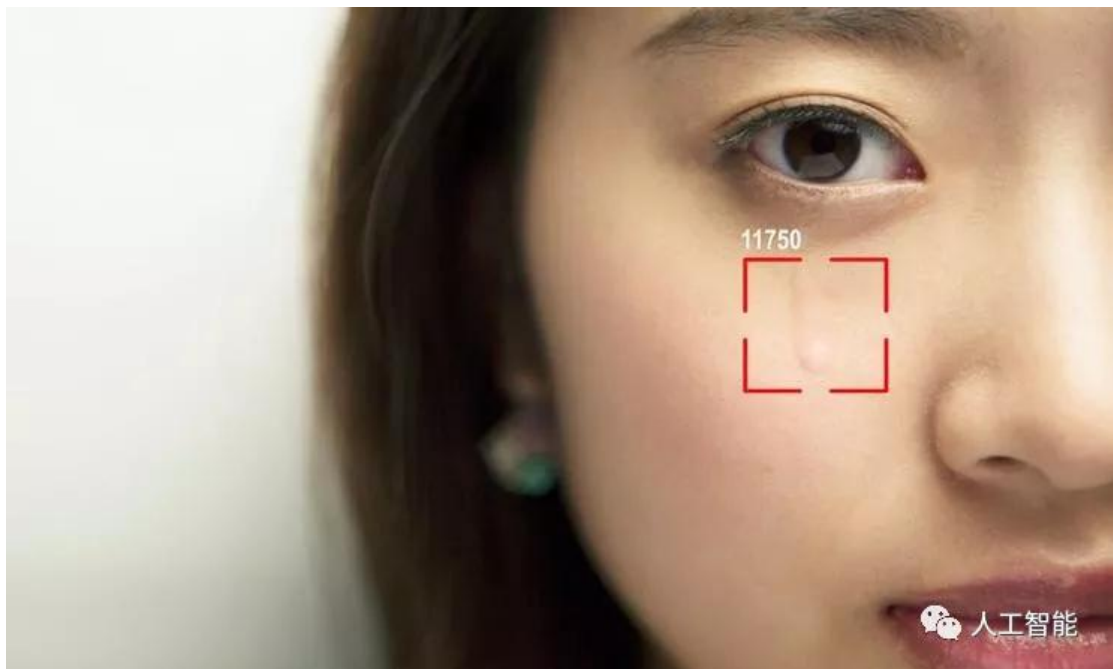


87 个国家、750 万个面孔数据，『情绪识别机器』催生 200 亿美元“财富梦”

人工智能 3 月 11 日

是否有一个程序可以通过读取潜在的恐怖分子的面部表情和行为来发现他们？

这是美国交通安全管理局（TSA）于 2003 年提出的一个假设，当时它开始测试一个新的监控程序，称为“通过观察技术筛查乘客”，简称 SPOT。



在制定该计划时，他们咨询了旧金山加利福尼亚大学心理学荣誉教授 Paul Ekman。

几十年前，Ekman 开发了一种识别微小面部表情并将其映射到相应情绪的方法。这种方法被用来训练“行为检测人员”，以扫描面部是否有欺骗迹象。

但是当该计划于 2007 年推出时，却遇到了问题。官员们或多或少地随意运用该技术对人进行审讯，而且少数逮捕行动是与恐怖主义无关。更令人担忧的是，该计划据称被用来证明种族相貌的合理性。

Ekman 试图将自己与 Spot 保持距离，声称他的方法被误用了。但其他人认为，该计划的失败是由于过时的科学理论支撑了 Ekman 的方法。

近年来，科技公司已经开始使用 Ekman 的方法来训练从面部表情中检测情绪的算法。一些开发人员声称，自动情绪检测系统不仅比人类更好地通过分析面部表情来发现真实的情绪，而且这些算法将适应我们内心的感受，极大地改善了与设备的交互。



但是，许多研究情绪科学的专家担心这些算法根据错误的科学理论会做出高风险的决定，也将可能再次失败。

识别你的脸：一个 200 亿美元的行业

情绪检测需要两种技术：

- 计算机视觉，精确识别面部表情；
-
- 机器学习算法，分析和解释这些面部特征的情绪内容。
-

通常，第二步采用一种称为监督学习的技术，通过这种技术，一个算法被训练来识别它以前见过的东西。基本的想法是，如果你在看到一张新的快乐脸谱图时，用标签“快乐”来显示成千上万张快乐脸谱图的算法，它会再次将其识别为“快乐”。

研究生 Rana el Kaliouby 是第一批开始尝试这种方法的人之一。2001 年，在从埃及搬到剑桥大学攻读计算机科学博士学位后，她发现自己花在计算机上的时间比其他人更多。她认为，如果她能教电脑识别情绪并对她的情绪状态做出反应，那么她远离家人和朋友的时间就不会那么孤独。

Kaliouby 致力于这项研究，并最终开发出一种辅助阿斯伯格综合症儿童阅读和回应面部表情的设备。她称之为“情绪助听器”。

2006 年，Kaliouby 加入麻省理工学院的情绪计算实验室，与实验室主任 Rosalind Picard 一起继续改进和完善该技术。然后，在 2009 年，他们共同创办了一家名为 Affectiva 的创业公司，这是第一家营销“人工情绪智能”的公司。

起初，Affectiva 将他们的情绪检测技术作为市场研究产品出售，为广告和产品提供实时的情绪反应。他们找到了客户，如 Mars、凯洛格和哥伦比亚广播公司等客户。

Picard 于 2013 年离开了 Affectiva，并加入了一家不同的生物识别创业公司，业务继续增长，围绕它的行业也是如此。

亚马逊、微软和 IBM 一直将“情绪分析”作为他们面部识别产品的主要功能进行宣传，并且许多小公司，如 Kairos 和 Eyeris 已经出现，为 Affectiva 提供类似的服务。



除了市场研究，情绪检测技术现在被用于监测和检测驾驶员损伤，测试视频游戏的用户体验以及帮助医疗专业人员评估患者的健康状况。

Kaliouby 看到情绪检测从一个研究项目发展成为一个价值 200 亿美元的行业，她对这种增长的持续性充满信心。她预测，在不久的将来，当这项技术无处不在并融入我们所有的设备中时，能够“利用我们的内心、潜意识做出瞬间的反应”。

来自 87 个国家的 750 万个面孔数据

与大多数机器学习应用程序一样，情绪检测的进展取决于访问更高质量的数据。

根据 Affectiva 的网站，他们拥有世界上最大的情绪数据存储库，来自 87 个国家的面孔，数量超过 750 万，其中大部分是从人们收看电视或开车每天上下班的选择性录像中收集的。



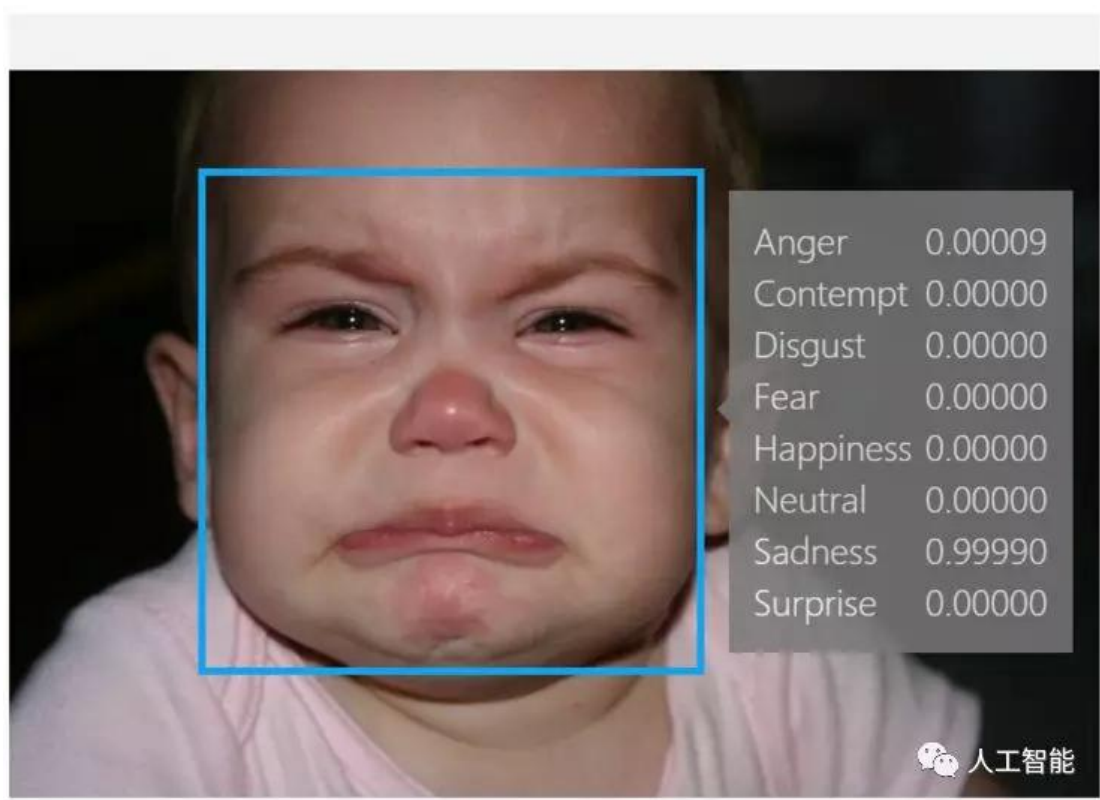
这些视频由位于开罗的 Affectiva 办公室的 35 名贴标人进行分类，他们观看镜头并将面部表情转化为相应的情绪。例如，如果他们看到低垂的眉毛、紧绷的嘴唇和凸出的眼睛，他们会附上标签“愤怒”。这个被标记的人类情绪数据集然后将用于训练 Affectiva 的算法，该算法学习如何将愤怒的脸、笑脸与幸福等联系在一起。

这种标记方法，被情绪检测行业的许多人认为是测量情绪的黄金标准，它由 Paul Ekman 和 Wallace V Friesen 在 20 世纪 80 年代开发的一种叫“情绪面部动作编码系统 (Emfac)”衍生而来。

这个系统的科学根源可以追溯到 20 世纪 60 年代，当时 Ekman 和两位同事假设有六种普遍的情绪——愤怒、厌恶、恐惧、快乐、悲伤和惊讶——这些情绪与我们有着紧密的联系，并且可以通过分析面部肌肉的运动在所有文化中被发现。

为了验证这一假设，他们向全世界不同的人群展示了面部照片，要求他们辨别他们所看到的情绪。他们发现，尽管存在巨大的文化差异，人类还是会将相同的面部表情与相同的情绪相匹配。一张眉毛低垂、嘴唇紧闭、眼睛凸出的脸对美国的银行家和巴布亚新几内亚的半游牧猎人来说都意味着“愤怒”。

在接下来的二十年里，Ekman 利用他的发现开发了一种识别面部特征并将其映射到情绪的方法。潜在的前提是，如果一个人的普遍情绪被触发，那么一个相关的面部动作就会自动出现在脸上。即使那个人试图掩饰他们的情绪，真实的本能的感觉也会“泄露”。



整个 20 世纪后半叶,这一理论被称为“经典情绪理论”,开始主导着情绪科学。Ekman 将他的情绪检测方法作为专利,并开始将其作为培训项目出售给 CIA、FBI、海关和边境保护局以及 TSA。真实情绪在脸上可读的观念甚至渗透到大众文化中,形成了“Lie to me”节目的基础。

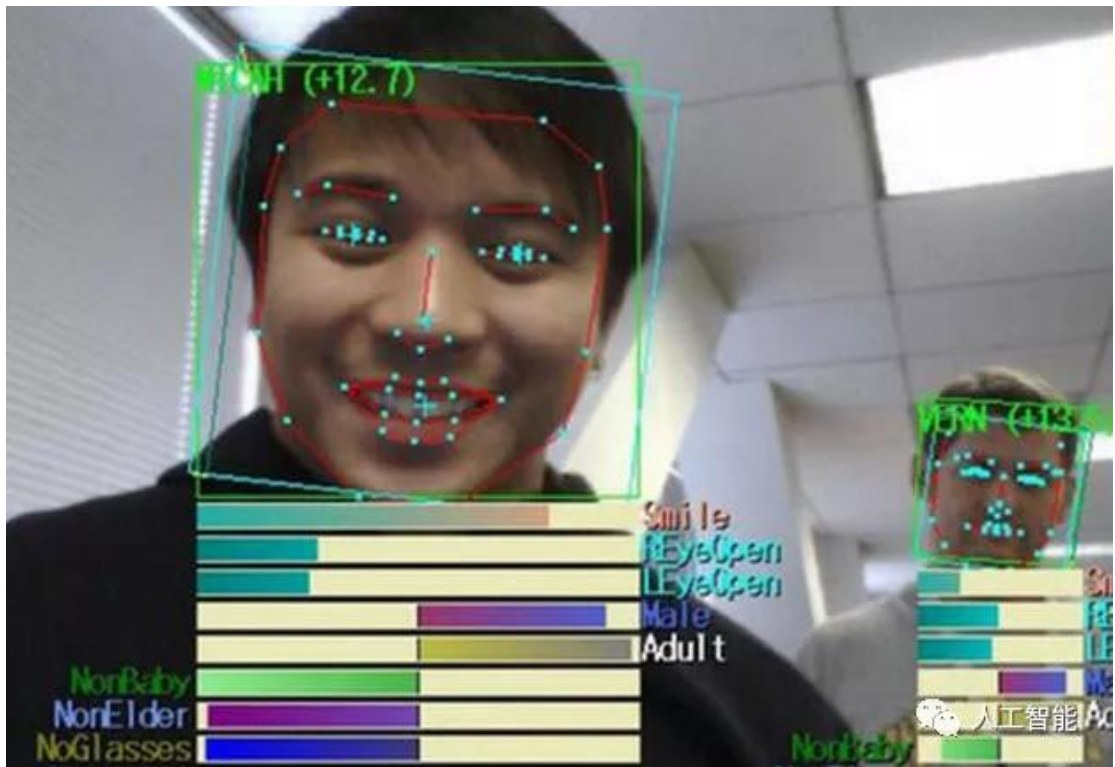
然而,许多研究情绪本质的科学家和心理学家质疑经典理论和 Ekman 的相关情绪检测方法。

情绪检测理论备受质疑

近年来,心理学教授 Barret 提出了一个特别强烈而持久的批评。

Barret 作为研究生首先遇到了古典理论。她需要一种客观地测量情绪的方法,并且遇到了 Ekman 的方法。在回顾文献时,她开始担心潜在的研究方法有缺陷——特别是,她认为通过为人们提供与照片匹配的预选情绪标签,Ekman 无意中“准备”了他们给出的某些答案。

她和一组同事通过重演 Ekman 的测试来检验这个假设,而不提供标签,让受试者可以自由地描述他们所看到的图像中的情绪。具体的面部表情和具体的情绪之间的关系直线下降。



从那时起，Barret 发展了她自己的情绪理论，这在她的书《情绪是如何产生的：大脑的秘密生活》中有所阐述。她认为，大脑中没有由外部刺激触发的普遍情绪。相反，每一次情绪体验都是由更基本的部分构成的。

她写道：

“它们是你身体物理特性的结合，是一个灵活的大脑，它将自己连接到它所处的任何环境中，以及你的文化和成长环境中。”

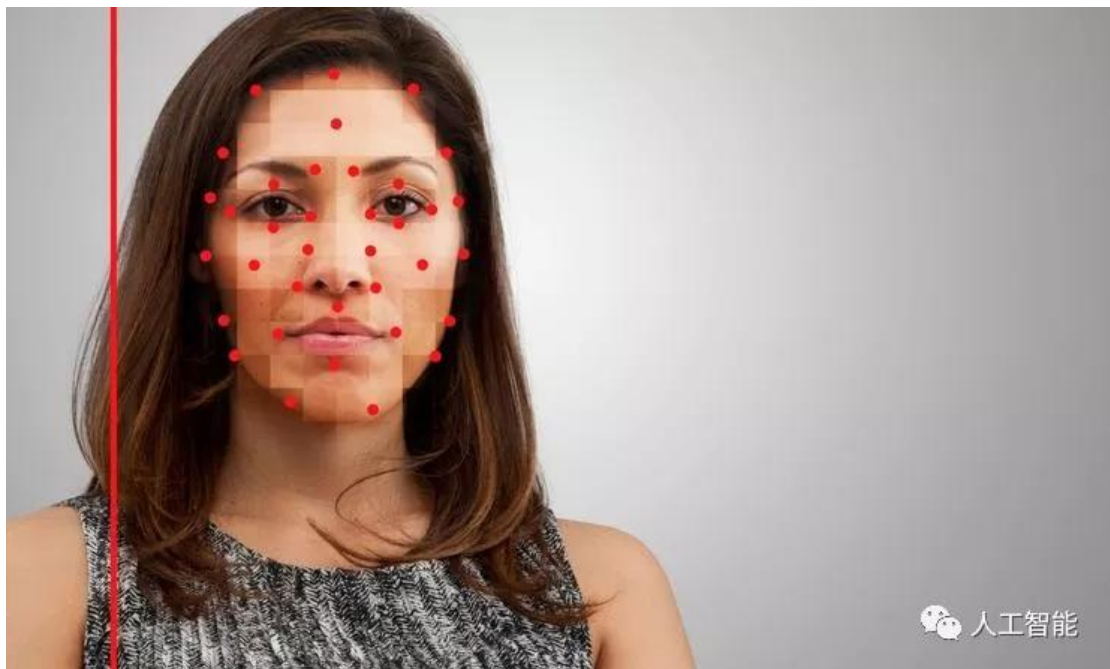
“情绪是真实的，但在客观意义上不是分子或神经元是真实的。它们的真实性与金钱的真实性是相同的，也就是说，这不是幻觉，而是人类一致同意的产物。”

Barret 解释说，把面部表情直接映射到所有文化和环境中的情绪是没有意义的。当一个人生气时可能会皱眉，而另一个人可能会在策划敌人倒台时礼貌地微笑。因此，评估情绪最好理解为一种动态实践，包括自动认知过程、人与人的互动、具体经验和文化能力。她说：“这听起来像是很多工作，但确实如此。”

“情绪是复杂的。”

Kaliouby 也同意情绪是复杂的，这就是为什么她和她所在的 Affectiva 的团队一直在努力提高他们数据的丰富性和复杂性的原因。除了使用视频而不是静止图像来训练算法外，他们还尝试着捕捉更多的上下文数据，比如声音、步态以及人类感知之外的面部细微变化。她相信更好的数据将意味着更准确的结果。一些研究甚至声称机器在情绪检测方面已经优于人类。

但据 Barret 说，这不仅与数据有关，而且与数据如何被标记有关。情绪检测公司和其他情绪检测公司用来训练算法的标记过程，只能识别 Barret 所说的“情绪刻板印象”，就像表情符号一样，这些符号符合我们文化中熟悉的情绪主题。



纽约大学人工智能研究所的联合主任 Whittaker 认为，基于 Ekman 过时的科学构建机器学习应用程序不仅仅是一种糟糕的实践，它还转化为真正的社会危害。

“你已经看到招聘公司使用这些技术来衡量应聘者是否是一个好员工。”

她说：“在学校环境中，你也可以看到一些实验性的技术被提出，以观察学生在课堂上是积极的、无聊的还是愤怒的。”“这些信息可以用来阻止人们获得工作或改变他们在学校的待遇和评估方式，如果分析不十分准确，那将是一种具体的物质伤害。”

Kaliouby 说，她意识到情绪检测可能被滥用的方式，并将认真对待她的工作道德。

“与公众就这一切的运作方式——在哪里应用以及在哪里不应用进行对话至关重要。”

Kaliouby 过去戴过头巾，她也敏锐地意识到建立不同数据集的重要性。她说：“我们确保当我们训练这些算法时，训练数据是多样的。”“我们需要代表白种人、亚洲人、肤色较深的人，甚至是戴头巾的人。”

这就是为什么 Affectiva 从 87 个国家收集数据的原因。通过这个过程，他们注意到在不同的国家，情绪表达似乎呈现出不同的强度和细微差别。例如，巴西人用宽而长的微笑来表达幸福，而在日本，微笑并不表示幸福，而是表示礼貌。

Affectiva 通过在系统中添加另一层分析来解释这种文化差异。汇编了 Kaliouby 所说的“基于种族的基准”，或编纂了关于不同种族文化中如何表达情绪的假设。

但正是这种基于种族等标记的算法判断，让 Whittaker 最担心情绪检测技术暗示了自动化“相面”的未来。事实上，已经有公司对某人成为恐怖分子或恋童癖者的可能性进行了预测，同时也有研究人员声称，他们拥有的算法能够单独从脸上检测到性行为。

最近几项研究还表明，面部识别技术会产生更可能伤害少数族裔群体的偏见。去年12月发表的一篇文章显示，与白人相比，情绪检测技术给黑人脸上的负面情绪更多。

Kaliouby 表示，情绪系统确实有一个“种族分类器”，但是他们现在没有使用它。相反，他们使用地理位置作为确定某人来自哪里的基准。这意味着他们将巴西的微笑与巴西的微笑进行比较，而日本的微笑与日本的微笑进行比较。

“如果巴西有一个日本人呢？难道系统能识别出礼貌微笑和幸福微笑的细微差别吗？”

Kaliouby 承认，在这个阶段，这项技术并不是百分之百的万无一失