Using TextCNN Method to Extract Sentiment Tendency and Material Perception Mapping in Leather Product Consumer Reviews

Qiaozheng Kang
School of Art and Design, Fuzhou University of International Studies and Trade, Fuzhou 350202, Fujian, China
Corresponding author: Qiaozheng Kang (e-mail: kangqiaozzz@163.com)

ABSTRACT Accurately identifying the relationship between consumer sentiment and material perception is essential for intelligent product evaluation and data-driven manufacturing optimization. In engineering scenarios involving electromagnetic sensing and smart inspection systems, reliable semantic understanding of material-related information can provide complementary support for multimodal perception and decision-making. To address the problems of mixed sentiment expressions, long-range semantic dependencies, and ambiguous sentiment-material associations in leather product reviews, this study proposes an improved TextCNN framework integrating gated attention and multi-task joint learning. The proposed method dynamically filters salient semantic features through a gating mechanism and employs global attention weighting to reconstruct long-distance contextual relationships while simultaneously optimizing sentiment classification and material perception mapping using shared semantic representations. Experimental results demonstrate that the proposed model achieves a Macro-F1 score of 89.4% for sentiment recognition, an average semantic association accuracy of 79.9%, and a mapping coverage of 81.6%, outperforming representative baseline methods. The framework effectively enhances the interpretability of sentiment-material coupling and provides a practical semantic analysis strategy for intelligent material evaluation, multimodal information fusion, and engineering applications associated with electromagnetic sensing and advanced manufacturing systems.

INDEX TERMS sentiment analysis, TextCNN, material perception, multi-task learning, electromagnetic sensing

I. INTRODUCTION

WITH the popularity of leather products in the consumer market, consumers' sentiment feedback on products and evaluation of material experience have increasingly become important bases for brand design and quality management [1-3]. There is often a long textual distance between sentiment words and material words, and they are often accompanied by cross-sentence, metaphor, or omission phenomena, which significantly increases the difficulty of semantic association modeling between the two [4-5]. Especially in the field of leather, consumer descriptions tend to be emotional expressions, such as "feeling high-end" and "soft like clouds". When a single review involves multiple materials, it becomes more difficult to match sentiment with specific materials, resulting in a decrease in the accuracy of existing models.

Traditional Word-of-Mouth (WOM) literature shows that emotions embedded in advertising appeals or refer rals/reviews influence consumer buying journey [6]. In recent years, researchers have paid extensive attention to how to accurately identify sentiment tendencies from consumer reviews to support product optimization and user behavior modeling [7]. Early studies mostly used traditional Machine Learning (ML) methods. Ramadhan F A [8]combined naive Bayes with Support Vector Machine (SVM) for sentiment classification. Kar daras D K [9] used fuzzy relations and fuzzy comprehensive evaluation methods to calculate the usefulness of reviews by incorporating user biases in terms of expressiveness and emotional polarity. The study shows that fuzzy logic can capture the subjectivity commonly found in reviews and can be used to quantify differences in user behavior and calculate the usefulness of reviews. Mitra S [10]and Almaqtari H [11]successively integrated Convolutional Neural Network (CNN) and Gated Recurrent Unit (GRU) networks to improve the context feature modeling capability. Yang M [12]further improved the model's context understanding capability based on Lite's Bidirectional Encoder Representations from Transformers (BERT) and bidirectional GRU structure . Although these methods have made progress in improving the accuracy of sentiment classification, most of them still focus on the modeling of sentiment polarity and ignore the modeling of the semantic carrier behind the generation of sentiment, resulting in the inability to achieve precise association between user sentiment and product attributes.

To solve the problem of the association between sentiment classification and semantic content extraction, some studies have begun to explore collaborative methods for sentiment recognition and product attribute modeling. Salim S[13] used BERT to model three types of consumer data with an accuracy rate of more than 98%, showing the powerful ability of the Transformer model in capturing user selection preferences. Alamsyah A [14] proposed a robustly optimized BERT multi-label learning mechanism and performed topic modeling to accurately extract the core topics related to clothing quality. In terms of attribute modeling, Choi J [15] proposed to establish a mapping relationship between material description and user perception through p-value test, correlation analysis, and Euclidean distance metric to achieve semantic modeling of product personality and material selection. Liu H [16] used the Deep Learning (DL) model under edge computing to perform feature sentiment analysis on comment content and effectively realized product attribute recognition. These have attempted to integrate sentiment analysis and semantic perception, but have failed to effectively integrate sentiment and attribute information for joint optimization, resulting in inaccurate recognition of sentiment-material mapping relationships, especially when faced with complex semantic expressions or fuzzy associations.

To address this issue, this paper uses the TextCNN method to study the extraction of sentiment tendencies and material perception mappings in consumer reviews of leather products. The main contributions of this paper are: 1) a gating mechanism is applied to dynamically screen convolutional features and enhance the model's response to core semantic fragments; 2) combined with the attention weighting strategy, each semantic fragment is globally fused to achieve the reconstruction and positioning of sentiment-material semantic pairs; 3) a multi-task learning structure with shared semantic representation is further constructed, and the sentiment classification and material perception mapping tasks are simultaneously optimized through a joint loss function to achieve modeling of the mapping relationship between the two. Compared with existing methods, the model in this paper performs better in terms of sentiment-material matching accuracy under complex and fuzzy semantics, and can more effectively mine cognitive relationships in consumer reviews.

SENTIMENT-MATERIAL SEMANTIC FEATURE MODELING BASED ON TEXTCNN

TEXT PREPROCESSING AND WORD VECTOR CONSTRUCTION

Due to the significant differences between Chinese and English language structures, and the frequent use of mixed Chinese and English expressions in actual e-commerce reviews for brand names, material abbreviations like "PU leather," and model numbers, this article adopts a mixed segmentation strategy. First, the Chinese portion is segmented using the precise mode based on the Jieba segmentation database. Subsequently, standard English segmentation is performed using the word_tokenize function from the Natural Language Toolkit (NLTK) library for any English words or phrases interspersed within the reviews. After word segmentation, each comment is converted into a sequence of words:

$$T = [w_1, w_2, \cdots, w_n] \tag{1}$$

w_i represents the i-th word, and n is the sentence length.

For Chinese vocabulary, Google's open-source Word2Vec (Skip-gram architecture, dimension set to 300) pre-trained model was used for embedding encoding to effectively capture Chinese semantic similarity relationships. For English vocabulary appearing in reviews, the GloVe (Global Vectors for Word Representation) pre-trained model [17] was used for vectorization, leveraging the rich contextual co-occurrence information trained on English corpus. The GloVe performance is better than other word embeddings because it applies to non-zero elements and a subset of a corpus, not a whole corpus or a separate window of the significant corpus [18]. The model is also trained on Wikipedia and Gigaword corpora, with the word vector dimension set to 300, which is strictly aligned with the Chinese Word2Vec vector dimension.

To address the inconsistency of Chinese and English word vectors in different semantic spaces, this paper uses a linear mapping alignment strategy to unify cross-lingual word vector spaces. First, we load Google's open-source Word2Vec Chinese pre-trained model and GloVe English pre-trained model. We select a bilingual dictionary containing approximately 1,500 common Chinese-English aligned word pairs as an anchor set to learn the linear transformation matrix from the GloVe English space to the Word2Vec Chinese space. This transformation is solved using orthogonal Procrustes analysis, with the goal of minimizing the distance between the anchor word pairs' vectors after the transformation:

$$\min_{W} \sum_{(e,c)\in\mathcal{P}} \| W \bullet v_e - v_c \|^2 \; s.t. \, W^\top W = I \tag{2}$$

Where P is the set of aligned Chinese and English word pairs, v_e is the vector of the English word in GloVe space, and v_c is the vector of the corresponding Chinese word in Word2Vec space. The resulting matrix W projects the English word vectors into a shared semantic space aligned with the Chinese vectors. For each word appearing in the comment, it is processed according to its language type: Chinese words are directly represented by their 300-dimensional vector v_c in the Word2Vec space; English words are first represented by their GloVe pre-trained vector v_e, and then spatially aligned using a learned orthogonal mapping matrix W to obtain the transformed vector Wv_e. This transformation projects the English word vectors into a semantic space consistent with the Chinese word vectors.

Finally, all Chinese and English words are represented as 300-dimensional vectors and reside in a unified shared semantic space. The vectors corresponding to each word in each comment are arranged sequentially to form an $L \times 300$ two-dimensional matrix (L is the sentence length, with a maximum truncation of 128, padded with zeros if insufficient, and truncated if exceeding), which serves as the input representation for the model. The word vector obtained by the Word2Vec model through Skip-gram training can effectively capture the semantic similarity relationship between words [19-20]:

$$v_w = f(w) \tag{3}$$

In Formula (3), $v_w \in \mathbb{R}^d$ represents the word vector representation of word w; d is the word vector dimension, which is set to 300 in this paper; function $f(\cdot)$ represents the mapping function from vocabulary to vector space.

TEXTCNN MULTI-SCALE CONVOLUTION

After completing the standardized preprocessing and word vector embedding of the comment text, this paper uses the TextCNN model as the main framework for feature extraction to capture the n-gram semantic structure in the local context fragment of the comment and explore the potential semantic association between sentiment tendency and material perception. The entire word vector matrix $X \in \mathbb{R}^{n \times d}$ is used as the input of the "text image", and multiple convolution kernels with different window sizes are applied to it for sliding extraction to obtain local feature representations at different n-gram levels, as shown in Figure 1.

For the convolution operation, a convolution kernel $W_h \in \mathbb{R}^{h \times d}$ with a window size of h is set. The convolution kernel slides on the input matrix to extract the feature vector of the local context area [21-22]. The word vectors between the sliding window positions i and $i+h-1$ are concatenated into a local submatrix $X_{i:i+h-1}$, and the convolution kernel is multiplied with it to obtain the feature representation c_i at that position, which is calculated as follows:

$$c_i = \mathrm{ReLU}(\langle W_h, X_{i:i+h-1} \rangle + b) \tag{4}$$

Feature maps are generated using convolutional kernels (windows) and nonlinear functions (filters). Pooling operations are then performed on these feature maps to select the optimal features [23]. In Formula (4), $\langle \cdot \rangle$ represents the matrix inner product operation; b is the bias term; ReLU is the activation function. In this paper, the ReLU (Rectified Linear Unit) function is used for nonlinear mapping:

$$\mathrm{ReLU}(z) = \max(0, z) \tag{5}$$

By sliding the convolution kernel at all possible positions in the sentence, a feature map can be obtained [24]:

$$c = [c_1, c_2, \cdots, c_{n-h+1}] \tag{6}$$

To further capture semantic structures of different granularities, this paper applies a multi-scale convolution kernel structure in the convolution layer design, that is, multiple convolution kernel groups with different window sizes $h \in \{2, 3, 4\}$ are set in parallel to learn n-gram features of different lengths respectively. In this paper, the convolution kernel window size is set to 2, 3, and 4, respectively, capturing the local semantic information of binary, tri-gram, and tetragram language combinations, respectively.

SEMANTIC ENHANCEMENT MECHANISM

In texts where "sentiment tendency" and "material description" are not continuously associated, it is difficult for the convolutional model to capture the potential coupling semantics between the two. To further enhance the model's ability to model the semantic association between sentiment words and material descriptions, this paper applies a context-aware mechanism into the TextCNN structure and constructs an improved TextCNN-Attention model. By connecting the gating mechanism and attention weighting strategy after the output of the convolutional layer, the global context modeling and dynamic weighting of the local n-gram feature sequence are performed to improve the dual performance of sentiment tendency recognition and material perception mapping, as shown in Figure 2:

For each convolution output feature vector $c_i \in \mathbb{R}^f$, its gating transformation is defined as:

$$\tilde{c}_t = c_i \odot \sigma(W_g c_i + b_g) \tag{7}$$

In Formula (7), $W_g \in \mathbb{R}^{f \times f}$ is the gating weight matrix; $b_g \in \mathbb{R}^f$ is the bias term; $\odot$ represents Hadamard multiplication.

The emotion classification branch and the texture perception branch each maintain independent gating weight matrices W_g^{sent} and W_g^{mat}, as well as bias terms b_g^{sent} and b_g^{mat}, to calculate the channel activation strengths for emotion and texture relevance, respectively. Sentiment classification and material perception tasks maintain independent gating mechanisms. For the i feature vector $h_i \in \mathbb{R}^f$ of the convolution output, its gating response in the sentiment task is:

$$z_i^{\mathrm{sent}} = \sigma(W_g^{\mathrm{sent}} h_i + b_g^{\mathrm{sent}}) \tag{8}$$

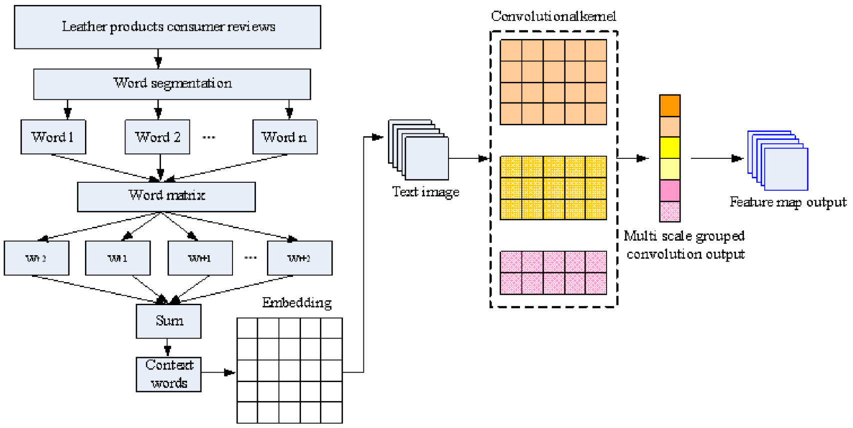

FIGURE 1. Multi-scale convolution feature extraction

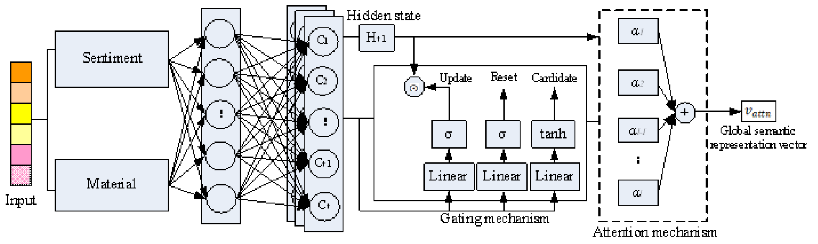

FIGURE 2. Gating mechanism and attention mechanism

With the corresponding enhancement feature being $\tilde{h}_i^{\text{sent}} = z_i^{\text{sent}} \odot h_i$; in the material perception task, the i gating response is:

$$z_i^{\text{mat}} = \sigma(W_g^{\text{mat}} h_i + b_g^{\text{mat}}) \tag{9}$$

With the corresponding enhancement feature being $\tilde{h}_i^{\text{mat}} = z_i^{\text{mat}} \odot h_i$. The gating feature sequences for the two tasks are retained independently; that is, the sentiment branch uses the sequence $\{\tilde{h}_1^{\text{sent}}, \cdots, \tilde{h}_L^{\text{sent}}\}$, and the material branch uses the sequence $\{\tilde{h}_1^{\text{mat}}, \cdots, \tilde{h}_L^{\text{mat}}\}$. Subse quently, each branch introduces an independent attention mechanism to calculate task-specific context weights and generate global semantic representations r_{sent} and r_{mat} for sentiment classification and material perception, respectively.

MULTI-TASK JOINT TRAINING AND LOSS OPTIMIZATION

This paper adopts a multi-task learning architecture based on hard sharing and supplemented by task-specific enhancements. The outputs of the multi-scale convolutional layers of TextCNN serve as shared low-level features, which are used by both tasks. Above this shared representation, sentiment classification and material awareness tasks introduce independent gating mechanisms and attention modules, respectively, to achieve task-oriented semantic enhancement. The shared portion ends at the convolutional features; starting from the gating mechanism, the two tasks enter their own independent processing paths, with each branch generating the final prediction through its own task-specific fully connected layer. Therefore, the model's "hard sharing" only applies to the convolutional feature extraction stage, while gating and attention belong to task-specific upper-level modules, as shown in Figure 3.

First, the multi-scale convolution output is subjected to a maximum pooling operation to compress the sequence length and retain the most representative local feature response value. Assuming that the convolution layer output sequence is $H = [h_1, h_2, \cdots, h_L] \in \mathbb{R}^{L\times f}$, where L is the sequence length, and F is the feature dimension, the maximum pooling is defined as:

$$z_j = \max_{i=1,\cdots,L}(h_{ij}), \quad j = 1, 2, \cdots, F \tag{10}$$

A fixed-length feature vector $z \in \mathbb{R}^f$ is obtained as the local semantic enhancement representation of the text. The feature vector z after maximum pooling is concatenated with the global semantic representation to form a joint feature representation:

$$F = \text{Concat}(z, r^{\text{sent}}, r^{\text{mat}}) \in \mathbb{R}^{2f} \tag{11}$$

On this basis, a dual-task output structure is constructed. First, a nonlinear transformation is performed on the joint feature vector F through a fully connected layer:

$$h_{\text{shared}} = \tanh(W_s F + b_s) \tag{12}$$

$W_s \in \mathbb{R}^{H\times 2f}$ is the shared weight matrix. Next, two independent output branches are derived from the shared

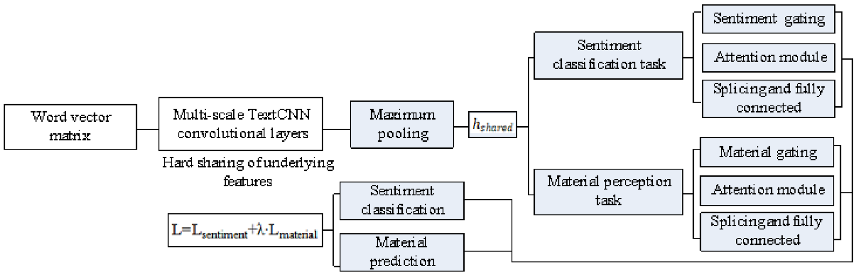

FIGURE 3. Dual-task learning

representation h_{shared}: sentiment classification branch and material perception branch. For the sentiment classification branch, there is:

$$y_{\text{sentiment}} = \text{softmax}(W_{\text{sent}} h_{\text{shared}} + b_{\text{sent}}) \tag{13}$$

$W_{\text{sent}} \in \mathbb{R}^{C_s \times H}$ is the sentiment classification weight matrix, and C_s is the number of sentiment categories, including positive, neutral, and negative. The Softmax function is used to output the probability distribution of each category. For the material perception task, the output layer is defined as:

$$y_{\text{material}} = \sigma(W_{\text{mat}} h_{\text{shared}} + b_{\text{mat}}) \tag{14}$$

$W_{\text{mat}} \in \mathbb{R}^{C_m \times H}$ is the material perception mapping weight matrix, and C_m is the number of material labels, including "genuine leather", "poly urethane (PU leather)", and "vegetable tanned leather". The activation function uses the Sigmoid function to support multi-label prediction. According to statistics on material labeling on product detail pages on most platforms, in major categories like leather shoes, bags, and leather clothing, the majority of products are clearly labeled as "genuine leather," "PU leather," or "vegetable tanned leather," with the remaining subcategories mostly present as sub-attributes. Therefore, these three categories cover the core material types that most consumers care about.

To achieve joint optimization of the two tasks, this paper adopts a multi-task loss function:

$$L = L_{\text{sentiment}} + \lambda \cdot L_{\text{material}} \tag{15}$$

$L_{\text{sentiment}}$ and L_{material} are the cross entropy loss functions of sentiment classification and material perception, respectively, and λ is the balance coefficient used to adjust the learning weights between the two tasks.

After completing the training of the multi-task joint learning framework, the context-aware driven sentiment-material mapping modeling is used to extract the semantically directed material-sentiment coupling relationship from unstructured comments. Based on the gated convolution-attention fusion representation, the attention weight sequence $\alpha = [\alpha_1, \alpha_2, \cdots, \alpha_L]$ of each convolution segment is generated. The original text segmentation result and the material perception label set $M = \{m_1, m_2, \cdots, m_k\}$ are combined. By matching the keyword position interval, the corresponding segment of each material word in the convolution sequence is located, and its weighted attention response value is calculated:

$$\omega_j = \sum_{i \in \text{span}(m_j)} \alpha_i \tag{16}$$

ω_j represents the semantic significance of the material word m_j in the entire comment, $\text{span}(m_j)$ is the token interval covering the material keyword in the convolution operation. On this basis, the local semantic segment with the highest sentiment intensity in the window where the material word is located is further extracted, and its dominant sentiment polarity s_j is determined by combining its corresponding Softmax classification probability. Finally, the material words and their associated sentiments identified in each comment are combined into a set of semantic pairs (s_j, m_j), and by aggregating such semantic pairs in all comments, a set of sentiment-material mapping relationships for the leather products field is constructed to achieve structured modeling and directional analysis of consumer perception characteristics. Emotional orientation and material perception belong to different semantic levels. In consumer reviews, not all material mentions are explicitly accompanied by emotional expression. If the model forcibly assigns a sentiment polarity to each material word, this may lead to the generation of false semantic pairs. To alleviate this problem, this paper

introduces an emotional salience threshold mechanism into the emotion-to-material mapping process. When extracting the dominant emotion s_j associated with a material word m_j, not only the maximum emotional probability value $p_{\max} = \max(p_{\text{pos}}, p_{\text{neu}}, p_{\text{neg}})$ of the convolutional segment to which it belongs is considered, but also a sentiment confidence threshold τ (set to 0.6 in this paper) is set. If $p_{\max} < \tau$, the material mention is considered to have no clear emotional orientation, and the emotion-to-material pair is not output; otherwise, the semantic pair is retained.

SEMANTIC RECOGNITION ACCURACY OF LEATHER REVIEWS

A. EXPERIMENTAL SETUP

To deeply explore the correlation between the sentiment tendencies and semantic features expressed by consumers in leather product reviews, this paper constructs and uses a Chinese e-commerce review dataset for leather products. This dataset mainly comes from public user reviews on JD.com and Taobao e-commerce platforms, covering four typical leather consumer categories: leather shoes, leather bags, leather clothes, and leather sofas. The data acquisition time range is from January 2023 to December 2023. A total of 52,629 original review data were collected. After deduplication, denoising, and manual screening, 20,568 high-quality Chinese reviews were retained for this study. All reviews are primarily in Chinese, with a small number of mixed Chinese and English expressions ("PU leather" and "soft feel") accounting for approximately 2.3%. However, the majority of the language is Chinese, consistent with the user language habits of mainstream e-commerce platforms. The dataset was randomly split into a training set: validation set: test set = 70%: 15%: 15% ratio. The split process used a fixed random seed of seed=42 to ensure experimental reproducibility. To prevent data leakage caused by multiple reviews of the same product appearing simultaneously in the training and test sets, this paper adopts a group-based split strategy when partitioning the dataset. All reviews belonging to the same product ID are collectively assigned to a subset of the training, validation, or test sets, ensuring that the model does not encounter product reviews seen during the training phase during the evaluation phase. The data category and attribute information are shown in Table 1.

TABLE 1. Dataset information

Item	Content
Total amount of data	20568
Average sentence length (number of words)	47.3 words
Sentiment distribution	Positive: 13214
Sentiment distribution	Neutral: 3076
Sentiment distribution	Negative: 4278
Material	Genuine leather
Material	PU leather
Material	Vegetable tanned leather

All data was obtained through public platforms' interfaces, strictly adhering to the respective platforms' terms of service and agreements. No restricted or authorized data resources were accessed. During the dataset partitioning phase, to prevent evaluation bias caused by the same user's comments being split across different sets, this paper uses anonymized user identifiers for user-level isolation: after data partitioning is completed, all information that can be used to identify an individual is removed, and only the anonymized plain text comment content is retained for subsequent model training and analysis. This research does not involve the storage or correlation analysis of user privacy data, complying with the requirements of the Personal Information Protection Law and the General Data Protection Regulation (GDPR), ensuring regulatory compliance. To construct a collection of sentiment-material semantic pairs, this paper employed a double-blind manual annotation process. Three researchers with backgrounds in natural language processing participated in the annotation task. All annotators received unified training, and the core rules included:

1) Material word identification: Only words that explicitly mentioned leather were annotated, excluding ambiguous terms.

2) Sentiment association: If a review included subjective evaluations of the material, a sentiment-material pair was created; if it contained only factual descriptions, no sentiment was associated.

3) Sentiment polarity annotation: Sentiment was categorized as positive, neutral, or negative, based on the overall contextual tendency.

To ensure annotation quality, this paper employs a structured double-blind annotation process. Each comment is first independently annotated with sentiment-material pairs by two annotators with natural language processing backgrounds. If the two annotators agree, the label is directly adopted as the final label; if there is disagreement, a third senior annotator is brought in for review. The final label is determined based on a majority vote of the three annotators' results. In rare cases (where all three annotators disagree), the annotation team reaches a consensus through focused discussion and makes the final decision. After annotation, inter-annotator agreement was calculated, using Cohen's Kappa (paired analysis) and Krippendorff's Alpha (overall agreement). The results showed an average Cohen's Kappa of 0.82 and a Krippendorff's Alpha of 0.79, indicating good reproducibility of the annotation process. Subsequent analysis was performed based on the consensus label set generated by majority voting, and divergent samples were determined after discussion. The model in this paper is implemented based on the PyTorch framework, and its specific parameter settings are shown in Table 2:

TABLE 2. Model parameter settings

Dimension	Parameter	Specification
Convolutional layer	Convolutional kernel window size	[2, 3, 4]
	Number of convolution kernels per class	100
Gating mechanism	Dimension of Gate Control Weight Matrix	Consistent with the number of channels
	Bias initialization method	Zero vector
Attention mechanism	Attention head count	1
	Context aware vector dimension	300
Training configuration	optimizer	Adam
	Learning rate	0.001
	Batch size	64
	Balance coefficient of loss function	0.6

In the multi-task joint loss function, a grid search was performed on the validation set for the balance coefficient λ. The evaluation metric comprehensively considered the Macro-F1 for sentiment classification and the Macro-F1 for texture perception, using a weighted harmonic mean as the selection criterion. Ultimately, the model achieved optimal overall performance on the validation set when $\lambda = 0.6$, so it was fixed at 0.6 and used in all comparative experiments and ablation studies.

The method is evaluated from three aspects: basic task performance, semantic modeling ability, and sentiment-material perception mapping quality. Three mainstream text classification models are selected as comparison methods: basic TextCNN (original structure), BiLSTM, and BERT. To ensure the fairness and scientific rigor of the comparative experiments, this paper independently tunes the hyperparameters of all baseline models within their typical configurations or standard practice ranges. Furthermore, all baseline models were trained in the same experimental environment to ensure the validity and comparability of the comparison results. Pre-trained language models such as BERT and RoBERTa typically employ a single-task fine-tuning

paradigm. However, to control for variation and ensure consistent output structure, this paper adapts them to a multi-task learning framework with a shared encoder and dual-task output heads, maintaining structural alignment with the improved TextCNN proposed in this paper. This aims to isolate differences in the backbone encoder and focus on comparing the performance of different feature extraction mechanisms within the same task framework. To ensure maximum fairness, all models were trained using identical hyperparameters: the Adam optimizer, a batch size of 64, a maximum sequence length of 128, and a dropout rate of 0.1.

EXPERIMENTAL RESULTS

(1) Basic task performance This paper constructs a ROC (Receiver Operating Characteristic) curve comparison experiment based on the predicted probability of the test set. By plotting the ROC curve of each model and calculating the AUC (Area Under Curve) value, its sentiment discrimination performance under imbalanced data is quantified, as shown in Figure 4:

As shown in Figure 4, the AUC value of this model is 0.843, which is significantly better than BERT (0.831), BiLSTM (0.817), and TextCNN (0.795). This model can more effectively perceive the dependency between sentiment words and non-continuous material descriptions in consumer reviews, and enhance the focus on key semantic fragments through global attention, achieving more precise sentiment tendency classification. Accuracy and Macro-F1 are further used as evaluation indicators to compare the performance of the model and verify the effectiveness of the method in capturing sentiment semantics. The results are shown in Table 3:

TABLE 3. Sentiment semantics capture results

Model	Accuracy (%)	Macro-F1 (%)
Basic TextCNN	86.2	84.5
BiLSTM	87.5	85.9
BERT	89.1	87.3
Improved TextCNN	90.6	89.4

As can be seen from Table 3, the proposed model outperforms the comparison model in the task of sentiment tendency recognition, with Accuracy and Macro-F1 reaching 90.6% and 89.4%, respectively, indicating that the applied attention mechanism and gating structure effectively enhance the modeling ability of complex semantics and improve the model's capture effect of key semantic fragments, thereby improving the overall sentiment semantic recognition performance.

To enhance the comparability and standardization of experimental results, this paper further uses standard metrics in the field of information extraction—precision and recall—to evaluate the performance of texture perception tasks and the joint extraction of emotion-texture semantic pairs, and compares them with baseline models. To enhance the cutting-edge nature and persuasiveness of the comparative experiments, in addition to the basic TextCNN, BiLSTM, and BERT, this paper further introduces RoBERTa-base-Chinese as a stronger Transformer baseline model. The experimental results are shown in Table 4.

TABLE 4. Standard performance comparison of material perception and semantic pair extraction

Model	Material Perception (Macro)	Sentiment-Material Pair Extraction
	P / R (Precision / Recall)	P / R
Basic TextCNN	72.1 / 74.9	68.3 / 65.2
BiLSTM	77.6 / 79.0	73.1 / 70.8
BERT	81.2 / 82.6	76.5 / 74.2
RoBERTa	83.6 / 83.9	77.4 / 76.1
Improved TextCNN	87.0 / 85.3	80.2 / 78.6

As shown in Table 4, our model achieves a precision and recall of 87.0% and 85.3% on the material perception task, significantly outperforming the baseline TextCNN (72.1% and 74.9%), BiLSTM (77.6% and 79.0%), and BERT (81.2% and 82.6%). The second-best RoBERTa achieves precision and recall of 83.6% and 83.9%, which are still lower than the model in this paper. In the emotion-material semantic pair extraction task, precision and recall reach 80.2% and 78.6% respectively, also achieving top performance. These results demonstrate that our method not only accurately identifies material information but also effectively establishes semantic associations between it and emotion.

(2) Semantic modeling ability To evaluate the model's ability to capture the semantic dependencies between sentiment words and material words, this paper constructs an evaluation metric based on manually annotated sentiment-material semantic pairs. For each comment, if the model successfully predicts a sentiment-material pair consistent with the manually annotated pair, it is considered a correct

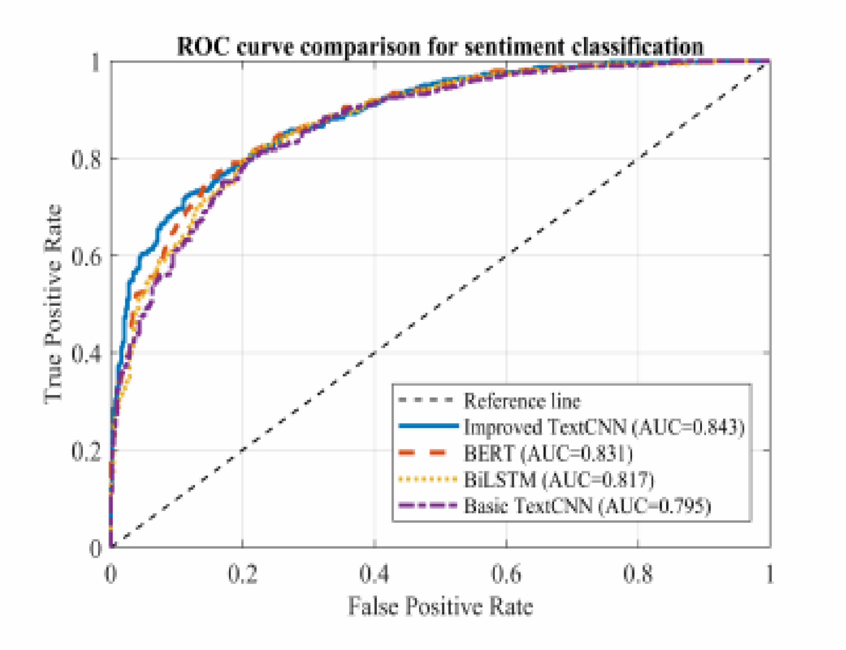

FIGURE 4. ROC curve comparison results

association recognition. To further analyze the model's performance across different contextual spans, this paper divides the word distance between sentiment words and material words into five preset intervals: ≤10, 11–15, 16–20, 21–25, and >25 words. These interval boundaries are pre-defined based on domain common sense and the typical length of Chinese comments, and the model's association recognition accuracy is calculated separately within each interval, as shown in Figure 5.

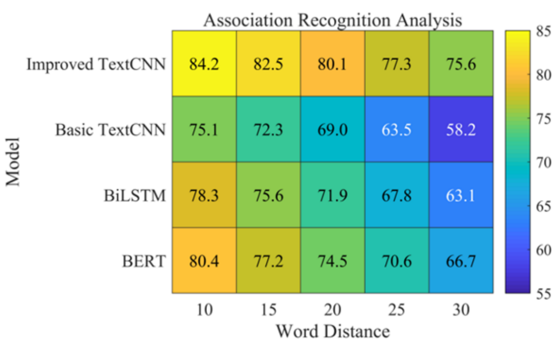

FIGURE 5. Comparison results of semantic association recognition

In Figure 5, as the word distance increases, the association recognition ability of all models shows a downward trend, indicating that long-distance context increases the difficulty of semantic modeling. The proposed model maintains a relatively higher recognition accuracy in all distance segments, and its semantic association recognition at each word distance reaches an average accuracy of 79.9%; the Basic TextCNN, BiLSTM, and BERT are 67.6%, 71.3%, and 73.9%, respectively. Basic TextCNN is limited by the local receptive field and is difficult to model long-distance semantic dependencies; although BiLSTM has sequence modeling capabilities, it is easily disturbed by noise; BERT pre-trained language models are powerful, but are not sensitive enough to the material-sentiment coupling relationship in specific domain fine-tuning. In contrast, the proposed method integrates global context information through the attention mechanism and the gating structure, significantly improving the recognition ability of key semantic segments and showing better association modeling effects.

(3) Sentiment-material perceptual mapping quality At different word distances, the sentiment-material matching accuracy metric was used to determine whether the model correctly extracted sentiment-material semantic pairs from reviews. This metric measures whether the model can accurately identify both the material label and its dominant sentiment polarity. A correct prediction was considered only when both the material and sentiment were correctly matched, reflecting the model's overall performance in the task of precise semantic alignment. Matching accuracy was measured by calculating the percentage of correctly predicted semantic pairs, and the Kappa coefficient was used to assess consistency with manual annotations, as shown in Figure 6.

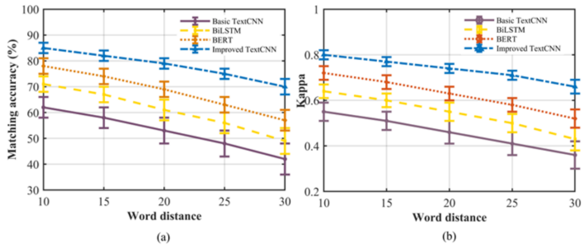

FIGURE 6. Sentiment-material matching comparison: (a) matching accuracy; (b) Kappa consistency coefficient

The results in Figure 6 show that Improved TextCNN has advantages in the long-distance sentiment-material association recognition task. In Figure 6(a), as the word distance increases, the sentiment-material matching accuracy of each model shows a downward trend. When the word distance reaches 30, the matching accuracy of the proposed model reaches 70%, and its average matching accuracy is 78.2%; the average accuracy of the three types of comparison models, Basic TextCNN, BiLSTM, and BERT, is 52.6%, 60.8%, and 68.2%, respectively. In Figure 6(b), the average Kappa consistency coefficient between the semantic pairs correctly predicted by the proposed model and the manual annotation is 0.74; while the comparison models are 0.46, 0.54, and 0.63, respectively.

To further test the completeness of the model's mapping structure, we calculated its mapping coverage by counting the proportion of the number of emotion-material combinations recognized by the model in the entire test set to the total number of manually annotated combinations. Mapping coverage measures the proportion of unique emotion-material combination types recognized by the model in the entire test set to the total number of manually annotated combinations, reflecting the model's ability to fully model user perception patterns. The results are shown in Figure 7:

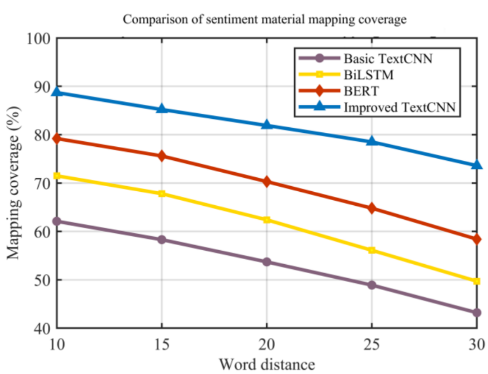

FIGURE 7. Comparison results of mapping coverage

In Figure 7, the improved TextCNN proposed in this paper outperforms the comparison model under all distance conditions. The average mapping coverage of the improved TextCNN reaches 81.6%, far exceeding BERT (69.7%), BiLSTM (61.5%), and basic TextCNN (53.2%), indicating that it has stronger ability in mapping structure integrity. By weighted fusion of contextual attention and based on the dual-task learning structure, the model's comprehensive understanding of the semantics of leather product reviews is improved on the basis of sharing the underlying semantic representation.

(4) Ablation Experiments To verify the independent contribution of each module proposed in this paper to the model performance, this paper designed a series of ablation experiments, gradually removing or replacing key components, and evaluating their impact on model performance on the same test set. The experiment took the complete model of this paper as the benchmark and conducted the following variant tests in sequence: (1) w/o Multi-scale removes the multi-scale convolution kernel; (2) w/o Gating removes the gating mechanism (3) w/o Attention removes the global attention module; (4) w/o Multi-task splits the dual-task framework and trains two independent models to complete the sentiment classification and material perception tasks respectively, and then associate the results through rule matching; (5) Full Model. The results are shown in Table 5:

TABLE 5. Comparison of ablation experiment results

Model variant	Sentiment classification macro average F1 value (%)	Semantic association recognition accuracy (%)	Fuzzy matching rate (%)	Mapping coverage rate (%)
Full Model	89.4	79.9	86.8	81.6
w/o Multi-scale	88.1	77.2	83.5	78.9
w/o Gating	87.6	76.1	81.3	77.4
w/o Attention	86.8	74.3	78.6	75.2
w/o Multi-task	87.9	75.8	80.1	76.8

Table 5 shows that after removing the multi-scale convolution, semantic association recognition accuracy dropped to 77.2%, indicating that n-gram features of different granularities complement each other in capturing the local semantics of material and sentiment words. Removing the gating mechanism resulted in a decrease in the Sentiment classification macro average F1 value to 87.6%, indicating its important role in suppressing irrelevant semantic fragments and enhancing key features; When using independent single-task training, various metrics also showed varying degrees of decline, validating the effectiveness of multi-task joint learning in enhancing the modeling of coupled sentiment and material. Overall, each component proposed in this paper significantly contributes to model performance, supporting its superior performance in complex semantic mapping tasks.

II. CONCLUSIONS

This paper addresses the problem of limited mapping modeling due to the long-range contextual dependency between sentiment words and material descriptions in consumer reviews of leather products. By constructing a TextCNN model that integrates multi-scale convolution with an attention mechanism, this paper considers TextCNN as an efficient local semantic detector and implements global semantic reconstruction based on the attention mechanism.

By dynamically screening and context-aware weighted fusion of local n-gram features, this paper achieves indirect modeling of non-continuous semantic segments, effectively improving the dual performance of sentiment recognition and material perception mapping. Experimental results show that in terms of semantic modeling, the average accuracy of the proposed model for sentiment-material matching are 78.2%, respectively, and the average mapping coverage is 81.6%, both of which are better than the comparison model. The proposed method provides a structured modeling path for product attribute-oriented sentiment analysis, which has good practical application value.

AUTHOR CONTRIBUTIONS

Qiaozheng Kang designed, collected and analyzed the data, and drafted the manuscript. Qiaozheng Kang conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Qiaozheng Kang participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] I. D'Adamo, M. Gagliarducci, M. Iannilli, and V. Mangani, "Fashion Wears Sustainable Leather: A So-cial and Strategic Analysis Toward Sustainable Production and Consumption Goals," Sustainability, vol. 16, no. 22, pp. 9971-10005, 2024, doi: 10.3390/su16229971.

[2] D. Rocklage M and H. Fazio R, "The enhancing versus backfiring effects of positive emotion in con-sumer reviews," Journal of Marketing Research, vol. 57, no. 2, pp. 332-352, 2020, doi: 10.1177/0022243719892594.

[3] B. Lis and M. Fischer, "Analyzing different types of negative online consumer reviews," Journal of Product & Brand Management, vol. 29, no. 5, pp. 637-653, 2020, doi: 10.1108/JPBM-05-2018-1876.

[4] K. Sharma, S. Trott, S. Sahadev, and R. Singh, "Emotions and consumer behaviour: A review and re-search agenda," International Journal of Consumer Studies, vol. 47, no. 6, pp. 2396-2416, 2023, doi: 10.1111/ijcs.12937.

[5] J. Alantari H, S. Currim I, Y. Deng, and S. Singh, "An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews," International Journal of Research in Marketing, vol. 39, no. 1, pp. 1-19, 2022, doi: 10.1016/j.ijresmar.2021.10.011.

[6] V. Srivastava, D. Kalro A, G. Raizada, & Sharma, and D, "Beyond stars: role of discrete emotions on online consumer review helpfulness," Journal of Marketing Theory and Practice, vol. 32, no. 2, pp. 137-157, 2024, doi: 10.1080/10696679.2022.2148111.

[7] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," Social network analysis and mining, vol. 11, no. 1, pp. 81-99, 2021, doi: 10.1007/s13278-021-00776-6.

[8] A. Ramadhan F, P. Ruslan R R, and A. Zahra, "Sentiment analysis of e-commerce product reviews for content interaction using machine learning," Cakrawala Repositori IMWI, vol. 6, no. 1, pp. 207-220, 2023, doi: 10.52851/cakrawala.v6i1.219.

[9] K. Kardaras D, C. Troussas, G. Barbounaki S, P. Tselenti, and K. & Armyras, "A Fuzzy Synthetic Evalua-tion Approach to assess usefulness of Tourism Reviews by considering bias identified in Sentiments and Articulacy," Information, vol. 15, no. 4, pp. 236-254, 2024, doi: 10.3390/info15040236.

[10] S. Mitra and M. Jenamani, "Helpfulness of online consumer reviews: A multi-perspective approach," Information Processing & Management, vol. 58, no. 3, pp. 1-17, 2021, doi: 10.1016/j.ipm.2021.102538.

[11] H. Almaqtari, F. Zeng, and A. Mohammed, "Enhancing Arabic Sentiment Analysis of Consumer Re-views: Machine Learning and Deep Learning Methods Based on NLP," Algorithms, vol. 17, no. 11, pp. 495-510, 2024, doi: 10.3390/a17110495.

[12] M. Yang, "An effective emotional analysis method of consumer comment text based on AL-BERT-ATBiFRU-CNN," International Journal of Information Technologies and Systems Approach (IJITSA), vol. 16, no. 2, pp. 1-12, 2022, doi: 10.4018/IJITSA.324100.

[13] S. Salim, Z. Iqbal, and J. Iqbal, "Emotion classification through product consumer reviews," Pakistan Journal of Engineering and Technology, vol. 4, no. 4, pp. 35-40, 2021, doi: 10.51846/vol4iss4pp35-40.

[14] A. Alamsyah and D. Girawan N, "Improving clothing product quality and reducing waste based on consumer review using RoBERTa and BERTopic language model," Big Data and Cognitive Compu-ting, vol. 7, no. 4, pp. 168-188, 2023, doi: 10.3390/ bdcc7040168.

[15] J. Choi, "Material selection: Exploring the reliability of material perception data and its relevance to materials," Metals and Materials International, vol. 27, no. 6, pp. 1599-1604, 2021, doi: 10.1007/s12540-019-00574-y.

[16] H. Liu, "Online review analysis on various networks' consumer feedback using deep learning," IET networks, vol. 11, no. 6, pp. 234-244, 2022, doi: 10.1049/ntw2.12045.

[17] Z. Labd, S. Bahassine, K. Housni, H. Aadi F Z A, & Benabbes, and K, "Text classification supervised algorithms with term frequency inverse document frequency and global vectors for word representa-tion: a comparative study," International Journal of Electrical & Computer Engineering (2088-8708), vol. 14, no. 1, pp. 589-599, 2024, doi: 10.11591/ ijece.v14i1.pp589-599.

[18] M. Mohammed S, K. Jacksi, and S. Zeebaree, "A state-of-the-art survey on semantic similarity for doc-ument clustering using GloVe and density-based algorithms," Indonesian Journal of Electrical Engineering and Computer Science, vol. 22, no. 1, pp. 552-562, 2021, doi: 10.11591/ijeecs.v22.i1.pp552-562.

[19] A. Fellah, A. Zahaf, and A. Elci, "Semantic similarity measure using a combination of word2vec and wordnet models," Indonesian Journal of Electrical Engineering and Informatics (IJEEI), vol. 12, no. 2, pp. 455-464, 2024, doi: 10.52549/ ijeei.v12i2.5114.

[20] K. Babic, F. Guerra, S. Martincic-Ipsic, and A. & Mestrovc, "A comparison of approaches for measuring the semantic similarity of short texts based on word embeddings," Journal of information and organiza-tional sciences, vol. 44, no. 2, pp. 231-246, 2020, doi: 10.31341/jios.44.2.2.

[21] Q. Li, X. Li, Y. Du, and X. & Chen, "A new sentiment-enhanced word embedding method for sentiment analysis," Applied Sciences, vol. 12, no. 20, pp. 10236-10253, 2022, doi: 10.3390/app122010236.

[22] H. Yang, S. Zhang, H. Shen, G. Zhang, Deng, X, J. Xiong, et al., "A multilayer feature fusion model based on convolu tion and attention mechanisms for text classification," Applied Sciences, vol. 13, no. 14, pp. 8550-8568, 2023, doi: 10.3390/ app13148550.

[23] J. Mutinda, W. Mwangi, and G. Okeyo, "Sentiment analysis of text reviews using lexicon-enhanced bert embedding (LeBERT) model with convolutional neural network," Applied Sciences, vol. 13, no. 3, pp. 1445-1458, 2023, doi: 10.3390/app13031445.

[24] K. Agbesi V, W. Chen, B. Yussif S, C. Ukwuoma C, H. Gu Y, and A. & Al-Antari M, "MuTCELM: An optimal multi-TextCNN-based ensemble learning for text classification," Heliyon, vol. 10, no. 19, pp. 1-21, 2024, doi: 10.1016/j.heliyon.2024.e38515.