REVIEW AND ANALYSIS ON DEEP LEARNING BASED APPROACH ON ARTIFICIAL EMOTIONAL INTELLIGENCE

Harikrushnareddy vangala, Master of information technology, Faculty of Information Technology, Monash University, Clayton

ABSTRACT

The goal of this study is to investigate how authors’ ability to recognize and convey emotions through words and expressions changes in an online, adaptive learning environment. The effectiveness of a method developed to improve a convolution neural network’s performance using Bidirectional Long Short-Term Memory is verified through a simulated exercise (CNN-BiLSTM). It was decided to do this after carefully examining the deep learning neural network methods already available. Based on the experimental results, the CNN-BiLSTM method provided in this article achieves an accuracy of up to 98.75%, which is at least 3.15 percentage points more than the accuracy of other algorithms. The book “Deep Learning-Based Artificial Emotional Intelligence” delves into the ways in which AI has been used to emotional intelligence, specifically to the goals of accelerating and augmenting emotional intelligence. Some examples of artificial intelligence applications that are investigated include machine learning and deep learning. It offers research on tools that may be utilized by system architects and designers to ease and streamline the process of developing deep learning systems. Specifically, we employ a deep learning approach for labelling feelings in our research. In this method, a huge number of sensor inputs from different modalities are added and subtracted in an iterative process. Our dataset was derived from information collected by multiple smartphones and wearables in a real-world investigation. It does so by displaying signal dynamics and the temporal correlations between the various sensor modalities, all of which are gathered from a unified model that incorporates on-body, ambient, and location sensors. The raw sensor data is fed into numerous learning methods, including a hybrid that employs both convolutional neural networks and long short-term memory recurrent neural networks (CNN-LSTM). Thus, feature engineering and extraction can be carried out automatically.

Keywords: Deep learning, Emotion recognition, Convolutional neural network, Long short-term memory, mobile sensing

1. INTRODUCTION

The term "emotional recognition" is used to describe the method of deducing emotional content from more primitive data. Facial expressions are a reliable indicator of human emotion. It is possible to recognize and classify a wide range of human emotions by examining certain characteristics of these displays. Tone and volume are two linguistic characteristics. The way the body is carried, for instance, is a part of the expression sent by the body. Subtle changes in blood pressure and brainwave activity can also be indicators of how someone is feeling. The face is the most direct route from one's mind to the outside world via which emotions can be communicated. Facial emotion detection, also referred to as facial expression recognition (FER) has numerous potential applications. This tech is used in a wide variety of virtual reality (VR), interactive gaming (IG), robotics (Robots), and digital entertainment (DE) applications that involve human-computer interaction (HCI) (DE). It is used for medical purposes (autism, mental disorder, and pain evaluation), as well as in surveillance and law enforcement.
Feelings, thoughts, and actions of whole humans are all part of what we call "emotion," which is also the psychological and physiological reaction to a wide range of stimuli. Feelings play a significant role in our professional and personal lives. Accurately identifying emotions is crucial in many contexts. Emotion detection research has recently found widespread use in fields as diverse as medicine, computer vision, artificial intelligence, psychology, and neuroscience (Fürbass et al., 2020). For instance, diagnosing mental illnesses like depression and schizophrenia can benefit from emotion recognition technology. Helps medical professionals get in touch with their patients' genuine feelings. Humans and computers can connect more effectively when computers are able to recognise and respond to human emotions.

In recent years, current models for emotion recognition have been split into those that use physiological cues and those that don't. Physiological signals can reveal the true nature of human emotion since they are not influenced by bias or other subjective influences. As a result, there are significant benefits in terms of trustworthiness and usability when using physiological signals to identify emotions. Academics are currently preoccupied with the role of physiological signalling in current challenges. There are a number of physiological signals utilised in emotion recognition, including Electroencephalography (EEG), facial expression, Eye Movement (EM), and Electrocardiography (ECG). Based on these indicators of internal state, we can make accurate assessments of the participants' emotional states. The electroencephalogram (EEG) is a physiological signal that has been studied extensively because of its ability to provide an objective reflection of human emotional states and to occur spontaneously and without bias (Mohamed et al., 2018). As a result, studying how the brain processes emotions using EEG has emerged as a promising area of study. Niemic (2004) confirmed the importance of EEG to the study of emotions by showing how the activity of various brain areas was linked to various emotional states.

In order to better identify emotions based on an individual's brainwave patterns, emotion identification algorithms that use EEG data are primarily developing in two machine learning subfields: conventional and advanced learning. Traditional machine learning-based approaches to emotion recognition rely on human feature extraction for classifiers like Naive Bayes (NB), Support Vector Machine (SVM), and others to work. Deep learning techniques may make feature extraction much easier, which is why they are being increasingly used for emotion recognition. These techniques use Recurrent Neural Network and Long Short-Term Memory (LSTM) models to automatically learn deep information and recognise emotions (RNN). The standard machine learning approach to EEG emotion detection was introduced by Lin et al. (2018).
This technique entails a number of steps, such as identifying the emotional trigger, acquiring the EEG signal, extracting features, and recognizing the appropriate categorization. Additionally, the flaw in conventional machine learning techniques was shown, illuminating the way forward for EEG emotion recognition. It is challenging to discriminate EEG signals using a linear technique due to the fact that EEG signals are non-linear and high-dimensional. Deep learning, which is helpful for addressing non-linear problems, is what enables end-to-end mapping to become a reality. Craik et al. (2019) investigate (CNN). It was established that deep neural networks performed better than other methods for EEG classification.

For the purpose of EEG emotion recognition, network models such as convolutional neural networks (Sharma et al., 2020), adversarial networks (Luo, 2018), and others have seen increased application in recent years. Using a combination of keywords, we searched Elsevier and Springer for relevant articles and retrieved 645 research for this topic. After weeding out duplicates and inconsistent studies, 44 were chosen for further examination. The chosen terms include "EEG emotion recognition," "deep learning," "classification," "EEG emotion classification," "machine learning," and "emotion recognition" or "EEG feature extraction."

2. EMOTION RECOGNITION

In recent years, as more and more individuals have access to various electronic gadgets, more and more time has been spent on online activities like social media, gaming, and shopping. Unfortunately, the majority of today’s human-computer interaction (HCI) systems lack the emotional intelligence required to analyse and comprehend emotional data. They can’t read emotional cues from people or act on them in any meaningful way. Resolving the disconnect between people and robots is an essential part of developing next-generation intelligent HCI systems. Any HCI system that fails to take into account users' emotional states would be unable to provide a satisfactory user experience. Machines that can read and respond to human emotions could be the key to solving this problem in HCI systems. It is vital to have an emotion recognition system that is dependable, accurate, flexible, and powerful [Zhang J, 2020]. For the development of high-quality HCI systems, this is an essential first step.

The computer that drives an intelligent HCI system needs to be highly flexible so that it can accommodate the many different types of research that go into HCI. Human communication patterns need to be correctly understood in order to produce suitable replies. A computer's flexibility is greatly enhanced by its understanding of human emotions and behaviour. As a result, understanding the user's emotional state is crucial for improving the efficacy of HCI tools.

Having a computer that can accurately read the mood of its human operator in real time is a major step toward creating a human-computer interface that is both intelligent and easy to use. Affective computing describes this new field of study (AC). The field of artificial intelligence known as affect detection in human-computer interaction (AC). A primary focus in the field of affective computing (AC) [Calvo RA, 2010] is the creation of methodologies by which machines can comprehend and respond to human emotions.

Emotions can be read from a person's actions, words, facial expressions, and even physiological markers [Yin Z, 2017]. The first three methods are quite open to interpretation. It's possible, for instance, that the people under study are intentionally suppressing their true emotions in order to appear more competent. Physiological signals provide more accurate and objective emotional identification [Torres EP, 2020].

Non-invasive and portable sensor technologies called brain-computer interfaces (BCIs) have been developed to record brain signals and transmit that information to machines that can assess emotional states based on EEG changes [He Z, 2020]. Brain-generated EEG signals are more attuned to emotional shifts than their peripheral nervous system counterparts. There is also evidence that EEG signals provide useful information for identifying emotions.

3. VARIOUS PROPOSED METHODS

Studies using the DEAP, SEED, and DREAMER public datasets for EEG emotion identification are discussed. There are a total of 35 publications and experimental methodologies presented, with the great majority relying on deep learning for feature extraction.
and emotion recognition. Finally, a summary of how various approaches affect sentiment classification is provided.

3.1 Emotion Recognition Based on Database for Emotion Analysis Using Physiological Signals

Using a strategy that combines lateralization of emotions and ensemble learning, Pane et al. (2019) present a solution. By employing four distinct channel sequences and combinations, we were able to recover the EEG signal’s time-domain properties, frequency-domain features, and wavelet features. This was followed by applying random forest to DEAP datasets for classification, which resulted in a 75.6% success rate.

Cheng et al. (2020) used deep forest to categorise EEG emotions; they preprocessed the data, utilised the spatial position link across channels to generate a 2D frame sequence, and evaluated their model on a series of functional magnetic resonance imaging scans of the brain. Overall, the DEAP dataset saw improvements in its ability to classify valence and arousal levels. The first one scored 97.69%, and the second one scored 97.53%.

When constructing feature blocks, Ya et al. (2021) gathered differential entropy before feeding each segment into a separate deep learning model. A novel deep learning model was created by fusing a Graph Convolutional Neural Network (GCNN) and a Long Short-Term Memory (LSTM). In the end, numerous tests are performed on the DEAP dataset. Experiments conducted on the subjects yielded promising outcomes.

In order to solve this problem, Huang et al. proposed a bi-hemispheric discrepancy convolutional neural network model (BiDCNN). Using a three-input, one-output network structure made up of three layers of convolutional neural networks, it is able to learn the visible differences in behaviour between the left and right sides of the brain. The model obtains an accuracy of 94.38 percent when used on the DEAP datasets to predict potency, and an accuracy of 94.72% when used to predict arousal.

Wavelet packet decomposition (WPD) was used by Mokatren et al. (2021) to segment an EEG signal into five sub-bands and then calculate wavelet energy and wavelet entropy. By employing the electrode placements, a channel mapping matrix may be constructed. The features are retrieved and fed into a convolutional neural network (CNN) for further analysis. Classification accuracy for the DEAP dataset was 91.85% for valuing, and 91.06% for arousal.

In 2020, Moon et al. suggested a novel classification approach that makes use of CNN to better understand how the brain's neural circuits interact with one another. After finishing the connectivity matrix, the method was put to the test on the DEAP dataset, where it was found to be successful.

The deep neural network proposed by Liu J. et al. (2020) for emotion recognition in EEG data combines convolutional neural networks (CNN), sparse autoencoders (SAE), and conventional neural networks. Valence recognition was 89.49% accurate and arousal recognition was 92.86% accurate on the DEAP dataset.

In 2020, Yin et al. introduced the novel LRFS (locally robust feature selection) technique. The first iteration of the algorithm represented the EEG data that was collected using a probability
density function. As a next step, we compared the extent to which all density functions were the same for any two subjects, and built locally robust EEG features based on that information. In the end, data from a large number of people was aggregated using ensemble learning. Accuracy for valence was 67.97% and for arousal it was 65.10% on the DEAP dataset.

Tan et al. (2021) developed a subject-related short-term EEG emotion recognition framework using a spiking neural network (SNN), optimising super parameters of the data representation via pulse coding, and a dynamically evolving SNN, all with the aim of detecting transient changes in facial markers via EEG segmentation (deSNN). On the DEAP dataset, valence (67.76%) and arousal (78.97%) classifications worked admirably.

Emotions were classified using extreme learning machines and neural network classifiers. A variety of experiments were conducted to test the effectiveness of the proposed methods, and the results demonstrate that they significantly enhanced the recognition of speech emotions over earlier efforts. The use of EEG devices for emotion classification motivates the development of novel convolutional neural networks, sparse autoencoders (SAEs), and deep neural networks [J. Liu, 2020].

Convolutional Neural Network (CNN) performance was investigated across multiple preprocessing methods [D. A. Pitaloka 2017]. The study and improvement of CNN for detecting six basic emotions.

Data preparation approaches such as face detection, cropping, and noise addition are compared with one another. The accuracy of face detection, at 85.087%, made it the most pivotal stage in the preprocessing pipeline, ahead of both the subsequent stage and the raw data. Combining these methods, however, allows CNN to achieve an accuracy rate of 97% [N. Sabri, 2020].

A face expression category can be identified from a static frontal image. Grayscale to contrast stretch, Haar cascade for facial recognition, face model for mouth and eye locations, skin-color categorization approach for image segmentation, GLCM feature extraction; these are just some of the steps that happen to an image after it has been collected. Classifying temperaments using support vector machine (SVM) regression. A channel-fusion-based LSTM hybrid algorithm is proposed as a workaround [K. M. Ravikumar 2020].

4. STUDY ON RESEARCH RESULTS

In machine learning and AI, deep learning (DL) is a subfield that makes use of the available data to improve its performance [Dong S, 2021]. The use of DL has proven effective in a wide range of classification and regression applications and datasets. Since it has wide-ranging applications in fields as diverse as healthcare, image recognition, text analytics, and even cybersecurity, it has recently risen to prominence in the computing community [Iqbal H. Sarker, 2021].

Neural networks with several hidden layers are used in deep learning to carry out the nonlinear processes that are necessary for the execution of deep learning. Employing a tree-like design of transformations and hidden layers, complex functions can be trained to recognise output classes in a classification job. Even though deep learning (DL) techniques are still relatively new in compared to the lengthy history of emotion study in psychophysiology, some recent publications have examined their potential for automated emotion recognition. Figure 2a shows the statistical growth of research on DL and EEG emotion recognition and categorization from 2014 to 2021, as obtained from the scopus database. If this look at Figure 2b, it can show how the DL is distributed in the study of EEG emotion recognition.
Due to the difficulties in extracting useful and consistent features from time series data, many researchers have turned their attention to DL approaches. Using deep learning, it’s simple to extract one-of-a-kind features for integration into ML applications. Instead, it can figure out how to express the feature in a hierarchy on its own. As a result, the steps of preparing data and reconstructing feature spaces, both standard in machine learning pipelines, are unnecessary. Deep learning gets its name from the many layers of an artificial neural network, therefore the “deep” (ANN). More than three layers of neural networks are used to process inputs and outputs in deep learning algorithms. DL started off in the industry in the 1980s. The neocognitron was one of the earliest examples of a “deep” artificial neural network that incorporated neurophysiological knowledge. Since Hinton and Salakhutdinov’s key work in 2006 [Chen M, 2015], feature extraction has seen significant development. Multiple research have demonstrated the effectiveness of multilayer NNs for representing and learning features via iterative and non-iterative methods, respectively.

Fine-grained estimation and display of emotion using EEG data was first made possible by a cGAN-based emotion identification system presented by Fu et al. The average classification accuracy for two distinct experiments performed on the SEED dataset was 92.02 and 82.14 percent, respectively. An EEG emotion recognition system based on residual networks (ResNet) was presented by Cheah et al. (2021), and it achieved 93.42% accuracy on the SEED dataset.

Table 1 summarises these aforementioned algorithms. Here we discuss the methods we used and the outcomes we saw when classifying EEG readings for emotions using the SEED dataset. Four different methods, including those based on the time domain, the frequency domain, the time-frequency domain, and deep learning, are used in the process of extracting characteristics from EEG. Signals are typically categorised into good, negative, and neutral categories based on the associated emotions.

Fig 2: Research on EEG emotion identification using deep learning techniques during the past decade (2014-2021) using the scopus database.
Dm et al. (2021) proposed a multi-channel EEG approach for emotion recognition that employs a convolutional neural network trained to recognise rhythms. Accuracy (97.17), (96.81), and Dominance (97.17) on the DREAMER dataset (97.24). A unique multi-channel dynamic graph convolutional neural network was presented by Song et al. (2018) for emotion identification (DGCNN). Average recognition accuracies for valence, arousal, and dominance classification in the DREAMER database were 86.23, 84.54, and 85.02%, respectively.

Table 2 displays the final results obtained by applying these methods. In this table, we show the methods used and the results of EEG emotion categorization on the DREAMER dataset. Current studies utilising this dataset are mostly comparative in nature and are sparse in number. Features of EEG are typically extracted using deep learning, which is the primary method used by the researchers.
TABLE 2 Research results based on DREAMER.

<table>
<thead>
<tr>
<th>References</th>
<th>Method</th>
<th>Emotion</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cui et al., 2020</td>
<td>Regional-Asymmetric Convolution Neural Network (RACNN), Asymmetric Differential Layer (ADL), softmax</td>
<td>Valence, Arousal</td>
<td>0.9555</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.9701</td>
</tr>
<tr>
<td>Cheng et al., 2020</td>
<td>2D frame sequences, spatial position relationship, deep forest</td>
<td>Valence, Arousal</td>
<td>0.8903</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.9041</td>
</tr>
<tr>
<td>Liu Y. et al., 2020</td>
<td>Multi-level features guided capsule network (MLF-CapsNet)</td>
<td>Valence, Arousal, Dominance</td>
<td>0.9459</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.9526</td>
</tr>
<tr>
<td>Wang Y. et al., 2020</td>
<td>SPD matrix network (daSPDnet)</td>
<td>Valence, Arousal, Dominance</td>
<td>0.6799</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.7657</td>
</tr>
<tr>
<td>Dm et al., 2021</td>
<td>Multi-channel, rhythm selection, CNN, softmax</td>
<td>Valence, Arousal</td>
<td>0.9717</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.9681</td>
</tr>
<tr>
<td>Song et al., 2018</td>
<td>Dynamic graph convolution neural network (DGCNN), Softmax</td>
<td>Valence, Arousal, Dominance</td>
<td>0.8623</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.8454</td>
</tr>
</tbody>
</table>

CONCLUSION

The results of the study showed that high levels of emotional intelligence are strongly linked to the success of SMEs in both professional and private settings. It’s possible that facial expression detection research will continue to pique the interest of engineers and scientists for quite some time. The goal of this study is to review the literature on EEG emotion recognition in its entirety. First, we provide an in-depth explanation of the EEG technique, the emotion trigger mode, and the classification model. Following this, we explore the current state of the art in EEG emotion identification algorithms, dissecting its components from the perspectives of feature extraction, feature selection, and classifier. As a last step, we conduct a literature review in which the outcomes of various emotion classification approaches are contrasted. Future research will concentrate on addressing the many open questions and problems in EEG emotion recognition, such as the absence of emotion categories and datasets.

REFERENCES

7. Craik A., He Y., Contreras-Vidal J. L. (2019). Deep Learning for...
Electroencephalogram (EEG) Classification Tasks: A Review. J. Neural Eng. 2019:5. 10.1088/1741-2552/ab0a5


