Multimodal Hierarchical CNN Feature Fusion for Stress Detection

Radhika Kuttala, Ramanathan Subramanian,
Venkata Ramana Murthy Oruganti
EE798 final assignment, Name: Amol Patil, Roll no.: 210711

Abstract

This study presents a hierarchical CNN-based feature fusion model for stress detection using EEG signals. By integrating low, mid, and high-level features, the model captures a comprehensive representation of physiological data, enhancing accuracy in classifying stressed versus non-stressed states. The proposed fusion approach consistently outperforms traditional methods by leveraging multiple feature levels for robust classification.

1 Introduction

Stress detection is crucial in mental health and human-computer interaction research. EEG signals, being sensitive to physiological changes, provide valuable data for identifying stress states. However, effective stress classification requires a model that can capture different levels of features. This study proposes a hierarchical CNN feature fusion model, which extracts low, mid, and high-level features from EEG data to improve classification accuracy. This method offers a comprehensive approach to modeling EEG signals by combining features across levels for robust stress detection.

2 Methodology

2.1 Feature Extraction

Feature extraction was performed across three levels to capture the full spectrum of information in EEG data:

- **Low-level features:** Basic statistical metrics such as mean, variance, and skewness, capturing signal distribution characteristics.
- **Mid-level features:** Frequency-based features like band power and entropy, capturing energy distribution in specific frequency bands.
- **High-level features:** Complex metrics, including fractal dimension and Hurst exponent, capturing self-similarity and long-term dependencies in EEG signals.

2.2 Fusion Strategy

The hierarchical fusion strategy involves encoding EEG data through three separate uni-modal models for each feature level. The outputs from these models are then concatenated, forming a comprehensive feature representation. This layered fusion approach ensures that each feature level contributes unique information to the final model, enhancing its ability to classify stress accurately.

2.3 Classification

A deep learning classifier, built on TensorFlow/Keras, was trained on the fused feature set to classify participants as stressed or non-stressed. This model, leveraging the multi-level feature fusion, is designed to capture and utilize both low-level signal characteristics and higher-level abstractions, making it highly effective in stress classification.

3 Dataset

The dataset consists of EEG signals recorded from 58 participants, each viewing 36 different video clips intended to elicit varying stress responses. EEG data was collected through 8 channels, providing a detailed physiological response to each video. Additionally, participants provided self-reported ratings on their perceived stress, which served as the ground truth for training the model.

4 Stress Classification

Stress classification was based on both the self-reports and the fused features extracted from EEG data. A binary classification approach was used, labeling participants as stressed or non-stressed. Classification was done through both conventional deep learning method and the proposed fusion method. The accuracy in the later was enhanced significantly, allowing for a more reliable stress detection system.

5 Results

The hierarchical CNN-based fusion model achieved higher accuracy compared to traditional single-level feature models. The conventional model showed an accuracy of 92-94%, while the proposed fusion model achieved an accuracy of 96-98%, demonstrating the effectiveness of combining features across multiple abstraction levels. This improvement of 2-4% suggests that hierarchical feature fusion enables a more comprehensive understanding of EEG data, resulting in better stress classification.

Method	Accuracy
	ASCERT
Proposed	97.6%

Figure 1: Snippet from the paper showing accuracy of the proposed model

method showed its effectiveness by outperforming existing models by 1-2%, respectively, on frequency band features. It is observed that the hierarchical feature set from all three levels performed better than all other combinations by 2-4%. As a result, this strategy can be a useful addition to stress detection.

Figure 2: Snippet from the paper determining increase in accuracy

```
Epoch 11/11
23/23 ━━━━━━━━━━━━━━━━━━━━ 0s 4ms/step - binary_accuracy: 0.9987 - loss: 0.0278 - val_binary_accuracy: 0.9448 - val_loss: 0.1499
8/8 ━━━━━━━━━━━━━━━━━━━━ 0s 3ms/step - binary_accuracy: 0.9197 - loss: 3.7919
Test Accuracy: 93.36%
```

Figure 3: Accuracy of the conventional model in the code

```
Epoch 11/11
23/23 ━━━━━━━━━━━━━━━━━━━━ 0s 4ms/step - binary_accuracy: 0.9978 - loss: 0.1048 - val_binary_accuracy: 0.9669 - val_loss: 0.1709
8/8 ━━━━━━━━━━━━━━━━━━━━ 0s 3ms/step - binary_accuracy: 0.9673 - loss: 0.1912
Test Accuracy: 96.02%
```

Figure 4: Accuracy of the proposed fusion model in the code

6 Conclusion

This study highlights the potential of hierarchical feature fusion in stress detection using EEG data. By combining features across multiple levels, our model offers a robust framework for classifying stress states, achieving higher accuracy than traditional methods. This approach not only improves classification performance but also lays the groundwork for future research in multimodal feature fusion for physiological data analysis.

References

- Radhika Kuttala, Ramanathan Subramanian, Venkata Ramana Murthy Oruganti, *Multimodal Hierarchical CNN Feature Fusion for Stress Detection.*