

REVIEW ARTICLE

Section: *Literature, Linguistics & Criticism*

Corpus linguistic-driven hybrid feature representation learning with graph convolutional network for text-based emotion detection

Mohammad Mahzari^{1*}, Noah N. Alsager¹ & Anwer Mustafa Hilal²

¹Department of English Language and Literature, College of Science & Humanities, Prince Sattam bin Abdulaziz University, Saudi Arabia

²Department of Computer and Self Development, Preparatory Year Deanship, Prince Sattam bin Abdulaziz University, Saudi Arabia

*Correspondence: m.mahzari@psau.edu.sa

ABSTRACT

Text-based emotion detection becomes a significant research area in natural language processing (NLP), supporting various domains such as diagnosing mental health status, customer feedback analysis, and social media monitoring. However, accurate detection of emotion classes is a challenging task as human emotional states are contextual, uncertain, and subtle. Existing emotion detection models typically struggle to fully understand these underlying emotional characteristics. To address these issues, this research introduces a Hybrid Attention Learning with Graph Convolutional Network for Text-Based Emotion Detection (HALGCN-TED) framework, aiming to develop a robust and automated model to improve the representation and detection of complex emotional patterns. The model uses a hybrid feature embedding technique by combining RoBERTa and FastText, where RoBERTa captures contextual semantic connections, while FastText learns sub-word and word-level patterns effectively. These representations are further fused through an attention-based feature fusion strategy to provide highly informative features and reduce the impact of less informative features. The model applies a graph convolutional network to learn complex relationships among these fused representations and facilitate more accurate separation of multiple emotional classes. Furthermore, the model employs AdamW optimization to improve the model's learning stability and support more effective convergence. An extensive experimental analysis is performed using two benchmark datasets across various performance metrics to verify the performance of the HALGCN-TED framework in emotion detection from text datasets. These results show that the proposed model achieves excellent outcomes over compared models and validates its potential in improved learning representations and accurate classification of multiple emotional classes.

KEYWORDS: emotion detection, graph convolutional network, hybrid feature embedding, deep learning, attention-based feature fusion

Research Journal in Advanced Humanities

Volume 7, Issue 4, 2026

ISSN: 2708-5945 (Print)

ISSN: 2708-5953 (Online)

ARTICLE HISTORY

Submitted: 29 July 2026

Accepted: 14 September 2026

Published: 18 October 2026

HOW TO CITE

Mahzari, M., Alsager, N. N., & Hilal, A. M. (2026). Corpus linguistic-driven hybrid feature representation learning with graph convolutional network for text-based emotion detection. *Research Journal in Advanced Humanities*, 7(4). <https://doi.org/10.58256/en4n0886>



Published in Nairobi, Kenya by Royallite Global, an imprint of Royallite Publishers Limited

© 2026 The Author(s). This is an open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

I. Introduction

Emotion is a significant multidisciplinary domain that is analysed in various fields, such as computer science, psychology, philosophy, neuroscience, and sociology. Emotions are considered psychological conditions that are interconnected with sentiments, subjective experiences, thoughts, and behavioural reactions [1]. Emotions can be recognized through text data, voice and speech recordings, video records, facial expressions, and behavioural patterns. Assessing emotions through textual data or documents remains challenging because textual expressions do not always contain direct emotion-based words, which makes it difficult to understand and interpret the concepts and emotions expressed in the text [2]. Emotional expressions are classified as positive and negative or neutral emotions. Positive emotions are considered warm tones, smiles, happiness, joy, triumph, laughter, and affection, while negative emotions are expressed as fear, sadness, anxiety, depression, and so on [3]. Furthermore, textual information or data are widely acquired from social media applications, including Instagram, YouTube, Twitter, Facebook, and other review platforms, in which people are posting their emotions and thoughts [4]. Analysing the text information available on the websites can be useful to recognize the severity of an individual's emotions. However, a lack of interpretation and understanding of each person's emotional condition can make it challenging to analyse their emotion accurately [5]. Text-Based Emotion Detection has been widely used by researchers for automated detection and identification of sentiments and thoughts expressed in textual format. Existing approaches mostly depend on manual inspection, which is time-consuming, problem-specific, and highly subjective [6]. Therefore, automated mechanisms and Artificial Intelligence (AI)-based models are required to capture the intricate patterns from small labeled datasets while focusing on multiple tasks and events.

Over the past few years, investigators have developed multiple techniques to identify the emotions of the text, including lexical affinity, learning-based, keyword-based, and hybrid approaches. Furthermore, AI-based techniques integrate Natural Language Processing (NLP) for text-based emotion detection that incorporates methods in linguistics and computation to support systems in understanding and delivering text and audio in the form of human-understandable languages [7]. In the context of information technology, NLP-based techniques have been widely adopted, mainly focusing on generating human-understandable languages, thereby enabling AI models to interpret and explain the decisions [8]. NLP-based applications, including virtual assistants, sentiment analysis, text summarization, machine translation, and autocorrect and predictive text, are undertaking social matters at an extraordinary rate. In recent times, Deep learning (DL) methods have gained significant attention and provided promising solutions to overcome these challenges, thereby capturing intricate patterns and feature representations from large-scale datasets [9]. DL-based approaches have been widely used for various NLP and text-assessment tasks, thereby offering remarkable solutions for text-based emotion detection [10]. The expanding volume of user-generated textual content and the involvement of textual content in various fields such as customer interaction, medical analysis, social media monitoring increased the demand has shaped a strong demand for more robust and automated emotion detection model. On the other hand, linguistically diverse, context-dependent content creates challenges in achieving more accurate detection outcomes. Although various AI-based models have been developed, they still face challenges in capturing deep semantic and contextual dependencies. These limitations motivate the development of the HALGCN-TED model by incorporating a standard preprocessing pipeline, a hybrid feature embedding strategy with an attention-based feature fusion module, a graph-based classifier, and AdamW optimization to offer more robust and reliable categorization of emotion classes. The key contribution of this research is presented below:

- Applies a hybrid feature embedding technique by combining the Robustly Optimized BERT approach (RoBERTa) to learn contextual and semantic relations from the text and FastText to provide sub-word local representation learning
- Develops an attention-based feature fusion module to fuse RoBERTa and FastText by assigning higher weights to the discriminative features and producing a more compact feature set for the classification model
- Employs a Graph Convolutional Network (GCN) to learn the complex relationship among the textual representations and enable the model to effectively categorize multiple emotional classes.
- Integrates AdamW optimizer to update the network parameters and support more stable model learning and effective model convergence.
- Performs experimental evaluation to measure the robustness of the proposed model using two benchmark

datasets across various performance metrics, while comparative experimental and ablation studies are also conducted to investigate the contribution of individual component to validate the model's stability and generalization capability in accurate emotion detection and classification.

II. Literature Review

This section presents a brief note about the previous research related to text-driven emotion detection. Pittala et al. [11] developed an innovative approach that integrates ML-based approaches with NLP in detecting emotion through textual documents and information. The authors mainly designed this framework due to the increase in digital platforms and the necessity for automated mechanisms to detect users' emotional states through textual data. This framework has the capacity to enhance machine-enabled models in understanding users' emotions through textual interactions. Bekmurat et al. [12] designed a novel approach by integrating NLP-based approaches with behavioural patterns to enhance the understanding of humans' psychological and emotional conditions. Additionally, the authors used text-based classification for better interpretation of users' emotional conditions. Furthermore, the framework integrates SHAP-based models to provide better interpretation of models' decision-making process. Ahmadipoor et al. [13] explored the role of NLP-based techniques in detecting emotions through text documents or information. The authors demonstrated that NLP-based approaches play a significant role in applications, including sentiment assessment, text summarization, machine translation, and virtual assistance. In addition, the authors developed deep operator networks to capture the intricate patterns in NLP tasks and provide accurate text analysis. Kabir et al. [14] introduced a novel framework by integrating NLP-based methods with ML-based techniques to identify and classify emotions through text documents. The authors analysed the framework by using a text preprocessing stage to attain precise classification and transparent data for emotion assessment. This framework has the capacity to analyse and capture intricate patterns from large-scale textual data, thereby providing precise outcomes for multiple applications.

Al-Hagery et al. [15] designed an innovative approach that integrates DL-based methods with NLP-based models for detecting emotions through textual documents. The authors demonstrated that these automated models can capture complex patterns from large-scale datasets, thereby providing sentiment assessment. This framework has the capacity to provide accurate, reliable, and robust solutions for text-based emotion detection when compared to traditional approaches. Bhaskaru et al. [16] developed an innovative approach by implementing NLP-based methods for reliable and accurate emotion identification through text-based documents. Furthermore, the authors integrated LSTM to capture the complex dependencies and features within the textual and emoji-based interaction. Rasool et al. [17] introduced NLP-based techniques with Large Language Models (LLMs) for emotion identification for psychological health care. The authors mainly focused on providing an automated mechanism for recognizing emotional features and monitoring patients' mental health, thereby providing better results. This framework utilized textual data or information to improve the mental health analysis. Kumar et al. [18] analysed the role of BERT and NLP for detecting emotion through text-based data. Furthermore, the authors demonstrate the exceptional generalization capacity of the framework under various emotional expressions, thereby emphasizing the model's efficiency for emotion-based NLP applications. The NL-based techniques demonstrated promising solutions for real-time use cases.

Existing text-based emotion detection models provide better solutions; however, various limitations continue to impact their accuracy and reliability. The existing models mostly use a single feature representation, which may not effectively learn all relevant information from the text data and leads to computational complexity. Most of the previous research uses standard CNN or transformer-based classification models, rarely incorporating graph-based techniques, which creates a gap in effectively capturing relational dependencies present in emotional text content. In addition, existing models employ standard optimization strategies that can lead to slow convergence, overfitting, and lower model classification accuracy. Furthermore, the existing model lacks an effective preprocessing pipeline, since real-world text data may contain special characteristics, emojis, abbreviations, and other additional linguistic noise that requires more effective preprocessing steps to provide clean and standardized input prior to model development. To address these gaps, this study develops the HALGCN-TED framework by leveraging strong preprocessing steps, hybrid feature embedding with an attention-based feature fusion strategy, an advanced graph-based classifier, along with AdamW optimization, to offer improved learning representation and accurate classification of emotional classes using textual data.

III. Proposed HALGCN-TED framework

In this work, a HALGCN-TED framework is proposed for emotion detection using textual data. The proposed model initially uses multiple steps in the preprocessing stage to provide more standardized input for further analysis. The model aims to improve the learned representation by integrating RoBERTa- and FastText-based feature embeddings; these are then fused through attention-based feature fusion to provide a more discriminative feature set. Then, the fused representation is processed using a GCN classifier to learn complex relationships among the features and perform emotion classification, whereas AdamW is further used for effective model optimization. Each stage of the proposed model is illustrated in Fig.1 and explained in the following subsections.

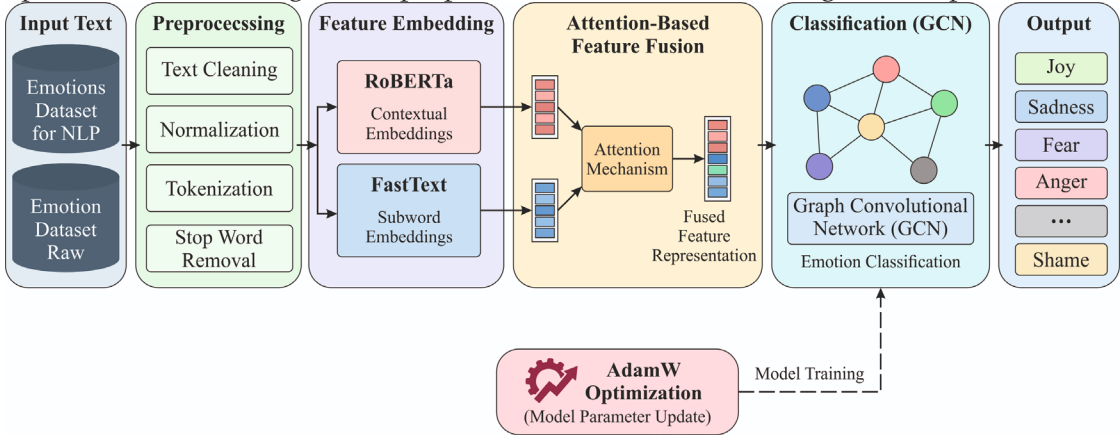


Fig. 1. Overall methodology of the proposed HALGCN-TED model

A. Dataset Descriptions

The empirical validation of the HALGCN-TED approaches can be examined using the Emotions dataset for NLP (Dataset 1) [19] and Emotion Detection from Text (Dataset 2) [19]. A brief explanation of the two databases is given below:

Dataset 1: The dataset is developed for NLP-based emotion classification, in which a model learns to predict the emotion expressed in an input text. The data is classified into training, validation, and testing phases and can be applied with ML- and DL-based approaches. The dataset consists of 20000 text samples with six emotional categories, as illustrated in Table 1.

TABLE I. DETAILS OF THE DATASET 1

Class Labels	Class Count
‘Joy’	6761
‘Sadness’	5797
‘Anger’	2709
‘Fear’	2373
‘Love’	1641
‘Surprise’	719
Total	20000

TABLE II. DETAILS OF THE DATASET 2

Class Labels	Class Count
‘Neutral’	8638
‘Worry’	8459
‘Happiness’	5209
‘Sadness’	5165
‘Love’	3842
‘Surprise’	2187
‘Fun’	1776

'Relief'	1526
'Hate'	1323
'Enthusiasm'	759
'Boredom'	179
'Anger'	110
Total	39173

Dataset 2: Emotion detection through text is one of the significant challenges in NLP-based models. The main reason is the insufficiency of labeled datasets and the multi-class nature of the problem. Humans have multiple emotions, and it is difficult to gather sufficient records for all emotions; thus, the complexity of class imbalance increases. Here we have labeled data for emotion detection, and the target is to develop a model for effective emotion detection. The dataset consists of 39173 text samples with various emotional categories, as represented in Table 2.

B. Text Pre-processing

Preprocessing acts as a significant part to ensure the integrity and quality of textual input prior to the model development [21]. The study handled the subsequent preprocessing phases, including text cleaning, normalization, tokenization, and stop word removal. The model starts with a text cleaning technique for the elimination of noisy data from the texts, such as emojis, special characters, URLs, and unnecessary whitespace that can be removed to standardize the input. Afterwards, the normalization phase is carried out by transforming text to lowercase, eliminating redundant spaces, and standardizing textual differences. It decreases redundant variances among the same words and increases representation reliability. Moreover, Tokenization includes splitting a sentence into its tokens or basic words. This method offers the simple textual elements needed for the following feature embedding method. Lastly, Common words like “the,” “is,” “and,” and “of” that provide inadequate discriminative information are eliminated. This support in reducing redundant information and preserving the most informative words that contribute more emotion-based patterns. These preprocessing phases convert the raw textual inputs into clean, normalized, and discriminative token sequences, which are then offered to the embedding process for efficient emotion classification.

C. Hybrid Feature Embedding

To learn the contextual meaning and fine-grained lexical patterns existing in emotion-based text, complementary word representations are more essential to offer richer linguistic information for the following classification process. The proposed framework uses a hybrid feature embedding method that combines RoBERTa and FastText, where RoBERTa learns contextual semantic connections, whereas FastText learns sub-word- and word-level patterns. RoBERTa is an enhanced form of BERT, it is utilized to learn the contextual meaning of words according to their surrounding text, supporting the identification of emotion-based semantic patterns. The mathematical form of RoBERTa representation can be determined by:

In the above-mentioned expression, the term $\mathbf{h}$ implies the contextual representation produced from the input text . Even though RoBERTa [22] efficiently learns contextual and semantic details, additional lexical and subword details are needed to improve the representation of emotion-based text. Hence, FastText is combined to learn word-level features and character n-gram patterns, thus offering complementary information for emotion recognition.

The FastText [23] embedding is an extensively used word representation model by Facebook AI. It has the ability to capture word representations efficiently. It can individually capture the structural form of every word from a text. This FastText embedding model follows the skip-gram model, which is used to learn sub-word and structural details. In this work, FastText is applied to learn lexical and character-level patterns, comprising word differences that can provide emotional expression. The FastText representation can be determined by:

Here, the term n refers to the number of characters -grams, parameter w refers to the word vector, C indicates the set of character -grams with respect to word w , and v implies the vector corresponding to character -gram c .

Accordingly, RoBERTa learns contextual semantics, whereas FastText captures subword and lexical-related characteristics, providing complementary representations for precise text-based emotion detection. Later, the hybrid feature representations are offered to the Attention-Based Feature Fusion model that integrates their complementary contextual and lexical details prior to GCN classification. Fig.2. represents the architecture of hybrid feature learning

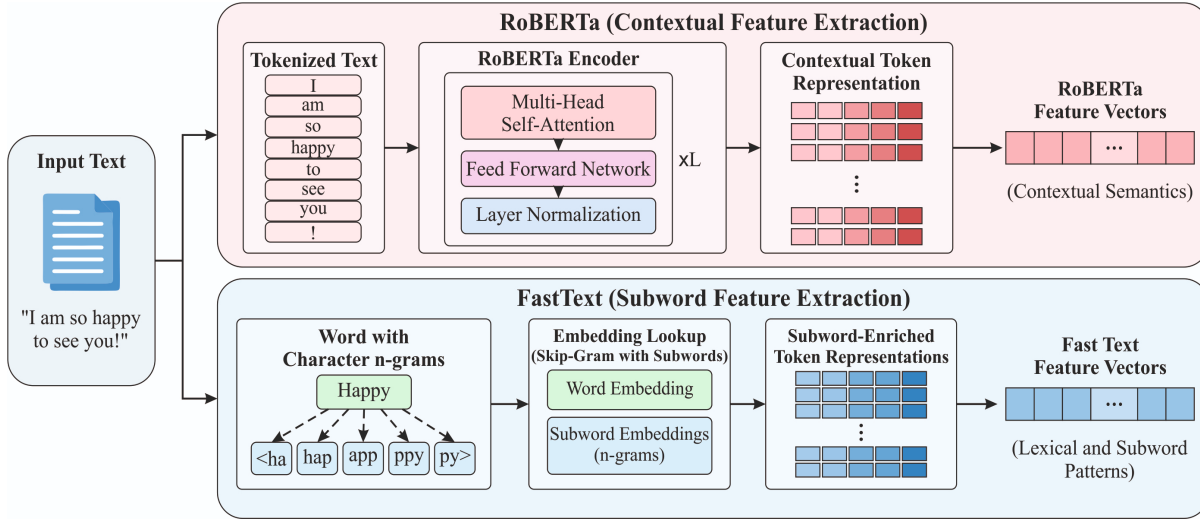


Fig. 2. Architecture of the hybrid feature learning

D. Attention-Based Feature Fusion

To combine the complementary information acquired from RoBERTa and FastText, an attention-based feature fusion method has been used. The attention mechanism allocates importance to these representations based on their contribution to emotion-based attributes. The mathematical expression of attention weights is measured by:

Here, the term u_i represents the importance score corresponding to the i -th feature representation, and the symbol α_i denotes the normalized attention weight. The weighted FastText and RoBERTa representations can be combined to form the new mathematical expression.

Now, the parameters u and α imply the FastText and RoBERTa representations, respectively, and the notations u and α denote their respective attention weights. Additionally, the resultant representation highlights informative lexical, contextual, and sub-word-related attributes while decreasing the impact of low-significance details. Afterwards, the fused representation is utilized as the input for graph construction and GCN-based emotion classification. Fig.3. Shows the working process of attention-based feature fusion mechanism

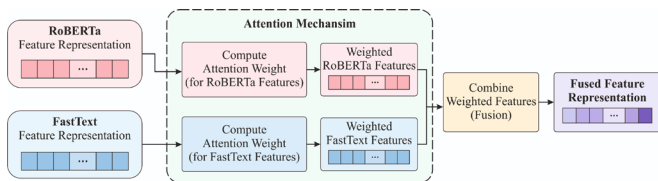


Fig. 3. Working process of attention-based feature fusion mechanism

E. GCN-based Classification Module

The fused representations learn contextual and semantic information from the input text, whereas GCN is used to model the connections among these fused textual representations and enrich the discrimination of diverse emotion categories.

Initially, the attention-based fused textual representations are organized into graph, in which each text representation is considered as a node and the edges denote connections among nodes. Additionally, the -nearest neighbor (K-NN) method is a technique used to identify and link nearest neighbors for each data instance, and finally construct a graph based on the similarity among data points. In the proposed model, every attention-fused text representation is perceived as a node, whereas the similarity between the fused representations is utilized to form the edges of the graph. It allows GCN to capture semantic interactions amongst textual representations. GCN effectively captures node representations by aggregating information from neighboring nodes whereas preserving the graph structure. The mathematical formulation of a GCN layer is expressed as:

In this expression, the parameter I implies the diagonal matrix of 1 , and the parameter $H^{(l)}$ refers to the node feature matrix of layer l . The notation σ implies an activation function (for example, rectified linear unit ReLU)). The mathematical representation A is the adjacency matrix of the graph that comprises connections to the node itself. The parameter $W^{(l)}$ denotes the trainable weight matrix of layer l . Furthermore, the GCN updates the node features in every individual layer through incorporation of them with the features of its neighboring nodes. In this study, the attention-fused RoBERTa and FastText features have been utilized as the initial node features. Later, the GCN captures semantic relationships amongst neighboring nodes for emotion classification.

As GCN involves graph data, K-NN is used to form the graph by interconnecting all the text representations to their K-NN by applying cosine similarity. Every representation is handled as a node, and the connections form the edges. The learned GCN representations are passed to the Softmax layer to acquire probabilities for the emotion classes:

In this expression, the term C implies the total count of emotion classes, and s_i specifies the output score for emotion class i . Additionally, the GCN parameters have been optimized by applying AdamW to improve the learning and generalization of the proposed framework. Accordingly, the GCN captures the semantic connections amongst the fused text representations, whereas softmax generates the final emotion-class probabilities for text-based emotion detection. Fig.4 illustrates the architecture of GCN-based emotion classification

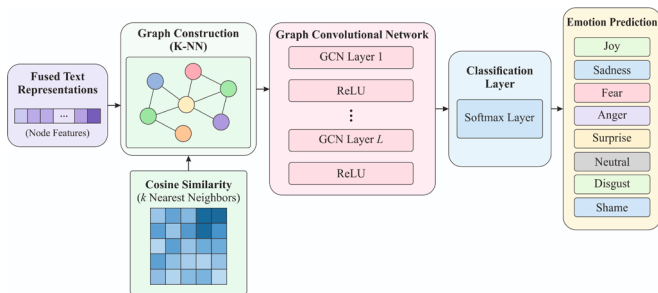


Fig. 4. Architecture of the GCN-based emotion classification

F. Optimization Function

During emotion classification, the AdamW optimizer is effectively selected to update the trainable parameters of the proposed framework; it is essential because fused emotion representations produced from RoBERTa and FastText can generate intricate and changing gradient patterns in training. Its adaptive gradient estimation allows constant parameter updates, although decoupled weight decay helps to minimize overfitting in model training. Model parameters were updated using the AdamW optimizer. Unlike standard Adam, AdamW decouples weight decay mechanisms from the gradient update process, helping to prevent adaptive learning behavior from being fundamental.

Here, θ_t indicates the updated parameters, and θ_0 depicts the model parameters at iteration 0. The term η refers to the learning rate, ϵ represents a small constant added for numerical stability, and $\hat{m}_t$ and $\hat{v}_t$ denote the bias-corrected first and second moment estimates of the gradients. In the final term, β_1 is employed by AdamW. This process adjusts the model parameters separately from the gradient-driven update, in order to enhance training

stability and better generalization. Thus, AdamW offers a stable and effective optimization, supporting the model to learn discriminative pattern and enhance its generalization performance.

IV. Experimental Results and Discussion

This section outlines the experimental result of the proposed model using two bechmark dataset along with detailed discussion

A. Classification Performance Analysis on Dataset 1

Figs. 5a and 5b depict the confusion matrices of the HALGCN-TED method with a 70:30 training/testing split for Dataset 1. Both matrices show that higher diagonal values indicate correct classification for each class, while off-diagonal values represent misclassified instances, demonstrating its effective and robust classification of the proposed text-based emotion detection.

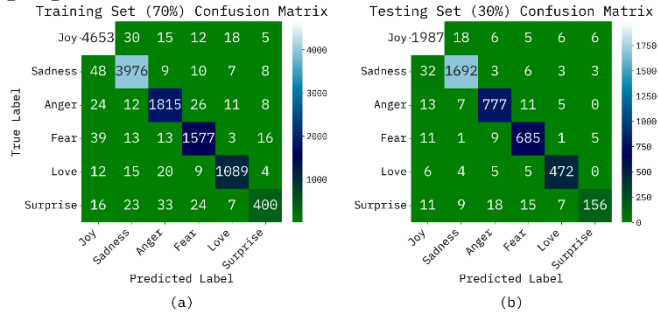


Fig. 5. Confusion matrices of the HALGCN-TED approach with (a-b) 70:30 training/testing split for Dataset 1 Table 3 portrays the textual emotion classification outcomes of the HALGCN-TED approach with a 70:30 training/testing split for Dataset 1. The proposed HALGCN-TED system achieved an overall accuracy of 96.50% and 96.15% for both training and testing phases. This performance indicates the reliability and capability of the proposed system for text-based emotion detection.

Figs. 6a and 6b indicate the ROC and PR curves for Dataset 1. The ROC curves show superior classification performance, with AUC values ranging from 0.987 to 0.997. Furthermore, the PR curves verify robust performance, with AP values ranging from 0.875 to 0.993. The XXX technique attains extremely consistent emotion classification, even though Surprise shows relatively lower performance when compared to other emotion categories.

TABLE III. TEXTUAL EMOTION CLASSIFICATION RESULTS OF THE HALGCN-TED METHOD WITH A 70:30 SPLIT FOR DATASET I

Classes	Accuracy	Precision	Recall	F1-Score	MCC
Training Set (70%)					
Joy	98.44	97.10	98.31	97.70	96.52
Sadness	98.75	97.71	97.98	97.85	96.97
Anger	98.78	95.28	95.73	95.50	94.79
Fear	98.82	95.11	94.94	95.03	94.36
Love	99.24	95.95	94.78	95.36	94.95
Sur- prise	98.97	90.70	79.52	84.75	84.41
Average	98.83 $\pm$	95.31 $\pm$	93.54	94.36 $\pm$	93.67
$\pm$ SD	0.24	2.26	$\pm$ 6.42	4.44	$\pm$ 4.24
Overall Accuracy: 96.50					
Testing Set (30%)					
Joy	98.10	96.46	97.98	97.21	95.78
Sadness	98.57	97.75	97.30	97.52	96.51
Anger	98.72	94.99	95.57	95.28	94.54

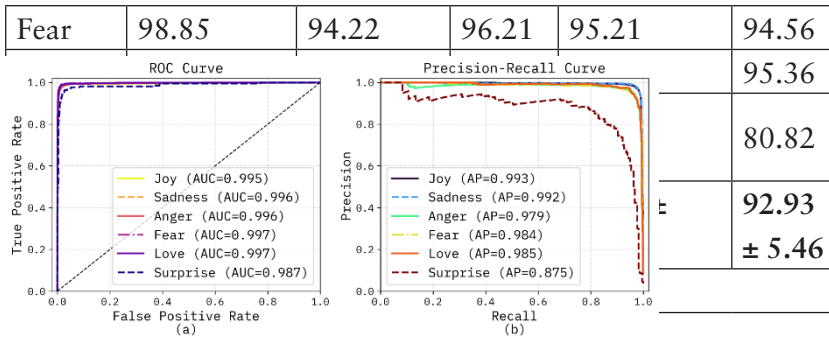


Fig. 6. Classification result of the HALGCN-TED approach with (a-b) ROC/PR curves for Dataset 1

B. Classification Performance Analysis on Dataset 2

Figs. 7a and 7b describe the confusion matrices of the HALGCN-TED approach with a 70:30 training/testing split for Dataset 2. These matrices exhibit maximal diagonal values, indicating that all categories are properly detected, whereas non-diagonal values imply misclassified instances, signifying efficient and strong classification of the presented text-based emotion detection.

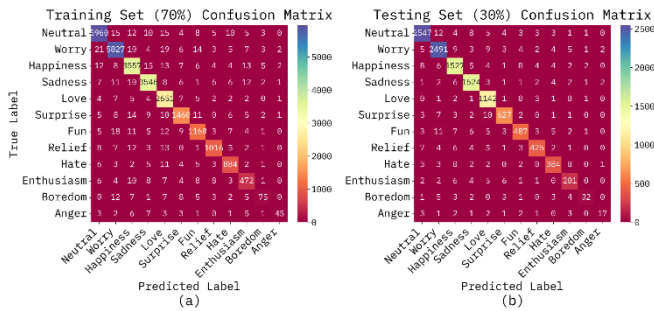


Fig. 7. Confusion matrices of the HALGCN-TED approach with (a-b) 70:30 training/testing split for Dataset 2

Table 4 reveals the textual emotion classification results of the HALGCN-TED approach with a 70:30 training/testing split for Dataset 2. The proposed HALGCN-TED system reached an overall accuracy of 97.23% and 97.04% for the training and testing phases. This result shows the consistency and ability of the presented framework for text-based emotion detection.

TABLE IV. TEXTUAL EMOTION CLASSIFICATION RESULTS OF THE HALGCN-TED METHOD WITH A 70:30 SPLIT FOR DATASET 2

Classes	Accuracy	Precision	Recall	F1-Score	MCC
Training Set (70%)					
Neutral	99.40	98.72	98.56	98.64	98.26
Worry	99.31	98.40	98.41	98.40	97.96
Happiness	99.31	97.29	97.56	97.43	97.03
Sadness	99.49	98.04	98.06	98.05	97.75
Love	99.43	95.74	98.55	97.12	96.82
Surprise	99.53	96.18	95.36	95.77	95.52
Fun	99.49	94.73	93.97	94.35	94.08
Relief	99.68	96.49	95.13	95.80	95.64
Hate	99.66	94.55	95.46	95.00	94.83
Enthusiasm	99.56	88.39	88.89	88.64	88.41
Boredom	99.74	78.95	60.00	68.18	68.70
Anger	99.86	86.54	58.44	69.77	71.05
Average $\pm$ SD	99.54 $\pm$ 0.16	93.67 $\pm$ 5.74	89.87 $\pm$ 13.96	91.43 $\pm$ 10.37	91.34 $\pm$ 9.93

Overall Accuracy: 97.23					
Testing Set (30%)					
Neutral	99.32	98.61	98.30	98.45	98.02
Worry	99.14	97.88	98.15	98.01	97.46
Happiness	99.29	96.95	97.70	97.32	96.91
Sadness	99.46	97.57	98.39	97.97	97.67
Love	99.49	95.81	99.13	97.44	97.17
Surprise	99.55	96.31	95.58	95.94	95.71
Fun	99.45	96.25	91.37	93.74	93.49
Relief	99.56	95.72	92.79	94.24	94.02
Hate	99.66	93.43	96.73	95.05	94.89
Enthusiasm	99.59	90.54	88.16	89.33	89.13
Boredom	99.75	82.05	59.26	68.82	69.62
Anger	99.82	77.27	51.52	61.82	63.01
Average $\pm$ SD	99.51 $\pm$ 0.19	93.20 $\pm$ 6.46	88.92 $\pm$ 15.41	90.68 $\pm$ 11.69	90.59 $\pm$ 11.19
Overall Accuracy: 97.04					

Figs. 8a and 8b show the ROC and PR curves for Dataset 2. The ROC curves show remarkable classification performance, with AUC values ranging from 0.995 to 0.999. Furthermore, the PR curves validate robust performance, with AP values ranging from 0.652 to 0.994. The XXX method attains highly consistent emotion classification, even though Surprise shows relatively lower performance when compared to other emotion categories.

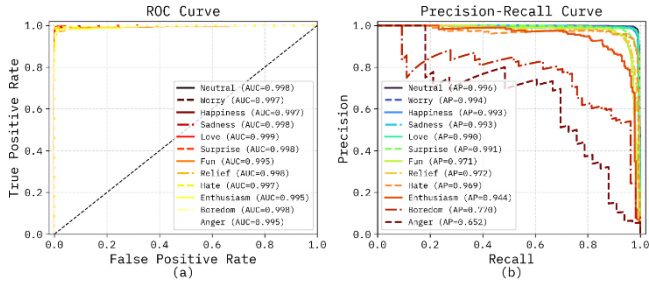


Fig. 8. Classification result of the HALGCN-TED approach with (a-b) ROC/PR curves for Dataset 2

C. Comparative Evaluation and Ablation Study

Table 5 and Fig. 9 reveal the comparison outcomes of the HALGCN-TED model with existing methodologies using several metrics for both datasets [26, 27]. In Dataset 1, the result implied that the proposed HALGCN-TED method achieves maximal performance with accuracy of 96.15%, precision of 95.12%, and recall of 92.54% compared with conventional approaches. Moreover, for Dataset 2, the outcome indicated that the TF-IDF + LR and CNN-LSTM methods attain lower performance, while the CNN + Word2Vec, DistilBERT, MJV, Bi-LSTM, and BERT-base algorithms slightly improve performance. However, the proposed HALGCN-TED system obtains higher performance with an accuracy of 97.04%, a precision of 93.20%, and a recall of 88.92%, thereby confirming its stable and effective performance in the presented text-based emotion classification.

TABLE V.

TABLE VI. COMPARATIVE RESULTS OF THE HALGCN-TED APPROACH WITH OTHER SYSTEMS

Methods	Dataset 1			Dataset 2		
	Accuracy	Precision	Recall	Accuracy	Precision	Recall
TF-IDF + LR [26]	76.80	76.70	76.80	74.92	74.35	74.18
CNN + Word2Vec [26]	81.20	81.10	81.00	79.46	78.92	78.71
DistilBERT [26]	88.40	88.30	88.40	86.73	86.21	86.45
BERT-base [26]	91.10	90.00	90.10	89.32	88.47	87.76
MJV [27]	84.28	84.10	83.20	82.61	81.94	81.53
CNN-LSTM [27]	77.59	75.60	74.70	75.28	73.84	72.91
Bi-LSTM [27]	86.00	83.30	84.40	84.51	82.17	82.93
HALGCN-TED	96.15	95.12	92.54	97.04	93.20	88.92

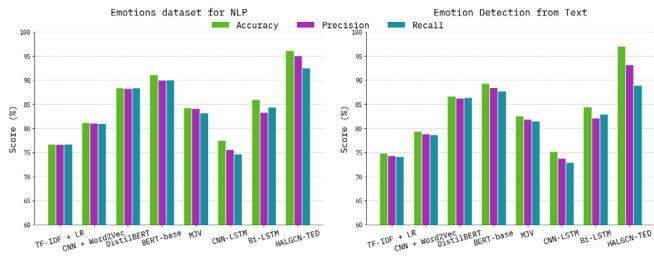


Fig. 9. Comparative results of the HALGCN-TED method with metrics using Dataset 1 and Dataset 2

Table 6 and Fig. 10 exhibit the ablation analysis of the HALGCN-TED method with various metrics using Dataset 1. The result implies that the baseline GCN model gains minimal performance with the combination of RoBERTa, while the successive integration of FastText slightly improves model performance. Adding the AFF technique further improves feature representation and fusion. The complete HALGCN-TED (RoBERTa + FastText + AFF + GCN+ AdamW) model achieves better performance with an accuracy of 97.04%, precision of 93.20%, recall of 88.92%, and an F1-score of 92.93, signifying the performance of each component and their integrated contribution to the presented system.

TABLE VII. ABLATION STUDY OF THE HALGCN-TED APPROACH WITH DATASET 1

Methods	Accuracy	Precision	Recall	F1-Score
GCN	89.62	86.74	82.36	84.49
RoBERTa + GCN	92.48	89.83	86.27	88.01
RoBERTa + FastText + GCN	94.16	92.14	89.48	90.79
RoBERTa + FastText + AFF + GCN	95.28	94.21	91.36	92.76
HALGCN-TED (RoBERTa + FastText + AFF + GCN+ AdamW)	96.15	95.12	92.54	92.93

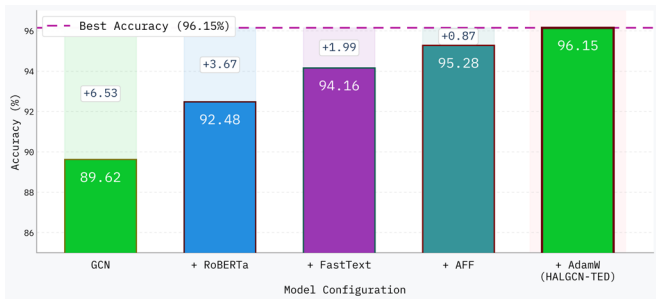


Fig. 10. Ablation study of the HALGCN-TED method with Dataset 1 using accuracy metrics

Table 7 and Fig. 11 show the ablation results of the HALGCN-TED approach with several metrics using Dataset 2. The outcome indicated that the baseline GCN model obtains lower performance, while the inclusion of RoBERTa improves the baseline model performance. The addition of FastText contributes to further enhancement in textual feature embedding. Integrating the AFF model enhances feature integration, resulting in more efficient classification results. The complete HALGCN-TED (RoBERTa + FastText + AFF +

GCN+ AdamW) model achieves better performance with an accuracy of 97.04%, precision of 93.20%, recall of 88.92%, and an F1-score of 90.59, emphasizing the performance of the incorporated contributions of all components in the proposed text-based emotion detection process.

TABLE VIII. ABLATION STUDY OF THE HALGCN-TED APPROACH WITH DATASET 2

Methods	Accuracy	Precision	Recall	F1-Score
GCN	90.18	85.72	80.46	83.01
RoBERTa + GCN	92.64	88.36	83.52	85.87
RoBERTa + FastText + GCN	94.52	90.74	85.76	88.18
RoBERTa + FastText + AFF + GCN	95.86	92.18	87.34	89.69
HALGCN-TED (RoBERTa + FastText + AFF + GCN+ AdamW)	97.04	93.20	88.92	90.59

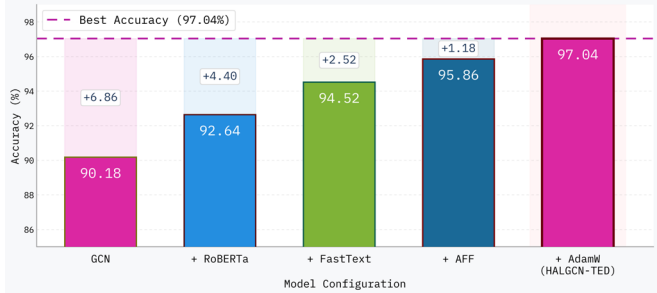


Fig. 11. Ablation study of the HALGCN-TED method with Dataset 2 using accuracy metrics

D. Discussion

The experimental outcome of the proposed HALGCN-TED model highlights its effectiveness in identifying multiple emotion classes from textual data by incorporating hybrid feature representation and graph-based learning. The preprocessing steps improved the data quality and provided more standardized input for model development. The hybrid RoBERTa and FastText feature learning effectively captured both contextual dependencies and sub-word-level semantic features, while the attention-based feature fusion stage effectively combined the learned feature representations and provided the most informative features for model learning. GCN-based classification learns the complex relationships among the interconnected textual features and performs reliable classification outcomes. Furthermore, AdamW supports effective parameter learning and provides better model generalization during training. Therefore, the proposed HALGCN-TED model enables effective categorization of emotion classes and validates its suitability for more accurate real-world text-based emotion detection.

V. Conclusion and Future Direction

In this research, a HALGCN-TED framework is proposed for accurate emotion detection and classification using textual data. The developed model aims to improve the learning representation by effectively capturing emotion-related features and enhancing the model accuracy in multi-class classification. Primarily, text preprocessing steps are applied to prepare the standardized data for model development by removing spelling errors, emojis, URLs, and other less informative details. For better feature learning, the model uses a hybrid strategy by combining RoBERTa and FastText to capture discriminative and sub-word-level semantic details, while an attention-driven fusion module further highlights the most relevant emotional features. Later, a GCN is used as a classifier to model relationships among the fused representations and facilitate accurate emotion classification, while AdamW is incorporated to attain stable model learning and faster convergence. Comprehensive experiments are performed on two benchmark datasets to assess the robustness of the presented HALGCN-TED model; these outcomes highlight the capability of the proposed model in delivering accurate and reliable text-based emotion classification over existing techniques. Although the proposed model achieves superior classification outcomes, several opportunities remain open for further improvement. Future work can explore a meta-optimization strategy to automatically tune model parameters and support more effective model convergence and classification accuracy. Future work can be assessed using a multimodal dataset by integrating text, visual, and video information to provide richer emotional cues and enhance the detection and classification performance.

Data Source and Availability : The datasets used to support the findings of this research are publicly available in the Kaggle repository at <https://doi.org/10.34740/kaggle/dsv/20875185> reference numbers [19].

Acknowledgment: The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2025/ 01/34640)

REFERENCES

- Khan, B., Usman, M. and Binsawad, M., 2026. A novel hybrid model for emotion detection in text through sequential and transformer-based approaches: LSTM-enhanced RoBERTa (LER). *Scientific Reports*, 16(1), p.2224.
- Al-Milli, N., Elsoud, E.A., Al-Shaikh, A.A., Omar, H. and AlNajjar, A., 2026. Emotion detection and sentiment analysis of textual data. *Egyptian Informatics Journal*, 34, p.100976.
- Kusal, S., Patil, S., Peerbhai, A., Kotecha, K., Selvachandran, G. and Abraham, A., 2026. Meta-learning ensemble for emotion detection in conversational text. *Neural Computing and Applications*, 38(4), p.54.
- Yamani, C., Thrylokya, K., Sirisha, B., Bhagyalatha, S., Eluri, R.K. and Moturi, S., 2025, March. Text-based emotion analysis: Approaches and evaluations. In *2025 IEEE International Conference on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI)* (Vol. 3, pp. 1-6). IEEE.
- Tamilkodi, R., Sujatha, B. and Leelavathy, N., 2025. Emotion detection in text: advances in sentiment analysis. *International Journal of System Assurance Engineering and Management*, 16(2), pp.552-560.
- Kusal, S., Patil, S. and Kotecha, K., 2025. Multimodal text-emoji fusion using deep neural networks for text-based emotion detection in online communication. *Journal of Big Data*, 12(1), p.32.
- Marchitan, T.G., Creangă, C. and Dinu, L.P., 2025, July. Team Unibuc-NLP at SemEval-2025 Task 11: Few-shot text-based emotion detection. In *Proceedings of the 19th International Workshop on Semantic Evaluation (SemEval-2025)* (pp. 468-475).
- Singh, B., Kaur, K. and Kaur, G., 2025, April. Optimizing emotion detection: An NLP-driven deep learning approach to sentiment encoding. In *2025 International Conference on Data Science and Business Systems (ICDSBS)* (pp. 1-6). IEEE.
- Alluhaidan, A.S., Alotaibi, S., Almubarak, H. and Al-Hagery, M.A., 2026. Mathematical modelling of attention-guided deep learning for real-time text emotion recognition in assistive interaction systems for disabled persons. *Disability and Rehabilitation: Assistive Technology*, pp.1-22.
- Ahamad, R. and Mishra, K.N., 2025. Exploring sentiment analysis in handwritten and E-text documents using advanced machine learning techniques: a novel approach. *Journal of Big Data*, 12(1), p.11.
- Pittala, R.B., Raj, R., Padmaja, D.L., Kiran, M.A., Thaile, M., Addanki, U.K. and Lakshmi, K.J., 2026, June. Emotion Detection in Text Messages Using NLP and Machine Learning (BERT). In *2026 6th International Conference on Intelligent Technologies (CONIT)* (pp. 1-6). IEEE.
- Bekmurat, O., Akpanbetov, D., Tursynkhan, A., Demeubayeva, L., Duisenbekkyzy, Z., Sansyzbay, K., Kadirkulov, S. and Bakhtiyarova, Y., 2026. Explainable and Computationally Efficient NLP Framework for Detecting Psycho-Emotional Risk Signals in Social Media. *Computers*, 15(5), p.327.
- Ahmadipoor, B., Mirzaei, A., Nouri-Moghaddam, B. and Mikaeilvand, N., 2026. DeepOpTED: An intelligent deep operator network for text emotion recognition. *Journal of Intelligent & Fuzzy Systems*, p.18758967261472291.
- Kabir, F., Khan, M.H. and Rabby, F., 2025, February. Emotion Detection From Textual Data Using Natural Language Processing and Machine Learning Techniques. In *2025 International Conference on Electrical, Computer and Communication Engineering (ECCE)* (pp. 1-6). IEEE.
- Al-Hagery, M.A., Alharbi, A.A., Alkharashi, A. and Yaseen, I., 2025. A novel hybrid attention based deep learning framework for textual emotion recognition using natural language processing technologies for disabled persons. *Scientific Reports*, 15(1), p.37611.
- Bhaskaru, O., Vali, M., Syed, K.V. and Mohammad, W., 2025, June. Leveraging NLP Techniques for Robust Emotion Recognition in Text. In *2025 6th International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 1572-1578). IEEE.
- Rasool, A., Aslam, S., Hussain, N., Imtiaz, S. and Riaz, W., 2025. nbert: Harnessing nlp for emotion recognition in psychotherapy to transform mental health care. *Information*, 16(4), p.301.
- Kumar, R.P., Ramya, P., Teja, G.S. and Krishna, V.R., 2025, March. Berting emotions: Exploring textual emotion recognition using nlp. In *2025 International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 888-895). IEEE.
- <https://doi.org/10.34740/kaggle/dsv/20875185>
- Alqarni, F., Sagheer, A., Alabbad, A. and Hamdoun, H., 2025. Emotion-Aware RoBERTa enhanced with emotion-

- specific attention and TF-IDF gating for fine-grained emotion recognition. *Scientific reports*, 15(1), p.17617.
- Faseeh, M., Jaleel, A., Iqbal, N., Ghani, A., Abdusalomov, A., Mehmood, A. and Cho, Y.I., 2024. Hybrid approach to automated essay scoring: Integrating deep learning embeddings with handcrafted linguistic features for improved accuracy. *Mathematics*, 12(21), p.3416.
- Ghosal, S. and Jain, A., 2023. Depression and suicide risk detection on social media using fasttext embedding and xgboost classifier. *Procedia Computer Science*, 218, pp.1631-1639.
- Lee, D., Choo, H. and Jeong, J., 2024. Gcn-based lstm autoencoder with self-attention for bearing fault diagnosis. *Sensors*, 24(15), p.4855.
- Sankoh, A.P., Raza, A., Parwez, K., Shishah, W., Alharbi, A., Javed, M. and Bilal, M., 2026. Automated Facial Pain Assessment Using Dual-Attention CNN with Clinically Calibrated High-Reliability and Reproducibility Framework. *Biomimetics*, 11(1), p.51.
- Azhar, M., Riaz, N., Azeem, W., Dewi, D.A., Amjad, A. and Arman, M., 2026. Explainable Transformer Models for Human Emotion Recognition: A Multi-Method Explainability Study in the Context of Mental Health. *Information*, 17(5), p.496.
- Bania, R.K., 2025. Novel text emotion classification model using long short-term memory network. *Journal of Data Science and Intelligent Syst*