

人工智能评论》手稿号。  
(将由编辑插入)

## 基于深度学习的人格检测的最新趋势

**Yash Mehta - Navonil Majumder - Alexander Gelbukh  
- Erik Cambria**

收到。2018-12-03 / 已接受。XXXX-XX-XX

### 摘要

最近，人格特征的自动预测受到了广泛的关注。具体来说，从多模态数据中预测人格特征已经成为情感计算领域的一个热门话题。在本文中，我们回顾了已被用于人格检测的重要机器学习模型，重点是基于深度学习的方法。本文概述了最流行的自动人格检测方法，各种计算数据集，其工业应用，以及用于人格检测的最先进的机器学习模型，并特别关注多模态方法。人格检测是一个非常广泛和多样化的话题：本调查只关注计算方法，而忽略了关于人格检测的心理学研究。

### 1 简介

人格是一个人的行为、情感、动机和思维模式的特点的组合。我们的个性对我们的生活有很大的影响，影响我们的生活选择、幸福、健康以及我们的喜好和欲望。因此，自动检测一个人的个性特征的能力有许多重要的实际应用。伍德沃思心理神经量表[71]通常被认为是第一个人格测试。它是在第一次世界大战期间为美国军队开发的，用于筛查新兵的炮击风险（创伤后应激障碍）。在今天，一个有着活跃的社区并被许多人使用的人格模型是过程交流模型（PCM）。

Y.Mehta  
伦敦大学学院盖茨比计算神经科学组

N.N. Majumder 和 A. Gelbukh  
国家政治学院计算机研究中心

E.Cambria  
新加坡南洋理工大学计算机科学与工程学  
院 电子邮件：cambria@ntu.edu.sg

它是由Taibi

Kahler在NASA的资助下开发的，最初用于协助航天飞机宇航员的选择。这项措施现在主要用于咨询目的，以帮助个人更有效地进行沟通，甚至避免或解决沟通偏离轨道的情况。

与其直接确定一个人的个性，人们可能还想知道他们是如何被周围的人所感知的。与自动识别人格的情况不同，感知人格分析的目标不是个人的真实人格，而是与他们交往的人所赋予他们的人格。周围的人对个人进行类似的人格问卷调查，然后确定这个人的感知人格。对感知到的人格进行评估时，使用的是与他们的“真实”人格相同的人格测量方法（如大五人格）。需要注意的一点是，感知的人格因人而异，受制于他们的心理使然或他们自己的个性。许多研究表明，人们在见到一个人的头十秒钟内就会通过其面部特征形成对一个人的第一印象。除了遇到的面孔的物理特征外，情绪表达、环境信息、观看者的心理形成也会影响这些印象。第一印象影响人们对新遇到的人或类似人类的代理人的行为。

### 1.1 人格测量

现代特质理论[77]试图通过设定一些分类维度（通常采用词汇法）并构建问卷来测量它们[65]。研究人员使用了各种人格模型方案，如16PF[14]、EPQ-R[67]和三特质人格模型PEN[31]，其中有超级因素精神病性、外向性和神经质（PEN）处于层次的顶端。迈尔斯-布里格斯类型指标（MBTI）[7]是世界上最广泛使用的人格测试之一，每年在数千家公司员工中进行数百万次测试。MBTI性格测试将人们分为四个方面：内向与外向，感觉与直觉，思考与感觉，以及判断与感知。

迄今为止，在自动人格检测的文献中，最流行的衡量标准是五大人格特征[25]，即以下的二进制（是/否）值。

- **外向性（EXT）**。这个人是外向、健谈、精力充沛，还是拘谨、孤独？
- **神经质（NEU）**。这个人是敏感和紧张，还是安全和自信？
- **认同度（AGR）**。这个人是否值得信赖、直率、慷慨和谦虚，与不可靠、复杂、微薄和夸夸其谈？
- **自觉性（CON）**。这个人是高效和有组织的，还是马虎和不关心的？
- **开放性（OPN）**。这个人是否具有创造性和好奇心，而不是教条主义和谨慎的态度？

由于大部分的研究都使用了大五人格测量法进行分类，除非另有说明，否则我们将默认使用该测量法。

## 1.2 应用

在今天的状况下，自动人格识别系统有各种不同的工业应用。我们推测在不久的将来会有一个巨大的市场，如果模型能够准确和稳定地测量人格，那么工业界对自动人格识别软件会有巨大的需求。随着在这一领域的重新探索，很快就会发现具有更高的准确度和可靠性的更好的人格预测模型。人工人格可以与几乎所有的人类计算机互动结合起来，向前发展。

任何计算设备都可以配备某种个性，使其能够对不同的人 and 情况作出不同的反应。例如，一部手机可以有不同的模式，有不同的可配置的个性。这将为更有趣和个性化的互动铺平道路。另一种可能性是将个性特征作为输入之一，以便在其他任务中实现更高的准确性，如讽刺检测、谎言检测或单词极性歧义系统。

- **增强的个人助理。** 当今的自动语音助手，如Siri、谷歌助手、Alexa等，可以自动检测用户的个性，从而给出定制的回答。此外，语音助手可以被编程为根据用户的个性显示不同的个性，以提高用户满意度。
- **推荐系统。** 拥有特定个性类型的人可能有类似的兴趣和爱好。推荐给一个人的产品和服务应该是那些被其他具有类似个性类型的用户积极评价过的产品和服务。例如，[108]提出根据客户的爱好和个性建立汽车购买意向的模型。[106]开发了一个系统，根据玩家的个性向他们推荐游戏，该系统是根据他们与其他玩家的聊天记录来建模的。
- **词的极性检测。** 人格检测可以被用于情感词典中的词极性歧义，因为同一概念对不同类型的人可以表达不同的意思。另外，将用户的个性特征和概况纳入讽刺和非讽刺内容的歧义辨别[82]，可以提高准确率。
- **专门的医疗保健和咨询。** 截至2016年，近三分之一的美国人曾因心理健康相关问题寻求专业咨询。这是人格特质预测存在巨大实际应用的另一个领域。根据一个人的个性，可以给予适当的自动咨询，或者心理医生可以利用这些信息来给予更好的咨询建议。
- **鉴证学。** 如果警方了解当时在犯罪现场的人的性格特征，可能有助于减少嫌疑人的圈子。人格检测也有助于自动欺骗检测领域，并能帮助建立具有更高精确度的测谎器。
- **工作筛选。** 在人力资源管理中，个性特征会影响一个人对某些工作的适应性。例如，某公司想招聘一个能够激励和领导某个团队的人。他们可以通过排除那些高度紧张和敏感的候选人，即那些具有高值的神经质特质，来缩小他们的筛选范围。[60]从心理学家和机器学习科学家的跨学科观点来讨论求职者筛选问题。

- **心理学研究**。自动人格特质检测将有助于发现人们的行为主义和人格特质之间更复杂和微妙的关系。这将有助于发现人类心理的新动态。
- **政治预测**。大规模的自动个性检测正在被用作政治家的指导方针，以提出更有效和有针对性的运动。如果一家分析公司能够获得有关选民的大规模行为数据，该公司就可以创建他们的心理档案。这些档案可以让人了解哪种广告最能有效地说服某个特定地点的某个政治事件的人。

### 1.3 公平和道德

在过去的一个世纪里，该领域的从业人员和研究人员已经开发并完善了各种访问人格的技术。已经制定了一些道德准则，以确保适当的程序得到遵循。然而，人格检测经常会出现有关适当利用和解释的伦理困境[68]。在侵犯隐私、文化偏见和保密性方面，人们对这些测试的不恰当使用提出了担忧。

最近，各政党正试图利用基于大规模机器学习的个性检测进行政治预测。在政治预测中使用个性检测的一个相当臭名昭著的例子是Facebook的剑桥分析公司数据丑闻。<sup>1</sup>该事件涉及收集8700万Facebook用户的个人身份信息。研究表明，选民的政治选择与他们的社会特征有很大关系[11]。据称，用户的社交数据被用来试图影响选民的意见，以支持2016年特朗普的竞选活动，确定对特定地区有效的说服性广告类型。它还涉及自动获取选民的个性特征，并以此预测谁会轻易被政治宣传所操纵。

还有人继续使用这些测试表示担忧，尽管它们缺乏有效的证据证明是对人格的准确测量。这些测试可能有固有的偏见，然后被机器学习算法无意中学习。减少机器学习算法中的偏见的一个步骤是通过算法影响评估（AIAs），正如纽约大学的AI Now研究所提出的。AIAs是从黑匣子方法导致恶性循环的想法中延伸出来的，不断远离对这些算法的理解，并削弱了解决这些算法可能产生的任何问题的能力。AI Now建议使用AIAs来处理机器学习在公共领域的使用，创建一套标准要求。通过AI As, AI Now旨在向公众提供清晰度，公开列出和解释所使用的算法系统，同时允许公众对这些系统提出异议，制定审计和评估程序，并提高公共机构的内部能力，以了解他们使用的系统。与使用AIA促进机器学习的透明度类似，国防高级研究计划局（DARPA）建议将可解释的人工智能（XAI）作为解决方案的一部分。其目标是产生更多可解释的模型，让用户能够理解和信任。

---

<sup>1</sup> The Great Hack, 最近一部关于剑桥分析公司数据丑闻的纪录片。

表1 用于每种模式的特征提取的各种流行工具。

提取的	模态工具特征	
文本	LIWC Receptiviti API Freeling SenticNet	心理和语言学特征 建立在LIWC之上。个性特征的视觉表现 POS标签 情感极性
音频	朴素 ĀĀĀ	强度、音调、响度、MFCC、抖动、闪烁、LSP
视觉	外交部 CERT 脸部 ++ EmoVu	面部表情, AU 8种面部表情（如：喜悦、悲伤、惊讶等）。 面部检测，106个面部标志，面部属性（如性别、年龄）。 人脸检测、情绪识别（含强度）、人脸属性

#### 1.4 组织机构

我们彻底回顾了基于机器学习的人格检测，重点是基于深度学习的方法。第二部分让读者了解了其他关于基于机器学习的自动人格检测的评论文章。我们根据输入的模式进行了分析，如文本、音频、视频或多模式。第3节简要讨论了这些模态中每一种模态中用于个性检测的流行技术和方法。在第4节中，我们深入分析了这些模式中每一种模式的多篇有趣的论文。第5节和第6节分别介绍了带有分析和结论的结果。

## 2 相关作品

自动人格检测是一个新的和即将到来的领域。在人格检测方面还没有很多全面的文献调查，我们的论文是第一篇给读者提供该领域最新趋势和发展的鸟瞰图的论文。

目前还没有最近的工作让读者对基于机器学习的自动人格检测的进展有一个整体的看法。第3节给出了每种模式下流行的机器学习模型，给读者提供了这个领域的广泛视角。[1]给出了使用Naïve

Bayes、kNN、MLR、高斯过程等浅层学习技术从文本中进行人格检测的趋势回顾。2014年后，端到端的深度神经网络架构和模型开始回升，并开始击败这些模型的最先进精度。[47]对各种图像处理技术和人格检测的面部特征分析进行了回顾，其中他们的重点是对机器学习技术和各种面部图像预处理技术的概述。

这个领域流行的图像预处理技术包括面部地标识别、面部动作编码系统（FACS），它给出了面部的AU，以及Viola

Jones的面部检测算法[101]。此外，他们没有使用标准的人格特征衡量标准（如Big-Five或MBTI），而是使用了稍微不同的衡量标准，如不快乐、奇怪、聪明、自信等等。这个调查只集中在视觉方面，而[100]给出了一个非常系统和清晰的关于从文本和音频中检测人格的调查。他们没有像本文那样按照模式来划分人格检测，而是选择了一个更广泛的主题，他们称之为人格计算。这又进一步分为3个部分

基本问题，即自动人格感知（APP）、自动人格合成（APS）和自动人格识别（APR）。本文主要讨论APR和APP（类似于感知人格）之后的方法。上面引用的论文对处理来自社交媒体（主要是Twitter和Facebook）的个性识别和感知的不同论文给出了非常详细的见解。他们对从文本、非语言交流（如人际距离、语言和身体动作）、社交媒体、手机、可穿戴设备以及最后从电脑游戏中识别人格进行了全面的分析。最近的一篇论文，[52]提供了一个关于基于计算机视觉的感知个性特征分析的全面调查。[29]从多媒体信息处理、模式识别和计算机视觉的角度对视频和图像中的明显个性自动分析的最新进展进行了汇编。然而，还有一些深度学习的技术、模型和论文，特别是针对多模态人格检测的技术、模型和论文，到目前为止还没有任何文献调查涉及。

### 3 基线方法

以下各小节总结了基于深度学习的人格检测中常用的各种流行模型、架构和技术。表3简要介绍了该领域的一些先驱性论文所描述的方法。

#### 3.1 文本

关于文本模式，数据预处理是一个非常重要的步骤，选择正确的技术可以产生明显更好的结果。通常，从文本中提取特征，如语言学调查和字数（LIWC）[73]、Mairresse、医学研究委员会（MRC）等，然后将其输入标准的ML分类器，如Se-quential Minimum Optimizer、支持向量机（SVM）、Na`ve Bayes等。学习单词嵌入并将其表示为矢量（使用GloVe或Word2Vec）也是一种非常普遍的方法。这些单词向量也可以通过将单词按字符送入长短期存储器（LSTM）或门控递归单元（GRU）来创建。据观察，将文本特征（LIWC、MRC）与其他东西如常识性知识、卷积等结合起来，可以获得更好的性能。

#### 3.2 音频

相对来说，只有少数方法专注于使用音频作为检测人格的唯一输入。它通常与视觉模式结合起来进行双模态人格检测。标准的音频特征，如Mel-Frequency Cepstral Coefficients (MFCC), Zero Crossing Rate (ZCR), Logfbank, 其他cepstral和spectral特征，可作为SVM和线性回归器的输入。

#### 3.3 视觉

正如大多数视觉深度学习任务一样，卷积神经网络（CNN）是最常用的，并在人格领域产生了最先进的结果。

探测也是如此。大多数方法分析面部特征，并试图找到一个函数，将这些特征映射到人格特征上。许多研究人员使用了预训练的深度CNN（如VGG-Face），并对其进行了微调以完成人格检测的任务（转移学习）。他们尝试了不同的提取面部特征的方法，如EigenFace、HOG（定向梯度直方图）[23]、FACS（提取AU，如凸起的眉毛、酒窝等）和Viola Jones算法（面部识别），以达到更高的准确性。

### 3.4 多式联运

大多数的多模态方法都是后期融合，也就是说，他们取的是音频和视觉模态预测的单个结果的平均值。深度双模态再回归通过使用稍加修改的深度残差网给出最先进的结果[42]。从每个模态中提取的特征可以一起使用，以得出人格预测结果。这种技术是早期融合。目前该领域的研究方向是寻找更有效的特征提取方法和多模态特征组合。我们看到很少有模型涉及到三模态特征的融合。

## 4 详细概述

在本节中，我们详细分析了个别论文，讨论了各种技术、方法和途径。平均而言，我们发现从视觉模态中提取的特征在单模态的人格检测中是最准确的。研究发现，正如人们所期望的那样，将一种以上的模式的输入结合起来，往往会有更高的预测准确性。

### 4.1 文本

在过去的十年中，许多研究将语言的使用与广泛的心理逻辑相关因素联系起来[72, 48]。最早的一些研究发现了写作风格（例如，用词频率）与人格的可靠关联[74]。例如，在外向性方面得分较高的人使用更多的积极情绪词（例如，伟大、惊人、快乐），而那些神经质方面得分较高的人被发现更频繁地使用第一人称单数（例如，我、我的、我）。利用这些语言使用上的差异进行人格检测被称为**封闭式词汇法**。

从文本（尤其是社交媒体）中检测个性的最常见的封闭词汇方法之一是LIWC。它将单词归类为各种心理相关的桶，如

"功能词"（如冠词、连接词、代词）、"情感过程"（如快乐、紧张、哭泣）和"社会过程"（如伴侣、谈话、朋友）。然后，它计算每个桶（超过60个心理学相关的桶）中的单词频率，然后预测文本作者的个性。由于这种方法从预定义的词语类别开始，它被描述为"封闭的词汇"，而且LIWC提取的大部分特征都依赖于语言。

#### Receptiviti

API[35]是一个建立在LIWC基础上的工具，用于从文本中分析人格。用户的个性预测是使用心理语言学特征进行的。Receptiviti API允许用户提交一个文本样本，该样本被分析并输出一个图形化的结果。

预测特征的表示，也可以在一个基于网络的界面上查看。然而，这个API在社交媒体衍生的文本上表现不佳，因为封闭式词汇法在处理一般的微型文本时准确率较低。

人格特征甚至可以从移动设备上不同来源的人际交流（如WhatsApp）中预测出来。[87]建立了一个应用程序，负责编译和发送关于用户的信息到一个服务器应用程序，然后运行分类算法。本文遵循的一条研究路线是检查人格是否与一个人的幸福感有关，然而并没有得出任何结论性的结果。这是为数不多的使用PEN理论进行分类的研究论文之一，而不是使用Big-

Five。测试了标准的机器学习算法，如J48、Random Forrest和SVM。

近年来，来自社交媒体的个性检测，特别是Twitter的情感分析已经获得了很多人的青睐。这可能部分是由于Twitter的数据收集是直接的，很容易通过Twitter的API获得。[53]将用户的社交行为（平均链接数、平均哈希标签数、平均提及数等）和语法信息（文本的平均长度、平均正负字数、平均逗号、问号等特殊字符数）结合起来用于用户个性检测。利用这些属性，构建一个特征向量，然后将其送入多层感知器（MLP）。这种方法的创新之处在于在Hadoop框架（Hadoop分布式文件系统和Map Reduce）上运行整个设置，从而能够同时对N个用户进行个性预测。

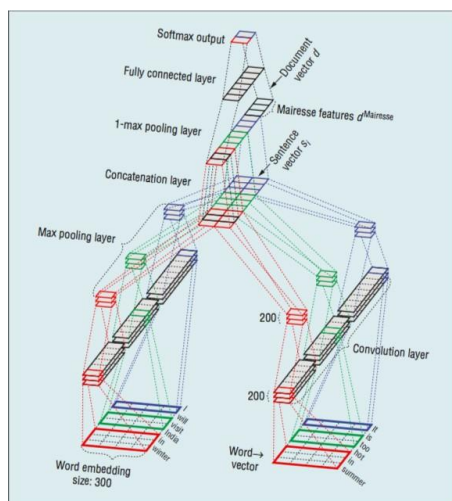


图1 CNN框架由[64]使用Word2Vec嵌入来进行文本中的个性检测。

除了社交媒体数据外，论文也是一种流行的文本模式，可以用于作者分析。一个流行的数据集是James Pennebaker和Laura King的意识流散文数据集[74]。它包含了2,468篇匿名文章，根据大五行特征对作者的个性进行了标记。[64]利用深度CNN进行文档级的人格检测任务。该CNN有助于提取单字、大字和小字。



表2受欢迎的数据集，根据输入模式划分。

模式	数据集	人格测量	描述
文本	论文一 [74]	大五	2,468篇匿名文章被贴上了作者的人格特质标签。志愿者在受控环境中写的意识流文章，文章的作者被要求标注他们自己的五大人格特征
	论文二 [95]	大五	2,400篇作文被手工贴上了五种不同性格特征的人格分数。然后对数据进行调整，并将回归分数转换成五种不同特征的类别标签
	MBTI Kaggle2	MBTI	包含8600多人的MBTI性格标签，以及他们在myPersonality cafe论坛上的50篇文章。myPersonality是一款Facebook应用程序，允许其用户通过填写人格问卷参与心理学研究。然而，从2018年起，他们决定停止与其他学者分享数据。
	我的个性 <sup>3</sup>	大五	748个意大利FriendFeed用户的样本（1,065个帖子）。该数据集从FriendFeed的公共URL中收集，那里的新帖子是公开的。
	意大利FriendsFeed [16]	大五	
音频	AMI	大五	独白、对话和多方讨论的视频和音频记录，并对通过BFI-10问卷调查发现的感知个性进行了注释。
	会议语料库[13]。	-	连接的数字语料库，包含8,440句干净的多条件训练数据和70,070句干净的噪声测试数据。
	Aurora2语料库[44]	-	与专业演讲者的实验数据库进行了录音，并通过使用NEO-FFI人格清单进行听力测试，生成了录音的大五因子分数。
	CMU自我记录的数据库[79]	大五	平衡语料库包括来自126个（以前不认识的）主题对的数据，构成了93.8小时的英语演讲
Five	哥伦比亚的欺骗语料库[58]	-	被300名Flickr用户收藏的前200张图片（因此共有6万张图片）被贴上了性格特征的标签。
	VisualPsychoFlickr数据集[21]	Big-	第一印象数据集包括从3000多个不同的YouTube高清视频中提取的10,000个片段（平均持续时间为15秒），这些视频是人们面对面用英语说话的。
	第一印象V2（CVP R'17） <sup>4</sup>		摄像机

该模型的结构如图1所示。据观察，去除中性句子后，预测的准确性有了明显的提高。每个词在输入中都用Word2Vec[66]表示为一个固定长度的特征向量，而句子则表示为一个可变数量的词向量。最后，文档级的Mairesse特征（LIWC、MRC等），共84个特征，与从深度CNN中提取的特征向量相连接。最后，这个串联的向量被送入全连接层，用于最终的人格特征预测。[43]试图通过将输入输入到循环神经网络（RNNs）来模拟句子间的时间依赖性。发现LSTM与vanilla RNN、GRU、bi-LSTM相比，效果更好。GloVe[75]50D的词嵌入被用来预测人们的MBTI性格，这些帖子是在一个叫做PersonalityCafe的特定社会性格讨论网站上发表的。

研究发现，将常识性知识与心理语言学特征相结合，可以显著提高准确性[83]。SenticNet[10]是一个流行的工具，用于从文本中提取常识性知识以及相关的情绪极性和情感标签。它是意见挖掘和情感分析的最有用的资源之一。这些特征被用作五个基于序列最小优化的监督分类器的输入，用于预测五个人格特征，由每个分类器独立预测。

虽然大多数研究者认为五种人格特征是相互独立的，但一些研究[90, 51]声称，这些特征之间存在一定的关联性，建立5个完全独立的分类器是不准确的。[114]对五种人格特征之间的相互关系进行了建模。

使用加权的ML-

KNN算法，该算法使用信息熵理论[49]为提取的特征分配权重。语言和相应的依赖性情感特征是从文本中提取的，然后可以根据Kohonen的特征地图算法[57]进行离散化。

**开放词汇**法不依赖于先前的词汇或类别判断，而是依赖于从文本中提取全面的语言特征集合。这些方法通过非单词符号（如标点符号、情感符号）的相对使用、单个未分类的单词、多单词短语以及通过无监督方法（话题）识别的语义相关的单词集群来描述样本文本。通常，Latent Dirichlet Allocation (LDA) [6]被用来形成语义相关词的聚类。

MyPersonality数据集是著名的文本社交媒体数据集之一，由66000多名Facebook用户的状态更新组成。这些志愿者中的每个人都参加了大五人格测试。[72]使用这个数据集从社交媒体中使用的语言进行自动人格检测。首先，使用文本的标准技术提取特征。然后，这些特征的维度被降低，接着进行Ridge回归[45]。有一个单独的回归器用于预测五种个人特征中的每一种。作者[109]在这个数据集的一个子集（Facebook用户状态）上训练了一个端到端的深度学习模型，作者[109]用n-grams（用CNN提取）和双向RNN（向前和向后GRU）进行了实验。他们没有使用预先训练好的Word2Vec模型，而是使用skip-gram方法[66]训练单词嵌入矩阵，以纳入网络俚语、表情符号和缩略语。为了提高准确性，他们对从文本中提取的非词汇特征和神经网络的倒数第二层进行后期融合。这个融合后的向量然后被送入一个softmax输出层。

[62]将基于深度学习的模型与文本的原子特征即字符结合起来。使用分层的、矢量的单词和句子表征来预测一个人的个性特征。这项工作的基础是[61]，它提供了一个独立于语言的人格检测模型。词向量表示不是利用标准的词嵌入（如GloVe或Word2Vec），而是由使用GRU[20]作为递归单元的双RNN形成。GRU的计算成本比LSTM低，但性能相似。这些词被送入另一个bi-RNN，它给出了句子的表示，然后由一个前馈神经网络来预测五个特征的值。这种基于深度学习的字符级词汇编码的最佳用途之一是推文的语料库。这些数据是有噪音的，由大量的微型文本组成，因此这个模型的表现比传统的LIWC/BoW模型好很多。

最近，NLP工具已经被开发出来，用于从未标记的文本中进行独立于语言的、无监督的特征识别（PR2[17]）。PR2系统利用了从LIWC和MRC中提取的独立于语言的特征，如标点符号、问号、引号、感叹号、数字等。只有输入文本中高于平均水平的特征值才会被映射到个性特征上。据观察，含有更多标点符号的文本与外向性和经验开放性呈负相关。这种方法并不像该领域的最新技术那样准确，但它的计算效率更高，而且不受语言影响。

诺基亚的研究人员开发了一个软件[19]，而不是仅仅依靠文本作为唯一的输入形式，该软件利用来自智能手机的数据，包括关于哪些应用程序被使用以及使用频率的信息、音乐偏好、匿名通话记录、短信记录和蓝牙扫描来进行个性检测。使用标准的相关性和多因素分析，系统地分析了人格特征和使用特征之间的关系。

寻求回归分析。SVM和C4.5分类器被用来对用户的个性特征进行分类。

社交媒体是人们意见和想法的一个巨大来源，分析它可以对人们的个性有重要的了解。[36]使用了文体特征，可分为两组：字符N-grams和POS N-grams。通过利用开源工具Freeling，从POS标签中获得作者的语法序列。字符N-grams和POS N-grams是从POS和类文件中提取出来的，以创建一个特征向量，然后将其输入linerSVM进行个性检测。

[94]引入了基于紧密联系的句子的抽象特征组合，他们称之为潜在句子组。作者使用双向LSTMs，与CNN结合，利用文本的结构检测用户的个性。

有时，输入文本可能是两个人之间的聊天记录集合。从对话中检测人格是一项比简单文本更难的任务，因为我们需要准确地建立时间依赖性模型。双人对话中的互动有不同程度的相互影响，由两个人之间的轮流对话引起。目前的网络在照顾语言依赖性方面很好。[92]利用虚构的RNNs为对话中的短期时间演变建模。一个255维的线性特征向量与隐马尔可夫模型（C-HMM）相配合，通过使用长期的轮流时间演化和跨说话人的语境信息来检测每个对话中两个说话人的性格。普通话转换对话语料库[96]被用来作为对话语料库来训练和评估模型。在许多研究中，个性的“表现”是近似的，而不是一个人的“真实”个性（在进行了个性问卷调查之后）。有时，训练有素的法官会根据一个人的行为给他的个性贴上标签，让我们了解他们是如何被周围的人感知的。

## 4.2 音频

我们注意到，最近大多数从音频模式中检测人格的模型都分两个独立的阶段，首先是特征提取，然后将特征输入分类器。音频描述符可以被细分为7个组，强度、音高、响度、声调、频谱、MFCC和其他特征。一些研究者使用这些特征组中的一个或多个来评估个性，而个性是在语篇层面上得出的[79]。这些音频特征（如强度、音高、MFCC）是用Praat声学分析程序提取的，并送入SVM分类器。这篇论文[79]的作者发现，五个人格特征是相互依存的，其中一个特征的变化会导致其他四个特征的可观察的变化。该数据集是通过记录话语，并通过让受过训练的说话人模仿特定类型的人格来进行标记。因此，训练和评估集的准确性完全取决于训练有素的演讲者在模仿某种性格的声音方面的能力。

一些研究者声称，在感知个性检测方面，非语言特征（语调、重叠、中断和语音活动）优于语言特征（对话行为和单词n-grams）。[98]在AMI语料库[12]的口语对话上测试了他们的模型，该语料库是在专门的仪器化会议室中捕获的会议集合，同时还有每个参与者的音频和视频。他们的方法的新颖之处在于将对话行为标签（捕捉讨论中说话者的意图）作为输入特征之一送入分类器。每个说话者都被贴上了14个对话行为标签，并与他们所感知的个性相关联。

观察到的特征（尤其是外向性）。语音活动特征（每个说话人的总时间和相对时间，每个说话人的平均停顿时间等）和语速特征（总句数、平均持续时间、最大持续时间等）也被纳入算法，使用多类提升（Booster）算法[89]。

最近，涉及训练深度神经网络的端到端方法已经获得了广泛的欢迎。它们在自动提取必要的声学特征的同时产生了最先进的精确度。[70]分析了CNN来理解在前两个卷积层之间建模的语音信息。第一个卷积层作为滤波器库，自动学习从未处理的音频波输入中提取相关特征。本文的关键区别在于，时间上的原始语音被直接送入CNN，而不是传统的两步程序，即先提取特征向量（如MFCC），然后再送入分类器。

人们经常注意到，了解一个人的个性特征可能反过来有助于其他任务，如欺骗检测。[59]使用AdaBoost[32]和随机森林进行个性检测，然后将其作为欺骗检测算法的输入。他们利用声音、语调和语言学特征（LIWC）进行人格预测。然后对每个LIWC特征的贡献进行了分析。结果发现，对于NEU-分数，金钱和权力维度是最有用的，对于EXT-分数，focusfuture和drives，对于CON-分数，时间和工作是最高的。这些大部分都是直观的，显示了使用LIWC特征进行个性检测的力量。该模型在纳入被试的个性特征后，在欺骗检测方面表现更好，并在哥伦比亚X-文化欺骗语料库中进行了评估。

#### 4.3 视觉

面相学[54]是一门从身体特征，特别是面部特征来确定个人性格或人格特征的艺术。研究人员发现，脸部为感知人格特征的推断提供了大部分的描述性信息[8, 105]。[54]给出了一种非深度学习的方法，它涉及到对人格预测的面部特征的分析。他们的重点是从图像中提取物理特征，如鼻子的形状，身体的形状（脂肪，肌肉，瘦），眉毛的形状，并找到它们与个性的相关性。有时，性格相似的人往往会有相似的偏好。[21]对这一观点进行了探讨，并研究了一个人的审美偏好（例如，他们喜欢的图片）和他们的个性之间是否有关系。从图片中提取的特征可以分为两类，审美（色彩丰富、光线的使用、GLCM-特征等）和内容（人脸、物体、GIST描述符），然后将其输入套索回归器。

[62]从一个人的推特简介照片的选择中预测了他的个性特征。该模型是在超过66,000名用户的数据上进行训练的，这些用户的性格类型是通过他们之前的推特内容预测的，使用的是最先进的文本模式技术。我们发现个人资料照片的美学和面部特征（使用face++和EmoVu API提取）与用户的个性特征之间存在关联性。认同感和自觉性的用户在他们的个人资料照片中表现出更多的积极情绪，而开放性高的用户则更喜欢有美感的照片。用户的自我表现特征由三维脸部姿势捕捉，其中包括脸部的俯仰、滚动和偏航角度以及眼睛的睁开程度（使用EmoVu API）。吸引人的图片往往有更高的对比度、清晰度、饱和度和更少的模糊，开放度高的人就是这种情况。

与其他许多计算机视觉问题一样，深度CNN在图像中的个性检测任务中表现得非常出色。然而，尽管在性能方面取得了成功，CNN被广泛认为是一个神奇的黑盒子。人们不知道在神经网络的各个隐藏层中出现的各种内部表示是什么。[99]对理解为什么CNN模型在这个复杂问题上表现得出奇的好进行了深入研究。他们利用类激活图来进行VI-

Sealization，并观察到激活总是围绕着脸部的，如图2所示。该神经网络特别关注面部区域，主要是与眼睛和嘴巴相对应的区域，以区分不同的个性特征并预测其价值。

我们在CNN中走得越深，各层就越擅长于识别高水平的特征，如眼睛、鼻子、眉毛等。几个动作单元（AU），如扬起的眉毛、酒窝等（FACS的子集），并观察到AU和感知的个性之间的良好关联。AU的特征向量被直接送入一个简单的线性分类器中，并取得了接近技术水平的准确度。

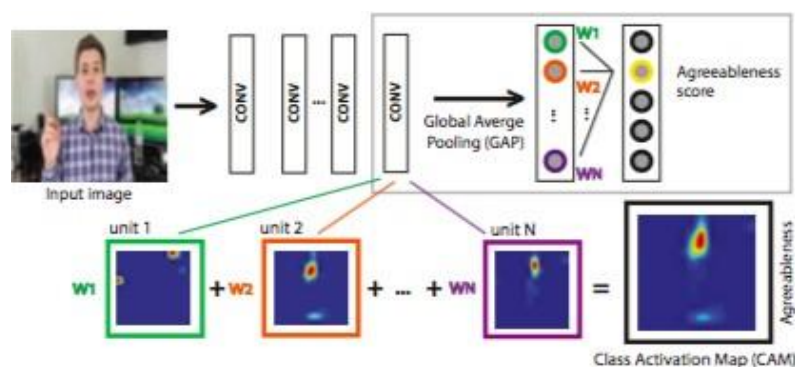


图2 类激活图，由[99]用于解释CNN模型。

[38]使用片段水平的闭塞分析研究了图像的不同部分对个性特征预测的贡献。用深度神经网络将视频帧分割成以下六个区域：背景区域、头发区域、皮肤区域、眼睛区域、鼻子区域和嘴巴区域。据观察，每个区域至少调节一个特征，不同的特征被不同的区域所调节。背景、皮肤和嘴部区域调节的性状最少。背景区域的闭塞增加了外向性特征，但减少了自觉性特征。皮肤区域的闭塞降低了合意性和自觉性特征。嘴部区域的闭塞增加了神经质和开放性特征。最后，眼睛区域对合意性特征有负向调节作用，而对神经质和开放性特征有正向调节作用。如图3所示，脸部区域和包括服装、姿势和身体部位在内的背景区域都被证明包含了关于感知的人格特征的信息。

心理学研究表明，人类在见到另一个人后，仅在几毫秒内就能推断出他的特质判断，也就是从第一印象开始。第一印象

表3 重要的人格探测模型的简要描述。

模态纸架构		
文本	[64] [83] [43] [15]	1D卷积提取n-grams与Mairesse特征相结合。 将LIWC与从文本中提取的MRC特征以及常识性知识结合在一起，使用Sentic计算技术。 GloVe嵌入到一个深度RNN（LSTM）中。 使用Mairesse特征的子集和无监督学习方法来衡量特征和人格特征之间的关联性
音频	[98] [79] [59]	提取了音频特征，如声调特征、语音活动特征、单词n-grams和对话行为标签，并将其输入简单的ML分类器中。 语音的epstral特征，如MFCC、零交叉率（ZCR）、强度、音调、响度、声调被送入SVM回归器。 将LIWC和发音特征结合起来，输入AdaBoost和RandomForest等ML分类器。
视觉	[86] [5] [39] [2]	用来建立脸部模型的直方图（HOG）、特征脸（EigenFace）和脸部的特定点，被送入SVM、梯度提升、bTree 用于提取面部表情的CERT 随后进行阈值处理以及隐马尔科夫模型 转移学习，预先训练VGG-Face和VGG-19来提取面部和背景特征，然后用内核ELM分类器进行正则化回归。 用于SVM的多项式和径向基函数（RBF）核子
多式联运	[38] [56]	通过VGG-Face提取面部特征，通过VGG-VD19提取场景特征，并通过基于随机森林的分数级融合与音频特征相结合。 同时学习领导力和外向性的多任务学习



图3 分段级的闭塞分析。每张图片显示了对应性状的预测变化与预定区域的函数关系[38]。

和感知个性实际上是非常相似的。在多模态感知个性识别领域，最重要和最广泛使用的数据集之一是ChaLearn First Impressions数据集。它包括从3000多个不同的YouTube高清视频中提取的10,000个片段（平均持续时间为15秒），这些片段是人们面对镜头说话。[18]提出了几种消除大型数据集标签中工人偏见的方法，并在ChaLearn数据集上测试了他们的方法。[39]利用该数据集的视觉模式来测试他们的模型，该模型结合了预先训练好的深度CNN所提取的面部和环境特征。这些特征然后被送入一个内核极端学习机器调节器[46]。这项工作的主要贡献是结合了面部和环境特征，同时有效地利用了迁移学习。

一个著名的检测图像中人脸的算法是Viola Jones算法。[2]在将图像转换为灰度、直方图均衡化（增加对比度）和重新缩放后，在预处理阶段使用该算法。然后，提取EigenFace[97]的特征，并将其输入到根据顺序最小优化方法训练的RBF SVM。该方法在一个数据库中进行了测试，该数据库是通过让法官独立地给FERET语料库中的图像贴上感知的个性特征标签而创建的。假设法官们没有民族或种族偏见。

[86]研究了面部各点之间是否存在特定的结构关系，以预先决定对面部特征的感知。他们对两种方法进行了比较，一种是使用面部的整体外观和纹理（整体方法）[85]，另一种是只使用特定的面部标靶点（结构方法）的localizations。整体方法以两种方式捕获面部外观信息。首先，通过分析像素信息的EigenFaces[97]方法；其次，利用定向梯度直方图（HOG），捕捉面部外观和形状。这些提取的特征随后被送入标准的ML分类器，如SVM或BTrees。本文没有使用五大人格模型，而是预测了9种不同的人格特征，如值得信赖的、主导的、外向的、威胁的、能力的、喜欢的、刻薄的、有吸引力的和可怕的。据观察，与EigenFace相比，HOG模型在面部特征和人格特征之间发现了更多的关联性。

Vlog是指主要以视频形式发布的博客。[5]分析了Vlogger的面部表情和个性之间的相关性。在这篇论文中，计算机表情识别工具箱（CERT）被用来确定博客作者的面部表情。CERT预测了7种普遍的面部表情（喜悦、厌恶、惊讶、中立、悲伤、蔑视和恐惧）。他们发现，具有积极价值的情绪的面部表情，如'喜悦'和'微笑'，几乎完全显示出与人格印象的积极影响。例如，"微笑"与外向性和顺从性等相关联。

尽管上述方法产生了良好的结果，但它们的计算成本非常高。[26]使用一种基于时间的面部纹理的方法来分析用于工作筛选的面部视频，这在计算上要经济得多。在预处理步骤之后，包括人脸识别、二维姿势校正和裁剪感兴趣区域（ROI），使用局部相位量化（LPQ）和二值化统计图像特征（BSIF）提取图像的纹理特征。这些特征被送入Pyramid多层，然后是五个非线性支持向量回归器，五大人格特征各一个。最后，这种感知到的个性预测被输入到高斯过程回归（GPR）中，该回归预测了一个二进制值，对应于候选人是否应该被叫去参加工作面试。

#### 4.4 二元制

大多数用于个性特征识别的多模态架构是双模态的，融合了音频和视觉模态的特征。[56]利用多任务学习从Emergent

Leader（ELEA）语料库的会议视频中提取的音频和视觉模式来识别外向和领导特征。预测外向性特征已经吸引了心理学家的注意，因为它能够解释广泛的行为，预测表现，也能评估个人的风险。该网络结构最初是在较大的VLOG语料库上训练的，然后在ELEA语料库上训练，因此有效地利用了迁移学习。特征选择技术，如最大相关度最小冗余（MRMR）过滤特征选择，使性能略有提高（约2%）。

深度剩余网络（DRN）用于感知个性分析的可视化是由[38]完成的，重点是解释这个端到端的深度架构的工作原理。这个DRN是一个双流网络，一个流用于视觉模式，另一个流用于音频模式。被感知的人格特征识别模型依靠什么信息来进行预测？[38]得出的结论是，所有特质水平都很高的女性面孔似乎

与那些水平较低的人相比，他们的颜色更均匀，对比度更高。此外，在观察单性别平均脸时，观察到高等级的脸偏向于女性，低等级的偏向于男性，在基于注释的平均脸的情况下更是如此。例如，所有特征的高水平都更加明亮和多彩，而且表情更加积极，与[103]观察到的结果一致。

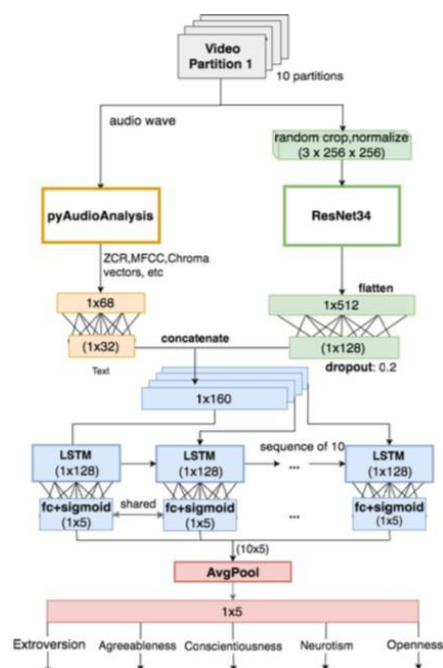


图4 [107]使用的深度双模态LSTM模型架构。

深度双模态回归（DBR）[111]框架在2016年ChaLearn挑战赛中取得了最高的准确性，用于感知个性分析。DBR由三个主要部分组成：视觉模态回归、音频模态回归和最后一个部分是融合两种模态信息的集合过程。该网络是在不使用任何特征引擎的情况下进行端到端的训练。传统的CNN架构通过放弃全连接层进行了修改。同时，最后一个卷积层的深度描述符被平均和最大池化，然后连接成一个1024D的向量（512个值来自最大池，其他512个值来自平均池）。最后，加入一个回归（fc+sigmoid）层进行端到端训练。这个修改后的CNN模型被称为描述符聚合网络（DAN）。对于音频模式，从每个视频的原始音频中提取对数滤波器库（logfbank）[24]特征。基于logfbank的特征，训练一个线性回归器以获得Big-Five特征预测。最后，通过对视觉和音频模型的五个预测分数进行平均，将这两种模式进行融合。



[84]提出了一个计算效率更高的双模回归框架，但准确性略有下降。从每个视频中提取多个音频片段（2-

3秒），而不是单一的15s音频片段，这增加了训练样本的数量，从而获得了更好的结果。FFMPEG用于从原始视频中提取音频片段，然后用OpenSMILE框架来提取音频特征。有许多研究[90,

51]声称五大人格特征并不是相互独立的。一个特定特质的强度是由两个成分的组合来量化的，即一个全局成分和一个特定特质成分。在这个模型中，每个特质模型都是通过全局成分和特质成分来学习的。我们创建了六个模型来预测五个人格特质，其中一个模型代表全局成分，其他五个代表特质的成分。

[3]将音频和文本模式结合起来进行个性检测。作者使用了声学-声音低级描述符（LLD）特征、来自LIWC的词类特征以及来自语言情感词典（DAL）的愉悦度、激活度和意象的词得分[104]。然后，这些特征的串联被送入一个MLP。他们将其与一个由词嵌入和LSTM组成的网络进行了比较，结果表明双模MLP网络在myPersonality语料库中表现最好。MLP在面临词汇外的情况下能更好地生成。这表明了声学-

发音特征的前景，它在语言方面更加稳健。他们还发现，早期和后期融合特征的模型取得了相似的性能。

[93]利用时间上有序的深度音频和随机的视觉特征进行双模的第一印象识别。本文的新想法是通过使用基于体积三维卷积[50]的深度神经网络工作来预测感知的个性特征。音频和视觉特征的时间模式是通过基于LSTM的深度RNN学习的。这两个模型在后期将从音频和视觉数据中提取的特征串联起来，最后送入一个全连接层。

[107]采用了类似的方法，试图通过使用LSTM对时间依赖性进行建模。音频特征如ZCR、MFCC和色度向量是用PyAudio Analysis提取的，而图像在随机裁剪和归一化后被送入ResNet34[42]架构。该架构如图4所示，在自觉性和开放性特征方面取得了最先进的性能。[40]通过将提取的特征向量送入极端学习机（ELM），然后融合它们的预测，在这个领域实现了最先进的性能。

声音特征是通过流行的openSMILE工具提取的，而预先训练好的深度卷积神经网络VGG-VD-

19[91]被用来从图像中提取面部情绪和环境信息。来自三个正交平面的局部Gabor二元模式（LGBP-

TOP）[112]与VGGFace一起被用于面部特征提取。用一个声学和三个视觉子系统的多级融合达到最佳效果，这些子系统被送入ELM，然后与分数级融合结合。

除了对五大特征的感知个性分析外，个人的态度，如娱乐、不耐烦、友好、热情、沮丧等，也可以用类似的架构进行分析。[63]利用声音和视觉特征分析，自动识别博客作者的态度。使用TCL/TK脚本提取音频的声音特征，如音调、强度、质量、话语的持续时间等。态度的声调分析显示不耐烦的音调最高，而沮丧的音调最低。然而，对视觉模式的分析只是通过将脸部地标（眉毛、嘴巴等）的绝对运动与态度相关联系起来完成。

[113]是唯一一篇通过计算研究人格特征对情绪影响的论文。作者将生理信号与人格特征联合起来进行建模。他们提出了顶点加权的模态多任务超图学习（VM2HL）学习模型，其中顶点是（主体，刺激物）对。Hyperedges制定了人格和生理信号之间的关系。通过引入在训练过程中学习的权重，赋予超图、模态和顶点不同的重要性。

#### 4.5 三角形

虽然大部分的工作都集中在音频和视觉输入的融合上，[37]创建了一个额外的流，纳入了来自文本模式的输入。这些个性预测然后被用于自动筛选求职者。他们在利用只有2个隐藏层的全连接神经网络的同时，取得了与最先进水平相当的性能。面部特征是用OpenFace[4]提取的，这是一个开源的面部特征提取工具箱。音频特征，如MFCC、ZCR、频谱能量分布和发言率，是用OpenSMILE[30]提取的，SenticNet用于文本的极性检测。

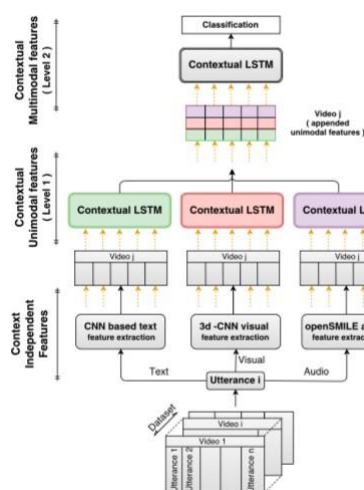


图5 用于提取依赖语境的多模态语篇特征的层次结构[81]。

[81]进行了多模态情感分析，并利用LSTMs来捕捉语境和语篇之间的依赖关系。最初，从每个模态中提取独立于环境的特征。使用3D-CNN提取视觉模态的特征，使用OpenSMILE从音频和Word2Vec嵌入中捕捉低级描述符，然后使用CNN提取文本模态。为了捕捉跨语料的信息触发流，在一个分层方案中使用了基于LSTM的RNN。最后，将其送入密集层以获得最终输出。这种结构表现得更好

表4 最先进的方法在流行的个性检测数据集上的表现。

测量平均最佳		模式 准确率论文	集个性	
文本	论文一 [74]	大五	57.99	[64]
	论文二 [95]	大五	63.6	[83]
	MBTI Kaggle	MBTI	67.77	[43]
	意大利FriendsFeed [16]	大五	63.1	[15]
音频	AMI会议语料库[13]	大五	64.84	[98]
	CMU自录数据库[79]	大五	-	[79]
	哥伦比亚的欺骗语料库[58]	-	-	[59]
视觉	面部评估 [69]	-	-	[86]
	YouTube博客	大五	0.092 (R-squared)	[5]
	ChaLearn第一印象数据集[80]。	大五	90.94	[39]
	彩色FERET数据库[78]	大五	65	[2]
多式联运	ChaLearn第一印象数据集[80]。	大五	91.7	[38]
	ELEA (Emergent LEAder) 语料库 [88]	领导力/外向性	81.3	[56]

比IEMOCAP[9]、MOUD[76]和MOSI[110]数据集上的技术水平更高，如图5所示。

在多模态数据上进行训练的深度网络一般都是非常密集的计算。[102]提出了一个多模态神经网络架构，该架构将不同类型、不同大小的输入数据与判别相关分析(DCA) [41]特征融合层相结合，计算复杂度低。融合后的特征被用作混合密度网络的输入层，以调整由于特征融合而造成的信息损失。最后的预测是使用先进的梯度提升算法建立的级联网络进行的。

而不是为5种人格特征中的每一种都有不同的训练过的神经网络。

[55]用多类方法训练一个单一的神经网络，有3个通道（每种模式各一个）。使用CNN从3种模式中的每个模式中提取特征。作者试验了使用早期融合（连接向量以获得输入数据的共享表示）和后期融合（每个模态进行单独预测，网络的预测是这些预测的一些加权平均）。

## 5 结果和讨论

表4提供了对流行的人格检测数据集的分类性能的系统概述。在过去的几年里，在人格检测领域发表的作品数量有了很大的增长。这是由于人格检测的大量工业应用，如第1.2节所述。最近，系统注释的数据集已经公开，因此，现在的研究可以集中在创建更好的模型和架构上，而不是在数据采购和预处理上。CVPR第一印象数据集[22]是个性检测中最流行的多模态数据集。我们看到有很多大五人格测量的数据集，但是其他的人格测量（如MBTI、PEN、16PF）却缺乏资源。人格测试的可靠性受到了质疑，因为有时会出现这样的情况：当一个人参加测试时，它给出了某种人格，然而当同一个人再次参加同样的测试时，它又给出了不同类型的人格。[28, 27]研究计算机视觉背景下深度学习模型的可解释性和可解释性

用于观察人的任务。MBTI人格测量是目前全世界最流行的人格测量。[33]研究了受试者的实际测试分数和他们对这些分数的估计之间的关系。结果显示，在大多数维度上都有明显的正相关关系。MBTI特质比Big-Five特质更难预测，也更复杂[34]，但我们看到目前的深度学习模型在这个领域也开始取得良好的准确性。

我们看到，个性检测方面的技术水平已经通过使用深度学习技术以及多模态融合的特征来实现。涉及双模态融合的技术目前非常流行。我们可以看到，在人格检测的三模态融合方面还有很大的研究和探索空间，因为目前很少有作品对所有三种模态的融合进行探索。如果一个人的个性可以被更可靠地预测，那么在几乎所有处理人机交互的代理中，如语音助手、机器人、汽车等，都有整合自动个性检测的空间。这个领域的研究正在从仅仅从文本数据中检测个性转向视觉和多模态数据。

我们已经看到，在当今的情况下，自动人格检测的实际潜力是巨大的。我们预计，本文讨论的许多深度学习模型将在未来几年内被用于工业应用。为了训练这些模型，需要有各种人格测量的大型公开数据集。在过去的5年里，我们看到用于自动人格检测的深度学习模型的数量有了极大的增长。大多数模式（尤其是视觉模式）的最佳性能都是由深度神经网络实现的。我们预计这一趋势将在未来继续下去，更智能、更强大的深度学习模型将随之而来。

## 6 总结

正如本文所讨论的，自动个性检测有许多不同的应用，可以在行业中使用，因此，它是一个非常热门和即将到来的领域。然而，机器学习模型的功能与用于训练它们的数据一样强大。在这个领域的大量案例中，没有足够的标记数据来训练庞大的神经网络。迫切需要更大、更准确和更多样化的数据集来进行人格检测。目前几乎所有的数据集都集中在大五人格模型上，很少有其他人格测量方法，如MBTI或PEN。通常情况下，人格是通过回答调查中的多个问题来测量的。假设每个参加调查的人都如实回答，那么这种调查在正确标示一个人的人格方面的可信度仍然是个问题。我们需要探索一种更准确、更有效的方法来标示人格特征。目前大多数创建人格检测数据集的方法都是通过使用亚马逊Mechanical Turk的众包来进行人工标注。

最近的多模态深度学习技术表现良好，开始进行可靠的人格预测。深度学习提供了一种利用大量数据和计算能力的方法，我们可以用很少的手工工程来处理。各种深度模型已经成为新的最先进的方法，不仅在人格检测方面，而且在其他领域也是如此。我们期待这一趋势继续下去，因为更深的模型和新的架构能够映射非常复杂的功能。我们希望看到更多的依赖于高效的多模态融合的个性检测架构。

## 7 鸣谢

我们要感谢Bharat M Deshpande教授的宝贵指导。A. Gelbukh感谢Instituto Politecnico Nacional通过Secretaria de Investigacion y Posgrado项目SIP 20196437和SIP 20196021的支持。

## 参考文献

1. Agarwal, B.: 文本中的个性检测。A review. *International Journal of Computer System* **1**, 1-4 (2014)
2. Al Moubayed, N., Vazquez-Alvarez, Y., McKay, A., Vinciarelli, A.: 基于面部的自动人格感知。在。第22届ACM国际多媒体会议论文集, 第1153-1156页。ACM (2014)
3. An, G., Levitan, R.: 用于人格识别的词汇和声学深度学习模型。In: *Proc. Interspeech 2018*, pp. 1761-1765 (2018). DOI 10.21437/Interspeech.2018-2263. URL <http://dx.doi.org/10.21437/Interspeech.2018-2263>
4. Baltrusaitis, T., Robinson, P., Morency, L.P.: Openface: 一个开源的面部行为分析工具箱。在。计算机视觉的应用 (WACV), 2016年IEEE冬季会议, 第1-10页。IEEE (2016)
5. Biel, J.I., Teijeiro-Mosquera, L., Gatica-Perez, D.: Facetube: 从在线对话视频中的面部情绪预示预测个性。在。第14届ACM国际多模态交互会议论文集, 第53-56页。ACM (2012)
6. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. *机器学习研究杂志* **3**(Jan), 993-1022 (2003)
7. Briggs Myers, I.: 类型介绍。了解你在迈尔斯-布里格斯类型指标上的结果的指南 (由LK KIRBY和KD迈尔斯修订)。帕洛阿尔托 (1993)。
8. Bruce, V., Young, A.: 了解人脸识别。英国心理学杂志 **77**(3), 305-327 (1986)
9. Busso, C., Bulut, M., Lee, C.C., Kazemzadeh, A., Mower, E., Kim, S., Chang, J.N., Lee, S., Narayanan, S.S.: Iemocap: 交互式情感双人动作捕捉数据库。语言资源和评估 **42**(4), 335 (2008)
10. Cambria, E., Poria, S., Hazarika, D., Kwok, K.: SenticNet 5: 通过上下文嵌入发现用于情感分析的概念原素。In: *AAAI*, 第1795-1802页 (2018)。
11. Caprara, G.V., Schwartz, S., Capanna, C., Vecchione, M., Barbaranelli, C.: 个性与政治。价值观、特质和政治选择。政治心理学 **27**(1), 1-28 (2006)
12. Carletta, J.: 释放杀手锏语料库: 创建多事之秋会议的经验 语料库。语言资源和评估 **41**(2), 181-190 (2007)
13. Carletta, J., Ashby, S., Bourban, S., Flynn, M., Guillemot, M., Hain, T., Kadlec, J., Karaiskos, V., Kraaij, W., Kronenthal, M., et al: Ami会议语料库。预告。在。国际研讨会 on Machine Learning for Multimodal Interaction, pp. 28-39. Springer (2005)
14. Cattell, H.E., Mead, A.D.: 16种人格因素问卷 (16pf)。The SAGE handbook of personality theory and assessment **2**, 135-178 (2008)
15. Celli, F.: 社交网站的无监督个性识别。In: 第六届国际数字社会会议 (2012)。
16. Celli, F., Di Lascio, F.M.L., Magnani, M., Pacelli, B., Rossi, L.: 社交网络数据和实践: friendfeed的案例。在。社会计算、行为建模和预测的国际会议, 第346-353页。Springer (2010)
17. Celli, F., Poesio, M.: Pr2: 一个独立于语言的无监督工具, 用于从文本中识别人格。arXiv预印本arXiv:1402.2796 (2014)
18. Chen, B., Escalera, S., Guyon, I., Ponce-López, V., Shah, N., Són, M.O.: 克服配对评级的模式标签的校准问题: 应用于人格特征。在。European Conference on Computer Vision, 第419-432页。Springer (2016)
19. Chittaranjan, G., Blom, J., Gatica-Perez, D.: Who's who with big-five: 用智能手机分析和分类personality特征。在。可穿戴计算机 (ISWC), 2011年第15届国际研讨会, 第29-36页。IEEE (2011)
20. Chung, J., Gulcehre, C., Cho, K., Bengio, Y.: 门控递归神经网络在序列建模上的经验评估。arXiv预印本arXiv:1412.3555 (2014)

21. Cristani, M., Vinciarelli, A., Segalin, C., Perina, A.: 揭示多媒体无意识：隐性认知过程和多媒体内容分析。在。第21届ACM国际多媒体会议论文集, 第213-222页。ACM (2013)
22. CVPR. Chalearn数据集看人挑战 (2017)。URL <http://chalearnlap.cvc.uab.es/dataset/24/description/>。
23. Dalal, N., Triggs, B.: 用于人类检测的定向梯度直方图。In: 计算机视觉和模式识别, 2005.CVPR 2005.IEEE计算机学会会议, 第一卷, 第886-893页。IEEE (2005)
24. Davis, S.B., Mermelstein, P.: 连续口语句子中单音节词识别的参数表示的比较。In: 语音识别的读物, 第65-74页。Elsevier (1990)
25. Digman, J.M.: 人格结构。五因素模型的出现。心理学年度评论 **41**(1), 417-440 (1990)
26. Eddine Bekhouche, S., Dornaika, F., Ouafi, A., Taleb-Ahmed, A.: 通过分析面部视频进行人格特征和求职者筛选。在。IEEE计算机视觉会议论文集和模式识别研讨会, 第10-13页(2017)
27. Escalante, H.J., Guyon, I., Escalera, S., Jacques, J., Madadi, M., ̇, X., Ayache, S., Viegas, E., ̇çlütürk, Y., ̇çli, U., 等人。设计一个可解释的机器学习挑战, 用于视频互看。在: 2017年国际神经网络联合会议 (IJCNN) 上, 第3688-3695页。IEEE (2017)
28. Escalante, H.J., Kaya, H., Salah, A.A., Escalera, S., Gucluturk, Y., Guclu, U., ̇, X., Guyon, I., Junior, J.J., Madadi, M., 等人。解释第一印象: 建模、识别和解释 视频中的明显个性。 arXiv preprint arXiv:1802.00745 (2018)
29. Escalera, S., ̇, X., Guyon, I., Escalante, H.J.: 客座编辑: 明显的个性分析。 IEEE Transactions on Affective Computing (3), 299-302 (2018)
30. Eyben, F., Wollmer, M., Schuller, B.: Opensmile: munich多功能和快速的开源音频特征提取器。在。第18届ACM国际多媒体会议论文集, 第145 9- 1462页。ACM (2010)
31. Eysenck, H.J.: 人格的模型。Springer Science & Business Media (2012)
32. Freund, Y., Schapire, R.E., et al: 一个新的提升算法的实验。In: Icml, vol. 96, pp. 148-156.意大利巴里 (1996)
33. Furnham, A.: 人们能准确估计自己的人格测试分数吗? 欧洲 人格杂志**4**(4), 319-327 (1990)
34. Furnham, A.: The big five versus the big four: the relationship between the myers-briggs type indicator (mbti) and neo-pi five factor model of personality.人格与个体差异**21**(2), 303- 307 (1996)
35. Golbeck, J.: 从社交媒体文本预测个性。AIS Transactions on Replication Research **2**(1), 2 (2016)
36. Gonzalez-Gallardo, C.E., Montes, A., Sierra, G., Nunez-Juarez, J.A., Salinas-Lopez, A.J., Ek, J.: 使用语料库相关标签、字符和位置n-grams进行推文分类。在。CLEF (Working Notes) (2015)
37. Gorbova, J., Lusi, I., Litvin, A., Anbarjafari, G.: 基于多模式视频处理的求职者自动筛选。In: IEEE计算机视觉和模式会议论文集识别研讨会, 第29-35页(2017)
38. Gucluturk, Y., Guclu, U., Perez, M., Balderas, H.J.E., Baro, X., Guyon, I., Andujar, C., Junior, J., Madadi, M., Escalera, S., et al: 用深度残差网络实现明显的个性分析的可视化。在。国际计算机视觉会议-ICCV 2017 (2017)
39. Gurpinar, F., Kaya, H., Salah, A.A.: 结合深度面部和环境特征进行第一印象。In: 欧洲计算机视觉会议, 第372-385页。斯普林格(2016)
40. Gurpinar, F., Kaya, H., Salah, A.A.: 音频、场景和面部特征的多模态融合, 用于第一印象的估计。In: 模式识别 (ICPR), 2016年第23届国际会议, 第 43-48。IEEE (2016)
41. Haghighat, M., Abdel-Mottaleb, M., Alhalabi, W.: 判别性相关分析。用于多模态生物识别的实时特征水平融合。IEEE Transactions on Information Forensics and Security **11**(9), 1984-1996 (2016)
42. He, K., Zhang, X., Ren, S., Sun, J.: 图像识别的深度残差学习。In: the IEEE conference on computer vision and pattern recognition, pp.770-778 (2016)
43. Hernandez, R.K., Scott, I.: 用文本预测Myers-Briggs类型指标
44. Hirsch, H.G., Pearce, D.: 噪声条件下语音识别系统性能评估的极光实验框架。在。ASR2000-自动语音识别。新千年的挑战 ISCA教程和研究研讨会(ITRW) (2000)

45. Hoerl, A.E., Kennard, R.W.: Ridge regression:非正交问题的有偏估计。Tech- nometrics **12**(1), 55-67 (1970)
46. Huang, G.B., Zhou, H., Ding, X., Zhang, R.: Extreme learning machine for regression and multiclass classification.IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) **42**(2), 513-529 (2012)
47. Ilmini, K., Fernando, T.L.: 使用机器学习算法进行人员'个性特征识别 和图像处理技术 (2016)
48. 爱尔兰, M.E., Pennebaker, J.W.: 写作中的语言风格匹配。作文、相关性 dence和诗歌中的同步性。人格与社会心理学杂志》**99**(3), 549 (2010)
49. Jaynes, E.T.: On the rationale of maximum-entropy methods.IEEE **70**(9), 939-952 (1982)
50. Ji, S., Xu, W., Yang, M., Yu, K.: 人类动作识别的三维卷积神经网络。IEEE 模式分析和机器智能交易**35**(1), 221-231 (2013)
51. Judge, T.A., Higgins, C.A., Thoresen, C.J., Barrick, M.R.: 五大人格特质、一般心理, 以及跨越人生的职业成功。人事心理学**52**(3), 621-652 (1999)
52. Junior, J., Jacques, C., Güçlürk, Y., Pérez, M., Güçli, U., Andujar, C., B, X., Escalante, H.J., Guyon, I., van Gerven, M.A., 等人。第一印象。关于基于计算机视觉的表象 人格特征分析的调查。arXiv预印本 arXiv:1804.08046 (2018)
53. Kalghatgi, M.P., Ramannavar, M., Sidnal, N.S.:基于大五模型的人格预测的神经网络方法。国际先进工程创新研究杂志 (IJIRAE) **2**(8), 56-63 (2015)
54. Kamenskaya, E., Kukharev, G.: 从面部识别心理特征。Metody Infor- matyki Stosowanej **1**(1), 59-73 (2008)
55. Kampman, O., Barezi, E.J., Bertero, D., Fung, P.:研究音频、视频和文本融合方法, 用于端到端的自动人格预测。在。计算语言学协会第56届年会论文集(第2卷: 短文), 第2卷, 第606-611页(2018)。
56. Kindiroglu, A.A., Akarun, L., Aran, O.:多领域和多任务预测会议视频中的外向性和领导力。EURASIP图像和视频处理杂志**2017**(1), 77 (2017)
57. Kohonen, T.: The self-organizing map.IEEE **78**(9), 1464-1480 (1990)
58. Levitan, S.I., An, G., Wang, M., Mendels, G., Hirschberg, J., Levine, M., Rosenberg, A.: 跨文化语音欺骗的产生和检测。在。2015年ACM关于 多模态欺骗检测研讨会的论文集, 第1-8页。ACM (2015)
59. Levitan, S.I., Levitan, Y., An, G., Levine, M., Levitan, R., Rosenberg, A., Hirschberg, J.: 从对话中识别性别、种族和个性的个体差异, 以进行欺骗检测。在。第二届欺骗检测的计算方法研讨会论文集, 第40-44页 (2016)
60. Liem, C.C., Langer, M., Demetriou, A., Hiemstra, A.M., Wicaksana, A.S., Born, M. P., König, C.J.: 精神病学与机器学习。关于求职者筛选算法的跨学科观点。在。计算机视觉和机器学习中的可解释和可诠释模型, 第197-253页。Springer (2018)
61. Ling, W., Luis, T., Marujo, L., Astudillo, R.F., Amir, S., Dyer, C., Black, A.W., Trancoso, I.: Finding function in form:arXiv preprint arXiv:1508.02096 (2015)
62. Liu, L., Preotiuc-Pietro, D., Samani, Z.R., Moghaddam, M.E., Ungar, L.H.: 通过社交媒体图片选择分析个性。In:ICWSM, 第211-220页 (2016)。
63. Madzlan, N.A., Han, J., Bonin, F., Campbell, N.: 自动识别视频博客中的态度--语调和视觉特征分析。在。国际语音 交流协会第十五届年会(2014年)。
64. Majumder, N., Poria, S., Gelbukh, A., Cambria, E.: 基于深度学习的文档建模, 用于从文本中进行每 sonality检测。IEEE智能系统**32**(2), 74-79 (2017)
65. Matthews, G., Deary, I.J., Whiteman, M.C.: 人格特质。剑桥大学出版社 (2003)
66. Mikolov, T., Chen, K., Corrado, G., Dean, J.: 矢量空间中词汇表征的高效估计。arXiv预印本arXiv:1301.3781 (2013)
67. Miles, J., Hempel, S.: Eysenck人格量表。艾森克人格问卷修订版(epq-r)和艾森克人格分析器(epp)。心理学评估综合手册- ment **2**, 99-107 (2004)
68. Mukherjee, S., Kumar, U.:人格评估中的伦理问题。The Wiley Handbook of Personality Assessment pp. 415-426 (2016)
69. Oosterhof, N.N., Todorov, A.: 人脸评估的功能基础。美国国家科学院院刊》**105**(32), 11087-11092 (2008)

70. Palaz, D., Magimai.-Doss, M., Collobert, R.: 使用原始语音作为输入的基于cnn的语音识别系统的分析。Tech.rep., Idiap (2015)
71. Papurt, M.J.:对伍德沃斯心理神经性清单的研究与建议的修订。该杂志的异常和社会心理学**25** (3), 335 (1930年)。
72. Park, G., Schwartz, H.A., Eichstaedt, J.C., Kern, M.L., Kosinski, M., Stillwell, D.J., Ungar, L.H., Seligman, M.E.: 通过社交媒体语言自动评估人格。人格杂志》 和社会心理学 **108**(6), 934 (2015)
73. Pennebaker, J.W., Francis, M.E., Booth, R.J.: 语言学调查和字数。Liwc 2001.Mahway:Lawrence Erlbaum Associates **71**(2001), 2001 (2001)
74. Pennebaker, J.W., King, L.A.: 语言学风格。语言使用是一种个体差异。人格和社会心理学杂志**77** (6), 1296 (1999)。
75. Pennington, J., Socher, R., Manning, C.: Glove:全局向量的单词表示。在。2014年自然语言处理经验方法会议 (EMNLP) 论文集, 第1532-1543页 (2014)
76. Perez-Rosas, V., Mihalcea, R., Morency, L.P.: 语段级多模态情感分析。In:计算语言学协会第51届年会论文集 (第1卷: 长篇论文), 第1卷, 第973-982页 (2013)。
77. Pervin, L.A., John, O.P.: 《人格手册》。理论与研究。Elsevier (1999)
78. Phillips, P.J., Wechsler, H., Huang, J., Rauss, P.J.: feret数据库和人脸识别算法的评估程序。图像和视觉计算**16**(5), 295-306 (1998)
79. Polzehl, T., Moller, S., Metze, F.: 从语音中自动评估人格。In:Semantic Computing (ICSC), 2010年IEEE第四届国际会议, 第134-140页。IEEE (2010)
80. Ponce-López, V., Chen, B., Oliu, M., Corneanu, C., Cs, A., Guyon, I., X., Escalante, H.J., Escalera, S.: Chalearn lap 2016:第一轮对第一印象的挑战-数据集和结果。In:欧洲计算机视觉会议, 第400-418页。斯普林格(2016)
81. Poria, S., Cambria, E., Hazarika, D., Majumder, N., Zadeh, A., Morency, L.P.: 用户生成的视频中的情境依赖性感觉分析。In:计算语言学协会第55届年会论文集 (第1卷: 长篇论文), 第1卷, 第873-883页 (2017)。
82. Poria, S., Cambria, E., Hazarika, D., Vij, P.: 使用深度卷积神经网络深入研究讽刺性推文。In:COLING, pp.1601-1612 (2016)
83. Poria, S., Gelbukh, A., Agarwal, B., Cambria, E., Howard, N.: 基于常识性知识的文本个性识别。In:Mexican International Conference on Artificial Intelligence, pp. 484-496.斯普林格(2013)
84. Rai, N.: 双模态回归的明显人格特征识别。In:模式识别 (ICPR), 2016年第23届国际会议, 第55-60页。IEEE (2016)
85. Roberto, B., Poggio, T.: 人脸识别。特征与模板。IEEE Transactions on pattern Analysis and machine Intelligence **15**(10), 1042-1052 (1993)
86. Rojas, M., Masip, D., Todorov, A., Vitria, J.: 面部特征判断的自动预测。外观 与结构模型。PloS one **6**(8), e23323 (2011)
87. Saez, Y., Navarro, C., Mochon, A., Isasi, P.: 一个个性和幸福感的检测系统。IJIMAI **2**(5), 7-15 (2014)
88. Sanchez-Cortes, D., Aran, O., Gatica-Perez, D.: 用于新兴领袖分析的音频视觉语料库。ICMI-MLMI), 机器学习的多模态语料库, 11月第14-18页 (2011)。
89. Schapire, R.E., Singer, Y.:Boostexter:一个基于提升的文本分类系统。机器学习- ing **39**(2-3), 135-168 (2000)
90. Shaver, P.R., Brennan, K.A.: 依恋风格和 "五大"人格特质。它们之间的联系以及与恋爱关系的结果。人格与社会心理学公报 **18**(5), 536-545 (1992)
91. Simonyan, K., Zisserman, A.: 用于大规模图像识别的深度卷积网络。 arXiv preprint arXiv:1409.1556 (2014)
92. Su, M.H., Wu, C.H., Zheng, Y.T.: 利用轮流发言的时间演化来实现双人对话中的个性特征感知。IEEE/ACM Transactions on Audio, Speech, and Language Processing **24**(4), 733-744 (2016)
93. Subramaniam, A., Patel, V., Mishra, A., Balasubramanian, P., Mittal, A.: 使用时间上有序的深度音频和随机视觉特征的双模式第一印象识别。In:European Conference on Computer Vision, pp.337-348.Springer (2016)
94. Sun, X., Liu, B., Cao, J., Luo, J., Shen, X.: 我是谁? 基于深度学习的个性检测, 用于 文本。In: 2018 IEEE International Conference on Communications (ICC), pp.1-6.IEEE (2018)
95. Tausczik, Y.R., Pennebaker, J.W.: 词语的心理学意义。Liwc和计算机化文本分析方法。语言和社会心理学杂志》 **29**(1), 24-54 (2010)



96. Tseng, S.C.: 处理普通话口语语料。 *Traitement Automatique des Langues* pp. 89-108 (2004)
97. Turk, M., Pentland, A.: Eigenfaces for recognition. *认知神经科学杂志* **3** (1), 71-86 (1991)
98. Valente, F., Kim, S., Motlicek, P.: 对阿米会议语料库中的口语交流中的人格特征进行注解和识别。在。第十三届国际演讲年会 交流协会 (2012)
99. Ventura, C., Masip, D., Lapedriza, A.: 解读Cnn模型的显式人格特质回归。In: 计算机视觉和模式识别研讨会 (CVPRW), 2017年IEEE E会议, 第 1705-1713。IEEE (2017)
100. Vinciarelli, A., Mohammadi, G.: A survey of personality computing. *IEEE Transactions on Affective Computing* **5**(3), 273-291 (2014)
101. Viola, P., Jones, M.J.: 健全的实时人脸检测。 *国际计算机视觉杂志* **57**(2), 137-154 (2004)
102. Vo, N.N., Liu, S., He, X., Xu, G.: 多模态混合密度提升网络的个性挖掘。In: *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pp.644-655. Springer (2018)
103. Walker, M., Vetter, T.: 改变一张脸的个性。感知的大二型和大五型人格, 在真实照片中建模。 *人格与社会心理学杂志* **110**(4), 609 (2016)
104. Whissell, C., Fournier, M., Pelland, R., Weir, D., Makarec, K.: A dictionary of affect in language: Iv. 可靠性、有效性和应用。 *知觉和运动技能* **62**(3), 875-888 (1986)
105. Willis, J., Todorov, A.: 第一印象。在接触到一张脸的100毫秒后就下定决心。 *心理科学* **17**(7), 592-598 (2006)
106. Yang, H.C., Huang, Z.R.: 从社交信息中挖掘个性特征, 用于游戏推荐系统- tems. *Knowledge-Based Systems* **165**, 157-168 (2019)
107. Yang, K., Mall, S., Glaser, N.: 用深度双模态LSTM预测个性第一印象 (2017)
108. Yin, H., Wang, Y., Li, Q., Xu, W., Yu, Y., Zhang, T.: A network-enhanced prediction method for auto- mobile purchase classification using deep learning (2018)
109. Yu, J., Markov, K.: 基于深度学习的脸书状态更新个性识别。In: *Awareness Science and Technology (iCAST)*, 2017 IEEE 8th International Conference on, pp.383-387. IEEE (2017)
110. Zadeh, A., Zellers, R., Pincus, E., Morency, L.P. 。 Mosi : 在线意见视频中的情感强度和主观性分析的多模态语料库。 *arXiv预印本arXiv:1606.06259* (2016)
111. Zhang, C.L., Zhang, H., Wei, X.S., Wu, J.: Deep bimodal regression for apparent personality analysis. In: 欧洲计算机视觉会议, 第311-324页。 Springer (2016)
112. Zhang, W., Shan, S., Gao, W., Chen, X., Zhang, H.: Local gabor binary pattern histogram sequence (lgbphs): a novel non-statistical model for face representation and recognition. In: 计算机视觉, 2005. ICCV 2005. 第十次IEEE国际会议, 第一卷, 第786-791页。 IEEE (2005)
113. Zhao, S., Gholaminejad, A., Ding, G., Gao, Y., Han, J., Keutzer, K.: 通过对生理信号的个性感知高阶学习实现个性化的情绪识别。 *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* **15**(1s), 14 (2019)
114. Zuo, X., Feng, B., Yao, Y., Zhang, T., Zhang, Q., Wang, M., Zuo, W.: 一个用于预测用户个性特征的加权ml-knn模型。 In: *Proc. Int. Conf. Inf. Sci. Comput. (ISCA)*, 第345-350页 (2013)