

Title

Respiration-based emotion recognition with deep learning

Paper Background

当前随着人工智能的发展，通过生理上的信号来进行情感计算是非常有前景的。生理信号可以通过一些可穿戴式装置，如心电图、脑电图、肌动电流图、血压等进行探测。这些信号是人体无意识的表现，无法特意去左右，而且它所包含的信息又与人类的情绪息息相关。因此依托于生理信号的情绪识别也是大有前途。

但是，情绪识别也存在着很多困难。首先人类的情绪会被很多因素所影响，并且不同的情绪之间的边界也很模糊。其次虽然不同的特征能够去尽力秒回生理信号的相关情绪特征，但是手工的特征提取也有很多的限制性，例如它需要依赖于工作人员的经验，匮乏的专业知识会导致不能捕捉正确的信号，不能描述一些信号的细节，会丢失一部分信息。

所幸当前的研究也完成了一些不同种类的情绪识别和测试它们的数据集。而深度学习则能够自动的从原生信号中去获得所需要的特征。

The Current Development

Model

2001年，Picard教授应用人工智能去识别人类的情绪情况，在八个情绪类别上达到81%的正确率。

Duan教授图书微分熵来代表EEG(脑电图学)特征相关的情绪情况，达到了81.17%的平均正确率。

Giakoumis教授提出Legendre and Krawtchouk moments 进行提取生物特征。

Yannakakis和Hallam教授使用近似熵特征和偏好学习

Lin教授应用机器学习算法来将EEG信号进行分类，并且在四个情绪类别中达到82.29%的平均正确率。

Wang教授为情绪分类系统的比较了三种EEG特征（功率谱特征，小波特征，非线性动力特征）。

Deep Learning as the Feature Extraction Method

2013年，Martinez教授介绍了卷积神经网络(CNN)来建立情感上的生理模型。Zheng教授训练了一个深度信念网络(DBN)去将EEG的数据分为两个情绪种类（高和低的 valence ）。Jirayucharoensak教授使用了稀疏自编码器，它的输入数据来自32位通道的EEG信号。Martin教授比较了手工的特征和结合深度信念网络(DBN)与隐藏马尔科夫模型(HMM)所得到的特征。

Paper Content

这篇文章主要研究的是在只通过呼吸信号去找到它检测心理活动的的能力。

Arousal-valence Emotion Theory

呼吸信号包含了丰富的信息而且很容易通过佩戴装置去探测。因此，在这篇文章中作者通过呼吸信号去关注情绪识别，他们使用了Russel的环状情绪理论。

该理论将情绪的情况在一个二维圆形空间中划分为arousal和valence维度。arousal是纵轴，刻画的是情感激活的强度；valence是横轴，刻画的是消极或积极的情绪。

- Low Valence , High Arousal : 左上象限
- High Valence , High Arousal : 右上象限
- Low Valence , Low Valence : 左下象限
- High Valence , Low Arousal : 右下象限

Deep Learning Framework

该文采用稀疏自编码器将原生信号转换为所提取的特征，再采用两个逻辑回归去预测情感情况。

Sparse Auto-Encoder

稀疏自编码器是用来从未标识的数据中去自动学习特征的深度学习方法。首先，作为SAE的基础元素，一个自动编码单元AE通过编码将输入数据转换成隐藏表现（提取特征）。然后AE也可以通过解码将隐藏表现退回到重构的输入数据，这与原生的输入数据很相似，尽管重构的输入数据会存在很小的误差。这个结构是为了在输入层中找到通用的表现。

编码器将输入数据映射到隐藏层中： $h(x)=f(W1*x+b1)$ ， $f(x)$ 是一个激活函数，实现非线性化（文章采用的是sigmoid函数）， $W1$ 为一个权重矩阵， $b1$ 是一个偏置顶。

解码器将隐藏层映射到重构的输入数据： $x=f(W2*h(x)+b2)$

为了最小化原始数据与重构后的差异，文章采用反向传播来调节 $W1$ 和 $W2$ ， $b1$ 和 $b2$ 的值，并且通过惩罚项来使隐藏单元的期望活动变得更稀疏。

Logistic Regression

在通过稀疏自编码器来提取特征后，文章使用两个逻辑回归来识别Arousal和Valence。文章期望的假设是去估计每一个数据在每一类的可能性（一共有四类），所以输出的是一个二维向量。向量中每一个值代表的是在每一类的权重值（可能性），所以之和为1。

Experiment

文章采用的训练数据集是使用生理信号的情绪分析的数据集（DEAP），将80%的样本作为训练集，20%的样本作为测试集。在完成模型的建立后，使用奥格斯堡大学的数据集去预测，来估计这个模型能力（准确度）

最后，模型训练的准确率为：Valence类别的准确率达到了73.06%,Arousal类别的准确率达到了80.78%。预测奥格斯堡大学的数据集的准确率为：Valence类别的准确率达到了85.89%,Arousal类别的准确率达到了83.72%。

Contribution

该文章的主要贡献主要有两个方面：

- 1、考虑到深度学习特征提取和特征选择的性能，作者介绍了稀疏自编码器去提取特征来进行情绪识别。
- 2、该实验证明了在情绪识别中仅凭呼吸信号是有能力去完成的，不一定非要依赖于复杂的脑电图学（EEG），呼吸信号的研究和应用在情绪识别中是非常有潜力的

Vocabulary

General Vocabulary

inappropriate：不适当的

depict：描述

derive：获得

feasibility：可行性

architecture：架构

reconstruct：重构

quadrant：象限

convert：转换

participant：参与者

preliminary：初步的

paradigm：范例

intrinsic：本质的

optimal：最佳的

manual：手工的

implement：实施

velocity：速率

validate：验证

intensity：强度

deliver：实现

spontaneous：无意识的

overlap：重叠

property：性能

seize：抓住

Specialized Vocabulary

sparse auto-encoder(SAE) : 稀疏自编码器
differential entropy : 微分熵
wavelet feature : 小波特征
动力特征
normalized magnitude : 归一化幅度
penalty term : 惩罚项
standard deviation : 标准差

logistic regression : 逻辑回归
approximate entropy : 近似熵
nonlinear dynamical feature : 非线性

back-propagation : 反向传播
cost function : 代价函数
threshold value : 阈值

Others

electrocardiogram(ECG) : 心电图
学
electromyogram(EMG) : 肌动电流图
galvanic skin response(GSR) : 皮电反应
photoplethysmogram(PPG) : 血管容积图
表现
power spectrum feature : 功率谱特征
cardiovascular system : 心血管系统
体系
plethysmography : 体积描记
cortical and subcortical reegion : 大脑皮层及皮下领域

electroencephalography(EEG) : 脑电图

blood volume pressure(BVP) : 血压
respiration pattern(RSP) : 呼吸模式
involuntary manifestation : 无意识的表

skin resistance : 表面阻力
neurophysiological system : 神经心理学

oscillation band : 振荡带