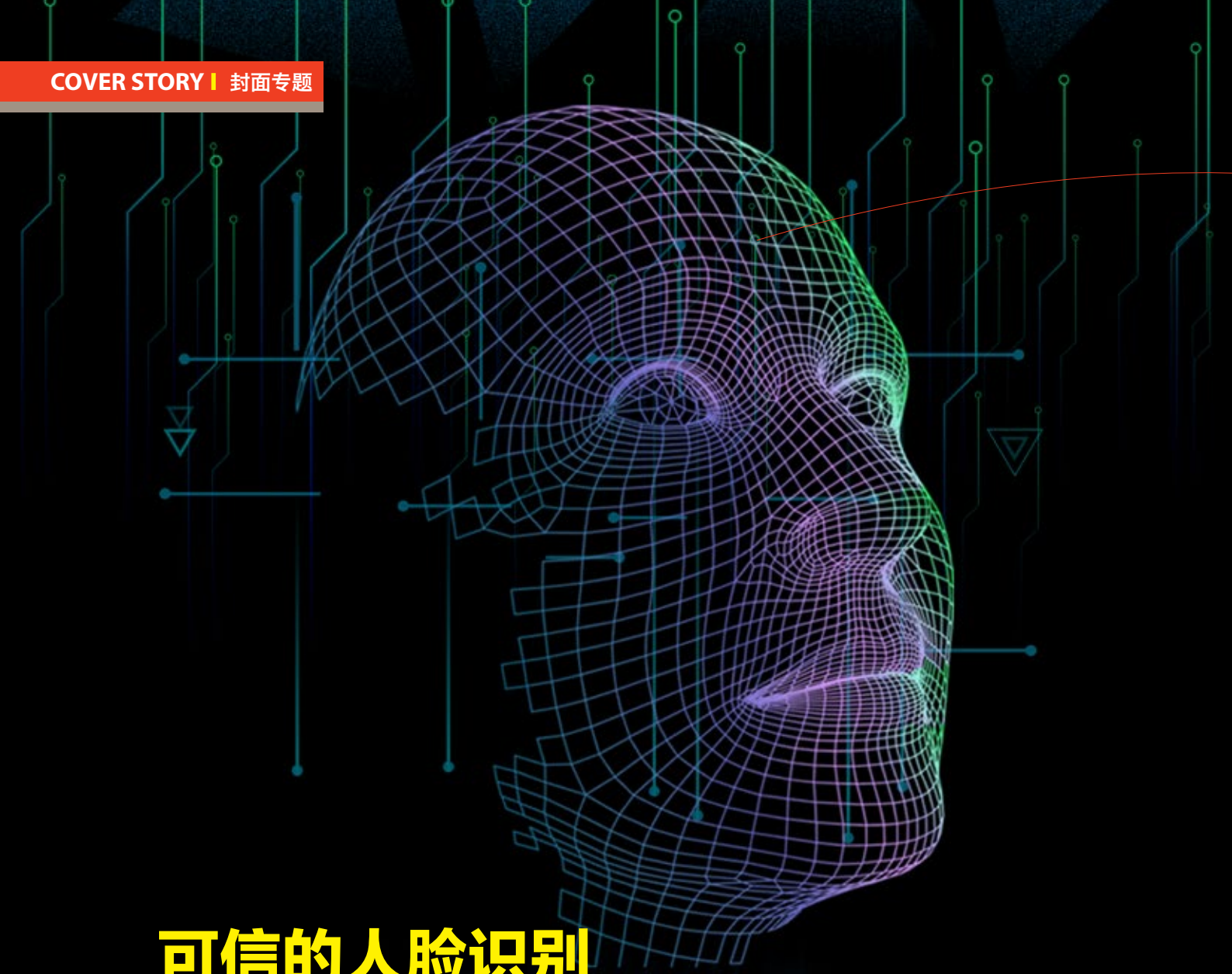

可信的人脸识别

文 / 汤寅航 张广鹏

近年来，“人工智能”被越来越多人熟知，“人脸识别技术”作为人工智能的重要分支领域也同样进行着飞跃式的发展。经过 30 年发展，特别是近年来深度学习技术的应用，人脸识别的精度获得了质的提高，人脸识别的应用也遍地开花。如何评判一个人脸识别系统是否可以信赖呢？我们认为判断人脸识别技术是否“可信”的三个重要指标包括“高精度”、“抗干扰”和“防伪”。“高精度”是指计算机在进行身份验证和识别时通过率高而误识率低。“抗干扰”指识别精度不易受光线条件、遮挡条件、姿态和表情变化的等因素的影响。在此之上，为了防止人脸识别系统被照片、视频、面具等人脸假体所欺骗，“人脸防伪”技术也是增强人脸识别技术安全性的重要部分。

人脸识别技术的核心思想是教会计算机分析理解人脸样本的方法，让计算机能够建立采集到的人脸样本和人脸身份的对应关系。该思想中的两个重要概念是“人脸样本采集”和“人脸样本特征提取”。前者决定了计算机看到了什么，后者决定了计算机如何理解看到的事物。

多模态人脸识别

在人脸样本采集方面，可见光人脸照片是最早被使用的人脸识别样本。可见光人脸照片易于采集，不依赖于特殊设备，符合人眼对物体的认知习惯。而且可见光人脸照片含有丰富的纹理和颜色信息，为计算机准确理解人脸图像奠定了基础。但是在不可控的光照环境中采集到的人脸图像会有不规则的光影效果，比如逆光、过曝、阴阳脸等，会使人脸识别准确率有所下降。

为了在图像采集阶段有效降低光照变化的影响，中科奥森首席科学家李子青教授提出了使用近红外波段的人脸图像进行人脸识别。通过设置近红外补光灯，并在采集设备上增加相应波段的带通滤光片，在最大程度上降低环境光照对人脸识别造成的影响。不仅如此，李子青教授还提出了将可见光人脸图像和近红外人脸图像进行异质图像融合，充分利用两者呈现的互补性纹理信息，来进一步提高人脸识别准确率。

除了二维图像中所能提供的人脸纹理信息，三维人脸模型提供的曲面几何信息同样有助于提升人脸识别的准确率。人脸的三维信息不受外在环境条件和采集设备的影响，对头部不可控的空间旋转也有极强的鲁棒性。 而随着采集设备的硬件技术和相应软件技术的发展，三维人脸样本的采集和三维人脸的模型重建技术也日益成熟，加快了三维人脸识别的发展进程。三维人脸模型的采集方式主要是 ToF（Time of Flight）和结构光（编码结构光、散斑结构光等）两种技术，二者均能直接获得人脸上稀疏点的三维坐标进而重建出三维人脸模型。三维人脸识别算法克服了二维人脸识别技术所遇到的人脸旋转、面部化妆、光照变化等问题，但容易受表情变化的影响。

基于构造特征的人脸识别

在完成人脸样本的采集后，如何设计一种计算机能够理解的人脸描述方法，则是人脸识别技术的第二个重要问题。为了能够提取人脸的本质表达，以抑制外界因素的影响，提高身份区分能力，研究人员构造了各种各样的特征。

最早的人脸识别算法可以追溯到 Eigenface 算法和 Fisherface 算法。两种方法的思路都是将每个人脸图像作为整体进行分析，并为所有待分类的样本找到一个具有可分性的子空间，使得投影后的样本能够各自聚到同一个身份标签下。EigenFace 方法是对训练集进行主成分分析（Principal Component Analysis，PCA），以此获得该训练集所对应一组“特征脸”（Eigenface）。该组特征脸构成了子空间的基向量，而其余所有人脸都可以表示为这些特征脸的线性组合。而为了进一步保证分类算法的准确性、稳定性和泛用性，Fisherface 算法提出一个优秀的子空间应该能使同一类的样本在子空间的投影点都应聚得紧密（类内距离最小化），而不同类样本的平均投影中心则应相互远离（类间间距最大化）。因此 Fisherface 依据这两个判别标准，在 Eigenface 的基础上加入了线性判别分析

（Linear Discriminant Analysis，LDA）来提高人脸识别技术的性能。由于基于原始的人脸图像，Eigenface 算法和 Fisherface 算法对于光照变化、旋转变化、尺度变化以及遮挡问题的鲁棒性较差，识别精度较低，但是这两种算法所采用的降维和提高特征区分度的思想被广泛应用到后来的人脸识别系统中。

为了克服基于原始图像的人脸特征容易受外界因素影响的问题，人们着力研究利用人工构造特征描述人脸图像。最广泛应用的局部特征有 Gabor 特征和 LBP 特征 (Local Binary Pattern)。Gabor 特征所使用的 Gabor 小波与人类视觉系统中简单细胞的视觉刺激响应非常相似，它的主要设计思路是指定一组多方向多频率的 Gabor 滤波器。它们对和滤波器有相近方向和频率的纹理信息会产生更强的相应。因而 Gabor 特征对于图像的边缘敏感能够提供良好的方向选择和尺度选择特性，而且对光照变化有良好的适应性。LBP 特征的设计思路是在 3×3 像素邻域内，将相邻的 8 个像素值与中心像素值进行比较。若周围像素值大于中心像素值，则该位置被标记为 1，反之为 0。这样周围 8 个点就会产生一个 8 位的二进制数，而其转换的十进制数就是中心像素点的 LBP 值，用来反映该区域的纹理特征。LBP 特征提取效率高，具有一定的抗旋转能力，并能够较好地解决均匀光照变化的问题。

在长达 20 余年的人脸识别研究中，人们花费了大量的精力构造了各种类型的特征，在一定程度上提高了人脸识别的精度，但在实际应用中仍难满足人们对精度的要求，使得人脸识别系统被局限在受控环境下，难以推广。

基于深度学习的人脸识别

深度学习技术是以人工神经网络为基础，通过模仿人类对于物体的多层感知模式，对数据逐层解析和抽象，从而形成精简高效的表征。近年来随着数据量、算法和硬件等条件的成熟，深度学习技术在图像识别、语音识别、语义理解等领域获得了极大的成功，引领了新一波人工智能应用的浪潮。

在人脸识别方面，2014 年 Facebook 在 CVPR 会议上发表文章，首次用实验证明了深度学习能够大幅提高人脸识别的准确率。不同于人工设计的局部特征，深度学习方法所得到的特征是计算机通过大规模神经元网络从大量训练数据中自主学习到的。这种特征的准确性和普适性主要受到两个因素的影响。第一个是人工神经元网络的规模，第二个是用于训练的样本数据量和多样性。通常而言，人工神经元网络的性能和它自身的深度和广度成正比。深度是从网络输入到输出所要经过的最长隐藏层路径，广度则是隐藏层的输出特征维度。当两者增加时，用以描述图像的参数和维度就会相应成指数级增加，对图像的描述就会更完备。使用这样的网络所需要的计算资源也会相应增加，因此近年来计算机硬件的发展成为了人脸识别技术的

强大助力。传统的人工构造特征会受制于自身设计的局限性难以充分利用训练样本的多样性，深度学习方法却能够广泛地利用和吸收训练数据所提供的多种信息并形成简洁高效的特征。提供给计算机用作训练的样本数量越大，样本的多样性越强，深度学习得到的用以描述人脸图像的特征就更准确更具有普适性。深度神经网络规模的增加，也提高了对训练数据的需求。为了保证在参数训练过程不会产生过拟合的错误，充足的数据量是深度学习的前提条件。以李子青教授的研究团队在 2014 年构建并公开的“CASIA-WebFace 数据库”为代表的一批人脸数据库的建立，保证了深度学习技术在人脸识别领域的应用。

深度学习技术的引入推动人脸识别的精度不断提高，误识率由传统特征的千分之一、万分之一降低到目前的百万分之一甚至亿分之一，对光照、表情、姿态等各种因素的影响也越来越鲁棒，比较好地满足了“高精度”和“抗干扰”这两个可信的人脸识别需要具备的指标。

人脸识别防伪技术

为了加强人脸识别技术对抗照片、视频、面具等人造假体的攻击行为，人脸防伪技术是增加人脸识别技术安全性的重要手段。早期的人脸防伪技术是以“动作配合式”为主，根据检测到的面部动作是否与指示的眨眼、张嘴、摇头等动作一致来验证人脸的真实性。这种防伪方式需要主动配合，在许多非受控场景无法使用，另外也容易收到人脸图像合成、人脸动画等方式的攻击。而为了推广人脸识别技术的应用，提高防伪强度，李子青教授提出了多光谱人脸防伪的专利技术。该技术采集人脸多光谱图像，使用纹理分析技术分析真实人脸和假体人脸之间的差别，无需用户主动配合，能够抵抗数十种攻击手段，满足了可信人脸识别“防伪”的要求，极大地扩展了人脸识别在实际场景中的应用。

人脸识别的应用

正因为近年来人脸识别技术识别精度和抗外界干扰能力不断提升，防伪能力的不断改善，这种“可信”的人工智能技术已经在安防领域、金融领域、商业领域、日常生活中全面落地开花，催生了大批以人脸识别技术为核心的实际应用。在安防领域，动态静态相结合的人证核验通关系统和公安路面监控黑名单系统保障了人们的生活安全。在金融领域，银行柜面机实名系统、ATM 自助业务、远程实名服务和移动支付业务都在保证客户的财产安全的前提下使各种金融业务更简便。在商业领域，人脸识别考勤系统、门禁系统、访客接待系统、广告定向推送系统等都已经在很多公司和商场得到了使用。在日常生活中，人脸识别小区门禁系统、刷脸自助取件服务、刷脸无人超市等融入了人脸识别技术的服务都在陆续涌现。总而言之，在计算机软硬件不断革新的今天，可信的人脸识别技术一定会让我们的生活更加智能。

作者单位：北京中科奥森数据科技有限公司