

模式识别技术的现状与发展

朱砾博 Arto Kaarna
(拉彭兰塔理工大学 芬兰拉彭兰塔市 53850)

摘要: 本文将从理论和应用两方面对模式识别研究领域最近的问题和未来的发展进行介绍, 并思考其未来的发展趋势。

关键词: 模式识别; 人工智能; 模式识别应用

1 引言

模式识别是人工智能领域的基础技术, 以寻找数据模式作为基本问题, 以机器识别和感知为基本结果。模式识别和机器学习作为基本方法在感知数据之外的数据中也有使用, 例如在大量数据中搜索信息的数据挖掘学科。近年来随着计算机技术的快速发展和算法技术的研究, 模式识别技术进入了一个新的高速发展期。无论是模式识别本身的理论和方法, 还是应用场景的拓宽, 都取得了长足的发展。

2 模式识别基础理论和方法的研究现状

模式识别的基础理论主要包括特征学习、分类器、聚类等, 为了达到更好分类的目的, 研究目标整体为高效性, 鲁棒性的改善。

2.1 分类器

分类器作用是在标记好类别的训练数据基础上判断一个新的观察样本所属的类别^[1]。分类器最热门的就是现在广泛使用的深度学习算法, 在整个人工智能领域都取得了应用成果, 远超先前相关技术。其作为一种模式识别分析手段, 包括卷积神经网络, 递归神经网络和无监督预训练三种方法。其中经典模型“卷积神经网络模型”灵感是来自视觉系统的结构。关于深度学习的研究虽然基本理论较为固定, 但是一直有着算法性能上的特化和改进, 偶尔有优化深度学习架构的尝试。在深度学习架构的优化中, 神经网络作为一种更加强大的模型, 能对更加丰富的函数簇建模^[2]。此外, 在深度学习大框架模型中, 可算性理论研究也很多。这些研究方法有助于提前估算所选择的具体方法和数据范围, 提高了效率。在研究具体问题时, 多种分类器合作结果也是常用的, 这种运用方式称之为集成学习, 其往往也能得到更好的训练结果^[3]。

2.2 聚类

聚类基本思想是定一个由样本点组成的数据集, 通过聚类分成的簇是一组数据集合, 这些对象与同一个簇中的对象彼此相似, 与其他簇中的对象相异, 其作为简化计算和加快收敛方式在研究中很常用。广泛使用的 K-均值(K-mean)无监督聚类法实现简单, 聚类效果也很好, 基本思想就是对于给定的样本集, 按照样本之间的距离大小, 将样本集划分为 K 个簇。让簇内的点尽量紧密的连在一起, 而让簇间的距离尽量的大。另一种聚类方式聚类树, 是将自上而下或自下而上的每个数据点视为单一聚类, 然后依次合并直到包含一个数据点的单一聚类。除此之外还有多种常用聚类模式如均值偏移聚类和高斯混合模型的期望最大化聚类。

聚类方式的发展研究主要是针对特定训练要求的优化和改动从而改善性能。这些热度不如深度学习, 但是经典的方法也一直有少量研究和改进。最近的研究表明 K-均值的算法比较稳定, 偶尔会有性能下降的情况, 改进的办法是重复优化初始化的操作来提高性能并带来更好的聚类精度^[4]。多视图聚类的思路最近也被提出, 通过多视图聚类可以得到比单视图聚类更准确的结果, 此方法从 2009 年提出到现在取得了长足发展。层次聚类也一直有最新的特化应用, 如在强调基于环境而行动的强化学习中, 利用蒙特卡罗搜索树中将树的返回值带入根, 可以达到降噪和加速收敛的目的^[5]。

聚类方法也有一定的局限性, 很难在众多聚类算法中找到较为优势或者说参数合适的聚类方法。最近新起的集成聚类可以将结果鲁棒性提高, 也是一个研究方向。现在更有数学上结合无限集成聚类的研究, 通过将自动编码器和降噪结合得到无限聚类下的期望^[5]。

3 模式识别应用的研究现状

模式识别技术的应用领域主要包括计算机视觉, 文字识别、语音识别、生物特征识别、医学分析、遥感数据分析等。其中计算机视觉是模式识别中较为重要也是最热门的研究方向, 其涉及的内容广泛, 问题也很多。模式识别技术广泛的应用价值受到了人们的极大重视, 使用范围也在不断扩大。

3.1 计算机视觉应用研究

计算机视觉是视觉是图像和视频内容在模式识别领域的应用, 是模式识别技术最早的研究方向和具体应用之一, 并且至今仍然是模式识别应用研究的主要方向。

物品检测其中的研究内容更是细化, 从分类, 定位, 检测, 分割各方面都有成果和应用。作为较为成熟发展的应用方向也已经大量商业化运用, 比如网络中的以图搜图或者以图识物。现在物品检测的研究更多是高压情况下的识别结果改善, 比如伪装物体检测。通常伪装物体与其背景之间具有高度相似性, 因此伪装物体检测更具挑战。在最近研究中, 模式识别技术在物品检测上甚至已经有超越人类的势头, 一些成熟算法甚至在人眼难以分别的图像中保持较高的鲁棒性^[7]。

计算机视觉也为自动驾驶研发中也有大量应用。人群检测, 轨迹规划和立体匹配等方面都有计算机视觉的参与。

二维视觉模式识别的长足发展给三维视觉模式识别打下了坚实的基础。三维视觉模式识别在特征匹配, 相机标定, 几何理论的加持下, 已经在三维重建领域前进。运动结构恢复作为流行策略, 提供了几何结构和相机运动的同时估计, 在无序图像中方法能够推进

三维重建。

计算机视觉领域另一个话题就是视频内容的模式识别，其基础与静态图像相似，但是视频内容作为动态变化，而且有更可能性要克服模糊，轨迹，形变等问题，从而带来更多计算压力。在运用深度卷积网络时也利用类似于视频压缩带来非必要帧的放弃和运动补偿是加速算法的一种思路^[8]。同时大规模视频数据库的建立也给视频内容的模式识别提供数据，但是得注意的是监控内容与此类数据往往质量相差甚远，因为电影电视剧的视频内容一般是不存在画质干扰的，而是有很多非写实画面。

3.2 其它应用研究

文字识别，语音识别和生物识别是方面发展最快也是较为成熟的三种应用方面。文字识别与基本的物品识别的基础上，除了将文字识别出来，还有字符切割，文本行识别，字符识别等核心问题。同时特殊文本识别如手写识别，古籍还原等也都要有特向化的技术改善。

语音识别总体基于隐马尔可夫模型和统计语言模型的语音识别技术^[9]。除了日常、商务的使用，在制造业的控检，直播和语音消息的内容过滤等方面都有参与。消除环境噪音是语音识别的重要课题之一，特征补偿一定程度上对平稳噪声有了适应能力，但是对应有变化的噪音则效果不佳。13年在深度神经网络的映射降噪算法下，语音技术有了较大突破。17年后各大商业公司的努力下更是构成系统，运用并行处理的算力解决了例如方言识别等普及方面的问题^[10]。生物特征识别技术从一开始的指纹检测，到脸部、虹膜、DNA检测，再到更多的新应用场景如眼动，步态，其发展一直需最关注的是检测效率 and 安全性。

4 发展趋势

模式识别是一门理论和实践紧密结合的学科，未来几年发展主要在于算法的改善、计算机算能的提升和应用范围的扩大。在理论基础方面，在分类器和聚类都面临着大数据的考验。在以往研究方向都在提高准确度追求高精度的现状下，如何将海量，多源，多类型，可靠性不一的数据高效统一分析成为关键。与此同时，如何利用好大数据带来的训练集数量的提升也是关乎效率的课题。应用方面，虽然模式识别技术已经有了不少成果，但是仍然与基础自然模式识别（人类、动物）能力差距甚大，比如人类通过味觉可以轻易分辨出苹果和苹果味道的糖。这其中不仅有味觉这种模式识别没有触及的领域，还有多类模式识别结果综合的过程。

模式识别技术大有可为但是暂时发展不够的方向就会是近未来发展的热门。一是自然语言的理解，当训练模型中句子，段落甚至文章内容的理解效果出色，意味着算法改进的达标。二是感情及其波动的识别，不仅包含多种生物特征识别结果如语音，视觉的综合技术，更意味着神经活动模式的分析与预测变为可能。三是医学分析，如果在模型，医生和病人之间达到信任，代表着整体模式识别预测结果鲁棒性的提升。稍加思考可知，模式识别技术想要长足发展，不仅在理论上需要不断完善和改进，仍要与工业，光电领域，生物学科和其他细节学科交叉合作综合应用，只有这样才能在更多

方向的模式识别中有所贡献。

当然在更未来高度信息化智能化的社会里，模式识别技术是无处不在的。理论会不断深入，使用场景会更复杂，当然也会给模式识别技术的发展带来更多的挑战和机遇。

参考文献

- [1] Duda, R. O., & Hart, P. E. (2006). Pattern classification. John Wiley & Sons.
- [2] Chen, R. T., Rubanova, Y., Bettencourt, J., & Duvenaud, D. K. (2018). Neural ordinary differential equations. In Advances in neural information processing systems (pp. 6571-6583).
- [3] 周飞燕, 金林鹏, & 董军. (2017). 卷积神经网络研究综述 [J]. 计算机学报, 40(6), 1229-1251.
- [4] Fränti, P., & Sieranoja, S. (2019). How much can k-means be improved by using better initialization and repeats?. Pattern Recognition, 93, 95-112.
- [5] Efroni, Y., Dalal, G., Scherrer, B., & Mannor, S. (2019). How to Combine Tree-Search Methods in Reinforcement Learning. AAAI.
- [6] Liu, H., Shao, M., Li, S., & Fu, Y. (2018). Infinite ensemble clustering. Data Mining and Knowledge Discovery, 32(2), 385-416.
- [7] Fan, D. P., Ji, G. P., Sun, G., Cheng, M. M., Shen, J., & Shao, L. (2020). Camouflaged object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 2777-2787).
- [8] Buckler, M., Bedoukian, P., Jayasuriya, S., & Sampson, A. (2018, June). EVA²: Exploiting Temporal Redundancy in Live Computer Vision. In 2018 ACM/IEEE 45th Annual International Symposium on Computer Architecture (ISCA) (pp. 533-546). IEEE.
- [9] Dahl, G. E., Yu, D., Deng, L., & Acero, A. (2011). Context-dependent pre-trained deep neural networks for large-vocabulary speech recognition. IEEE Transactions on audio, speech, and language processing, 20(1), 30-42.
- [10] 侯一民, 周慧琼, & 王政一. (2017). 深度学习在语音识别中的研究进展综述 [J]. 计算机应用研究, 34(8), 2241-2246.

作者简介

朱砺博 (1994-), 男, 安徽省舒城县人。硕士研究生学历。研究方向为模式识别。

Arto Kaarna (1956-), 男, 芬兰国籍, 博士学位, 副教授。研究方向为计算科学与工程。