**Project Information**

This project presents a comprehensive investigation into a novel framework for emotion recognition that combines multimodal deep learning with principles from quantum cognition. The research explores the viability and challenges of integrating quantum machine learning (QML) into a complex, real-world classification task that uses data from four distinct modalities: facial images, speech, text, and physiological EEG signals.

The core of the research involves the design, training, and comparative analysis of several advanced models:

1. A robust **Classical Baseline Model** using intermediate fusion.
2. A higher-performing **Classical Hybrid Fusion Model**.
3. An experimental **Variational Quantum Classifier (VQC)**.
4. A comparative **Quantum Support Vector Machine (QSVM)**.

The final results demonstrate that while the classical hybrid fusion model provides a stable and effective solution, the variational quantum model suffers from fundamental instability. This work contributes a valuable analysis of the practical challenges of applying current VQCs and provides a clear benchmark for future research in hybrid quantum-classical AI.

**Model Architecture: The Classical Hybrid Fusion Model**

This is the best-performing classical architecture developed in this research. It is designed to mimic human perception by processing multiple data streams in parallel, intelligently fusing them first in pairs and then globally to make a final, context-aware prediction.

* **Four Parallel Feature Extractors**

The model has four separate "streams," each specialized in understanding one type of data.

* + 🖼️ **Image Stream:** A **Convolutional Neural Network (CNN)** analyzes the facial expression images from the FER-2013 dataset.
  + 🗣️ **Speech Stream:** Another **CNN** processes the acoustic features (MFCCs) extracted from the RAVDESS audio dataset.
  + 📜 **Text Stream:** A pre-trained **Transformer model (BERT)** reads and understands the semantic meaning of the text data.
  + 🧠 **EEG Stream:** A **Multi-Layer Perceptron (MLP)** processes the physiological brainwave feature vectors from the DEAP dataset.
* **Paired "Mini-Fusion" Layers**

To capture the rich interplay between key modalities, the model first creates specialized "paired" features before the final fusion.

* + **Audio-Visual Fusion:** A dedicated layer combines the features from the **Image CNN** and **Speech CNN** to learn from simultaneous visual and auditory cues.
  + **Text-Acoustic Fusion:** Another layer combines the features from the **BERT Model** and **Speech CNN** to learn from the interplay between words and tone.
* **Final Global Fusion & Classifier**

All available features—the four original unimodal features and the two new paired features—are combined into a single, rich feature vector through concatenation. This vector is then passed to a final **Multi-Layer Perceptron (MLP)**, which acts as the classifier head to make the final prediction across the seven emotion classes.

**Experimental Analysis and Key Findings**

The research involved a series of rigorous experiments to validate the proposed architectures.

* **Ablation Study**

An ablation study on the classical model revealed that the visual and language-based modalities (Image, Text, Speech) were the most dominant contributors to performance. However, the full four-modality model achieved the highest F1-score, indicating that it produced the most balanced and nuanced predictions across all emotion classes.

* **Comparative Model Performance**

The final results provide a clear comparison of the different architectural paradigms:

* + **Classical Baseline:** Achieved a stable accuracy of **41.12%**.
  + **Classical Hybrid Fusion:** The best-performing model, achieving a final accuracy of **43.49%** with stable training.
  + **Variational Quantum Classifier (VQC):** Proved to be fundamentally unstable in all configurations, consistently suffering from "mode collapse" where it would only predict a single emotion.
  + **Quantum Support Vector Machine (QSVM):** Was shown to be a stable quantum learner, achieving **48.60%** accuracy on a simplified, unimodal task.
* **Conclusion on Quantum Models**

The key finding of this research is that while the **QSVM** was a stable quantum learner, the **VQC** suffered from a difficult optimization problem, as evidenced by its "bumpy" loss landscape. This highlights a critical and open challenge in the practical application of current variational quantum algorithms for complex, real-world tasks.