CROSS CULTURAL FACIAL EMOTION RECOGNITION USING TRANSFER LEARNING

ANNOTATED BIBILOGRAPHY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS OF THE DEGREE OF

BACHELOR OF THE SCIENCE IN ENGINEERING

Submitted by:

Kanistan K. (2021/E/064)

DEPARTMENT OF COMPUTER ENGINEERING

FACULTY OF ENGINEERING

UNIVERSITY OF JAFFNA

[JANUARY] [2025]

1. Multi frame transfer learning framework for facial emotion in e-learning context

(1) Pordoy, J. et al. (2024) 'Multi-frame Transfer Learning Framework for Facial Emotion Recognition in e-learning contexts', IEEE Access, 12, pp. 151360–151381. doi:10.1109/access.2024.3478072.

(2) This study introduces a method combining multi-frame analysis and transfer learning to enhance emotion recognition, addressing challenges in detecting nuanced e-learning emotions like boredom, confusion, and engagement. **(3)** The study aims to improve FER performance by combining transfer learning with sequential frame analysis to extract dominant emotions. The methods involve fine-tuning pre-trained models on AffectNet and FER2013 datasets. **(4)** The research focuses on using FER to dynamically recognize students' emotions during e-learning, contributing to personalized learning experiences and improved engagement through video-based emotion detection techniques. **(5)** This research provides helpful ideas for my work on cross-cultural FER by showing how transfer learning and multi-frame methods can be used effectively in systems that recognize dynamic emotion. **(6)** The limitation of the proposed framework lies in its susceptibility to biases stemming from variable data collection conditions, cultural differences, and limited datasets, impacting its generalizability and performance in diverse contexts. **(7)** The proposed framework, with an overall average accuracy of 0.878, improves e-learning by detecting emotions over time, helping create adaptive e-learning and enhancing students' learning outcomes. **(8)** This paper supports my research on cross-cultural FER by offering methods to adapt transfer learning frameworks, addressing global variations in emotional expressions for enhanced applicability in diverse cultural contexts.

2. Facial emotion recognition using conventional machine learning and deep learning methods: Current achievements, analysis, and remaining challenges

(1) Khan, A.R. (2022) 'Facial emotion recognition using conventional machine learning and deep learning methods: Current achievements, analysis and remaining challenges', *Information*, 13(6), p. 268. doi:10.3390/info13060268.

(2) The author discusses the importance of facial emotion recognition (FER) in pattern recognition, its applications in various fields, and highlights the growing importance of non-verbal communication and the role of FER systems in understanding human emotions in different situations. **(3)** The article aims to evaluate FER technologies by looking at different datasets, classification methods, and performance metrics. It uses a systematic review of studies that apply classifiers like SVM, CNN, and LSTM and it also discusses about FER datasets like CK+ and JAFFE, explaining their unique features and how they help improve FER technologies. **(4)** The scope of the research includes a wide range of traditional ML techniques and advanced DL methods used in FER, along with an analysis of publicly available datasets for evaluating these approaches. **(5)** This article is highly useful for my research, as it provides fundamental knowledge about existing FER methodologies and explains the challenges faced in automating emotion detection. **(6)** Primarily focuses on traditional ML and DL methods without going into detail about cross-cultural differences in emotion recognition. **(7)** The study highlights FER advancements, achieving an average accuracy of 0.99 Future research should explore 3D data and IoT integration to enhance accuracy and expand real-world applications. **(8)** This article helps my research by providing useful information about cross-cultural challenges but does not directly talk about transfer learning

3. Facial emotion recognition using Handcrafted features and CNN

(1) Gautam, C. and Seeja, K.R. (2023) 'Facial emotion recognition using handcrafted features and CNN', *Procedia Computer Science*, 218, pp. 1295–1303. doi:10.1016/j.procs.2023.01.108.

(2) The paper discusses challenges in recognizing emotions from facial expressions and explain the importance of identifying emotions accurately and proposes combining handcrafted features extraction techniques with convolutional neural networks (CNN) for improved accuracy. **(3)** The research aims to introduce a new method for emotion recognition that combines handcrafted feature extraction methods (HOG, SIFT) with CNN to enhance accuracy and efficiency. The study used CK+ and JAFFE datasets for preprocessing, feature extraction, and classification using CNN-based sequential models and performance evaluation metrics. **(4)** This study investigates improving emotion recognition by using HOG and SIFT features with CNNs, focusing on small datasets**. (5)** This research is important for understanding emotions across different cultures as it shows effective ways to extract features for emotion recognition using varied datasets. **(6)** Depending on CK+ and JAFFE datasets limits how well the findings apply to wider since these datasets are controlled and do not reflect the diversity or variety of cross-cultural situations. **(7)** The suggested HOG-CNN model performed better than other models and the results show that handcrafted feature extraction methods are useful for emotion recognition, especially when combined with CNN, as they achieving an average of accuracy 0.93. **(8)** This paper helps my research by showing how handcrafted feature extraction and CNN work together in emotion recognition.

4. CNN-BiLSTM based Facial Emotion Recognition

(1) Lamichhane, A. and Karn, G. (2024) 'CNN-BILSTM based facial emotion recognition', *International Journal on Engineering Technology*, 2(1), pp. 227–236. doi:10.3126/injet.v2i1.72579.

(2) This article discusses the importance of facial emotion recognition in human-computer interaction and introduces a CNN-BiLSTM hybrid model to improve accuracy in emotion classification from facial expressions. **(3)** The research aims to enhance emotion recognition accuracy by developing a hybrid model combining CNN and BiLSTM to effectively process spatial and temporal features from facial expression images. The study trains and evaluates the hybrid model using the FER2013 dataset, with CNN handling feature extraction and BiLSTM capturing time-related patterns for better emotion classification. **(4)** The study's scope covers the application of deep learning techniques in recognizing emotions from facial expressions, specifically evaluating the effectiveness of the CNN-BiLSTM model on a larger dataset. **(5)** This study is an important contribution to understanding hybrid deep learning models and supports using transfer learning to recognize emotions across different cultures by identifying important image and time-related features. **(6)** The primary limitation lies in the dataset's class imbalance, reducing the model's effectiveness for lower dataset emotions such as disgust. **(7)** The study concludes that the CNN-BiLSTM model excels in sequence-based tasks but shows minimal accuracy gains for static images, achieving an accuracy of 0.79. **(8)** The results show how combined models can handle both spatial and time-based data, and provide ideas for adapting transfer learning models to better understand cultural differences in emotional expressions, which aligns with my research.

5. Facial Expression Recognition Using Visible, IR, and MSX Images by Early and Late Fusion of Deep Learning Models

(1) Naseem, M. T., Lee, C.-S., & Kim, N.-H. (2024). Facial Expression Recognition Using Visible, IR, and MSX Images by Early and Late Fusion of Deep Learning Models. *IEEE Access*, 12, 20692-20704. DOI: 10.1109/ACCESS.2024.3362247.

(2) This article introduces a new way to recognize facial emotions by using early and late fusion methods with visible, infrared (IR), and multispectral dynamic imaging (MSX) data. **(3)** The research aims to improve emotion recognition accuracy by combining different types of data using fusion techniques, and this helps classify facial expressions more effectively in various lighting and environmental conditions. The study used early and late fusion techniques with ResNet-18 architecture, transfer learning, and fine-tuning across visible, IR, and MSX datasets to evaluate model performance. **(4)** The research focuses on using visible, infrared (IR), and multispectral (MSX) images to improve facial emotion recognition through multi-modality and different training strategies. **(5)** This study provides valuable insights into how fusion techniques can improve accuracy in emotion recognition, and it also supports using transfer learning to study emotions across different cultures. **(6)** The primary limitation is the small dataset size, even after augmentation, which led to overfitting issues, particularly for infrared images, limiting the model's real-world applicability. **(7)** It concludes that early and late fusion methods with ResNet-18 significantly improve classification accuracy, achieving an accuracy of 0.84. **(8)** The findings provide useful ideas for combining multimodal fusion in facial emotion recognition, aligning with my research by combining spatial and temporal data effectively.

This document contains eight different points as follows:

1. Citation

2. Introduction

3. Aims and Research methods

4. Scope

5. Usefulness (to your research/to a particular topic)

6. Limitations

7. Conclusions

8. Reflections (explain how this work illuminates your topic or how it will fill in with your research).