具有互补特征学习框架和注意力特征融合模块的语音情感识别模型*

黄佩瑶[1]，程慧慧[2]，唐小煜[1,2†]

(1. 华南师范大学工学部 电子与信息工程学院，广东 佛山 528225；2. 华南师范大学 物理学院，广东 广州 510006)

摘　要： 针对深度学习的特征提取方法无法全面提取语音中的情感特征，也无法有效地融合这些特征的问题，提出了一种集成互补特征学习框架和注意力特征融合模块的语音情感识别模型. 该互补特征学习框架包含两条独立的表征提取分支和一条交互互补表征提取分支，能够全面覆盖情感特征的独立性表征和互补性表征. 为了进一步优化模型性能，引入注意力特征融合模块，该模块能够根据不同表征对情感分类的贡献程度分配合适的权重，使模型能最大程度地关注对情感识别最有助的特征. 基于两个公开情感数据库（Emo-DB和IEMOCAP）的仿真实验结果，验证了所提模型的鲁棒性和有效性.

关键词： 语音情感识别；深度神经网络；情感特征表征；特征提取器；特征融合；注意力机制；人工智能

DOI： 10.13568/j.cnki.651094.651316. 2023.07.05.0002

中图分类号： TN912；TP391　**文献标识码：** A　**文章编号：** 2096-7675(2024)01-0052-07

引文格式： 黄佩瑶，程慧慧，唐小煜. 具有互补特征学习框架和注意力特征融合模块的语音情感识别模型[J]. 新疆大学学报(自然科学版)(中英文)，2024，41(1)：52-58.

英文引文格式： HUANG Peiyao，CHENG Huihui，TANG Xiaoyu. Speech emotion recognition with complementary feature learning framework and attentional feature fusion module[J]. Journal of Xinjiang University(Natural Science Edition in Chinese and English)，2024，41(1)：52-58.

Speech Emotion Recognition with Complementary Feature Learning Framework and Attentional Feature Fusion Module

HUANG Peiyao[1], CHENG Huihui[2], TANG Xiaoyu[1,2]

(*1. School of Electronics and Information Engineering, Faculty of Engineering, South China Normal University, Foshan Guangdong 528225, China; 2. School of Physics, South China Normal University, Guangzhou Guangdong 510006, China*)

Abstract： Addressing the limitations of deep learning feature extraction methods, which fail to comprehensively extract and effectively integrate emotional features from speech, this paper proposes a novel speech emotion recognition model. It integrates a complementary feature learning framework and an attention feature fusion module. The complementary feature learning framework consists of two independent representational extraction branches and an interactive complementary representational extraction branch, thoroughly covering both independent and complementary representations of emotional features. To further optimize model performance, an attention feature fusion module is introduced. This module allocates appropriate weights based on the contribution level of different representations to emotion classification, enabling the model to focus maximally on features most beneficial for emotion recognition. Simulation experiments conducted on two public emotion databases (Emo-DB and IEMOCAP) validate the robustness and effectiveness of the proposed model.

Key words： speech emotion recognition; deep neural networks; emotional feature representation; feature extractor; feature fusion; attention mechanism; artificial intelligence

---

* **收稿日期：** 2023-07-05

基金项目： 国家自然科学基金“基于深度学习和能量采集的无线体域网高能效技术研究”（62001173）；广东省大学生科技创新人才培养专项资金项目“基于深度学习的水下目标检测系统”（pdjh2022a0131）.

作者简介： 黄佩瑶（1999—），女，硕士生，从事语音情感识别的研究，E-mail：pyaooo@qq.com.

† **通讯作者：** 唐小煜（1980—），男，博士，副教授，主要从事信息系统开发、机器视觉、智能控制和物联网的研究，E-mail：tangxy@scnu.edu.cn.

0　引 言

语言是人类最重要的交流媒介，除了语言信息以外，语音信号还承载着许多反映说话者情感的信息. 在人机交互[1−2]（Human-Machine Interaction, HMI）中，通过用户的语音信号对用户的情感进行识别是一个关键环节. 从语音信号中提取情感特征以进行情感分类的语音情感识别（Speech Emotion Recognition, SER）是人机交互中广泛应用的技术[3]. 语音情感识别面临的一大挑战是从语音信号中提取有效的情感特征，情感特征的有效性很大程度上影响了最终情感识别的准确率[4]. 当前许多语音情感识别的研究都面临缺乏具有可辨别性的情感特征的问题，这限制了整体模型的情感识别能力. 故本文针对先前研究中情感特征提取研究的不足，提出了互补特征学习框架（Complementary Feature Learning Framework, CFLF）和基于注意力机制的注意力特征融合模块（Attentional Feature Fusion Module, AFFM），该模块可从语音信号中获得更加全面的情感表征，提升整体模型情感识别的能力.

本文主要贡献包括三个方面：

1） 提出了CFLF，将梅尔倒谱系数[5]（Mel-Frequency Cepstral Coefficients, MFCCs）和使用openSMILE[6]提取的手工特征（Hand-Crafted Features, HCFs）分别输入卷积神经网络（Convolutional Neural Network, CNN）和循环神经网络（Recurrent Neural Network, RNN）分支中，以获得独立性表征；再将MFCCs和HCFs同时输入交互处理通道，以捕捉这两类特征的通道相关性和标记（token）相关性，从而获得高级的交互互补特征表征.

2）提出了基于注意力机制的AFFM. CFLF输出的表征通过注意力机制关注跨通道和跨token的信息，并生成注意力特征融合权重，最终得到融合特征.

3） 交互式情感二元动作捕捉数据库（IEMOCAP）和柏林情感数据库（Emo-DB）中的仿真实验证实所提SER模型具有优异的性能，其中非加权精度（Unweighted Accuracy, UA）和加权精度（Weighted Accuracy, WA）均得到了提升.

1　相关工作

提取语音信号中的情感特征是SER模型中十分重要的环节. 传统的SER模型常常使用低级HCFs[7−10]，这种特征是基于经验设计的，不足以表征情绪状态.

近年来深度学习方法被广泛应用于生成高级的情感特征表征，SER中常用方法有CNN、长短期记忆网络（Long Short Term Memory, LSTM）和RNN等. Jiang等[11]提出了具有频谱特征的并行卷积循环神经网络（PCRNN），捕捉情绪的细微变化. 为了充分利用不同特征的情感信息，许多研究者使用了双通道结构[11−14]，但未考虑不同特征的独立性. Zhong等[15]针对此问题，提出了独立训练框架并利用深度学习自动学习特征和经验特征的互补优势，但未考虑两种特征的相关性，且使用简单的连接操作融合独立表征. Liu等[16]使用全连接层进行特征融合，Jung等[17]在面部情绪识别任务中使用与Liu等[16]同样的全连接层进行融合，但使用联合微调方法独立训练全连接层，这两种融合方法优于简单的加权求和融合. Woo等[18]提出了轻量级通用前馈CNN注意力模块，能够有效融合情感特征，让模型集中在对分类贡献更大的特征上.

为了提升SER模型性能而使用多类特征，却缺乏对不同特征互补性的关注从而损失有效情感信息的问题[15]，和融合多路表征时未考虑不同表征对后续情感分类的贡献程度的问题[16−17]. 本文提出具有CFLF和AFFM的SER模型，能将互补特征（MFCCs和HCFs）的独立特征表征和具有交互性表征提取出来，再有效地特征融合. 这不仅保留了不同特征之间的独立性、通道相关性和token相关性，也通过注意力机制考虑了不同表征对情感分类的贡献程度，从而在表征融合时为不同表征分配适当权重. 由于本文着重于特征提取部分的优化，故后续的情感分类器使用支持向量机（Support Vector Machine, SVM）[19].

2　具有CFLF和AFFM的SER模型

本节介绍了具有CFLF和AFFM的SER模型的整体结构，包括输入的特征MFCCs和HCFs，特征提取部分的CFLF和AFFM. 图1展示了SER模型整体网络架构.

2.1　特征选择

SER模型输入时，许多研究者会使用不同的特征进行情感分类，如基于人耳听觉敏感性提出的MFCCs[20]和使用openSMILE提取的HCFs. 本文使用的MFCCs特征大小为400×13. 首先对语音信号进行预加重和平滑处

理，提高高频，其次利用Hamming窗函数将语音信号分割成帧，再对语音信号的能谱进行离散傅里叶变换，将能谱传递到梅尔尺度三角滤波器组，最后利用离散余弦变换（DCT）获得MFCCs. 而本文使用的HCFs为384维，其中共包含32个低级描述符：过零率、能量、$F0$、MFCCs、语音概率等，所获得的特征集涵盖了主要的语音情感特征. 而每一类特征都具有不同的情感域分布，图2展示了MFCCs和HCFs的情感域分布，和两类特征联合后的情感域分布. 由图2可知，特征的情感域分布具有显著重叠，而不同类型特征的重叠区域具有差异. 由图2（a）可知，MFCCs的易混淆情感域是悲伤、无聊和中性；由图2（b）可知，HCFs的易混淆情感域是厌恶和中性；两类特征联合后，图2（c）的易混淆部分为无聊和中性. 情感域分布的重叠差异说明不同类型特征之间具有互补性，所以当不同类型特征独立处理和联合处理时获得的特征表征是具有显著差异的. 这便是我们提出CFLF的主要动机，学习不同特征的互补性和独立性，以提取出更全面的情感表征.

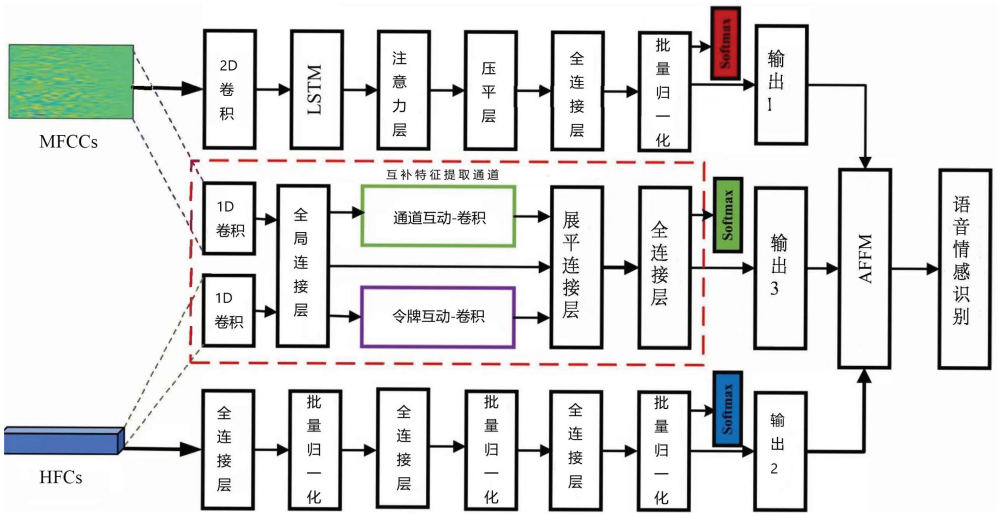
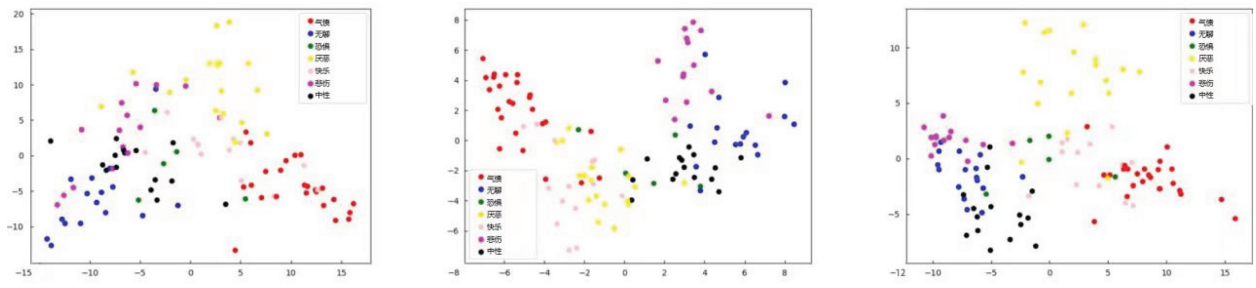


图 1 SER模型的整体网络架构



(a) MFCCs (b) HCFs (c) COMBINE

图 2 不同类型特征的情感域分布图

2.2 CFLF

为学习不同特征的互补性和独立性，本文提出了CFLF. 框架包括三条分支：一条独立处理MFCCs的CNN特征提取分支；一条独立处理HCFs的DNN特征提取分支；一条处理联合的MFCCs和HCFs的交互特征提取分支. 独立CNN特征提取分支中采用了四个卷积块以挖掘MFCCs的时频域内的空间关系，卷积块由卷积层、最大池化层和批量归一化层组成[21]. 在卷积层最后加入注意层，以找出MFCCs的显著情感区域，该分支的输出称为$F1$.

独立CNN特征提取分支包含三个全连接网络[22]和一个批处理归一化层，从而有效捕捉CHFs之间的线性特征，该分支的输出称为$F2$. 为在交互特征提取分支中提取MFCCs和HCFs之间的交互互补特征表征，使用1D卷积块分别处理MFCCs和HCFs，输出式（1）：

$$O(x)=\delta[B(Conv1\mathrm{D}(x))],\quad O\in R^{W\times H\times C} \tag{1}$$

其中：δ为非线性激活函数，B为批量归一化层. 再使用全局拼接层将MFCCs和HCFs的1D卷积输出组合在一起，获得的输出$F(x)$包含全局上下文信息，公式为：

$$F(x)=[O(\text{MFCCs}),O(\text{HCFs})],\quad F\in R^{W\times H\times 2C} \tag{2}$$

拼接完成后，该模型通过交互卷积学习通道和空间感知上下文，即在信道交互卷积过程中，沿着通道轴进行卷积，公式为：

$$C_O(x)=\delta[B(Conv1\text{D}(F(x)))],\quad O\in C_O^{W\times H\times 2C} \tag{3}$$

空间交互卷积时，首先将$G(x)$重塑为$G'(x)$，新的形状为$W\times 2C\times H$，通过沿H轴的卷积得到空间感知特征，公式为：

$$S(x)=\delta[B(Conv1\text{D}(G'(x)))],\quad S\in R^{W\times 2C\times H} \tag{4}$$

最后，将生成的全局、通道和空间感知特征聚合在平坦的级联层，并后接一个全连接层. 将交互特征提取分支的输出称为$F3$.

$$Flatten=[Ft(G(x)),Ft(C(x)),Ft(S(x))] \tag{5}$$

$$F3=\delta(B(Fc(Flatten))) \tag{6}$$

其中：Ft为平坦的级联层，Fc为全连接层.

2.3　AFFM

受前人工作[4,18]启发，通过学习不同输出之间的跨通道和跨token的信息生成注意力特征融合权重. 为有效融合CFLF中输出的三个分支$F1$、$F2$、$F3$，并充分利用MFCCs和HCFs中的独立表征和交互互补表征中的情感信息，使用了基于注意力机制[23]的AFFM. 图3为AFFM结构图.

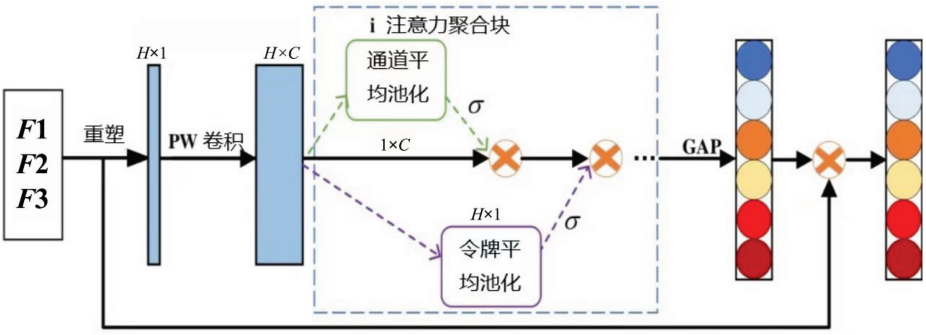


图 3　AFFM结构图

将CFLF输出的$F1$、$F2$、$F3$特征组合成一个全局向量$F(x)$:

$$F(x)=[F1,F2,F3] \tag{7}$$

将$F(x)$作为AFFM的输入. $F(x)$首先被重塑为$F'(x)$，其形状为$B\times H\times 1$. 使用逐点卷积（Point-Wise Convolution, PWConv）聚合通道和跨token交互，逐点卷积的输出为：

$$O(x)=\delta[B(\text{PWConv}(F(x)))],\quad O(x)\in R^{B\times H\times C} \tag{8}$$

经逐点卷积后，获得的输出分别沿通道轴和token轴进行平均池化，再经sigmoid函数，公式为：

$$C_O(x)=\sigma(\frac{1}{C}\sum_{i=1}^{C}O_{[:,i]}),\quad C_O\in R^{1\times H\times 1} \tag{9}$$

$$T_O(x)=\sigma(\frac{1}{T}\sum_{j=1}^{T}O_{[j,:]}),\quad T_O\in R^{1\times 1\times C} \tag{10}$$

其中：σ是sigmoid函数. 为生成跨通道和跨token上下文信息，将$O(x)$、$C_O(x)$和$T_O(x)$相乘：

$$G(x)=O(x)\times C_O(x)\times T_O(x),\quad G\in R^{B\times H\times C} \tag{11}$$

AFFM中使用了两个PWConv层，每个PWConv层具有大小为3×3的内核. 假设两个PWConv层的输出均为$G(x)$，在$G(x)$后应用全局平均池化（Global Average Pooling, GAP）生成通道注意力权重，公式为：

$$W(x)=\sigma(\frac{1}{C+H}\sum_{i=1}^{C}\sum_{j=1}^{H}O_{[:,i,j]}),\quad W(x)\in R^{B\times H} \tag{12}$$

经GAP后，全局跨通道上下文信息被压缩为一个标量，从而强调CFLF中三个支路的输出对后续情感识别的不同贡献，最后使用跳跃连接对特征进行细化. 为了尽可能地保持已提取的情感特征并强调通道的可选择性，将AFFM中短跳跃连接看作是模型输出的映射. 给定注意力特征融合权重，细化后的特征表示为：

$$F'=F(x)\times W(x),\quad F'\in R^{W\times H} \tag{13}$$

最后获得的F'是输入后续情感分类器的情感表征，它包含了MFCCs和HCFs的独立性表征和交互互补表征，并根据不同表征在情感识别中的贡献被分配了不同的权重.

3 实验设置与数值结果

3.1 实验设置

使用IEMOCAP和Emo-DB数据库[24−25]测试所提SER模型. IEMOCAP由5个环节组成，每个环节由两位演讲者（1女1男）完成. 共包含10 039个话语，其持续时间从3秒到15秒不等. 此外，本文只选择了中性、愤怒、悲伤和快乐四种情绪标签的话语. Emo-DB由10位专业演员制作的535个话语组成，涵盖7个情感类别，以16千赫频率采样，平均持续时间为2.7秒.

试验中80%的数据用于训练，20%的数据用于测试. 由于CNN的输入必须保持长度一致，故我们对所有的样本进行了填充或切割操作以保证每条语音长度一致. 采用UA和WA性能指标评价实验结果. 为对比不同文献中的特征表征提取、特征融合方法的性能，证实本文提出的CFLF和AFFM的SER模型的有效性，我们设计了四个SER模型：

1）基线模型[15]：将两种输入特征使用两条独立分支进行处理，提取独立性表征，输入情感分类器.

2）全连接模型[16]：在基线模型的两条独立分支后增加一个全连接层，以融合两个支路输出. 将融合输出和两个独立性表征联合以输入情感分类器.

3）联合微调模型：受联合微调方法[17]的启发，提出了联合微调模型，该模型和全连接模型结构相同，但是在两条分支权重冻结情况下重新训练全连接层来微调.

4）本文模型：使用CFLF和AFFM，得到最终的加权表征以进行情感分类.

以上模型均在IEMOCAP和Emo-DB上进行训练，选取的情感特征均为MFCCs和openSMILE提取的HCFs，末端情感分类器均使用SVM. 此外，与近期研究[11−12,14]中提出优化特征提取方法的模型进行了对比.

3.2 数值结果

表1展示了四个SER模型、仅使用CFLF块的本文模型及其它研究实验的数值结果. 相比前人[15−17]的提取互补特征方法，本文所提CFLF结合AFFM的SER模型取得了更好的情感识别结果，即使仅使用CFLF也比大多数模型效果好. 可见采用CFLF获取到的不同特征的独立性和互补性表征能够包含更加充分的情感信息，使用AFFM来融合不同表征能够让模型有效地选择更具有影响力的情感表征进行识别.

为了探究SER模型中AFFM和CFLF的有效性，进行了消融实验. 使用基线模型、仅使用CFLF的模型和使用CFLF结合AFFM的模型进行实验. 由表2可知，仅使用CFLF时，模型性能也优于基线模型，可见提取出的交互互补特征表征的加入有助于提升情感识别性能. 而同时使用CFLF和AFFM时，模型性能进一步提高，说明AFFM能够合理地为独立情感表征和交互互补情感表征分配权重，从而有效地帮助模型关注到对情感识别贡献度更大的情感表征.

表 1　不同模型比较

模型	数据集	分类器	WA/%	UA/%
文献[11]中模型	IEMOCAP	-	-	-
	Emo-DB	Softmax	86.44	84.53
文献[12]中模型	IEMOCAP	FC	68.73	70.29
	Emo-DB	-	-	-
文献[14]中模型	IEMOCAP	-	-	-
	Emo-DB	CNN&BLSTM注意力模型	86.67	86.03
基线模型[15]	IEMOCAP	SVM	75.36	69.63
	Emo-DB	SVM	83.54	83.65
全连接模型[16]	IEMOCAP	SVM	73.67	69.17
	Emo-DB	SVM	84.02	84.55
联合微调模型	IEMOCAP	SVM	74.16	68.33
	Emo-DB	SVM	83.03	83.22
仅使用CFLF	IEMOCAP	SVM	75.49	69.63
	Emo-DB	SVM	83.03	85.92
CFLF+AFFM（本文模型）	IEMOCAP	SVM	**76.10**	69.54
	Emo-DB	SVM	**87.10**	**87.34**

表 2　消融实验

数据集	模型	WA/%	UA/%
IEMOCAP	基线模型[15]	75.36	69.62
	仅使用CFLF	75.49	69.63
	CFLF+AFFM（本文模型）	**76.10**	69.54
Emo-DB	基线模型[15]	83.54	83.65
	仅使用CFLF	83.03	85.92
	CFLF+AFFM（本文模型）	**87.10**	**87.34**

4　总结与展望

为提取出不同特征之间的互补信息，使用了具有交互特征提取分支和两个独立性特征提取分支的CFLF，获得了独立性和互补性的特征表征. 这有利于全面提取语音信号中的情感信息. 而AFFM则是根据不同表征的贡献来为表征分配权重，让SER模型更集中注意在有效的情感特征上. 然而本文仅集中在特征的互补性上和权重分配上，对分类器的研究仍有欠缺，后续研究中会考虑使用深度学习框架来设计分类器.

参考文献：

[1] DE LOPE J，GRAÑA M. An ongoing review of speech emotion recognition[J]. Neurocomputing，2023，528：1-11.

[2] STOCK-HOMBURG R. Survey of emotions in human-robot interactions：Perspectives from robotic psychology on 20 years of research[J]. International Journal of Social Robotics，2022，14(2)：389-411.

[3] LIU Z Y，HU B，LI X Y，et al. Detecting depression in speech under different speaking styles and emotional valences[C]//International Conference on Brain Informatics. Cham：Springer，2017：261-271.

[4] CHEN M Y，HE X J，YANG J，et al. 3-D convolutional recurrent neural networks with attention model for speech emotion recognition[J]. IEEE Signal Processing Letters，2018，25(10)：1440-1444.

[5] DAHAKE P P，SHAW K，MALATHI P. Speaker dependent speech emotion recognition using MFCC and support vector machine[C]//2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT). Pune，India. IEEE，2016：1080-1084.

[6] EYBEN F，WÖLLMER M，SCHULLER B. OpenEAR-introducing the Munich open-source emotion and affect recognition toolkit[C]//2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops. Amsterdam，Netherlands. IEEE，2009：1-6.

[7] ZHENG L, LI Q, BAN H, et al. Speech emotion recognition based on convolution neural network combined with random forest[C]//2018 Chinese Control and Decision Conference (CCDC). Shenyang, China. IEEE, 2018: 4143-4147.

[8] ZHOU P, LI X P, LI J, et al. Speech emotion recognition based on mixed MFCC[J]. Applied Mechanics and Materials, 2012, 249/250: 1252-1258.

[9] LALITHA S, MUDUPU A, NANDYALA B V, et al. Speech emotion recognition using DWT[C]//2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC). Madurai, India. IEEE, 2015: 1-4.

[10] RAO K S, KOOLAGUDI S G, VEMPADA R R. Emotion recognition from speech using global and local prosodic features[J]. International Journal of Speech Technology, 2013, 16: 143-160.

[11] JIANG P X, FU H L, TAO H W, et al. Parallelized convolutional recurrent neural network with spectral features for speech emotion recognition[J]. IEEE Access, 2019, 7: 90368-90377.

[12] CHEN Q P, HUANG G M. A novel dual attention-based BLSTM with hybrid features in speech emotion recognition[J]. Engineering Applications of Artificial Intelligence, 2021, 102: 104277.

[13] GUO L L, WANG L B, DANG J W, et al. Exploration of complementary features for speech emotion recognition based on kernel extreme learning machine[J]. IEEE Access, 2019, 7: 75798-75809.

[14] HE J R, REN L Y. Speech emotion recognition using XGBoost and CNN BLSTM with attention[C]//2021 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/IOP/SCI). Atlanta, GA, USA. IEEE, 2021: 154-159.

[15] ZHONG S M, YU B X, ZHANG H. Exploration of an independent training framework for speech emotion recognition[J]. IEEE Access, 2020, 8: 222533-222543.

[16] LIU J X, LIU Z L, WANG L B, et al. Speech emotion recognition with local-global aware deep representation learning[C]//ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Barcelona, Spain. IEEE, 2020: 7174-7178.

[17] JUNG H, LEE S, YIM J, et al. Joint fine-tuning in deep neural networks for facial expression recognition[C]//2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile. IEEE, 2015: 2983-2991.

[18] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//European Conference on Computer Vision. Cham: Springer, 2018: 3-19.

[19] KE X X, ZHU Y J, WEN L, et al. Speech emotion recognition based on SVM and ANN[J]. International Journal of Machine Learning and Computing, 2018, 8(3): 198-202.

[20] SAHOO S, ROUTRAY A. MFCC feature with optimized frequency range: An essential step for emotion recognition[C]//2016 International Conference on Systems in Medicine and Biology (ICSMB). Kharagpur, India. IEEE, 2016: 162-165.

[21] WU S, LI G Q, DENG L, et al. $L1$-norm batch normalization for efficient training of deep neural networks[J]. IEEE Transactions on Neural Networks and Learning Systems, 2019, 30(7): 2043-2051.

[22] LI R N, WU Z Y, JIA J, et al. Dilated residual network with multi-head self-attention for speech emotion recognition[C]//ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Brighton, UK. IEEE, 2019: 6675-6679.

[23] HAN K, XIAO A, WU E H, et al. Transformer in transformer[EB/OL]. 2021: arXiv: 2103.00112. https://arxiv.org/abs/2103.00112.pdf.

[24] BUSSO C, BULUT M, LEE C C, et al. IEMOCAP: Interactive emotional dyadic motion capture database[J]. Language Resources and Evaluation, 2008, 42: 335-359.

[25] RAMDINMAWII E, MOHANTA A, MITTAL V K. Emotion recognition from speech signal[C]//TENCON 2017 - 2017 IEEE Region 10 Conference. Penang, Malaysia. IEEE, 2017: 1562-1567.

责任编辑：赵新科　刘 敏