

Discrete Emotion Classification Based on Skin Conductance: A Comparative Study of Machine Learning Models Across Contexts

Xinyuan Ye *

Department of Biomedical Engineering, Southern University of Science and Technology,
Shenzhen, China

* Corresponding author: 12313455@mail.sustech.edu.cn

Abstract. With the rapid development of the field of affective computing, the accurate identification and classification of human emotions have become increasingly important. This study aims to classify neutral emotion (neu) and other emotions using peripheral physiological signals such as Skin Conductance Response (SCR), Skin Conductance Level (SCL), and Instantaneous Skin Conductance Response (iSCR) through Support Vector Machine (SVM), Multilayer Perceptron (MLP), and Convolutional Neural Network (CNN) models. The study compares the classification effects under different emotion induction contexts (imaginary and video contexts) and evaluates the generalization ability across these contexts. A total of 43 subjects were included in the study, and the accuracy of emotion recognition and the generalization of models were explored by analyzing the physiological signal characteristics in different emotional states. By comparing the classification effect in different emotion-inducing situations, the research helps to find more accurate emotion recognition models, enhance the generalization ability of models, and promote the development and application of emotion computing technology.

Keywords: Emotion classification, Skin conductance, Support Vector Machine (SVM), Multilayer Perceptron (MLP), Convolutional Neural Network (CNN).

1. Introduction

In the field of affective computing, the accurate identification and classification of human emotions are crucial for enhancing the naturalness and intimacy of human-computer interaction. This study utilizes peripheral physiological signals such as Skin Conductance Response (SCR), Skin Conductance Level (SCL), and Instantaneous Skin Conductance Response (iSCR) to explore the performance of Support Vector Machine (SVM), deep learning models (including Gated Recurrent Units GRU and Transformer), and time series analysis methods in the task of classifying discrete emotions pairwise. It compares the classification effects under different emotion induction contexts (imaginary and video contexts) and the generalization ability across contexts. The study involved 43 subjects and analyzed the characteristics of physiological signals in different emotional states to explore the accuracy of emotion recognition and the generalization of models.

2. Dataset Overview

In this study, the data used is derived from the EmoEEG-MC dataset [1], a multi-context emotional dataset specifically designed for cross-context emotion decoding. The dataset includes 64-channel EEG data from 60 participants under two different emotion-inducing scenarios, video-induced and imagination-induced, as well as peripheral physiological data from 43 subjects, including skin conductance response (SCR), skin conductance level (SCL), and instantaneous skin conductance response (iSCR). Notably, our focus is on SCR, SCL, and iSCR, which were collected by wristband devices at sampling rates of 40Hz and 20Hz. Participants experienced seven different emotional

categories, including joy, inspiration, tenderness, fear, disgust, sadness, and neutrality, with three trials for each emotion in each context. The construction of the dataset was reviewed and approved by the Southern University of Science and Technology Medical Ethics Committee (Project Number: 20220174), and informed consent was obtained from the participants. This study only used the peripheral physiological signal data from the dataset, extracting a feature sample every three seconds to explore the physiological response patterns under different emotional states and to assess the potential of these signals in emotion classification. All related data processing, feature extraction, and emotion classification analysis were conducted based on this foundation.

3. Related Work

In the field of emotion recognition, research based on physiological signals is gradually deepening, especially in the study of skin conductance (GSR) signals. This section reviews five studies that explore how to use GSR signals to identify and analyze emotional states from different perspectives.

Firstly, the first study is based on GSR signals and is dedicated to identifying six basic emotions [2]: calm, happy, surprised, angry, disgusted, and sad. Researchers recruited first-year students from Southwest University as subjects, collected GSR signals by having them watch emotionally stimulating videos, and established a database of emotional physiological signals. The signals were preprocessed using a Butterworth filter, and 30-dimensional statistical features were extracted. The study employed particle swarm optimization algorithms and their variants, combined with Fisher classifiers, to achieve feature selection and emotional state classification recognition of GSR signals, achieving good classification results and verifying the feasibility of establishing an emotion recognition model. The study results indicate that GSR signals are most effective in recognizing surprise and fear emotions, while the recognition effect for disgust emotions is the worst. In addition, the application of immune particle swarm optimization algorithms and simulated annealing immune particle swarm optimization algorithms has improved the recognition rate of emotional states and reduced the dimensionality of the feature subset.

The second study aims to improve the accuracy and generalization ability of dimensional model physiological signal emotion recognition [3]. Based on the DEAP dimensional emotional physiological dataset, researchers extracted time-domain statistical features, power spectrum features, and wavelet packet entropy features of skin electrical signals and used decision trees and random forest algorithms for emotion classification. By selecting appropriate emotional label thresholds, the accuracy and robustness of machine emotion recognition were improved. The normalization processing of the difference between skin electrical signals in emotional induction state and natural state was used to eliminate individual differences and improve the generalization ability of the model. In the two emotional dimensions of arousal and valence, the classification accuracy and F1 value both performed excellently.

The third study proposes an emotion recognition and regulation method based on skin electrical signals [4], which is of great significance for the realization of human-computer emotional interaction systems. By analyzing skin electrical signals, MLP neural networks optimized by genetic algorithms were used to identify emotional categories, and curve fitting methods were used to evaluate emotional intensity, determining the real-time emotional state of users. According to the user's emotional state, when the user needs emotional regulation, the computer can automatically select appropriate materials from the emotional regulation material library to present to the user for emotional regulation, achieving human-computer emotional interaction.

The fourth study focuses on the time-encoded images of electrodermal activity (EDA) segments to identify significant patterns for an emotion recognition system [5]. Researchers obtained EDA signals from two publicly accessible datasets, namely, Continuously Annotated Signals of Emotions (CASE) and Wearable Stress and Affect Detection (WESAD). These signals were preprocessed and decomposed into phasic signals through a convex optimization approach. Subsequently, time-encoded image representations of these windowed phasic signals were generated using a Gramian

Angular Summation Field, Markov Transition Field, Recurrence Plot (RP), and a fusion of these images. A variety of machine learning models were developed to classify two-class emotions associated with arousal and valence from CASE, as well as three-class emotions (amusement, neutral, and stress) from WESAD, considering three different approaches: the first half, the second half, and the whole phasic signal. The study achieved high classification accuracy, highlighting the potential of the proposed methodology for analyzing emotions in healthcare, with the ability to accurately classify emotions holding promising implications for improving patient care, mental health assessment, and overall well-being.

Finally, the fifth study optimized the electrodermal activity (EDA) segment for an emotion recognition system [6]. Researchers obtained EDA signals from two publicly available datasets: the Continuously Annotated Signals of Emotion (CASE) and Wearable Stress and Affect Detection (WESAD) for 4-class dimensional and three-class categorical emotional classification, respectively. These signals were preprocessed and decomposed into phasic signals using the 'convex optimization to EDA' method. Further, the phasic signals were segmented into two equal parts, each subsequently segmented into five non-overlapping windows. Spectrograms were then generated using short-time Fourier transform and Mel-frequency cepstrum for each window, from which 85 features were extracted. Four machine learning models were built for the first part, second part, and whole phasic signals to investigate their performance in emotion recognition. The study demonstrated superior performance in three-class emotions classification, attaining high accuracy for both whole phasic and second part phasic segments. As a result, the second part of EDA is strongly recommended for optimal outcomes.

In summary, these studies demonstrate the progress of GSR signal-based emotion recognition technology in terms of accuracy, generalization ability, and real-time performance, providing new directions and methods for future research and application of emotion recognition technology.

4. Experiment

4.1. Emotion Classification Models

This study employs three advanced machine learning and deep learning models for emotion classification research based on physiological signals: Support Vector Machine (SVM), Multilayer Perceptron (MLP), and Convolutional Neural Network (CNN). These models are widely used in various pattern recognition and classification tasks due to their unique advantages and applicability. Below is a brief introduction to these models and their application in emotion classification tasks.

4.1.1. Support Vector Machine (SVM)

Support Vector Machine is a supervised learning algorithm primarily used for classification problems. Its core idea is to find an optimal hyperplane in the feature space to distinguish data points of different categories and maximize the boundary (i.e., margin) between the two categories. SVM is particularly suitable for dealing with small samples and high-dimensional data, and for nonlinear problems, it can be processed through kernel tricks (such as the Radial Basis Function kernel RBF). In emotion classification, SVM can identify subtle differences in physiological signals, providing an effective means for accurate classification of emotional states.

4.1.2. Multilayer Perceptron (MLP)

The Multilayer Perceptron (MLP) is a type of neural network with multiple layers, including input, hidden, and output layers. It uses backpropagation for training, adjusting weights to minimize error. MLPs are adept at capturing complex patterns, making them suitable for emotion classification from physiological signals. They can be tuned by changing the number of layers and neurons to better fit the data. Activation functions like ReLU help introduce non-linearity, while techniques such as dropout can prevent overfitting. MLPs are flexible and can learn to classify emotions effectively from subtle physiological changes.

4.1.3. Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is a deep learning model particularly suitable for processing data with grid-like topological structures, such as images. It extracts local features through convolutional layers and reduces the spatial dimension of features through pooling layers, thereby reducing computational load. CNN has performed exceptionally well in image recognition, speech recognition, and other fields, capable of automatically learning hierarchical features of data. In emotion classification based on physiological signals, CNN can effectively extract the temporal features of physiological signals, providing a powerful feature representation for the recognition of emotional states.

4.2. Feature Selection

Feature selection plays a pivotal role in emotion recognition models, as it can significantly impact the classification performance by identifying the most relevant and informative features from a set of physiological signals. In this study, we focus on three key features: Skin Conductance Response (SCR), Skin Conductance Level (SCL), and Incremental Skin Conductance Response (iSCR), which are closely associated with emotional states.

SCR, a measure of rapid changes in skin conductance, is indicative of emotional arousal and valence. It is a sensitive marker for emotional responses, particularly in the context of stress and excitement [7]. SCL, on the other hand, represents the baseline level of skin conductance and is associated with the intensity and duration of emotional states. It provides a stable measure that can be used to differentiate between various emotional states [8]. iSCR refers to the incremental changes in skin conductance over specific time intervals, offering dynamic information about emotional fluctuations.

The selection of features is crucial for the classification performance of emotion recognition models. Features are considered statistically significant if they exhibit substantial differences across various emotional states, which can enhance the model's ability to distinguish between classes [9]. Information gain, a measure of how much uncertainty is reduced by a feature, is another critical factor in feature selection. Features with high information gain provide more insight into the classification problem, thereby improving the model's predictive power [10].

In our analysis, we employed both statistical significance and information gain to evaluate the relevance of SCR, SCL, and iSCR features. We found that SCR's statistical features, particularly its standard deviation or variance and the derivative of the SCR signal, performed better in the time domain compared to other SCR-related features [7]. This suggests that these features are more sensitive to changes in emotional states and thus are more informative for classification purposes. SCL and iSCR also showed promise in distinguishing between emotional states, with SCL providing a stable baseline measure and iSCR capturing the dynamics of emotional responses [8].

In conclusion, feature selection based on statistical significance and information gain is essential for optimizing the performance of emotion recognition models. SCR, SCL, and iSCR features, when selected appropriately, can significantly enhance the accuracy and efficiency of emotion classification.

4.3. Experimental Design

Initially, we segmented the 43 subjects from both imaginary (ima) and video (vid) scenarios into training and testing sets, allocating 20% for the testing set. We then normalized the training and testing sets to boost the efficiency and performance of model training. After that, we carved out a 25% validation set from the normalized training data. Ultimately, we applied three different models—Support Vector Machine (SVM), Multilayer Perceptron (MLP), and Convolutional Neural Network (CNN)—to analyze the 43 subjects from both imaginary (ima) and video (vid) scenarios. The models were designed to classify neutral and six additional emotional states based on three physiological indicators: Skin Conductance Response (SCR), Skin Conductance Level (SCL), and Instantaneous Skin Conductance Response (iSCR).

4.4. Experimental Results

Present the emotion classification accuracy of different models under different contexts. Compare the classification effects between imaginary and video contexts. The six rows of this table display the accuracy rates of the models for neutral emotion and the following six emotions: sadness, disgust, fear, joy, tenderness, and inspiration. Each row corresponds to an emotion, indicating the accuracy rate of the respective emotion recognition model in binary classification tasks.

Table 1. Emotion Recognition Accuracy under Imaginary Context by SVM, MLP, and CNN Models

	Emotion	SVM	MLP	CNN	avg
1	Sadness	0.58	0.55	0.54	0.557
2	Disgust	0.52	0.55	0.54	0.537
3	Fear	0.54	0.54	0.54	0.540
4	Joy	0.59	0.61	0.59	0.597
5	Tenderness	0.56	0.57	0.57	0.567
6	Inspiration	0.51	0.52	0.57	0.533
	avg	0.550	0.557	0.558	
	STD	0.032	0.030	0.021	

Table 2. Emotion Recognition Accuracy under Video Context by SVM, MLP, and CNN Models

	Emotion	SVM	MLP	CNN	avg
1	Sadness	0.61	0.60	0.57	0.593
2	Disgust	0.59	0.59	0.59	0.590
3	Fear	0.60	0.57	0.61	0.593
4	Joy	0.62	0.53	0.61	0.587
5	Tenderness	0.62	0.62	0.58	0.607
6	Inspiration	0.64	0.63	0.62	0.630
	avg	0.615	0.590	0.597	
	STD	0.018	0.036	0.020	

4.5. Analyse and Discussion

By analyzing the data in Table1 and table 2, we can discuss the performance of the models in different contexts from two dimensions: the average accuracy and variance of the model accuracy rates, and the discrimination between neutral emotions and other emotions.

4.5.1. Analysis of model accuracy, averages and variance

In the imaginary context (ima context), the CNN model has the highest average accuracy, reaching 0.558, and at the same time, the smallest variance, indicating that the CNN model not only performs the best overall in this context but also has the highest consistency and stability of results. This may mean that the CNN model can more effectively capture and utilize emotional features when dealing with emotion recognition tasks in the imaginary context, thus achieving higher accuracy and smaller fluctuations.

In contrast, in the video context (vid context), the SVM model has the highest average accuracy, 0.615, and the smallest variance, showing that the SVM model has the best overall performance and result stability in the video context. This may be related to the SVM model's ability to better identify and distinguish emotional features in video data.

4.5.2. Analysis of discrimination between neutral and other emotions

In the imaginary context, among the discrimination between neutral emotions and other emotions, the accuracy of joy emotion recognition is the highest, reaching 0.597, followed by tenderness, with an accuracy of 0.567. This may indicate that in the imaginary context, the model has a stronger ability

to recognize joy and tenderness emotions, and can more accurately distinguish between neutral emotions and these two emotions.

In the video context, among the discrimination between neutral emotions and other emotions, the accuracy of inspiration emotion recognition is the highest, reaching 0.630, followed by tenderness, with an accuracy of 0.607. This indicates that in the video context, the model has a stronger ability to recognize inspiration and tenderness emotions, and can more accurately distinguish between neutral emotions and these two emotions.

5. Conclusion

From the perspective of model accuracy averages and variance, the CNN model performs the most stably and accurately in the imaginary context, while the SVM model performs the most stably and accurately in the video context.

From the perspective of emotion recognition discrimination, whether in the imaginary context or the video context, the accuracy of tenderness emotion recognition is ranked second, which may indicate that the features of tenderness emotions are more apparent in both contexts and are easily captured and recognized by the model.

In general, the selection and optimization of models need to consider specific application contexts and emotion types. In practical applications, the most suitable model can be selected based on the model's performance in different contexts to improve the accuracy and discrimination of emotion recognition. At the same time, the analysis of the discrimination between neutral emotions and other emotions also provides valuable information about the difficulty of recognizing emotional features, which helps to further optimize and adjust the model.

Acknowledgement

The author gratefully acknowledges the support from the Neural Computing and Control Lab (NCC Lab) at the Southern University of Science and Technology.

References

- [1] XU Xin, SHEN Xinke, CHEN Xuyang, ZHANG Qingzhu, WANG Sitian, LI Yihan, LI Zongsheng, ZHANG Dan, ZHANG Mingming, LIU Quanying. EmoEEG-MC: A Multi-Context Emotional EEG Dataset for Cross-Context Emotion Decoding. ScienceDB, 2024.
- [2] Zhou Yuting. Research on Emotional Feature Extraction and Classification Recognition of Skin Conductance Signals. Southwest University, 2024.
- [3] He Guozhu, Qiao Xiaoyan. Emotion Recognition Based on Tree Model Machine Learning of Skin Conductance Signals. Journal of Test Technology, 2021, 35 (6): 508 - 514528.
- [4] Liu, F. Research on emotion recognition and regulation based on skin conductance signals [Dissertation]. Chongqing: Southwest University, 2015.
- [5] P. Sriram Kumar and J. F. A. Ronickom, Emotion Classification through Optimal Segments of EDA and Texture Analysis of Time-Encoded Images with Artificial Intelligence, in IEEE Transactions on Instrumentation and Measurement, 2024.
- [6] Kumar, P. S., & Ronickom, J. F. A. Optimal Electrodermal Activity Segment for Enhanced Emotion Recognition Using Spectrogram-Based Feature Extraction and Machine Learning. International Journal of Neural Systems, 2024, 34 (5), 2450027.
- [7] Research on Skin Electrophysiology: Feature Extraction and Selection for Emotion Recognition using EDA Sensors. CSDN Blog. <https://blog.csdn.net/cc1609130201/article/details/127183390>. Accessed on 2022 - 10 - 06.
- [8] Chen Shali, Zhang Liuyi, Jiang Feng, et al. Research on Emotion Recognition Based on Multiple Physiological Signals. Chinese Journal of Medical Devices, 2020, 44 (4): 283 - 287.

- [9] Nie, D., Wang, X., Duan, R., & Lv, B. (2012). A Survey on EEG Based Emotion Recognition. Chinese Journal of Biomedical Engineering, 2012, 31 (4), 595 - 606.
- [10] Wang Xiaolin, Zhao Lei, Zhang Wei, Fu Yun. Emotion Feature Extraction and Recognition Based on HHT and Kernel Function Selection. Journal of Nanjing University (Natural Sciences), 2021, 57 (3): 502 - 511.