S13 (2)

北京鑫海锦江大酒店大连厅

第二届医工韶光创新分论坛
- 医学人工智能技术新进展与应用

The Second Medical Shaoguang Innovation Forum - New Progress and Application of Medical AI

主席：孔德兴

主席：董　骧

2020 年 11 月 22 日　星期日　8:30-12:15		
时间	**演讲者姓名和单位**	**演讲题目**
8:30-8:35	领导致辞	
基础技术研究		
8:35-8:55	**吴　丹** 浙江大学生物医学工程与仪器科学学院	胎儿与婴儿大脑磁共振及人工智能算法的应用
8:55-9:15	**张光磊** 北京航空航天大学	基于人工智能技术的生物医学成像方法研究
9:15-9:35	**李世娟** 北京大学	虚拟现实技术在医疗中应用的理论与实践
9:35-9:55	**许　燕** 微软亚洲研究院 / 北京航空航天大学	弱监督和无监督学习在医学及病理图像分析上的探索
9:55-10:15	**郭志昌** 哈尔滨工业大学	非线性扩散方程和深度学习对超声图像斑块噪声的移除
应用场景示范		
10:15-10:35	**倪广健** 天津大学	智能医学提升听觉障碍人群生活质量
10:35-10:55	**朱　磊** 迈瑞医疗超声影像系统开发部	超声引导肝癌消融规划和术后即时评估解决方案
10:55-11:15	**钱天翼** 腾讯觅影	医疗影像国家新一代人工智能开创新平台建设进展
11:15-11:35	**许言午** 百度智慧医疗事业部	循证 AI 赋能基层医疗
11:35-11:55	**甄　浩** 北京元影科技有限公司	高标准医疗数据集建设的思路、挑战和实践
11:55-12:15	**徐　明** 北京天明创新数据科技有限公司	尘肺病人工智能标准研究—数据和标注质量控制的专家共识

主席：孔德兴

Email:dkong@zju.edu.cn

浙江大学求是特聘教授、博士生导师；浙江大学应用数学研究所所长，浙江大学理学部图像处理研发中心主任，浙江大学附属第一医院双聘教授；国家卫健委《国家医学图像数据库》项目领导小组副组长；国家药监局医疗器械审评中心 - 中国生物医学工程 AI 平台技术专家组组长；中国医学装备人工智能联盟专家委员会成员；兼任：中国人民解放军总医院（北京 301 医院）、中国人民解放军医学院客座教授、中国人民解放军国防科技大学客座教授、英国 Northumbria University 客座教授等职。

主席：董　骧

Email:harrydx@hotmail.com

2002 年毕业于四川大学，获得生物医学工程硕士学位，长期从事生物力学和生物材料基础及应用研究工作。2009 年加入纳通医疗集团，作为项目负责人和主要研发人员研发的人工膝关节、人工髋关节、创伤及脊柱内固定系统 10 余个产品已成功应用于临床，获得数项发明专利。作为主要研究人员参与过多项国家 863 及科技支撑计划课题。在骨科、齿科内植物的优化设计和功能评价方面有丰富产品及技术开发经验。近年来，组织和参与了多项骨科手术用智能医疗器械和骨科智能康复医疗器械的研发和产业化，具有丰富的产业转化经验。

吴　丹

Email:danwu.bme@zju.edu.cn

浙江大学生物医学工程与仪器科学学院研究员，国家“青年千人”。美国约翰霍普金斯大学生物医学工程系博士，曾任约翰霍普金斯大学助理教授。研究方向包括弥散磁共振成像序列研发与微结构建模、多模态神经影像分析、胎儿与婴幼儿成像等。近 6 年在影像领域顶级期刊发表了论文约 50 余篇（2/3 以上为一作或通讯），主持国家自然科学基金青年、面上、重点项目子课题、国家科技部重点专项等，曾主持美国国家卫生所 R01、R21、R03 基金项目。入选 MIT 科技评论 35 岁以下科技创新 35 人。

张光磊

Email:guangleizhang@buaa.edu.cn

男，北京航空航天大学“医工百人”计划特聘副研究员，生物医学工程高精尖中心研究员，博士生导师。研究方向：生物医学影像。现任 IEEE member、中国生物医学工程学会会员、生物医学光子学分会会员、医学物理分会青年委员、中国图象图形学学会会员。

李世娟

Email:shijuan.li@pku.edu.cn

北京大学信息管理系副教授。现任中国图书馆学会图书馆学教育委员会委员、阅读推广委员会图书馆与家庭阅读专业委员会委员、中国生物医学工程学会医学人工智能分会青年委员。

许　燕

Email:xuyan@buaa.edu.cn

北京航空航天大学，生物与医学工程学院，长聘副教授，硕士导师。微软亚洲研究院访问研究员。获得国家自然科学基金优秀青年基金，北京市青年英才和微软亚洲研究院铸星计划。中国生物医学工程学院人工智能分会青年委员；中国睡眠协会老年分会青年委员。在 IEEE TMI, MIA, TBME, JAMIA 等本领域权威期刊在内的重要期刊和 CVPR, ICCV, MICCAI, ICASSP 等本领域重要会议上共发表多篇论文。在各类期刊上发表的总论文数 40 多篇，谷歌学术总引用接近 2000，H 指数 22。以责任人身份负责多项政府、业界资助的项目。资助单位包括国家自然科学基金，国家重点研发计划，北京市自然科学基金，深圳市自然科学基金以及微软亚洲研究院的项目等。研究成果已经被应用在多家三甲医院及企业产品中。

郭志昌

Email:mathgzc@gmail.com

男，1982 年 07 月出生，准聘副教授，哈尔滨工业大学数学学院计算数学系副主任，中国生物医学工程学会医学人工智能分会青年委员，主要从事偏微分方程图像处理和深度学习等方面的研究。在图像去噪、图像恢复、图像分割、图像压缩等问题有一系列原创成果。已发表论文 20 余篇，在 SIAM Journal on Imaging Sciences，Journal of Nonlinear Science，Neurocomputing，IEEE Transactions on Image Processing 等高水平期刊上有一系列研究成果。主持国家青年基金 1 项，作为主要参与人参与面上基金 2 项，主持省基金 1 项，主持博士后新教师基金 1 项，主持博士后面上基金 1 项。现阶段主要研究方向为，分数阶方程的数值理论和在图像恢复中的建模，深度学习卷积神经网络的部分解释，基于 PDE 和深度学习卷积神经网络的融合模型。

朱　磊

Email:zhulei@mindray.com

高级工程师，迈瑞医疗超声影像系统研发总监。先后主持多项超声成像系统的开发，其中的 Resona7 系列产品性能达到国际主流水平，实现了超高端超声诊断仪的进口替代，同时还主持了公司多项超声成像系统的预研与平台开发工作，达到国际主流及领先水平。

许言午

Email:xuyanwu@baidu.com

百度智慧医疗科学家，WHO 数字健康专委会委员，中国科学院慈溪医工所客聘研究员，新加坡眼科研究所客聘研究员，IEEE 高级会员，全国智能眼科学组常委。他持续从事计算机视觉、机器学习理论及其应用研究，共发表了 110 余篇国际期刊及会议论文，谷歌引用 2700 余次。

徐　明

Email:Xmas@maxdiag.com

中国生物医学工程学会 医学人工智能分会青年委员 职业健康与职业病学组 组长

清华大学 EMBA，北京天明创新数据科技有限公司创始人 CEO。曾任惠普等跨国企业高管，先后从事互联网、电信运营商、通讯、网络、计算机设备、软件等多个行业。对政府、医疗、科技等多个产业有深入的认识与研究。

倪广健

Email:niguangjian@tju.edu.cn

博士研究生导师。天津大学医学部主任助理、智能医学工程教育部工程研究中心副主任、天津市智能人机交互康复工程技术中心主任。2012 博士毕业于英国南安普顿大学。2012 年至 2017 年在南安普顿大学分别担任 Research Fellow、Senior Research Fellow。主要研究领域为听觉原理、听觉假体、听觉康复客观评估。先后主持或参与多项欧盟、英国 MRC、EPSRC 及国家自然科学基金项目。在 Hearing Research、Journal of the Royal Society Interface、The Journal of the Acoustical Society of America 等高水平期刊、学术会议发表论文 60 余篇，先后获得欧盟 Marie Curie Fellowship、The Spoendlin Award、The Sir James Lighthill Award 等奖励。

钱天翼

Email:

腾讯觅影总经理

腾讯医疗技术委员会主任，医疗影像国家新一代人工智能开放创新平台负责人，广东省智慧医疗平台负责人。全国人工智能医疗器械标准化技术归口单位专家组成员，IEEE 医疗 AI 标准组专家。在国内多所大学及机构担任研究生校外导师及特聘研究员。在国内外知名学术期刊以及会议发表论文 150 余篇。

甄　浩

Email: zhenhao@metaimage.cn

现任：北京元影科技有限公司 创始人

曾任：上海复星智健信息科技有限公司创始人、中电数据服务有限公司 首席数据官、阿里巴巴集团 - 阿里云 数据生态总监、IBM Watson 早期核心成员。

胎儿与婴儿大脑磁共振及人工智能算法的应用

吴　丹
浙江大学
Email:danwu.bme@zju.edu.cn

胎儿与婴儿大脑的磁共振成像是研究大脑发育与发育过程中神经系统疾病的重要手段。然后胎儿与婴儿由于剧烈与不规则运动，成像与图像处理的难度极大，因而国内外该方向的研究仍处于初期。本次报告将针对胎儿大脑的成像与图像分析技术，包括高分辨率成像技术、胎儿大脑发育图谱的建立、基于深度学习的胎儿大脑运动伪影校正和脑龄估计方法；婴儿大脑的成像与图像分析技术，包括婴儿大脑成像序列优化、婴儿大脑发育的非对称性分析、大脑发育的统计分析模型研究等方向介绍领域内的前沿发展和本团队的近期研发成果。

Fetal and Neonatal MRI and the Application of Learning-based Techniques

Magnetic resonance imaging (MRI) of fetal and infant brains is an important tool to study the early brain development and neurological disorders during development. However, because of the uncontrolled movement of the fetuses and infants, the imaging and image processing is extremely difficult, and therefore, research in this direction is still quite preliminary. This presentation will focus on 1) fetal brain imaging and image analysis technologies, including high-resolution 3D image acquisition and reconstruction, fetal brain atlas generation, deep learning based fetal brain motion correction and brain age estimation method; and 2) infant brain imaging and image analysis techniques, including optimization of infant brain imaging sequence, asymmetric analysis of infant brain development, statistical models of brain development, etc. I will introduce the frontier development in the field and the recent research and development achievements of our team.

基于人工智能技术的生物医学成像方法研究

张光磊
北京航空航天大学
Email:guangleizhang@buaa.edu.cn

近年来，人工智能技术在生物医学成像领域实现了全链条应用，包括采集、重建、后处理等，推动了生物医学成像技术的发展。在本报告中，报告人将介绍人工智能技术应在图像重建和后处理领域的最新研究进展，具体如下：（1）荧光分子断层成像技术面临病态性强、重建速度慢的瓶颈，报告人摒弃了传统数值迭代理论，提出了基于深度学习理论的高精度、超快速重建方法，利用深度网络强大的拟合能力提高了成像精度、利用端到端的高效运算模式提升了成像速度；（2）癌症的分类分型对于临床治疗至关重要，但“金标准”活检技术却存在有创、易受空间异质性影响等问题。报告人进行了医学影像人工智能分析研究，基于 CT 影像实现了无创、整体的肺癌分类分型。

Researches on biomedical imaging method based on artificial intelligence

In recent years, artificial intelligence technology has been widely applied in the field of biomedical imaging, including acquisition, reconstruction, post-processing, etc., which promotes the development of biomedical imaging technology. In this report, the speaker will introduce the latest research progress of artificial intelligence technology in the field of image reconstruction and post-processing. The details are as follows: (1) Fluorescence molecular tomography (FMT) faces the bottleneck of strong ill-conditioness and slow reconstruction speed. Instead of the traditional numerical iteration theory, the speaker proposed a high-precision and ultra-fast reconstruction method based on deep learning theory, and used the powerful fitting ability of deep network to improve the imaging accuracy, the efficient end-to-end computing mode to improve the imaging speed; (2) The subtype classification of cancer is very important for clinical treatment, but the “gold standard” biopsy technology has some problems, e.g., it is invasive and vulnerable to spatial heterogeneity. The speaker carried out medical image analysis based artificial intelligence, and realized the non-invasive and overall classification of lung cancer based on CT images.

虚拟现实技术在医疗中应用的理论与实践

李世娟
北京大学
Email: shijuan.li@pku.edu.cn

“虚拟现实对医疗健康的好处可以用一个词来概括：革命性。”虚拟现实（VR）经常提出这样的宏大的主张，但“真实”的现实是，当前的应用主要是以研究为基础，具有一定的局限性，而且往往远离常规实践。本文介绍了虚拟现实技术的背景，国内外发展现状及其在医疗健康中应用的分类。并综述虚拟现实技术在医疗健康中应用的理论与实践的最新进展。

VR in Healthcare

“The benefits of Virtual Reality to healthcare can be summarized in a single word: revolutionary.” Virtual reality (VR) has often made such grand claims but the ‘true’ reality is that the current state of activity is still primarily research based, extremely localized and often remote from routine practice. This work provides an overview and taxonomy of existing VR health applications, indicating those that are in the research domain and those that are used in practice. And summarized the state of art of applications of VR in healthcare from the perspective of both theory development and implementation in practice.

弱监督和无监督学习在医学及病理图像分析上的探索

许　燕
微软亚洲研究院 / 北京航空航天大学
Email:xuyan@buaa.edu.cn

医学图像配准在临床医学和研究中都有着十分重要的应用，但是由于医学图像的空间结构复杂，图像的信息量较大，给算法的配准带来了一定的困难。虽然目前已经出现了多种可以用于医学图像配准的算法，但是这些算法在细微的组织结构的配准中表现并不理想。我们提出一种弱监督或无监督深度学习的可形变医学图像配准算法，能够同时提取图像的特征并生成光流场，可用于配准细微的组织结构。在 3D 的医学图像和 2D 的病理图像上都获得了很好的准确率。

weakly supervised and unsupervised learning for medical image analysis

3D medical image registration is of great clinical importance. However, supervised learning methods require a large amount of accurately annotated corresponding control points (or morphing), which are very difficult to obtain. Unsupervised learning methods ease the burden of manual annotation by exploiting unlabeled data without supervision. In this article, we propose a new unsupervised learning method using convolutional neural networks under an end-to-end framework, Volume Tweening Network (VTN), for 3D medical image registration. We propose three innovative technical components: (1) An end-to-end cascading scheme that resolves large displacement; (2) An efficient integration of affine registration network; and (3) An additional invertibility loss that encourages backward consistency. Experiments demonstrate that our algorithm is 880x faster (or 3.3x faster without GPU acceleration) than traditional optimization-based methods and achieves state-of-the-art performance in medical image registration.

非线性扩散方程和深度学习对超声图像斑块噪声的移除

郭志昌
哈尔滨工业大学数学学院计算数学系
Email:mathgzc@gmail.com

在这次演讲中，我将介绍一个非线性扩散方程模型和一个基于深度学习模型在超声散斑噪声去除。

散斑噪声污染了超声图像，给进一步的图像分析和临床诊断带来了困难。在本文中，我们从非线性扩散方程理论的角度来讨论这个问题。在考虑图像梯度信息和灰度信息的基础上，建立了基于扩散方程的非线性扩散模型。利用区域指标作为变量指数，可以根据图像灰度水平自适应地控制在裴洛纳 - 马利克扩散和 Charbonnier 扩散之间的扩散类型。此外，我们对所提出的模型进行了理论和数值分析。实验表明，与传统的去斑方法相比，该方法具有更好的散斑抑制和边缘保护效果，特别是在低灰度和低对比度区域。

另一种去噪模型是基于重尾分布的新型结构的深度学习模型。超声散斑噪声服从伽马分布，属于重尾分布。为了解决输入和输出都可能被重尾化的问题，我们提出了一个新的带截断极小化的损失函数。与以往使用平方损失来衡量性能的研究不同，新的损失函数根据绝对损失和平方损失的优点赋予不同的系数。一方面，利用经截断函数作用的绝对损失处理噪声的重尾部分。另一方面，采用平方损失平滑噪声。此外，我们使用 ResNet 网络框架进行模型训练。数值算例表明了该方法的有效性和有效性。实验结果表明，该模型能够很好地去除乘性噪声。

A Nonlinear Diffusion Equation-Based model and A Deep Learning-Based model in Ultrasound Speckle Noise Removal

In this talk, I will introduce a nonlinear diffusion equation model and a deep learning-based model in Ultrasound Speckle Noise Removal.

Ultrasound images are contaminated by speckle noise, which brings difficulties in further image analysis and clinical diagnosis. In this paper, we address this problem in the view of nonlinear diffusion equation theories. We develop a nonlinear diffusion equation-based model by taking into count not only the gradient information of the image, but also the information of the gray levels of the image. By utilizing the region indicator as the variable exponent, we can adaptively control the diffusion type which alternates between the Perona–Malik diffusion and the Charbonnier diffusion according to the image gray levels. Furthermore, we analyze the proposed model with respect to the the oretical and numerical properties. Experiments show that the proposed method achieves much better speckle suppression and edge preservation when compared with the traditional despeckling methods, especially in the low gray level and low-contrast regions.

The other denoising model is a deep learning model, based on a novel architecture with heavy-tailed distributions. Ultrasonic speckle noise obeys gamma distribution, which is a heavy-tailed distribution. To address the challenge that both the input and output could be heavy-tailed, we propose a new loss function with the truncated minimization. Different from previous studies that use the squared loss to measure the performance, the new loss function assigns different coefficients according to the advantages of absolute loss and square loss. On one hand, truncated function is used to absolute loss to deal with the heavy tail part of the noise. On the other hand, the square loss is used to smooth the noise. In addition, we used the ResNet network framework for training. Numerical examples are presented to demonstrate the effectiveness and efficiency of the proposed method. Experimental results show that the proposed model can handle multiplicative noise removal quite well.

智能医学提升听觉障碍人群生活质量

倪广健
天津大学
Email:niguangjian@tju.edu.cn

无论是先天性听力障碍还是后天听觉系统的受损，均是不可逆的过程。听力损失对听障人群的生活带来了严重的负面影响，是一个重要的全球健康问题。人工耳蜗、助听器等设备会为不同程度听损人群提供一定听觉补偿，但是在复杂环境下的听觉感知能力至今无法满足要求。随着人工智能的出现，智能化助听技术为进一步提高听障人群的生活质量带来了希望。本次报告将以基本听觉原理为切入点，介绍听觉感知的过程及听觉损失的原因。随后从“智能 + 医学”的角度，对远程调试、声学信号处理、状态监测等方面涉及的智能专家系统、机器学习、生理信号检测等技术进行介绍。

Intelligent Medicine Helps Improve the Quality of Life for the Hearing Impaired

Both congenital and acquired hearing impairment are irreversible processes. Hearing loss has a serious negative impact on the quality of life for hearing-impaired, which is an important global health problem. Cochlear implants, hearing aids and other devices can provide certain hearing compensation for people with different levels of hearing loss, but the auditory perception ability in complex environment cannot meet the requirements. With the emergence of artificial intelligence, intelligent hearing aid technology has brought hope to further improvement. This talk will begin with an introduction to the process of auditory perception and the causes of hearing loss. Then, from the perspective of “intelligence + medicine”, the intelligent expert system, machine learning, physiological signal detection and other technologies involved in remote care, acoustic signal processing and condition monitoring are discussed.

超声引导肝癌消融规划和术后即时评估解决方案

朱　磊
迈瑞医疗超声影像系统开发部
Email: zhulei@mindray.com

随着介入超声技术的广泛普及，超声引导的肿瘤消融手术逐渐成为微创治疗肝癌的主流方法，然而超声成像因分辨率相对较低、视野小等局限为该微创手术的应用带来诸多挑战。迈瑞医疗针对临床痛点，开发出多模态融合成像技术以及单模态消融规划和评估技术。多模态融合成像技术能够有效将术前 CT/MR 等三维影像与术中二维超声实时精准融合，其中的呼吸补偿功能可实时矫正肝脏因呼吸运动导致的位置偏移。另外，针对术前 CT/MR 影像获取不便、大肿瘤消融布针复杂以及术后评估消融效果不理想的临床情况，单模态消融规划和评估技术能够实现术中二维实时超声与三维超声体数据的快速精准融合，三维布针规划和术前术后三维超声造影数据自动配准功能可针对肝癌提供包含术中规划及术后评估的完整解决方案。

Ultrasound-guided ablation planning and postoperative evaluation solution for liver cancer

Ultrasound-guided ablation therapy for tumors has become one of the principal means for the minimally invasive treatment of liver cancer. However, due to the limitations of relatively low resolution and small field of view, the application of the ultrasound-guided ablation therapy is facing many challenges. Mindray has developed multi-modality fusion imaging technology and mono-modal ablation planning and evaluation technology for clinical pain points. Multi-modality fusion imaging can effectively fuse preoperative CT/MR images and intraoperative 2D ultrasound in real-time. The function of respiratory compensation is able to correct the position change of liver caused by respiratory movement. Moreover, considering the inconvenient acquisition of preoperative CT/MR images, the complexity of needle arrangement for large tumors and unsatisfactory evaluation of postoperative ablation effect, the mono-modal ablation planning and evaluation can realize rapid and accurate fusion of intraoperative 2D real-time ultrasound and 3D ultrasound volume data. The 3D virtual needle placement planning and the auto-registration of preoperative and postoperative 3D contrast-enhanced ultrasound data can provide a complete solution including intraoperative planning and postoperative evaluation for liver cancer.

医疗影像国家新一代人工智能开创新平台建设进展

钱天翼
腾讯觅影

腾讯觅影负责建设的医疗影像国家新一代人工智能开放创新平台，目前已经完成了数据管理、项目管理、标注工具、人工智能开发模块的开发上线。现通过腾讯云向行业开放，为产医研的紧密合作，提供平台工具，赋能医学影像人工智能产业。

National open innovation platform of new generation AI in Medical image

Tencent AIMIS is responsible for the construction of the National Open Innovation Platform of new generation artificial intelligence in medical imaging. At present, it has completed the development of data management, project management, annotation tools and artificial intelligence development modules. Based on the power of Tencent cloud, the platform is now online and open to the industry. Tencent AIMIS providing platform tools for the close collaboration of innovation companies, hospitals and researchers, empower the medical imaging artificial intelligence ecosystem.

循证 AI 赋能基层医疗

许言午
百度智慧医疗事业部
Email:xuyanwu@baidu.com

百度智慧医疗在布局之初，提出了“循证 AI”概念，让 AI 技术与医学更为紧密的结合，既要解决“是什么”的结果性问题，更要善于解决“为什么”的解释性问题。在此基础上，百度智慧医疗沉淀了两个系列产品。CDSS 产品，将医疗自然语言处理、医疗知识计算等 AI 技术与循证医疗知识有机结合，打造了一系列可循证的辅助决策能力。影像筛查产品，通过符合临床的循证逻辑架构和可解释深度学习算法结合，进一步强化基于体征和临床规则的疾病分类和分级能力。现如今，“循证 AI”已经得到了业界的广泛认可，成为了行业热点和驱动力。本次报告将以这两个产品为例，解析百度智慧医疗的愿景——“循证 AI 赋能基层医疗”背后的技术框架和实践成果。

Evidence-based AI empowers primary healthcare

Baidu Intelligent Healthcare came up with the concept of "evidence-based Artificial Intelligence (AI)" as its inception to increasingly integrate AI technologies and medical science. Both "what" and "why" need to be elaborated, representing results and explanations, respectively. On this basis, Baidu Intelligent Healthcare developed two product lines. The Clinical Decision Support System (CDSS) products, which combine AI technology (e.g., natural language processing, medical knowledge computing, and knowledge graph) with evidence-based medicine, build up the abilities to produce reliable and explainable clinical decision recommendations. The intelligent medical image analysis products, which combine evidence-based architecture with interpretable deep learning algorithms, further strengthen the classification and grading of diseases. Today, "evidence-based AI" has been widely recognized by the industry and has become a hot topic and driving force in the field. This report takes these two products as examples to introduce the technical frameworks and practical experience behind the vision of Baidu Intelligent Healthcare —— "evidence-based AI empowers primary healthcare."

高质量医疗数据集的建设思路和实践

甄　浩
北京元影科技有限公司
Email: zhenhao@metaimage.cn

医疗数据是指是患者就医过程中所产生的数据。电子化的医疗数据方便了数据的存储和传输，但当其应用于真实世界研究、人工智能开发、临床辅助决策支持等特定领域时，由于数据标准不统一、数据质量差、关键数据项缺失等问题，往往应用效果不佳。通过技术手段、推倒重来或行政干预等方法解决以上问题，往往代价高又达不到效果。本报告将结合讲者在实际建设高质量医疗数据集的实践，提出一种应用场景和临床指南双驱动的高质量数据集的建设思路，旨在建设低成本、高质量、可持续的符合专业领域应用要求的医疗数据集。

Medical data refers to the data generated during the patient's medical treatment. Electronic medical data facilitates the storage and transmission of data, but when it is applied to specific fields such as real-world research, artificial intelligence development, clinical decision support, etc., due to inconsistent data standards, poor data quality, and missing key data items, etc. The application effect is often poor. Solving the above problems through technical ways, overthrowing or administrative intervention is often costly and ineffective. This Report will combine the speaker's practice in the actual construction of high-quality medical data sets, and propose a construction idea of a high-quality data set driven by application scenarios and clinical guidelines, aiming to build a low-cost, high-quality and sustainable application that meets the requirements of professional fields Medical data set.

尘肺病人工智能标准研究—数据和标注质量控制的专家共识

徐　明
北京天明创新数据科技有限公司
Email: Xmas@maxdiag.com

尘肺病指劳动者在职业活动中接触生产性矿物性粉尘所致的以弥漫性肺间质纤维化为基本病理改变的职业性肺部疾病，是我国最主要及危害最严重的职业病。我国是世界上尘肺病人数量最多和年报告新发病例最多的国家。我国人工智能产业发展迅猛，利用人工智能技术研发用于尘肺病筛查，进而发展成人工智能尘肺病诊断产品符合科技创新的要求。将人工智能应用于尘肺病筛检和诊断，可有效提高职业病诊断和基层职业健康检查人员的读片效率，降低人工阅片误差，有效进行质量控制。研制高性能人工智能尘肺病 DR 阅片系统（国家药品监督管理局三类医疗器械）的技术关键是建立明确的尘肺病人工智能诊断标准，其中的关建技术基础框架是数据集的管理与标注质量控制。通过研究、实践、总结、讨论、完善尘肺病 DR 胸片及相关信息的数据采集、筛选、处理过程，以集成数据标注的思路、方法、质量控制过程来为尘肺病人工智能产品（模型）标准打好基础，以期形成严格、合理、符合医学规律并且技术上可达的产品技术标准规范。为此中国生物医学工程学会医学人工智能分会胸部影像及职业病标准组组织国内公共卫生、职业医学与职业病、呼吸系统疾病以及医学影像、人工智能等各方面专家就如何开展尘肺病胸部 DR 数据标注与质量控制进行了专门的研究和深入的讨论，各方专家就尘肺病 DR 胸片数据数的采集、筛选、质量控制、标注内容、标准方法、标注规则、标注流程以及质量判定达成了共识。

Chinese expect consensus on pneumoconiosis data labeling specifications and quality control (2020 edition)

Pneumoconiosis refers to occupational pulmonary disease with diffuse pulmonary interstitial fibrosis as the basic pathological change caused by workers' exposure to productive mineral dust during professional activities. It is the most important and most harmful occupational disease in China. It is also the country with the largest number of pneumoconiosis patients and the largest number of new cases reported annually. The consensus pointed out that China's artificial intelligence industry is developing rapidly. The use of artificial intelligence technology to develop artificial intelligence products for pneumoconiosis screening and development to pneumoconiosis diagnosis meets the requirements of scientific and technological innovation. Applying artificial intelligence to screening and diagnosis of pneumoconiosis can effectively improve the reading efficiency of occupational disease diagnosis and grass-roots occupational health checkers, reduce manual reading errors, and effectively perform quality control. The key to the development of a high-performance artificial intelligence pneumoconiosis DR imaging system (three types of medical devices of the State Drug Administration) is to establish clear artificial intelligence diagnostic standards for pneumoconiosis. Among them, the basic technical framework of Guan Jian is the management of data sets and the quality control of annotation. Through research, practice, summary, discussion, and improvement of data collection, screening, processing of DR chest radiographs and related information of pneumoconiosis, and ideas, methods, and quality control processes for labeling the data, the standards for pneumoconiosis artificial intelligence products (models) are set Good foundation, with a view to forming strict, reasonable, medical laws and technically reachable product technical standards and specifications. To this end, the Chinese Society of Biomedical Engineering Medical Artificial Intelligence Branch of the Chest Imaging and Occupational Diseases Standards Group organized domestic public health, occupational medicine and occupational diseases, respiratory diseases, medical imaging, artificial intelligence and other experts on how to carry out pneumoconiosis chest DR data annotation and Special research and in-depth discussions were conducted on quality control. Experts from all parties reached consensus on the collection, screening, quality control, labeling content, standard methods, labeling rules, labeling process, and quality judgment of DR chest radiograph data for pneumoconiosis.