



Emotion Recognition With Knowledge Graph Based on Electrodermal Activity

Hayford Perry Fordson¹, Xiaofen Xing¹, Kailing Guo^{1*} and Xiangmin Xu^{1,2*}

¹ School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China, ² School of Future Technology and the School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China

OPEN ACCESS

Edited by:

Chee-Kong Chui,
National University of Singapore,
Singapore

Reviewed by:

Mohammad Khosravi,
Persian Gulf University, Iran
Yizhang Jiang,
Jiangnan University, China

*Correspondence:

Kailing Guo
guokl@scut.edu.cn
Xiangmin Xu
xmxu@scut.edu.cn

Specialty section:

This article was submitted to
Perception Science,
a section of the journal
Frontiers in Neuroscience

Received: 03 April 2022

Accepted: 25 April 2022

Published: 09 June 2022

Citation:

Perry Fordson H, Xing X, Guo K and
Xu X (2022) Emotion Recognition With
Knowledge Graph Based on
Electrodermal Activity.
Front. Neurosci. 16:911767.
doi: 10.3389/fnins.2022.911767

Electrodermal activity (EDA) sensor is emerging non-invasive equipment in affect detection research, which is used to measure electrical activities of the skin. Knowledge graphs are an effective way to learn representation from data. However, few studies analyzed the effect of knowledge-related graph features with physiological signals when subjects are in non-similar mental states. In this paper, we propose a model using deep learning techniques to classify the emotional responses of individuals acquired from physiological datasets. We aim to improve the execution of emotion recognition based on EDA signals. The proposed framework is based on observed gender and age information as embedding feature vectors. We also extract time and frequency EDA features in line with cognitive studies. We then introduce a sophisticated weighted feature fusion method that combines knowledge embedding feature vectors and statistical feature (SF) vectors for emotional state classification. We finally utilize deep neural networks to optimize our approach. Results obtained indicated that the correct combination of Gender-Age Relation Graph (GARG) and SF vectors improve the performance of the valence-arousal emotion recognition system by 4 and 5% on PAFEW and 3 and 2% on DEAP datasets.

Keywords: affective computing, electrodermal activity, knowledge graph, emotion recognition, MLP

1. INTRODUCTION

Emotions are vital for humans because they influence affective and cognitive processes (Sreeshakthy and Preethi, 2016). Emotion recognition (Li et al., 2021) over the years has received large attention from academic researchers and industrial organizations and has been applied in numerous sectors including transportation (De Nadai et al., 2016), mental health (Guo et al., 2013), robotics (Tsiourti et al., 2019), and person identification (Wilaiprasitporn et al., 2020). Emotions are accompanied by physical and psychological reactions when an external or internal input is introduced. Representative methods for emotion recognition can be categorized into two sectors, physical-based and physiological-based. Physical-based signals include facial expressions (Huang et al., 2019), body gestures (Reed et al., 2020), audio (Singh et al., 2019), etc. All these signals are easy to collect and show good emotion recognition performance. However, the dependability of the physical-based data can be equivocal as it can be easy for people to intentionally alter their

reactions, which results in a false reflection of their real emotions. Physiological-based signals on the other hand include electroencephalogram (EEG), electrocardiogram (ECG), electromyogram (EMG), electrodermal activity (EDA)/galvanic skin response (GSR), blood volume pulse (BVP), temperature, photoplethysmography (PPG), respiration (RSP), and so on (Chen et al., 2021). While people may know the reasons why their signals are being collected, physiological signals relating to emotional states are hard to manipulate because they are controlled by the autonomic nervous system (ANS) (Shu et al., 2018). Compared to physical-based methods, emotion recognition based on physiological signal is less sensitive to societal and cultural differences amongst users and are acquired in natural emotional states (Betella et al., 2014). Physiological signals also perform better in detecting sounds than on beeps as they primarily reflect emotional states (Stuldreher et al., 2020).

Emotion should be well defined and approached in a quantifiable manner. Psychologists model emotions in two ways: discrete and multi-dimensional continuous (Liu et al., 2018; Mano et al., 2019; Yao et al., 2020). Discrete emotion theory claims that there is a small number of core emotions. They are usually limited and can barely differentiate heterogeneous emotions and composite mental states. Multi-dimensional emotion model on the other hand finds correlation among different discrete emotions which correspond to a higher level of a particular emotion. Valence-Arousal (V-A) space model has been widely used in affective computing research as a type of multi-dimensional continuous emotion model. Valence stands for a negative to a positive level of emotion that ranges from unpleasant to pleasant feelings. Arousal stands for low to high level of emotion that ranges from drowsiness to intense human excitement. The combination of valence and arousal space model enables the creation of continuous emotion models that captures both moderate and complex emotions needed for building accurate emotion recognition systems (Zhang et al., 2020). This study adopts the V-A space model for final emotion state classification.

The electrodermal activity also known as galvanic skin response is a non-stationary signal. **Figure 1** shows the EDA signal of participant #39 from the PAFEW dataset. Skin conductance response occurs when there is an induction of stimuli and may occur many times in short instances as shown in the figure. Portable low-cost EDA sensors (Milstein and Gordon, 2020) have been developed and applied in a physiological research context as signals become easier to collect. Prior works have connected electrodermal activities with biofeedback in diffusing brain activation as its cost effective and easy to apply treatment options are with no known risk (Gebauer et al., 2014). Researchers have begun combining and comparing the sufficiency of traditional machine learning and deep learning algorithms to predict users' mental states from EDA (Ayata et al., 2017).

Knowledge graph (KG) (Hogan et al., 2021) involves the acquisition and integration of information from data, creating relevant entities and relations, and predicting a meaningful output. As a new type of representation, KG has gained a lot of attention in the field of cybersecurity, natural language

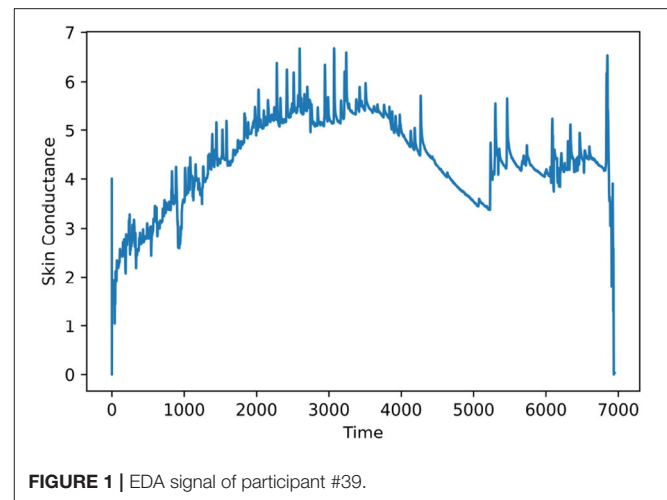


FIGURE 1 | EDA signal of participant #39.

processing, recommendation systems, and human cognition (Jia et al., 2018; Guan et al., 2019; Chen et al., 2020; Qin et al., 2020; Yang et al., 2020). Wang (Wang et al., 2019) represented cognitive relations between different emotion types using a knowledge graph. Yu et al. (2020) studied a framework made up of non-contact intelligent systems, knowledge modeling, and reasoning to represent heart rate (HR) and facial features to predict emotional states. Farashi and Khosrowabadi (2020) used the knowledge of minimum spanning tree (MST) graph features derived from computational methods to classify emotional states. Wenbo et al. used knowledge embedding in a deep relational network to capture and learn relationships between cartoon, sketch, and caricature face recognition (Zheng et al., 2020). Gender and age are important factors that affect human emotion. To the best of our knowledge, there is no existing work on emotion recognition that attempts to combine gender and age with EDA/GSR statistical features (SF) to accurately model emotional states. In this paper, we propose an effective knowledge embedding graph model based on observed participants' gender and age to capture the relations between given entities. This has been lacking in EDA-Based Emotion Recognition Systems where researchers only focus on time, frequency, and time-frequency statistical features (SFs). We further propose a sophisticated feature fusion technique that exploits the knowledge embedding vectors as a weight to SFs. We then utilized a deep neural network to capture relevant complex information from participants' age and gender and predict emotional states.

The main contributions of this article is summarized as follows:

- 1) We propose an effective knowledge embedding graph model based on observed participants' gender and age to capture the relations between given entities.
- 2) We propose a sophisticated feature fusion technique that exploits the knowledge embedding vectors as a weight to SFs.
- 3) The proposed model shows better recognition accuracy when compared to other methods.

The overall implementation framework of this paper is illustrated in **Figure 2**.

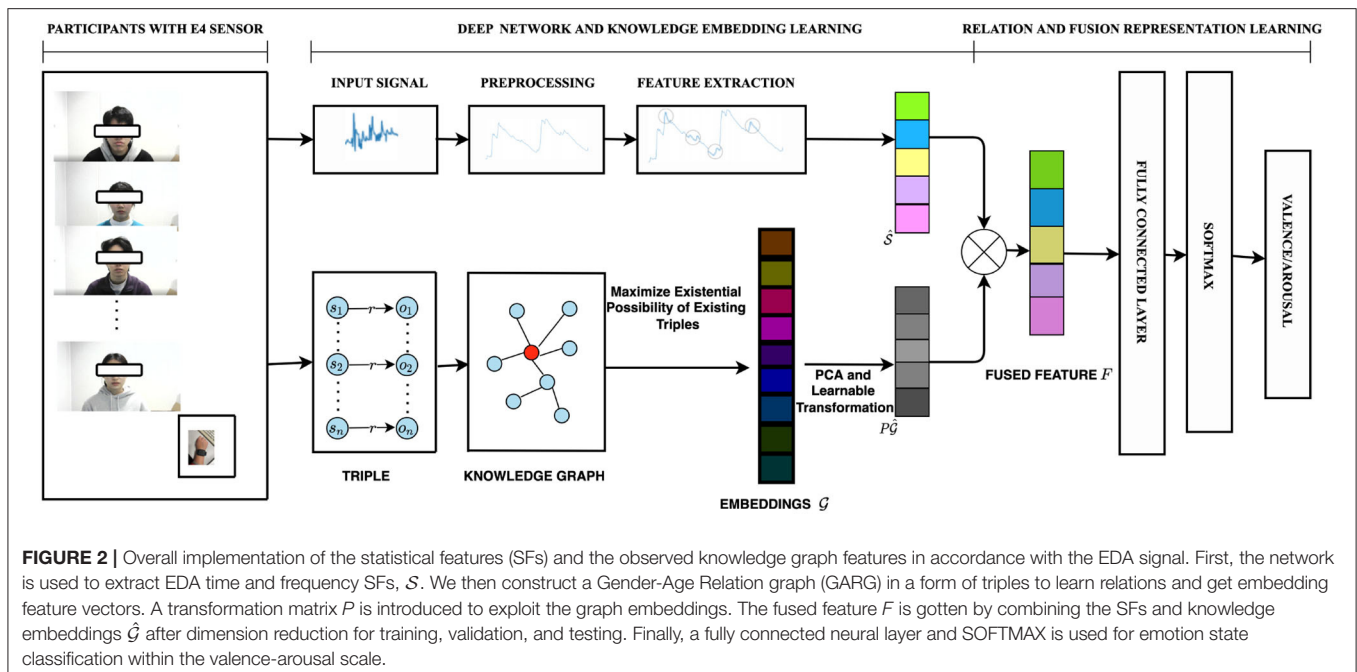


FIGURE 2 | Overall implementation of the statistical features (SFs) and the observed knowledge graph features in accordance with the EDA signal. First, the network is used to extract EDA time and frequency SFs, S . We then construct a Gender-Age Relation graph (GARG) in a form of triples to learn relations and get embedding feature vectors. A transformation matrix P is introduced to exploit the graph embeddings. The fused feature F is gotten by combining the SFs and knowledge embeddings \hat{G} after dimension reduction for training, validation, and testing. Finally, a fully connected neural layer and SOFTMAX is used for emotion state classification within the valence-arousal scale.

The rest of the paper is organized as follows: Section 2 presents related background and approaches to EDA signals and knowledge graph and also presents our novel methodology and algorithms. Section 3 reports the experimental results. Section 4 discusses the novel approach and compares them with previous works. Finally, Section 5 concludes the study.

2. MATERIALS AND METHODS

2.1. Related Background

Electrodermal activity-based emotion recognition has become popular in affective computing and cognitive developmental studies. Knowledge graphs have also been applied to embed entities in attempts to get more information from available data. Both ideas are frequently used nowadays in a variety of contexts and researchers have attempted to design frameworks to accurately solve today's problems.

2.2. EDA-Based Emotion Recognition

Electrodermal activity is a cheaper, easily collected physiological signal that reflects internal reactions to yield exhilaration. Deep learning algorithms (Yang et al., 2022) have become a catalyst in emotion recognition to capture time, frequency, and time-frequency information in EDA signals. Multilayer Perceptron (MLP), Convolutional neural network (CNN), Deep Belief Networks (DBN), Attention-Long Short Term Memory (A-LSTM) are often used. These algorithms with processors in the suggestive connection between neurons and layers can learn to extract relevant features for a reference task. In the study by Hassan et al. (2019), signals from EDA, Zygomaticus Electromyography (zEMG), and Photoplethysmogram (PPG) are fused using DBN for in-depth feature extraction to gear up a feature fusion vector. The vector is used to classify five discrete

basic emotions, relaxed, happy, sad, disgust, and neutral. The study also used Fine Gaussian Support Vector Machine (FGSVM) together with radial basis kernel function to help classify the non-linearity in the human emotion classification. Song et al. also proposed an attention-long short term memory (A-LSTM) to extract discriminative features from EDA and other physiological signals to strengthen sequence effectiveness (Song et al., 2019). These studies did not use the gender and age information of participants as features to validate their results. Also, the performance of their approach did not quantify the contribution of the EDA signal but proves deep learning is effective in extracting emotional features.

2.3. Knowledge Graph

An effective way to learn representation from data is through embedding a KG into a vector space while maintaining its properties (Li et al., 2020). It involves knowledge triples h, r, t compiled of two entities h and t and a relation r . Face recognition studies have tried to study representational gaps using deep learning techniques to derive useful facial information (Cui et al., 2020). KG embedding can also be used to find vector representation of known graph entities by regarding them as a translation of entities in space. Human knowledge allows for a formal understanding of the world (Ji et al., 2022). For cognition and human level intelligence, knowledge graphs that represent structural relations between entities have become an increasingly popular research direction. The performance of these approaches shows that when KG is integrated, model performance is enhanced. Graph convolutional network has been studied for drug prediction in computational medicine (Nguyen et al., 2022). KG has been used for image processing and facial recognition but has never been used with physiological signals for emotion recognition. We attempt to address this issue.

2.4. Datasets

2.4.1. The PAFEW Dataset

For the PAFEW dataset (Liu et al., 2020), there are 57 healthy students that participated in the experiment and each view about 80–90 video clips in all 7 emotion categories. Participants are aged between 18 and 39 years and were subjectively rated to keep them focused. The dataset is summarized in **Table 1**. Participants watched videos with one label continuously. This encourages them to be fully dissolved into one emotion before moving on to watching another video with a different label. The initial dataset was collected in a multi-class fashion (shown in **Figure 3**). From the figure, we can conclude that the PAFEW dataset is highly overlapped resulting in class imbalance. To tackle the imbalance issue, we subset target labels into the valence and arousal dimensions (shown in **Figure 4**). This is to more effectively capture both moderate and complex emotions. For example, fear and anger are two separate emotions which may differ from person to person but fall under high arousal and negative valence. Also, to be able to assess the performance metrics, we defined the total number of samples per emotion used to train and test the performance of the proposed method (seen in **Table 2**). Overall, we have a total of 3,554 data points.

2.4.2. DEAP Dataset

The DEAP dataset (Koelstra et al., 2012) is made up of 32 participants that watched 40 video clips while their physiological signals and facial expressions are recorded. Participants are aged between 19 and 37 (mean age 26.9). Participants' rating was also collected. We used the GSR data/channel to design an emotion recognition system. A summary of the DEAP dataset is shown in **Table 3**.

2.5. Methodology

In this section, we first illustrate the Gender-Age Relation Graph (GARG) and SF structures that are used in conducting the experiment. We also describe the weighted fusion representation learning which is the keystone of our solution.

2.5.1. GARG Feature Learning Structure

Knowledge bases represent data by a directed graph with labeled edges (relations) between nodes (entities). We attempt to solve the problem of valence-arousal emotion recognition. Usually, the directed graph is represented by triples in the form (subject, predicate, and object). In our case, we use the gender and age information of a participant to construct the graph, so we have triples like (*Participant*, *AgeIs*, 28) and (*Participant*, *GenderIs*, *F/M*).

Formally speaking, we define our knowledge graph $\mathcal{G} = \mathcal{E} \times \mathcal{R} \times \mathcal{E}$:

- \mathcal{E} : a set of entities;
- \mathcal{R} : a set of relations.

Let $\mathcal{E} = (e_1, e_2, \dots, e_{N_e})$ represent the set of entities in the knowledge graph and let $\mathcal{R} = (r_1, r_2, \dots, r_{N_r})$ represent the set of all relation types. We represent each triple as $x_{a,b,c} = (e_a, r_c, e_b)$ and model its presence with a binary random variable $y_{a,b,c} \in \{0, 1\}$ indicating whether a triple exists or not. In this paper, we

TABLE 1 | Summary of the PAFEW dataset.

Attributes	PAFEW	
Length of sequence	480–6,040 ms	
No. of sequence	4023	
No. of participants	57	
Max. participants per clip	9	
Min. participants per clip	2	
Mean participants per clip	5.2	
No. of sequence per expression	Happy	699
	Surprise	362
	Disgust	436
	Anger	475
	Fear	434
	Sad	472
	Neutral	676

extract latent embedding features by maximizing the existential possibility of existing triples.

Following (Trouillon et al., 2016), we use complex embedding features. The complex embeddings of the subject, object, and their relation are denoted by $e_s, e_o,$ and $w_r \in \mathbb{C}^k$, respectively. The scoring function of a triple is defined by:

$$\begin{aligned}
 f(s, o, r; \Theta) &= \text{Re}(\langle w_r, e_s, \bar{e}_o \rangle) \\
 &= \text{Re}(\langle \sum_{k=1}^K w_{rk}, e_{sk}, \bar{e}_{ok} \rangle) \\
 &= \langle \text{Re}(w_r), \text{Re}(e_s), \text{Re}(e_o) \rangle \\
 &\quad + \langle \text{Re}(w_r), \text{Im}(e_s), \text{Im}(e_o) \rangle \\
 &\quad + \langle \text{Im}(w_r), \text{Re}(e_s), \text{Im}(e_o) \rangle \\
 &\quad - \langle \text{Im}(w_r), \text{Im}(e_s), \text{Re}(e_o) \rangle, \tag{1}
 \end{aligned}$$

where $\text{Re}(\cdot)$ and $\text{Im}(\cdot)$, respectively, extract the real and imaginary parts of a complex vector, and Θ denotes the parameters corresponding to the model.

It is expected to score the correct triples (e_1, r, e_2) higher than incorrect triples (e', r, e_2) and (e', r, e_1) which is not equal to correct triples by one entity. Following Toutanova and Chen (2015), the conditional probability $p(e_2 \mid e_1, r)$ for the object entity given the relation and the subject entity is defined as:

$$p(e_2 \mid e_1, r_c; \Theta) = \frac{e^{f(x_{e_1, e_2, r_c}; \Theta)}}{\sum_{e'_2 \in \text{Neg}(e_1, r_c, ?)} e^{f(x_{e_1, e'_2, r_c}; \Theta)}} \tag{2}$$

where $(e_1, r, ?)$ is a set of triples that do not match the object position in the relation triple. Since the entity number of such a set is large, the negative triples are randomly sampled from the full set. Similarly, given the relation and object, the conditional probability of the subject is defined as:

$$p(e_1 \mid r_c, e_2; \Theta) = \frac{e^{f(x_{e_1, e_2, r_c}; \Theta)}}{\sum_{e'_1 \in \text{Neg}(?, r_c, e_2)} e^{f(x_{e'_1, e_2, r_c}; \Theta)}} \tag{3}$$

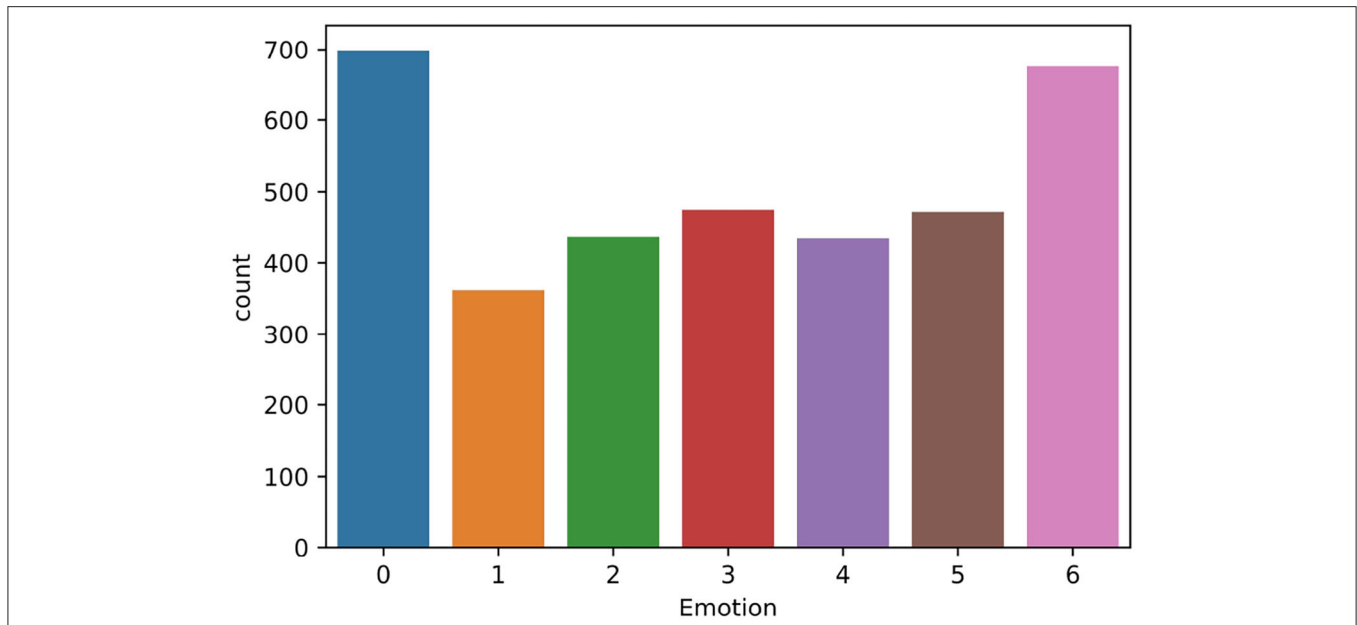


FIGURE 3 | Imbalanced emotion class distribution in the PAFEW dataset.

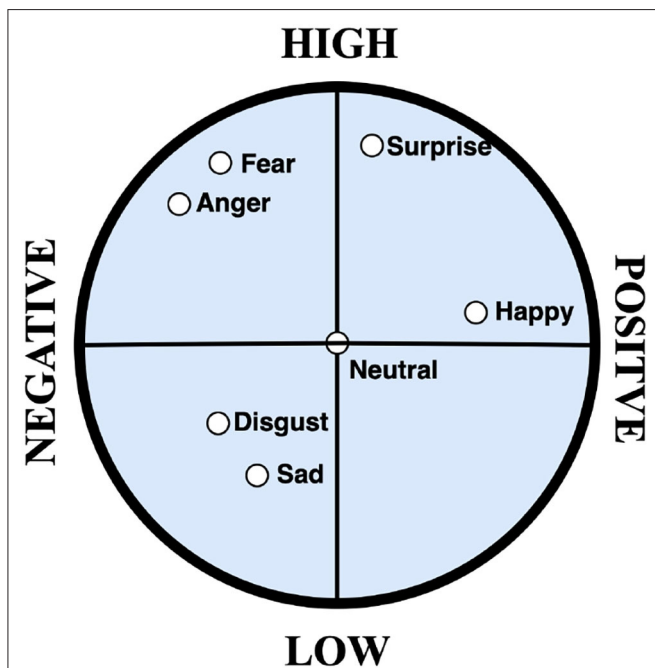


FIGURE 4 | Valence-Arousal 2D Model: This is put into two separate groups. The first group is high arousal and low arousal. High arousal contains surprise, fear, anger, and happy. Low arousal contains disgust, neutral, and sad. Second group is positive valence and negative valence. Positive valence contains neutral, surprise, and happy. Negative valence contains fear, anger, sad, and disgust (Liu et al., 2020).

defined as the sum of the negative log-probabilities of observed triples with $L2$ penalty. Suppose X denotes the set of observed triples and λ is the trade-off parameter, the training loss is given by:

$$\begin{aligned} \mathcal{L}(X, \Theta, \lambda) = & - \sum_{x_{e_1, e_2, r_c} \in X} \log p(e_2 | e_1, r_c : \Theta) \\ & - \sum_{x_{e_1, e_2, r_c} \in X_{train}} \log p(e_1 | r_c, e_2 : \Theta) \\ & + \lambda \Theta^T \Theta. \end{aligned} \tag{4}$$

The algorithm of graph feature extraction is summarized in **Algorithm 1**.

2.5.2. SF Learning Structure

2.5.3. Preprocessing

In this paper, we used PAFEW and DEAP databases for analysis. E4 wristband is used to collect EDA signals at a rate of 4 Hz in the PAFEW dataset. Participants were asked to watch videos of the same emotional label consecutively. We treat each emotion category as continuous in time for normalization. Data obtained varies significantly from 0 to 6.68 among participants. Min-max normalization is applied to reduce in-between participant differences. The 11-point median filter is used to remove noise, smoothen signal sequence. For the DEAP dataset, we used the preprocessed data. We also normalized the data to reduce intra- and inter-individual differences associated with emotional responses. The data is downsampled to 128 Hz. After, it is segmented into 60 s trials and 3 s pre-trials, baseline is removed.

Given the definition of conditional probability, we can maximize the probability of existing triples. Our training loss function is

TABLE 2 | List of independent data points for valence arousal classification.

	Valence		Arousal				
	Positive	Negative	High	Low			
Happy	699	Disgust	436	Happy	699	Disgust	436
Surprise	362	Anger	475	Fear	434	Sad	472
Neutral	676	Fear	434	Anger	475	Neutral	676
-	-	Sad	472	Surprise	362	-	-
Total	1737	Total	1817	Total	1970	Total	1584

TABLE 3 | Summary of the DEAP dataset.

Array name	Array shape	Array contents
data	40 × 1 × 8,064	trial × channel × data
labels	40 × 2	trials × label

Algorithm 1 | GARG.**Input:** Triples $G = \{(\mathcal{E} \times \mathcal{R} \times \mathcal{E})\}$.**Output:** Embedding Feature Vector.**Procedure:**

- 1: Create entities (s and o) that will form a graph.
- 2: Initialize $x_{a,b,c}$ based on relationship between entities and ensure its presence using $y_{a,b,c} \in \{0, 1\}$.
- 3: **loop**
- 4: **for** each ϕ_{abc} **do**
- 5: Define the score triple using Equation (1).
- 6: Corrupt to generate Neg object triples $G^- = (\mathcal{E} \times \mathcal{R} \times \mathcal{E}')$ at training for positive, or true triple.
- 7: Evaluate by defining filters to ensure no negative statements generated by the corruption procedure are actually positives by concatenating train and text sets.
- 8: **end for**
- 9: Update entity and relation embedding w.r.t gradients of Equation (4).
- 10: **end loop**

2.5.4. Feature Extraction

We extract time and frequency signals which is listed in **Table 4**. Features were extracted for the entire duration as participants watch emotional videos. For the PAFEW data, extracted features were categorized into three groups: basic statistical variables, first-order differential variables, and second-order differential variables. For the DEAP dataset, we used the toolbox for emotional feature extraction from physiological signals (TEAP) (Soleymani et al., 2017) to extract SFs. These features are in line with the work of Shukla et al. (2019) that extensively studied and reviewed EDA features across statistical domains relevant for emotion recognition. Our choice of features is easy to compute online, thus, making them advantageous in the future for real-time recognition.

2.5.5. Graph Feature Weighted Fusion Representation Learning

We present our proposed weighted feature fusion mechanism for representation learning which effectively improves the performance of our model.

After training with the loss function (4), we extract the real part of the complex feature vector of the subjects as a graph embedding feature, and the feature matrix of all entities is denoted by \mathcal{G} . High-dimensional space is required to embed the entities so that they are adequately distinguished, but the dimension is incompatible with the SF. Here, we use principal component analysis (PCA) to reduce the dimension to be the same as SF. The feature matrix after dimension reduction is denoted by $\hat{\mathcal{G}}$.

To fuse SF S and graph embedding feature $\hat{\mathcal{G}}$, one naive method is concatenation. Since they are extracted based on a totally different viewpoint, simple concatenation may not fuse them well. Here, we use a more sophisticated method by exploiting the graph embedding feature as the weight of the SF. Because the dimension order of the graph embedding feature does not naturally match the dimension order of SF, we introduce a learnable transformation matrix P . The fused feature is given by:

$$F = S \otimes (P\hat{\mathcal{G}}), \quad (5)$$

where \otimes denotes element-wise product. By taking F as the input of a neural network, we can obtain the transformation matrix P through training. Here, we use an MLP with the architecture of three hidden layers and an output layer. The dimension for the number of the hidden units is 64, 128, and 64. The output of the network is 2-dimension vector. We use the mean square error (MSE) criterion and Adam optimizer. Dropout is set at 0.2, batch-normalization and ReLU are applied after each hidden layer after which Softmax is employed for final classification. The whole procedure of our feature fusion learning method is summarized in **Algorithm 2**.

3. RESULTS

We reported experimental results separately for each experiment to have a clearer assessment of our knowledge graph EDA-based approach.

Tables 5, 6 present the average results of classification recall, precision, accuracy, and F1-scores on PAFEW and DEAP. We also report the standard deviation of accuracy.

The results are based on leave-one-out training and testing, i.e., training all but one subject data, and using that one subjects data for testing. The results for SF features are consistent with previous studies. However, the GARG feature embeddings give a boost in model performance, hence improvement in classification results. This is because our graph embedding feature vectors can effectively learn the representation between participants' gender and age and their respective emotional labels. Also, because the GARG features are extracted differently and require a high dimensional space to effectively mine related information.

TABLE 4 | Features used in this research.

Dataset	Features	Parameters	Discription
PAFEW	Basic statistical features	MeanEDA	Mean of signal
		KurtEDA	Kurtosis of signal
		SkewEDA	Skewness of signal
		MaxEDA	Maximum of signal
		MinEDA	Minimum of Signal
		StdEDA	Standard deviation of signal
	First order differential features	VarEDA	Variance of Signal
		MeanAbsDiff	Mean of first order
	Second order differential features	MeanNegativeDiff	Negative mean of first order differential
		MeanSecAbsDiff	Mean of first order
Graph embedded features	MeanSecNegativeDiff	Negative mean of second order differential	
	Gender	Gender of participant	
DEAP	Statistical Features	Age	Age of participant
		No. of Peaks	Number of peaks in resistance
		Amplitude of Peaks	GSR peak amplitude
		Rise time	Time taken to reach peak
		Statistical moments	Mean and Standard deviation
	Graph Embedded Features	Local minima	No. of local minima in GSR signal
		Gender	Gender of participant
		Age	Age of participant

Algorithm 2 | Weighted SF-GARG.**Input:** Graph embedding feature \hat{G} and statistical feature \mathcal{S} **Procedure:**

- 1: Extract embedding feature using **Algorithm 1** in high dimension and denoted as \mathcal{G}
- 2: Use PCA to reduce dimension of feature matrix then denote as \hat{G}
- 3: Extract statistical feature and denote as \mathcal{S}
- 4: Introduce a learnable feature transformation matrix P
- 5: Compute the fused feature F using Equation (5)
- 6: Define the neural network architecture
- 7: Take the fused feature F as input and train the network

Output: P, W

Furthermore, directly combining them with the SF features will be inappropriate. We, therefore, use PCA to match their dimensional space to the SFs. Hereby, we used them as a weight to the statistical features to enhance model performance. The tables also clearly show that, a weighted combination of SF-GARG features yield better performance than those of the SF features alone. MLP clearly shows the highest performance in all tables compared to SVM, KNN, RF, and NB.

The Nemenyi test (Demšar, 2006) and Friedman test (Friedman, 1937) are used to show statistical significance of our method. To detect the differences between multiple methods across multiple test results, the Friedman test, which is a non-parametric statistical test, is used. Its null hypothesis states

that - all methods have the same performance. The Nemenyi test is used to distinguish whether the performances of the methods are significantly different if the null-hypothesis is rejected. In the PAFEW dataset, on the arousal scale, we calculated $F_{(5,)} = 3.57$ for accuracy and $F_{(5,)} = 3.75$ for F1 score. On the valence scale, we calculated $F_{(5,)} = 3.69$ for accuracy and $F_{(5,)} = 3.89$ for F1 score. This showed $p < 0.05$. Similarly, in the DEAP dataset on the arousal scale, we calculated $F_{(5,)} = 3.63$ for accuracy and $F_{(5,)} = 3.76$ for F1 score. On the valence scale, we calculated $F_{(5,)} = 3.68$ for accuracy and $F_{(5,)} = 3.74$ for F1 score which showed $p < 0.05$. The critical distance (CD) of the Nemenyi test is defined as follows: $CD = q_{\alpha} \sqrt{\frac{k(k+1)}{6N}}$, where q_{α} is a default critical value of 0.05, k denotes the number of methods ($k = 5$) in this work, and N the number of result groups. For PAFEW, $N = 57$ and for DEAP, $N = 32$. This resulted in no overlaps with methods suggesting that our proposed methods are statistically different in performance with compared methods.

Figure 5 is the confusion matrix for the classification results on the PAFEW dataset and shows a clearer picture of our model performance.

Figures 6, 7 show the training and validation loss on the PAFEW dataset for both valence and arousal classification. Our method reduces the complexity of the model with respect to the reduction of dimension of the complex embeddings and the number of layers in our neural network model. In the training phase, the time taken to train the model is relatively fast. The figures show that our model is able to reduce variance of bias and finds real relationships between

TABLE 5 | Classification results (%) on the PAFEW dataset.

Features	Model	Valence				Arousal			
		Recall	Precision	Accuracy	F1-Score	Recall	Precision	Accuracy	F1-Score
SF	SVM	0.623	0.623	0.634 ±0.064	0.623	0.714	0.733	0.614 ±0.059	0.723
	KNN	0.752	0.752	0.763 ±0.057	0.752	0.713	0.713	0.690±0.060	0.713
	Random forest	0.754	0.739	0.754±0.058	0.746	0.724	0.741	0.724±0.058	0.732
	Naive bayes	0.612	0.663	0.612±0.063	0.636	0.691	0.691	0.654±0.061	0.691
	MLP	0.812	0.812	0.794±0.052	0.812	0.753	0.753	0.744±0.057	0.753
SF-GARG	SVM	0.741	0.741	0.652±0.058	0.741	0.734	0.724	0.633±0.059	0.729
	KNN	0.789	0.789	0.745 ±0.054	0.789	0.745	0.745	0.745 ±0.058	0.745
	Random forest	0.778	0.778	0.778±0.055	0.778	0.756	0.767	0.745±0.056	0.761
	Naive bayes	0.717	0.732	0.690±0.059	0.724	0.724	0.704	0.678 ±0.060	0.714
	MLP	0.847	0.862	0.828±0.046	0.854	0.815	0.793	0.770±0.054	0.804

The bold values are the highest accuracy obtained.

TABLE 6 | Classification results (%) on the DEAP dataset.

Features	Model	Valence				Arousal			
		Recall	Precision	Accuracy	F1-Score	Recall	Precision	Accuracy	F1-Score
SF	SVM	0.690	0.700	0.590±0.081	0.695	0.705	0.732	0.690±0.078	0.718
	KNN	0.750	0.690	0.700±0.082	0.719	0.767	0.745	0.734 ±0.077	0.756
	Random forest	0.754	0.702	0.733±0.081	0.727	0.752	0.752	0.713±0.076	0.752
	Naive bayes	0.614	0.589	0.623±0.087	0.601	0.754	0.744	0.691±0.077	0.749
	MLP	0.774	0.774	0.723±0.074	0.774	0.789	0.783	0.761±0.729	0.786
SF-GARG	SVM	0.745	0.749	0.770 ±0.077	0.747	0.700	0.720	0.690±0.079	0.710
	KNN	0.740	0.733	0.714 ±0.078	0.736	0.745	0.767	0.734±0.075	0.756
	Random forest	0.740	0.740	0.720 ±0.078	0.740	0.750	0.756	0.714±0.076	0.753
	Naive bayes	0.720	0.710	0.680 ±0.080	0.715	0.740	0.740	0.689 ±0.078	0.740
	MLP	0.805	0.802	0.793 ±0.070	0.803	0.804	0.801	0.802±0.071	0.798

The bold values are the highest accuracy obtained.

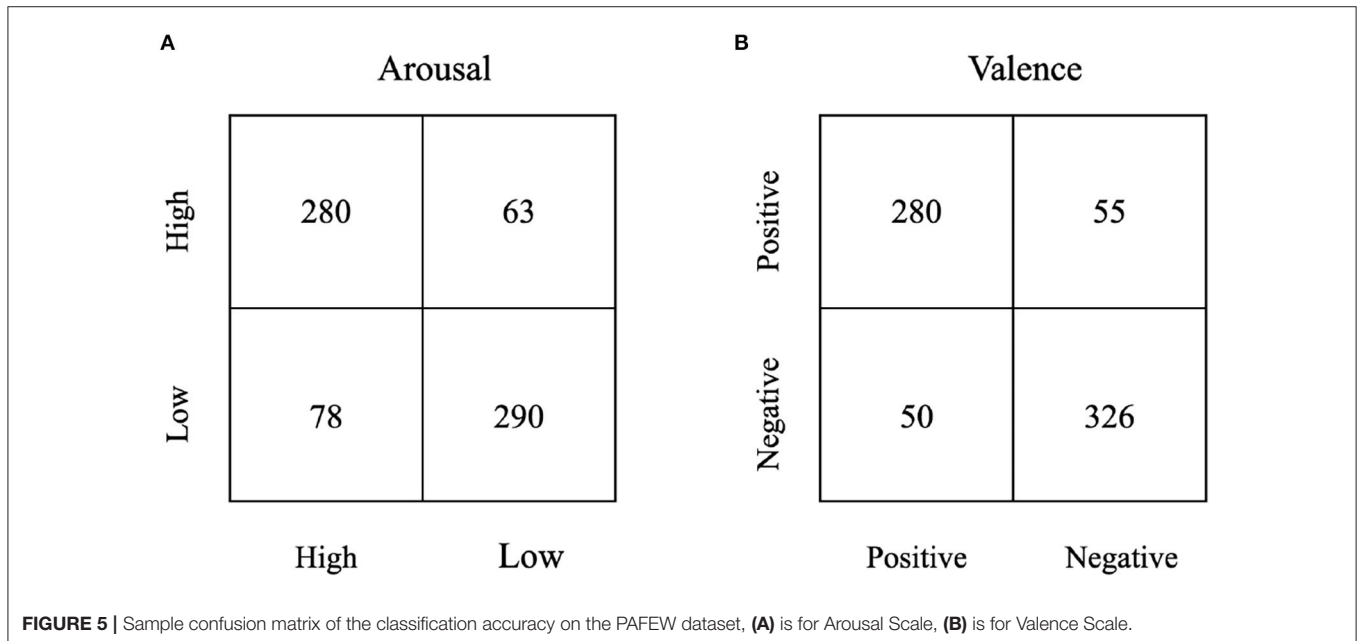


FIGURE 5 | Sample confusion matrix of the classification accuracy on the PAFEW dataset, **(A)** is for Arousal Scale, **(B)** is for Valence Scale.

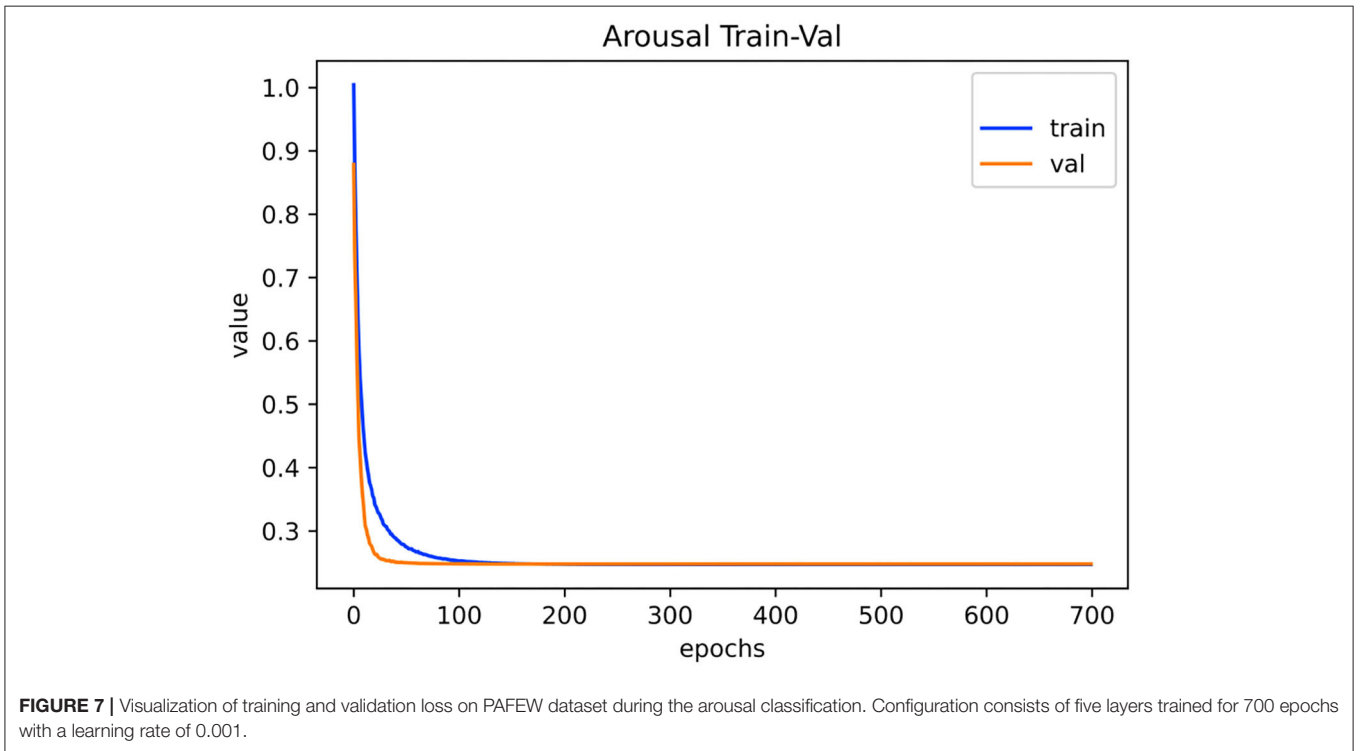
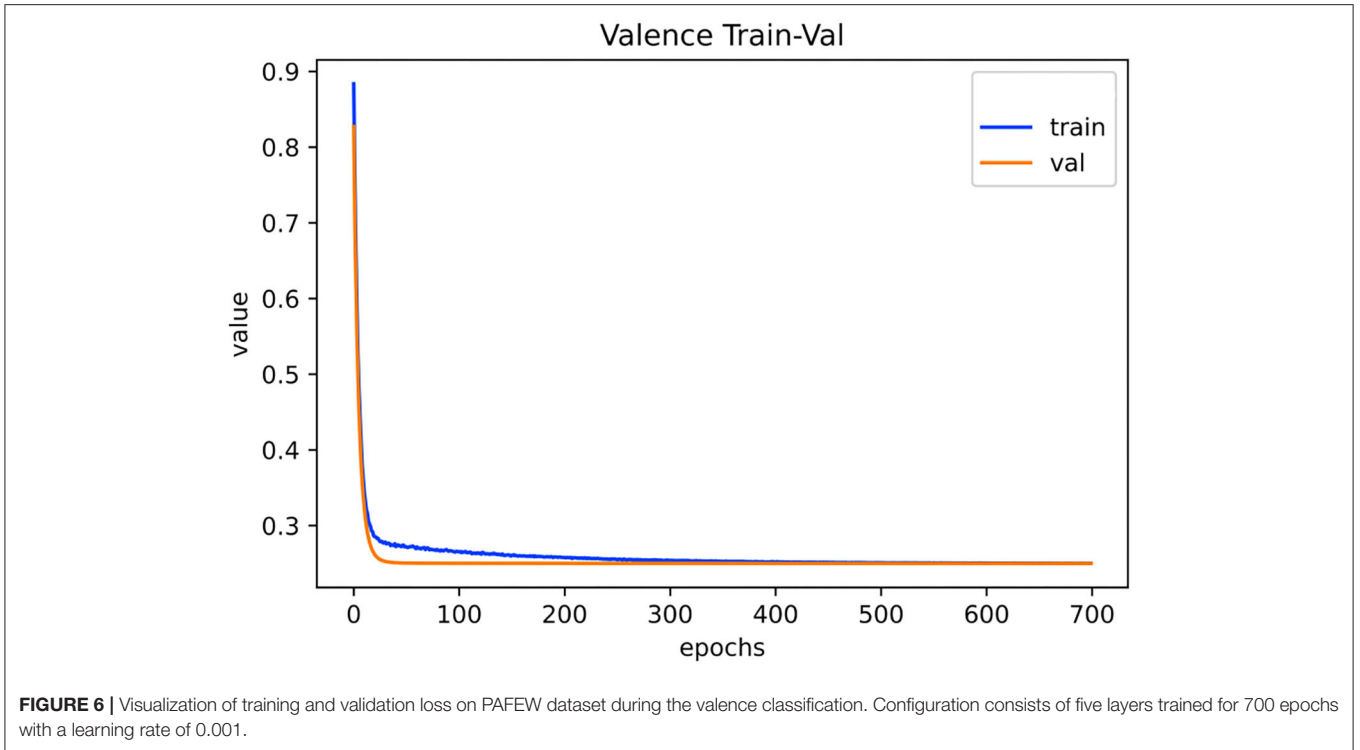


TABLE 7 | Accuracy and F1-Score comparison with state-of-the-art-research using electrodermal activity (EDA) signals.

Paper	Dataset	Input signals	Val	Valence score		Arousal score	
				Accuracy	F1-Score	Accuracy	F1-Score
Ganapathy et al., 2020	DEAP	EDA	Leave-one-out	0.721	0.713	0.754	0.791
Koelstra et al., 2012	DEAP	EEG, EOG, EMG, SKT, BVP, EDA, RSP	Leave-one-out	0.631	0.613	0.574	0.533
Soleymani et al., 2017	DEAP	EDA	Leave-one-out	0.591	0.574	0.560	0.521
Zhang et al., 2020	DEAP	EEG, EMG, EDA, RSP	Leave-one-out	0.645	0.696	0.631	0.701
Liu et al., 2020	PAFEW	EDA	Leave-one-out	0.727	0.727	0.666	0.666
Our Work	DEAP	EDA, graph embedding	Leave-one-out	0.793	0.803	0.802	0.798
Our Work	PAFEW	EDA, graph embedding	Leave-one-out	0.828	0.854	0.770	0.804

The bold values are the highest accuracy obtained.

TABLE 8 | Taxonomy of the compared approaches to EDA-Based emotion recognition.

Method	Knowledge graph embedding	Deep learning	Wearable devices
Ganapathy et al., 2020	No	Yes	No
Koelstra et al., 2012	No	No	No
Soleymani et al., 2017	No	No	No
Zhang et al., 2020	No	Yes	No
Liu et al., 2020	No	Yes	Yes
SF-GARG	Yes	Yes	Yes

graph and SF vectors from the EDA signal and their emotional labels.

4. DISCUSSION

In this paper, we focused on two kinds of issues. 1) Knowledge graph generation using participant’s gender and age information and 2) Weighted feature fusion issue for EDA emotion classification system. In generating a knowledge graph using gender and age, we created triple entities and drew relationships between EDA signals, participants’ gender and age, and their respective emotion labels. The weighted feature fusion issue is an algorithmic idea that exhibits the performance of the proposed method in an effective way for EDA emotion applications. We also deliberated the advantages of MLP over other machine learning methods.

Regarding knowledge graph generation, results indicated that our deep learning approach can capture relevant complex information from participants’ gender and age and predict emotional states correctly. On the contrary, SVM, KNN, RF, and NB could not fully mine such complex information but averaged consistent results higher than only SF features which also proves GARG features as effective for learning representation in EDA. The deep learning approach can also capture invisible features that traditional methods cannot. The results also show that our model can effectively identify true positive results (recall) than previous EDA-based approach.

Pertaining to the weighted feature fusion issue, the proposed sophisticated algorithm outruns existing methods. Unlike previous methods where features from other signals are directly concatenated, we introduced a transformation matrix that exploits the gender-age graph embedding features as a weight to the SF and performed element-wise multiplication. **Table 7** compares our work with previous studies. Our model is able to reach an F1-measure of 85.4% and 80.4% for valence and arousal, respectively, for PAFEW dataset. On the DEAP dataset, our model reaches an F1-measure of 80.3% on the valence scale and 79.8% on the arousal scale. In comparison with Koelstra et al. (2012) and Soleymani et al. (2017) that uses machine learning algorithms, the features extracted did not particularly reflect participants’ affective states. Our deep learning approach is data driven and based on training data. Also, the GARG embedded features reflect participants’ affective state in the valence-arousal scale. Hence, the robustness of our method compared to theirs. Additionally, Ganapathy et al. (2020), Zhang et al. (2020), and Liu et al. (2020) uses DL but our model was manifestly superior with regard to training speed with low SD in both valence and arousal dimension. **Table 8** summarizes all compared methods with respect to some qualitative features. It is easily observed that our proposed approach possesses two kinds of qualities—integrating knowledge graph embedding vectors and EDA features and excellent complexity order those other methods do not.

The proposed knowledge graph and EDA-based method take advantage of a single EDA module and graph embedding tricks to improve fusion performance. Our approach does not consume more computation time compared with other methods that use multiple kernel combination techniques. The introduction of the learnable transformation matrix enables the feature space to form an implicit combination. This is less computationally expensive compared to other methods. The multilayer neural network employed exhibits strength and flexibility in representation learning hence superior performance regarding fusion capabilities.

Finally, none of the previous works utilized participants’ GARG embeddings as weight to statistical features to optimize their approach. This is an important advantage of our proposed approach. Investigating GARG embeddings and accurately

combining them with SFs lead to a 4% and 5% increase in performance on the PAFEW dataset and 3% and 2% on the DEAP dataset for valence and arousal, respectively.

In the future, we can try to explore more physiological signals like EEG, EMG, and BVP with knowledge graph embeddings and devise new fusion techniques to improve emotion classification results.

5. CONCLUSION

This paper investigates the feasibility of employing GARG embeddings features and EDA/GSR Statistical Features (SF) to build an emotional state classification system. We proposed an effective knowledge graph method that uses gender and age information that is mostly ignored in feature analysis to predict participants' emotional states. We then extract SFs from the EDA/GSR data consistent with cognitive research. Finally, we introduced a weighted fusion strategy that uses GARG embeddings as a weight to SFs to improve classification performance. We evaluated our study on PAFEW and DEAP datasets. The results obtained show the superiority of our approach when compared to previous works.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. The data can be found here: <https://sites.google.com/view/emotiw2020> and <https://www.eecs.qmul.ac.uk/mmv/datasets/deap/>.

REFERENCES

- Ayata, D., Yaslan, Y., and Kamasak, M. (2017). "Emotion recognition via random forest and galvanic skin response: comparison of time based feature sets, window sizes and wavelet approaches," in *2016 Medical Technologies National Conference, TIPTEKNO 2016* (Anatalya: Institute of Electrical and Electronics Engineers Inc.), 1–4.
- Betella, A., Zucca, R., Cetnarski, R., Greco, A., Lanatà, A., Mazzei, D., et al. (2014). Inference of human affective states from psychophysiological measurements extracted under ecologically valid conditions. *Front. Neurosci.* 8, 1–19. doi: 10.3389/fnins.2014.00286
- Chen, J., Li, H., Ma, L., Bo, H., Soong, F., and Shi, Y. (2021). Dual-threshold-based microstate analysis on characterizing temporal dynamics of affective process and emotion recognition from EEG signals. *Front. Neurosci.* 15, 689791. doi: 10.3389/fnins.2021.689791
- Chen, X., Jia, S., and Xiang, Y. (2020). A review: knowledge reasoning over knowledge graph. *Expert. Syst. Appl.* 141, 112948. doi: 10.1016/j.eswa.2019.112948
- Cui, Z., Song, T., Wang, Y., and Ji, Q. (2020). "Knowledge augmented deep neural networks for joint facial expression and action unit recognition," in *34th Conference on Neural Information Processing Systems (NeurIPS 2020)* (Vancouver, BC), 1–12.
- De Nadai, S., D'Inca, M., Parodi, F., Benza, M., Trotta, A., Zero, E., et al. (2016). "Enhancing safety of transport by road by on-line monitoring of driver emotions," in *2016 11th Systems of Systems Engineering Conference*, (Konsberg), 1–4.
- Demš, J. (2006). Statistical comparisons of classifiers over multiple data sets. *J. Mach. Learn. Res.* 7, 1–30. doi: 10.5555/1248547.1248548
- Farashi, S., and Khosrowabadi, R. (2020). EEG based emotion recognition using minimum spanning tree. *Phys. Eng. Sci. Med.* 43, 985–996. doi: 10.1007/s13246-020-00895-y

AUTHOR CONTRIBUTIONS

HP, XXi, KG, and XXu proposed the idea. HP conducted the experiment, analyzed the results, and wrote the manuscript. XXi was in charge of technical supervision and provided revision suggestions. KG analyzed the results, reviewed the article, and was in charge of technical supervision. XXu was in charge of technical supervision and funding. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported in part by the National Natural Science Foundation of China under Grant U1801262, Guangdong Provincial Key Laboratory of Human Digital Twin (2022B1212010004), in part by the Key-Area Research and Development Program of Guangdong Province, China, under Grant 2019B010154003, in part by Natural Science Foundation of Guangdong Province, China, under Grant 2020A1515010781, and Grant 2019B010154003, in part by the Guangzhou key Laboratory of Body Data Science, under Grant 201605030011, in part by Science and Technology Project of Zhongshan, under Grant 2019AG024, in part by the Fundamental Research Funds for Central Universities, SCUT, under Grant 2019PY21, Grant 2019MS028, and Grant XYMS202006, and in part by Guangdong Philosophy and Social Sciences Planning Project, under Grant GD20CYS33.

- Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *J. Am. Stat. Assoc.* 32, 675–701. doi: 10.1080/01621459.1937.10503522
- Ganapathy, N., Veeranki, Y. R., and Swaminathan, R. (2020). Convolutional neural network based emotion classification using electrodermal activity signals and time-frequency features. *Expert. Syst. Appl.* 159, 113571. doi: 10.1016/j.eswa.2020.113571
- Gebauer, L., Skewes, J., Westphael, G., Heaton, P., and Vuust, P. (2014). Intact brain processing of musical emotions in autism spectrum disorder, but more cognitive load and arousal in happy vs. sad music. *Front. Neurosci.* 8:192. doi: 10.3389/fnins.2014.00192
- Guan, N., Song, D., and Liao, L. (2019). Knowledge graph embedding with concepts. *Knowledge Based Syst.* 164, 38–44. doi: 10.1016/j.knsys.2018.10.008
- Guo, R., Li, S., He, L., Gao, W., Qi, H., and Owens, G. (2013). "Pervasive and unobtrusive emotion sensing for human mental health," in *Proceedings of the 2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops* (Venice: PervasiveHealth), 436–439.
- Hassan, M. M., Alam, M. G. R., Uddin, M. Z., Huda, S., Almogren, A., and Fortino, G. (2019). Human emotion recognition using deep belief network architecture. *Inf. Fusion* 51, 10–18. doi: 10.1016/j.inffus.2018.10.009
- Hogan, A., Blomqvist, E., Cochez, M., D'Amato, C., Melo, G. D., Gutierrez, C., et al. (2021). Knowledge graphs. *ACM Comput. Surveys* 54, 1–37. doi: 10.1145/3447772
- Huang, Y., Chen, F., Lv, S., and Wang, X. (2019). Facial expression recognition: a survey. *Symmetry* 11, 1110–1189. doi: 10.3390/sym11101189
- Ji, S., Pan, S., Cambria, E., Martinen, P., and Yu, P. S. (2022). A survey on knowledge graphs: representation, acquisition, and applications. *IEEE Trans. Neural Netw. Learn. Syst.* 33, 494–514. doi: 10.1109/TNNLS.2021.3070843
- Jia, Y., Qi, Y., Shang, H., Jiang, R., and Li, A. (2018). A practical approach to constructing a knowledge graph for cybersecurity. *Engineering* 4, 53–60. doi: 10.1016/j.eng.2018.01.004

- Koelstra, S., Mühl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., et al. (2012). DEAP: a database for emotion analysis using physiological signals. *IEEE Trans. Affect. Comput.* 3, 18–31. doi: 10.1109/T-AFFC.2011.15
- Li, C., Xian, X., Ai, X., and Cui, Z. (2020). Representation learning of knowledge graphs with embedding subspaces. *Scientific Program*. 2020, 10. doi: 10.1155/2020/4741963
- Li, J., Li, S., Pan, J., and Wang, F. (2021). Cross-subject EEG emotion recognition with self-organized graph neural network. *Front. Neurosci.* 15, 611653. doi: 10.3389/fnins.2021.611653
- Liu, Y., Gedeon, T., Caldwell, S., Lin, S., and Jin, Z. (2020). *Emotion Recognition Through Observer's Physiological Signals*. Canberra, ACT: ACM.
- Liu, Y. J., Yu, M., Zhao, G., Song, J., Ge, Y., and Shi, Y. (2018). Real-time movie-induced discrete emotion recognition from EEG signals. *IEEE Trans. Affect. Comput.* 9, 550–562. doi: 10.1109/TAFFC.2017.2660485
- Mano, L. Y., Mazzo, A., Neto, J. R., Meska, M. H., Giancristofaro, G. T., Ueyama, J., et al. (2019). Using emotion recognition to assess simulation-based learning. *Nurse Educ. Pract.* 36, 13–19. doi: 10.1016/j.nepr.2019.02.017
- Milstein, N., and Gordon, I. (2020). Validating measures of electrodermal activity and heart rate variability derived from the empathica E4 utilized in research settings that involve interactive dyadic states. *Front. Behav. Neurosci.* 14, 148. doi: 10.3389/fnbeh.2020.00148
- Nguyen, T., Nguyen, G. T., Nguyen, T., and Le, D. H. (2022). Graph convolutional networks for drug response prediction. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 19, 146–154. doi: 10.1109/TCBB.2021.3060430
- Qin, C., Zhu, H., Zhuang, F., Guo, Q., Zhang, Q., Zhang, L., et al. (2020). A survey on knowledge graph-based recommender systems. *SCIENTIA SINICA Inform.* 50, 937–956. doi: 10.1360/SSI-2019-0274
- Reed, C. L., Moody, E. J., Mgrublian, K., Assaad, S., Schey, A., and McIntosh, D. N. (2020). Body matters in emotion: restricted body movement and posture affect expression and recognition of status-related emotions. *Front. Psychol.* 11, 1961. doi: 10.3389/fpsyg.2020.01961
- Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., et al. (2018). A review of emotion recognition using physiological signals. *Sensors* 18, 2074. doi: 10.3390/s18072074
- Shukla, J., Barreda-Angeles, M., Oliver, J., Nandi, G. C., and Puig, D. (2019). Feature extraction and selection for emotion recognition from electrodermal activity. *IEEE Trans. Affect. Comput.* 3045, 1–1. doi: 10.1109/TAFFC.2019.2901673
- Singh, L., Singh, S., and Aggarwal, N. (2019). Improved TOPSIS method for peak frame selection in audio-video human emotion recognition. *Multimed Tools Appl.* 78, 6277–6308. doi: 10.1007/s11042-018-6402-x
- Soleymani, M., Villaro-Dixon, F., Pun, T., and Chanel, G. (2017). Toolbox for emotional feature extraction from physiological signals (TEAP). *Front. ICT* 4, 1. doi: 10.3389/fict.2017.00001
- Song, T., Zheng, W., Lu, C., Zong, Y., Zhang, X., and Cui, Z. (2019). MPED: A multi-modal physiological emotion database for discrete emotion recognition. *IEEE Access* 7, 12177–12191. doi: 10.1109/ACCESS.2019.2891579
- Sreeshakthy, M., and Preethi, J. (2016). Classification of human emotion from deap EEG Signal using hybrid improved neural networks with cuckoo search. *Broad Res. Artif. Intell. Neurosci.* 6, 60–73.
- Stuldreher, I. V., Thammasan, N., van Erp, J. B., and Brouwer, A. M. (2020). Physiological synchrony in EEG, electrodermal activity and heart rate detects attentionally relevant events in time. *Front. Neurosci.* 14, 575521. doi: 10.3389/fnins.2020.575521
- Toutanova, K., and Chen, D. (2015). “Observed versus latent features for knowledge base and text inference,” in *Proceedings in 3rd Workshop on Continuous Vector Space Models and Their Compositionality* Beijing, 1–10.
- Trouillon, T., Welbl, J., Riedel, S., Ciaussier, E., and Bouchard, G. (2016). “Complex embeddings for simple link prediction,” in *33rd International Conference on Machine Learning*, vol. 5, (New York, NY), 2071–2080.
- Tsiourti, C., Weiss, A., Wac, K., and Vincze, M. (2019). Multimodal integration of emotional signals from voice, body, and context: effects of (in)congruence on emotion recognition and attitudes towards robots. *Int. J. Soc. Rob.* 11, 555–573. doi: 10.1007/s12369-019-00524-z
- Wang, Z. M., Hu, S. Y., and Song, H. (2019). Channel selection method for EEG emotion recognition using normalized mutual information. *IEEE Access* 7, 143303–143311. doi: 10.1109/ACCESS.2019.2944273
- Wilaprasitporn, T., Dithaporn, A., Matchaparn, K., Tongbuasirilai, T., Banluesombatkul, N., and Chuangsuwanich, E. (2020). Affective EEG-based person identification using the deep learning approach. *IEEE Trans. Cognit. Dev. Syst.* 12, 486–496. doi: 10.1109/TCDS.2019.2924648
- Yang, J., Wang, R., Guan, X., Hassan, M. M., Almogren, A., and Alsanad, A. (2020). AI-enabled emotion-aware robot: the fusion of smart clothing, edge clouds and robotics. *Future Generat. Comput. Syst.* 102, 701–709. doi: 10.1016/j.future.2019.09.029
- Yang, K., Zheng, Y., Lu, K., Chang, K., Wang, N., Shu, Z., et al. (2022). PDGNet: predicting disease genes using a deep neural network with multi-view features. *IEEE/ACM Trans. Comput. Biol. Bioinform.* 19, 575–584. doi: 10.1109/TCBB.2020.3002771
- Yao, Z., Wang, Z., Liu, W., Liu, Y., and Pan, J. (2020). Speech emotion recognition using fusion of three multi-task learning-based classifiers: HSF-DNN, MS-CNN and LLD-RNN. *Speech Commun.* 120, 11–19. doi: 10.1016/j.specom.2020.03.005
- Yu, W., Ding, S., Yue, Z., and Yang, S. (2020). “Emotion recognition from facial expressions and contactless heart rate using knowledge graph,” in *Proceedings-11th IEEE International Conference on Knowledge Graph, ICKG 2020* (Nanjing: IEEE), 64–69.
- Zhang, X., Liu, J., Shen, J., Li, S., Hou, K., Hu, B., et al. (2020). Emotion recognition from multimodal physiological signals using a regularized deep fusion of kernel machine. *IEEE Trans. Cybern.* 51, 4386–4399. doi: 10.1109/TCYB.2020.2987575
- Zheng, W., Yan, L., Wang, F.-Y., and Gou, C. (2020). Learning from the past: meta-continual learning with knowledge embedding for jointly sketch, cartoon, and caricature face recognition. 736–743.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Perry Fordson, Xing, Guo and Xu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.