

报告编号: 202536000Z08X036

# 科技查新报告

项目名称: 基于 Deepseek 大模型的适用于全腔镜外科术中实时手术图文报告生成系统

委托人: 西安交通大学第一附属医院  
彭子洋, 王志博

委托日期: 2025 年 3 月 10 日

查新机构: 教育部科技查新工作站(Z08)

完成日期: 2025 年 3 月 11 日

教育部科技发展中心

二〇一三年制

查新项目 名称	中文：基于 Deepseek 大模型的适用于全腔镜外科术中实时手术图文报告生成系统			
	英文：A Real-Time Image-Text Report Generation System for Total Endoscopic Surgery Based on the Deepseek Large Model			
查新机构	名称	教育部科技查新工作站(Z08)		
	通信地址	陕西省西安市咸宁西路 28 号	邮政编码	710049
	负责人	强自力	电话	029-82668105
	联系人	翟中会	电话	029-82655067
	电子邮箱	transmission@mail.xjtu.edu.cn		
一、查新目的 成果鉴定				
<p><b>二、项目的科学技术要点</b></p> <p>本项目针对术中影像记录的诸多难题，设计了基于 Deepseek 大模型的适用于全腔镜外科术中实时手术图文报告系统。该系统以 Deepseek 大模型为核心，利用其强大的自然语言处理和多模态数据融合能力，实现术中影像数据的高效处理与分析。Deepseek 大模型采用先进的 Transformer 架构和 Mixture-of-Experts (MoE) 架构，能够并行处理大量数据，并通过多阶段训练方式不断提升模型性能。</p> <p>系统可以将全腔镜产生的影像数据通过微创术中成像设备记录并自动传输至医院影像归档和通信系统 PACS 处，再由 PACS 分配至核心交换机，经由网闸传输至医院端边缘计算 AI 加速服务器与具有更强算力的云端。云端利用 Deepseek 大模型对新到达的数据进行计算，并将数据反馈至医院服务器。医院服务器根据云端数据对影像进行可视化标注，并回传至智康微创术中影像记录模块解码，实时在手术室成像设备上显示。</p> <p>本项目采用三位一体的设计，具体分为微创术中影像记录模块、术后图文报告生成模块和术中影像数据分析模块。</p> <p>微创术中影像记录模块基于现有微创术中成像设备与实时音视频传输网络 (RTN)，对术中影像编码并传输至医院核心交换机，同时能够解码经过计算后的影像数据，实时对术中影像进行 AI 分析。该模块使用的记录与传输设置可兼容现有国内主流内镜设备，并内置弹性存储实现影像数据动态传输，实现完整术中医疗影像的稳定传输。视频编码采用 H.265 视频高性能传输标准，不仅极大减少传输的</p>				

视频大小体积，还可以在低于 2Mbps 的网络环境下实现 4K 高清数字图像传送，减轻医疗影像传输与存储负担。传输使用的 RTN 网络可以将网络延迟控制在 300m 以内，并采用实时音视频传输网络以有效解决中间传输以及最后一公里传输两大难题，满足传输术中影像实时分析图像的低延迟传输需求。此外，该模块还具备高精度的图像预处理功能，能够对术中影像进行实时增强和去噪处理，确保影像质量不受传输过程的影响，为后续的 AI 分析提供更清晰、准确的图像基础。

术后图文报告生成模块立足于传统记录模式精确度低、术中视频资料不方便获得等临床痛点，设计了更高效的术后图文文书报告生成系统。该系统使术中主刀医师直接通过按压脚踏板的方式，自主对于术中影像进行截取，不干扰手术操作。在视频记录手术过程的同时，可以通过术中影像识别的方式更加彻底地了解病人情况，给出关键术中影像帧并进行进一步分析。通过知识图谱技术，进行自动术后图文文书报告生成。本产品以知识图谱技术为基础，开发出针对于术后文书辅助决策系统（CDSS），可以对术后文书数据进行挖掘和分析，为医生提供医学实践参考。同时，该模块还引入了自然语言处理技术，能够自动识别和提取术中语音记录中的关键信息，并将其转化为结构化的文本内容，与术中影像数据相结合，生成更加全面、详细的术后报告。此外，系统还具备智能校对功能，能够对生成的报告进行语法检查和医学术语准确性校验，确保报告的质量和专业化。

术中影像数据分析模块基于已有微创术中影像记录与传输装置，对术中微创影像进行实时处理与分析。本项目采用低延迟、高精度的语义分割技术处理术中影像视频，并实时回传至手术室屏幕，协助术中临床医师决策，提升术中外科医生手术决策准确度。为了在术中实时分析低延迟的基础上保证算法的先进性与准确性，本项目设计了边缘计算 AI 加速服务器。以低能耗分布式云端训练系统、高效的分布式云端 AI 训练系统和新型 Yun-Transform 语义分割模型对算法模型进行改良。同时，该模块还集成了多模态影像融合技术，能够将术中影像与其他医学影像数据（如 CT、MRI 等）进行融合分析，为医生提供更全面的患者病情信息，辅助制定更加精准的手术方案。此外，系统还具备实时数据监控和预警功能，能够对术中影像数据中的异常情况进行实时检测，并及时向手术团队发出警报，有效降低手术风险。

本项目通过微创术中影像记录模块、术后图文报告生成模块和术中影像数据分析模块三个核心服务，构建了全腔镜手术中医疗数据资产获取、挖掘、应用的全流程服务链条，旨在以术中数据驱动减轻医生术中负担、提升医患信任度、加速临床医生培养，将医工交叉工具融入到医学生日常教学中。同时，本项目还注重系统的安全性和稳定性，多重采用了数据加密技术和冗余备份机制，确保术中影像数据的保密性和完整性，为医疗信息安全提供有力保障。

### 三、查新点

1、术中影像记录与传输采用 H.265 视频高性能传输标准，在相同的图像质量下，相比于传统的 H.264 传输方式，通过 H.265 编码的视频大小体积将减少大约 39-44% 并可以在低于 2Mbps 的网络环境下实现 4K 高清数字图像传送，减轻医疗影像传输与存储负担。此外，H.265 标准还具备更好的压缩效率和容错能力，能够在复杂的网络环境下保持稳定的传输性能，确保术中影像的实时性和连续性，为手术的顺利进行提供可靠的技术支持。

2、选择基于院内边缘计算的算法加速设计，能够降低术中医疗图像分析辅助、临床自动图文病例生成等服务的延迟及推理速度，同时减轻术中影像分析计算对云平台或数据中心服务器造成的大额能耗负担。边缘计算 AI 加速服务器的引入，不仅提高了算法的运行效率，还实现了数据的本地化处理，有效保护了患者的隐私信息。此外，该设计还结合了智能调度算法，能够根据手术的实时需求动态分配计算资源，进一步优化系统的性能和能耗比，为术中影像分析提供更加高效、节能的解决方案。

3、术后文书辅助决策系统（CDSS）基于现有临床术中文书相关信息，构建适用于辅助术后图文报告诊断的知识图谱，通过临床相关主题分面树、临床术后文书碎片知识、临床知识关系挖掘等技术，实现临床辅助决策术后报告自动生成。在临床信息抽取与数据分析方面产品性能远高于行业主流产品，辅助信息分析能力实现 60% 的提升，人力成本平均降低 45%。该系统还具备自我学习和持续优化的能力，能够随着临床数据的不断积累和更新，自动调整知识图谱的结构和内容，提高辅助决策的准确性和可靠性，为医生提供更加精准、个性化的医学实践参考。

4、实现适用于肝胆胰、胃肠等微创术中图像分割的实时 AI 医疗辅助系统。该系统汇集了大规模相关临床术中影像数据的 AI 预训练模型和基于 Transform 的 AI 图像分割算法。该算法将 Transformer 与 CNN 融合，依靠 Transformer 自带的长依赖特性，相比于传统的 CNN 解决了模型臃肿、计算量大、梯度消失、网络整体无法收敛的问题。Transformer 模型通常在大型数据集上预训练，针对较小的下游任务进行微调，比传统的 CNN 网络更具效率和可扩展性。同时，该系统还结合了多任务学习和注意力机制，能够同时处理多种术中影像分析任务，并自动聚焦于图像中的关键区域，提高诊断的准确性和效率，为微创手术的实时图像辅助提供了全新的技术手段。

#### 四、查新范围要求

要求查新机构通过查新，证明在所查范围内国内外有无相同或类似的文献报道。

#### 五、文献检索范围及检索策略

##### 中文数据库检索范围

1. 中国学术期刊网络出版总库	1915---2025年2月
2. 中国博士学位论文全文数据库	1984---2025年2月
3. 中国优秀硕士学位论文全文数据库	1984---2025年2月
4. 中国重要会议论文全文数据库	1953---2025年2月
5. 国家科技成果数据库	1978---2025年2月
6. 重庆维普中文科技期刊数据库	1989---2025年2月
7. 万方期刊论文数据库	1998---2025年2月
8. 万方学位论文数据库	1980---2025年2月
9. 万方会议论文数据库	1985---2025年2月
10. 国家科技图书文献中心	1984---2025年2月
11. 中国专利	

##### 外文数据库检索范围

12. Medline	1950 ---2025年2月
13. SCIE (Science Citation Index Expanded 科学引文索引)	1900 ---2025年2月
14. CPCI-S (ISTP 科学技术会议录索引)	1996 ---2025年2月
15. BIOSIS (BA 网络版)	1994 ---2025年2月
16. SciFinder Scholar (CA 网络版)	1907 ---2025年2月
17. ScienceDirect (Elsevier)	1995 ---2025年2月
18. PQDD 博硕士论文文摘数据库	1861 ---2025年2月
19. OVID 医学期刊数据库	1860 ---2025年2月
20. Wiley InterScience	1996 ---2025年2月
21. Springer Link	1990 ---2025年2月
22. PubMed	1950 ---2025年2月

#### 六、检索词和检索式

##### 检索词：

Deepseek、手术图文报告

Deepseek、Real-Time Image-Text Report

##### 检索式：

Deepseek AND 手术图文报告

Deepseek AND Real-Time Image-Text Report

## 七、检索结果

依据上述文献检索范围和检索式，检索到密切相关文献 5 篇。

1. 汤博, 崔晓海, 关正, 张晓刚, 吴荣谦, 吕毅. 外科手术语音图文记录报告系统的研发及临床应用研究[J]. 中国医学教育技术, 2016, 30 (02): 194-197.  
外科手术语音图文记录报告系统主要由硬件系统及软件系统组合而成。
2. 吕毅, 汤博, 仵正, 刘学民, 樊林. 西安交通大学第一附属医院, 外科手术语音图文记录报告工作系统[Z]. 鉴定日期:2016-03-01  
提供一种应用于外科手术过程的语音图文记录工作系统,既可实现术中实时图片的采集,又可配有主刀操作的步骤说明,通过软、硬件处理后,最终生成外科手术图文报告
3. Truhn, D., et al., Extracting structured information from unstructured histopathology reports using generative pre-trained transformer 4 (GPT-4). J Pathol, 2024. 262(3): p. 310-319. 使用生 GPT-4 从非结构化组织病理学报告中提取结构化信息
4. Chen, X., et al., ICGA-GPT: report generation and question answering for indocyanine green angiography images. Br J Ophthalmol, 2024. 108(10): p. 1450-1456. ICGA-GPT: 吲哚青绿血管造影图像的报告生成和问答。
5. Chen, K., W. Xu and X. Li, The Potential of Gemini and GPTs for Structured Report Generation based on Free-Text (18)F-FDG PET/CT Breast Cancer Reports. Acad Radiol, 2025. 32(2): p. 624-633. Gemini 和 GPT 在基于 Free-Text (18) F-FDG PET/CT 乳腺癌报告的结构化报告生成中的潜力。

## 八、查新结论

受西安交通大学第一附属医院委托，针对基于 Deepseek 大模型的适用于全腔镜外科术中实时手术图文报告生成系统课题的查新要点，我们利用国内数据库进行了查新检索。从国内外 22 个数据库及网站共检索到 5 篇相关文献。

文献 1 讨论了外科手术语音图文记录报告系统主要由硬件系统及软件系统组合而成。

文献 2 提供一种应用于外科手术过程的语音图文记录工作系统,既可实现术中实时图片的采集,又可配有主刀操作的步骤说明,通过软、硬件处理后,最终生成外科手术图文报告。

文献 1、2 为本课题成员发表的相关文献。

文献 3 使用生 GPT-4 从非结构化组织病理学报告中提取结构化信息

文献 4 讨论了 ICGA-GPT: 吲哚青绿血管造影图像的报告生成和问答。

文献 5 讨论了 Gemini 和 GPT 在基于 Free-Text (18) F-FDG PET/CT 乳腺癌症报告的结构化报告生成中的潜力。

文献 3-5 与本课题基于 Deepseek 大模型的适用于全腔镜外科术中实时手术图文报告生成系统内容不同。

#### 综上所述:

**国内外未见与本课题查新点** (1、术中影像记录与传输采用 H.265 视频高性能传输标准, 在相同的图像质量下, 相比于传统的 H.264 传输方式, 通过 H.265 编码的视频大小体积将减少大约 39-44%, 并可以在低于 2Mbps 的网络环境下实现 4K 高清数字图像传送, 减轻医疗影像传输与存储负担。此外, H.265 标准还具备更好的压缩效率和容错能力, 能够在复杂的网络环境下保持稳定的传输性能, 确保术中影像的实时性和连续性, 为手术的顺利进行提供可靠的技术支持。2、选择基于院内边缘计算的算法加速设计, 能够降低术中医疗图像分析辅助、临床自动图文病例生成等服务的延迟及推理速度, 同时减轻术中影像分析计算对云平台或数据中心服务器造成的大额能耗负担。边缘计算 AI 加速服务器的引入, 不仅提高了算法的运行效率, 还实现了数据的本地化处理, 有效保护了患者的隐私信息。此外, 该设计还结合了智能调度算法, 能够根据手术的实时需求动态分配计算资源, 进一步优化系统的性能和能耗比, 为术中影像分析提供更加高效、节能的解决方案。3、术后文书辅助决策系统 (CDSS) 基于现有临床术中文书相关信息, 构建适用于辅助术后图文报告诊断的知识图谱, 通过临床相关主题分面树、临床术后文书碎片知识、临床知识关系挖掘等技术, 实现临床辅助决策术后报告自动生成。在临床信息抽取与数据分析方面产品性能远高于行业主流产品, 辅助信息分析能力实现 60%的提升, 人力成本平均降低 45%。该系统还具备自我学习和持续优化的能力, 能够随着临床数据的不断积累和更新, 自动调整知识图谱的结构和内容, 提高辅助决策的准确性和可靠性, 为医生提供更加精准、个性化的医学实践参考。4、实现适用于肝胆胰、胃肠等微创术中图像分割的实时 AI 医疗辅助系统。该系统汇集了大规模相关临床术中影像数据的 AI 预训练模型和基于 Transform 的 AI 图像分割算法。该算法将 Transformer 与 CNN 融合, 依靠 Transformer 自带的长依赖特性, 相比于传统的 CNN 解决了模型臃肿、计算量大、梯度消失、网络整体无法收敛的问题。Transformer 模型通常在大型数据集上预训练, 针对较小的下游任务进行微调, 比传统的 CNN 网络更具效率和可扩展性。同时, 该系统还结合了多任务学习和注意力机制, 能够同

时处理多种术中影像分析任务，并自动聚焦于图像中的关键区域，提高诊断的准确性和效率，为微创手术的实时图像辅助提供了全新的技术手段。) 内容相同的文献报道。

查新员（签字）：

查新员职称：副研究员

审核员（签字）： 张小曼

审核员职称：副研究员

（科技查新专用章）

2025年3月11日

### 九、查新员、审核员声明

1. 报告中陈述的事实是真实和准确的；
2. 我们按照科技查新规范进行查新、文献分析和审核，并做出上述查新结论；
3. 我们获取的报酬与本报告中的分析、意见和结论无关，也与本报告使用无关；

查新员（签字）：

审核员（签字）： 张小曼

2025年3月11日

2025年3月11日

### 十、附件清单

相关文献

### 十一、备注

- 1.本查新报告无“查新专用章”和签字无效；
- 2.本查新报告涂改、部分复印无效；
- 3.检索结果及查新报告结论仅供参考；
- 4.查新报告有效期半年。

#### 相关文献如下：

1. 汤博, 崔晓海, 关正, 张晓刚, 吴荣谦, 吕毅. 外科手术语音图文记录报告系统的研发及临床应用研究[J]. 中国医学教育技术, 2016, 30 (02): 194-197.

摘要:外科手术语音图文记录报告系统主要由硬件系统及软件系统组合而成。硬件主要包括:图像采集单元,用于采集术中图像信息;可活动臂单元,用于安装图像采集单元调整采集角度;PC 计算机单元,配有语音接收插口和 USB 插口,语音接收插口与语音收发单元相连接,USB 插口与彩色打印机单元相连接;工作站单元,与图像采集单元连接接收其采集的图像

信息;术中监视器单元,与工作站单元连接对其接收的图像信息进行实时显示;无线路由器单元,用于建立网络;彩色打印机单元,将 PC 计算机单元下发的打印内容进行打印;语音收发单元,接收语音描述以及相关解释并发送至 PC 计算机单元。软件主要是自主开发具有自主知识产权的外科手术语音图文软件。通过初步使用,显示该系统既可实现术中实时图片采集,又可配有主刀操作步骤说明,通过软硬件处理后生成手术图文报告。

2. 吕毅, 汤博, 仵正, 刘学民, 樊林. 西安交通大学第一附属医院, 外科手术语音图文记录报告工作系统[Z]. 鉴定日期:2016-03-01

成果简介:随着计算机网络信息技术的发展,信息技术、网络技术以及云计算等技术在医疗健康领域的也得到了广泛应用和长足的发展,使得依托互联网技术的云医院已现端倪,便捷高效的结果提供将颠覆当前传统的医疗模式。从最初使用计算机记录病历和开处方,到 PACs(医学影像系统)、uS(检验系统)、HIS(医院信息系统)、AI(医院管理系统)等相继进入,管理效率得到了极大快速的提升。医疗文档作为医生进行临床活动主要依据的,许多早已采用了图文报告形式,如各类临床病理学诊断、影像学检查、胃镜检查等报告单,并且都附有详细的图文报告之规范要求,这种简洁的图文并茂形式不仅有助于医患的沟通与医师诊断,而且有助于减少医疗资源的重复与浪费。外科手术报告作为临床最为重要的医疗文书之一,目前仅使用文字描述,缺少必要的纪实照片,成为医疗文书全面信息化、数字化“最后的死角”。传统手术图文记录,是由下级医师术后手绘完成的,其缺点第一,不精确,非主刀执笔;第二,不明了,文字不直观;第三,负担重,导致病例文字大量的复制粘贴。据查阅文献和相关报道,国内、外专家对于外科手术图文记录产生过共鸣,世界著名肝胆外科泰斗二村雄次教授在其史诗级论著《胆道外科:要点与盲点》中指出:“手绘图文手术记录虽费时费力,但极有利于医师的成长与医患交流,应设法在临床推广。”;2015年1月9日,由我国著名肝胆外科专家吕毅教授,在中华医学会外科学术学组会议上发起了“倡用图文手术记录专家共识”,与会专家(赵玉沛院士等)形成会议共识指南,并且在《中华肝脏外科手术学电子杂志》2015第5期已经发表。总而言之,术中图文资料对疾病诊疗技术革新、年轻医师成长具有重要意义,同时还可以帮助患者了解诊疗过程缓解医疗纠纷的压力。研究的目的在于提供一种应用于外科手术过程的语音图文记录工作系统,既可实现术中实时图片的采集,又可配有主刀操作的步骤说明,通过软、硬件处理后,最终生成外科手术图文报告。与现有技术相比,本系统旨在解决传统文字手术记录过程中所引发的信息不准确、医患交流不畅等难题。在使用过程中,一方面可以实时记录术中图片,另一方面可根据术中术者语音表达的实时描述及解释,并且依托互联网百度语音进行在线翻译文字,最终完成手术过程采集的图文信息一一配对,整理后即可形成手术过程记录的图文报告,与传统手术的手绘、文字术中过程描述相比,要可靠、真实、便捷的多。外科手术图文报告,不论对医生处于教学、交流...

3. Truhn, D., et al., Extracting structured information from unstructured histopathology reports using generative pre-trained transformer 4 (GPT-4). J Pathol, 2024. 262(3): p. 310-319.

Deep learning applied to whole-slide histopathology images (WSIs) has the potential to enhance precision oncology and alleviate the workload of experts. However, developing these models necessitates large amounts of data with ground truth labels, which can be both time-consuming and expensive to obtain. Pathology reports are typically unstructured or poorly structured texts, and efforts to implement structured reporting templates have been unsuccessful, as these efforts lead to perceived extra workload. In this study, we hypothesised that large language models (LLMs), such as the generative pre-trained transformer 4 (GPT-4), can extract structured data from unstructured plain language reports using a zero-shot approach without requiring any re-training. We tested this hypothesis by utilising GPT-4 to extract information from histopathological reports, focusing on two extensive sets of pathology reports for colorectal

cancer and glioblastoma. We found a high concordance between LLM-generated structured data and human-generated structured data. Consequently, LLMs could potentially be employed routinely to extract ground truth data for machine learning from unstructured pathology reports in the future. © 2023 The Authors. The Journal of Pathology published by John Wiley & Sons Ltd on behalf of The Pathological Society of Great Britain and Ireland.

4. Chen, X., et al., ICGA-GPT: report generation and question answering for indocyanine green angiography images. *Br J Ophthalmol*, 2024. 108(10): p. 1450-1456.

**BACKGROUND:** Indocyanine green angiography (ICGA) is vital for diagnosing chorioretinal diseases, but its interpretation and patient communication require extensive expertise and time-consuming efforts. We aim to develop a bilingual ICGA report generation and question-answering (QA) system. **METHODS:** Our dataset comprised 213 129 ICGA images from 2919 participants. The system comprised two stages: image-text alignment for report generation by a multimodal transformer architecture, and large language model (LLM)-based QA with ICGA text reports and human-input questions. Performance was assessed using both qualitative metrics (including Bilingual Evaluation Understudy (BLEU), Consensus-based Image Description Evaluation (CIDEr), Recall-Oriented Understudy for Gisting Evaluation-Longest Common Subsequence (ROUGE-L), Semantic Propositional Image Caption Evaluation (SPICE), accuracy, sensitivity, specificity, precision and F1 score) and subjective evaluation by three experienced ophthalmologists using 5-point scales (5 refers to high quality). **RESULTS:** We produced 8757 ICGA reports covering 39 disease-related conditions after bilingual translation (66.7% English, 33.3% Chinese). The ICGA-GPT model's report generation performance was evaluated with BLEU scores (1-4) of 0.48, 0.44, 0.40 and 0.37; CIDEr of 0.82; ROUGE of 0.41 and SPICE of 0.18. For disease-based metrics, the average specificity, accuracy, precision, sensitivity and F1 score were 0.98, 0.94, 0.70, 0.68 and 0.64, respectively. Assessing the quality of 50 images (100 reports), three ophthalmologists achieved substantial agreement ( $\kappa=0.723$  for completeness,  $\kappa=0.738$  for accuracy), yielding scores from 3.20 to 3.55. In an interactive QA scenario involving 100 generated answers, the ophthalmologists provided scores of 4.24, 4.22 and 4.10, displaying good consistency ( $\kappa=0.779$ ). **CONCLUSION:** This pioneering study introduces the ICGA-GPT model for report generation and interactive QA for the first time, underscoring the potential of LLMs in assisting with automated ICGA image interpretation.

5. Chen, K., W. Xu and X. Li, The Potential of Gemini and GPTs for Structured Report Generation based on Free-Text (18)F-FDG PET/CT Breast Cancer Reports. *Acad Radiol*, 2025. 32(2): p. 624-633.

**RATIONALE AND OBJECTIVE:** To compare the performance of large language model (LLM) based Gemini and Generative Pre-trained Transformers (GPTs) in data mining and generating structured reports based on free-text PET/CT reports for breast cancer after user-defined tasks. **MATERIALS AND METHODS:** Breast cancer patients (mean age, 50 years  $\pm$  11 [SD]; all female) who underwent consecutive (18)F-FDG PET/CT for follow-up between July 2005 and October 2023 were retrospectively included in the study. A total of twenty reports from 10 patients were used to train user-defined text prompts for Gemini and GPTs, by which structured PET/CT reports were generated. The natural language processing (NLP) generated structured reports and the structured reports annotated by nuclear medicine physicians were compared in terms of data extraction accuracy and capacity of progress decision-making. Statistical methods,

including chi-square test, McNemar test and paired samples t-test, were employed in the study. RESULTS: The structured PET/CT reports for 131 patients were generated by using the two NLP techniques, including Gemini and GPTs. In general, GPTs exhibited superiority over Gemini in data mining in terms of primary lesion size (89.6% vs. 53.8%,  $p < 0.001$ ) and metastatic lesions (96.3% vs 89.6%,  $p < 0.001$ ). Moreover, GPTs outperformed Gemini in making decision for progress ( $p < 0.001$ ) and semantic similarity (F1 score 0.930 vs 0.907,  $p < 0.001$ ) for reports. CONCLUSION: GPTs outperformed Gemini in generating structured reports based on free-text PET/CT reports, which is potentially applied in clinical practice. DATA AVAILABILITY: The data used and/or analyzed during the current study are available from the corresponding author on reasonable request.

西安交通大学