantly higher than those of the maximum entropy and conditional random field models. This is due to the fact that these two methods only provide simple statistics on the occurrence of words in a sentence, and cannot directly represent the semantic relationship of the clothing metaphor words in a sentence, whereas the model in this paper integrates semantic vectors, which can well represent the semantic information of the words, so that the text has certain semantic features, which makes the classification and identification better than the Maximum Entropy and Conditional Random Field Models. The experiments in the paper also show obvious advantages in accuracy, recall and F-value, and the F-value of each group of experiments is improved, up to 81.82%.

Table 2. The experimental results were compared with the comparison experiment

Costume metaphor	MEM/%	CRF/%	This method/%
Square towel	67.54	77.49	78.74
Wide sleeve	58.72	69.98	77.52
Long skirt	68.76	83.75	88.75
Cardigan	59.94	76.26	78.78
Uniform	67.51	75.03	78.71
Round collar	68.27	76.25	79.99
Mean value	65.12	76.46	80.42

Table 3. Precision Recall rate and F value

Costume metaphor	MEM/%			CRF/%			This method/%		
	P	R	F	P	R	F	P	R	F
Square towel	67.83	92.71	75.24	81.27	80.29	78	80.16	91.51	83.79
Wide sleeve	66.67	39.49	44.53	72.38	83.35	74.43	90.06	77.36	81.41
Long skirt	63.9	95.81	76.17	78.27	87.34	79.63	86.68	78.01	80.45
Cardigan	69.98	90.59	77.21	85.02	89.34	86.6	83.73	98.76	89.57
Uniform	83.52	47.85	57.13	75.68	77.14	75.44	76.94	84.2	78.85
Round collar	41.17	42.51	40.09	70.85	76.05	70.14	80.83	80.48	76.67
Mean value	65.54	68.14	61.77	77.25	82.23	77.37	83.1	85.07	81.82

Ming Dynasty novels are rich and diverse in genres, this paper selects more representative historical novels, heroic and legendary novels, divine and magical novels, worldly affairs novels, public case novels, and proposed novels in the corpus of the Chinese Language Research Center for the study of costume metaphor recognition of the text, and the same comparative model mentioned above is adopted to compare with the model of this paper, and the correct and accurate rate of classification of the different novels, the R-value, and the F-value are shown in Table 4 and Table 5, respectively. The results are shown in Table 4 and Table 5 respectively. From the point of view of the correct classification rate of different novel types, the overall effect of the clothing metaphor recognition method in this paper is significantly better than the experimental effect of the MEM model and the CRF model. Especially in the recognition of clothing metaphors in romance novels, the model in this paper performs very well, and the correct rate of clothing metaphor recognition is improved by 17.97% and 7.28% compared with the MEM model and CRF model, respectively. This is due to the fact that there are more romance novels based on social reality, especially family life, which portray the world, and involve more depictions of clothing metaphors of characters in the family, mostly projections of characters' character, rank, identity and roles. And the language style of this kind of novels is usually rich in metaphor and symbolism, in which dress metaphors are widely used and semantics are more stable. This stylistic feature makes the model of this paper, which integrates semantic vectors, able to better recognize and understand the contextual semantics of dress metaphors. For example, “gorgeous clothes” in romance novels may be a metaphor for a character's vanity and status, and the model can capture this metaphor more accurately through the fusion of semantic vectors.

Table 4. Performance of model data on different novel types

Costume metaphor	MEM/%	CRF/%	This method/%
Historical novel	67.86	77.08	79.33
Heroic novels	58.78	70.28	76.88
Magic novel	68.66	83.37	88.55
World fiction	73.56	84.25	91.53
Public fiction	67.66	75.57	79.01
Quasi fiction	68.48	76.74	80.06
Mean value	67.50	77.88	82.56

Table 5. Performance of model data on different novel types

Costume metaphor	MEM/%			CRF/%			This method/%		
	P	R	F	P	R	F	P	R	F
Historical novel	92.21	68.56	93.72	74.94	81.15	79.36	78.17	80.89	91.27
Heroic novels	38.72	67.62	39.86	44.44	72.95	82.86	74.9	89.99	76.74
Magic novel	95.68	63.43	92.32	75.52	78.37	87.22	79.84	86.56	77.54
World fiction	90.55	69.65	95.53	89.65	84.8	96.08	92.35	84.34	98.73
Public fiction	47.16	83.34	48.05	57.11	75.99	77.37	75.31	76.67	84.36
Quasi fiction	43.02	42.13	41.7	39.66	71.1	75.68	69.49	80.2	80.53
Mean value	67.89	65.79	68.53	63.55	77.39	82.10	78.34	83.11	84.86

5.2. Performance of the method in different models

In order to illustrate the effectiveness of the proposed method in this chapter, this section adds the proposed method to the previous model, i.e., without changing the model, the multi-head mechanism is used to obtain the feature representation of the metaphorical semantic information from different angles.

1) *DiLSTM: This method adds the multi-head attention mechanism after the context encoder layer of DiLSTM.

2) *LSTM-CRF: This method adds the attention mechanism to the input and LSTM layers.

3) *CNN-LSTM: This model replaces the embedding layer of CNN-LSTM with a multi-attention mechanism.

The experimental results are shown in Table 6, which show that the F-score of the model improves by 3.58%, 5.85% and 3.13% respectively after adding this paper's strategy compared to the previous dress metaphor detection model. In the previous methods, they ignored the multi-level semantic information of words and phrases. The proposed model in this paper can use content information from multiple inputs and utilize the multi-attention mechanism to learn the importance of each word, which effectively mines the hierarchical features of the text to obtain richer semantic features. Overall, compared with the previous models, the application of this strategy helps greatly in improving the model performance, proving the effectiveness of the proposed strategy.

Table 6. Adds performance comparisons after different strategies

Model	P	R	F	Inc-F
DiLSTM	51.13	64.44	56.97	-
DiLSTM*	55.73	66.48	60.55	3.7
LSTM-CRF	64.42	45.91	53.53	-
LSTM-CRF*	68.61	52.29	59.33	5.8
CNN-LSTM	60.78	70.04	65.18	-
CNN-LSTM*	65.93	70.83	68.31	3.3

5.3. Effects of Hierarchical Characteristics

In addition, to further illustrate the effectiveness of the proposed model, this section conducts a set of experiments under different settings, i.e., comparing the performance of the model using only word-level features, only sentence-level features, and both features at different number of iterations. For each iteration, the experiments give the performance of the model in the test set. In all these experiments, the rest of the model was kept constant.

The results of the experiments are shown in Figure 3, which shows that the performance of the model using both word-level features and sentence-level features is optimal. This is because by using two levels of features, the model can learn deeper semantic representations and fully utilize the semantic information of the text. In addition, the number of iterations of the model proposed in this chapter is set to 70 since it is observed that the model achieves optimal performance when the number of iterations is 70.

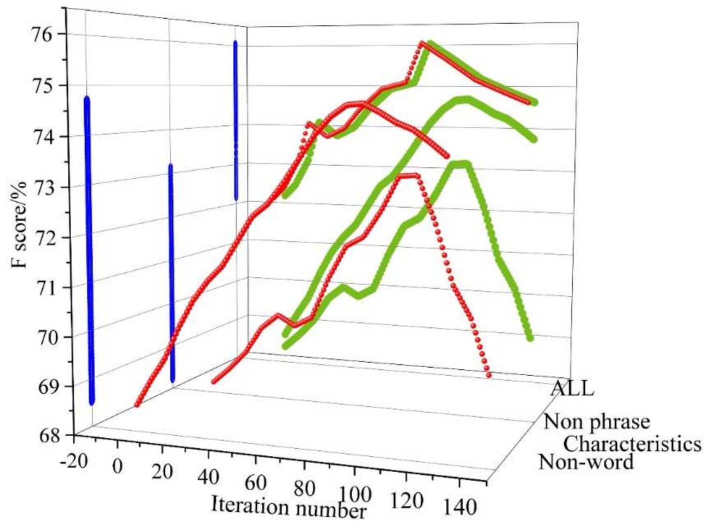

Fig. 3. The experimental results of the model at different Settings

6. Results and analysis of clothing metaphor interpretation

6.1. Quality Analysis of Clothing Metaphor Interpretation

We tested the performance level of this paper's model and the masked language interpretation model (MLD) in the MOH X and TroFi datasets using the metaphor types of dress metaphors such as metaphors of rank and identity, metaphors of culture and etiquette, metaphors of gender and role, metaphors of ornamentation and symbolism, and metaphors of region and fashion, and the performance results are shown in Table 7. This paper's model is able to generate longer responses in all the five types of dress metaphors, and the average length of the interpretations generated by this paper's metaphor interpretation model is 84.84 and 89.52 in the MOH X and Tro Fi datasets respectively, which is 17.1 and 16.14 higher than that of the interpretations generated by the masked linguistic paraphrase (MLD) model.The responses are not only the basic interpretations and analyses of the sentences, but they also These responses are not only the basic interpretation and analysis of the sentences, but also contain deeper metaphorical content of the costumes. The length and depth of the responses are attributed to the metaphorical understanding approach of the metaphor interpretation model based on the Seq2seq framework, which not only guides the model to learn sufficiently, but also ensures that the generated responses reflect a true understanding of the clothing metaphor phenomenon, confirming that the responses are not illusions generated by the model, but real insights into the language.

Table 7. Generate quality analysis

Type	MOH X		Tro Fi	
	MLD	Based on the Seq2seq model	MLD	Based on the Seq2seq model
The metaphor of rank and identity	66.9	82.6	71.5	88.9
The metaphor of culture and etiquette	67.9	85.4	72.6	91.2
The metaphor of gender and role	68.9	86.1	74.1	87.6
The metaphor of the grain and the meaning	65.6	84.2	73.6	89.3
Geographical and fashion metaphor	69.4	85.9	75.1	90.6
Mean	67.74	84.84	73.38	89.52

6.2. Results of Clothing Metaphor Interpretation

Table 8 shows some of the paraphrasing results of this paper's dress metaphor paraphrasing model and masked language paraphrasing (MLD) model on the multi-word dress metaphor paraphrasing task, where the context of the last sample depicts the Tang monk's dress and outlines the image of the Tang monk. The metaphor model based on the Seq2seq framework performs sequence modeling of the context in the calculation of contextual cooperation strength, which can capture the sequence information of the sentences, and therefore can better interpret chapter-level metaphors. The final result is “the Tang monk carries the mission and piety of Buddhism”. The interpretation result obtained by the mask language interpretation model is more superficial, i.e. “the Tang monk is very pious”.

Table 8. The meaning of the metaphor of the dress

Input content	The interpretation of the MLD method	The interpretation of this method
After the doctor added, Simon qing rode his great white horse, his head in a gauze hat, and a colorful thread, and a lion, and a lion, and a pair of pink soap shoes with a pair of pink soap shoes	Simon qing is rich	Simon is rich and has the pursuit of the upper class
With a black oil hair on the head, a single flower and a head flower, the upper body in white summer cloth, peach and red skirt, blue and blue, the bottom of the white satin, the red yarn lacquer	Pan jinlian beauty	Pan jinlian is beautiful and wants to live in luxury
The head is wrapped in a thousand fingers, and it is worn with a blue cloth shirt, a red towel in the waist, and a knee and knee in the lower leg	Wuzi is brave	Wuzi wu yi gao qiang, brave and brave
In the golden dress, the gold crown is worn with gold corona, and the hand is a golden stick, and the foot is matched	Monkey King	Sun wukong was untalized to be filial
In the majestic doya show, the Buddha dress is like a cut, the glow is full of flowers and the time, the color of the universe	Tang's piety	Tang's monk shoulders the mission and the piety of the dharma

7. Conclusion

In this paper, we constructed a metaphorical dress recognition model based on Transformer and graph convolutional neural and a metaphor comprehension model based on Seq2seq framework to conduct a relevant research on dress metaphors in Ming Dynasty novel texts.

The metaphorical dress recognition model based on Transformer and graph convolutional neural combines the properties of Transformer and graph neural network, using Transformer to extract global contextual semantic information and combining it with graph convolutional neural network to deal with sentence dependencies, which is capable of recognizing multi-word dress metaphorical expressions more effectively. Better performance is achieved on different metaphors and types of Ming Dynasty novels, exceeding the best results achieved by the comparison model. In the task of recognizing dress metaphors in worldly love novels, the correct rate of dress metaphor recognition of this paper's model is improved by 17.97% and 7.28% compared with the MEM model and CRF model. It shows that the metaphorical dress recognition model based on Transformer and graph convolutional neural can effectively combine the contextual semantic and syntactic structural information when dealing with dress metaphors in Ming Dynasty novel texts, and significantly improve the metaphor recognition accuracy.

The metaphor understanding model based on the Seq2seq framework is able to better model the generation and understanding of metaphors end-to-end through the encoder-decoder structure, which is especially suitable for the understanding of complex clothing metaphors, and is able to generate profoundly meaningful interpretation content of clothing metaphors, and the average length of the interpreted content is higher than that of the comparison model by 17.1 and 16.14. It is able to accurately interpret clothing metaphors in the novels of the Ming dynasty, and reveal their metaphorical meaning in the cultural and social contexts. It can accurately interpret the clothing metaphors in Ming Dynasty novels, revealing their metaphorical meanings in the cultural and social contexts.

Future research can further explore the combination of the metaphor recognition model and the metaphor interpretation model in this paper to achieve a more comprehensive metaphorical computational linguistics modeling.

References

[1] Prihandoyo, W. B. (2022). Roland Barthes semiotic analysis of Rimpu Bima costume. Journal of Positive Psychology and Wellbeing, 6(1), 2889-2896.

[2] Özdil, M. A. (2021). The effect of clothing as a marker on identity. Motif Akademi Halkbilimi Dergisi, 14(33), 117-130.

[3] Tajuddin, F. N. (2018). Cultural and Social Identity in Clothing Matters Different Cultures, Different Meanings. European Journal of Behavioral Sciences, 1(4), 21-25.

[4] Miller-Spillman, K. A., & Reilly, A. (2019). The meanings of dress. Fairchild Books.

[5] Shefer, T., Ratele, K., & Clowes, L. (2017). “Because they are me”: Dress and the making of gender. South African Review of Sociology, 48(4), 63-81.

[6] Kornilieva, L. (2022). The Signifying Role of the Dress in the Novel “In Search of Lost Time”(A La Recherche Du Temps Perdu) by Marcel Proust. Bibliotekarz Podlaski, 54(1), 181-191.

[7] LI, W. Y. (2018). IN THE BOOK OF SONGS. The Homeric Epics and the Chinese Book of Songs: Foundational Texts Compared, 329.

[8] Ren, C., & Yang, X. (2020). Research on Costume Ideology of the Pre-Qin Period from Perspective of Chu Ci. Asian Social Science, 16(4), 59.

[9] Plaks, A. H. (2025). The Four Masterworks of the Ming Novel: Ssu ta ch'i-shu. Princeton University Press.

[10] Meilan, Z., & Hao, Z. (2019). The differentiation between elegance and popularity in the orthography of popular literature of the Ming Dynasty and its causes. Journal of Chinese Writing Systems, 3(1), 29-44.

[11] Jinju, W., & Zhiwei, D. (2020). Innovation in Temporal and Spatial Narration of The Plum in the Golden Vase and Its Significance to the History of Fiction. Theoretical Studies in Literature and Art, 40(3), 107-116.

[12] Liu, Y. (2022, July). Social and Aesthetic Values of the Novels in the Ming and Qing Dynasties. In 2022 3rd International Conference on Language, Art and Cultural Exchange (ICLACE 2022) (pp. 290-296). Atlantis Press.

[13] Xu, Y. (2023). The Relationship between Literati Livelihoods and the Development of Novels and Operas in the Canal Region during the Ming and Qing Dynasties. Journal of chinese humanities, 9(2), 196-213.

[14] Lu, C., & Ismail, N. H. (2023). THE DESIGN CHARACTERISTIC OF LATE 15TH CENTURY MING DYNASTY COSTUMES IN THE JIANGNAN REGION. Jurnal Tuah, 4(1), 179-188.

[15] Choi, Y. H., Kim, S. E., & Lee, K. H. (2021). The Relationship Between Character and Costume in literary Work using Semantic networks-The novel “Norwegian Wood”. Journal of Digital Convergence, 19(1), 307-314.

[16] Ichien, N., Stamenković, D., & Holyoak, K. (2024). Interpretation of Novel Literary Metaphors by Humans and GPT-4. In Proceedings of the Annual Meeting of the Cognitive Science Society (Vol. 46).

[17] Peter Uhe,Chris Lucas,Laurence Hawker,Malcolm Brine,Hamish Wilkinson,Anthony Cooper... & Christopher Sampson. (2025). FathomDEM: an improved global terrain map using a hybrid vision transformer model. Environmental Research Letters(3),034002-034002.

[18] Romy Sauvayre,Jessica S. M. Gable,Adam Aalah,Melvin Fernandes Novo,Maxime Dehondt & Cédric Chauvière. (2024). The Impact of Autonomous Vehicle Accidents on Public Sentiment: A Decadal Analysis of Twitter Discourse Using roBERTa. Technologies(12),270-270.

[19] Burak Gülmez. (2025). GA-Attention-Fuzzy-Stock-Net: An optimized neuro-fuzzy system for stock market price prediction with genetic algorithm and attention mechanism. Heliyon(3),e42393-e42393.