·影像信息学专题·

自然语言处理在医学影像中的应用

马帅，王霄英

【摘要】 随着计算机技术快速发展，与影像诊断密切相关的软件研发有望超过传统硬件成为影像信息学发展的主流，自然语言处理(NLP)作为新兴技术在医学影像领域表现出良好的应用前景。本文概述 NLP 原理及其在医学影像中的应用，并对未来发展方向进行展望。

【关键词】 自然语言处理；医学标准术语；文本挖掘；结构化

【中图分类号】R814.41；R814.42；R445.2 【文献标识码】A 【文章编号】1000-0313(2016)12-1120-04

DOI:10.13609/j.cnki.1000-0313.2016.12.002

医学影像报告是电子健康病历(electronic health record，EHR)中包含大量数字信息的重要组成部分。但影像报告多以自由文本形式出现在 EHR 中，这种非结构式数据不利于信息的提取和利用。由于对报告信息进行人工提取耗时且难于操作，所以自然语言处理(natural language processing，NLP)技术成为医学影像报告信息化的重要工具[1]。NLP 通过计算机智能分析自由文本，并自动完成数据挖掘任务，将人类自然语言翻译成结构化形式[2]，从而有效地利用了报告中信息。本文概述 NLP 的原理及其在医学影像中的应用，并对未来发展方向进行了展望。

NLP 工作原理

NLP 从自然语言数据中推导出规则和模型，将文本转化为结构化的编码信息，从而可进行快速查询和分析。在 NLP 工作过程中涉及了语言学方法(如语法、语义和语境等)和统计方法。

虽然多种 NLP 的具体目标、技术、操作过程不尽相同，但主要工作原理基本相似，均可分为特征提取、特征加工、系统训练和验证几个步骤，现分述如下。

1. 特征提取

特征提取是指 NLP 分割文本、识别单个概念，并定义识别出的概念与其它医学概念的关系，输出结构式的数据。在特征提取过程中，先进行词汇分割，再进行词汇的语义分析。

按从大到小的尺度进行词汇分割。先将整个影像报告分割为若干段落，再分为句子、词组、词汇。在词汇层面上，判定词根、纠正拼写错误以及把缩略语扩充完整。

按从局部到整体的尺度进行词汇的语义分析。词汇的特征从局部到整体可分为：概念、词典和知识体系(ontology，计算机术语为“本体”)。“概念(concept)”指的是每个词汇被赋予的独特含义(如某种疾病)。“词典(lexicon)”指的是一组有相同含义的概念及其同意词、衍生词和相关术语等，如一体化医学语言系统(unified medical language system，UMLS)词典[3]或者 RadLex 词典[4]。“知识体系(ontology)”指的是每个概念与其它不同概念之间的相互关系，如本概念对其它概念所起到的限定、修饰作用等，如 SNOMED-CT。

通过特征提取，报告中的自然语言被分割为结构式的概念，且每个概念都被定义了与其它概念的关系，进一步用于后续的特征加工。

2. 特征加工

判断从报告中提取出的结构化数据是否包含目标概念，进而判断能否通过提取出的数据推导出某种临床结局。进行特征加工必须依据某种规则，通常有两种生成规则的方法：一种是专家制定规则；另一种是通过统计或机器学习方法从数据中自动推导制定规则。也可以联合制定规则，如先由机器学习产生规则，再由专家对其判断和校正。无论何种方法进行特征加工，所设定的规则均应进行训练和验证，才可进一步应用。

3. 系统训练和验证

完成特征加工后的系统，要进行训练和验证。在此过程中，应提供给系统足够的分类“标准答案”。通常情况下，训练时使用越大量的标准数据，并对其进行验证，越可保证实际使用中系统的稳定运行。但考虑到训练和验证的成本，在实际操作中用于各类学习任务的训练数据量可有一定差异，通常几百例数据对于大部分任务是足够的。

文献报告 NLP 验证的结果通常较好，在许多系统中其敏感度和特异度均超过 90%。在不同软件系统、不同应用目的、不同时间点进行测试，其性能未表现出明显差异。

4. 部分 NLP 相关资源

近年来，不同机构发布了多种 NLP 工具[1-2]，其目的、任务有所不同，可根据不同的研究目的来选用，具体见表 1。

NLP 临床应用概述

根据信息提取的对象和目的不同，NLP 可用于患者个体信息分析、患者群体信息分析和医学影像流程信息分析等。

1. 患者个体影像诊断信息提取和分析，对患者个体疾病处理提供帮助

提示“危急发现(critical findings)”：NLP 检出影像报告中描述的、可能导致严重后果的影像征象，提醒处理该患者的医师注意[5]。目前 NLP 可提示的危急情况有阑尾炎、急性肺损伤、肺炎、血栓栓塞性疾病及各类潜在恶性病变等[6]。如 NLP 在报告中发现危急情况，会提示影像医师及时与临床医师交流。

提示随访建议：NLP 检出报告中应提示临床进行后续操作

作者单位：100034 北京，北京大学第一医院医学影像科
作者简介：马帅(1987－)，男，山东庆云人，博士研究生，主要从事影像诊断研究工作。
通讯作者：王霄英，E-mail：cjr.wangxiaoying@vip.163.com

表 1 部分 NLP 相关资源

名称	描述	链接
Apache OpenNLP	基于机器学习，可支持常规 NLP 任务的工具包	http://open nlp. apache. org/
BRAT	文本结构化协同标注，免费开源	http://brat. nlplab. org/
BROK	基于 Java，确定 BI-RADS 分类	http://www. brighamandwomens. org/Research/labs/cebi/BROK/default. aspx
cTAKES	提取临床电子病历自由文本信息，开源	http://ctakes. apache. org/
dtSearch	搜索引擎和索引生成器，付费使用	http://www. dtsearch. com/PLF_Features_2. html
eHOST	人工标注临床文本，支持编码标准临床词汇（如 SNOMED-CT），开源	http://code. google. com/p/ehost
GATE	Java 工具套件，含信息提取系统（ANNIE）	https://gate. ac. uk/
I2E	文本挖掘软件，可查询非结构性文本，付费使用	http://www. linguamatics. com/welcome/software/I2E. html
iSCOUT	利用知识体系检索影像报告特殊发现的工具包	http://sourceforge. net/projects/iscout/
LEXIMER	对非结构化影像报告进行提取、结构化处理和分类，付费使用	http://www. nuance. com/index. htm
LifeCode	编码临床叙述性报告，财务用途，付费使用	http://www. optum360. com/hospital/coding-documentation/clinical-documentation-improvement. html
MALLET	基于 Java，用于统计、文本分类、集群、主题建模、信息提取及其它针对文本的机器学习	http://mallet. cs. umass. edu/
MEDINA	用于法语电子病历的去标识	https://medina. limsi. fr/index-en. html
MedLEE	提取和编码临床叙述性文本（包括影像报告、出院记录和病理报告），付费使用（Health Fidelity, PaloAlto, Calif）	http://healthfidelity. com
MetaMap	将生物医学文本匹配到 UMLS，自动或半自动索引 NLM 生物医学文献	http://metamap. nlm. nih. gov
NegEx	模式匹配，检测语句中索引词组的否定状态	https://code. google. com/p/negex/
ONYX	整合语法、语义相关知识以解释自由文本，可经训练应用于特定领域文献，开源	http://aclweb. org/anthology/W09-1303
OpenNLP	机器学习，处理自然语言文本	https://opennlp. apache. org/
Porter stemming algorithm	英文词干的分析算法	http://tartarus. org/～martin/ PorterStemmer/index. html
Protégé	显示和翻译标注，免费开源	http://protege. stanford. edu/
RadLex	影像信息资源的标准化索引和检索词典	http://www. radlex. org/
Render	在线搜索影像数据库，可查询影像报告和图像	Not publicly available
SAPHIRE	以概念匹配、自动索标、概率检索和层级关系为特征的信息检索系统	http://dx. doi. org/10. 1016/0010-4809(90)90031-7
SMILE Text Ana-lyzer/Stemmer	简易在线词汇标记和词干分析	https://smile-pos. appspot. com/; http://smile-stemmer. appspot. com
SNOMED-CT	综合性多语言临床术语集，包括：临床表现、症状、诊断、治疗经过、身体结构、微生物及其他病因、机体、用药、设备和标本等术语；概念以分级结构呈现，展示概念间的连接关系	http://www. ihtsdo. org/snomed-ct
Stanford University NLP	统计学工具和资源的综合列表；包括 Stanford Log-linear Part-of-Speech Tagger	http://nlp. stanford. edu/links/statnlp. html
SymText	集成语法和语义分析的医学领域工具	http://www. ncbi. nlm. nih. gov/pmc/articles/PMC2579100/
UIMA	非结构化信息分析	https://uima. apache. org/
UMLS	生物医学科学领域的词汇一览表及分类系统	http://www. nlm. nih. gov/research/umls/
WEKA	借助机器学习算法进行数据挖掘，可作为独立程序使用或从 Java 语言程序中调用，开源	http://www. cs. waikato. ac. nz/ml/index. html
YTEX	耶鲁大学 cTAKES 的扩展：临床 NLP、语义相似、数据挖掘以及特征管理	https://code. google. com/p/ytex/

注：ACR = American College of Radiology，美国放射学会；BI-RADS = Breast Imaging-Reporting and Data System，乳腺影像报告与数据系统；BROK = BI-RADS Observation Kit，乳腺影像报告与数据系统观察工具；cTAKES = Apache clinical Text Analysis and Knowledge Extraction System，Apache 临床文本分析及知识提取系统；GATE = General Architecture for Text Engineering，文本设计的整体构建；HOST = Extensible Human Oracle Suite of Tools，可拓展人类 Oracle 套件；LEXIMER = Lexicon Mediated Entropy Reduction，词典介导的熵约简；NLM = National Library of Medicine，美国国家医学图书馆；MedLEE = Medical Language Extraction and Encoding System，医学语言提取和编码系统；MALLET = Machine Learning for Language Toolkit，机器学习语言工具包；MEDINA = Medical Information Anonymization，医学信息匿名化；UIMA = Unstructured Information Management applications，非结构化信息管理应用；UMLS = Unified Medical Language System，一体化医学语言系统；WEKA = Waikato Environment for Knowledge Analysis，怀卡托智能分析环境系统。

的内容，自动生成随访建议，提示后续检查或治疗[7]。

提示偶然发现[8]：利用机器学习分类器可检出有临床意义的偶然发现，避免临床医生忽略该发现而造成延迟诊治[9]。

检查报告中的逻辑错误：根据预先设定好的逻辑，可检出有明确逻辑矛盾的内容，提示报告医师是否可能为误读、误判或误操作。

2. 患者群体影像诊断信息提取和分析，构建患者队列，用于流行病学研究、行政管理等

流行病学研究队列的构建：使用 NLP 可高效率地分析大数量、患者群体的影像报告，得到群体的特征性数据。使用传统方法构建流行病学研究患者队列，需耗费大量时间和人力才能筛选出合适病例，而 NLP 可提高流行病学研究效率，为循证影像医学研究提供帮助[11-15]。

在医院或社会群体水平监控公共卫生情况：NLP 可用于评估区域健康情况[10]。利用从图像中提取的群体 NLP 特征值和其它结构化电子病历数据来监控公共健康水平，进行决策分析[16-17]。

3. 医学影像流程信息的提取和分析，用于医学影像报告质量评价和改进

报告质量评价和报告规范的建立：NLP 可识别医学影像学的流程和质量指标，判断影像报告是否符合相关指南或诊断规则[18]。对大量影像报告中的海量数据进行自动内容分析可反映影像科日常工作运行情况。目前 NLP 系统已可用于评价报告的完整性和规范性，是否给出正确的建议，是否及时进行危急情况的预警，报告信息是否用于疾病的诊断等方面[5,19-21]。利用 NLP 结果，对建立报告规范可起到指导作用。

医师个人表现评价和改进建议：NLP 可针对医师完成报告的表现进行评价，用于诊断医师个人的质量评价[22-23]。在对诊断医师表现进行评判的基础上提出改进建议[24]。

影像检查全流程的改进：NLP 可对各类影像的综合信息进行分析，将报告中的检查结果和建议等信息与全面的临床信息相互关联，如检查适应证、疾病种类、患者年龄、性别、申请科室、申请医师及患者类型（住院或门诊）等[25]。这种大规模的数据分析在经过验证后，可得到预测模型，形成适合本地情况的临床决策支持系统（clinical decision support system，CDSS），应可应用到计算机医嘱系统（computerized physician order entry，CPOE）中去[26]，对影像检查从申请开始、到临床应用结束的全过程进行高质量、高效率和标准化的管理。

行政和财务方面的应用：NLP 可将影像报告结论自动匹配到医疗编码系统（比如将影像报告结论和 ICD-10 编码实现自动匹配），对于医政管理、财务及决策制定等工作有帮助[27]。

NLP 应用中分析

医学影像中使用 NLP 的总体目标是挖掘诊断报告中结构化信息，并将其应用于临床诊治过程。尽管应用目的多种多样，但目前大部分 NLP 系统都用于判定影像报告中是否包含了特定影像学发现（如某疾病表现或特定发现）[28]。NLP 的主要优势就是自动化，减少甚至免除人工审阅报告的精力并实现对大量数据资料的评估，因此之前难于操作的任务如今也能轻松实现。NLP 的另一个优势是可对影像报告书写过程进行监测，直接对诊断或临床医师提出建议。

但 NLP 实际应用中尚待解决的问题仍很多，主要有以下几点。①NLP 虽有不少临床应用案例，但基本上仍处于初步探索阶段，关键性问题尚待解决。NLP 中各类技术指标的确定，并非由临床需求本身决定，而是取决于医疗机构可获得的技术工具，特别是 NLP 系统开发者的专业知识和业务水平。客观地讲，NLP 仍处于概念验证阶段，对实际临床问题的解决效能、以及解决问题所带来的实际临床价值，尚未获得足够的证据支持。②NLP 处理信息的规则不明晰，使得 NLP 不易被接受。在特征分析过程中，如使用了专家制定的规则，则较易于被理解，其结论也易于被使用者所接受。但如果使用了机器学习方法，系统的逻辑规则是由计算机系统通过数据分析得到的，这个过程很复杂，不可能明确说明其内在逻辑，此时临床医师常常不愿意接受自动算法的结果。③原始影像报告未达标准化，也使得 NLP 结果不易推广。在传统工作模式下，影像报告的书写与医师个人知识、工作习惯有关，也与本单位的规则、管理要求等有关，但目前多数报告尚不符合标准化要求，对 NLP 的应用效果造成不利影响[29]。实际上，报告本身的标准化并不是主要困难，整个影像链的非标准化造成了 NLP 应用更大的困难。只有遵循全影像链的规范化标准操作，NLP 的结果才能最终得到推广普及。

NLP 未来发展方向

在影像工作过程中，亟待进一步研发和探索 NLP 更有价值的临床应用：从一系列影像报告中自动做出疾病病程的评价；挖掘临床信息与影像信息的内在联系；对影像报告的综合结果进行编码，用于特定患者的队列构建，进行病历自动审核功能等。NLP 系统不仅可用于影像报告生成之后，还应在报告生成过程中直接对报告提供帮助，辅助影像医师更加高效、高质量、规范化地书写报告；将诊断报告中的结构式信息与相关指南和临床信息进行关联，自动提供处理建议等。

在临床工作过程中，NLP 应把其它相关医学信息与影像信息整合，使影像工作从检查申请开始、到临床发布过程中都明显受益，如：NLP 在海量电子病历信息中识别出某些临床诊断和临床需求，将这些信息提供给影像医师，使得影像检查开始之前即可明确检查目的，使检查全过程的效率和质量都得到明显提高。

总之，随着 NLP 在医学应用中的推广，这项技术手段将进入到医学影像常规工作中，其价值也将会逐渐被发现和确认。以 NLP 为代表的信息学工具的使用，终将改变医学影像工作的流程、效率和质量，使医学影像工作模式因此而发生明显的转变。

参考文献：

[1] Cai T, Giannopoulos AA, Yu S, et al. Natural language processing technologies in radiology research and clinical applications[J]. Radiographics, 2016, 36(1): 176-191.

[2] Pons E, Braun LM, Hunink MG, et al. Natural language processing in radiology: a systematic review[J]. Radiology, 2016, 279(2): 329-343.

[3] U. S. National Library of Medicine. Unified Medical Language System (UMLS)[D/OL]. July 29, 2016. http://www.nlm.nih.gov/research/umls/.

[4] Radiological Society of North America. RadLex[D/OL]. 2016 October 10. http://www.rsna.org/RadLex.aspx. Accessed,.

[5] Lakhani P, Kim W, Langlotz CP. Automated extraction of critical test values and communications from unstructured radiology reports: an analysis of 9.3 million reports from 1990 to 2011[J]. Radiology, 2012, 265(3): 809-818.

[6] Chapman WW, Fizman M, Chapman BE, et al. A comparison of classification algorithms to automatically identify chest X-ray reports that support pneumonia[J]. J Biomed Inform, 2001, 34(1): 4-14.

[7] Zingmond D, Lenert LA. Monitoring freetext data using medical language processing[J]. Comput Biomed Res, 1993, 26(5): 467-481.

[8] 杜婧, 王霄英. ACR 关于腹盆部检查中偶然发现的处理原则[J]. 放射学实践, 2015, 30(12): 1225-1231.

[9] Dutta S, Long WJ, Brown DF, et al. Automated detection using natural language processing of radiologists recommendations for additional imaging of incidental findings[J]. Ann Emerg Med, 2013, 62(2): 162-169.

[10] Liu V, Clark MP, Mendoza M, et al. Automated identification of pneumonia in chest radiograph reports in critically ill patients [J]. BMC Med Inform DecisMak, 2013, 13(90): 1-8.

[11] Sada Y, Hou J, Richardson P, et al. Validation of case finding algorithms for hepatocellular cancer from administrative data and electronic health records using natural language processing[J]. Med Care, 2016, 54(2): 9-14.

[12] Carrell DS, Halgrim S, Tran DT, et al. Using natural language processing to improve efficiency of manual chart abstraction in research: the case of breast cancer recurrence[J]. Am J Epidemiol, 2014, 179(6): 749-758.

[13] Do BH, Wu A, Biswal S, Kamaya A, et al. Informatics in radiology: RADTF-a semantic search-enabled, natural language processor-generated radiology teaching file[J]. RadioGraphics, 2010, 30(7): 2039-2048.

[14] Chang EK, Yu CY, Clarke R, et al. Defining a patient population with cirrhosis: an automated algorithm with natural language processing[J]. J Clin Gastroenterol, 2016, 50(10): 889-894.

[15] Masino AJ, Grundmeier RW, Pennington JW, et al. Temporal bone radiology report classification using open source machine learning and natural langue processinglibraries[J/OL]. BMC Med Inform Decis Mak, 2016, 16: 65. DOI: 10.1186/S12911-016-0306-3.

[16] Pham AD, Névéol A, Lavergne T, et al. Natural language processing of radiology reports for the detection of thromboembolic diseases and clinically relevant incidental findings[J/OL]. BMC Bioinformatics, 2014, 15(1): 266. DOI: 10.1186/1471-2105-15-266.

[17] Jensen PB, Jensen LJ, Brunak S. Mining electronic health records: towards better research applications and clinical care[J]. Nat Rev Genet, 2012, 13(6): 395-405.

[18] Dutta S, Long WJ, Brown DF, et al. Automated detection using natural language processing of radiologists recommendations for additional imaging of incidental findings[J]. Ann Emerg Med, 2013, 62(2): 162-169.

[19] Ip IK, Mortele KJ, Prevedello LM, et al. Focal cystic pancreatic lesions: assessing variation in radiologists' management recommendations[J]. Radiology, 2011, 259(1): 136-141.

[20] Duszak R, Nossal M, Schofield L, et al. Physician documentation deficiencies in abdominal ultrasound reports: frequency, characteristics and financial impact[J]. J Am Coll Radiol, 2012, 9(6): 403-408.

[21] Ip IK, Mortele KJ, PrevedelloLM, et al. Repeat abdominal imaging examinations in a tertiary care hospital[J]. Am J Med, 2012, 125(2): 155-161.

[22] 杜婧, 王霄英. ACR 影像报告及数据系统介绍[J]. 放射学实践, 2016, 31(4): 331-335.

[23] Dreyer KJ, Kalra MK, Maher MM, et al. Application of recently developed computer algorithm for automatic classification of unstructured radiology reports: validation study[J]. Radiology, 2005, 234(2): 323-329.

[24] Sevenster M, Bozeman J, Cowhy A, et al. A natural language processing pipeline for pairing measurements uniquely across free-text CT reports[J]. J Biomed Inform, 2015, 53(1): 36-48.

[25] Dang PA, Kalra MK, Blake MA, et al. Natural language processing using online analytic processing for assessing recommendations in radiology reports[J]. J Am Coll Radiol, 2008, 5(3): 197-204.

[26] Patel TA, Puppala M, Ogunti RO, et al. Correlating mammographic and pathologic findings in clinical decision support using natural language processing and data mining methods[J]. Cancer, 2016 Aug 29. doi: 10.1002/cncr.30245.

[27] Farkas R, Szarvas G. Automatic construction of rule-based ICD-9-CM coding systems[J/OL]. BMC Bioinformatics, 2008, 9(Suppl 3): S10. DOI: 10.1186/1471-2105-9-S3-S10.

[28] Sevenster M, Buurman J, Liu P, et al. Natural language processing techniques for extracting and categorizing finding measurements in narrative radiology reports[J]. Appl Clin Inform, 2015, 6(3): 600-610.

[29] 秦岫波, 王蕊, 高歌, 等. 前列腺多参数 MRI 报告进展: 基于第 2 版前列腺影像报告和数据系统的结构式报告的构建[J]. 肿瘤影像学, 2016, 25(2): 111-116.

(收稿日期: 2016-10-20 修回日期: 2016-11-05)