

DOI:10.3969/j.issn.1000-9760.2018.05.016

医学图像数据的多元分析

任晓丽

(山西医科大学汾阳学院, 汾阳 032200)

摘要 医学图像是当代医学的重要分支,以无创伤方式将人体内的信息采集并予以形象化显示。本文立足于图像的数字化优势,在当今各类医学影像数据的迅速积累,即大数据背景下,阐释了医学图像数据的多元特征,就图像处理进行了多层面分析。

关键词 医学图像数据;信息;内容;处理

中图分类号:TP183 文献标识码:A 文章编号:1000-9760(2018)10-367-05

Multivariate analysis of medical image data

REN Xiaoli

(Fenyang College of Shanxi Medical University, Fenyang 032200, China)

Abstract: Medical image is an important branch of modern medicine, which collects and visualizes the information in human body without trauma. Based on digital image edge in today's rapid accumulation of all kinds of medical image data namely the background of big data, we illustrates the multiple features of medical image data, and the analysis was carried out in the image processing level.

Keywords: Medical image data; Information; Content; Processing

图像是目标的复杂映射。数字化条件下,图像是由离散图像元素组成的多维数据。数据是介于目标和图像间的信息载体,医学数据加上意义就是医学信息的描述。基于物理因子与人体组织作用机制,依据具体用途经适合的数据处理系统而生成的医学图像,突破了原先听诊、触摸等临床诊断方式的局限,以其形象化、无损化方式被广泛使用,不断发展为医学界的一枝独秀。其主要过程包括初始数据的采集获取,针对对象特征信息的处理及最后目标图像的再现。

1 医学图像数据的产生

1.1 成像概述

通过视觉方式获取机体信息是临床诊断的重要途径。自 19 世纪末伦琴发明 X 射线摄影方法,100 多年来分别产生了超声(US)、计算机断层扫描(CT)、核素显像(RI)及共振(MRI)等电离辐射和非辐射型成像方式。共同之处都是由源到像的映射,将体内不可见的信息进行“可视化”的表达过程。医学数据的可视化和分析是医学研究的一

个实际方向,初始数据的获得及处理是良好诊断结果的基本前提^[1]。人体内不同组织脏器的位置形状等性质迥异,同一病灶在不同病理期的参数亦有变化,经一定的能量波(电磁波、机械波)作用后,客体的特征逐渐被主体化,被赋予了处理描述的含义。

1.2 作用机理

医学图像生成的前提是体内的目标物与输入图像的图像结构(区域、线等)之间的“正确”映射。不同模式的图像因其成像机制不同各有特点和优势。X 线图像占临床总影像的 70% ~ 80%,主要基于 X 光子和体内物质粒子的光电、康普顿作用。图像是体内某一特性对 X 线辐射强度的吸收衰减描述,是射线透过被检组织、器官的投影。普通的 X 线成像将立体的目标物沿 X 线方向压成了二维图像,产生骨骼组织重叠,深度信息被掩埋。用来获取观测目标的断层图像技术,即 CT 解决了目标源到图像的重叠映射的问题。CT 技术起源于医学成像,也逐渐应用到其他图像领域。X-CT 基于计算机,采用精密的扫描、探测系统,获得人体的高密

度分辨力和无重叠体层图像。同时 CT 的出现也促进了不同医学成像方式的有机融合。广义的医学 CT 是通过某种辐射源(X 射线、放射性核素、超声波、磁场等)作用于信息源(目标),检测投影数据,并进行图像重建、数据处理^[2]。如发射型计算机断层成像(ECT),基于放射性药物经人体代谢,在正常组织与病变部位间形成浓度差异,探测该差异进行处理成像,是利用诊断物体内部发射的能量波成像,着重提供脏器的功能信息。磁共振断层成像(MRI),基于组织的氢质子磁化矢量差异,由射频波作用于三维人体生成的截面投影数据重建而形成二维图像,具有多方位、多参数、无辐射、无骨伪影等优势。总之由不同的成像设备进行功能的融合可获取特定的断层数字图像序列。

2 图像数据及信息层面特征

2.1 信息数据化

信息是对物质世界的描述,是将数据经过加工处理后所获取的有用知识。数据是对客观事物记录下来的可以鉴别的符号,包括数据内容和形式。两者的联系如下:1)数据只有经过加工(解释),才有意义,才能形成信息;信息是经过加工并对客观世界产生影响的数据。2)数据是信息的具体表现形式,信息是数据的有意义的表达。故信息是数据的本质与含义,由数据加上意义构成。如通过 $CT = k \frac{u - u_w}{u}$,将物质对一定强度的 X 线的衰减量化为 CT 值,100HU 的 CT 值仅代表数据,超过 100HU,对于颅内可以确定为钙化,此时才会形成信息。另外数据也具备物理属性,数据主要是对图像的像素和矩阵的描述^[2]。除此,数据也包括成像过程中难以克服的各种扰动信息,整个数据空间就是所有可能数据的集合。

医学成像的根本目的是将人体内各种生理、病理、生化活动等信息转化为图像对象的特征:如颜色、纹理、形状、边界等传统的一般特征。除此之外,还有三维物体的深度、表面法线方向、动态景物的运动参数等信息。医学图像由计算机中的二维点阵构成,即空间坐标离散。其视野面积取决于点阵的大小和数目。各点的特征值用整数值表示,即性质离散。众多点阵组成的数据集合并不能明确表示对象的特征,远低于于图像所表示的逻辑语义层次上的意义。依据人们对图像认知处理的阶梯性

可将信息划分为如下层次:1)基于初始图像信息的像素层;2)基于特定领域的识别对象及一致性区域的对象层;3)高层次的语义概念层:运用应用领域知识经验可生成基于对象级别所识别出的对象或区域信息。

2.2 数据的高维性

由于受成像源或模态、参数、成像时间等因素的影响,医学图像数据复杂、多样及上下文具有丰富的关联^[3]。图像数据具有多个变量,反映同一对象的不同方面的细节特征,即高维性。对客体信息的详尽描述离不开对象的各个方面的特征及特征间的关系,需要有不同的角度的数据描述支撑。因而,产生了多个变量抽象描述复杂对象,组成了矢量数据,形成了抽象的高维数据空间,这种高维性是对图像对象的多重属性及相互联系描述的基础,所有的属性维度可能上升到指数级的组合情况。

高维数据空间对应多维特征空间,作用是表达空间实体的多个属性或特征,一个空间实体看作特征空间的一个点,成千上万的数据记录是特征空间成千上万的点。高维的矢量数据隐显不同层次深度的信息,如基于像素级的微观粒度,对应分析层次的最小粒度;基于像素集团级的中观粒度;基于对象级的宏观粒度,对应分析层次的大粒度,表现为空间对象间的空间关系。不同层次的数据具有不同的特征,如浅层特征表示边缘等信息,深层特征具有更高层的语义等。

2.3 数据的挖掘性

由图像与数据源的关系可将信息分为内容信息(内部)和源信息(外部)。后者指客观直接的无需推理分析的所属性信息(如患者年龄、性别等)及所用成像仪器种类情况等。内部信息对应图像的实质性内容。具有可抽取、可挖掘性。即从一幅(或多幅)图像数据中提取出有用、规律性的信息、图像数据关系、空间模式等。

图像数据是对全局图像特征(灰度、颜色等)和图像对象基于内容及彼此间关系的描述(如逻辑和语义等),即数据的逻辑性,隐显着各种预期的目标医学信息。除了反映客体最基本、最直观的信息外(如器官层次:心脏、脑、肝脏等;系统层次:心血管、神经系统等;及细胞层次^[4]),还可提取出不同程度的图像特征数据(像元层次、特征层次、对象层次)。如从海量的数据中可挖掘出不同种类疾病间的关系以及疾病与职业、血型等因素的关

系。挖掘过程旨在抽取不同层次的概念,可理解为从图像数据中形成概念,将数据赋予语义化信息联系,可提取语义及关联等特征,挖掘需要综合利用领域相关的属性数据,而不是仅仅针对图像数据本身,对图像背后隐藏的需求进行挖掘,并通过与周边环境信息进行关联计算^[5]。如基于纹理特征,从海量的图像数据集中发现重要的关联规则,进而发现乳房 X 光图像中的隐藏信息^[6]。甚至深入挖掘特征空间的底层结构,探究特征之间的关系,进行特征选择,这同时是一种重要的降维技术^[7]。

2.4 数据的融合性

每个模态成像具有其特定的特征或功能,以获得形态学信息结构和生物功能,两者在器官、组织或分子水平上关于图像的内容和解释是不同的^[3]。单一化的特征信息在图像上的显现、描述,不足以成为准确判定某一病症的依据。往往需要将不同来源、不同模式、不同表示的医学图像数据有机结合,即图像融合,综合不同成像机理的图像特征,保留并加强各自的靶区特征信息,最终显示在一张图像上^[8],以得到目标对象的更精确描述,即医学数据融合。这是类似于脑组织收集并综合处理各类信息的过程^[9]。低层次的融合主要针对初始的多源数据加权相加。较高层次的特征融合事先对各源数据提取特征信息,形成特征向量再融合分析,如基于主元分析的特征级数据融合可以解决特征向量维数高的问题。融合在临床上应用的关键是多模态图像的有效相似度度量^[10]。数据融合有利于目标对象特征的全面分析和抽取。如在 MRI 和 PET 图像的融合结果中可以反映出 MRI 脑软组织结构信息与 PET 的脑内部代谢情况^[8];在肿瘤定位中 PET/CT 既能反映组织细胞的代谢信息,又能准确定位病灶^[11];PET 和 SPET 可以提供大量功能性、生化类和生理学信息。

3 图像数据的分析处理

医学图像的生成是针对体内特征进行二维投影及“信息处理”的过程。

图像处理通常概括为从不完整的数据中恢复真实的目标信息的问题,处理目的是通过输入图像提取出适应的图像结构与线索等视觉元素,以确定图像特征的真实表现,是为了获得更好的解释。

3.1 映射误差

成像过程中不可能做到无误差数据的映射。

加之人体是一个高度复杂的非线性动力学系统,即发生在确定性系统中的貌似随机的不规则运动^[4],导致图像会产生一定的退化现象。受微观粒子特性的随机影响,局部体效应(同一体素含有不同目标对象)的存在,及患者的自主性运动或本能性生理活动等,都会使图像中盘现各类噪声和伪影,边缘失锐和信号强度异常衰减。加之投影算子不唯一可逆^[1]。很可能存在病态物理求解情况:成像环节中所进行的各种模型化处理与实际物理过程总会产生误差,一旦数学方程组的系数发生变化,即便是微小的,也会导致待测量的较大波动,最终会影响到图像特征的正确显现。当然求解的病态程度取决于人们对图像精细程度的要求,是相对的。因为图像数据值本身具有相对性,取决于图像与周围背景的关系。由此需要扩充各种先验知识(医学领域、成像工程),建立并改进与目标对象吻合的处理模型,从测量的数据出发,试图恢复所需要的信息,努力还原图像的“原创性”。

3.2 计算机视觉与图像分析、理解

针对初始数据的缺憾(不完整、含噪、重复等),在主要的图像处理之前进行一定预处理,如传统的图像增强、滤波,可提高对比度及信噪比等。该过程输入、输出均为图像,基本没有知识的参与指导,属于低级处理。

由于图像诊断的广泛利用及成像设备的不断改进,医学图像数据大规模增加,为提高医生的判读效率与质量,借助计算机视觉(CV),模拟人的视觉机理对图像目标进行识别、跟踪、测量,从多维数据中获取信息形成“感知”的能力。

将图像转化为更适合人眼观察或便于仪器检测的形式,属于生物视觉的一种模拟。其中根本的问题是判定一组图像数据中是否包含某个特定的物体特征或运动状态,可通过图像分割将具有特殊意义的不同区域分开,即用具有生物学意义的标签(如解剖结构和组织类型)标记图像像素或体素的过程,这是生物医学图像分析的基本问题之一^[12],也是计算机视觉识别研究的根本所在。

早期的图像分割有边界提取、基于灰度的区域增长算法等,是经典的直接的分割处理。随着计算机视觉理论方法及图像分析的不断完善,对于图像兴趣区的分割方法得到很大改进。针对不同的脏器及成像模态,形成了基于不同成像设备、参数和解剖结构等先验知识的专门分析算法。如基于不

同组织结构共振成像时有不同的成像参数,同种组织不同参数灰度表现不同,对应不同的加权图像, T_1 图像主要显示解剖结构 $[I = kB_0\rho(1 - e^{-T_R/T_1})]$, T_2 图像主要对病变组织进行定性分析 $(I = kB_0\rho e^{-T_E/T_2})$,基于上述共振成像原理的先验知识,针对脑磁共振图像,使用不同频谱,而非单一的 T_1 图像,在分开脑白质和灰质之后,可有效分割各自的损伤区域。

图像分析着重构造图像描述方法,并利用各种有关知识进行推理。可将原始的高维数据通过线性或非线性变换映射到一个低维空间,获得一个关于原数据集紧凑的低维表示,即降低高维医学数据的维数(维度约简)。由更具代表的属性集合处理高维性的问题;基于深度学习的多层面数据处理,组建多个处理层对原始数据进行逐层抽象,从数据中获得不同层面的抽象特征^[13]。如基于形态特征 ROI 提取可获得具有诊断意义的像素区域信息;基于局部、全局、边缘的纹理粗细度及方向性变化的纹理特征提取等^[14];将研究对象固有的、本质的属性量值化,有效提取特征参数;分解对象以形成各区域的特征向量,相关的特征选择是从高维数据属性中选取重要属性的子集,得到该图像的某个或某些特征的信息或描述,从而形成图像分析结果,输出数据。

把分析的结果作为输入,进一步对图像进行理解和识别。图像理解就是对图像语义的理解,即以图像为对象、知识为核心、研究图像中有什么目标及目标之间的关系。故处理的信息分为视觉数据信息和人类知识信息,属于高级计算机视觉范畴,是在人类先验知识的参与指导下,基于人类对客观世界的认知机理,借助计算机系统将某种成像模式数据,利用一定的知识信息进行加工以期获得解释。输入的是图像的初级特征,输出的是高层次信息描述。这是人的先验性领域知识与计算机图像处理有机契入的过程,如基于专家知识,对图像数据集进行标注,提取形态学特征^[15]。

综上,依据人们对图像的递进认知,完整的图像分析理解层别有:1) 获取图像数据的数据层(灰度图等),主要针对像素操作;2) 提取特征的描述层,将像素表示形式化;3) 以有一定强度的知识作为支撑的认知层以及应用层,得益于计算机技术,人们对医学图像的解释终将克服人工解释的有限、主观、差异和疲劳等缺憾^[16]。

4 结语

科学证明人们通过视觉获得的信息占总信息的 70% 以上,医学图像以其百闻不如一见的优势在现代医学中起着越来越重要的作用。在影像大数据条件下,基于医学图像数据的巨量性、高维性、挖掘及融合特征,人们将拓宽知识信息获取的途径,增加信息推理的多样性,改进数据生成的随机性和物理求解的病态性。在数据驱动的机器学习算法基础上精确检测解剖标记^[17],正确提取图像特征。不断丰富知识领域,将复杂的感知和概念处理用于图像分析的基础上,将图像像素转换成有意义的内容,对医生诊断观察行为进行建模^[18];利用计算机视觉领域的技术选择——深度卷积网络^[19],进行更多隐含层的信息分层处理,构筑稳固的计算机视觉能力;充分利用深度学习开创图像重建的革新模式^[20]。总之在不断丰富图像模态的同时也提高图像显示的精确性,最终将促进医学图像与目标物之间的精准映射。

参考文献:

- [1] Evgeny, Semenishchev, Viacheslav Voronin, et al. Reducing noise component on medical images [DB/OL]. <https://www.spiedigitallibrary.org/conference-proceedings-of-spie>, 2018-04-13.
- [2] 张勇, 宋余庆, 严振. 基于特征的医学图像数据存储模型[J]. 计算机工程, 2011, 37(12): 209-211.
- [3] Zhang JG. Big Data Issues in Medical Imaging and Informatics [DB/OL]. <http://proceedings.spiedigitallibrary.org>, 2015-07-28.
- [4] 胡广书, 汪梦蝶. 生物医学信号处理研究综述[J]. 数据采集与处理, 2015, 30(5): 915-932.
- [5] 吴辉群, 翁霞, 王磊, 等. 医学影像大数据的存储与挖掘技术研究[J]. 中国数字医学, 2016, 11(02): 2-6.
- [6] Jyoti Deshmukh, Udhav Bhosle. Image mining using association rule for medical image dataset [DB/OL]. <https://core.ac.uk/download/pdf/81951354.pdf>, 2018-06-29.
- [7] Liu L, Wang Q, Adeli E, et al. Feature selection based on iterative canonical correlation analysis for automatic diagnosis of Parkinson's disease [DB/OL]. https://link.springer.com/chapter/10.1007/978-3-319-46723-8_1, 2016-10-2.
- [8] 杨杰, 乔宇. 医学影像分析和三维重建及其应用[M]. 上海: 上海交通大学出版社, 2015: 242-243.

(下转第 374 页)

观察,主动学习,主动总结。5) 建议科室制订激励机制,根据科室新技术、新业务使用情况给予一定的奖励,促使进修人员进修后必须持续主动学习。

全程无缝隙追踪管理模式运用了系统理论,对护理人员院外进修实行全程无缝隙追踪管理。严格选拔和评估是护理人员合理选择专业的前提,严格监管是护理人员完成计划的基础,严格督导及评价是实现进修目标的保障。全程无缝隙追踪管理加速了护理人员专业能力的培养,推动了护理科研的发展,促进了护理水平的提高。

参考文献:

- [1] 凌曦,任佰玲,朱似珏. 对我院近十年进修教育的分析[J]. 继续医学教育,1997,11(3):4-6.
- [2] 陈红,成翼娟. 教与学策略在高等护理教育中的应用研究[J]. 护理研究,2004(14):1223-1225. DOI:10.3969/j. issn. 1009-6493. 2004. 14. 001.
- [3] 刘于,李秀云. 专科护士培养方法探讨[J]. 护理学杂志,2006,21(9):42-43.
- [4] 安鸿章. 现代企业人力资源管理[M]. 北京:中国劳动出版社,2002:300-316.
- [5] 陈红,曾继红,李芸. 进修护士培训需求调查分析[J]. 现代护理,2006,12(20):1952-1953. DOI:10.3760/cma. j. issn. 1674-2907. 2006. 20. 058.
- [6] 吴秀云,李曼春,郭洪伟. 社区卫生技术人员知识与技能需求及影响因素分析[J]. 中国农村卫生事业管理,2004,24(8):17-19. DOI:10.3969/j. issn. 1005-5916. 2004. 08. 008.
- [7] 徐宁,周慧,赵存凤. 军队医院护理科研现状分析与对策[J]. 解放军医院管理杂志,2006,13(1):89-90. DOI:10.3969/j. issn. 1008-9985. 2006. 01. 052.
- [8] 刘建华,史晓娟,贾瑛. 护理实践与护理科研关系初探[J]. 中国实用护理杂志,2006,22(2):45-46.
- [9] 陈长香,吴庆文,李建民,等. 教学医院临床护理人员护理科研现状调查与分析[J]. 护士进修杂志,2005,20(3):253-255. DOI:10.3969/j. issn. 1002-6975. 2005. 03. 031.
- [10] 杨春. 护理人员院外进修的三严管理[J]. 护理学杂志,2009,24(15):68-69. DOI:10.3870/hlxzz. 2009. 15. 068.
- (收稿日期 2018-04-15)
(本文编辑:甘慧敏)
-
- (上接第 370 页)
- [9] 于晓青,曹慧,魏德建. 数据融合技术及其在医学领域的应用[J]. 中国医疗设备,2017,32(3):99-102.
- [10] Cheng X, Zhang L, Zheng YF. Deep similarity learning for multimodal medical images [DB/OL]. <https://www.tandfonline.com/doi/abs/10.1080/21681163.2015.1135299>, 2016-04-06.
- [11] 魏兴瑜,陆惠玲,周涛. 两模态 PET/CT 图像融合研究进展[J]. 重庆医学,2015,44(14):1979-1982.
- [12] Iglesias JE, Sabuncu MR. Multi-atlas segmentation of biomedical images: A survey[J]. Med Image Anal, 2015, 24(1):205-219. DOI:10.1016/j.media.2015.06.012.
- [13] 张巧丽,赵地,迟学斌. 基于深度学习的医学影像诊断综述[J]. 计算机科学,2017,44(S2):1-7.
- [14] 王保加,潘海为,谢晓芹,等. 基于多模态特征的医学图像聚类方法[J]. 计算机科学与探索,2018,12(03):411-422.
- [15] Zhang J, Liu M, Shen D. Detecting anatomical landmarks from limited medical imaging data using two-stage task-oriented deep neural network [DB/OL]. <https://ieeexplore.ieee.org/abstract/document/7961205/>, 2017-10-10.
- [16] Greenspan H, Summers RM. Guest editorial deep learning in medical imaging: overview and future promise of an exciting new technique [DB/OL]. <https://ieeexplore.ieee.org/document/7463094/>, 2016-05-05.
- [17] Gernot Riegler, Martin Urschler, Matthias Rütger, et al. Anatomical landmark detection in medical applications driven by synthetic data [DB/OL]. <https://ieeexplore.ieee.org/document/7406370/>, 2016-02-15.
- [18] Li R, Shi PC, Anne R, et al. Image understanding from experts' eyes by modeling perceptual skill of diagnostic reasoning processes [DB/OL]. <http://www.docin.com/p-1557550516.html>, 2016-05-02.
- [19] Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis [J]. Med Image Anal, 2017, 42:60-88. DOI:10.1016/j.media.2017.07.005.
- [20] Wang G. A perspective on deep imaging [DB/OL]. <https://ieeexplore.ieee.org/document/7733110/>, 2016-11-03.
- (收稿日期 2018-07-11)
(本文编辑:林琳)