Task 1

Emotion Detection Using Images and Audio

1. Introduction

In today's digital age, understanding human emotions plays a crucial role in various applications, ranging from customer sentiment analysis to human-computer interaction. The Emotion Detection project aims to harness the power of deep learning for accurately identifying emotions in both images and audio files. By leveraging state-of-the-art neural networks, this project contributes to the advancement of emotion recognition technology.

2. Background

Emotion detection has witnessed significant advancements with the advent of deep learning techniques. Traditional methods often relied on handcrafted features, but deep learning models excel at automatically learning relevant features from data. This project builds upon this paradigm shift and incorporates convolutional neural networks (CNNs) for image analysis and for audio processing. Previous research in the field has shown promising results, motivating our exploration into the integration of these methods.

3. Learning Objectives

- Develop a deep learning model capable of accurately classifying emotions in images and audio.
- Explore the use of convolutional neural networks (CNNs) for image and audio processing, respectively.
- Enhance proficiency in data pre-processing techniques to optimize model performance.

- Understand the challenges and intricacies of integrating multiple modalities for comprehensive emotion detection.

4. Activities and Tasks

Data Collection:

Curated datasets containing diverse images and audio clips labelled with corresponding emotions. For this project I have used many datasets combined.

For images I used FER-2013 dataset and for audio

1.Ravdess-Emotional Speech Audio

2.toronto emotional speech set

3.surrey-audiovisual-expressed-emotion-savee

4.Crowd Sourced Emotional Multimodal Actors Dataset (CREMA-D)

Model Development:

Implemented a CNN architecture for image emotion classification.

Developed an CNN model for audio emotion analysis.

Integrated the two models for a multimodal emotion detection system.

Testing and Evaluation:

Split datasets into training and testing sets. Fine-tuned models and conducted rigorous testing for performance evaluation.

5. Skills and Competencies

- Proficiency in Python for deep learning implementation (TensorFlow).
- Mastery of image and audio pre-processing techniques.

- Understanding of CNNs for image and audio analysis.

6. Feedback and Evidence

To provide continuous feedback, regular code reviews and milestone evaluations was conducted. I was expected to submit well-documented code, and evidence of successful completion of tasks will include trained models, performance metrics, and a functioning drowsiness detection system.

7. Challenges and Solutions

Challenges:

Addressing overfitting in the CNN and RNN models.

Solutions:

Tuned the model for better results.

8. Outcomes and Impact

The implemented model achieved an accuracy of over 94% in emotion classification for both audio and for images it was about 85%. This project sets the stage for real-world applications, such as improving user experience in human-computer interaction and enhancing sentiment analysis in multimedia content.

9. Conclusion

The Emotion Detection project successfully demonstrates the feasibility of employing deep learning techniques for multimodal emotion analysis. The project not only enhances technical skills in deep learning but also showcases the potential impact of emotion recognition in diverse domains. Recommendations for future work include expanding the dataset and exploring advanced architectures for improved model performance.