

上都说明对训练集的样品应严格检查, 并应采取分层分类, 以减小误识率。

2. 实验结果证明, 用 Fisher 法把多维分布投影成一维分布后, 更加接近正态, 这时可按给定的假阴性率选择判决阈值, 以降低对异常细胞的误识率。但投影后的分布与正态分布的适

合程度如何, 由于现有数据太少, 尚难肯定。

参 考 文 献

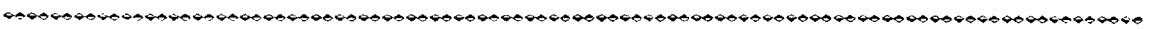
[1] Duda, R. O. and Hart, P. E.: "Pattern Classification and Scene Analysis", Chap. 4, 1973.
[2] 中国科学院计算中心概率统计组: 《概率统计计算》, 科学出版社, 1979 年, 第一章, 第 4 节。

结 语

以上介绍全部工作中, 除第一个专题是为整个研究工作提供可以满足机器识别特殊要求的细胞涂片, 其余四个专题提出了数字图象的预处理、特征抽提、特征选择、分类判别函数, 再加上各个步骤间相互连接的软件, 所有这些内容就构成了一个完整的算法系统。掌握了这一算法系统, 标志着对模式识别概率统计方法全过程认识有了良好开端, 并为开展下步工作打下了可靠的基础。有了这个基础, 只要具备必要的硬设备, 食管癌细胞的自动分类工作就立刻可以应用于临床; 同时对其他种类细胞进行自动分类的应用性研究工作也就较易开展, 因为完整的算法系统已为按照给定的指标 (即模式识别术语的“特征”) 进行细胞自动分析和类别判决提供了现成的贮存数字化图象的数据库和分析图象的程序包。从提高的角度来看, 有了这一基础就可以进一步开展有关扩大判别、分类特征种类, 以及分层判别、增加分类类别以提高识别率等研究工作, 使这种新的分析方法可以解决生物学现有的常规方法所无法解决的生物医学图象分类问题。考虑到现已掌握的只是算法系统的最基本内容, 这种提高工作就更加必要, 而这正是我们今后工作的重点。

为了取得更多的成果, 在研究中学科间相互渗透、紧密配合是十分重要的。回顾开展工作的初期, 往往甲提出的问题乙不理解, 乙阐述的概念甲又感到陌生。但是在共同的目标下, 经过相互学习, 共同实践, 彼此之间的共同语言愈来愈多, 工作进展也随之加快。如果今后从事这方面研究的工作者在开展工作时更自觉地注意取得共同语言, 同时争取在国际上有一定水平的我国数学界的有力配合, 以发展新概念, 提出新方法, 并密切结合我国所特有的研究对象, 虽然我国起步较晚, 硬设备也有一定限制, 但可以预期将作出具有我国特色的工作来。

[本栏各文于 1979 年 12 月 12 日收到]



学 科 简 介

细 胞 图 象 分 析

模式识别 (Pattern Recognition) 是近二十年来随着电子计算机的广泛应用而产生和发展的新兴学科。它企图用计算机连同特殊输入装置模拟或部分代替人类的视觉、听觉或其他智能, 以分析、描述与识别图象、声音或其他模式。这一学科当前在机器人视觉, 生物医学图象处理、遥感图象的解释和文字、声音识别等方面均有广泛应用并取得了卓越的成果。

应用模式识别方法于细胞学研究实际上是将现代技术中的信息理论用于分析细胞。它既可科学地总结临床细胞病理学家的诊断经验, 又可充分发挥机器视觉分辨率高 (可达 256 灰阶)。抽提特征方式灵活多样

(可按各种数学模型形成特征, 不一定要每个特征具有明确的直观意义。例如美国对一细胞可提取 200 多个特征, 加拿大可提取 300 个以上)。随着细胞图象自动分析方法的完善与工业基础的发展, 不难实现癌细胞的自动筛选与诊断, 同时也可为难症病症的鉴别诊断和对从癌前到早期癌演变过程的研究提供客观指标与有效手段。

癌细胞学自动化与细胞图象分析第三次国际会议今年五月将在西德慕尼黑召开。

(中国科学院生物物理研究所陈传渭)