
警惕缺乏科学论证的新冠肺炎人工智能误导当前的科学研究

作者：田晓婷 来源：学术头条

本文原地址：<https://www.iikx.com/news/topnews/9527.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

警惕缺乏科学论证的新冠肺炎人工智能误导当前的科学研究

。近年来，人工智能在处理医学影像方面取得了实质性的进展，但新冠疫情大爆发期间，似乎出现了这样一种危险的趋势：一些人仓促地使用存在瑕疵的、有问题的数据来训练针对 COVID-19 的人工智能解决方案，这一做法不仅对患者和医生没有任何帮助，而且还会损害人工智能的声誉。

随着新冠肺炎(COVID-19)疫情的全球爆发，围绕着“病毒从哪来”“病毒如何传播”“中间宿主是什么”“疫苗在哪儿”“如何快速检测患者”等问题，生物医学、计算机科学等领域的研究人员开展了广泛的研究，尤其是人工智能(AI)在助力新冠肺炎患者快速检测方面显得备受关注。

我们不能因为封锁隔离而放弃基本的科学原则，人工智能既不是呼吸机，也不是疫苗，更不是药丸。值得注意的是，根据一些大多数由没有临床经验或不具备医学能力的工程师和计算机科学家创建的医学影像小数据集，有些网站甚至提供了关于如何从 X 射线扫描中高精度地检测出 COVID-19 的建议，甚至还有一些未经同行评议的论文，用“COVID-Net”这样的夸张名字为他们的解决方案进行“洗礼”。这类论文往往未经证实，也未经放射科医生的实验指导，甚至在数据集的创建者还没来得及对其收集数据的过程进行足够解释的情况下，这些研究就匆匆公之于众了。

近日，滑铁卢大学工程系教授 Hamid Tizhoosh 发出呼吁，**要警惕缺乏科学论证的 COVID-19 人工智能解决方案误导当前的科学研究。**

Tizhoosh 表示，AI 技术的发展可以在一定程度上辅助医生进行诊断，但还需理性认识其根本，切不可将之“神话”。

可以肯定地说，我们都对 COVID-19 大流行深感担忧。这种新型冠状病毒已经彻底改变了当前的社会，在全球疫情形势严峻的城市中，人们正在经历着面临压力、被限制、被隔离;正在见证包括科研人员、护士和医生在内的护理人员英勇牺牲的事迹;有些人正在失去亲人;人们面临着经济困难，以及未来几个月可能将面临的巨大不确定性。

在这种情况下，许多人都在思考如何以最快的方式，为这场疫情防控提供一些帮助。当然，人工智能领域的研究人员也不例外。

尽管正在应对一场大流行疾病，但是我们不能摒弃基本的科学原则。数据必须经由医学专家整理，并执行充分、严格的验证程序，特别是在社会面临许多不确定性的时候，任何解决方案或提议付诸实施之前，结果都必须经过同行的审核。

机器学习方法主要依赖于数据，通过对标签数据进行学习，从而对数据进行分类、预测和估计，任何人工智能方法的质量和可靠性直接取决于所采用的的标签数据的质量和可靠性。

在计算机科学中，低质量的输入会产生不可靠的输出，也就是我们所讲的“输入的是垃圾，输出的也是垃圾(Garbage In Garbage Out, GIGO)”。特别是，当我们在处理类似医学影像这样高度复杂的数据形式时，通常需要高度专业化的知识来对数据作出恰当的解释，这一点就变得更加重要。

在人工智能领域，我们完全依赖数据。只要不是金融、医疗卫生、监控等敏感领域，我们通常会

使用各种方法来搜集数据集，小到人工采集样本，大到通过高度复杂的网络爬虫来解析互联网和其他公开资源库的数据。

然而，在医学影像学领域，我们所处理的是一个高度敏感的数据领域，这通常需要一个漫长的过程来整理和访问一组标签图像。毫无疑问，数据整理工作必须在医院内进行，这不仅是因为有专家在场，而且为了遵守隐私规定，需要对图像进行去身份处理。

但有时研究人员会变得不耐烦，会在线期刊上手动收集公开数据源，来创建小数据集。

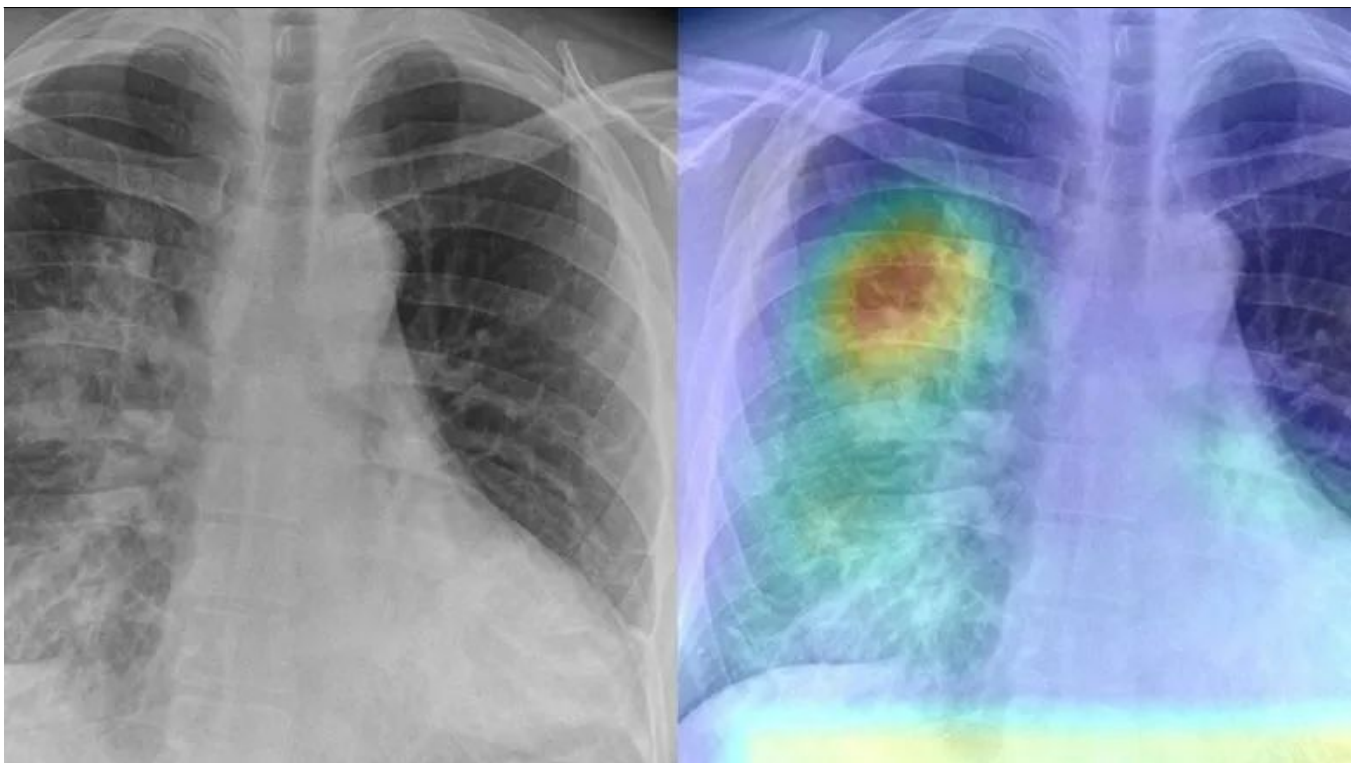
缺少放射科医生的“AI+医疗影像”研究。

Tizhoosh 表示，值得注意的是，大多数时候医学影像领域的小数据集是由没有临床经验或不具备医学能力的工程师和计算机科学家创建的，而不是由医生和医学专家创建的。

Tizhoosh 认为，可以理解的是，当前世界各地的放射科医生都很忙。如果你是一个雄心勃勃的人工智能研究人员，想要在这个时候提供一些帮助，明显目前不是与放射科医生合作的最佳时机。

而一些研究人员已经开始收集和创建自己的数据集，为未来的任务做准备。从网上搜集来的 X 光照片和 CT 图像似乎随处可见，而且随着创作者不断添加图像，网上这些图像也在不断增加。

由于这类数据的可获得性，以及基础人工智能知识和工具存在的普遍性，许多人工智能爱好者和初创公司已经冲动地开始研发从 X 光照片中检测 COVID-19 的解决方案。



有些网站和博客甚至提供了关于如何从 X 射线扫描中高精度地检测出 COVID-19 的建议，还有一些人给出了从 X 光照片中检测 COVID-19 的教程。

甚至还有一些未经同行评议的论文，更进一步，用“COVID-Net”这样的夸张名字为他们的解决方案进行“洗礼”。这种类型的工作通常缺乏许多实验细节，来解释如何通过极少数患者的医学影像图像满足深度神经网络的需求。

这类论文往往未经证实，也未经放射科医生的实验指导，甚至在数据集的创建者还没来得及对其收集数据的过程进行足够解释的情况下，这些研究就匆匆公之于众了。

为了克服数据量小的问题，AI 爱好者和初创公司将他们拥有的为数不多的 COVID-19 图像与其他类似肺炎数据集这样的公共数据集混合在一起。这是一个很聪明的做法，但也会产生一些麻烦的后果，比如 COVID-Nets 会将一到五岁的小儿肺炎当成新冠病例，与成人 COVID-19 患者进行比较。

Tizhoosh 表示，当我们把放射科医生排除在需要专家监督的研究之外时，就会出现这种情况。

保持理性，尊重科学原则。

为什么会有研究人员在使用混杂的医学影像小数据集、没有放射学医生支持、没有验证的情况下，发布错误的人工智能结果呢？他们是真的想帮助 COVID-19 患者吗？

或许，一些初创公司为了争取融资机会，以及研究可能获得更多曝光率，正在误导科研人员做出错误的研究行为。

我们不能因为封锁隔离而放弃基本的科学原则，人工智能既不是呼吸机，也不是疫苗，更不是药丸。在疫情期间的中国武汉、伊朗库姆或意大利贝加莫，精疲力竭的放射科医生不太可能仅仅为了获得一个有缺陷的补充性意见，而去下载我们根据不充分、不恰当的数据训练出来的 Python 代码所形成的结果快速撰写成文的论文。

毋庸置疑，大家都想为全球大流行疾病贡献力量。但是，请等待医院提供真实的数据，经过伦理审批和去鉴定，让我们和放射科医生一起制定关于未来胸部问题的解决方案。

否则，我们可能会给人留下这样的印象，那就是我们在做耸人听闻的研究，更多的是在自我推销，而不是为了患者的健康。

为了了解这种病毒在医学影像中的表现，放射科医生们正在日以继夜地工作，让我们和他们一起努力，向他们学习，真正释放出人工智能在未来对抗病毒感染的潜力。

Hamid R. Tizhoosh 博士自 2001

年起担任滑铁卢大学工程系教授，并领导KIMIA实验室(医学图像分析知识推理实验室)。自 1993 年以来，他的研究领域包括人工智能、计算机视觉和医学成像。他著有两本书，14

个书刊章节，以及 150 多份期刊和会议论文。Tizhoosh 博士拥有丰富的行业经验，并与许多公司合作过。他也是加拿大多伦多向量研究所的委任教授，以及加拿大滑铁卢大学滑铁卢人工智能研究所的成员。

原文资料：

<https://venturebeat.com/2020/04/24/the-surge-of-sensationalist-covid-19-ai-research/>

<https://www.news-medical.net/health/The-Surge-of-Sensationalist-COVID-19-AI-Research.aspx>

更多 科研头条 请访问 <https://www.iikx.com/news/topnews/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](https://www.iikx.com)转发