

# 情绪识别技术介绍

## I. 基本介绍

情绪是一个人在某种环境状态下的行为和心理综合表现。从人体的表现角度来看,情绪的体现主要分为神经生理层面和外在行为层面。

情绪表达方式多种多样,情绪的体现途径也非常丰富。威廉·詹姆斯被认为是美国心理学的奠基人之一,他提出情绪是人们对自身身体变化的感知。情绪的感知往往伴随着特定的行为表现,包括面部表情、肌肉紧张和内脏活动等。因此,通过观察这些行为表现,我们可以推测情绪的变化。

目前,情绪识别主要采用两种方式。一种是基于非生理信号,例如面部表情、语音声调、身体姿态等,但这些信号有时可能受到人为伪装的影响,难以准确反映真实情绪状态。另一种方式则是基于生理信号,如脑电信号、眼动数据、心电信号、肌电信号、皮肤电反应等,这些信号是由人体的神经系统和内分泌系统自发产生的,不容易受到人为因素的干扰,因此更有利于准确地反映人类的情绪状态。基于生理信号的情绪识别方法能够提供更客观真实的结果,因此在实际应用中具有广泛的潜力。

### 基于非生理信号的情绪识别方法

#### 1. 文本情绪识别

文本情绪识别是一个广泛研究的领域,自然语言处理(NLP)一直是一个不断发展的领域。在线情绪分析是通过用户的书面文本和网络活动来解释用户心理状态的最传统和流行的方法之一。

传统上,通过文本进行情绪识别是通过选择情绪关键词、词袋和 N-gram 来完成的。但是关键字可能会也可能不会出现在给定的句子中。因此,为了克服这个问题,引入了一种称为基于知识的人工神经网络的替代方法,其中使用本体中的单词含义作为特征。随着自然语言处理(NLP)的不断发展,文本情绪识别也变成了一个广泛研究的领域。在线情感分析是通过用户在网络

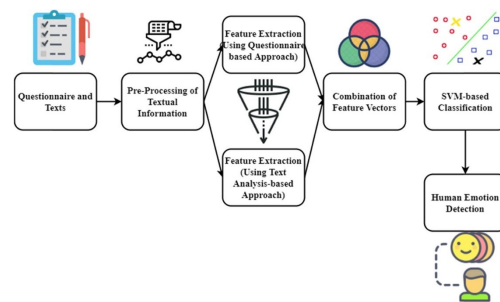


图 1. 文本情绪识别处理过程

上的书面文本和活动来解读用户心理状态的一种常规和流行的方式之一。

#### 2. 语音信息与情绪识别

通过语音进行的情绪识别是通过音频信号的变化来完成的。语音情绪识别具有多种应用,如语音识别、呼叫中心、客户服务等。语音情绪识别基本上分为特征提取和分类两个步骤。

卷积神经网络(CNN)方法: 在[107]中提出的方法使用了卷积神经网络,分为两个阶段。首先,它使用稀疏自编码器(SAE)和重构分离来计算未标记样本的局部不变特征(LIFs)。然后,在第二阶段,将LIF作为输入进行特征提取。

半卷积神经网络(semi-CNN)方法: [108]中的方法首先使用半卷积神经网络,然后使用收缩卷积网络进行特征提取。特征选择方法: [113]中描述了一种用于减少特征维度的特征选择方法,使用粒子群优化辅助的生物地理学算法在不同的数据集上进行模拟,获得高精度的分类结果。

#### 3. 面部表情与情绪识别

面部表情与情绪识别是一种利用人类的面部肌肉运动来检测和分类人类情绪的技术。面部表情是人类表达情绪的一种重要方式,它可以反映出人类的内在感受和 psychological 状态。面部表情与情绪识别的一般流程包括以下几个步骤:

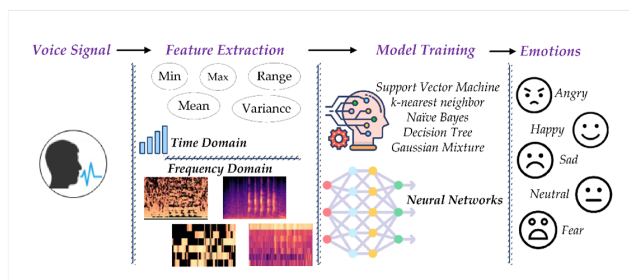


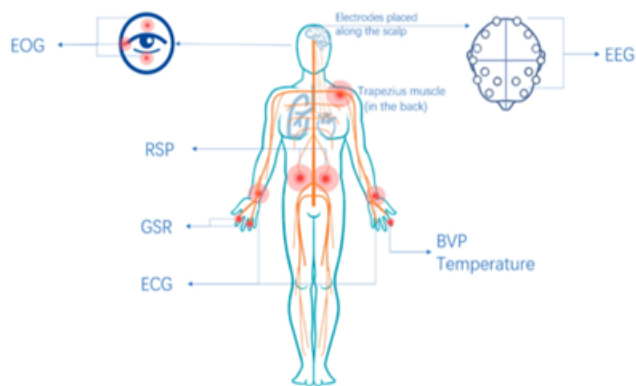
图 2. 语音情绪识别处理过程

- 1) 面部检测：从图像或视频中定位和分割出人脸区域，可以使用一些常见的算法，如Haar特征分类器、级联卷积神经网络等。
- 2) 面部对齐：对检测出的人脸区域进行旋转、缩放、裁剪等操作，使其与标准模板对齐，可以使用一些常见的算法，如仿射变换、活动外形模型等。
- 3) 面部特征提取：从对齐后的人脸区域中提取有助于区分不同情绪的特征，例如面部特征点、局部二值模式、深度学习特征等。
- 4) 情绪分类：使用机器学习或深度学习等方法，根据提取的面部特征，将人脸区域分为不同的情绪类别，如高兴、悲伤、愤怒、恐惧等。



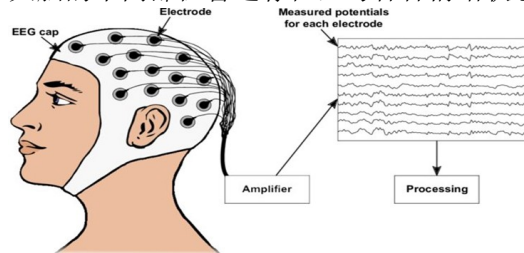
### 基于生理信号的情绪识别方法

人类的情绪是一种自发的心理状态，通过人体的生理变化来反映，并且显著影响我们的意识。威廉·詹姆斯首先提出了情绪生理学理论，他认为外部刺激会引发自主神经系统的活动，并在大脑中产生生理反应。例如，当我们感到高兴时，我们会笑；当我们感到害怕时，我们的汗毛会竖立。在情绪识别中，涉及到许多生理信号，如脑电图 (EEG)、心电图 (ECG)、皮肤电反应 (GSR)、血容量脉搏 (BVP) 和眼电图 (EOG)，如图所示。

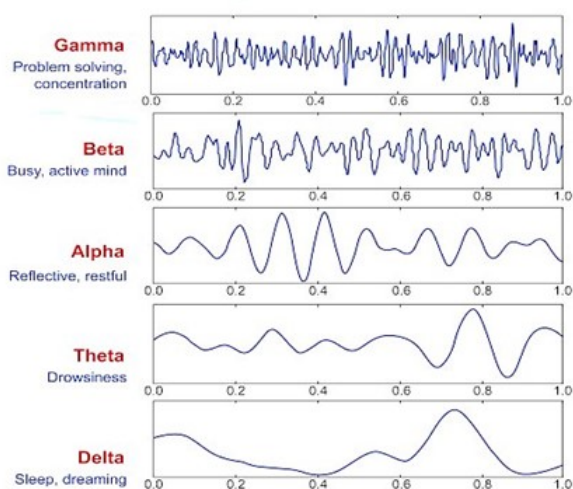


生理信号有助于更好地理解参与者表达的潜在情感反应，这些与来自中央和自主神经系统的多通道记录相对应。在大脑上记录的测量数据也使我们能够观察到人们感受到的情绪。脑电图因为其无创、快速、廉价等优点，使其成为研究大脑对情绪刺激的首选方法。如今，由于其可穿戴性、价格、便携性和易用性，新的无线脑电图设备正在进入市场。本部分就主要介绍基于EEG的情绪识别。

EEG 脑电图 (EEG) 是一种电生理学的无创技术，用于记录来自人脑的电活动。该技术的首次应用报告，由德国精神病学家汉斯·贝格尔于 1924 年在人类身上进行。通常使用一种特殊设备称为脑电图仪来收集 EEG 信号。该设备的主要部分是特殊的金属板电极，应放置在人的头皮上。为了评估人类情绪，通常通过 EEG 信号中的五个频率范围来测量和分析大脑对各种刺激的反应，即： $\delta$  波、 $\theta$  波、 $\alpha$  波、 $\beta$  波和 $\gamma$  波。这些频带波在大脑的不同部位普遍存在，与各种情绪状态相关。



一般来说，高兴的情绪会增加 $\beta$  波和 $\gamma$  波的幅度和频率，表明大脑处于高度警觉和认知的状态；悲伤的情绪会增加 $\alpha$  波和 $\theta$  波的幅度和频率，表明大脑处于放松和潜意识的状态；愤怒的情绪会增加 $\delta$  波和 $\theta$  波的幅度和频率，表明大脑处于深度睡眠或昏迷的状态；恐惧的情绪会增加 $\beta$  波和 $\gamma$  波的幅度和频率，表明大脑处于高度警觉和认知的状态。



下面选择了两篇较为经典的关于EEG-based emotion recognition的论文。

#### A. Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks

这篇文章 [2]是关于使用一种新颖的动态图卷积神经网络(DGCNN)来识别多通道EEG信号中的情绪状态在2015年发表于*IEEE Transactions on Autonomous Mental Development*, 应用次数1340次

该论文的作者是上海交通大学计算机科学与工程系郑伟龙和陆宝梁教授, 陆教授曾任亚太神经网络大会\*(APNNA)主席, 并于18年担任第2011届神经信息处理国际会议主席。

这篇文章是关于使用深度信念网络(DBN)来构建基于EEG信号的三类情绪(积极、中性和消极)识别模型的。文章的主要贡献和创新点有:

通过分析训练后的DBN的权重分布, 研究了不同的电极数量对情绪识别的影响, 确定了比全通道更优的电极位置和关键频带。实验结果表明, 使用12个通道的电极布局可以达到最高的识别精度和最低的标准差, 且与积极、中性和消极情绪相关的神经特征确实存在, 而且他们在个体中具有共性。

将DBN模型与基于特征的浅层模型(如SVM、LR和KNN)进行了比较, 发现DBN模型具有更高的识别性能和更低的计算成本, 因为它可以绕过人工的特征提取和选择, 利用无监督学习自动进行特征选择和提取。

文章的主要流程如下:

首先, 设计了一种实验方案, 要求15名被试在观看三种类型的情感电影时引出自己的情绪, 并记录了他们的62通道EEG信号和正面面部视频。

然后, 对原始的EEG信号进行预处理, 包括降采样、滤波、去除肌电和眼电干扰等操作, 并将每个通道划分为1秒等长无重叠的片段。

接着, 对每个EEG片段进行小波变换, 提取不同频带( $\delta$ 、 $\theta$ 、 $\alpha$ 、 $\beta$ 和 $\gamma$ )下的微分熵(DE)特征, 以及DE特征的差值和比值(DASM、RASM和DCAU)。

然后, 将所有通道的DE特征向量拼接成一个矩阵, 作为DBN模型的输入。DBN模型由三个受限玻尔兹曼机(RBM)和一个逻辑回归层组成, 使用无监督预训练和有监督微调两个步骤来训练。

最后, 在测试集上进行情绪识别, 并分析训练后的DBN的权重分布, 选择不同数量和位置的电极布局, 并与其他分类器进行比较。

#### B. EEG Emotion Recognition Using Dynamical Graph Convolutional Neural Networks

该文章 [1]发表于2018年*IEEE Transactions on Affective Computing*, 作者为东南大学的宋腾飞。引用次数726

该文章肯定了上海交通大学的深度神经网络(DNN)在EEG情感分类中的成功应用。

文本提到了在文献中有许多方法用于处理EEG情感分类问题, 其中使用深度神经网络(DNN)的方法被证明是最成功的之一。卷积神经网络(CNN)在图像、视频和语音分类中的应用: CNN是一种著名的DNN方法, 广泛用于图像分类、对象检测、跟踪和分割等各种分类问题。

然而, CNN主要用于连续变化的图像、视频和语音等本地特征学习问题。传统CNN方法在离散和不连续空间领域的限制: 传统的CNN方法在处理离散和不连续的信号数据, 如交通网络和脑网络数据时, 可能不太适用。与传统的图卷积神经网络(GCNN)方法不同, 该文章所提出的DGCNN方法可以通过训练神经网络, 动态学习由邻接矩阵表示的不同脑电图(EEG)通道之间的内在关系, 从而有利于更具判别性的脑电特征提取。该文章使用了上一篇文章创建的SJTU Emotion EEG database测试, 还与上交的实验结果进行了对比分析。文章的主要贡献和创新点有:

Feature	Method	$\delta$ band	$\theta$ band	$\alpha$ band	$\beta$ band	$\gamma$ band	all ( $\delta, \theta, \alpha, \beta, \gamma$ )
DE	DBN [18]	64.32 / 12.45	60.77 / 10.42	64.01 / 15.97	78.92 / 12.48	79.19 / 14.58	86.08 / 8.34
	SVM [18]	60.50 / 14.14	60.05 / 10.20	66.64 / 14.41	80.76 / 11.56	79.56 / 11.38	83.99 / 9.72
	GCNN	72.75 / 10.85	<b>74.40 / 8.23</b>	73.46 / 12.17	83.24 / 9.93	83.36 / 9.43	87.40 / 8.20
	DGCNN	<b>74.25 / 11.42</b>	71.52 / 5.99	<b>74.43 / 12.10</b>	<b>83.65 / 10.17</b>	<b>85.73 / 10.64</b>	<b>90.40 / 8.49</b>
PSD	DBN [18]	60.05 / 16.66	55.03 / 13.88	52.79 / 15.38	60.68 / 21.31	63.42 / 19.66	61.90 / 16.65
	SVM [18]	58.03 / 15.39	57.26 / 15.09	59.04 / 15.75	73.34 / 15.20	71.24 / 16.38	59.60 / 15.93
	GCNN	69.89 / 13.83	70.92 / 9.18	73.18 / 12.74	76.21 / 10.76	76.15 / 10.09	81.31 / 11.26
	DGCNN	<b>71.23 / 11.42</b>	<b>71.20 / 8.09</b>	<b>73.45 / 12.25</b>	<b>77.45 / 10.81</b>	<b>76.60 / 11.83</b>	<b>81.73 / 9.94</b>
DASM	DBN [18]	48.79 / 9.62	51.59 / 13.98	54.03 / 17.05	69.31 / 15.22	70.06 / 18.14	72.73 / 15.93
	SVM [18]	48.87 / 10.49	53.02 / 12.76	59.81 / 14.67	<b>75.03 / 15.72</b>	<b>73.50 / 16.57</b>	72.81 / 16.57
	GCNN	<b>57.07 / 6.75</b>	54.80 / 9.09	62.97 / 13.43	74.97 / 13.40	73.28 / 13.67	76.00 / 13.32
	DGCNN	55.93 / 9.14	<b>56.12 / 7.86</b>	<b>64.27 / 12.72</b>	73.61 / 14.35	73.50 / 16.6	<b>78.45 / 11.84</b>
RASM	DBN [18]	48.05 / 10.37	50.62 / 14.02	56.15 / 15.28	70.31 / 15.62	68.22 / 18.09	71.30 / 16.16
	SVM [18]	47.75 / 10.59	51.40 / 12.53	60.71 / 14.57	74.59 / 16.18	74.61 / 15.57	74.74 / 14.79
	GCNN	<b>59.70 / 5.65</b>	<b>55.91 / 8.82</b>	59.97 / 14.27	<b>79.45 / 13.32</b>	79.73 / 13.22	84.06 / 12.86
	DGCNN	57.79 / 6.90	55.79 / 8.10	<b>61.58 / 12.63</b>	75.79 / 13.07	<b>82.32 / 11.54</b>	<b>85.00 / 12.47</b>
DCAU	DBN [18]	54.58 / 12.81	56.94 / 12.54	57.62 / 13.58	70.70 / 16.33	72.27 / 16.12	77.20 / 14.24
	SVM [18]	55.92 / 14.62	57.16 / 10.77	61.37 / 15.97	75.17 / 15.58	76.44 / 15.41	77.38 / 11.98
	GCNN	62.60 / 12.88	<b>65.05 / 8.35</b>	66.41 / 11.06	77.28 / 11.55	78.68 / 13.00	79.02 / 11.27
	DGCNN	<b>63.18 / 13.48</b>	62.55 / 7.96	<b>67.71 / 10.74</b>	<b>78.08 / 10.81</b>	<b>80.05 / 13.03</b>	<b>81.91 / 10.06</b>

图 3. 实验数据对比

- 1) 提出了一种基于神经网络的方法来动态地学习不同EEG通道之间的关系，用一个邻接矩阵来表示。这样可以有效地处理不同个体之间的 $EEG$ 信号差异，提高情绪识别的准确性和鲁棒性。
- 2) 将学习到的邻接矩阵作为图卷积神经网络( $GCNN$ )的输入，利用图结构来提取更具判别力的 $EEG$ 特征。 $GCNN$ 是一种能够处理非欧几里得数据()的深度学习模型，可以捕捉 $EEG$ 信号中的空间和频域信息；
- 3) 它在两个公开的情绪数据集( $SEED$ )和 $DREAMER$ 上进行了广泛的实验，与现有的方法进行了对比。实验结果表明，该方法在主观依赖和主观独立的情况下都取得了更好的识别性能，分别达到了90.4

## 参考文献

- [1] Tengfei Song, Wenming Zheng, Peng Song, and Zhen Cui. Eeg emotion recognition using dynamical graph convolutional neural networks. *IEEE Transactions on Affective Computing*, 11(3):532–541, 2018.
- [2] Wei-Long Zheng and Bao-Liang Lu. Investigating critical frequency bands and channels for eeg-based emotion recognition with deep neural networks. *IEEE Transactions on autonomous mental development*, 7(3):162–175, 2015.