

A decorative graphic on the left side of the page features a circular pattern of concentric arcs and dots in shades of blue, green, and grey, resembling a stylized sunburst or a data visualization.

计算决定未来

医疗大数据应用技术研讨班

从工业界诞生的大数据理论、方法和技术引起了社会各界的广泛关注。李克强说：“信息数据‘深藏闺中’是极大浪费”。医院信息系统属于迄今世界现存的企业级信息系统中最复杂的一类。如何使用大数据的理论、方法和技术处理深藏于医院信息系统的医疗大数据，提高医疗机构的运行效率，帮助解决老百姓“看病难、看病贵”的问题，是相关从业人员普遍关注的热点。

广东省高性能计算学会邀请上海交通大学高性能计算中心的科学计算总监 Simon See 教授、中组部“青年千人”陆遥教授、曙光信息产业股份有限公司总工程师宋怀明博士、“国家优青”郑伟诗教授分别从工业界和学术界的不同视角介绍医疗大数据应用技术。

一、人数及注意事项

拟限 50 人，采取先报先得、额满即止的原则确定报名学员的培训资格。
学员自带笔记本电脑。

二、时间、地点

时间：2016 年 7 月 16-17 日（周六、日）

地点：广州市海珠区新港西路 135 号中山大学 336 栋旧数学楼 104 室



三、组织单位

主办：广东省高性能计算学会 协办：广东省计算数学学会

四、课程设置

时间	课程名称	主讲人
16 日上午	医疗大数据特点及应用实践	宋怀明
16 日下午	Deep Learning for Modern Healthcare	Simon See
17 日上午	放射组学及医学大数据分析	陆 遥
17 日下午	大规模数据在线学习处理	郑伟诗

五、费用标准

在职人员 1800 元、学生 1500 元，前 20 名（以报名先后为准）按标准的 80%收费。

六、交通食宿安排

参加研讨班的人员交通、住宿费用自理，餐饮由学会统一安排。承办单位可为有需要的学员协助联系住宿。

七、报名方式和录取

1、报名学员需于 2016 年 6 月 30 日之前填写“医疗大数据应用技术研讨班”报名表（请从资料下载栏目下载），并将表格电子版发送至报名邮箱：hpcoffice@gdhpcs.org，邮件主题为：姓名_单位_医疗大数据应用技术研讨班报名表。

2、采取先报先得、额满即止的原则确定申请人参加资格。通过资格确定的申请人，将会收到来自主办单位的确认邮件。

八、奖励和证书

参加培训的学员将获得由广东省高性能计算学会、广东省计算数学学会联合颁发的结业证书。此外，主办单位将对学员在学习期间的表现进行评定，表现优异者将获得现金奖金和奖励证书。

九、联系人、联系电话

季老师 020 -84111426、84110128

十、汇款信息

汇款户名：广东省高性能计算学会；

帐 号：4400 1430 0460 5250 0994；

开 户 行：中国建设银行广州中山大学支行。

十一、主讲老师简介

Simon See 教授简介

Simon See, graduated from University of Salford (UK) with a Ph.D. in electrical engineering and numerical analysis in 1993. Prof. See is currently the High Performance Computing Technology Director and Chief Solution Architect for Nvidia Inc, Asia, a Professor and Chief Scientific Computing Officer in Shanghai Jiao Tong University, Chief Scientific Computing Advisor for BGI (China) and has a position in Nanyang Technological University (Singapore) and King-Mong Kung University of Technology (Thailand). His research interests are in the area of High Performance Computing, Artificial Intelligence, machine learning, computational science, Applied Mathematics and simulation methodology. He has published over 100 papers in these areas and has won various awards.



陆遥教授简介

陆遥，中山大学教授，博士生导师。现为广东省计算数学学会秘书长、广东省计算科学重点实验室副主任、广东联通“智慧城市”专家组成员和广东生物超算与健康产业大数据应用联盟专家委员会成员。长期从事医学影像、医学图像分析、机器学习和数据挖掘领域的研究工作。曾与美国通用电器(GE)全球研发中心合作开发了新一代乳腺层析成像设备和乳腺癌计算机辅助诊断系统；现与美国哈佛大学医学院及密西根大学医学院合作进行乳腺癌计算机辅助诊断算法研究。主持过世界著名的乳腺癌基金会 Susan Komen 基金会研究基金、国家自然科学基金项目、广东省前沿与关键技术创新重大专项基金各一项，2013 年入选中组部“青年千人计划”项目，及中山大学“百人计划”。在医学影像及图像分析领域国际一流期刊 *Applied and Computational Harmonic Analysis*, *Inverse Problems*, *Medical Physics*, *Physics in Medicine and Biology*, *IEEE Transactions on Circuits and System I* 等上发表 SCI/EI 论文 16 篇，以及国际主要学术会议论文 16 篇。主办过三次科学计算与图像分析领域国际会议，为 16 个国际著名学术期刊审稿人。



宋怀明总工程师简介

宋怀明，曙光信息产业股份有限公司大数据总工程师。主要研究方向为海量数据存储和处理技术，2009 年博士毕业于中国科学院计算技术研究所，师从中科院计算技术研究所所长、中国计算机学会高性能计算专业委员会主任孙凝辉教授。2009-2011 年间，在美国伊利诺伊理工大学和美国阿贡国家实验室从事博士后研究，在著名国际会议 SC、HPDC、CCGRID、CLUSTER、ISPA、PDSW 等以第一作者发表多篇论文，申请专利 34 项。2011 年 8 月回国即加入中科曙光公司，担任大数据研发经理和技术方向负责人，从事大数据产品的研发工作。



此外，作为项目负责人，承担了科技部 863 项目“基于内存计算的实时大数据处理关键技术和系统”，发改委信息安全专项“大数据安全管理平台”，北京市十百千工程“通用大数据存储和处理一体机”等重大科研项目的研发工作。现为科技部国家重点研发计划项目评审专家、广东生物超算与健康产业大数据应用联盟专家组成员。

郑伟诗教授简介

郑伟诗，中山大学教授，国家自然科学基金优秀青年基金获得者、英国皇家学会牛顿高级学者基金获得者、广东省杰出青年科学基金获得者、微软亚洲研究院青年学者铸星计划入选者。主要面向智能视频监控，展开视频图像信息与信号的处理研究。目前的主要研究方向是视频监控下的行人身份识别与行为理解。其中，在国内外较早和持续开展行人再标识的研究，并发表一系列以跨视域度量学习为主线的研究工作。已发表/接受含 5 篇 PAMI 在内的 80 余篇主要学术论文，其中 40 余篇发表在图像识别和模式分类 IEEE TPAMI、IEEE TIP、IEEE TNN、PR、IEEE TCSVT、IEEE TSMC-B 等国际主流权威期刊和 ICCV、CVPR、IJCAI 等计算机学会推荐 A 类国际学术会议。在所发表的论文中，曾发表 ESI 高被引论文 3 篇。近 5 年来，与国内外同行一道，在计算机学会推荐 A 类国际学术会议 ICCV 和 CVPR 上以及其它著名国际学术会议期间做 Tutorial。曾担任 IEEE AVSS 的 Area Chair，担任 2012 年、2015 年和 2016 年全国生物特征识别学术会议的联合程序委员会主席。



十二、课程介绍

（一）Deep Learning for Modern Healthcare（授课人：Simon See 教授）

在世界最发达国家美国，人口老龄化与过度医疗和重复医疗使医保支出成为财政的沉重负担！为此奥巴马政府提出并大力推进“精准医疗”计划希望能够彻底变革医疗卫生系统，提高医疗机构的运行效率，实现病人的个性化治疗，控制医疗费用支出。医疗信息系统是医疗卫生系统的基础组成部分之一。上海交通大学科学计算总监施忠伟教授，将从工业界的视角介绍世界最高效率的深度学习系统及其处理医疗大数据的工程实践和在医学影像学研究中的成功案例。通过本课程的学习您可以快速使用深度学习高效的开展有关工程实践和课题研究。

（二）放射组学及医学大数据分析（授课人：陆遥教授）

医学影像能够直观地提供患者功能性、结构性的信息，通过医学影像大数据分析，在大的病人数据集合里找出疾病的共性，建立肿瘤的具有可表征性和可预测性的分析模型，从而利用大数据分析实现医学数据的充分挖掘和机器学习，为临床诊断和治疗提供辅助性的证据。

放射组学从大量临床 CT、PET 和 MRI 等影像数据中提取高代表性定量特征，并对

其进行分析，从而对病例数据进行病变诊断、分级、预后和预测评估等工作。重要的是，所有数据均采自标准治疗的临床数据，从而构建一个巨大的潜在病例池。在此基础上，将数据特征和临床数据特征联系起来，建立临床上可应用的数学模型。

本课程深入的探讨放射组学及医学影像大数据分析。医学大数据作为“智慧医疗”的一个重要课题，随着临床医学和计算机科学的发展，对医学数据分析提出了新的需求。我们将介绍医学影像数据获取和重建，图像分割和绘制，特征提取和定量分析，数据库建立和数据分享以及信息学分析等相关的医学影像大数据分析的技术，并探讨医学成像在指导治疗、监测疾病的发生发展以及治疗后反应等方面的潜力。

（三） 医疗大数据特点及应用实践（授课人：宋怀明总工程师）

随着大数据应用技术的飞速发展，大数据技术已深入到各行业，推动了行业数据的汇聚和融合、提高了信息的综合利用水平。在医疗信息化领域，经过多年的独立建设和发展，医疗信息系统已形成数据种类繁多、结构复杂、访问多样化等应用特点，多种类别的医疗数据相对孤立的存在于不同业务系统，对数据的融合和综合利用提出了新的挑战。大数据技术的发展为医疗行业的数据融合和业务融合提供了很好的契机。加速对多源、异构以及不同类型的医疗数据的采集、存储组织、权限管理、综合检索、分析挖掘、全景展示，是医疗大数据系统建设的核心技术。讲座从医疗数据的结构特点和访问特点出发，结合实际应用需求和建设经验，讲述医疗大数据系统建设中的数据采集、清洗、存储、处理、展示等方法，探讨医疗大数据的应用和实践。

（四） 大规模数据在线学习处理（授课人：郑伟诗教授）

医疗互联网和智能终端的广泛推广，使得人们在医疗活动中产生的数据速度越来越快，如医疗社交网络产生大量文本、图像数据，对计算机的计算能力提出了更高的要求。在线学习是处理以顺序到来的（流式）数据的一类重要算法，可以很大程度减少在大规模数据下的计算负荷。特别地，一遍式在线学习算法用来预测当前新样本的类标，并且根据当前预测结果和实际结果的差距程度来更新当前模型，其中一遍式是指每个样本仅用一次并且不被存储下来，以减少内存使用量。目前，已有的一遍式在线学习算法都是建立全局的线性模型，并不把数据分布的局部结构考虑进去，而数据分布的局部结构常常蕴含着非线性可分情形下的重要信息。在本次的课程里，我们将介绍在线分类器学习（局部线性学习算法）和在线快速检索模型（在线哈希学习模型）的学习，同时将探讨在线学习的未来发展方向。