

DOI: <https://doi.org/10.17816/DD321420>

Технология распознавания речи в лучевой диагностике

Н.Д. Кудрявцев¹, К.А. Бардасова², А.Н. Хоружая¹¹ Научно-практический клинический центр диагностики и телемедицинских технологий, Москва, Российская Федерация² Уральский государственный медицинский университет, Екатеринбург, Российская Федерация

АННОТАЦИЯ

Устройства, способные распознавать речь, являются перспективным инструментом для системы здравоохранения. Технология распознавания речи имеет довольно длинную историю применения в западных системах здравоохранения (с 1970-х годов), однако широкое распространение она получила лишь в начале XXI века, заменив медицинских транскрипционистов. Для отечественного здравоохранения данная технология относительно новая. Её активная разработка началась лишь в начале 2010-х годов, а повсеместное внедрение в здравоохранение — в конце 2010-х годов. Такая задержка связана с особенностями русского языка и ограничением вычислительных мощностей, присутствующих в начале XXI века.

В настоящее время комплексы устройств и программного обеспечения для распознавания речи используются в голо-совом заполнении медицинской документации и позволяют сократить время подготовки протоколов рентгенологиче-ских исследований при сравнении с традиционным (клавиатурным) вводом текста.

В литературном обзоре отражена краткая история развития и применения технологии распознавания речи в лучевой диагностике. Отражены ключевые научные исследования, подтверждающие эффективность её использования в за-падных системах здравоохранения. Продемонстрирован отечественный опыт применения технологии распознавания речи и оценена её эффективность. Описаны перспективы дальнейшего развития данной технологии в российском здравоохранении.

Ключевые слова: научный обзор; технология распознавания речи; голосовой ввод; лучевая диагностика; рентгенология; заполнение медицинской документации.

Как цитировать

Кудрявцев Н.Д., Бардасова К.А., Хоружая А.Н. Технология распознавания речи в лучевой диагностике // *Digital Diagnostics*. 2023. Т. 4, № 2. С. 185–196.DOI: <https://doi.org/10.17816/DD321420>

DOI: <https://doi.org/10.17816/DD321420>

Speech recognition technology in radiology

Nikita D. Kudryavtsev¹, Kristina A. Bardasova², Anna N. Khoruzhaya¹

¹ Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation

² Ural State Medical University, Ekaterinburg, Russian Federation

ABSTRACT

Speech recognition devices are promising tools for the healthcare system. Speech recognition technology has had a relatively long history of use in Western healthcare systems since the 1970s. However, it became widely used at the beginning of the 21st century, replacing medical transcriptionists. This technology is relatively new in home healthcare. Its active development began only in the early 2010s, and its implementation in healthcare started in late 2010. This delay is due to the idiosyncrasies of the Russian language and the limited computational power present at the beginning of the 21st century.

Currently, complexes of devices and software for speech recognition are used in the voice filling of medical records and can reduce the time for preparing reports for radiological examinations compared with traditional (keyboard) text input.

The literature review provides a brief history of speech recognition technology development and application in radiography. Key scientific studies showing its efficacy in Western healthcare systems are reflected. Voice recognition technology in the home is demonstrated, and its effectiveness is evaluated. The prospects for further development of this technology in Russian healthcare are described.

Keywords: medical records; radiation diagnostics; radiology; speech recognition software; voice input.

To cite this article

Kudryavtsev ND, Bardasova KA, Khoruzhaya AN. Speech recognition technology in radiology. *Digital Diagnostics*. 2023;4(2):185–196.

DOI: <https://doi.org/10.17816/DD321420>

Received: 17.03.2023

Accepted: 19.04.2023

Published: 25.05.2023

DOI: <https://doi.org/10.17816/DD321420>

语音识别技术在放射诊断中的应用

Nikita D. Kudryavtsev¹, Kristina A. Bardasova², Anna N. Khoruzhaya¹

¹ Moscow Center for Diagnostics and Telemedicine, Moscow, Russian Federation

² Ural State Medical University, Ekaterinburg, Russian Federation

简评

能够进行语音识别的设备是保健系统的一个有前途的工具。语音识别技术在西方医疗系统中有相当长的使用历史（自20世纪70年代以来），但它在21世纪初才得到了广泛推广，取代了医疗抄写员。对于国内的医疗保健来说，该技术是相对较新的。它的积极开发是在2010年代初才开始，并2010年代末才在保健事业广泛采用的。这种延迟是由于俄语的特点和21世纪初计算能力的限制而导致的。

语音识别的设备和软件包现在被用于通过语言输入填写病历，此外，与传统（用键盘）文本输入相比，减少了准备X射线学协议所需的时间。

本文综述简要介绍了语音识别技术在放射诊断中的发展和应用的历史。介绍了证实其在西方医疗系统中使用的有效性的主要科学研究。展示了国内使用语音识别技术的经验，并对其有效性进行了评估。描述了该技术在俄罗斯保健事业进一步发展的前景。

关键词：科学综述；语音识别技术；语音输入；放射诊断；X射线学；填写病历。

引用本文：

Kudryavtsev ND, Bardasova KA, Khoruzhaya AN. 语音识别技术在放射诊断中的应用. *Digital Diagnostics*. 2023;4(2):185–196.

DOI: <https://doi.org/10.17816/DD321420>

收到: 17.03.2023

接受: 19.04.2023

发布日期: 25.05.2023

绪论

如今，声音控制已成为许多智能家居设备的标准功能。这要归功于语音识别技术的发展。基于语音识别技术的系统能够分析用户的语音并将其转换为数字数据。除了控制智能设备外，语音识别技术在电话通信中也得到了广泛应用。在许多政府和商业机构的电话中都使用了电话答录机。它可识别来电者的语音请求并将其转接给选定的专家。2019年以来，在莫斯科实行一个项目，在语音助手的帮助下通知市民预约挂号并提醒他们要接受临床检查。在这样的通话中，市民可以挂号看病，取消或重新安排就诊时间。该系统还询问病人哪儿不舒服¹。

在医疗保健领域，语音识别系统已被广泛应用于医疗记录的语音填写。这是因为医生的大部分工作时间都花在准备病历上[1-4]。这一因素对医疗质量产生了负面影响，尤其是在接待病人时间有限的情况下。使用技术，如在放射诊断中，可以减少填写诊断检查总结的时间，将更多的时间用于研究诊断图像、随附医疗文件以及与患者交流。正因为如此，语音输入系统在放射科最受欢迎。放射科中的工作流程组织最便于引进这种技术。现代系统综述 [5-7]表明了，在这些条件下使用语音识别系统是有效的。良好的接受能力是由于大量的文本信息。放射科医生应该将这些信息记录在总结中。

语音识别技术在放射学中的应用历史 早期应用

语音识别技术在医疗领域的首次尝试是在20世纪70年代和80年代。1975年，一个语音识别系统的原型被描述出来了。该系统可以分析非常有限的医学词汇和术语[8]。1981年，在放射诊断部门对该技术首次进行了测试[9]。在以色列的Beth Israel Hospital，该系统开始被用于准备放射检查协议书。起初，该医院的医生使用了CLIP (Coded Language Information Processing) 系统。CLIP是一种分层标准化的医学术语语言。它由M. Simon和B. W. Leeming 开发[10]。该语言的结构包含医学术语。这些术语以特定方式编码。例如，值A表示anatomic area (解剖区域)，A6表示the lower limb (下肢)，A61表示the thigh (大腿)。而值B、B6、B61、B611和B612分别表示bone (骨头)、the lower limb bones (下肢骨)、the femur (股骨)、the lesser trochanter (小转子)和the grosser trochanter (大转子)。放射师使用键盘输入了各种代码组合，能够写任何复杂度和规模的总结[11]。

当时的语音识别系统有一个很大的缺点，即词汇量有限。词汇量包括大约200个独特的医学术

语。这是由于内存容量较小。这一因素导致无法全面描述诊断检查。将语音识别技术与CLIP系统结合使用是解决方案之一。这样就可以规避工作记忆容量的限制。

1981年的一项研究比较了键盘输入和语音输入[9]。对60份放射检查协议书的准备速度和质量进行了评估。结果表明了，协议书的规模与填写方法无关。这间接证实了所述研究的复杂度相似。使用语音输入填写总结的持续时间超过键盘输入的4倍。语音输入的协议书平均包含12个识别错误。而键盘输入的协议书则没有错误。这项研究还描述了该技术的其他局限性。最初的系统不能完全抑制外部噪音，因此对医生语音的识别质量较低。这导致了错误的发生。填写一份医疗文件所需的时间较长，这是因为系统只能识别单独说出的单词。医生必须在每个词之后停顿一下，或者在CLIP系统中，在每个代码之后停顿一下。这种语音填写方法对人类来说既不舒服也不自然。值得注意的是，语音输入系统需要3到6个小时的预先调整和应用户（医生）的特定语音。

所有这些限制都阻碍了当时的语音识别系统在医疗保健领域的普及。因此，当时医疗转录员的参与是合理的。然而，在随后的发展阶段中，出现了一个基础，可以将技术的“弱点”考虑在内。

随着放射检查的日益普及，计算机和磁共振断层扫描仪的出现，以及从模拟介质向数字介质的过渡，导致了放射科医生的工作量增加，总结准备时间延长。20世纪80年代中期，西方国家建立了记录中心，以解决这一问题并优化人力资源。放射师向录声系统口述放射检查结果。这些结果是在诊断图像解读过程中发现的。录音被传送到录音中心。医学转录员译解了录音并以文本形式编写检查总结。准备好的总结由医生检查并签署。在某些情况下，转录员会将其收回来以纠正错误[12]。记录中心在海外医疗机构一直流行到2010年代[5, 13]。

一系列作者比较了医疗转录员准备与使用语音识别技术准备的放射检查总结的有效性。做出了如下结论：医疗转录员的主要优势在于能够发现语法错误并考虑到患者的上下文信息。这些优势有助于正确理解并同意医生的录音，即使录音质量很差[14]。然而，由于医疗转录员短缺，经常出现口述16小时后才将打字的总结交还给医生的情况[15]。根据记录中心的能力，录音的译解需要6至24小时。

语音识别系统的进一步发展最终取代了医疗转录员[14]。语音识别技术在识别准确性上不如医疗转录员好，但它允许将放射科每月的经济成本降低81%，并显著减少准备检查总结所需的时间[5, 13, 16]。

¹ 莫斯科市长官方网站[互联网]。语音助手将在综合诊所门诊之前通过电话收集关于病人哪儿不舒服的信息。访问方式：<https://www.mos.ru/news/item/89302073/>。

也有人尝试在俄罗斯医疗保健系统中引入记录中心²。然而，评估该中心的有效性和应用的科学研究没有开放获取。现代国内医疗保健系统中没有这种中心，因此可以认为使用这种中心被证明是不合适的。

20世纪90年代：新一轮创新发展

20世纪末，语音识别系统的内存和词典容量增加到19,000个放射学术语。预设时间缩短至几分钟。识别准确率得到了提高。1995年，美国推出了第一个自然语言译解程序。它们能够确定连贯的英语发音。现在，医生可以以舒适的对话速度口述，每个词之间没有停顿。当时这些程序的识别准确率低于分词输入系统[17]。技术的进一步发展和语音识别准确率的提高使得连贯语音输入系统取代分词输入系统。

21世纪

本世纪初，西方国家开始在放射科大规模引进和应用自动语音识别技术。国外的研究中比较了英文医疗文件的语音填写速度与键盘输入、医疗文件量和医生满意度。获得的数据显示，技术的应用使填写文件的速度提高了26%，协议书的数量也有所增加。语音输入也优化了工作流程。准备医疗文件的时间缩短了。协议书的内容质量得到了提高。医生对文件工作的满意度也提高了[16, 18]。错误发生的频率也降低了。主要错误是标点符号错误[7]。上述因素有助于检查总结准备的总时间从16小时减少到5小时[15]。评估总结准备时间的研究结果、评估每分钟平均字符数的结果、错误的数量和频率反映了放射科中技术的发展和引进的积极动态。在一小时内准备好总结的百分比从26%增加到58%。协议书在内容上更加结构化[19]。

时间一长经济成本也有所下降。因此，7篇关于成本估算的系统综述中的5篇[5]包含了成本降低的信息。只有2篇文章报道了经济成本的增加。内分泌学和精神病学也有类似的结果。语音识别系统的使用提高了医生工作的生产力和效率[20]。在外科学中使用这些系统有助于将准备手术总结的时间从4天缩短到3天。在一天内准备好手术总结的百分比从22%增加到37%[21]。国外2019-2020年的研究作者认为，语音识别技术节省

了医生的时间，提高了他们的工作效率，以及允许在填写医疗文件时记录更多重要细节[22-24]。然而，医生使用语音输入系统的主要障碍是人为因素，即对变革的抵触和对新技术的恐惧[25]。

2016年，Microsoft research研究表明了，语音识别系统的准确率已达到人类能力的94%[26]。目前，在英语国家，这项技术现已深入医疗实践。放射科的语音输入采用率已达到85%³。目前，医疗保健软件市场的份额约占全球市场的25%⁴。Nuance Communications、IBM和Philips⁵都是语音识别软件开发领域的领导公司。

大约45年来，语音识别在英语国家的医疗领域变得越来越普遍。从初级医疗和急诊到高度专业化科室，语音识别已深入到医疗保健的各个层面。现代医学英语语音识别系统的准确率高达99%。它们能够适应不同的口音，并且不需要事先对医生的语音模式进行训练⁶。

现代语音识别系统如何工作？

如前所述，语音识别技术的基础是利用计算机将人的口语翻译成文本的过程。人工智能算法被用于现代语音输入系统。人工智能算法允许大大提高用户语音识别的质量和速度[27]。识别过程由几个阶段组成。每个阶段都有自己的特点（图1）。

1. **采集音频信号。**语音识别的第一阶段是声波记录。声波记录可以借助麦克风完成。麦克风可内置在任何音频记录设备中，如智能手机。然后将声波转换为数字格式，以进行计算机处理。
2. **预处理音频信号。**在此阶段，对接收到的音频数据进行预处理。处理过程可以消除无关的背景噪声，突出用户的语音。这将提高语音识别的质量。
3. **分割成片段。**然后将音频文件分割成长度约为10-25毫秒的小片段。这是必要的，以优化音频数据分析。语音识别系统分别对每个片段进行分析。
4. **提取语音特征。**在此阶段，计算语音特征（即声音的频率、持续时间和振幅）。语音特征被用于识别音素。这些音素组成口语单词。
5. **进行识别模型匹配。**音素匹配是使用声学模型进行的，在大量语音样本上对这些模型进

² 国家预算卫生机构“莫斯科市卫生管理局传染科临床医院第1”的官方网站[互联网]。医院历史。访问方式：<https://ikbl.ru/about/>。

³ Reaction Data[互联网]。Speech Rec in Radiology-State of the Market。2019[引用日期：2019.12.23]。访问方式：<https://www.reactiondata.com/report/speech-recognition-in-radiology-state-of-the-market/>。

⁴ Grand View Research[互联网]。Voice And Speech Recognition Market Size Report, 2030。访问方式：<https://www.grandviewresearch.com/industry-analysis/voice-recognition-market>。

⁵ Nuance Communications。Healthcare Clinical Documentation AI Solutions & Services for the NHS (<https://www.nuance.com/en-gb/healthcare.html>)；Philips。Healthcare—Philips (<https://www.dictation.philips.com/gb/industries/industry/healthcare-professionals/>)；IBM。Watson Speech to Text (<https://www.ibm.com/cloud/watson-speech-to-text>)。

⁶ Nuance Communications[互联网]。Dragon Medical One—#1 Clinical Documentation Companion。访问方式：<https://www.nuance.com/healthcare/provider-solutions/speech-recognition/dragon-medical-one.html>。

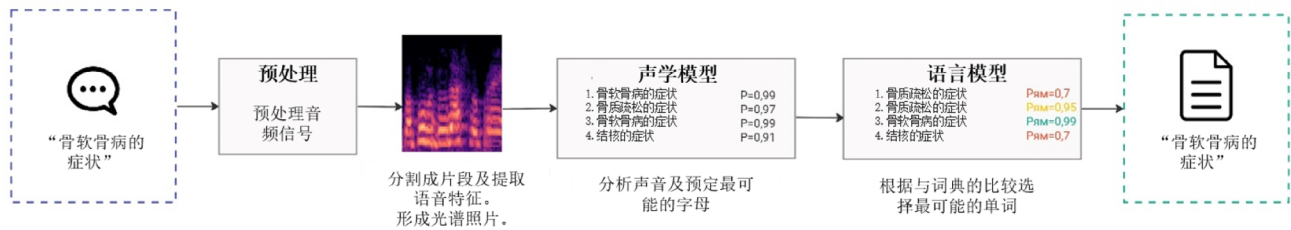


图1. 经典语音识别系统的简化操作方案。图1给出识别短语“骨软骨病的症状”的算法。

注：ЯМ——来自俄语的 языковая модель（语言模型）。

行训练。这些模型可以使用多种方法，包括隐马尔可夫模型、神经网络和其他机器学习算法[28-30]。

6. **组合单词和文本。**在这一阶段，语言模型将音素组合成单词和短语，然后组合成完整的文本。该文本可能需要额外处理，以纠正拼写错误和其他错误。
7. **插入组合的文本。**这是最后一阶段。将组合的文本插入医疗文件中。由于语音识别系统的准确率尚未达到100%，在某些情况下，医生必须对识别出的文本进行手修正。

声学 and 语言模型使用人工智能模块，以将声音更准确地转换成字母。无论口述者的性别、年龄和语调如何，都必须识别声音。开发人员在数据集上训练神经网络。同时，该数据集包括大量录音和文本的示例。当接收到语音信号时，神经网络在数据集中“搜索”匹配的语音信号。神经网络会继续学习，并在使用时创建新的声音与字母组合。这使其更有可能重现预期文本，而不会出现拼写错误。在训练过程中，计算机可识别音素发音的最重要特征。计算机以用户配置文件的形式记录获得的数据⁷[31]。

最近，建立语音识别系统的所谓端到端方法得到了广泛应用。语音识别技术中的端到端方法是指自动执行整个语音信号处理的方法[32]。无需执行单独的处理步骤（如特征提取和语音识别模型创建）[32]。深度学习（deep learning）是最流行的端到端方法之一。在这种情况下，在原始音频数据上直接对神经网络进行训练，无需进行预处理[32]。端到端方法的另一个例子是使用循环神经网络（recurrent neural network, RNN）或卷积神经网络（convolutional neural networks, CNN）进行语音识别。在这种情况下，神经网络在输入信号和输出文本上学习。在这种情况下，神经网络使用监督学习技术[32]。

端到端方法的优势是它们可以提供更高的识别准确率，因为所有语音信息都被用于训练神经网络。然而，这些方法的实现可能更加复杂，并且需要更多数据进行训练[33]。

国内在放射诊断中使用语音识别技术的经验

最早的俄语语音识别系统出现在2000年代中期[34, 35]，但由于使用通用词汇，无法在医疗实践中使用该语音识别系统。又过了几年，才开发出能够译解包含医学术语的俄语语音输入系统[36]。语音识别技术的积极发展直到2010年代后半期才出现。这一延迟是由于难以识别俄语语音。

与英语相比，俄语的构词结构更为复杂，因为俄语是一种合成语言，有大量的词形。为了识别单词，必须使用较大的词典。这降低了系统的性能[37]。例如，现代英语语音识别系统使用的词典包含多达30万个单词和术语，而俄语词典可包含500多万个单词、词形和词组[33, 38, 39]。此外，同一单词的大多数词形只有词尾不同，而用户往往发音不清。这就导致了识别整个词组的错误，于是需要对最终文档进行修正。并且俄语在句子中的单词排列变体较多，而英语则使用严格的语法结构。这导致为语音识别系统创建语言模型的过程变得困难，并降低了其工作的准确性。

STC (Speech Technology Center, 语音技术中心) 集团是当今俄罗斯医疗保健领域语音输入系统开发的领导公司⁸。2020年，对放射科使用语音识别技术的有效性进行了首次研究。该研究是在莫斯科市卫生局的七家城市综合医院进行的。研究中使用的是Voice2Med语音输入系统（STC集团）的早期版本。识别准确率为93%。研究包括使用键盘输入和语音识别系统填写医疗文件的速度比较。放射科医生填写了电子计算机断层扫描和磁共振成像检查的协议书。一项时间研究表明了，使用键盘输入描述一项检查的平均时间为10分15秒。同时，使用语音识别描述一项检查的平均时间为8分2秒。在现在阶段中，俄语医学术语的识别准确率达到98%。这归功于医学术语词典的形成和对放射科医生反馈意见的分析

⁷ Cloud.mts.ru[互联网]。语音识别技术及其对实业的重要性。访问方式：<https://cloud.mts.ru/cloud-thinking/blog/tehnologiya-raspoznavaniya-rechi/>。

⁸ STC集团[互联网]。语音的合成和识别，录音和分析，人脸和语音的识别。访问方式：<http://www.speechpro.ru/>。

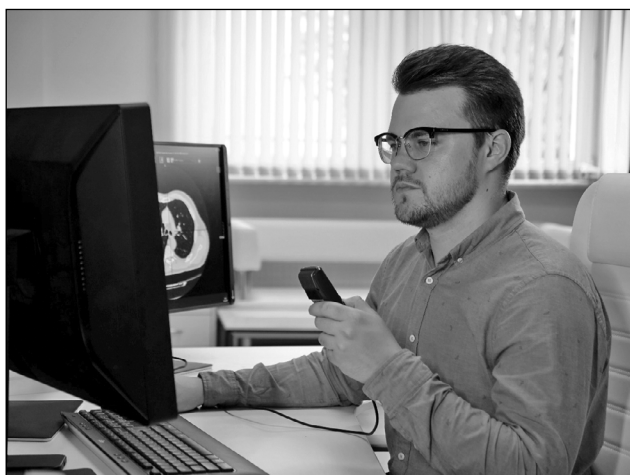


图2. 配有语音识别系统的莫斯科放射诊断参考中心放射科医生的工作场所。填写医疗文件的过程。

成为可能⁹（图2）。该词典是在250万份放射检查协议书的基础上编制的。

2022年对放射科医生进行了一项调查。调查显示了，62.8%的受访者表示使用语音识别系统后工作效率有所提高。大多数医生将放射术语的识别质量评为良好至优秀。这些医生在日常工作中使用语音输入。受访者指出，有一些特殊情况：其他人的语音也会被识别，而且会出现词尾识别错误。外来背景噪声（诊断设备的工作、医务人员与患者或同事的交流）和劣质录音设备会对识别质量产生负面影响。影响是否坚持使用新技术的一个重要因素是医生的年龄及其对创新的兴趣。年轻的专科医生对该技术保持更开放的态度。30-40岁的专家更倾向于在工作中使用语音输入系统。调查结果显示了，在使用语音识别技术的两年内，放射科医生对语音识别技术的态度呈现积极动态变化[40]。

语音识别技术的发展前景

识别准确率的进一步提高将进一步缩短准备电子医疗文件所需的时间。语音识别系统开发人员面临的主要挑战之一是，在录音包含大量噪音或陌生人声音的困难声学条件下，如何确保语音分析的高度准确性。由于俄语的特殊性，用户面临的另一个问题是词尾的识别。因此，俄语语音识别系统最重要的任务是开发语言模型。该模型应该能准确预测和匹配句子中的单词。

语音输入程序与医疗信息系统的整合将允许实现远程填写结构化的电子医疗文件。改进后的系统不仅能识别医生的用语，还能理解所识别的文本应放在医疗文件的哪个部分。这种功能的实现将允许超声医生、病理形态学家、内窥镜医生、外科医生在医疗操作过程中直接填写医疗文件，

而不是事后填写。这将对文件的质量和准备速度产生重大影响。

语音识别技术在医疗文件（包括放射检查协议书）词汇的标准化和统一化方面也具有巨大潜力。迄今为止，还没有一个通用的术语列表来描述放射学中的相同病理状态[41]。即使是同一科室的两位不同放射科医师，在准备检查总结时也可能使用不同的同义词来描述相同的病理检查所见。一系列研究指出，使用结构化和标准化的协议书以及统一的术语，不仅使放射科医生更容易理解和获得必要的信息，而且也使其他专科医师更容易理解和获得必要的信息[42-44]。

迄今为止，人们一直尝试建立一个专门的术语库。该术语库将使电子计算机断层扫描检查中发现的病理变化的描述标准化。该术语库包含120个俄语放射术语及其描述示例[45]。然而，术语库的开发是一项艰巨的任务。它需要大量专家和放射学界对术语进行统一化。

结论

文献综述包括放射诊断中语音识别技术发展的简要历史参考、其演变的详细描述以及基于文献数据的创新优缺点评估。我们特别注意语音识别技术在俄罗斯放射科的应用。一些论文描述了俄语医学术语识别准确率的显著提高。未来，语音识别技术的应用可以缩短医疗文件的准备时间，并更多的时间被用于与患者交流和研究其病史。这种方法为个性化医疗的发展提供新机会。然而，识别词尾和句中词语一致的错误仍然存在。这就需要医生花费更多的时间来纠正这些错误。未来，这些问题将通过应用新人工智能算法得到解决。

一些研究结果表明，放射科医生对语音识别系统保持积极态度。这表现在他们在工作中更常使用语音识别系统。毫无疑问，该技术应该在俄罗斯医疗领域继续发展。国内试点和国外成熟的实施案例证明了其使用的积极态势。进一步提高医疗术语识别的准确性将使语音识别技术在医疗专家中得到更多的支持。

ADDITIONAL INFORMATION

Funding source. This article was prepared by a group of authors as a part of the research and development effort titled "Theoretical and methodological framework for digital transformation in radiology" (USIS No. 123031400118-0) in accordance with the Order No. 1196 dated December 21, 2022 "On approval of state assignments funded by means of allocations from the budget of the city of Moscow to the state budgetary (autonomous) institutions subordinate to the Moscow Health Care Department, for 2023 and the planned period of 2024 and 2025" issued by the Moscow Health Care Department.

⁹ STC集团[互联网]。Voice2Med: 语音填写病程序。访问方式: <https://www.speechpro.ru/product/programmy-dlya-raspoznavaniya-rechi-v-tekst/voice2med>。

Competing interests. The authors declare that they have no competing interests.

Authors' contribution. All authors made a substantial contribution to the conception of the work, acquisition, analysis, interpretation of data for the work, drafting and revising the work, final approval of the

version to be published and agree to be accountable for all aspects of the work. N. Kudryavtsev — concept and design of the paper, text writing, approval of the final version of the paper; K. Bardasova — data collection and analysis; A. Khoruzhaya — data interpretation, text editing.

СПИСОК ЛИТЕРАТУРЫ

1. Вечорко В.И. Распределение рабочего времени на амбулаторном приеме врача-терапевта участкового с медицинской сестрой в поликлинике города Москвы (фотохронометражное наблюдение) // Социальные аспекты здоровья населения. 2016. № 6. С. 4.
2. Каплиева О.В., Марера Л.А., Воробьева Л.П. Хронометраж рабочего времени врачей детского консультативно-диагностического отделения // Дальневосточный медицинский журнал. 2018. № 4. С. 72–76.
3. Ryabchikov I.V., Zagafarov R.R., Mukhina V.V., et al. Distribution of the traumatologist-orthopaedician's working time with outpatients // *Moscow Sur J.* 2018. N 6. P. 38–43. doi: 10.17238/issn2072-3180.2018.6.38-43
4. Kudryavtsev N.D., Sergunova K.A., Ivanova G.V., et al. Evaluation of the effectiveness of the implementation of speech recognition technology for the preparation of radiological protocols // *VIT.* 2020. Vol. 6, N S1. P. 58–64. doi: 10.37690/1811-0193-2020-S1-58-64
5. Blackley S.V., Huynh J., Wang L., et al. Speech recognition for clinical documentation from 1990 to 2018: A systematic review // *J Am Med Inf Association.* 2019. Vol. 26, N 4. P. 324–338. doi: 10.1093/jamia/ocy179
6. Motyer R.E., Liddy S., Torreggiani W.C., Buckley O. Frequency and analysis of non-clinical errors made in radiology reports using the National Integrated Medical Imaging System voice recognition dictation software // *Ir J Med Sci.* 2016. Vol. 185, N 4. P. 921–927. doi: 10.1007/s11845-016-1507-6
7. Hodgson T., Coiera E. Risks and benefits of speech recognition for clinical documentation: A systematic review // *J Am Med Inf Association.* 2016. Vol. 23, N e1. P. e169–e179. doi: 10.1093/jamia/ocv152
8. Itakura F. Minimum prediction residual principle applied to speech recognition // *IEEE Trans Acoust Speech Signal Process.* 1975. Vol. 23, N 1. P. 67–72. doi: 10.1109/TASSP.1975.1162641
9. Leeming W., Porter D., Jackson J.D., et al. Computerized radiologic reporting with voice data-entry // *Radiology.* 1981. Vol. 138, N 3. P. 585–588. doi: 10.1148/radiology.138.3.7465833
10. Simon M., Leeming B.W., Bleich H.L., et al. Computerized radiology reporting using coded language // *Radiology.* 1974. Vol. 113, N 2. P. 343–349. doi: 10.1148/113.2.343
11. Vogel M., Kaisers W., Wassmuth R., Mayatepek E. Analysis of documentation speed using web-based medical speech recognition technology: Randomized controlled trial // *J Med Internet Res.* 2015. Vol. 17, N 11. P. e247. doi: 10.2196/jmir.5072
12. Ramaswamy M.R., Chaljub G., Esch O., et al. Continuous speech recognition in MR imaging reporting // *Am J Roentgenol.* 2000. Vol. 174, N 3. P. 617–622. doi: 10.2214/ajr.174.3.1740617
13. Poder T.G., Fiset J.F., Déry V. Speech recognition for medical dictation: Overview in quebec and systematic review // *J Med Systems.* 2018. Vol. 42, N 5. P. 89. doi: 10.1007/s10916-018-0947-0
14. Sankaranarayanan B., David G., Vishwanath K.R., et al. Would technology obliterate medical transcription? // *Proceedings of the 2017 ACM SIGMIS Conference on Computers and People Research.* New York, NY, USA: ACM, 2017. P. 97–104. doi: 10.1145/3084381.3084414
15. Houston J.D., Rupp F.W. Experience with implementation of a radiology speech recognition system // *J Digital Imaging.* 2000. Vol. 13, N 3. P. 124–128. doi: 10.1007/BF03168385
16. Saxena K., Diamond R., Conant R.F., et al. Provider adoption of speech recognition and its impact on satisfaction, documentation quality, efficiency, and cost in an inpatient EHR // *AMIA Jt Summits Transl Sci Proc.* 2018. Vol. 2017. P. 186–195.
17. Schwartz L.H., Kijewski P., Hertogen H., et al. Voice recognition in radiology reporting // *Am J Roentgenol.* 1997. Vol. 169, N 1. P. 27–29. doi: 10.2214/ajr.169.1.9207496
18. Vogel M., Kaisers W., Wassmuth R., Mayatepek E. Analysis of documentation speed using web-based medical speech recognition technology: Randomized controlled trial // *J Med Int Research.* 2015. Vol. 17, N 11. P. e247. doi: 10.2196/jmir.5072
19. Hammana I., Lepanto L., Poder T., et al. Speech recognition in the radiology department: A systematic review // *Health Inf Manag.* 2015. Vol. 44, N 2. P. 4–10. doi: 10.1177/183335831504400201
20. Mohr D.N., Turner D.W., Pond G.R., et al. Speech recognition as a transcription aid: A randomized comparison with standard transcription // *J Am Med Inf Association.* 2003. Vol. 10, N 1. P. 85–93. doi: 10.1197/jamia.m1130
21. Singh M., Pal T.R. Voice recognition technology implementation in surgical pathology: Advantages and limitations // *Arch Pathol Laboratory Med.* 2011. Vol. 135, N 11. P. 1476–1481. doi: 10.5858/arpa.2010-0714-OA
22. Goss F.R., Blackley S.V., Ortega C.A., et al. A clinician survey of using speech recognition for clinical documentation in the electronic health record // *Int J Med Inf.* 2019. N 130. P. 103938. doi: 10.1016/j.ijmedinf.2019.07.017
23. Blackley S.V., Schubert V.D., Goss F.R., et al. Physician use of speech recognition versus typing in clinical documentation: A controlled observational study // *Int J Med Inform.* 2020. N 141. P. 104178. doi: 10.1016/j.ijmedinf.2020.104178
24. Yang L., Ene I.C., Belaghi R.A., et al. Stakeholders' perspectives on the future of artificial intelligence in radiology: A scoping review // *Eur Radiol.* 2022. Vol. 32, N 3. P. 1477–1495. doi: 10.1007/s00330-021-08214-z
25. European Society of Radiology (ESR). Impact of artificial intelligence on radiology: A EuroAIM survey among members of the European Society of Radiology // *Insights Imaging.* 2019. Vol. 10, N 1. P. 105. doi: 10.1186/s13244-019-0798-3
26. Szymański P., Żelasko P., Morzyet M., et al. WER we are and WER we think we are // *arXiv.* arXiv:2010.03432.2020. doi: 10.48550/arXiv.2010.03432
27. Li J. Recent advances in end-to-end automatic speech recognition // *arXiv.* arXiv:2111.01690.2022. doi: 10.48550/arXiv.2111.01690
28. Juang B.H., Rabiner L.R. Hidden markov models for speech recognition // *Technometrics.* 1991. Vol. 33, N 3. P. 251–272.

29. Graves A., Mohamed A., Hinton G. Speech recognition with deep recurrent neural networks // 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver, BC, Canada: IEEE, 2013. P. 6645–6649. doi: 10.48550/arXiv.1303.5778
30. Deng L., Li X. Machine learning paradigms for speech recognition: An overview // IEEE Trans Audio Speech Lang Process. 2013. Vol. 21, N 5. P. 1060–1089.
31. Казачкин А.Е. Методы распознавания речи, современные речевые технологии // Молодой ученый. 2019. № 39. С. 6–8.
32. Kamath U., Liu J., Whitaker J. Deep learning for NLP and speech recognition. Cham: Springer International Publishing, 2019. 621 p.
33. Wang D., Wang X., Lv S. An overview of end-to-end automatic speech recognition // Symmetry. 2019. Vol. 11, N 8. P. 1018. doi: 10.3390/sym11081018
34. Zhozhikashvili V.A., Farkhadov M.P., Petukhova N.V., Zhozhikashvili A.V. The first voice recognition applications in Russian language for use in The Interactive Information Systems // Speech and Computer. Saint-Petersburg, SPECOM, 2004. P. 304–307.
35. Карпов А.А., Ронжин А.А., Ли И.В. SIRIUS система дикторнезависимого распознавания слитной русской речи // Известия Южного федерального университета. Технические науки. 2005. Т. 54, № 10. С. 44–54.
36. Ирзаев М.Г. Использование голосового ввода информации в медицинских учреждениях для заполнения электронных карт и историй болезней пациентов // Новые технологии и техника в медицине, биологии и экологии: сборник научных трудов. 2013. № 3. С. 149–154.
37. Vazhenina D., Markov K., Karpov A., et al. State-of-the-art speech recognition technologies for Russian language // Proceedings of the 2012 Joint International Conference on Human-Centered Computer Environments. Aizu-Wakamatsu Japan: ACM, 2012. P. 59–63. doi: 10.1145/2160749.2160763
38. Kamvar M., Chelba C. Optimal size, freshness and time-frame for voice search vocabulary // arXiv. arXiv:1210.8436.2012.
39. Kipyatkova I.S., Karpov A.A. An analytical survey of large vocabulary Russian speech recognition systems // SPIIRAS Proceedings. 2014. Vol. 1, N 12. P. 7. doi: 10.15622/sp.12.1
40. Kudryavtsev N.D., Semenov D.S., Kozhikhina D.D., Vladzimirskyy A.V. Speech recognition technology: Results of a survey of radiologists at the Moscow reference center for diagnostic radiology // Healthcare Management. 2022. Vol. 8, N 3. P. 95–104. doi: 10.33029/2411-8621-2022-8-3-95-104
41. Сеницын В.Е., Комарова М.А., Мершина Е.А. Протокол рентгенологического описания: прошлое, настоящее, будущее // Вестник рентгенологии и радиологии. 2014. № 3. С. 35–40.
42. Sobez L.M., Kim S.H., Angstwurm M., et al. Creating high-quality radiology reports in foreign languages through multilingual structured reporting // Eur Radiol. 2019. Vol. 29, N 11. P. 6038–6048. doi: 10.1007/s00330-019-06206-8
43. Ganeshan D., Duong P.A., Probyn L., et al. Structured reporting in radiology // Academic Radiology. 2018. Vol. 25, N 1. P. 66–73. doi: 10.1016/j.acra.2017.08.005
44. Dos Santos P.D., Hempel J.M., Mildemberger P., et al. Structured reporting in clinical routine // Rofo. 2019. Vol. 191, N 01. P. 33–39. doi: 10.1055/a-0636-3851
45. Андрианова М.Г., Кудрявцев Н.Д., Петряйкин А.В. Разработка тезауруса рентгенологических терминов для голосового заполнения протоколов диагностических исследований // Digital Diagnostics. 2022. Т. 3, № S1. С. 21–22. doi: 10.17816/DD105703

REFERENCES

1. Vechorko VI. Distribution of working time at an outpatient appointment of a district therapist with a nurse in a polyclinic in Moscow (photochronometric observation). *Social Aspects Public Health*. 2016;(6):4. (In Russ).
2. Kaplieva OV, Marega LA, Vorobyeva LP. Timekeeping of working hours of doctors of the children's consultative and diagnostic department. *Far Eastern Med J*. 2018;(4):72–76. (In Russ).
3. Ryabchikov IV, Zagafarov RR, Mukhina VV, et al. Distribution of the traumatologist-orthopaedician's working time with outpatients. *Moscow Sur J*. 2018;(6):38–43. (In Russ). doi: 10.17238/issn2072-3180.2018.6.38-43
4. Kudryavtsev ND, Sergunova KA, Ivanova GV, et al. Evaluation of the effectiveness of the implementation of speech recognition technology for the preparation of radiological protocols. *VIT*. 2020;6(S1):58–64. (In Russ). doi: 10.37690/1811-0193-2020-S1-58-64
5. Blackley SV, Huynh J, Wang L, et al. Speech recognition for clinical documentation from 1990 to 2018: A systematic review. *J Am Med Inf Association*. 2019;26(4):324–338. doi: 10.1093/jamia/ocy179
6. Motyer RE, Liddy S, Torreggiani WC, Buckley O. Frequency and analysis of non-clinical errors made in radiology reports using the National Integrated Medical Imaging System voice recognition dictation software. *Ir J Med Sci*. 2016;185(4):921–927. doi: 10.1007/s11845-016-1507-6
7. Hodgson T, Coiera E. Risks and benefits of speech recognition for clinical documentation: A systematic review. *J Am Med Inf Association*. 2016;23(e1):e169–e179. doi: 10.1093/jamia/ocv152
8. Itakura F. Minimum prediction residual principle applied to speech recognition. *IEEE Trans Acoust Speech Signal Process*. 1975;23(1):67–72. doi: 10.1109/TASSP.1975.1162641
9. Leeming W, Porter D, Jackson JD, et al. Computerized radiologic reporting with voice data-entry. *Radiology*. 1981;138(3):585–588. doi: 10.1148/radiology.138.3.7465833
10. Simon M, Leeming BW, Bleich HL, et al. Computerized radiology reporting using coded language. *Radiology*. 1974;113(2):343–349. doi: 10.1148/113.2.343
11. Vogel M, Kaisers W, Wassmuth R, Mayatepek E. Analysis of documentation speed using web-based medical speech recognition technology: Randomized controlled trial. *J Med Internet Res*. 2015;17(11):e247. doi: 10.2196/jmir.5072
12. Ramaswamy MR, Chaljub G, Esch O, et al. Continuous speech recognition in MR imaging reporting. *Am J Roentgenol*. 2000;174(3):617–622. doi: 10.2214/ajr.174.3.1740617

13. Poder TG, Fisetle JF, Déry V. Speech recognition for medical dictation: Overview in quebec and systematic review. *J Med Systems*. 2018;42(5):89. doi: 10.1007/s10916-018-0947-0
14. Sankaranarayanan B, David G, Vishwanath KR, et al. Would technology obliterate medical transcription? In: Proceedings of the 2017 ACM SIGMIS Conference on Computers and People Research. New York, NY, USA: ACM; 2017. P. 97–104. doi: 10.1145/3084381.3084414
15. Houston JD, Rupp FW. Experience with implementation of a radiology speech recognition system. *J Digital Imaging*. 2000;13(3):124–128. doi: 10.1007/BF03168385
16. Saxena K, Diamond R, Conant RF, et al. Provider adoption of speech recognition and its impact on satisfaction, documentation quality, efficiency, and cost in an inpatient EHR. *AMIA Jt Summits Transl Sci Proc*. 2018;2017:186–195.
17. Schwartz LH, Kijewski P, Hertogen H, et al. Voice recognition in radiology reporting. *Am J Roentgenol*. 1997;169(1):27–29. doi: 10.2214/ajr.169.1.9207496
18. Vogel M, Kaisers W, Wassmuth R, Mayatepek E. Analysis of documentation speed using web-based medical speech recognition technology: Randomized controlled trial. *J Med Int Research*. 2015;17(11):e247. doi: 10.2196/jmir.5072
19. Hammana I, Lepanto L, Poder T, et al. Speech recognition in the radiology department: A systematic review. *Health Inf Manag*. 2015;44(2):4–10. doi: 10.1177/183335831504400201
20. Mohr DN, Turner DW, Pond GR, et al. speech recognition as a transcription aid: A randomized comparison with standard transcription. *J Am Med Inf Association*. 2003;10(1):85–93. doi: 10.1197/jamia.m1130
21. Singh M, Pal TR. Voice recognition technology implementation in surgical pathology: Advantages and limitations. *Arch Pathol Laboratory Med*. 2011;135(11):1476–1481. doi: 10.5858/arpa.2010-0714-OA
22. Goss FR, Blackley SV, Ortega CA, et al. A clinician survey of using speech recognition for clinical documentation in the electronic health record. *Int J Med Inf*. 2019;(130):103938. doi: 10.1016/j.ijmedinf.2019.07.017
23. Blackley SV, Schubert VD, Goss FR, et al. Physician use of speech recognition versus typing in clinical documentation: A controlled observational study. *Int J Med Inform*. 2020;(141):104178. doi: 10.1016/j.ijmedinf.2020.104178
24. Yang L, Ene IC, Belaghi RA, et al. Stakeholders' perspectives on the future of artificial intelligence in radiology: A scoping review. *Eur Radiol*. 2022;32(3):1477–1495. doi: 10.1007/s00330-021-08214-z
25. European Society of Radiology (ESR). Impact of artificial intelligence on radiology: A EuroAIM survey among members of the European Society of Radiology. *Insights Imaging*. 2019;10(1):105. doi: 10.1186/s13244-019-0798-3
26. Szymański P, Żelasko P, Morzyet M, et al. WER we are and WER we think we are. *arXiv*. arXiv:2010.03432.2020. doi: 10.48550/arXiv.2010.03432
27. Li J. Recent advances in end-to-end automatic speech recognition. *arXiv*. arXiv:2111.01690.2022. doi: 10.48550/arXiv.2111.01690
28. Juang BH, Rabiner LR. Hidden markov models for speech recognition. *Technometrics*. 1991;33(3):251–272.
29. Graves A, Mohamed A, Hinton G. Speech recognition with deep recurrent neural networks In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. Vancouver, BC, Canada: IEEE; 2013. P. 6645–6649. doi: 10.48550/arXiv.1303.5778
30. Deng L, Li X. Machine learning paradigms for speech recognition: An overview. *IEEE Trans Audio Speech Lang Process*. 2013;21(5):1060–1089.
31. Kazachkin AE. Speech recognition methods, modern speech technologies. *Young Scientist*. 2019;(39):6–8. (In Russ).
32. Kamath U, Liu J, Whitaker J. Deep learning for NLP and speech recognition. Cham: Springer International Publishing; 2019. 621 p.
33. Wang D, Wang X, Lv S. An overview of end-to-end automatic speech recognition. *Symmetry*. 2019;11(8):1018. doi: 10.3390/sym11081018
34. Zhzhikashvili VA, Farkhadov MP, Petukhova NV, Zhzhikashvili AV. The first voice recognition applications in Russian language for use in The Interactive Information Systems. In: Speech and Computer. Saint-Petersburg, SPECOM; 2004. P. 304–307. (In Russ).
35. Karpov AA, Ronzhin AA, Li IV. SIRIUS system of dictation-independent recognition of the merged Russian speech. *Izvestia Southern Federal University. Technical Sci*. 2005;54(10):44–54. (In Russ).
36. Irzaev MG. The use of voice input of information in medical institutions to fill in electronic charts and patient medical histories. *New technologies and techniques in medicine, biology and ecology: Collection of scientific papers*. 2013;(3):149–154. (In Russ).
37. Vazhenina D, Markov K, Karpov A, et al. State-of-the-art speech recognition technologies for Russian language. In: Proceedings of the 2012 Joint International Conference on Human-Centered Computer Environments. Aizu-Wakamatsu Japan: ACM; 2012. P. 59–63. doi: 10.1145/2160749.2160763
38. Kamvar M, Chelba C. Optimal size, freshness and time-frame for voice search vocabulary. *arXiv*. arXiv:1210.8436.2012.
39. Kipyatkova IS, Karpov AA. An analytical survey of large vocabulary Russian speech recognition systems. *SPIIRAS Proceedings*. 2014;1(12):7. (In Russ). doi: 10.15622/sp.12.1
40. Kudryavtsev ND, Semenov DS, Kozhikhina DD, Vladzymyrskiy AV. Speech recognition technology: Results of a survey of radiologists at the Moscow reference center for diagnostic radiology. *Healthcare Management*. 2022;8(3):95–104. (In Russ). doi: 10.33029/2411-8621-2022-8-3-95-104
41. Sinityn VE, Komarova MA, Mershina EA. Protocol of radiological description: Past, present, future. *Bulletin Radiology Radiology*. 2014;(3):35–40. (In Russ).
42. Sobez LM, Kim SH, Angstwurm M, et al. Creating high-quality radiology reports in foreign languages through multilingual structured reporting. *Eur Radiol*. 2019;29(11):6038–6048. doi: 10.1007/s00330-019-06206-8
43. Ganeshan D, Duong PA, Probyn L, et al. Structured reporting in radiology. *Academic Radiology*. 2018;25(1):66–73. doi: 10.1016/j.acra.2017.08.005
44. Dos Santos PD, Hempel JM, Mildenerger P, et al. Structured reporting in clinical routine. *Rofo*. 2019. Vol. 191, N 01. P. 33–39. doi: 10.1055/a-0636-3851
45. Andrianova MG, Kudryavtsev ND, Petryaykin AV. Development of a thesaurus of radiological terms for voice filling of diagnostic research protocols. *Digital Diagnostics*. 2022;3(S1):21–22. (In Russ). doi: 10.17816/DD105703

AUTHORS' INFO

*** Nikita D. Kudryavtsev;**

address: 24/1 Petrovka street, 127051 Moscow, Russia;

ORCID: 0000-0003-4203-0630;

eLibrary SPIN: 1125-8637;

e-mail: KudryavtsevND@zdrav.mos.ru

Kristina A. Bardasova;

ORCID: 0009-0002-4310-1357;

eLibrary SPIN: 1156-7627;

e-mail: bardasovakris@mail.ru

Anna N. Khoruzhaya;

ORCID: 0000-0003-4857-5404;

eLibrary SPIN: 7948-6427;

e-mail: KhoruzhayaAN@zdrav.mos.ru

ОБ АВТОРАХ

*** Кудрявцев Никита Дмитриевич;**

адрес: Россия, 127051, Москва, ул. Петровка, д. 24, стр. 1;

ORCID: 0000-0003-4203-0630;

eLibrary SPIN: 1125-8637;

e-mail: KudryavtsevND@zdrav.mos.ru

Бардасова Кристина Алексеевна;

ORCID: 0009-0002-4310-1357;

eLibrary SPIN: 1156-7627;

e-mail: bardasovakris@mail.ru

Хоружая Анна Николаевна;

ORCID: 0000-0003-4857-5404;

eLibrary SPIN: 7948-6427;

e-mail: KhoruzhayaAN@zdrav.mos.ru

* Corresponding author / Автор, ответственный за переписку