Development of Emotion Stress Relief Recommender Tool using Machine Learning

M. Mohammed Mustafa
Sri Krishna College of Engineering & Technology, Coimbatore, Tamilnadu, INDIA

Article Info

Article history:
Received: April 5, 2024
Revised: June 18, 2024
Accepted: June 28, 2024

Keywords:
API;  
Face Detection;  
Image Processing;  
Machine Learning;  
Stress.

Abstract

In today's society, individuals face unprecedented levels of stress, with varying indicators and effects. Music has long been recognized for its calming effects on the brain and body, especially slow and calm classical music, which can lower heart rate, blood pressure, and stress hormone levels. However, the effectiveness of music in stress management may be reduced if the music does not match the listener's current emotional state. To overcome this problem, this research aims to develop an emotion-based stress reduction recommendation tool using machine learning. This research employs experimental methods. The research results show this tool will analyze the user's emotions and recommend music most likely to reduce stress. By tailoring music recommendations to an individual's emotional state, we hope to increase the overall effectiveness of music as a stress management tool. Implications of this research include the potential for creating intelligent music players that provide more personalized and effective stress reduction.


INTRODUCTION

Nowadays, the number of people who experience stress has risen significantly due to obligations, better object value, a lousy economy, excessive costs, and so on [1]. In 2017, the Department of Emotional Wellbeing gathered information from telephone management for mental wellbeing publications and discovered that Thai people's stress was increasing. The number of calls reached 30,000, which might be doubled in 2019.

There are several programs available to track participants and remedy stress. However, no software could select melodies, movies, and books depending on the purchasers' emotions [2], [3]. To address this limitation, this flexible tracking software that could prescribe melodies depending on the purchasers' emotions may be utilized [4]. Initially, the software snaps a picture of the purchaser [5]. When the software receives the purchaser's face image from the flexible camera, it examines the purchaser's emotions (dismal, cheerful, livid, or impartial) [6]. To identify the purchaser's emotion from the face image, the software makes use of the Google Vision API [7].

Utilizing Google Cloud, the sensation is broken down from the face photograph. Google's Emotion acknowledgments API verifies the sensation of the customers depending on their face photograph on the software [8][9]. After getting the emotion subtleties from the Google API, the software offers numerous guidelines to decrease the stress degree to the customer. Based on the customer's emotions, the software will advise books, songs, movies, and so forth.

- Corresponding author:
  M. Mohammed Mustafa, Sri Krishna College of Engineering & Technology, Department of Information Technology, Coimbatore, Tamilnadu, INDIA. mohammedmustafa@skcet.ac.in
  © 2024 The Author(s). Open Access. This article is under the CC BY-SA license [https://creativecommons.org/licenses/by-sa/4.0/].
Detection of Image Properties

This is a conventional instance of computer vision that offers the extraction of textual content from a photo. The Vision API accommodates many state-of-the-art processes [10]. With Vision API, you may retrieve a photo's attributes and functions, including the dominant color.

Face Detection

The task is to identify the faces in a picture or a hard copy image. This includes many large-scale programs, such as Surveillance Systems [11]. These are some of the best use cases for Vision API, and you can incorporate any of them into your programs in a significantly shorter amount of time. Numerous programming languages, including Go, C#, Java, PHP, Node.js, Python, and Ruby, are supported by Vision API. The subsequent sections will demonstrate how to use the Vision API in Python.

Pattern Matching

Identifying styles using a device-learning algorithm is known as pattern popularity [12]. Pattern recognition is data classification based on existing knowledge or statistical information derived from patterns and their representation. The utility potential is a key factor in determining the recognition of the sample. Pattern recognition involves the classification and grouping of patterns.

Existing Work

Electroencephalogram (EEG) signal-based emotion recognition has recently gained a lot of attention and has been widely used in clinical, affective computing, and other important fields. In any case, most research has focused on the accuracy of relationships while neglecting the understandability of emotional expression. Currently, we are advocating for using AI and EEG signals to enhance the recognition of various interpretable emotions. This paper introduces a new concept called the enthusiastic actuation bend to explain how emotions are triggered [13]. The calculation initially focuses on extracting key features from EEG signals and analyzing the emotions of individuals using AI techniques. Different parts of a dataset are then used to create a proposed model and evaluate its effectiveness in recognizing emotions. Secondly, we construct new ways of experiencing emotions based on the results of associations, and these ways are determined by emotion coefficients, which are the coefficients that represent the strength of relationships and the level of uncertainty. The initiation bend can evoke emotions and somewhat reveal the emotional motivation system. Finally, we receive a weight coefficient from the two coefficients to improve emotion recognition accuracy. To evaluate the proposed technique, assessments were conducted using the DEAP and SEED datasets [14], [15]. The findings support the idea that emotions are influenced by logical reasoning during the study, and the weighting coefficients, which depend on the relationship coefficient and the entropy coefficient, can greatly improve the accuracy of EEG-based emotion recognition [16], [17].

Let us discuss the problem of most traditional exams lacking an understanding of emotional stimulation. To address this issue, we suggest using a coefficients-based approach that relies on AI and EEG signals. This method not only surpasses the benchmark calculations in terms of accuracy but also understands the process of emotional initiation.

At first, we removed the important parts from EEG signals [Figure 1] and organized emotions using artificial intelligence methods. We also found that the last portion of EEG signals has stronger connections with emotions, so using the second half of the initial

Figure 1. The EEG Signals
signals for training can lead to better classifier performance [Figure 2]. Furthermore, considering the results of the grouping, the relationship curves and information curves of emotions are formed, which, to some extent, indicate the emotional activation process. It has been observed that emotions are constantly being experienced. The proposed method provides a more concrete understanding of the emotional motivation system. For instance, it helps explain why the latter half of an initial activation leads to better overall results. In the end, we apply the obtained connection and entropy coefficients to construct weight coefficients to improve the association’s accuracy compared to modern benchmark algorithms. Considering that the burden coefficient depends on the connection coefficient and entropy coefficient, the validity of the burden coefficient also implies the validity of the proposed connection coefficient and entropy coefficient from another perspective. This indicates that the proposed hypothesis of dynamic regulation of emotions is reasonable.

![Figure 2. States of the Human Stress Behaviour](image)

Measuring the EEG signals is complicated because:

- Recognizing the range of signals for each emotion can be quite complex.
- This system requires many hardware components, making it not very compact.
- The initial cost of implementing this module is quite expensive.
- The learning algorithm can make the system more complex and difficult to maintain.

Previous studies have focused on using machine learning to predict individuals’ emotions based on contextual information [18], analyze emotions to forecast emotional labels [19], use emotion detection from speech [20]–[22], and recognize emotions through various multi-modal data, such as text and facial expressions [23]–[25]. However, these studies have not yet extended to developing systems that predict emotions and provide recommendations, such as books or music, to alleviate stress for individuals whose emotions are predicted. Therefore, this research aims to complement the shortcomings of previous studies by providing additional features, such as recommendations for reading materials or music to manage stress associated with predicted emotions.

**METHODS**

This research employs experimental methods [26] to assess the efficacy of utilizing machine learning. Through controlled experiments and systematic data collection, the research aims to evaluate the performance and accuracy of machine learning algorithms in predicting emotional states based on facial expressions. By comparing the predictions made by machine learning models with real-world emotional responses, the research seeks to determine the effectiveness and reliability of these algorithms in capturing and interpreting human emotions. Through rigorous experimentation and analysis, this study contributes to advancing our understanding of the capabilities and limitations of machine learning in emotion recognition. It provides valuable insights for the development of emotion-aware systems and applications. Figure 2 shows the working stages of machine learning.
The tool captures the user's facial image using a flexible camera. At that point, the facial image is uploaded to the image processing application and sent to the Google Vision cloud for image recognition. The face image is processed using the VISION API, which utilizes the organized informational collections of the Google Vision Cloud to interpret the facial image's emotions. Then, it will determine the current emotion of the user, whether they are sad, happy, neutral, or angry, and send it back to the application. The administrator manages the music, movies, books, and information collections within the Firebase Database.

Figure 3. The Stages of the Machine Learning Tool
Furthermore, depending on the diagnosed emotions of the users, the utility shows the music, books, and movies [Figure 4]. This process will help the purchaser with lowering his/her stress.

- Image recognition Module
- Emotion location (Vision API) Module
- Songs and Books Dataset Module
- User Interface Module

**Image Recognition Module**

The software receives a facial photograph of the users from a flexible camera or the Gallery. The facial image is uploaded to the software for image processing. It is sent to the Google Cloud for image recognition. The Vision API includes modules that can perform face and emotion recognition. The face recognition module retrieves the photo from the user's device. Once uploaded to Google's server, it analyzes the facial features in the uploaded photo. The Vision API can detect at least one human face and analyze various characteristics, such as age, emotion, gender, posture, smile, and facial hair. It can also identify 27 markers on each face in the image.
Emotion Detection (Vision API) Module
Among the properties, such as age, emotion, sexual orientation, posture, smile, and facial hair, the trait necessary for the application (i.e., emotion) is collected by making an API call to the Google server. The API server can provide eight emotions: angry, contemptuous, disgusted, fearful, happy, neutral, sad, and surprised. The eight styles of emotions have been condensed into four categories (Sad, Happy, Angry, and Neutral) to enhance clarity and simplify understanding. The programming interface will bring back the subtle nuances of emotions to the application, allowing the user to experience a more authentic emotional response.

Training Set
The training set DEAP is a tool for assembling a model. The set of images can be used to teach the tool. The training regulations and algorithms used provide useful information on combining input data with output decisions. The device uses professional methods to apply algorithms to the dataset, extracting relevant information and obtaining results. Figure 6 Typically, about 80% of the data in the dataset is used for training purposes.

Figure 6. Dataset Classification

Figure 7. Represents the API with the Keys
Music and Books Dataset & User Interface Module

Music, movies, books, and informational indexes are stored within the Firebase Database. The music, movies, and books are organized into categories associated with various emotions. Each page within the application is linked to a specific emotion [27]. A collection of songs, movies, and books are linked to their respective media within the application. The application includes different actions for each emotion, and the landing page is designed to gather the image from the user's device and then send the image to the Google server. After receiving the sensory information from API [Figure 17], a web page that aligns with their emotions is shown to the user. This page includes stress relief tips, such as recommended books and music that match the user's current mood.

![Image of Firebase interface](https://via.placeholder.com/150)

**Figure 8. The Accuracy of Using a Real-time Database**

Vision API plays out the AI calculations utilizing a prepared informational index, which makes the framework straightforward and effective.

- The accuracy of the emotion acknowledgment utilizing VISION API is 0.92199 (92.19%) [Table 1]
- This framework does not require some other external tools except cell phones.
- The framework joins different emotions into significant emotions (sad, happy, angry, and neutral) to give better proficiency.
- The framework interface and support are easy to comprehend by the client.

RESULTS AND DISCUSSION

The various emotions of stress are represented using various techniques and their accuracy can be seen in Figures 9 – 16

![Bar chart showing comparison of techniques](https://via.placeholder.com/150)

**Figure 9. The Comparison of Anger with Different Techniques**
Figure 10. The Comparison of Disgust with Different Techniques

Figure 11. The Comparison of Fear with Different Techniques

Figure 12. The Comparison of Happy with Different Technique

Figure 13. The Comparison of Neutral with Different Techniques

Figure 14. The Comparison of Surprise with Different Techniques

Figure 15. The Comparison of Sad with Different Techniques
Figure 16. The Comparison of Accuracy with Different Techniques

The results show that the various techniques play a significant role in the various features of human stress.

Table 1. The Accuracy of Various Techniques

<table>
<thead>
<tr>
<th>Technique</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feed-Forward Neural Network</td>
<td>82%</td>
</tr>
<tr>
<td>PCA and SVM</td>
<td>92%</td>
</tr>
<tr>
<td>Neural network</td>
<td>90%</td>
</tr>
<tr>
<td>Convolutional Neural Networks (CNN)</td>
<td>60%</td>
</tr>
<tr>
<td>k-Nearest Neighbour Algorithm</td>
<td>84%</td>
</tr>
<tr>
<td>Support Vector Machine (SVM)</td>
<td>91%</td>
</tr>
<tr>
<td>Relevance Vector Machines (RVM)</td>
<td>90.84%</td>
</tr>
<tr>
<td>Vision API</td>
<td>92.19%</td>
</tr>
</tbody>
</table>

Figure 17. The Developed Application

Figure 18. Comparison of the Accuracy Using Various Techniques
After analyzing the data from Table 1, it is evident that the classification techniques exhibit differences in accuracy. Principal Component Analysis (PCA) and Support Vector Machines (SVM), along with the Vision API, achieved the highest accuracy rates of 92% and 92.19%, respectively. Support Vector Machines (SVM) performed exceptionally well, achieving an accuracy of 91%, which aligns with the research findings. Relevance Vector Machines (RVM) and Neural Networks also demonstrated strong performance, with 90.84% and 90% accuracy, respectively. The Feed-Forward Neural Network and k-nearest Neighbor Algorithm have impressive accuracy rates of 82% and 84%, respectively, which is quite high compared to other research findings. However, Convolutional Neural Networks (CNN) tend to perform less, achieving only 60% accuracy. Based on the analysis conducted in this study, we can conclude that Principal Component Analysis (PCA), Support Vector Machines (SVM), and the Vision API are the most accurate techniques for classification in this particular context.

The results of the CNN in this study showed a minimum accuracy of 60%. This is consistent with previous research [28], which also found results of 66% that tended to be on the lower side. The k-nearest Neighbor Algorithm network has a fairly high accuracy of 84%, which differs from the findings of a previous study [29] that reported a k-nearest Neighbor Algorithm value of 75%. In this research, the SCM model demonstrated a high level of accuracy, specifically 90%, reaching into the 90’s. This also happened with previous research. In one study, SVM achieved a 95% accuracy rate [30] Another study conducted by [31] and [32] also reached 96%.

Several main factors can cause variations in precision between different techniques in data classification in this study. Firstly, it is important to consider that the type of data and features being used can impact how effective a specific method is. Methods like Principal Component Analysis (PCA) and Support Vector Machines (SVM), along with the Vision API, are more effective in identifying and utilizing significant characteristics of the dataset. This leads to improved precision, as mentioned in reference [33]. Additionally, the complexity of the model is a significant factor. Convolutional Neural Networks (CNN) are more complex models that may need more data to achieve the best possible performance [34].

Furthermore, it is important to note that algorithm performance greatly depends on proper hyperparameter tuning. Some machine learning models, such as Neural Networks and SVMs, can be well optimized, while CNNs may not be as easily optimized. Another factor that affects performance is the amount of training data. Techniques like CNNs, which require a large amount of data, may not perform well when limited data is available [34]. Lastly, it is important to consider the impact of non-linear features and the chosen architecture and implementation models. Models like the Vision API are designed with advanced architectures and various training techniques, which leads to exceptional performance.

CONCLUSION

The main idea is to present stress relief suggestions to clients by tapping into their emotions and experiences. The framework utilizes the sensation acknowledgment API, which enhances its specificity and simplicity at the same time. Furthermore, we estimated the accuracy and compared it with several other techniques, finding it 92.19%. Figure 18 Instead of using the cutting-edge framework, which uses a complex method of utilizing EEG to detect emotions, we can use the image processing APIs to perform this task. It can be challenging to achieve those goals when using various techniques to understand the human experience, which requires a significant amount of data and makes it more complicated. This tool enhances the accuracy and maintainability of the process, providing greater comfort to the user.

REFERENCES


