# BGEK: External Knowledge-Enhanced Graph Convolutional Networks for Rumor Detection in Online Social Networks

Xiaoda Wang, Chenxiang Luo, Tengda Guo, Zhangrui Liu, Jiongyan Zhang, and Haizhou Wang$^{(\boxtimes)}$

School of Cyber Science and Engineering, Sichuan University, Chengdu 610207, China
{wangxiaoda,iridescense,guotengda, liuzhangrui,zhangjiongyan}@stu.scu.edu.cn, whzh.nc@scu.edu.cn

**Abstract.** Nowadays, social media has become a dominant platform for disseminating news and information. However, it also accelerated the spread of rumors, which causes great impacts on the real world. Therefore, the detection of rumors plays a crucial role in controlling the diffusion of misinformation. Cantonese is widely spoken across the world, yet limited research has been conducted on Cantonese rumor detection. Moreover, current detection methods primarily focus on extracting textual and propagation structure characteristics without utilizing external knowledge to enhance the performance of identifying rumors. In this paper, we propose a novel model using Bidirectional Graph Convolutional Networks Embedded with External Knowledge, namely BGEK, for Cantonese rumor detection. Specifically, we first construct a directed heterogeneous knowledge graph using official statements and related entity descriptions from Wikipedia to obtain external knowledge embeddings. Secondly, we use BERT (Bidirectional Encoder Representations from Transformers) model to extract text features from source tweets and obtain the correlation vectors of external knowledge and text features through the Comparison Network. Thirdly, we utilize Bidirectional Graph Convolutional Networks to extract rumor propagation structural features. Finally, we fuse text features, structural features and comparison features to construct a Cantonese rumor detection model. To the best of our knowledge, we are the first to apply external knowledge to Cantonese rumor detection. Experimental results demonstrate that the BGEK model outperforms existing state-of-the-art detection models.

**Keywords:** Cantonese rumors · rumor detection · external knowledge · Bidirectional Graph Convolutional Networks · feature fusion

## 1 Introduction

With the rapid development of the Internet, social media has become an extremely popular source for individuals to obtain news information and express

their opinions. According to the statistics in Digital 2022[1], social media users have escalated to 4.62 billion accounting for 58.4% of the global population, with an annual growth rate exceeding 10%. For instance, Facebook has emerged as the most influential social media platform, with a worldwide registration of 2.96 billion individuals[2]. At the same time, the rapid growth of social media platforms has facilitated the dissemination of rumors which can cause serious social panic [1].

Cantonese originated in Guangdong Province of China, and is presently one of the most extensively spoken languages by native and overseas Chinese. With over 82.4 million speakers[3], it has spread throughout overseas Chinese communities. The rapid dissemination of Cantonese rumors through social media has resulted in serious harm to society. Therefore, it is essential to develop an intelligent method to identify rumors automatically.

Most existing rumor detection works focus on detecting rumors in Chinese and English [1,13,16,19,21], but there are few works considering Cantonese rumors [3,7,10]. They use the textual information of tweets and feature engineering for Cantonese rumor detection [3,7]. However, Cantonese rumor detection still faces two critical challenges that still need to be resolved. **Firstly, existing studies in Cantonese rumor detection mainly rely on text and user statistic features** [3,7]**.** These approaches don't take into account the dissemination structures of retweets and comments, and there is no available structural Cantonese rumor dataset that includes source tweets, retweets, and comments. Therefore, the absence of a benchmark Cantonese rumor dataset makes the detection of Cantonese rumors a daunting task. **Secondly, despite the successful works achieved by current rumor detection methods, external knowledge graphs are rarely integrated into their existing methods.** Existing knowledge graph-based approaches mainly extract the structural triplets (head, relation, tail) from the tweet text to compare with faithful triples from knowledge graph to predict the truthfulness of tweets [5,14]. But they don't consider the information in the retweets, and comments and don't embed external knowledge into source tweets to assist in the judgment of rumors. Meanwhile, existing research has applied Graph Convolutional Networks for detecting rumors, but there is no research on combining external knowledge with graph convolutional networks for rumor detection.

To address the above two challenges in Cantonese rumor detection, firstly, based on Facebook, we build and label a structural Cantonese rumors benchmark dataset, including source tweets, retweets, and comments. Secondly, we propose a novel feature extraction method based on Graph Convolutional Networks for external knowledge integration. This approach involves constructing a heterogeneous graph using official statements and Wikipedia entity descriptions and utilizing BERT to capture the text features of the source tweet. To generate

---

correlation vectors, a Comparison Network is employed to compare the external knowledge with source text. The Bidirectional Graph Convolutional Networks is then utilized to extract top-down and bottom-up propagation features of tweets during the rumor propagation process. Finally, these three types of features are fused to train a Cantonese rumor detection model. The main contributions of this paper are summarized as follows:

- **To the best of our knowledge, we are the first to construct a structural Cantonese rumor dataset containing source tweets, retweets and comments in social networks, which is publicly available on GitHub**[4]. Specifically, we crawled 2,721 source tweets and 91,246 comments and retweets. After data cleaning and labeling, a Cantonese rumor dataset was constructed including 1,925 source tweets, 64,221 comments and retweets.
- **We construct a novel method for extracting external knowledge features based on Graph Convolutional Networks and obtaining correlation features through Comparison Network**. First, we construct a heterogeneous knowledge graph based on official statements and entity descriptions from Wikipedia. Then, we use a Heterogeneous Graph Convolutional Networks to extract the embedded features of external knowledge. Finally, we obtain correlation vectors from external knowledge embedded and tweet text features by Comparison Network.
- **We propose a novel Cantonese rumor detection model BGEK for Cantonese rumor detection, which integrates the text features, comparison features, and structural features of tweets. To the best of our knowledge, we are the first to apply external knowledge to Cantonese rumor detection.** Our experimental results demonstrate that the BGEK detection model achieves remarkable detection results and outperforms other state-of-the-art baseline models.

## 2   Related Work

Traditional rumor detection methods are based on machine learning and they mainly focus on text and statistical features, aiming to train a classifier for rumor detection through supervised learning. A series of methods have been proposed [2,9,20], such as Random Forest [9], Decision Tree [2] and Support Vector Machine (SVM) [20]. The above methods rely heavily on feature engineering, which takes plenty of time, and the detection performance is not ideal.

In order to automatically extract features of rumors, a series of methods based on deep learning have been proposed [12,13,15,17,21]. They mostly use text features for rumor detection, for example, Recurrent Neural Network (RNN) [12], Convolutional Neural Network (CNN) [21], Structure Recurrent Neural Network (RvNN) [13]. Song et al. [15] proposed an adversarial awareness rumor detection framework. Sun et al. [17] applied contrastive learning to rumor detection. However, the currently available methods prove to be inefficient in their ability to learn the features of the propagation structure, and they exhibit a lack of consideration for the global structural characteristics of rumor dispersion.

---

[4] https://github.com/yiyepianzhounc/BGEK.

Compared with deep learning-based rumor detection mentioned above, Graph Convolutional Networks [8] has been applied in the field of rumor detection in recent years [1,11,19]. Bian et al. [1] used the Graph Convolutional Networks for the first time in the social network rumor detection field. Wei et al. [19] proposed an Edge-enhanced Bayesian Graph Convolutional Networks to obtain robust node feature representations. Lu et al. [11] proposed a Graph-aware Co-Attention Network (GCAN) for interpretable disinformation detection.

In recent studies, the significance of external knowledge for rumor detection has been recognized, leading to the development of rumor detection models that incorporate knowledge graph enhancement [5,6,14,16]. The methods [5,14] utilized tuples (head, relationship, tail) from the text to compare with equivalent elements in knowledge graphs. They mainly extracted the structural triplets (head, relation, tail) from the tweet text to compare with faithful triples from knowledge graph to predict the truthfulness of tweets.

Existing research on rumor detection in Cantonese [3,7,10] utilizes text and user information for rumor detection. Our previous work [3,7] combined Cantonese tweets' semantic and statistical features to detect Cantonese rumors on the Twitter platform. Lin et al. [10] proposed an annotation system that facilitates manual fact-checking. Existing Cantonese rumor detection methods mainly focus on text information and statistical features and do not use external knowledge graphs which can provide corresponding evidence for rumor detection. They also don't utilize the structural characteristics of the propagation and dispersion of rumors and limit the improvement of detection performance.

## 3   Methodology

In this section, we will describe the Cantonese rumor detection model based on the Bidirectional Graph Convolutional Network embedding external knowledge.

### 3.1   Dataset Construction

Existing Cantonese rumor datasets [3,7] mainly focus on source tweets and user information. Deep learning models can be built based on these existing datasets. However, there is no structural Cantonese rumor dataset including retweets and comments which is of great significance for the research of Graph Neural Networks based Cantonese rumor detection. Facebook has become one of the essential social media websites for users to obtain news information and a large number of Cantonese rumors spread on Facebook. So we choose Facebook social platform as the research object and construct a wholly structural Cantonese rumor dataset. To obtain the structural information, the Selenium framework is utilized to crawl data due to the Facebook's restrictions on data crawling. Then the data is manually labeled according to the method of [3]. After data cleaning and screening, we finally construct a Cantonese rumor dataset, named Facebook-C-Dataset, which contains 1,925 source tweets, 64,221 comments and retweets. The tweets are further classified into three major domains, namely

society, health, and information technology, and 49 specific topics, such as cancer, chronic diseases, radiofrequency radiation, and COVID-19 vaccine. Table 1 shows the detailed composition of the dataset.

**Table 1.** Statistics of the Facebook-C-Dataset

|  | Number of Tweets | Number of Retweets and Comments | Date of Tweets | Domains | Topics |
|---|---|---|---|---|---|
| Rumor | 565 | 20,625 | 01/2020-03/2023 | 3 | 49 |
| Non-rumor | 1,360 | 43,596 |  |  |  |
| Total | 1,925 | 64,221 |  |  |  |

## 3.2 BGEK Rumor Detection Model

The Cantonese rumor detection model BGEK is shown in Fig. 1. In order to better utilize external facts, we use a Comparison Network to embed external knowledge into the text representation of tweets. At the same time, we use a Bidirectional Graph Convolutional Networks, which fully uses the contribution of retweets and comments generated in the process of tweet dissemination to rumor detection. The specific implementation details are as follows:

**Tweet Propagation Graph Construction.** Given the source tweet, retweet, and comment information of a event, we can represent the total rumor dataset as $C = \{c_1, c_2, ..., c_m\}$, where $c_i$ denotes the $i-th$ event and $m$ is the number of events. $c_i$ can be denoted as $c_i = \{r_i, t_1^i, t_2^i, ..., t_{n_i-1}^i, G_i\}$, where $t_j^i$ denotes the $j-th$ responsive tweet, $n_i$ denotes the total number of comments and retweets contained in the event $c_i$, $G_i$ represents the rumor propagation graph composed of event $c_i$. $G_i$ is defined as $G_i = <V_i, E_i>$, where the node set is $V_i = \{r_i, t_1^i, t_2^i, ..., t_{n_i-1}^i\}$, $r_i$ is the root node in the propagation graph, and $E_i = \{e_{st}^i | s, t = 0, 1, ..., n_i - 1\}$ is the edge set. Each $e_{st}^i$ represents the directed relationship among tweets, retweets, and comments. For an adjacency matrix $A_i \in \mathbb{R}^{n_i \times n_i}$ , the initial value can be calculated as:

$$a_{st}^i = \begin{cases} 1, & if \quad e_{st}^i \in E \\ 0, & otherwise \end{cases}. \tag{1}$$

For each event $c_i$, there is a corresponding label $y_i \in Y$, where $Y$ represents different categories of events, our goal is to train a classifier $f : C \rightarrow Y$.

**Structural Feature Extraction.** Based on the relationship between source tweets, retweets, and comments, we constructed a propagation graph $G_i = <V_i, E_i>$ for each event $c_i$ and then built the adjacency matrix $A_i \in \mathbb{R}^{n_i \times n_i}$. We constructed text features $x_i$ for each node in the graph, and the feature matrix can be represented as $X = \{x_1, x_2, x_3, ..., x_{n_i}\}$, where $n_i$ represents the total number of comments and retweets in the event $c_i$. We used a Bidirectional
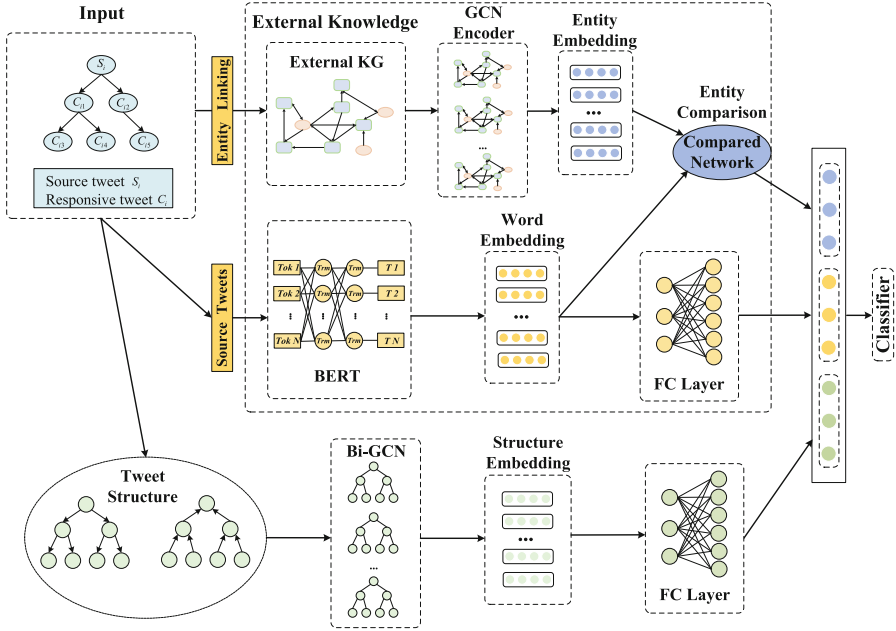
**Fig. 1.** BGEK rumor detection model

Graph Convolutional Networks (Bi-GCN) to calculate the node representations in the graph, which includes a top-down Graph Convolutional Networks (TD-GCN) and a bottom-up Graph Convolutional Networks (BU-GCN). The adjacency matrices for TD-GCN and BU-GCN can be represented as $A^{TD} = A$ and $A^{BU} = A^T$ respectively. The top-down and bottom-up propagation features can be obtained by two layers of GCN as follows:

$$H_1^{BU} = \sigma(\tilde{A}^{BU} X W_0^{BU}) \tag{2}$$

$$H_2^{BU} = \sigma(\tilde{A}^{BU} H_1^{BU} W_1^{BU}) \tag{3}$$

where $\tilde{A}^{BU}$ is the regularized adjacency matrix of $A^{BU}$, $H_1^{BU}$, $H_2^{BU}$ and $W_0^{BU}$, $W_1^{BU}$ are the hidden features and weight matrix respectively, and $\sigma$ is the activation function. Similarly, the top-down hidden features $H_1^{TD}$, $H_2^{TD}$ can be obtained by the above equation. Meanwhile, in order to make full use of the features of the source tweets, we concatenate the root node features of the $k-1$ layer with the hidden layer features of the $k$ layer as follows:

$$\tilde{H}_k^{BU} = concat(H_k^{BU}, (H_{k-1}^{BU})^{root}) \tag{4}$$

Through the propagation and dispersion features $\tilde{H}_2^{TD}$, $\tilde{H}_2^{BU}$ obtained above, the propagation features and dispersion features can be connected to obtain the structural features:

$$T = concat(\tilde{H}_2^{TD}, \tilde{H}_2^{BU}) \tag{5}$$

**External Knowledge Extraction**

*External Knowledge Graph Construction.* We construct a heterogeneous graph $\omega = < V, E >$ for the types of source tweets, including official statements and entity descriptions. The graph contains two different types of nodes: official statements $R = \{V_1^r, V_2^r, V_3^r, ..., V_x^r\}$ and entity descriptions $D = \{V_1^d, V_2^d, V_3^d, ..., V_y^d\}$, where $x$ represents the number of official statements and $y$ represents the number of entity descriptions. And the set of edges $E$ includes bidirectional connections and unidirectional connections. The construction method for the external knowledge graph is outlined as follows:

The source tweet contains M-specific aspects, which can be expressed as $S = \{s_1, s_2, s_3, ..., s_M\}$. The source tweet may belong to multiple aspects, and the content of tweets under the same aspect has a particular content similarity. First, we bidirectionally connect the official statements constructed under each aspect. Then for the entities contained in the official statement and the source tweet, we connect the entity to the entry on Wikipedia and select the content of the first paragraph as the entity description. Because the entity description and the official statement are related, we link the official statement and the entity description under the same aspect bidirectionally. Considering that the official statement corresponding to the same type of aspect has a certain similarity, we link the same type of official statement bidirectionally. Because an original tweet may belong to multiple aspects, we create an undirected connection edge between entity descriptions of the same type and other entity descriptions.

*Heterogeneous Graph Convolutional Networks Construction.* Through the directed heterogeneous graph $\omega = <V, E>$ constructed above, we use a Heterogeneous Graph Convolutional Networks to represent and learn official statements and entity descriptions. We use the Cantonese corpus we constructed to fine-tune the BERT model based on Chinese pre-training. Then we use BERT to obtain the node embedding feature matrix $X \in \mathbb{R}^{|V| \times D}$, where $X = \{x_1, x_2, x_3, ..., x_{|V|}\}$ includes the features of all nodes on the heterogeneous graph and $x_i$ represents the feature of the $i-th$ node. We define $A$ as the adjacency matrix and $D$ as the degree matrix. Then the heterogeneous graph convolutional layer updates the $i+1-th$ layer clustering features by clustering the features of the $i-th$ layer adjacency matrix:

$$A' = D^{-\frac{1}{2}}(A + I)D^{-\frac{1}{2}} \tag{6}$$

$$H^{(i+1)} = \sigma(A'H^iW^i) \tag{7}$$

where $I$ is the identity matrix of $|V|$ dimension, $A'$ is the adjacency matrix after self-connection and regularization, $W^i$ is the weight matrix of $i-th$ layer, $H^i$ is the feature matrix of $i-th$ layer, $\sigma$ is the activation function.

**Comparison Feature Extraction.** We get the embedded representation of external knowledge $K = \{k_1, k_2, k_3, ..., k_{|V|}\}$ through the above-mentioned Heterogeneous Graph Convolutional Networks. The text of the source tweet can be expressed as

$$T = \{t_1, t_2, t_3, ..., t_{|C|}\} \tag{8}$$

where $|C|$ represents the number of source tweets in the dataset. We fine-tuned the BERT model based on Chinese pre-training through the constructed Cantonese corpus, and then the text features can be obtained through the features as follows:

$$B = BERT(T) . \tag{9}$$

where $B = \{b_1, b_2, b_3, ..., b_{|C|}\}$ is the text feature of the source tweet, and then we get the comparison vector by comparing the text feature $b_n$ of the source tweet with the knowledge embedding feature $k_n$:

$$C_n = f_{cmp}(b_n, k_n) \tag{10}$$

where $f_{cmp}()$ is the comparison function. Based on [6], we designed the comparison function as $f_{cmp}(x, y) = W[x - y, x \odot y]$. Where $W$ is the dimension transformation matrix, $x$ and $y$ are the text features of the source tweets and knowledge embedding feature vectors, and $\odot$ represents the Hadamard product.

**Feature Concatenation.** Initially, we perform concatenation of various feature sets including source tweet text features $B_n$, comparison features $C_n$ and structural features $T_n$ based on retweets and comments to obtain vector

$$F_n = concat(B_n + C_n + T_n) \in \mathbb{R}^{|B_n|+|C_n|+|T_n|} \tag{11}$$

Then $F_n$ is subsequently fed as input to the Softmax layer, which can be represented as

$$Z = Softmax(WF_n + b) \tag{12}$$

Here, $W$ is the parameter matrix of the fully connected layer, and $b$ is the bias matrix of the fully connected layer.

## 4    Experiments

### 4.1    Experiment Settings

In our work, all experiments are conducted on an NVIDIA A100-SXM4 workstation with 80G of memory. The dataset used in the experiments is the Facebook-C-Dataset. To extract structural features, TF-IDF scores are applied to the top 5,000 words in the source tweets, retweets, and comments. In evaluating the model's performance, accuracy, precision, recall, and F1 scores are used, along with ten-fold cross-validation, to provide an average evaluation metric value.

### 4.2    Performance Comparison with Baselines

To assess the efficacy of our proposed BGEK model, we employ a comparative analysis approach with various state-of-the-art baselines. The results of different models for Cantonese rumor detection are shown in Table 2. Text-based features are represented by T, external knowledge-based features are represented by K,

and structural features are denoted by S. Among the baseline models, BERT outperforms other models with only text information, which means text features of source tweets are important for identifying rumors. Our proposed BGEK model integrates the text features, comparison features, and structural features and achieves the best performance in terms of all metrics.

**Table 2.** Results of comparison among different models in Facebook-C-Dataset

| Model | Features | | | Acc. | Non-Rumor | | | Rumor | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | T | S | K | | Prec. | Rec. | F1 | Prec. | Rec. | F1 |
| RFC [9] | ✓ | | | 0.8029 | 0.7915 | 0.9795 | 0.8755 | 0.8833 | 0.3759 | 0.5274 |
| TextCNN [21] | ✓ | | | 0.8672 | 0.8800 | 0.9642 | 0.9119 | 0.8214 | 0.6571 | 0.7302 |
| RNN_Att [22] | ✓ | | | 0.8750 | 0.8812 | 0.9570 | 0.9175 | 0.8519 | 0.6571 | 0.7419 |
| SVM [20] | ✓ | | | 0.7884 | 0.8405 | 0.8651 | 0.8526 | 0.6489 | 0.6028 | 0.6250 |
| BERT [4] | ✓ | | | 0.9135 | 0.9196 | 0.9522 | 0.9325 | 0.8448 | 0.7703 | 0.7924 |
| Transformer [18] | ✓ | | | 0.8047 | 0.8400 | 0.9032 | 0.8705 | 0.6786 | 0.5429 | 0.6032 |
| RvNN [13] | ✓ | ✓ | | 0.7332 | 0.8099 | 0.8244 | 0.8171 | 0.5196 | 0.4953 | 0.5072 |
| BiGCN [1] | ✓ | ✓ | | 0.7813 | 0.7824 | 0.7501 | 0.7635 | 0.6204 | 0.6662 | 0.6342 |
| EBGCN [19] | ✓ | ✓ | | 0.8223 | 0.8679 | 0.8855 | 0.8758 | 0.4608 | 0.4230 | 0.4374 |
| BGEK (Ours) | ✓ | ✓ | ✓ | **0.9727** | **0.9807** | **0.9792** | **0.9791** | **0.9536** | **0.9606** | **0.9528** |

### 4.3   Ablation Experiment

Our proposed BGEK model integrates external knowledge features, text features, and propagation structure features. To assess the individual contribution of each feature on the BGEK model, we conducted ablation experiment using the variants shown in Table 3. T represents the text features extracted from the source tweets, K represents the external knowledge features, and S represents the structural features originating from the source tweets, retweets and comments.

**Table 3.** The description of different variants

| Feature set | Features categories |
|---|---|
| BGEK | Source text feature, Structural feature, External knowledge feature |
| BGEK/K | Source text feature, Structural feature |
| BGEK/S | Source text feature, External knowledge feature |
| BGEK/T | Structural feature, External knowledge feature |

The evaluation metrics obtained from the experiment on different variants are shown in Fig. 2. The results shows all variants perform worse than complete BGEK model integrating the text features, comparison features, and structural

features. External knowledge feature plays the most important role in the identification of Cantonese rumors, which also illustrates the necessity of external knowledge information for better performance.

### 4.4   Embedding Visualization

To visually represent the feature embedding, Fig. 3 exhibits the embedding outcomes of diverse baselines on the Facebook-C-Dataset. From the plots, we find that our proposed BGEK model outperforms the other methods, by effectively segregating rumor and non-rumor information.
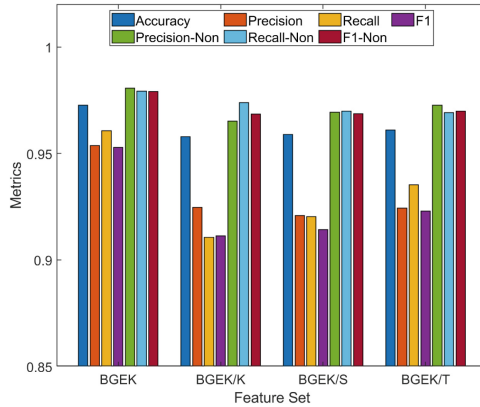


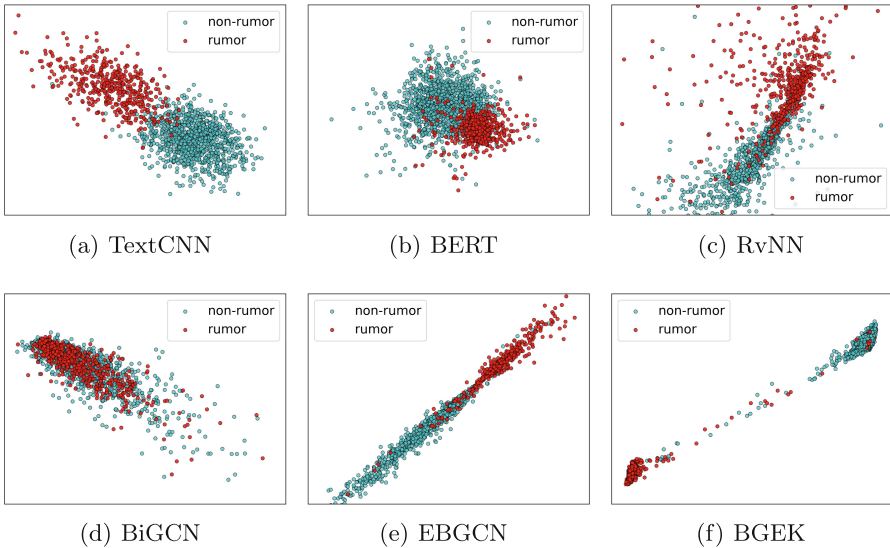**Fig. 2.** Results of ablation experiment among different variants



(a) TextCNN               (b) BERT               (c) RvNN

(d) BiGCN               (e) EBGCN               (f) BGEK

**Fig. 3.** Visualizing embeddings among different models

### 4.5 Robustness Experiment

In this experiment, a randomized portion of labels from the training set is intentionally mislabeled at varying ratios ranging from 5% to 45%. Following this, the model is re-trained on the modified training set and tested for its ability to withstand noise at different levels. The experimental results are shown in Fig. 4, which show that as the noise rate increases, the performance of all models (F1 score) decreases. Notably, our proposed model exhibits the most robust performance, displaying the smallest decline compared to other baseline models.
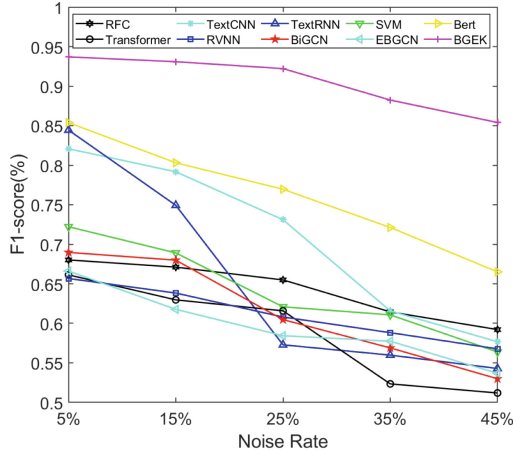


**Fig. 4.** Results of robustness experiment

### 4.6 Transferability Experiments on Twitter15 and Twitter16

To validate the efficacy of our proposed BGEK model for detecting rumors, we conducted comparative experiments on the Twitter15 and Twitter16 datasets using the aforementioned baselines.

**Table 4.** Results of comparison among different models in Twitter15 and Twitter16

| Model | Twitter15 | | | | | Twitter16 | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Acc. | NR | F | T | U | Acc. | NR | F | T | U |
| | | F1 | F1 | F1 | F1 | | F1 | F1 | F1 | F1 |
| DTC [2] | 0.455 | 0.933 | 0.355 | 0.317 | 0.415 | 0.465 | 0.643 | 0.393 | 0.419 | 0.403 |
| SVM-TS [20] | 0.544 | 0.796 | 0.472 | 0.404 | 0.483 | 0.544 | 0.796 | 0.472 | 0.404 | 0.483 |
| BERT [4] | 0.896 | 0.680 | 0.794 | 0.921 | 0.767 | 0.896 | 0.692 | 0.840 | 0.928 | 0.739 |
| RvNN [13] | 0.852 | 0.840 | 0.846 | 0.884 | 0.837 | 0.852 | 0.840 | 0.846 | 0.884 | 0.837 |
| BiGCN [1] | 0.871 | 0.860 | 0.867 | 0.913 | 0.836 | 0.885 | 0.829 | 0.899 | 0.932 | 0.882 |
| EBGCN [19] | 0.892 | 0.869 | 0.897 | 0.934 | 0.867 | 0.915 | 0.879 | 0.906 | 0.947 | 0.910 |
| BGEK (Ours) | **0.906** | **0.882** | **0.907** | **0.939** | **0.886** | **0.927** | **0.893** | **0.929** | **0.955** | **0.931** |

Each event is labeled as Non-rumor (NR), False Rumor (F), True Rumor (T), or Unverified Rumor (U). Due to the lack of external knowledge graphs, our proposed model does not use comparison features. The experimental outcomes are presented in Table 4. Among all baseline models, BERT achieves the best accuracy in Twitter15 and EBGCN achieves the best accuracy in Twitter16. By fusing text features and structural features, our model achieves the best detection results. And it unequivocally shows the generalization ability of our model.

## 5   Conclusion

In this paper, we construct a structural Cantonese rumor dataset based on the Facebook platform. Additionally, we propose a novel approach for extracting embedding features that incorporate official statements and entity descriptions of external knowledge by utilizing a Heterogeneous Graph Convolutional Networks. A Comparison Network is proposed to generate comparison features by comparing external knowledge with source tweets. Subsequently, we propose a novel Cantonese rumor detection model named BGEK (Bidirectional Graph Convolutional Networks Embedded with External Knowledge) that integrates text features, comparison features, and structural features. Five main experiments are conducted to evaluate the performance of our proposed BGEK model and experimental results demonstrate that our proposed model outperforms other state-of-the-art models.

## References

1. Bian, T., et al.: Rumor detection on social media with bi-directional graph convolutional networks. In: Proceedings of the 34th AAAI Conference on Artificial Intelligence, pp. 549–556 (2020)
2. Castillo, C., Mendoza, M., Poblete, B.: Information credibility on twitter. In: Proceedings of the 20th International Conference on World Wide Web, pp. 675–684 (2011)
3. Chen, X., Wang, H., Ke, L., Lu, Z., Su, H., Chen, X.: Identifying cantonese rumors with discriminative feature integration in online social networks. Expert Syst. Appl. **215**, 119347 (2023)
4. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018)
5. Fionda, V., Pirrò, G.: Fact checking via evidence patterns. In: Proceedings of the 27th International Joint Conference on Artificial Intelligence, pp. 3755–3761 (2018)

6. Hu, L., et al.: Compare to the knowledge: graph neural fake news detection with external knowledge. In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, pp. 754–763 (2021)

7. Ke, L., Chen, X., Lu, Z., Su, H., Wang, H.: A novel approach for cantonese rumor detection based on deep neural network. In: Proceedings of 33rd IEEE International Conference on Systems, Man, and Cybernetics, pp. 1610–1615 (2020)

8. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016)

9. Kwon, S., Cha, M., Jung, K., Chen, W., Wang, Y.: Prominent features of rumor propagation in online social media. In: Proceedings of 13th IEEE International Conference on Data Mining, pp. 1103–1108 (2013)

10. Lin, Z.H., Wang, Z., Zhao, M., Song, Y., Lan, L.: An AI-based system to assist human fact-checkers for labeling cantonese fake news on social media. In: Proceedings of 10th IEEE International Conference on Big Data, pp. 6766–6768 (2022)

11. Lu, Y.J., Li, C.T.: GCAN: graph-aware co-attention networks for explainable fake news detection on social media. In: Proceedings of 58th Annual Meeting of the Association for Computational Linguistics, pp. 505–514 (2020)

12. Ma, J., et al.: Detecting rumors from microblogs with recurrent neural networks. In: Proceedings of the 25th International Joint Conference on Artificial Intelligence, pp. 3818–3824 (2016)

13. Ma, J., Gao, W., Wong, K.F.: Rumor detection on twitter with tree-structured recursive neural networks. In: Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, pp. 1980–1989 (2018)

14. Pan, J.Z., Pavlova, S., Li, C., Li, N., Li, Y., Liu, J.: Content based fake news detection using knowledge graphs. In: The Semantic Web-ISWC : 17th International Semantic Web Conference, pp. 669–683 (2018)

15. Song, Y.Z., Chen, Y.S., Chang, Y.T., Weng, S.Y., Shuai, H.H.: Adversary-aware rumor detection. In: Proceedings of the 59th Findings of the Association for Computational Linguistics: ACL-IJCNLP, pp. 1371–1382 (2021)

16. Sun, M., Zhang, X., Zheng, J., Ma, G.: DDGCN: dual dynamic graph convolutional networks for rumor detection on social media. In: Proceedings of the 36th AAAI Conference on Artificial Intelligence, vol. 36, pp. 4611–4619 (2022)

17. Sun, T., Qian, Z., Dong, S., Li, P., Zhu, Q.: Rumor detection on social media with graph adversarial contrastive learning. In: Proceedings of the 31st ACM Web Conference, pp. 2789–2797 (2022)

18. Vaswani, A., et al.: Attention is all you need. In: Advances in Neural Information Processing Systems, vol. 30 (2017)

19. Wei, L., Hu, D., Zhou, W., Yue, Z., Hu, S.: Towards propagation uncertainty: Edge-enhanced bayesian graph convolutional networks for rumor detection. arXiv preprint arXiv:2107.11934 (2021)

20. Wu, K., Yang, S., Zhu, K.Q.: False rumors detection on sina weibo by propagation structures. In: Proceedings of the 31st IEEE International Conference on Data Engineering, pp. 651–662 (2015)

21. Yu, F., Liu, Q., Wu, S., Wang, L., Tan, T.: A convolutional approach for misinformation identification. In: Proceedings of the 26th International Joint Conference on Artificial Intelligence, pp. 3901–3907 (2017)

22. Zhou, P., et al.: Attention-based bidirectional long short-term memory networks for relation classification. In: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics, pp. 207–212 (2016)