# VAD Emotion Distribution Augmented BERT for Fine-Grained Emotion Recognition

Sufen Chen, Chunyang Li, and Xueqiang Zeng*

***Abstract*—The fine-grained emotion recognition model, which delineates human emotions into multiple distinct categories, excels in capturing the subtleties of emotional expression. The large number of emotion categories, coupled with the intricate relationships between them, presents challenges for the advancement of fine-grained emotion recognition models. A promising approach to enhancing the performance of existing models is the integration of VAD (Valence-Arousal-Dominance) psychological theory, which encapsulates emotions through three continuous dimensions. This approach offers a more granular depiction of emotional nuances compared to traditional discrete emotion models. However, most existing fine-grained emotion recognition models directly integrate the three-dimensional VAD scores of affective words without adequately considering the interrelations between emotions in the VAD space. To address this issue, this paper proposes the VAD Emotion Distribution Augmented BERT for fine-grained emotion recognition (EDA-BERT). EDA-BERT leverages the emotional correlations in the VAD space through emotion distribution and integrates these with contextual representations to refine emotion classification. The EDA-BERT architecture is composed of three modules: a semantic information module, an emotion distribution module, and a fusion prediction module. The semantic information module utilizes a pre-trained BERT model to distill contextual text embeddings. The emotion distribution module computes the emotional proximities using VAD-based distance metrics, thereby constructing an emotion distribution map for the sentiment annotations of affective words. The fusion prediction module, equipped with an attention mechanism, synthesizes the contextual embeddings with the emotion distribution data to yield accurate emotion predictions. Empirical evaluations conducted on the GoEmotions and EmpatheticDialogues datasets confirm the superior performance of the EDA-BERT model over existing baseline methods in the realm of fine-grained emotion recognition.**

***Index Terms*—emotion recognition, emotion distribution, VAD emotion model, fine-grained emotion, BERT**

## I. Introduction

With the advent of the social media era, characterized by unprecedented connectivity and the rapid proliferation of user-generated content, the recognition of textual emotions has emerged as a rapidly evolving and crucial research direction within the expansive field of Natural Language Processing (NLP) [1]. Text emotion recognition has seen extensive application across various sectors, including the healthcare sector, online public sentiment monitoring, and consumer behavior analysis [1]-[4]. Initial scholarly efforts predominantly concentrated on discerning the sentiment valence of textual content, aiming to classify it as either positive or negative. However, in recent times, the pursuit of fine-grained emotion analysis has gained prominence, characterized by an expansion in the number of emotional categories and the intricate interconnections among them. This development has considerably heightened the inherent complexity associated with the task of emotion recognition in textual data, posing new challenges and opportunities for researchers in the field.

Traditional emotion recognition models generally rely on basic emotion models for emotion classification. Psychologists Ekman [5] and Plutchik [6] categorized human emotions into six or eight basic emotions. For instance, Zhang et al. [7] employed word2vec to train word embeddings and constructed a CNN model to identify seven emotions in a Weibo corpus; Khanpour et al. [3] used CNN and LSTM models to detect six basic emotions in health-related texts; Lei et al. [8] utilized graph convolutional networks and Bi-LSTM to extract semantic features and recognize seven emotions from Chinese corpora. Basic emotion models, which encompass only six or eight emotions, are inadequate for fully capturing the broad spectrum of emotions that humans experience and express. Consequently, researchers have refined and expanded the scope of basic emotion models, adopting sophisticated fine-grained emotion models to enhance the precision and depth of their study into human emotional complexities. For instance, Keltner et al. [9] proposed an emotion model comprising 34 emotions, while Cowen et al. introduced 24 emotions expressed in human speech prosody [10] and 28 emotions conveyed through facial expressions [11]. Demszky et al. [12] proposed the fine-grained text emotion dataset GoEmotions, comprising 54,000 Reddit comments, which includes 27 emotion labels and a "neutral" label.

Fine-grained emotion models utilize a large number of emotion categories to represent individuals' emotional states, offering enhanced ability to express emotions compared to traditional emotion models. However, the expanded number of emotion categories, coupled with the interrelations and ambiguities between fine-grained emotions, renders fine-grained emotion recognition more challenging. To address these challenges, some studies have enhanced emotion prediction models by incorporating external knowledge, such as psychological models. For instance, Dhar

Manuscript received January 20, 2025; revised July 21, 2025. This work was supported in part by the Natural Science Foundation of China under grant no. 62266021.

Sufen Chen is an associate professor at Jiangxi University of Water Resources and Electric Power, Nanchang, China (email: csf@nit.edu.cn).

Chunyang Li is a graduate student of Jiangxi Normal University, Nanchang, China (email: lichunyang@jxnu.edu.cn).

Xueqiang Zeng is a professor at Jiangxi Normal University, Nanchang, China (Corresponding author, tel: 86-18970085460, email: xqzeng@jxnu.edu.cn).

et al. [13] employed sentiment lexicons to classify emotions in corporate tweets; Zhong et al. [14] introduced the VAD (Valence-Arousal-Dominance) psychological emotion model [15] and proposed a dynamic attention model based on affective words VAD scores; Suresh et al. [16] proposed a knowledge-embedded attention mechanism that combines the VAD scores of affective words with the contextual representations provided by pre-trained models. By incorporating external knowledge, these methods have achieved some success in fine-grained emotion recognition tasks. However, current fine-grained emotion recognition models that incorporate VAD knowledge typically use the three-dimensional VAD scores of affective words directly, neglecting the interrelationships between emotions in the VAD space. In fact, the distances between fine-grained emotions in the psychological emotion model reflect the degree of correlation among emotions. Constructing emotion distributions based on emotional distances can quantitatively capture the degree to which affective words express distinct emotions, providing a significant advantage in emotion recognition tasks involving emotional ambiguity [17].

By introducing emotion distribution to model the correlations between fine-grained emotions in the VAD space, this paper proposes the VAD Emotion Distribution Augmented BERT for Fine-grained Emotion Recognition (EDA-BERT) model. The EDA-BERT model computes the distances between emotions in the continuous emotional space based on the VAD model, constructs emotion distributions from these distances, and integrates the emotion distribution information with contextual representations for emotion prediction. The EDA-BERT model consists of three modules: the semantic information module, the emotion distribution information module, and the fusion prediction module. The semantic information module extracts contextual representations from the pre-trained BERT language model; the emotion distribution information module calculates the distances between emotions with the VAD psychological model and constructs emotion distributions, quantifying the extent to which affective words express different emotions across various emotion categories; the fusion prediction module combines emotion distribution information with contextual representations using attention mechanisms [18] and concatenation [19], which are used for emotion prediction. Experimental results on the GoEmotions fine-grained emotion dataset demonstrate that the EDA-BERT model outperforms baseline models in fine-grained emotion recognition tasks.

## II. Related Work

### A. Text Based Emotion Recognition

Emotion recognition is a critical task in text-based intelligent systems, finding widespread applications across various domains [20]. In recent years, text emotion recognition has predominantly relied on neural networks, including CNN, LSTM, and RNN [21]-[24]. With the advancement of deep learning, the Transformer-based pre-trained language model Bidirectional Encoder Representations from Transformers (BERT) [25], proposed by Google, has led to groundbreaking advancements in the field of NLP. Given BERT's outstanding performance across various NLP tasks, an increasing number of researchers have applied pre-trained models to emotion recognition studies [26]. Traditional emotion recognition models often examine basic emotion models, such as those proposed by psychologists Ekman [5], who identified six basic emotions (happiness, anger, fear, surprise, sadness, and disgust), and Plutchik [6], who proposed eight basic emotions. For instance, Li et al. [27] utilized a Chinese affective words database to identify six emotions in Weibo text; Akhtar et al. [28] proposed a multi-task learning framework with CNN and LSTM, leveraging different feature representations for emotion recognition; Khanpour et al. [3] utilized CNN and LSTM models to detect six basic emotions in health-related texts.

However, basic emotion models, which encompass only six to eight emotions, often fail to capture the full complexity of human emotions. Fine-grained emotion models, which utilize a larger number of emotion categories, offer a greater capacity to represent emotions. Psychologist Keltner et al. [9] proposed a fine-grained emotion model comprising 34 emotions. Building on this, Demszky et al. [12] developed the GoEmotions fine-grained text emotion dataset, which includes 54,000 Reddit comments and contains 27 emotion labels, along with a 'neutral' label. While fine-grained emotion models provide richer emotional expression, the large number of emotion categories, along with their interrelationships and ambiguities, poses significant challenges for fine-grained emotion recognition models. To address these challenges, existing fine-grained emotion recognition research has enhanced model performance by incorporating psychological emotion knowledge. For instance, Bruyune et al. [19] integrated the contextual representations of BERT with dictionary scores and performed classification using Bi-LSTM; Dhar et al. [13] applied the VADER lexicon for sentiment classification of tweets; Zhong et al. [14] leveraged the VAD scores of affective words from the NRC-VAD lexicon as external knowledge to boost fine-grained emotion recognition performance. However, these methods rely on the three-dimensional VAD scores of affective words, neglecting the correlations between emotions in the VAD space.

### B. VAD Emotion Knowledge

Dimensional models in psychological emotion theory suggest that the structure of emotions is based on multiple evaluative dimensions, such as valence and arousal, and that emotions can be described using quantitative variables, forming a high-dimensional continuous emotion space

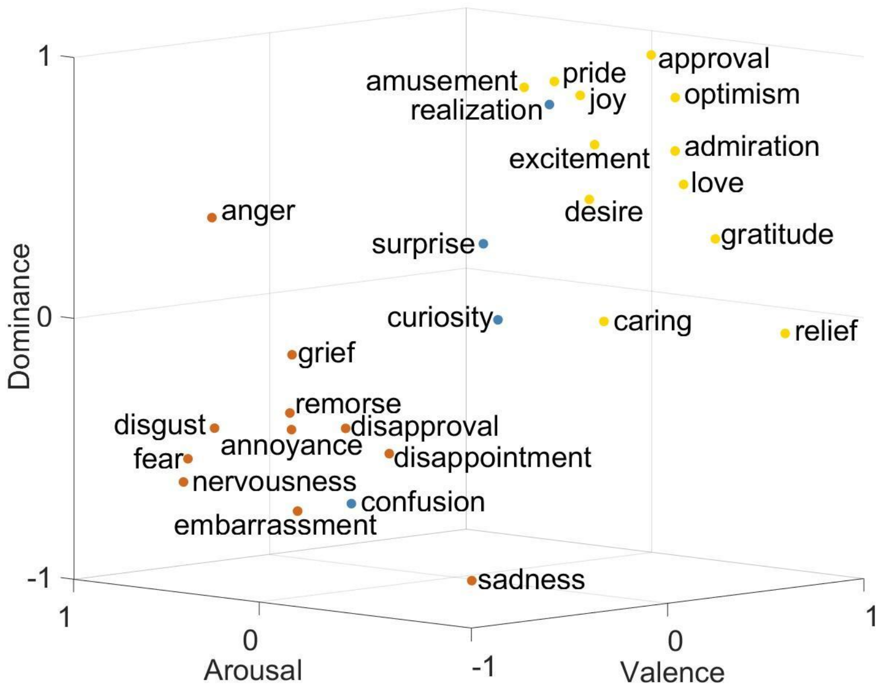


Fig. 1. Expression of 27 emotions in the VAD model space

system [28]. Continuous-space emotion models define emotions with continuous values, facilitating the identification of latent connections between emotions and providing a more accurate description of the relationships between different emotions. The VAD model of emotion, introduced by Russell [15], posits that valence, arousal, and dominance are the fundamental dimensions for defining emotions. The VAD model defines 151 emotional states across these three dimensions, using mean and standard deviation values. As illustrated in Figure 1, the VAD model is applied to 27 distinct emotions.

The VAD model employs Euclidean distance to quantify the similarity between emotions. A smaller distance indicates greater similarity between emotions, with stronger correlations, while a larger distance signifies greater emotional differences and weaker correlations. Because the distance between emotions is represented as continuous values, the VAD model captures the subtle differences between emotions more effectively than discrete emotion models. Figure 2 presents the emotional distance matrix for 27 distinct emotions based on VAD emotional knowledge.

For instance, in emotionally similar categories, such as excitement and pride, Figure 1 shows their distribution in the VAD space, suggesting that the two emotions are closely related. In Figure 2, the VAD distance between excitement and pride is 0.4, smaller than the distance between excitement and other emotion categories. This information enables the model to differentiate between semantically similar but emotionally distinct categories. Although excitement and pride are semantically close, the VAD model, by incorporating dimensions such as arousal and dominance, enhances the model’s ability to differentiate between emotions prone to confusion.

### *C. Knowledge-Enhanced Text Representations*

External knowledge offers invaluable emotional insights for fine-grained emotion recognition tasks, which can significantly enhance the representations learned by deep learning models, thereby effectively complementing their emotion recognition capabilities [29]. Conventional methods typically fuse external knowledge with the contextual embeddings of pre-trained models through techniques like attention mechanisms [18] and concatenation [19]. For example, Wang et al. [30] used an adapter model to pre-train knowledge data separately and then integrated it with contextual representations; Bruyune et al. [19] fused the contextual representations of BERT with external knowledge from sentiment lexicons via a Bi-LSTM; and Suresh et al. [16] designed a knowledge-embedded attention mechanism, combining affective words scores from emotion lexicons with the contextual representations of the pre-trained BERT model. These methods have achieved notable success by incorporating additional knowledge sources [31]-[32]. Nevertheless, the straightforward application of emotion scores derived from sentiment lexicons as external knowledge neglects the inherent nature of the VAD psychological emotion model, which exists within a three-dimensional, continuous space. Within this VAD space, the proximity between emotions signifies their interconnections and nuances. Consequently, the direct utilization of VAD scores does not adequately harness the intricate associations among fine-grained emotions, and it does not provide a thorough external knowledge base for the enhancement of fine-grained emotion recognition models.

## III. Methodology

### *A. Model Description*

To address these challenges, this paper proposes the VAD Emotion Distribution Augmented BERT for Fine-grained Emotion Recognition (EDA-BERT), which improves

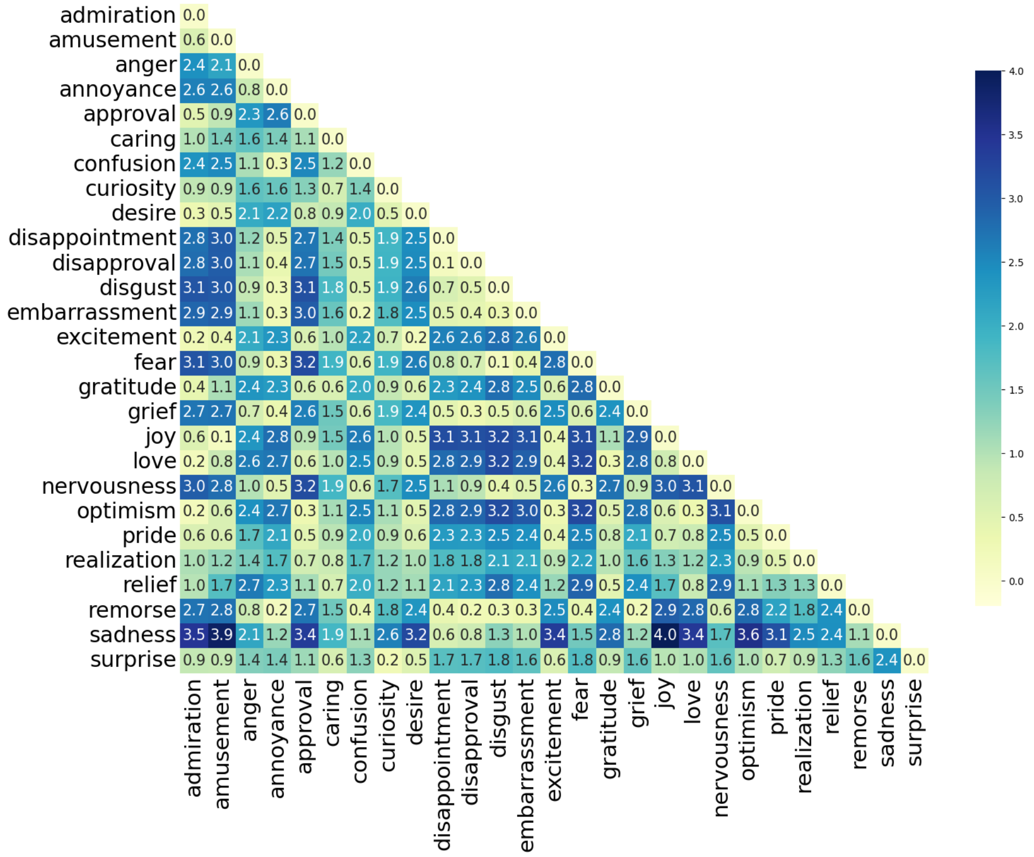


Fig. 2. Emotional distance matrix of 27 emotions in the VAD model space knowledge

fine-grained emotion recognition by quantifying the expression strength of each affective word across different emotion categories in the VAD space through emotion distribution. Algorithm 1 outlines the operational procedure of the EDA-BERT model for clarity.

The architecture of the EDA-BERT model is illustrated in Figure 4, comprising three modules: the semantic information module, the emotion distribution information module, and the fusion prediction module. The corresponding pseudo-code is given in Algorithm 1.

### *B. The Semantic Information Module*

The semantic information module extracts semantic features from text using the pre-trained BERT model. The pre-trained model effectively leverages the contextual information of each word by encoding its relationships with surrounding context, thus improving the model's understanding of textual content [25].

The process of semantic information extraction is as follows: Let the training set be denoted as $X$ , where $x_i\left(i\in\{1,2,3,\ldots,m\}\right)$ represents the $i$-th sentence, and $m$ denotes the total number of sentences in the training set. First, the sentence $x_i$ is fed into the model; then, the pre-trained BERT model converts the input text into a word vector; finally, this word vector is input into BERT to obtain the semantic information of the sentence $H_c=\{h_0,h_1,\ldots,h_n\}$, where $n$ is the number of words in the sentence, $H_c\in R^{n\times l_c}$ , and $l_c$ is the output dimension of the hidden state layer ($l_c$=768 in the BERT-base version).

---

**Algorithm 1: EDA-BERT Model**

**Input:** $X=\{x_i, y_i\}$ // training set
**Output:** $\hat{y}=\{\hat{y}_1,\hat{y}_2,\ldots,\hat{y}_i\}$ // the set of predictive labels
**for** $i=\{1,2,\ldots,m\}$ **do**
  **if** $x_i$ **then**
    $H_c=\{h_0,h_1,\ldots,h_n\}\leftarrow$ BERT( $x_i$ );
    Calculate first token '[CLS]' representations $h_0$ ;
  **end if**
  **if** $x_i$ **then**
    Extracting affective words $W=\{w_{i,k}\}_{k=1}^{n_i}$ from $x_i$ ;
    Obtain all emotion labels $q_{i,k}^t$ of $w_{i,k}$ ;
    Generate emotion label distribution $f_{q_{i,k}^t}$ for each $q_{i,k}^t$ ;
    Obtain emotion distribution sequence $D_e=\{d_0,\ldots,d_k\}$
  **end if**
  $H_e$=FC（$D_e$）
  $s=\mathrm{Softmax}\left(h_0^T\cdot H_e\right)$ //calculate attention scores
  $h_a=s^T\cdot H_e$
  $H_{ce}=\mathrm{concat}(H_c,h_a)$
  Use $H_{ce}$ to predict result of classification $\hat{y}$
**end for**
Back propagation and update parameters in EDA-BERT Model;
return $\hat{y}=\{\hat{y}_1,\hat{y}_2,\ldots,\hat{y}_i\}$ ;

---

### *C. The Emotion Distribution Information Module*

The purpose of the emotion distribution module is to transform the emotion labels of affective words within a sentence into an emotion distribution. The emotions expressed in a sentence are considered as a combination of

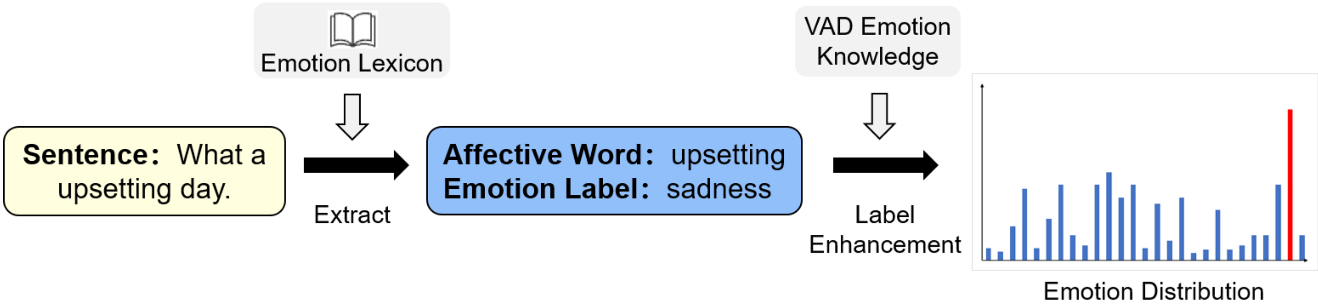


Fig. 3. Emotional distribution label enhancement example

multiple basic emotions, each with varying intensities. This module leverages the VAD emotion model to quantify the expression intensities of all emotion categories, collectively constructing an emotion distribution. As illustrated in Figure 3, the construction process comprises three main steps: 1) Extracting affective words from the sentence. 2) Generating an emotion distribution for each emotion label.

Step 1: Using a sentiment lexicon, extract the $k$-th affective word $w_{i,k}$ from the input sentence $x_i$ and construct the corresponding affective word set $W=\{w_{i,k}\}_{k=1}^{n_i}$, where $n_i$ represents the number of affective words in the sentence $x_i$. Each affective word corresponds to at least one sentiment label, specifically the $t$-th sentiment label $q_{i,k}^t$ of $w_{i,k}$.

Step 2: Based on each emotion label, the corresponding emotion distribution $f_{q_{i,k}^t}$ is generated. The emotion distribution of each affective word is represented by $d_k$, where $C$ denotes the number of emotion labels for each affective word.

$$d_k=\frac{1}{C}\cdot\sum_{t=1}^{C} f_{q_{i,k}^t} \tag{1}$$

The primary emotion $\alpha$ is determined based on the emotion label of the affective word, while other emotion categories in the VAD emotion space are treated as secondary emotions $e$. The correlation between two emotions is measured by the distance between the primary emotion $\alpha$ and the secondary emotion $e$ in the VAD space. This distance is then transformed into the corresponding emotion distribution using a Gaussian kernel. The rules for generating the emotion distribution are as follows: 1) To emphasize the primary emotion $\alpha$, it is assigned the maximum score in the emotion distribution. 2) The scores of secondary emotions are inversely proportional to their VAD distance from the primary emotion $\alpha$. The emotion distribution is calculated using the following formula:

$$f_\alpha^e=\frac{1}{Z}\exp\left(-\frac{(\|\mu_e-\mu_\alpha\|_2+b)^2}{2\tau^2}\right) \tag{2}$$

$$Z=\sum_e\exp\left(-\frac{(\|\mu_e-\mu_\alpha\|_2+b)^2}{2\tau^2}\right) \tag{3}$$

where, $\mu_e=\left[V^e,A^e,D^e\right]$ represents the coordinates of emotion $e$ in the three dimensions of the VAD emotion model. The term $b$ is a bias parameter, set to $b=1$, $\tau$ to control the locality and weight decay of the emotion distribution. The variable $Z$ is a normalization factor ensuring that $\sum f_\alpha^e=1$; $\|\mu_e-\mu_\alpha\|_2$ denotes the Euclidean distance between the secondary emotion $e$ and the primary emotion $\alpha$ in the VAD space. Based on the emotion labels of each affective word, the final emotion distribution sequence of the sentence is obtained, denoted as $D_e$, $D_e=\{d_0,d_1,\ldots,d_k\}$, where $k$ represents the number of affective words in the sentence.

### *D. The Fusion Prediction Module*

The fusion prediction module combines semantic information with emotion distribution data, achieving information fusion and prediction through attention mechanisms and concatenation operations. As shown in Figure 4, the process begins with the extraction of affective words from the sentence via the emotion distribution module, converting them into an emotion distribution sequence $D_e$. Next, the emotion distribution sequence $D_e$ is transformed into emotion knowledge encoding $H_e$ via a fully connected layer, where $H_e\in R^{l_e\times l_c}$. A self-attention mechanism is then applied, with $H_e$ serving as the Key(K) and $h_0$ as the Query(Q). Matrix multiplication between K and Q is passed through a Softmax function to yield attention scores $s$. Subsequently, $H_e$ is used as the Value(V) and multiplied by the attention score $s$ through matrix multiplication.

$$s=\mathrm{Softmax}\left(h_0^T\cdot H_e\right) \tag{4}$$

$$h_a=s^T\cdot H_e \tag{5}$$

$$H_{ce}=\mathrm{concat}(H_c,h_a) \tag{6}$$

$$\hat{y}=\mathrm{sigmoid}(H_{ce}) \tag{7}$$

Finally, the semantic information $H_c$ is concatenated with the output of the attention mechanism $h_a$ to form $H_{ce}$, which is subsequently input into a fully connected layer. The output is passed through a Sigmoid activation function to produce the emotion prediction probability. The loss for

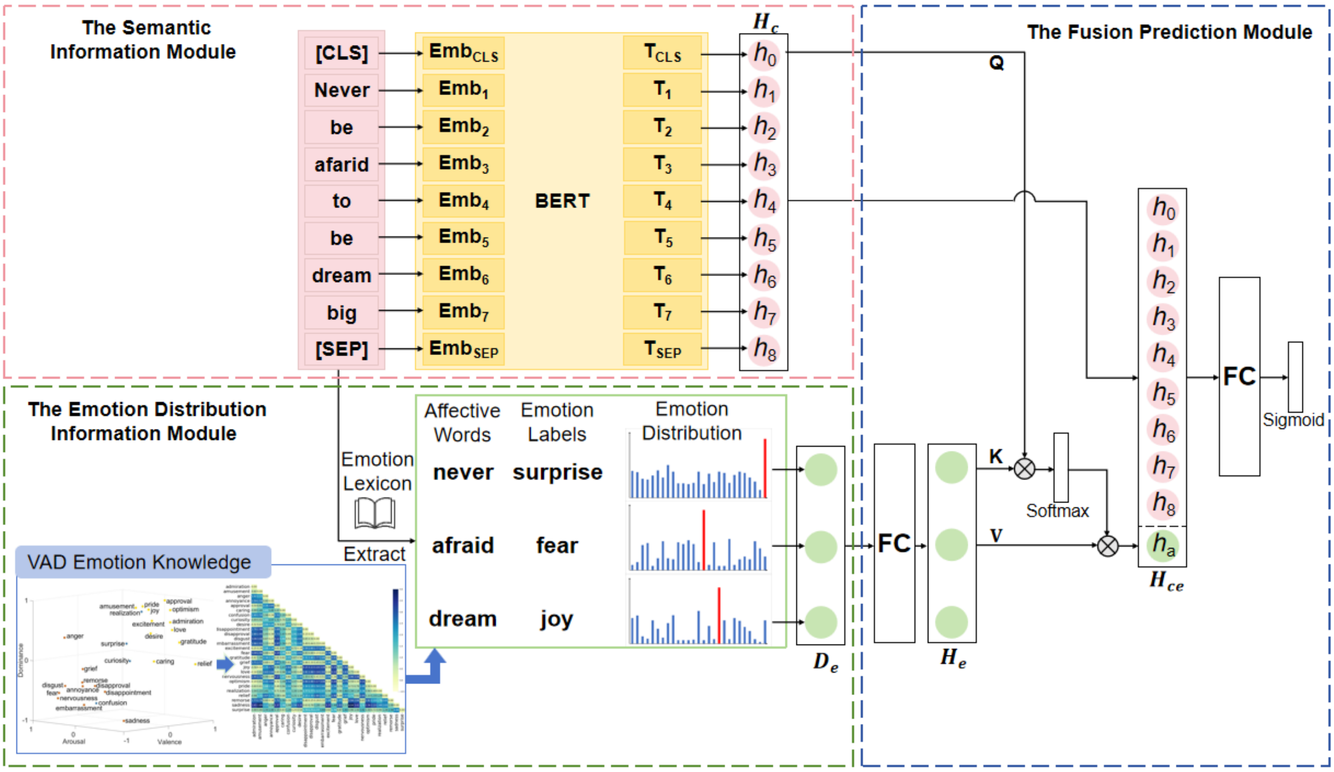


Fig. 4. Architecture diagram of VAD Emotion Distribution Augmented Fine-grained Emotion Recognition. EDA-BERT comprising three modules: the semantic information module, the emotion distribution information module, and the fusion prediction module.

multi-label classification is calculated using the binary cross-entropy loss.

$$L=-\sum_{i=1}^{N}\sum_{j=1}^{C}(y_{i,j}\log(\hat{y}_{i,j})+(1-y_{i,j})\log(1-\hat{y}_{i,j})), \quad (8)$$

where $y_{i,j}$ and, $\hat{y}_{i,j}$ is the predicted probability, respectively, for the $j$ label of the $i$ sample.

## IV. EXPERIMENT

### A. Experimental Setup

To demonstrate the effectiveness of the proposed model, we rigorously conduct comparative experiments on two diverse datasets. These datasets are commonly used and encompass a wide array of fine-grained emotion categories, ensuring a comprehensive assessment of our models' performance across various text domains. Specifically, one dataset comprises forum posts, while the other consists of conversations.

- **GoEmotions:** This dataset released by Google, which comprises approximately 54,000 English language comments collected from Reddit. The GoEmotions dataset includes 28 fine-grained emotion categories, which are organized into a three-level hierarchical structure, as illustrated in Figure 5. Within the dataset, 83% of the samples are annotated with a single emotion, 15% with two emotions, and 2% with three or more emotions. During data preprocessing, the 54,263 samples in the GoEmotions dataset were split into training, validation, and test sets in an 8:1:1 ratio. The distribution of samples across emotion categories is imbalanced. For instance, the "neutral" category contains 14,219 training samples, whereas categories such as "grief," "pride," "relief," and "nervousness" have fewer than 200 training samples each.
- **EmpatheticDialogues (ED):** This dataset is an English language dialogue corpus specifically designed for emotion-aware dialogue modeling. The dataset contains 24,850 two-person conversations, with each dialogue averaging 4.31 utterances, mirroring the emotional nuances typically observed in everyday interpersonal communication. Each conversation is annotated with one of 32 predefined emotion categories, such as grateful, anxious, and lonely, covering a broad emotional spectrum from positive to negative sentiments. A notable strength of the ED dataset is its balanced distribution of emotion classes, which helps reduce the risk of model bias due to class imbalance and improves the generalizability of emotion recognition systems across a diverse emotional range. The dataset is divided into three standard subsets for experimentation: 19,533 conversations for training, 2,770 for validation, and 2,547 for testing. For model input, utterances are concatenated using the [SEP] token as a separator.

The English sentiment lexicon employed in this study was constructed by merging the NRC Emotion Lexicon [33] and EmoSenticSpace [34]. The intersection of emotion labels from both lexicons was preserved. For affective words present in both lexicons, their emotion labels were combined as the union of the original labels.

To evaluate the performance of the model, we use the Macro-F1 score as the evaluation metric. The Macro-F1

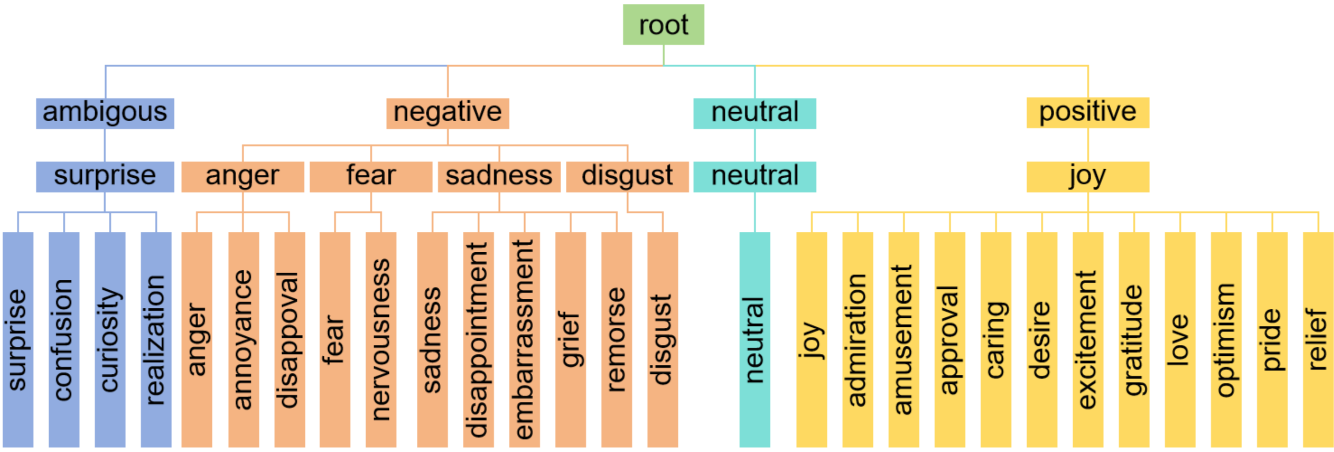


Fig. 5. GoEmotions Dataset Emotion Taxonomy Hierarchy.

score is calculated based on true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Precision (P) is the ratio of correctly predicted positive samples to the total predicted positive samples, P=TP/(TP+FP). Recall (R) is the ratio of correctly identified positive samples to the total actual positive samples, R=TP/(TP+FN). The F1 score is the harmonic mean of precision and recall, $F1=2\times P\times R/(P+R)$, The Macro-F1 score is the average F1 score across all categories, calculated as: $\text{Macro-F1}=\sum_{i=1}^{K}F1/K$, where $K$ is the number of categories.

To evaluate the performance of the EDA-BERT model on fine-grained emotion recognition tasks, the following models are used as baseline models for comparison:

- TextRCNN Model: The TextRCNN model combines convolutional neural network (CNN) and recurrent neural network (RNN) for text classification, followed by a max-pooling layer for feature selection and a fully connected layer for classification.
- BERT Model: BERT is a pre-trained language model that captures deep bidirectional representations of language through large-scale textual pre-training. It is fine-tuned for various natural language processing tasks. In this study, a fully connected layer is added to BERT for emotion recognition.
- KEA-BERT [16] Model: The KEA-BERT model is designed for emotion recognition by incorporating sentiment lexicon knowledge. It employs BERT to extract contextual representation from text and integrates emotional information from sentiment lexicons. A knowledge-embedded attention mechanism combines emotional information with contextual representations for fine-grained emotion recognition. KEA-BERT has two variants based on the integration approach: sentence-level fusion ($\text{KEA-BERT}_{\text{sentence}}$) and word-level fusion ($\text{KEA-BERT}_{\text{word}}$).
- HGCN-EC [35] Model: The HGCN-EC model is an emotion recognition framework based on hierarchical graph convolutional networks. It consists of a text encoder and a hierarchical structure encoder. The model employs a graph convolutional network to integrate textual feature representations with label hierarchy knowledge, which are subsequently fed into a fully connected layer for classification.

The experiments were conducted on hardware consisting of an Intel Core i9-11900K CPU, 16GB RAM, and an NVIDIA GeForce RTX 3070 GPU, utilizing PyTorch 1.7 and Python 3.7. For the TextRCNN model, 300-dimensional GloVe vectors were employed to initialize the word embeddings. The RNN layer utilized a bidirectional GRU (Bi-GRU) with a hidden layer size of 100. The training batch size and learning rate were set to 64 and 1e-4, respectively, with the Adam optimizer employed.

For the BERT, KEA-BERT, HGCN-EC, and EDA-BERT models, the training batch size and learning rate were set to 10 and 3e-5, respectively, also using the Adam optimizer. Additionally, a dropout parameter with a value of 0.2 was applied during training to randomly deactivate certain neurons, thereby enhancing the models' generalization ability. The detailed comparative experimental results are presented in Table 1, with the highest score for each emotion category highlighted in bold.

### *B. Main Experimental Results*

As demonstrated by the experimental results in Table I, no single model exhibits superior performance across all categories. Overall, the proposed EDA-BERT model achieves optimal performance, with its Macro-F1 score surpassing those of the other baseline models. Specifically, the EDA-BERT model outperforms the TextRCNN, BERT, KEA-BERTsentence, KEA-BERTword, and HGCN-EC models by 8.32%, 2.62%, 0.86%, 3.62%, and 0.6%, respectively.

In comparison to the TextRCNN and BERT models, which do not incorporate external knowledge, the performance of the EDA-BERT, HGCN-EC, and KEA-BERTsentence models, which integrate external knowledge, shows a marked improvement. This finding aligns with the results of many studies, suggesting that the introduction of external knowledge enriches text representation and effectively enhances the model's fine-grained emotion recognition performance.

TABLE I
COMPARISON OF F1 SCORES OF MULTIPLE EMOTION RECOGNITION MODELS ON THE GOEMOTIONS DATASET (%)

| Emotion | TextRCNN | BERT | KEA-BETR$_{sentence}$ | KEA-BERT$_{word}$ | HGCN-EC | EDA-BERT |
|---|---|---|---|---|---|---|
| neutral | 67.02 | 67.37 | 68.00 | 65.33 | **68.33** | 66.02 |
| admiration | 66.15 | 69.14 | 68.18 | 67.40 | 69.25 | **70.28** |
| amusement | 80.80 | 81.80 | 82.59 | 82.93 | **83.39** | 80.75 |
| anger | 48.45 | 46.80 | 50.45 | 51.01 | **51.38** | 49.18 |
| annoyance | 32.54 | 34.79 | 34.49 | 33.90 | 35.26 | **36.30** |
| approval | 37.75 | 38.86 | 38.59 | 38.82 | 38.98 | **40.81** |
| caring | 25.71 | 41.87 | 41.15 | 39.04 | 40.16 | **42.48** |
| confusion | 34.34 | 47.05 | 43.20 | 44.71 | **47.27** | 45.21 |
| curiosity | 52.75 | 56.05 | 57.49 | 55.23 | 58.13 | **59.26** |
| desire | 45.20 | 49.66 | **52.41** | 51.53 | 49.62 | 51.66 |
| disappointment | 15.95 | 29.39 | 33.60 | 30.82 | 32.64 | **34.84** |
| disapproval | 35.91 | 40.81 | 41.54 | 38.61 | 40.76 | **42.51** |
| disgust | 48.62 | 47.24 | 48.32 | 46.02 | 47.56 | **48.67** |
| embarrassment | 40.00 | 45.61 | 46.37 | 46.67 | 48.48 | **49.18** |
| excitement | 45.97 | 46.07 | 44.10 | 39.59 | 47.89 | 44.21 |
| fear | 64.47 | 66.30 | 66.66 | 66.29 | 66.74 | **67.80** |
| gratitude | 90.96 | 90.47 | 92.48 | 92.13 | 91.67 | **92.50** |
| grief | 0.0 | 0.0 | 0.0 | 0.0 | 0 | 0.0 |
| joy | 61.04 | 62.58 | **63.12** | 57.75 | 62.34 | 61.54 |
| love | 79.92 | 80.80 | 81.08 | 80.31 | 80.23 | **81.10** |
| nervousness | 0.0 | 31.58 | 37.20 | 34.44 | 35.33 | **40.00** |
| optimism | 57.72 | 54.68 | **55.93** | 53.40 | 55.22 | 54.45 |
| pride | 0.0 | 43.47 | **45.45** | 38.10 | 44.78 | **45.45** |
| realization | 19.21 | 23.14 | **25.70** | 21.30 | 24.11 | 23.11 |
| relief | 0.0 | 0.0 | 30.76 | 0.0 | 35.49 | **42.86** |
| remorse | 66.66 | 64.82 | 67.10 | 64.52 | 64.39 | **67.18** |
| sadness | 49.66 | **58.30** | 54.54 | 54.54 | 55.11 | 56.19 |
| surprise | 49.26 | **57.23** | 54.29 | 53.15 | 57.08 | 55.33 |
| Macro-F1 | 43.43 | 49.13 | 50.89 | 48.13 | 51.15 | **51.75** |

Among the models incorporating external knowledge EDA-BERT, HGCN-EC, KEA-BERTsentence, and KEA-BERTword the proposed EDA-BERT model performs better on the GoEmotions dataset. Specifically, the EDA-BERT model achieves a Macro-F1 score that is 0.86% and 3.62% higher than KEA-BERTsentence and KEA-BERTword, respectively. This indirectly demonstrates that the EDA-BERT model excels at extracting emotional features from text. Thus, for fine-grained emotion recognition tasks, using emotion distribution to quantitatively measure the interrelations between emotions effectively enhances emotion distinguishability, proving to be more beneficial than traditional sentiment lexicon knowledge in improving model performance. Compared to the HGCN-EC model, the EDA-BERT model achieves a Macro-F1 score 0.6% higher and outperforms HGCN-EC in several emotion categories, including admiration, approval, and relief. We attribute the advantage of EDA-BERT to its use of emotion distribution, constructed from emotional distances in the VAD model, as external knowledge input. In contrast to HGCN-EC, which relies on hierarchical knowledge between emotion categories, EDA-BERT excels at capturing the subtle differences between categories, thereby enhancing its ability to recognize fine-grained emotions.

Additionally, all models perform poorly on the four emotion categories: pride, relief, grief, and nervousness. We attribute this primarily to the limited number of samples in these categories in the training dataset (all fewer than 200 samples), coupled with significant differences between the training and testing sets. A small number of samples in emotion recognition tasks leads to limited information, hindering the models' ability to effectively recognize rare emotions. Specifically, for the grief category, with only 39 samples, the extremely low sample count results in a classification score of 0 for all models in this category. External knowledge can enrich the information available for rare emotion categories. The proposed EDA-BERT model

TABLE II
PERFORMANCE COMPARISON ON SECONDARY EMOTION CLASSIFICATION IN THE GOEMOTIONS DATASET (%)

| Emotion | TextRCNN | BERT | KEA-BETR$_{sentence}$ | KEA-BERT$_{word}$ | HGCN-EC | EDA-BERT |
|---|---|---|---|---|---|---|
| neutral | 66.71 | 68.22 | 69.49 | 67.29 | 69.13 | **70.58** |
| anger | 47.82 | **56.41** | 55.67 | 54.62 | 56.38 | 56.01 |
| disgust | 38.61 | 45.74 | 50.61 | 49.16 | 50.56 | **50.89** |
| fear | 32.54 | 64.22 | 64.28 | 64.62 | 65.24 | **66.16** |
| joy | 79.19 | 79.88 | 81.27 | 81.66 | 82.43 | **83.33** |
| sadness | 48.81 | **61.70** | 61.05 | 56.98 | 61.11 | 61.63 |
| surprise | 55.48 | 61.15 | 62.21 | 60.78 | **64.18** | 63.68 |
| Macro-F1 | 52.73 | 62.45 | 63.51 | 62.16 | 64.14 | **64.61** |

integrates emotion distribution, which models the correlations between emotions in the VAD space as external knowledge. This enables the model to learn related knowledge about rare emotions from the VAD distances between different emotions in the VAD space, thereby increasing the amount of available information. By utilizing this additional information, the model can better distinguish between the subtle differences in fine-grained emotions. Our EDA-BERT model significantly outperforms the other models in the pride, relief, and nervousness categories. For instance, in the relief category, all models except for EDA-BERT, HGCN-EC, and KEA-BERT$_{sentence}$ achieved a score of 0. Moreover, the EDA-BERT model's Macro-F1 score is 7.37% and 12.1% higher than those of the HGCN-EC and KEA-BERT$_{sentence}$ models, respectively.

Table II provides a comprehensive comparison of the performance metrics of various models on the seven-category secondary emotion classification tasks using the GoEmotions dataset. Among these models, EDA-BERT achieved the best overall performance by a notable margin. Its Macro-F1 score exceeded those of TextRCNN, BERT, KEA-BERTsentence, KEA-BERTword, and HGCN-EC by 11.88%, 2.16%, 1.1%, 2.45%, and 0.45%, respectively. In particular, EDA-BERT demonstrated superior performance in four emotion categories—neutral, disgust, fear, and joy—compared to the other models. These findings suggest that consolidating the original 28 fine-grained emotion categories into 7 broader secondary-level classes effectively reduces inter-class confusion, thereby significantly improving model performance.

### *C. Conversation Dataset Experiment*

In Table III, EDA-BERT demonstrates the strongest overall performance on the ED dataset, achieving a Macro-F1 score of 53.46%, which constitutes a 3.39 percentage point improvement over the standard BERT baseline (50.07%). In terms of precision, EDA-BERT achieved the highest score (59.71%) among all models, outperforming the second-best (KEA-BETRsentence) by a margin of 3.46 percentage points. This indicates that the model makes more reliable positive predictions, minimizing false positives in emotion classification. Although its recall score (48.41%) is slightly lower than that of HGCN-EC (50.38%), it still ranks among the top performers and demonstrates a favorable trade-off between precision and recall. The consistent improvements across multiple evaluation metrics validate the effectiveness of integrating VAD emotion knowledge into the model architecture. By embedding fine-grained affective information into the representation learning process, EDA-BERT enhances its emotional understanding capabilities, leading to more contextually appropriate and emotionally aware dialogue modeling. These findings demonstrate that the proposed enhancements are not only theoretically sound but also empirically beneficial in advancing the performance of emotion recognition in dialogue systems.

TABLE III
PERFORMANCE COMPARISON ON THE ED DATASET (%)

| Model | Precision | Recall | Macro-F1 |
|---|---|---|---|
| TextRCNN | 48.04 | 40.25 | 43.65 |
| BERT | 52.11 | 48.17 | 50.07 |
| KEA-BETR$_{sentence}$ | 56.25 | 49.52 | 52.94 |
| KEA-BERT$_{word}$ | 52.68 | 49.65 | 51.30 |
| HGCN-EC | 54.26 | **50.38** | 52.06 |
| EDA-BERT | **59.71** | 48.41 | **53.46** |

### *D.Ablation Experiment*

The VAD Emotion Distribution Augmented fine-grained emotion recognition, which integrates VAD knowledge through emotion distribution and an attention mechanism, was evaluated via ablation experiments on the GoEmotions dataset. These experiments aimed to verify the specific impact of these two techniques on the performance of the EDA-BERT model. To validate that emotion distribution captures more detailed relationships between emotions than directly using VAD scores, we replaced the emotion distribution module in EDA-BERT with a method that directly utilizes the VAD scores of affective words from the sentiment lexicon. To verify that combining the attention mechanism with external knowledge leads to more effective contextual representations, we averaged the emotion distribution knowledge $H_e$ and directly concatenated the semantic information $H_c$ for final emotion prediction. The ablation experiment results are shown in Table IV (where "✓" indicates the method was used).The experimental results in Table IV demonstrate that both emotion distribution and the attention mechanism enhance the model's performance. When both techniques are applied, the EDA-BERT model achieves optimal performance, with a 1.93% improvement in the Macro-F1 score. Using only emotion distribution results in a 1.4% improvement in the Macro-F1 score, whereas using only the attention mechanism leads to a 0.57% improvement. In comparison, the performance improvement is more substantial when emotion distribution is used.

When emotion distribution is used to quantitatively record the expression degree of affective words across each emotion category, compared to directly using VAD scores, the model's Macro-F1 score improves by 1.36%. This validates that, by measuring the similarity between emotions based on VAD distance in the VAD emotion space and modeling emotion distribution using this as the affective word label, the model can more precisely reflect the correlations between emotions, offering more effective external knowledge for fine-grained emotion recognition tasks.

Additionally, when the attention mechanism is employed, compared to directly concatenating averaged emotion distribution knowledge with semantic information, the model's Macro-F1 score improves by 0.53%. This demonstrates that the attention mechanism helps integrate external knowledge with the contextual representations of pre-trained language models, creating a more effective data representation for fine-grained emotion recognition, thereby enhancing the model's ability to distinguish fine-grained emotions.

TABLE IV
ABLATION STUDY OF THE EDA-BERT MODEL

| Emotion distribution | Attention mechanism | Macro-F1 |
|---|---|---|
| | | 49.82 |
| | ✓ | 50.39 |
| ✓ | | 51.22 |
| ✓ | ✓ | 51.75 |

*E. Case Study*

In this case study, we delve deeper into model performance using texts drawn from the GoEmotions dataset, specifically focusing on the nuances within its fine-grained emotion hierarchy. A significant challenge arises for models tasked with fine-grained emotion recognition when different specific emotions are subsumed under the same broader secondary emotion category. This inherent similarity, often characterized by overlapping linguistic cues and contextual triggers, makes it difficult for models to accurately discriminate between them. For instance, consider the emotion pairs "annoyed" and "angry," where the distinction might lie in the intensity or duration of the sentiment expressed; "nervousness" and "fear," which might both manifest through expressions of anxiety but differ in their underlying causes or future orientation; or "joy" and "excitement," often sharing exuberant language but potentially differing in the stability or expected outcome of the positive event. These emotion pairs, despite their subtle differences, share the same secondary-level classification within the GoEmotions taxonomy, presenting a formidable test for model discernment. Specifically, both "nervousness" and "fear" are categorized under the secondary emotion of "fear," implying a shared core sentiment that can obscure the finer distinctions. To illustrate this challenge concretely, Table V presents two illustrative example texts: the first is labeled with "nervousness," likely reflecting anticipatory anxiety or worry about an upcoming event, while the second is labeled with "fear," possibly indicating a more immediate, intense response to a perceived threat or danger. Analyzing how models handle such examples is crucial for understanding their limitations and for guiding the development of more sophisticated emotion recognition techniques capable of capturing these subtle emotional gradations.

TABLE V
TEXT EXAMPLES FROM THE GOEMOTIONS DATASET

| Text | Emotion label |
|---|---|
| I have a job interview tomorrow and I can't stop thinking about it. | nervousness |
| The thought of losing my job keeps me up at night. | fear |

Figure 6 illustrates segments of the predicted probability distributions produced by two models for two example texts from the GoEmotions dataset, which contains 28 emotion labels. Our focus is specifically on the classification of ambiguous emotions. In the example involving the emotion pair "nervousness" and "fear," it was observed that the BERT model tends to predict both texts as "fear." In contrast, the EDA-BERT model demonstrates improved prediction performance for the "nervousness" category. This improvement may be attributed to the modeling of affective distributions, where the VAD (Valence-Arousal-Dominance) space indicates a distance of 0.3 between "nervousness" and "fear." This suggests that the additional affective information enables the EDA-BERT model to more effectively distinguish between fine-grained emotions.

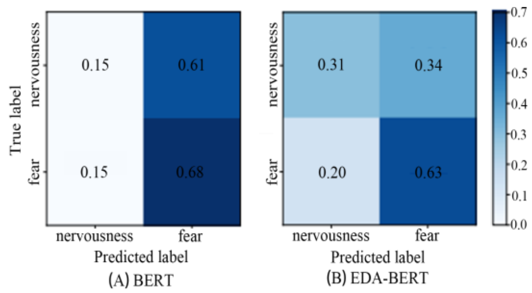


Fig. 6. Probability Distribution Comparison of Model Predictions.

## V. CONCLUSION

To address the challenge of fine-grained emotion recognition, this paper proposes the EDA-BERT model, which integrates VAD psychological model knowledge and enhances fine-grained emotion recognition via emotion distribution. By quantitatively modeling the expression intensity of affective words across various emotions, the EDA-BERT model effectively captures the correlations between fine-grained emotions in the psychological VAD emotion space. Experimental results on the GoEmotions dataset demonstrate that the EDA-BERT model outperforms other baseline models, confirming that emotion distribution contains richer emotion-related information.

In future research, we will explore methods to extract more relevant features from text to enrich the representation of emotion models. Additionally, we plan to investigate the use of more complex graph network structures or knowledge graphs to model the relationships between fine-grained emotions, further enhancing the model's ability to recognize these emotions.

## References


[1] B. Cardone, F. Di Martino, and S. Senatore, "Improving the Emotion-Based Classification by Exploiting the Fuzzy Entropy in FCM Clustering," *International Journal of Intelligent Systems*, vol. 36, no. 11, pp.6944-6967, 2021.

[2] H. J. Hu, Y. Y. Yang, Y. Yi, and M. F. Liu, "An Emotion Classification Method Based on Fine-Grained Information Perception BERT-EEP," *Computer Engineering & Science*, vol. 45, no. 4, pp.751-760, 2023.

[3] Y. P. Zou, D. Wang, D. Wang, C. Zheng, Q. Song, Y. Zhu, C. Fan, and K. Wu, "Research on Wearable Emotion Recognition Based on Multi-Source Domain Adversarial Transfer Learning," *Chinese Journal of Computers*, vol. 47, no. 2, pp.266-286, 2024.

[4] X. Deng, Z. Li, T. Zhang, and W. Zhou, "Portfolio Decision and Risk Analysis with Disappointment and Regret Emotions," *IAENG International Journal of Applied Mathematics*, vol. 55, no. 2, pp.407-421, 2025.

[5] P. Ekman, "An Argument for Basic Emotions," *Cognition & Emotion*, vol. 6, no. 3-4, pp.169-200, 1992.

[6] R. Plutchik, "The Nature of Emotions," *American Scientist*, vol. 89, no. 4, pp.344-350, 2001.

[7] H. T. Zhang, D. Wang, H. L. Xu, and S. Y. Sun, "Sentiment Classification of Micro-Blog Public Opinion Based on Convolution Neural Network," *Journal of the China Society for Scientific and Technical Information*, vol. 37, no. 7, pp.695-702, 2018.

[8] Z. Lei, Y. Yang, M. Yang, and Y. Liu, "A Multi-Sentiment-Resource Enhanced Attention Network for Sentiment Classification," in *Proceedings 56th Annual Meeting of the Association for Computational Linguistics (ACL)*, Melbourne, Australia, pp.758-763, 2018.

[9] D. Keltner, and J. S. Lerner, *Handbook of Social Psychology*. New Jersey: Wiley, pp.317-352, 2010.

[10] A. S. Cowen, H. A. EIfenbein, P. Laukka, and D. Keltner, "Mapping 24 Emotions Conveyed by Brief Human Vocalization," *American Psychologist*, vol. 74, no. 6, pp.698-712, 2019.

[11] A. S. Cowen, and D. Keltner, "What the Face Displays: Mapping 28 Emotions Conveyed by Naturalistic Expression," *American Psychologist*, vol. 75, no. 3, pp.349-364, 2020.

[12] D. Demszky, D. Movshovitz-Attias, J. Ko, A. Cowen, G. Nemade, and S. Ravi, "GoEmotions: A Dataset of Fine-Grained Emotions," *in Proceedings 58th Annual Meeting of the Association for Computational Linguistics (ACL)*, pp.4040-4054, 2020.

[13] S. Dhar, and I. Bose, "Emotions in Twitter Communication and Stock Prices of Firms: The Impact of Covid-19 Pandemic," *Decision*, vol. 47, no. 4, pp.385-399, 2020.

[14] P. Zhong, D. Wang, and C. Miao, "Knowledge-Enriched Transformer for Emotion Detection in Textual Conversations," in *Proceedings 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, Hong Kong, China, pp.165-176, 2019.

[15] J. A. Russell, and A. Mehrabian, "Evidence for a Three-Factor Theory of Emotions," *Journal of Research in Personality*, vol. 11, no. 3, pp.273-294, 1977.

[16] V. Suresh, and D. C. Ong, "Using Knowledge-Embedded Attention to Augment Pre-Trained Language Models for Fine-Grained Emotion Recognition," in *Proceedings 2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII)*, Nara, Japan, pp.1-8, 2021.

[17] J. S. Lai, Z. Y. Wan, and X. Q. Zeng, "The Image Emotion Distribution Learning Based on Emotion Wheel and Multi-task Convolutional Neural Network," *Journal of Jiangxi Normal University (Natural Science)*, vol. 46, no. 4, pp.363-371, 2022.

[18] K. Margatina, C. Baziotis, and A. Potamianos, "Attention-Based Conditioning Methods for External Knowledge Integration," in *Proceedings 57th Annual Meeting of the Association for Computational Linguistics (ACL)*, Florence, Italy, pp.3944-3951, 2019.

[19] L. De Bruyne, P. Atanasova, and I. Augenstein, "Joint Emotion Label Space Modeling for Affect Lexica," *Computer Speech & Language*, vol. 71, Art no. 101257, 2022.

[20] N. Alswaidan, and M. E. B. Menai, "A Survey of State-of-the-Art Approaches for Emotion Recognition in Text," *Knowledge and Information Systems*, vol. 62, no. 8, pp.2937-2987, 2020.

[21] E. Batbaatar, M. Li, and K. H. Ryu, "Semantic-Emotion Neural Network for Emotion Recognition from Text," *IEEE Access*, vol. 7, pp.111866-111878, 2019.

[22] R. S. EI-Sayed, "A Hybrid CNN-LSTM Deep Learning Model for Classification of the Parkinson Disease," *IAENG International Journal of Applied Mathematics*, vol. 53, no. 4, pp.1427-1436, 2023.

[23] N. H. Ho, H. J. Yang, S. H. Kim, and G. Lee, "Multimodal Approach of Speech Emotion Recognition Using Multi-Level Multi-Head Fusion Attention-Based Recurrent Neural Network," *IEEE Access*, vol. 8, pp.61672-61686, 2020.

[24] J. A. Adisa, S. Ojo, P. A. Owolawi, A. Pretorius, and S. O. Ojo, "The Effect of Imbalanced Data and Parameter Selection via Genetic Algorithm Long Short-Term Memory (LSTM) for Financial Distress Prediction," *IAENG International Journal of Applied Mathematics*, vol. 53, no. 3, pp.796-809, 2023.

[25] J. Devlin, M. W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding," in *Proceedings 2019 Conference North American Chapter of the Association for Computational Linguistics*: *Human Language Technologies(NAACL-HLT)*, Minneapolis, USA, pp.4171-4186, 2019.

[26] F. A. Acheampong, H. Nunoo-Mensah, and W. Chen, "Transformer Models for Text-Based Emotion Detection: A Review of BERT-Based Approaches," *Artificial Intelligence Review*, vol. 54, no. 8, pp.5789-5829, 2021.

[27] Q. Li, C. Wei, J. Dang, L. Cao, and L. Liu, "Tracking and Analyzing Public Emotion Evolutions During COVID-19: A Case Study from the Event-Driven Perspective on Microblogs," *International Journal of Environmental Research and Public Health*, vol. 17, no. 18, pp.6888, 2020.

[28] E. Troiano, L. Oberländer, and R. Klinger, "Dimensional Modeling of Emotions in Text with Appraisal Theories: Corpus Creation, Annotation Reliability, and Prediction," *Computational Linguistics*, vol. 49, no. 1, pp.1-72, 2023.

[29] A. Roy, and S. Pan, "Incorporating Extra Knowledge to Enhance Word Embedding," in *Proceeding of the 29th International Joint Conference on Artificial Intelligence (IJCAI)*, Yokohama, Japan, pp.4929-4935, 2021.

[30] R. Wang, D. Tang, N. Duan, Z. Wei, and X. Huang, "K-Adapter: Infusing Knowledge into Pre-trained Models with Adapters," in *Proceedings Findings of the Association for Computational Linguistics (Findings of ACL)*, Online, pp.1405-1418, 2021.

[31] A. R. Abas, I. Elhenawy, M. Zidan, and M. Othman, "BERT-CNN: A Deep Learning Model for Detecting Emotions from Text," *Computers, Materials & Continua*, vol. 71, no. 2, pp.2943-2961, 2022.

[32] P. Kumar, and B. Raman, "A BERT Based Dual-Channel Explainable Text Emotion Recognition System," *Neural Networks*, vol. 150, pp.392-407, 2022.

[33] S. M. Mohammad, And P. D. Turney, NRC Emotion Lexicon(Version 0.92). Ottawa: National Research Council of Canada, 2013.

[34] S. Poria, A. Gelbukh, E. Cambria, A. Hussain, and G. B. Huang, "EmoSenticSpace: A Novel Framework for Affective Common-Sense Reasoning," *Knowledge-Based Systems*, vol. 69, pp.108-123, 2014.

[35] X. F. Nie, Y. X. Tan, X. Q. Zeng, and J. L. Zuo, "Emotion Recognition Based on Hierarchical Graph Convolution Networks," *Journal of Chinese Information Processing*, vol. 38, no. 6, pp.139-150, 2024.


**Sufen Chen** received the B.S. and M.S. degrees from Jiangxi Normal University in 2002 and 2005 respectively. Currently, she is an associate professor at the Jiangxi University of Water Resources and Electric Power.

Her current research interests include natural language processing and affective learning.

**Chunyang Li** is a graduate student at Jiangxi Normal University. His main research interest is multiple sentiment analysis.

**Xueqiang Zeng** received the B.S., M.S., and Ph.D. degrees from Nanchang University, Jiangxi Normal University and Shanghai University, in 1999, 2005, and 2009, respectively. He is currently a Professor with the School of Computer and Information Engineering, Jiangxi Normal University.

His current research interests include natural language processing, educational data mining and affective learning. He has published more than 60 refereed technical papers in journals and conferences.