**RESEARCH PROPOSAL TEMPLATE /**

**THUYẾT MINH ĐỀ CƯƠNG NGHIÊN CỨU**

|  |  |
| --- | --- |
| **Research title** | A Survey about Speech emotion recognition using Graph Neural Network |
| **Members**  *(2 -3 students)* | Nguyen Minh Nhut |
| Full name 2 |

**Abstract/Tóm tắt**

*Research objectives, research questions, research problems, research methodology (maximum 250 words) / Giới thiệu tóm tắt về đề tài (nêu mục tiêu nghiên cứu, câu hỏi nghiên cứu, vấn đề nghiên cứu và phương pháp nghiên cứu, không quá 250 từ).*

The primary objective of this study is to investigate current methods for speech emotion recognition (SER) and evaluate how Graph Neural Networks (GNNs) can be applied to enhance these techniques. This study explores existing SER methodologies and evaluates how GNNs can be integrated to enhance performance. By leveraging the graph structure, GNNs provide a comprehensive understanding of the data's spatial and relational context, potentially outperforming traditional methods. The analysis will include comparing the strengths and weaknesses of GNNs against CNNs, RNNs, and transformers, highlighting their ability to address data heterogeneity and interdependencies. Despite their advantages, employing GNNs in SER presents challenges such as model complexity and computational demands. This investigation will also identify potential hurdles, including managing diverse data types and optimizing GNN architectures, thereby offering insights into the practical application of GNNs for advancing SER technologies.

**Key words: Speech emotion recognition, Graph Neural Network**

1. **Introduction / Giới thiệu** 
   1. **1.1.Literature review / Tình hình nghiên cứu trong và ngoài nước**

*Review and discuss the related research work(s) done by Vietnamese and foreign scientists. / Nêu rõ những nghiên cứu do các nhà khoa học nước ngoài, trong nước tiến hành theo hướng nghiên cứu của đề tài và các kết quả đạt được, đồng thời nhận xét về những kết quả đã nêu.*

*Review the applications of related works in reality / Nêu rõ ứng dụng của các nghiên cứu này trên thực tế.*

Yan Li et al. [1] establish the speech-graph using feature similarity and introduce a novel architecture for graph neural networks that leverages an LSTM aggregator and weighted pooling. The unweighted accuracy of 65.39% and the weighted accuracy of 71.83% are obtained on the IEMOCAP dataset, achieving performance comparable to or better than existing graph baselines. This method can improve the interpretability of the model to some extent, and identify speech emotion features effectively.

Chenquan Gan et al. [2] present a novel Graph Neural Network approach for conversational emotion recognition at the utterance level. To efficiently capture the semantic content of the conversations, they leverage the zero-shot feature-extraction capabilities of pre-trained large-scale language models and then integrate two key contributions into the graph neural network to ensure competitive recognition results. The first is a novel context filter that establishes meaningful utterance dependencies for the graph construction procedure and removes low-relevance and uninformative utterances from being used as a source of contextual information for the recognition task. The second contribution is a feature-correction procedure that adjusts the information content in the generated feature representations through a gating mechanism to improve their discriminative power and reduce emotion-prediction errors. The accuracy and F1-score on IEMOCAP dataset are respectively 70.22%, 69.77% and MELD dataset is 61.72%, 59.74%.

Chenyu Liu et al. [3] have a survey, their categorization reveals the commonalities and differences of existing approaches under a unified framework of graph construction. We analyze and categorize methods from three stages in the framework to provide clear guidance on constructing GNNs in EEG-based emotion recognition. In addition, we discuss several open challenges and future directions, such as Temporal full-connected graph and Graph condensation.

Junghun Kim et al [4] propose a cosine similarity-based graph structure as an ideal graph structure for SER, and present the Cosine similarity-based Graph Convolutional Network (CoGCN), as a Graph Convolutional Network (GCN) variant for SER. Finally, they show that our method outperforms state-of-the-art methods or provides competitive results with a significant model size reduction with only 1/30 parameters. In the experiment result of this study, the weighted accuracy is 62.64% and the unweighted accuracy is 63.67%.

* 1. **1.2. The limitation of current works / Những hạn chế của các nghiên cứu hiện tại**

*Clearly explain the limitations of current work and the necessity of your research. This helps establish the significance and originality of your study.*

*Nêu rõ những hạn chế của các nghiên cứu hiện tại để từ đó nêu lên được sự cần thiết của nghiên cứu này.*

* 1. **1.3. The necessity of the research / Sự cần thiết tiến hành nghiên cứu**

*Clearly define the problem or hypothesis to be addressed. / Nêu rõ vấn đề mà đề tài tập trung giải quyết.*

*Originality, relevance, and scientific significance of the question under investigation. / Phân tích tính mới, tính thời sự, ý nghĩa khoa học và sự cần thiết của vấn đề cần nghiên cứu.*

1. **Research objectives / Mục tiêu của đề tài**

*Clearly outline the research objectives as a basis for determining research content and implementation plan. / Nêu rõ mục tiêu cần đạt được của đề tài, làm cơ sở xác định nội dung nghiên cứu và kế hoạch triển khai.*

1. **Research scope / Phạm vi nghiên cứu**

*Clearly define the key study topics and the research scope. / Nêu rõ những nội dung nghiên cứu chính, phạm vi bao quát của đề tài.*

1. **Feasibility of research / Tính khả thi của đề tài**

*Analyze the feasibility of research / Phân tích tính khả thi của đề tài.*

1. **Approach and Method / Cách tiếp cận và phương pháp nghiên cứu**

*Describe research methodology and techniques appropriate for each issue of the research. / Mô tả chi tiết cách tiếp cận, phương pháp nghiên cứu, kỹ thuật sử dụng phù hợp với từng nội dung của đề tài nhằm đạt được mục tiêu của đề tài.*

1. **Research plan / Kế hoạch thực hiện nghiên cứu**

*Describe the research plan* (*Give* ***a concise description of each task****, and their expected results,* ***schedule,*** *and* ***person in charge*** *as in the following table)* / Mô tả kế hoạch thực hiện của toàn bộ dự án nghiên cứu.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No.** | **Date** | **Task** | **Output** | **Person in charge** |
| 1 | dd/mm/yyyy – dd/mm/yyyy |  |  |  |
| 2 |  |  |  |  |
| ... |  |  |  |  |

**Computational Resource Requirements (if any)**

1. **Expected results / Dự kiến kết quả đề tài**

*For example: theoretical, methodological, and practical implications; and plans to apply this research in practical. / Nêu dự kiến kết quả nghiên cứu (Đề xuất về mặt lý thuyết, phương pháp và ứng dụng, và khả năng ứng dụng vào thực tế, …);*

1. **Publication Plans / Kế hoạch công bố kết quả nghiên cứu**

*If applicable, mention plans for disseminating your research, such as conference presentations or journal submissions. / Dự kiến công bố kết quả nghiên cứu ở hội thảo quốc tế hay tạp chí chuyên ngành.*

*Conference paper will be submitted to International Conference on ICT Convergence*

*Journal paper will be submitted to Expert Systems With Applications*

**References**

**Appendix (*if any*)**

# References

|  |  |
| --- | --- |
| [1] | Y. a. W. Y. a. Y. X. a. I. S.-K. Li, Speech emotion recognition based on Graph-LSTM neural network, Springer, 2023. |
| [2] | C. a. Z. J. a. Z. Q. a. J. D. K. a. {. V. Gan, A graph neural network with context filtering and feature correction for conversational emotion recognition, Elsevier, 2024. |
| [3] | C. a. Z. X. a. W. Y. a. Y. R. a. Z. L. a. J. Z. a. L. Y. Liu, "Graph Neural Networks in EEG-based Emotion Recognition: A Survey," *arXiv preprint arXiv:2402.01138,* 2024. |
| [4] | J. a. K. J. Kim, "Representation learning with graph neural networks for speech emotion recognition," *arXiv preprint arXiv:2208.09830,* 2022. |