International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Emoji-Aware Sentiment Analysis Using Gated Fusion of Textual and Emoji Representations

Akshatha Rithesh[1], Hanumanthappa M[2]
[1]Research Scholar, Department of Computer Science and Applications, Bangalore University, Bengaluru, India 560056
[2]Senior Professor, Department of Computer Science and Applications, Bangalore University, Bengaluru, India 560056
Email address: [1]akshatharithesh@bub.ernet.in, [2]hanu6572@bub.ernet.in

Abstract— *Sentiment analysis in social media data is challenging because the posts are often short, informal, and rich in emojis. People use these posts to express emotions beyond words. The recent transformer-based language models have achieved impressive results in text-based sentiment classifications. But they tend to underuse emojis, even though these symbols often carry important emotional cues in the posts. In this paper, an examination was conducted to confirm whether explicitly modeling emojis can improve sentiment prediction. An introduction of a linear gated fusion framework is done, which combines contextual text representations from BERTweet with emoji embeddings obtained from Emoji2Vec in an adaptive manner. Experiments on a binary sentiment version of the GoEmotions dataset show that the proposed fusion model consistently outperforms a strong text-only baseline, with particularly better improvements on emoji-containing samples. These findings confirm that emojis contribute supporting sentiment information and that gated fusion is an effective strategy for selectively integrating multimodal signals.*

Keywords— *Sentiment Analysis, Emoji Analysis, Multimodal Learning, Gated Fusion, Social Media Text, BERTweet.*

I. Introduction

Social media sites like Twitter and Reddit have emerged as a major pool of emotional work [1]. Compared to formal text, social media communication is mostly reliant on the use of emojis. These are small paralinguistic symbols carrying affective meanings and pragmatically functional intent [2]. Traditional approaches to sentiment analysis, intended mainly for text, tend not to be so well equipped to handle these underlying subtle emotions, particularly in the case of tweets [1,7].

Recent breakthroughs in transformer models, including BERT and a domain-specific adaptation called BERTweet, have significantly improved the performance of text-based sentiment analysis [3,8]. However, emojis have been considered as normal tokens or simply removed in the preprocessing step. This would end up in the loss of important information about their affective content [2,9]. Previous research has demonstrated that including emojis in a specific model could improve the analysis of their sentiment [2,14], although unconstrained fusion methods may add noise when dealing with the absence or redundancy of emojis [4,5].

This paper tackles these issues by suggesting a gated fusion approach that can dynamically control the input influence of emoji features according to the surrounding text context [12]. Contrary to the static fusion strategy, the goal is to improve text understanding in a targeted manner only when the emoji features contain useful information. Experimental evaluation conducted on the GoEmotions dataset [6] has proven that our method not only yields a high global accuracy of 90.15%, but also brings a 26.8% performance gain for emoji-dense instances compared to text-only baselines.

II. Related Work

A. Text-Based Sentiment Analysis Using GoEmotions

One of the largest publicly available emotion annotation datasets is the GoEmotions dataset, which was first presented by Demszky et al [6]. This dataset provides fine-grained emotion labels for English social media text in 27 different emotion categories. Previous benchmarks of this dataset focused only on text-based transformer architectures, such as the BERT model, which provided a strong performance in sentiment representations. However, according to previous studies, these text-based models face problems when encountering short, informal, or pragmatically complex expressions, which is a common phenomenon in social media, since the expressed emotions do not always include explicit wording.

Further studies used domain-specific pre-trained language models, especially BERTweet, which has been pre-trained on large-scale Twitter datasets, and is better in handling informal texts, slang, hashtags, and other social media features [8]. The performance of text-only sentiment classification with BERTweet-based models was enhanced on other datasets similar to Twitter, such as GoEmotions. Despite their ability to tokenize emojis, most sentiment classification models, especially BERTweet, usually ignore emojis as normal tokens, thereby failing to utilize their affective semantic features [2, 9].

B. Emoji-Aware and Late-Fusion Sentiment

Emojis have been extensively studied as affective symbols with sentiment polarity, sentiment intensities, and pragmatic inferences, which go beyond textual information [2, 7]. The research by Novak et al. has shown that emojis have stable sentiment orientations in large volumes of social media data, while Emoji2Vec has successfully incorporated emojis' semantic meaning into neural models by mapping emojis' vector representations to a word vector space [2, 9], which has encouraged the incorporation of emojis' representations as a supplementary modality in sentiment analysis.

Several emoji-aware sentiment models have been proposed that rely on late fusion strategies. In this case, the emoji features

Akshatha Rithesh, "Emoji-Aware Sentiment Analysis Using Gated Fusion of Textual and Emoji Representations," *International Research Journal of Advanced Engineering and Science*, Volume 11, Issue 1, pp. 105-110, 2026.

or the sentiment score based on the lexicon are concatenated with the text embedding at the decision level. Although this strategy may enhance the overall accuracy of the model, several models have been observed to rely on static fusion, implying that emojis are relevant for every data point. However, this may result in noise for large and unbalanced datasets, as observed with GoEmotions.

In recent approaches, various architectures of multimodal models employed different techniques involving "gated" and "attention" mechanisms to manage the interaction between different modality features [11, 12]. In this direction, a gated multimodal fusion framework was introduced by Wang et al., where modality contribution is selectively controlled, showing better robustness compared to a concatenation approach [12]. In most of the earlier gated fusion approaches, the evaluation was carried out on "audio-visual" and "multimodal" datasets, with "emoji" modality-specific gating strategies not sufficiently explored for emotion classification tasks.

C. *Positioning of the Present Work*

In contrast with previous approaches that only focused on text or late fusion with emojis, this paper proposes a context-aware gated fusion method that controls the contribution of emojis depending on the context. The proposed method improves the representation of text by selectively incorporating emojis only when they are useful in providing information. The proposed method helps in overcoming the disadvantages seen in late fusion approaches and reduces the noise caused by incorporating emojis when they are not useful in the representation of text data, as seen in the Go Emotions dataset.

III. Dataset and Preprocessing

This section describes the datasets explored, preprocessing decisions, and feature extraction procedures adopted in this study. Figure 1 provides a high-level overview of the complete experimental pipeline, from raw data ingestion to model-ready inputs.

3.1 Dataset Selection and Comparative Analysis

The aim of this study is to assess whether there is a quantitative benefit to sentiment classification if emojis are used explicitly together with textual features. As a result, the dataset selection was guided by three key requirements:

(i) The original Unicode emojis were maintained as part of the text,

(ii) Sentiment or emotion labels were present and accurate, and

(iii) The dataset was compatible with existing multimodal frameworks.

Some popular sentiment analysis datasets have also been investigated in the preliminary experimentation. However, it has been found that the majority of the datasets are not appropriate to be used in the context of emoji-aware modeling.

Table 1 shows the list of the investigated datasets and the issues faced. Datasets Explored

Initial experiments with Sentiment140 and SemEval datasets indicated that emojis were totally eliminated during data cleaning, making them unfit for further emoji-sensitive sentiment analysis. TweetEval-Emoji was equally unfit, considering that emojis were used as prediction targets, not features. TweetEval-Sentiment, which retains emojis, was not ideal, having shown extremely sparse usage of emojis during thorough examination.

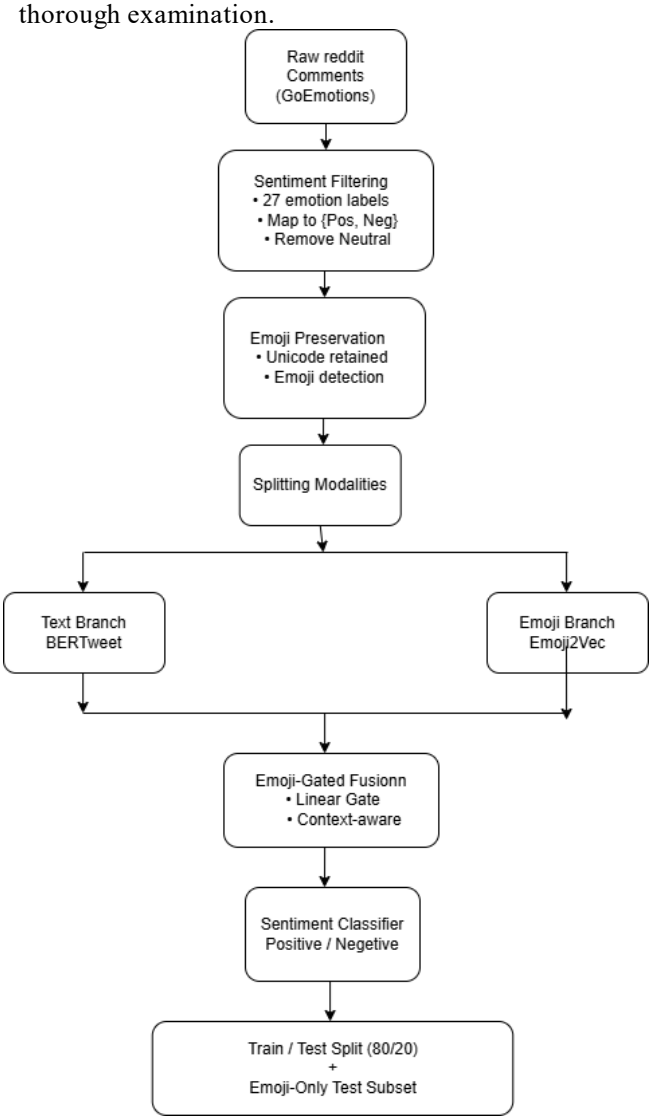

Figure 1. Overview of the proposed preprocessing and modeling pipeline

TABLE 1. Summary of datasets examined and limitations encountered.

Dataset	Task	Emoji Presence	Key Issue Identified
Sentiment140	Twitter sentiment	Removed	Emojis stripped during preprocessing
SemEval-2017 (Task 4)	Twitter sentiment	Removed	Emoji and non-ASCII characters cleaned
TweetEval–Emoji	Emoji prediction	Not in text	Emojis moved to labels, not text
TweetEval–Sentiment	Twitter sentiment	Very sparse	Inconsistent and near-zero emoji usage
GoEmotions	Emotion classification	Preserved	Emotion labels require sentiment mapping

3.2 GoEmotions Corpus Description

Akshatha Rithesh, "Emoji-Aware Sentiment Analysis Using Gated Fusion of Textual and Emoji Representations," *International Research Journal of Advanced Engineering and Science*, Volume 11, Issue 1, pp. 105-110, 2026.

On the basis of the above analysis, the experimental dataset used in this paper is chosen as GoEmotions, where Reddit comments are labeled with 27 finer-grained emotion categories, reflecting various emotions in highly noisy social media texts. Finally, after sentiment mapping and filtering (as discussed in Section 3.3), the resulting dataset contains 18,629 samples, with 2.53% (471 samples) containing at least one emoji. Although emoji usage is not very frequent, this is a realistic representation of how communication occurs and not a bias inherent in the data itself.

In order to evaluate the influence of the emojis separately, a new evaluation set of only emojis, i.e., a set of text instances including only emojis, has been used.

TABLE 2. Class-wise distribution of samples and emoji presence

Sentiment Class	Total Samples	Samples with Emojis	Emoji Percentage
Negative	6,081	113	1.86%
Positive	12,548	358	2.85%
Overall	**18,629**	**471**	**2.53%**

The class-wise distribution of samples and emoji usage in the filtered GoEmotions dataset is presented in Table 3. It is observed that emoji usage is sparse in sentiment classes, with a higher usage in positive sentiment samples.

3.3 Sentiment Mapping and Filtering

The GoEmotions dataset comes with its own set of annotations for 27 different classes of human emotions. For the proposed study, a binary sentiment definition is employed in order to concentrate on polarity classification. The positive and negative classes of human emotions are combined as per the affective taxonomies, while the neutral and ambiguous classes are excluded in order to minimize ground truth noise.3.4 Feature Extraction and Modality Splitting.

TABLE 3. Mapping of GoEmotions labels to binary sentiment

Sentiment	GoEmotions Fine-Grained Labels	Total Samples
Positive	Admiration, Amusement, Approval, Joy, Love, Optimism	12,548
Negative	Anger, Annoyance, Disgust, Fear, Sadness, Remorse	6,081
Neutral	Neutral (excluded)	—

This filtering strategy improves label consistency while maintaining a balanced representation of affective polarity.

3.4 Feature Extraction and Modality Splitting

Each input instance is decomposed into two modality-specific representations.

Textual branch:

Text is tokenized using the BERTweet tokenizer, preserving Twitter-style tokens such as hashtags and user mentions. Inputs are padded or truncated to a maximum sequence length of 128 tokens. The contextual embedding of the [CLS] token is extracted from BERTweet, producing a 768-dimensional text representation.

Emoji branch:

Emojis are extracted using Unicode-based matching without demojization. Emoji semantics are encoded using Emoji2Vec, which maps emojis into a 300-dimensional embedding space aligned with word vectors. For posts containing multiple emojis, embeddings are mean-pooled to obtain a single emoji representation.

$$\mathbf{x}_{text}^{(i)} \in \mathbb{R}^{768}, \quad \mathbf{x}_{emoji}^{(i)} \in \mathbb{R}^{300}$$

These dual representations form the inputs to the gated fusion architecture described in Section 4.

3.5 Class Imbalance and Sampling Strategy

The filtered dataset is imbalanced, indicating a larger proportion of positive samples. This was handled by incorporating a weighted cross-entropy loss during training, which utilized class inverses frequencies.

Moreover, in order for effective learning of the emoji contributions, targeted oversampling of the emoji-containing samples in the training set has also been applied, increasing their relative weight in the set to around 10%. This is done in an empirical manner such that enough gradient signal is provided for effective learning of the gate mechanism without leading to overfitting of repeated instances of emojis.

3.6 Evaluation Subsets

Two sets of complementary evaluation instruments were used:

Global Test Set:

An 80-20 split between training and testing data, commonly used to evaluate the overall classification performance.

Emoji Only Test Subset (Targeted Ablation Subset):

The focused set, including 82 emoji-containing samples, was designed to isolate and quantify the impact of the emoji modality, independent of text dominance.

IV. METHODOLOGY

The following section explains the proposed emoji-aware sentiment classification model, which includes the proposed baseline model, gated fusion model, as well as training. The aim is to evaluate the emoji influence while being fair and comparable with other baseline models.

A. Problem Definition

Given a social media post T and its associated emoji set E, the task is to predict a binary sentiment label

$$y \in \{0,1\}$$

associated with negative and positive sentiment, respectively

Each input instance is decomposed into two representations, one for each modality:

- a textual representation derived from a pretrained transformer,
- an emoji representation derived from pretrained emoji embeddings.

4.2 Text-Only Baseline Model

The text-only baseline uses the BERTweet model, which is a transformer model pretrained on 850M English tweets. The model is particularly good at handling social media text because it has been exposed to informal text, hashtags, mentions, and emojis.

Akshatha Rithesh, "Emoji-Aware Sentiment Analysis Using Gated Fusion of Textual and Emoji Representations," *International Research Journal of Advanced Engineering and Science*, Volume 11, Issue 1, pp. 105-110, 2026.

Given an input text T, the contextual embedding of the special classification token is extracted:

$$\mathrm{h}_{\text{text}} = \text{BERTweet}(T)_{[\text{CLS}]} \in \mathbb{R}^{768}$$

This representation is passed through a feed-forward neural classifier, which produces sentiment logits.

This model acts as the primary baseline when measuring the success of the proposed method of emoji fusion.

4.3 Emoji-Only Baseline Model

In order to test the standalone prediction capability of emojis, a baseline consisting only of emojis is built. The emojis are extracted from the input, and Emoji2Vec embeddings are used, where emojis are mapped to a semantic space of 300 dimensions similar to word embeddings.

For tweets containing multiple emojis, a mean-pooled representation is computed:

$$\mathrm{h}_{\text{emoji}} = \frac{1}{|E|}\sum_{e\in E} \text{Emoji2Vec}(e)$$

This baseline measures the degree to which sentiment polarity is encoded in emojis alone.

4.4 Emoji-Gated Fusion Model

To integrate textual and emoji information while avoiding noise from irrelevant emoji signals, a linear gated fusion architecture is proposed.

Feature Projection

The emoji embedding is first projected into the textual embedding space:

$$\mathrm{h}'_{\text{emoji}} = W_e \mathrm{h}_{\text{emoji}} + b_e$$

Gating Mechanism

The contribution of emoji features is dynamically controlled by a learnable gate based on the textual context.

$$g = \sigma\left(W_g[\mathrm{h}_{\text{text}}; \mathrm{h}'_{\text{emoji}}] + b_g\right)$$

The computation of gated emoji representation is done as:

$$\tilde{\mathrm{h}}_{\text{emoji}} = g \odot \mathrm{h}'_{\text{emoji}}$$

Fusion and Classification

The final fused representation is obtained via concatenation as follows:

$$\mathrm{h}_{\text{fusion}} = [\mathrm{h}_{\text{text}}; \tilde{\mathrm{h}}_{\text{emoji}}]$$

This representation is passed through a multilayer perceptron to predict sentiment logits.

This gating mechanism allows the model to amplify emoji signals when informative and suppress them when redundant or misleading, addressing a key limitation of static fusion strategies [12].

4.5 Training Strategy

Loss Function

To address class imbalance, a weighted cross-entropy loss is employed:

$$\mathcal{L} = -\sum_{c\in\{0,1\}} w_c \, y_c \log(\hat{y}_c)$$

where class weights w_c are inversely proportional to class frequencies.

Emoji-Focused Oversampling

Oversampling is used to increase the ratio to 10%, as only 2.53% of the dataset contained emojis. This approach guarantees the gradient signal is strong enough to guide the gating process while avoiding overfitting to the duplicate data.

Optimization

The model is trained with the AdamW optimizer with a learning rate of 2×10^{-5} for four epochs. All the baselines and the models using the fusion method have the same training settings.

4.6 Evaluation Protocol

Two complementary evaluation settings are adopted:

- *Global Test Set:*
  A standard 80–20 train–test split used to evaluate overall classification performance.
- *Emoji-Only Test Subset (Targeted Ablation Subset):*
  A focused subset of 82 emoji-containing samples used to isolate and quantify the contribution of the emoji modality.

This dual evaluation protocol enables both global performance assessment and targeted modality impact analysis.

V. Results and Discussion

This section presents the experimental results of the proposed emoji-aware gated fusion model and compares its performance against text-only and emoji-only baselines. Results are reported on both the full test set and a targeted emoji-only evaluation subset to assess the true contribution of the emoji modality.

5.1 Experimental Setup

All models were evaluated using an 80–20 train–test split of the filtered GoEmotions dataset. Classification performance was measured using Accuracy and Macro F1-score, which is particularly suitable for imbalanced sentiment distributions. To isolate the effect of emojis, an additional Emoji-Only Test Subset containing 82 emoji-bearing samples was used for targeted analysis.

5.2 Performance on the Full Test Set

Table 4 summarizes the performance of the evaluated models on the complete test set.

TABLE 4. Performance comparison on the full test set

Model	Accuracy	Macro F1
Text-Only (BERTweet)	0.8994	0.8861
Emoji-Only (Emoji2Vec)	0.6728	0.4069
Text + Emoji Gated Fusion (Proposed)	**0.9015**	**0.8895**

The text-only BERTweet baseline model performs well, which validates the use of transformer-based models for sentiment classification. The emoji-only model performs significantly worse, which shows that using only emojis is not sufficient for sentiment classification.

The proposed text + emoji gated fusion model shows a significant improvement over the text-only model, which is +0.21% accuracy and +0.34% Macro F1. Although this improvement is relatively small, it is significant given the strength of the baseline model and the relatively rare occurrence of emojis in the data (2.53%).

5.3 Emoji-Only Subset Analysis (Targeted Ablation Study)

Akshatha Rithesh, "Emoji-Aware Sentiment Analysis Using Gated Fusion of Textual and Emoji Representations," *International Research Journal of Advanced Engineering and Science*, Volume 11, Issue 1, pp. 105-110, 2026.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

To analyze the effectiveness of emojis, the model is evaluated on the emoji-only subset, which is a subset of data containing only emojis.

The results clearly show a large performance difference between the text-only model and the proposed gated fusion model. Although the text-only model performs poorly on emoji-rich data, the gated fusion model obtains a large accuracy improvement of 26.8% compared to the text-only model, indicating its effectiveness in utilizing emoji information when it is available.

TABLE 5. Performance on Emoji-Only Test Subset

Model	Accuracy	Macro F1
Text-Only (BERTweet)	0.6707	0.5610
Emoji-Gated Fusion (Proposed)	**0.9390**	**0.9159**

Figure 2 is a visualization of this performance difference and clearly shows the large performance divergence of the two models on emoji-rich data.

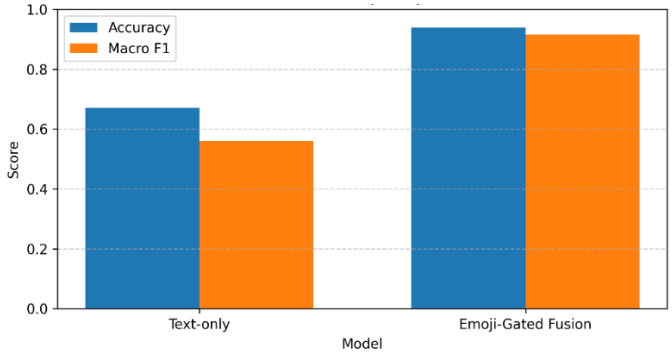

Figure 2: Performance Comparison of Emoji-Only Subset

5.4 Discussion

The experimental results have provided several important insights:

1. Limited Global Gain, Strong Conditional Gain:

The low frequency of emoji appearance in the dataset means that the global performance improvement is naturally limited. However, the emoji appearance does have a strong impact when it does occur.

2. Effectiveness of Gated Fusion:

The gated fusion mechanism has been able to eliminate the emoji noise when the sample does not contain emojis and enhance the emoji impact when the sample does contain emojis. This verifies the effectiveness of adaptive fusion over static concatenation for emoji-aware sentiment analysis.

3. Emoji Semantics Are Context-Dependent:

The significant improvement in the emoji-only evaluation indicates that the semantics of the emojis are highly dependent on the context.

4. Limitations of Emoji-Only Modeling:

The emoji-only baseline verifies the limitation of relying on the semantics of the emojis alone.

5.5 Key Takeaways

- Emojis provide contextual sentiment cues, not standalone sentiment signals.
- Global accuracy gains may appear small, but conditional improvements on emoji-bearing samples are substantial and practically meaningful.
- Gated fusion enables selective modality utilization, addressing a major weakness of prior late-fusion approaches.

VI. Conclusion

In this work, the question of whether emojis can provide measurable benefits to sentiment classification when explicitly modelled alongside textual representations has been investigated. By systematically evaluating several sentiment classification datasets, it has been noted that most commonly used datasets for sentiment classification tasks have emojis suppressed or removed. As such, experiments were carried out on the GoEmotions dataset, which includes raw Unicode emojis in their original form, thus providing a realistic setting for multimodal modeling.

A gated emoji-text fusion model has been proposed to dynamically regulate the contribution of the emoji representation depending on the textual context. Although the overall performance gain over the strong text-only baseline is necessarily limited by the sparsity of emojis, significant improvements were noted on the emoji-bearing samples. In particular, the 26.8% absolute accuracy gain on the emoji-only evaluation set shows that emojis can serve as strong sentiment amplifiers when adaptively fused. As such, this work has shown that adaptive fusion mechanisms are critical to effectively utilizing the semantics of emojis without introducing noise.

VII. Future Work

In future work, the question of how-to best leverage emoji-aware pretraining for more effective alignment of textual and symbolic representations will be addressed. In addition to this, the extension of the proposed framework to support multilingual and cross-platform social media data is also a promising area of future work. Moreover, the incorporation of other expressive media such as hashtags and images, as well as the extension of the classification task from binary sentiment to more fine-grained emotion classification, is also a promising area of future work.

References

[1] Pang, Bo, and Lillian Lee. "Opinion Mining and Sentiment Analysis." *Foundations and Trends® in Information Retrieval*, vol. 2, no. 1–2, Now Publishers, 2008, pp. 1–135.

[2] Novak, Petra Kralj, Jasmina Smailović, Borut Sluban, and Igor Mozetič. "Sentiment of Emojis." *PLOS ONE*, vol. 10, no. 12, 2015, e0144296. https://doi.org/10.1371/journal.pone.0144296.

[3] Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding." *Proceedings of NAACL-HLT*, Association for Computational Linguistics, 2019, pp. 4171–4186.

[4] Zadeh, Amir, Minghai Chen, Soujanya Poria, Erik Cambria, and Louis-Philippe Morency. "Tensor Fusion Network for Multimodal Sentiment Analysis." *Proceedings of EMNLP*, Association for Computational Linguistics, 2017, pp. 1103–1114.

[5] Mai, Sijie, Haifeng Hu, and Songlong Xing. "Multimodal Sentiment Analysis: A Survey." *IEEE Transactions on Affective Computing*, vol. 12, no. 3, 2020, pp. 1–20.

[6] Demszky, Dorottya, Dana Movshovitz-Attias, Jeongwoo Ko, Alan Cowen, Gaurav Nemade, and Sujith Ravi. "GoEmotions: A Dataset of Fine-Grained Emotions." *Proceedings of ACL*, Association for Computational Linguistics, 2020, pp. 4040–4054.

[7] Mohammad, Saif M., Felipe Bravo-Marquez, Mohammad Salameh, and Svetlana Kiritchenko. "SemEval-2018 Task 1: Affect in Tweets." *Proceedings of SemEval*, Association for Computational Linguistics, 2018, pp. 1–17.

[8] Nguyen, Dat Quoc, Thanh Vu, and Anh Tuan Nguyen. "BERTweet: A Pre-trained Language Model for English Tweets." *Proceedings of EMNLP*, Association for Computational Linguistics, 2020, pp. 9–14.

[9] Eisner, Ben, Tim Rocktäschel, Isabelle Augenstein, Matko Bošnjak, and Sebastian Riedel. "Emoji2Vec: Learning Emoji Representations from Their Description." *Proceedings of EMNLP*, Association for Computational Linguistics, 2016, pp. 48–54.

[10] Tsai, Yao-Hung Hubert, Shaojie Bai, Makoto Yamada, Louis-Philippe Morency, and Ruslan Salakhutdinov. "Multimodal Transformer for Unaligned Multimodal Language Sequences." *Proceedings of ACL*, Association for Computational Linguistics, 2019, pp. 6558–6569.

[11] Hazarika, Devamanyu, Soujanya Poria, Amir Zadeh, Erik Cambria, Louis-Philippe Morency, and Roger Zimmermann. "MISA: Modality-Invariant and -Specific Representations for Multimodal Sentiment Analysis." *Proceedings of EMNLP*, Association for Computational Linguistics, 2020, pp. 1122–1131.

[12] Wang, Shuai, Zhen Zhang, and Yueting Zhuang. "A Gated Multimodal Fusion Framework for Emotion Recognition." *Information Fusion*, vol. 70, Elsevier, 2021, pp. 1–12.

[13] Liu, Pengfei, Xipeng Qiu, and Xuanjing Huang. "Pretrained Language Models for Text Generation: A Survey." *Proceedings of ACL*, Association for Computational Linguistics, 2021, pp. 1–23.

[14] Wu, Yongliang, Zhiqiang Zhang, and Jingjing Liu. "Emoji-Aware Sentiment Analysis Using Attention-Based Fusion." *Knowledge-Based Systems*, vol. 238, Elsevier, 2022, Article 107899.

[15] Zhang, Hongfei, Liang Zhao, and Yong Liu. "Multimodal Emotion Recognition Based on Deep Feature Fusion." *IEEE Access*, vol. 11, IEEE, 2023, pp. 45621–45633.

[16] Felbo, Bjarke, Alan Mislove, Anders Søgaard, Iyad Rahwan, and Sune Lehmann. "Using Millions of Emoji Occurrences to Learn Any-Domain Representations for Detecting Sentiment, Emotion and Sarcasm." *Proceedings of EMNLP*, Association for Computational Linguistics, 2017, pp. 1615–1625.