CUBE.AI: A Multi-agent AI Framework For Autonomous End-to-end Radiology Research

Anonymous Author(s)

Anonymous Institution

Abstract. Radiology research requires integrating clinical reasoning with complex, fragmented technical workflows. We present ***CUBE.AI*** , a multi-agent system that autonomously orchestrates the complete radiology research lifecycle through specialized agents: a Planning Agent for hypothesis formalization and literature retrieval, a Cohort Selection Agent for PACS-integrated patient assembly, a Segmentation Agent with hierarchical fallback across automatic and interactive methods, a Feature Extraction Agent for extracting meaningful representations and feature analysis, a Modeling Agent for AI models training and development, and a Reporting Agent that compiles findings into a LaTeX-formatted manuscript. Agents communicate with medical infrastructure through the Model Context Protocol (MCP), ensuring safe and standardized tool access. Evaluation via an LLM as a judge demonstrates reliable end-to-end research automation. ***CUBE.AI*** achieved planning and reporting quality scores of 8.90/10 and 7.43/10, respectively.

Keywords: Multi-Agent Systems · Radiology Research · Clinical Workflow Integration · Medical Imaging · PACS Integration · Large Language Models

1 Introduction

Accurate diagnosis, prognosis, and treatment planning in radiology increasingly depend on quantitative information that is not fully captured by visual interpretation alone, yet radiologists routinely lack an efficient way to translate this quantitative potential into validated, publishable research findings [11]. Extracting and validating imaging biomarkers requires radiologists to manually coordinate cohort curation, segmentation, feature extraction, statistical modeling, and manuscript writing across disconnected tools, a process that is time-consuming, error-prone, and inaccessible to clinicians without dedicated computational support [19,1]. Therefore, we aim to automate the end-to-end radiology research pipeline, from clinical hypothesis to manuscript, within a single autonomous system that removes the need for manual tool-switching and technical expertise at each stage.

Existing work has focused on automating individual stages of this pipeline rather than the pipeline as a whole. Image visualization and annotation platforms and radiomics feature-extraction libraries provide strong functionality for specific

Table 1: Capability comparison across agentic systems relevant to radiology research. ✓ = supported; ✗ = not supported; ∼ = partial. Columns correspond to the seven stages of the ***CUBE.AI*** pipeline.

System	Study Design	Cohort Planning	Segmentation	Feature Extraction	ML Modelling	Reporting	Agentic
AI Scientist [13]	✓	✗	✗	✗	∼	✓	✓
Agent Laboratory [18]	✓	✗	✗	✗	✓	✓	✓
InternAgent-1.5 [5]	✓	✗	✗	✗	✓	✓	✓
RadFabric [2]	✗	✗	✗	✗	✗	✓	✓
CT-Agent [14]	✗	✗	✗	✗	✗	✗	✓
Radiologist Copilot [22]	✗	✗	✗	✗	✗	✓	✓
M^3Builder [4]	✗	✗	✗	✗	✓	✗	✓
mAIstro [20]	✗	✗	∼	∼	✓	✗	✓
CUBE.AI **(ours)**	✓	✓	✓	✓	✓	✓	✓

tasks such as ROI delineation or handcrafted feature computation [19], but each is designed in isolation: the absence of shared data standards and interoperable interfaces across tools means every hand-off between stages, from segmentation to feature extraction to modeling, requires manual export, reformatting, and transfer, which increases workflow complexity and delays the research outcome [9,10].

A different line of research has focused on agentic AI systems for scientific discovery. AI Scientist [13] demonstrated the feasibility of end-to-end automated research, and Agent Laboratory [18] introduced a structured multi-stage research pipeline and showed that human feedback substantially improves research quality. However, as general-purpose research agents, both lack the domain-specific components, including PACS/DICOM-aware data ingestion, ROI segmentation utilities, and radiomics-compliant feature extraction and reporting, so neither can be applied to imaging research without substantial redesign.

Radiology-focused agentic systems address only parts of the workflow, targeting either clinical image interpretation and report generation [2,22,14] or model-development pipelines [4,20]. None integrates study design, cohort assembly, modeling, and reporting into a unified workflow.

As shown in Table 1, existing systems remain fragmented: general research agents omit imaging-specific stages, whereas radiology agents omit either upstream study planning or downstream analysis and reporting. Nevertheless, all of these methods suffer from a shared limitation: none supports the complete radiology research pathway, from hypothesis generation to manuscript writing, within a single unified, radiology-specialized environment.

To this end, we propose ***CUBE.AI*** , a radiology-specialized multi-agent research system for end-to-end radiology research automation. ***CUBE.AI*** consists of a Planner Agent that translates a natural-language clinical hypothesis into a literature-grounded, executable study protocol, and a set of execution-stage agents (Cohort Selection, Segmentation, Feature Extraction, Modeling, and Reporting). Specifically, the Planner Agent solves the gap of disconnected study design by grounding hypothesis structuring and workflow planning in retrieved literature, and the execution agents solve the gap of manual hand-offs by operating on shared, validated upstream artifacts rather than researcher-mediated file transfer. Our contributions are as follows:

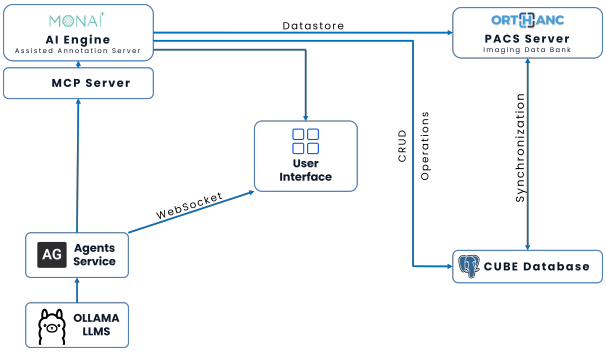


Fig. 1: Physical deployment architecture illustrating integration with MONAI [3], Orthanc [7], Ollama, MCP and the CUBE database through a unified orchestration layer.

- We introduce ***CUBE.AI*** , the first agentic AI system to automate the complete radiology research lifecycle, from clinical hypothesis to final manuscript, within a single unified environment.
- We design a literature-grounded Planner Agent that produces executable, clinically valid research protocols, coupled with specialized execution agents.
- We provide a three-level evaluation framework (planning, per-agent, and end-to-end) and quantify the factual reliability of the generated reports through a hallucination audit against execution traces.

2 Methods

CUBE.AI is a multi-agent radiology research system built as an augmented medical imaging viewer over OHIF [23]. Unlike conventional fragmented workflows, ***CUBE.AI*** coordinates the entire process through a single orchestration layer that connects specialized agents to clinical and computational infrastructure. As illustrated in Fig. 1, the platform integrates: (i) Picture Archiving and Communication Systems (PACS) for direct access to medical imaging data; (ii) assisted annotation and automatic segmentation tools; (iii) Python-based libraries for radiomic feature extraction, statistical modeling, and data visualization; and (iv) a LaTeX engine for automated scientific report generation.

The system operates in two stages (Fig. 2): a **literature-grounded Planning Stage**, in which a Planner Agent converts a clinician's high-level hypothesis into a structured experimental design; and an **Execution Stage**, in which specialized agents autonomously assemble the study dataset and carry it through feature extraction, modeling, and reporting, sharing state through a centralized scratchpad, while maintaining detailed observability, allowing the researcher to inspect intermediate agent decisions and outputs at every step.

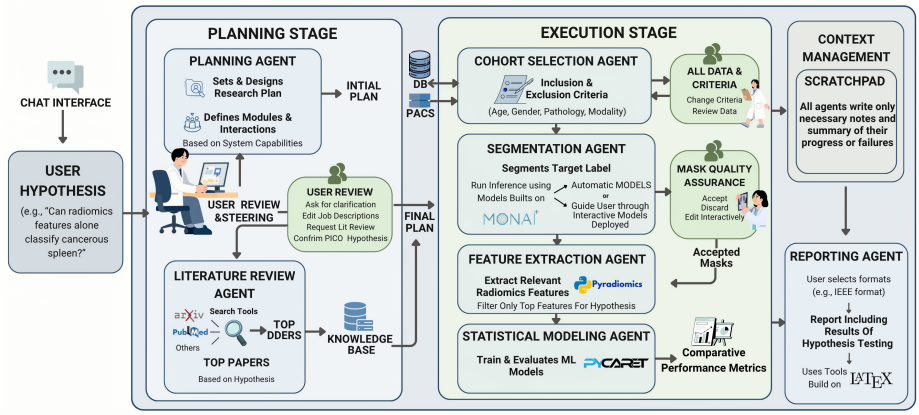


Fig. 2: Architecture of the proposed CUBE.AI agentic radiology research system. The workflow is divided into three modules: (i) a Planning Stage that performs literature-guided study design and user validation, (ii) an Execution Stage comprising specialized agents for cohort construction, segmentation, radiomics, statistical analysis, and reporting, and (iii) a Context Management module that provides shared memory and coordination through a centralized scratchpad.

2.1 Literature-Grounded Planning Stage

Unlike prior agentic systems, which plan without evidence grounding [13] or retrieve literature only for post-hoc manuscript writing [18], ***CUBE.AI*** places literature retrieval at the start of the pipeline, so the resulting knowledge base guides protocol generation before execution begins.

The Planner Agent converts the user's request into a structured PICO hypothesis, augments it with evidence retrieved by the Literature Review Agent, and decomposes the overall research objective into an ordered sequence of steps, selects the appropriate agents required for each step, and generates detailed task-specific instructions tailored to the study design. Researchers may iteratively refine the plan before approving execution.

The Literature Review Agent constructs a structured knowledge base by decomposing the research hypothesis into searchable concepts and issuing Boolean queries across multiple literature sources. Retrieved studies are ranked according to their relevance to the PICO hypothesis before entering an iterative refinement loop that expands the search space and improves retrieval quality. The refinement process terminates once the incremental knowledge gain falls below a predefined threshold ($\Delta_t < \delta$), where the knowledge gain at iteration t is defined as

$$\Delta_t = 0.3\frac{|T_t \setminus T_{t-1}|}{\max(|T_t|, 1)} + 0.7\max\left(\frac{Q_t - Q_{t-1}}{\max(Q_{t-1}, 1)}, 0\right),$$

where T_t denotes the set of discovered domain concepts at iteration t, Q_t denotes the number of validated high-relevance search queries, the first term

quantifies semantic expansion through newly discovered concepts, and the second quantifies retrieval-quality improvement through additional validated search queries.

Finally, the agent produces a structured literature review summarizing current evidence, methodological precedents, research gaps, recommended imaging modalities, and candidate analysis strategies. The resulting knowledge base is subsequently used to justify cohort criteria, imaging endpoints, and downstream analytical decisions, ensuring that study design remains aligned with established evidence throughout the research lifecycle.

2.2 Autonomous Dataset Assembly

Existing research pipelines assume that the analysis dataset has already been constructed, requiring researchers to manually query the PACS, define the cohort, and obtain segmentation masks before analysis [4,20]. ***CUBE.AI*** removes this assumption through an autonomous Dataset Assembly stage that constructs the study cohort directly from the PACS and adaptively performs segmentation when dedicated models are unavailable.

The Cohort Selection Agent derives inclusion and exclusion criteria from the experimental protocol and retrieves the study cohort by querying the PACS, formally defining the selected cohort as $U = \{u \in \text{PACS} \mid u \models C_{\text{inc}} \wedge u \not\models C_{\text{exc}}\}$, where u denotes a candidate imaging study, C_{inc} and C_{exc} represent the inclusion and exclusion criteria, respectively, and U denotes the final cohort satisfying all eligibility constraints. The agent then verifies whether the assembled cohort satisfies the statistical requirements of the study and, if necessary, recommends additional imaging characteristics required to complete the cohort while remaining decoupled from direct patient data access.

The Segmentation Agent completes dataset assembly by generating standardized segmentation masks for the selected cohort. For each study, segmentation follows the prioritized fallback policy $M_{\text{PACS}} \rightarrow M_{\text{MONAI}} \rightarrow M_{\text{VoxTell}} \rightarrow M_{\text{HITL}}$, selecting the first available segmentation mask M. Existing masks are reused when available; otherwise, an organ-specific MONAI Label model is invoked, followed by VoxTell [17], a vision-language model for zero-shot text-prompted segmentation, using offline precomputed text embeddings to reduce inference memory and enable unattended execution, and finally a Clinician-in-the-Loop workflow when no automatic solution exists.

2.3 Execution Stage

Once the dataset is assembled, execution proceeds sequentially through feature extraction, modeling, and reporting, streaming intermediate outputs to the shared scratchpad throughout.

The Feature Extraction Agent transforms the segmented cohort into a radiomic feature matrix $F \in \mathbb{R}^{n \times d}$, where each row corresponds to one patient and each column to a quantitative imaging descriptor. Following the mAIstro framework [20], the agent automatically generates modality-specific extraction

configurations and computes standardized radiomic features using `PyRadiomics` [6].

The Modeling Agent preprocesses the extracted feature matrix, performs feature selection, and learns the predictive model $f^* = \arg\max_{f \in \mathcal{F}} S(f)$, where S denotes the task-specific evaluation metric. The agent dynamically selects classification or regression algorithms according to the research objective and evaluates performance using AUC, precision, recall, and F1-score for classification, and MAE, MSE, RMSE, and R^2 for regression.

The Reporting Agent retrieves execution logs, intermediate artifacts, and analytical results from the shared scratchpad and synthesizes the final manuscript as $R = \mathcal{G}(S)$, where S denotes the complete execution state. The generated report summarizes the methodology, experimental results, and interpretation before being exported as a LaTeX manuscript and compiled into PDF format.

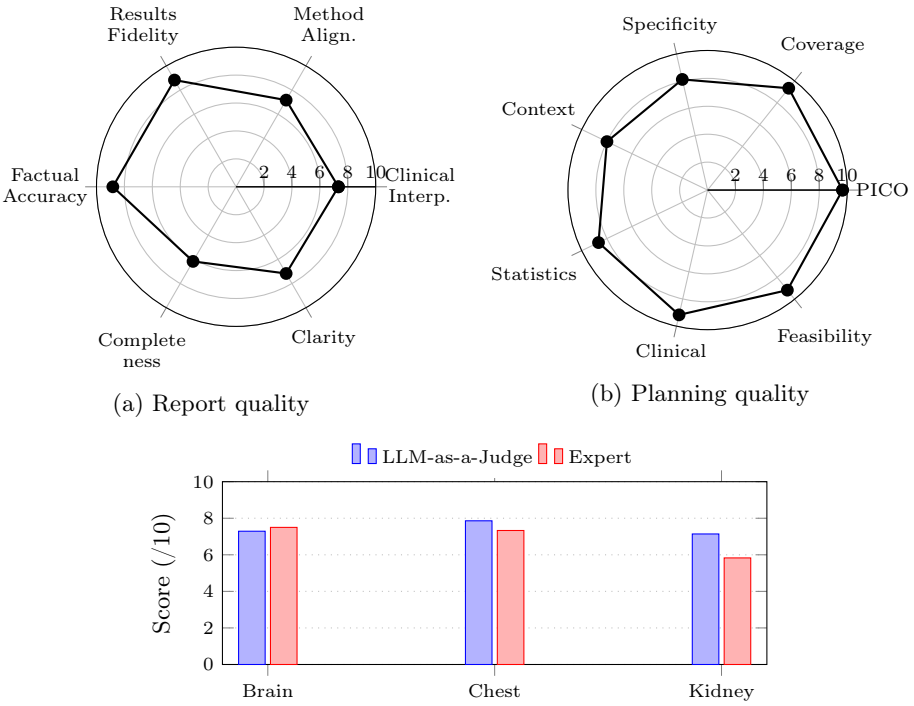
3 Experimental Results

Evaluation spanned pathology classification and regression benchmarks across multiple organ systems, each derived from a published study serving as the ground-truth reference. ***CUBE.AI*** was evaluated at three complementary levels: (i) planning-stage quality, (ii) individual agent performance, and (iii) end-to-end autonomous execution. Since the primary objective of ***CUBE.AI*** is full research automation, we present the end-to-end evaluation first, followed by agent-level validation. All experiments were conducted on a commodity workstation (Intel Core i7-13620H, NVIDIA RTX 4050 Laptop GPU with 6 GB VRAM, 32 GB RAM) using GPT-OSS-120B [16] as the reasoning engine, representing a conservative lower bound on performance with stronger frontier models.

3.1 End-to-End Autonomous Evaluation

CUBE.AI was evaluated autonomously on three representative studies (brain [15], chest [8], and kidney [12]), each receiving only a natural-language clinical hypothesis. The primary evaluation uses an LLM-as-a-Judge framework following Agent Laboratory [18], employing MIMO-V2.5 [21] with rubrics derived from reference publications. As complementary validation, three independent medical-imaging researchers evaluated the generated manuscripts on six qualitative dimensions.

The Planning Stage achieved a mean score of 8.90, with the strongest performance in PICO formulation (9.67) and workflow coverage (9.33). End-to-end report quality reached 7.43/10, while expert reviewers assigned 6.89/10, showing close agreement with the automated evaluation. Results Fidelity and Factual Accuracy achieved the highest scores (both 8.83), whereas Completeness (6.17) remained the primary limitation, indicating that generated manuscripts were generally correct but occasionally omitted methodological details. Across all studies, the mean hallucination rate was 10.4% which was derived from comparing the claims in the final report with the agents traces and intermediate results.



(a) Report quality

(b) Planning quality

Fig. 3: End-to-end evaluation. Top: planning and report quality across evaluation dimensions. Bottom: report quality under LLM-as-a-Judge and preliminary expert evaluation. Mean scores: Planning 8.90 ± 0.21, LLM 7.43 ± 0.37, Expert 6.89 ± 0.87. Mean hallucination rate: 10.4%.

3.2 Agent-Level Evaluation

Each execution-stage agent was evaluated independently using ground-truth upstream inputs to isolate intrinsic performance from error propagation (Table 2).

Table 2: Agent-level evaluation.

Agent	**Function**	**Metric**	**Result**
Population	Cohort selection	F1, adherence	F1=0.73 ± 0.46; 75% adherence
Segmentation	ROI generation	DSC, routing	DSC=0.899 ± 0.031; 5/5 routing
Feature Extraction	Radiomics	Completion, recovery	0.64 completion; 83.3% recovery
Modeling	ML pipeline	Completion, recovery	100%; 100% recovery
Reporting	Manuscript	Quality	$7.43 \pm 0.37/10$

The Segmentation Agent achieved perfect routing with a Dice score of 0.899, exceeding the radiomics reliability threshold (DSC$\geq$ 0.88). The Modeling Agent

completed all workflows with full recovery from execution failures. Feature Extraction exhibited the lowest robustness (64% completion, 83.3% recovery), identifying workflow ordering as the primary bottleneck. Population Agent errors were mainly attributable to heterogeneous dataset naming rather than reasoning failures.

4 Discussion

The evaluation reveals that the primary challenge is scientific *completeness* rather than *correctness*. Reports consistently preserved factual accuracy (Factual Accuracy and Results Fidelity both 8.83/10) while Completeness (6.17/10) emerged as the principal weakness, reflecting methodological omissions rather than factual errors, a distinction with direct clinical relevance, as incomplete-but-accurate drafts require only editorial augmentation by a domain expert. The preliminary expert evaluation (6.89/10) corroborated the automated profile: both protocols identified the same strengths (clarity, internal consistency) and the same weakness (methodological depth), increasing confidence that the LLM-judge scores are directionally reliable despite the limited reviewer sample. The plan-to-report gap (8.90 vs. 7.43/10) indicates that $\sim$83% of planned methodology was executed and reflected in the final draft; the residual loss arose primarily in literature contextualisation and statistical reporting. That these results were obtained with a non-frontier model (GPT-OSS) indicates performance derives primarily from the multi-agent architecture rather than LLM capacity; frontier models are expected to yield further gains without architectural modification.

4.1 Limitations

Primarily, report quality relied on a single LLM-as-a-Judge model (MIMO V2.5), which may introduce systematic bias undetected without a multi-judge ensemble. Furthermore, the complementary expert evaluation ($n = 3$, 3 manuscripts) provides a corroborating signal but is insufficient for inter-rater reliability analysis (e.g., Cohen's κ); expanding to a structured expert cohort with formal reliability measurement is the highest-priority evaluation objective. Beyond the evaluation, completeness remains the most actionable technical gap, motivating improvements in literature-grounded discussion generation and automated figure synthesis. Finally, generalisation across scanner heterogeneity, additional modalities, and rare pathologies requires larger multi-organ benchmarks.

5 Conclusion

CUBE.AI demonstrates that human-supervised multi-agent systems can execute end-to-end radiology research with factual fidelity sufficient for expert-reviewable drafts. The system reduces workflow fragmentation and lowers the technical barrier to clinical imaging research, establishing the feasibility of the proposed architecture and motivating multi-site validation as the next step toward real-world deployment.

References

1. Cè, M., Chiriac, M.D., Cozzi, A., Macrì, L., Rabaiotti, F.L., Irmici, G., Fazzini, D., Carrafiello, G., Cellina, M.: Decoding radiomics: A step-by-step guide to machine learning workflow in hand-crafted and deep learning radiomics studies. Diagnostics **14**(22), 2473 (2024). https://doi.org/10.3390/diagnostics14222473
2. Chen, W., Dong, Y., Ding, Z., Shi, Y., Zhou, Y., Zeng, F., Luo, Y., Lin, T., Su, Y., Wu, Y., Zhang, K., Xiang, Z., Liu, T., Liu, N., Sun, L., Yuan, Y., Li, X.: Radfabric: Agentic ai system with reasoning capability for radiology (2025), local PDF: RadFabric.pdf
3. Diaz-Pinto, A., Alle, S., Nath, V., Tang, Y., Ihsani, A., Asad, M., et al.: Monai label: A framework for ai-assisted interactive labeling of 3d medical images. Medical Image Analysis **95**, 103207 (2024). https://doi.org/10.1016/j.media.2024.103207
4. Feng, J., Zheng, Q., Wu, C., Zhao, Z., Zhang, Y., Wang, Y., Xie, W.: M3builder: A multi-agent system for automated machine learning in medical imaging (2025), local PDF: Agentic AI M3Builder.pdf
5. Feng, S., Ma, R., Yan, X., Fan, Y., Hu, Y., Huang, S., Zhang, S., Cao, Z., Peng, T., Yuan, J., et al.: InternAgent-1.5: A unified agentic framework for long-horizon autonomous scientific discovery. arXiv:2602.08990 (2026), https://arxiv.org/abs/2602.08990
6. van Griethuysen, J.J.M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., et al.: Computational radiomics system to decode the radiographic phenotype. Cancer Research **77**(21), e104–e107 (2017). https://doi.org/10.1158/0008-5472.CAN-17-0339
7. Jodogne, S.: Orthanc: A lightweight, restful dicom server for healthcare and medical research. BioMedical Engineering OnLine **13**(147) (2014). https://doi.org/10.1186/1475-925X-13-147
8. Kermany, D.S., Goldbaum, M., Cai, W., Valentim, C.C.S., Liang, H., Baxter, S.L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M.K., Pei, J., Ting, M.Y.L., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., Shi, A., Zhang, R., Zheng, L., Hou, R., Shi, W., Fu, X., Huu, V.A.N., Wen, C., Zhang, E.D., Zhang, C.L., Li, O., Wang, X., Singer, M.A., Sun, X., Xu, J., Tafreshi, A., Lewis, M.A., Xia, H., Zhang, K.: Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell **172**(5), 1122–1131.e9 (2018). https://doi.org/10.1016/j.cell.2018.02.010
9. Kocak, B., Baessler, B., Bakas, S., Buckler, A.J., Cuocolo, R., Fedorov, A., Maier-Hein, K., Rubin, D.L., Strawbridge, R., Taieb, J., Uhlig, J.: CheckList for EvaluAtion of Radiomics research (CLEAR): a step-by-step reporting guideline for authors and reviewers endorsed by ESR and EuSoMII. Insights into Imaging **14**(1), 75 (2023). https://doi.org/10.1186/s13244-023-01415-8
10. Lambin, P., Kinahan, P.E., Siegel, E.L., Shankar, L.K.: Criteria for the translation of radiomics into clinically useful tests. Nature Reviews Clinical Oncology **20**(2), 69–82 (2023). https://doi.org/10.1038/s41571-022-00707-0
11. Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R.G., Granton, P., Zegers, C.M., Gillies, R., Boellaard, R., Dekker, A., Aerts, H.J.: Radiomics: Extracting more information from medical images using advanced feature analysis. European Journal of Cancer **48**(4), 441–446 (2012). https://doi.org/10.1016/j.ejca.2011.11.036

12. Li, Y., Gao, X., Tang, X., Lin, S., Pang, H.: Research on automatic classification technology of kidney tumor and normal kidney tissue based on computed tomography radiomics. Frontiers in Oncology **Volume 13 - 2023** (2023). https://doi.org/10.3389/fonc.2023.1013085, https://www.frontiersin.org/journals/oncology/articles/10.3389/fonc.2023.1013085
13. Lu, C., Lu, C., Lange, R.T., Foerster, J., Clune, J., Ha, D.: The AI Scientist: Towards fully automated open-ended scientific discovery. arXiv:2408.06292 (2024), https://arxiv.org/abs/2408.06292
14. Mao, Y., Xu, W., Qin, Y., Gao, Y.: Ct-agent: A multimodal-llm agent for 3d ct radiology question answering (2025), local PDF: Agentic AI - VLM - CT-Agent.pdf
15. Menze, B.H., Jakab, A., Bauer, S., Kalpathy-Cramer, J., Farahani, K., Kirby, J., Burren, Y., Porz, N., Slotboom, J., Wiest, R., Lanczi, L., Gerstner, E., Weber, M.A., Arbel, T., Avants, B.B., Ayache, N., Buendia, P., Collins, D.L., Cordier, N., Corso, J.J., Criminisi, A., Das, T., Delingette, H., Demiralp, Ç., Durst, C.R., Dojat, M., Doyle, S., Festa, J., Forbes, F., Geremia, E., Glocker, B., Golland, P., Guo, X., Hamamci, A., Iftekharuddin, K.M., Jena, R., John, N.M., Konukoglu, E., Lashkari, D., Mariz, J.A., Meier, R., Pereira, S., Precup, D., Price, S.J., Raviv, T.R., Reza, S.M., Ryan, M., Sarikaya, D., Schwartz, L., Shin, H.C., Shotton, J., Silva, C.A., Sousa, N., Subbanna, N.K., Szekely, G., Taylor, T.J., Thomas, O.M., Tustison, N.J., Unal, G., Vasseur, F., Wintermark, M., Ye, D.H., Zhao, L., Zhao, B., Zikic, D., Prastawa, M., Reyes, M., Van Leemput, K.: The multimodal brain tumor image segmentation benchmark (brats). IEEE Transactions on Medical Imaging **34**(10), 1993–2024 (2015). https://doi.org/10.1109/TMI.2014.2377694
16. OpenAI: gpt-oss-120b gpt-oss-20b model card (2025), https://arxiv.org/abs/2508.10925
17. Rokuss, M., Langenberg, M., Kirchhoff, Y., Isensee, F., Hamm, B., Ulrich, C., Regnery, S., Bauer, L., Katsigiannopulos, E., Norajitra, T., Maier-Hein, K.: Voxtell: Free-text promptable universal 3d medical image segmentation (2025), https://arxiv.org/abs/2511.11450
18. Schmidgall, S., Su, Y., Wang, Z., Sun, X., Wu, J., Yu, X., Liu, J., Moor, M., Liu, Z., Barsoum, E.: Agent laboratory: Using llm agents as research assistants (2025), local PDF: Agent Laboratory_ Using LLM Agents as Research Assistants.pdf
19. van Timmeren, J.E., Cester, D., Tanadini-Lang, S., Alkadhi, H., Baessler, B.: Radiomics in medical imaging—"how-to" guide and critical reflection. Insights into Imaging **11**(1), 91 (2020). https://doi.org/10.1186/s13244-020-00887-2
20. Tzanis, E., Klontzas, M.E.: maistro: an open-source multi-agentic system for automated end-to-end development of radiomics and deep learning models for medical imaging (2025), local PDF: mAistro.pdf
21. Xiaomi MiMo Team: Mimo-v2.5: Native omnimodal foundation model with strong agent capabilities. https://mimo.mi.com (2026), official technical release and model documentation
22. Yu, Y., Huang, Z., Mu, L., Zhang, S., Zhang, X.: Radiologist copilot: An agentic assistant with orchestrated tools for radiology reporting with quality control (2025), local PDF: 3D Agentic - Radiologist Copiolt.pdf
23. Ziegler, E., Urban, T., Brown, D., Petts, J., Pieper, S.D., Lewis, R., Hafey, C., Harris, G.J.: Open health imaging foundation viewer: An extensible open-source framework for building web-based imaging applications to support cancer research. JCO Clinical Cancer Informatics (4), 336–345 (2020). https://doi.org/10.1200/CCI.19.00131