Bronchoscopy as a Spatial Reasoning Problem: A Position on XR and AI as Cognitive Infrastructure in the Operating Room

Ezequiel R. Zorzal[*]
ICT/UNIFESP, Universidade Federal de São Paulo
São José dos Campos, Brazil

Catarina Moreira[†]
Data Science Institute, University of Technology Sydney
Sydney, Australia

Joaquim Jorge[‡]
INESC-ID/Instituto Superior Técnico, Universidade de Lisboa
Lisboa, Portugal

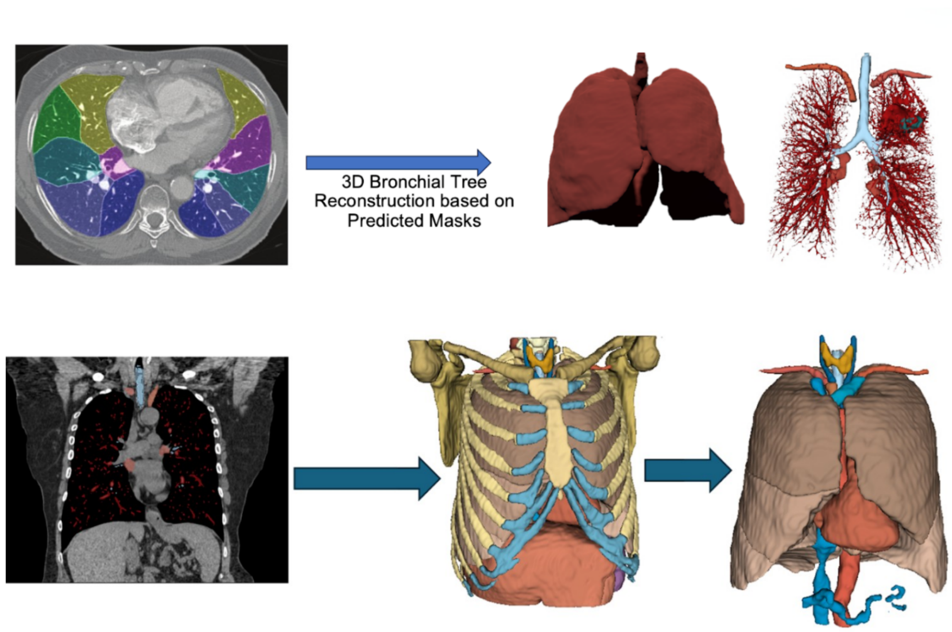

Figure 1: Virtual Bronchoscopy: a Segmentation and Planning Approach.

Abstract

Bronchoscopic procedures targeting peripheral lung lesions pose significant challenges due to patient-specific airway anatomy, uncertainty in lesion accessibility, and high cognitive demands on clinicians. While Artificial Intelligence (AI) has enabled reliable segmentation of the bronchial tree from computed tomography (CT) data, and commercial systems provide intraoperative navigation support, current workflows remain fragmented and offer limited assistance for pre-procedural planning and spatial reasoning, motivating us to frame bronchoscopy as a spatial and topological reasoning problem and to propose immersive XR combined with assistive AI as cognitive infrastructure in the operating room (Figure 1). In this paper, we argue that bronchoscopy should be viewed as a spatial and topological reasoning problem and that Extended Reality (XR), combined with assistive AI, can act as cognitive infrastructure in the operating room. We outline how immersive, patient-specific XR environments can support bronchoscopic planning, lesion localization, and procedural rehearsal while preserving clinical decision-making authority and maintaining a low cognitive load. We further identify open research challenges and opportunities for integrating XR and AI into bronchoscopic workflows to improve procedural confidence, safety, and effectiveness.

Index Terms: Bronchoscopy, Operating Room, Extended Reality, Artificial Intelligence

1 Introduction

Lung cancer remains the leading cause of cancer-related mortality globally, driving a critical clinical imperative for early diagnosis through minimally invasive interventions [3]. Consequently, modern flexible bronchoscopy has evolved to target increasingly peripheral pulmonary lesions, which, while offering the highest potential for curative treatment, present significant accessibility challenges due to the intricate and variable nature of patient-specific airway anatomy [9]. The efficacy of these procedures is fundamentally constrained by the high cognitive demands placed on the clinician, who must navigate a complex bronchial "maze" where segments

*e-mail: ezorzal@unifesp.br
†e-mail: catarina.pintomoreira@uts.edu.au
‡e-mail: jorgej@tecnico.ulisboa.pt

often appear visually indistinguishable [7]. Traditionally, this task requires the bronchoscopist to mentally reconstruct three-dimensional (3D) anatomical structures from two-dimensional (2D) CT slices, a process that relies heavily on spatial ability and experience and is prone to interpretation errors [9]. This disconnect between the 2D planning data and the dynamic 3D operative environment contributes to inter-operator variability and limits diagnostic completeness, even among experienced practitioners [7].

While technological advancements such as electromagnetic navigation and robotic-assisted bronchoscopy have provided valuable intraoperative guidance, a critical distinction must be drawn between navigation and spatial planning. Current commercial systems primarily function as execution aids, offering turn-by-turn guidance during the procedure, yet they often fail to support the pre-procedural cognitive process required to fully comprehend the topology of the patient's specific anatomy [9, 11]. Navigation assists in following a path, whereas planning involves the strategic assessment of lesion accessibility, the identification of anatomical barriers, and the selection of optimal approaches before the procedure begins. Existing workflows remain fragmented, frequently isolating these planning phases from the operative context and providing limited support for the spatial reasoning required to effectively target distal airways.

We are implementing a modular XR system that integrates AI-driven bronchial-branch segmentation with interactive visualization on head-mounted displays. The pipeline consists of three key components: (1) CT data preprocessing and segmentation via pre-trained foundation models[14] on annotated airway datasets; (2) automatic inference of lesion-to-trachea paths using graph traversal over fused segmentation and voxel evidence, explicitly designed to tolerate missing airway segments; and (3) immersive visualization within Unity3D. AI algorithms now enable robust, automated segmentation of the bronchial tree from CT data, creating detailed anatomical maps that can be visualized immersively through XR devices to enhance depth perception and spatial understanding beyond what standard monitors allow [6]. However, segmentation alone is insufficient in distal and pediatric airways, where limited spatial resolution, respiratory motion, and partial-volume effects frequently break topological continuity precisely in the regions most critical for lesion targeting. In such cases, planning fails not because anatomy is absent, but because connectivity is uncertain. We therefore argue that bronchoscopic planning requires task-conditioned topological inference: the ability to reason about plausible anatomical paths between a lesion and the central airway tree, even when explicit airway segments are missing. These limitations are particularly acute in pediatric bronchoscopy, where airway diameters approach imaging resolution limits and missing segments are common rather than exceptional.

Our approach spans the full procedural continuum, from training to planning to intraoperative assistance. In the training phase, we aim to enhance spatial reasoning and anatomical familiarity through immersive, patient-specific simulations that allow trainees to explore airway structures and practice navigation strategies in a risk-free environment. In the pre-operative planning phase, our focus is on reducing inter-operator variability and improving lesion targeting accuracy by enabling clinicians to assess accessibility, identify anatomical challenges, and rehearse procedures within a high-fidelity XR interface. Finally, while still under development, our vision for intraoperative assistance includes real-time, heads-up visualization of segmented anatomy and AI-suggested pathways, designed to integrate seamlessly into the clinical workflow without disrupting the sterile field. By addressing all three phases as part of a unified cognitive support framework, we aim to standardize performance, improve safety, and accelerate learning curves in bronchoscopy [8, 12]. Finally, we outline key challenges and opportunities for embedding this cognitive infrastructure into clinical workflows to improve effectiveness, and clinician confidence.

2 Background and Current Practice

AI advancements have revolutionized medical imaging by enabling deep learning algorithms to rapidly convert CT data into high-fidelity "digital twins" of patient anatomy, including bronchial trees and pulmonary vasculature [11, 13]. In practice, this data is operationalized through robotic platforms and electromagnetic navigation systems, which provide virtual pathways to guide clinicians toward peripheral nodules [6, 1]. However, a significant disconnect persists; the pre-procedural planning phase remains decoupled from these intraoperative tools, forcing clinicians to rely on the manual mental reconstruction of 3D relationships from 2D slices [11, 6].

This fragmentation is exacerbated by the fact that current commercial systems prioritize trajectory execution over genuine cognitive support. Much like a GPS, they offer turn-by-turn guidance without enhancing the clinician's intrinsic topological understanding of the anatomy [12]. Furthermore, these visualizations often ignore critical factors like tissue deformation and the divergence between static pre-operative images and the dynamic operative environment, potentially leading to navigational errors due to a lack of real-time uncertainty estimation [5].

3 Bronchoscopy as a Spatial Reasoning Problem

Navigating the tracheobronchial tree requires intense spatial and topological reasoning, as clinicians must correlate a limited endoscopic view with complex, bifurcating anatomy susceptible to respiratory deformation. The procedure's success depends on maintaining a robust mental model of sequential bifurcations to reach peripheral targets [4]. Currently, the "mental reconstruction" of 2D CT slices into a 3D path imposes a high cognitive load, forcing operators to shift focus between the patient and multiple 2D monitors [6]. This format necessitates constant mental rotation to align exocentric planning models with the egocentric operative view, creating a spatial reasoning bottleneck [9, 11].

XR and advanced HCI paradigms address this by externalizing spatial reasoning through Head-Mounted Displays (HMDs), merging anatomical maps directly with the operative field [6]. This shift provides intuitive depth perception and anatomical clarity unattainable via 2D screens [9]. Supported by AI-driven segmentation, the system functions as a "cognitive scaffold," offloading the burden of tracking airway segments. By transforming bronchoscopy from a memory-intensive mapping task into a supported navigational process, this integration standardizes performance and allows clinicians to prioritize procedural safety and diagnostic accuracy [1].

4 Open Research Challenges

Despite promising advances in integrating XR and AI into bronchoscopic workflows, key research challenges remain before these technologies can become part of routine clinical care. A central technical limitation is the inadequate handling of anatomical uncertainty, as current navigation systems rely on static pre-operative models that do not reflect respiratory motion or tissue deformation caused by bronchoscope interaction [5]. This discrepancy creates a false sense of security, necessitating the development of probabilistic visualization models that can communicate registration errors and potential anatomical shifts to the clinician in real-time, effectively mitigating the risks associated with "brain shift" phenomena observed in other neurosurgical contexts [5]. Furthermore, robustly linking this 3D anatomical model with the live endoscopic video feed presents significant computer vision challenges. While techniques like Simultaneous Localization and Mapping (SLAM) and Neural Radiance Fields (NeRF) show promise for markerless tracking [15], maintaining accurate registration in the presence of visual artifacts such as mucus, blood, or bubbles remains an open problem [2]. Future work must develop AI-driven registration resilient to intraoperative occlusions, ensuring alignment between the anatomical "map" and operative "territory".

Beyond technical accuracy, the integration of XR introduces complex human factors that require rigorous evaluation, particularly regarding cognitive load. While the goal of XR is to function as cognitive infrastructure, there is a risk that overlaying excessive information can lead to "crowding" of the visual field or "attention bias," in which the operator focuses more on the digital interface than on the patient [5]. More granular, real-time assessment methods—potentially leveraging eye-tracking or biometric sensors—are needed to ensure that these systems reduce rather than exacerbate mental demand during critical procedural steps. Concurrently, evaluation metrics must move beyond diagnostic yield or procedure time. Recent studies highlight process-oriented measures, such as Structured Progress (SP) and Mean Intersegmental Time (MIT), which capture the efficiency and systematic quality of navigation rather than only final outcomes [3]. Adopting such metrics in clinical trials will be essential to demonstrate that AI and XR not only help reach the target but also standardize the navigation.

Finally, the widespread acceptance and adoption of these technologies hinge on overcoming practical implementation barriers. Ergonomic limitations, such as the weight of HMDs causing discomfort during prolonged procedures, and the limited battery life of current hardware, continue to hinder routine use [6]. Moreover, the high cost of advanced XR equipment and the steep learning curve associated with mastering gesture or voice-controlled interfaces present significant hurdles for resource-limited settings [12]. Future work must prioritize the development of lightweight, cost-effective hardware and intuitive user interfaces that integrate seamlessly into the sterile operating room workflow, ensuring that the cognitive benefits of XR and AI are accessible to a broader range of clinicians and institutions [1],

5 Discussion and Outlook

The position's relevance to the XR4OR community lies in reframing bronchoscopy as a challenge of spatial reasoning rather than just manual dexterity. By replacing 2D cross-sections with 3D holographic structures, we align the operative interface with innate spatial cognition, reducing the mental effort required to translate radiological data into motion [9]. This establishes XR as a critical bridge between preoperative planning and intraoperative reality, enabling clinicians to identify targets more accurately and refine surgical strategies in real time [11].

The success of this paradigm depends on the intersection of AI and HCI. AI acts as the computational engine, automating the segmentation of complex airways to create "digital twins" that standardize performance across experience levels [3]. However, adoption relies on HCI designs that respect the sterile field and ergonomic constraints; while HMDs allow hands-free interaction, barriers such as weight, battery life, and visual fatigue must be addressed to ensure routine clinical viability [6].

Future research must pivot toward the dynamic intraoperative environment, specifically addressing tissue deformation caused by respiration and instrument contact [11]. Developing real-time, AI-driven registration algorithms is essential to keep the 3D "map" aligned with the changing "territory" [15]. Furthermore, evaluation must move beyond diagnostic yield to quantify cognitive load reduction and situational awareness [1]. By fostering collaboration and open-source validation protocols, we can transform AI and XR into essential cognitive infrastructure that democratizes advanced surgical skills and enhances clinical precision [5, 15, 10]. By shifting the focus from complete anatomical reconstruction to task-conditioned topological reasoning, XR systems can better reflect how bronchoscopists think under uncertainty, rather than forcing them to adapt to brittle computational models.

Disclosure

During the preparation of this work, the author(s) used AI Tools including ChatGPT 5.2 and Gemini 3, for language refinement and improving clarity. After using these tools, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

Acknowledgments

This work was partially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) under Grants 423521/2021-7 and 304588/2025-3, and by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) under Grant 2025/08366-1, and also from the Portuguese Science Foundation (FCT), under grant 2022.09212.PTDC, and under the auspices of the UNESCO Chair on AI&XR.

References

[1] K. O. Agbontaen et al. Artificial Intelligence-Guided Bronchoscopy is Superior to Human Expert Instruction for the Performance of Critical-Care Physicians: A Randomized Controlled Trial. *Critical Care Medicine*, 53, 2025. doi: 10.1097/ccm.0000000000006629 2, 3

[2] J.-C. Chien, J.-D. Lee, E. Su, and S.-H. Li. A Bronchoscope Localization Method Using an Augmented Reality Co-Display of Real Bronchoscopy Images with a Virtual 3D Bronchial Tree Model. *Sensors*, 20, 2020. doi: 10.3390/s20236997 2

[3] K. M. Cold, K. Agbontaen, A. O. Nielsen, C. S. Andersen, S. Singh, and L. Konge. Artificial intelligence improves bronchoscopy performance: a randomised crossover trial. *ERJ Open Research*, 11, 2024. doi: 10.1183/23120541.00395-2024 1, 3

[4] K. M. Cold, S. Xie, A. O. Nielsen, P. F. Clementsen, and L. Konge. Artificial Intelligence Improves Novices' Bronchoscopy Performance. *CHEST*, 165, 2024. doi: 10.1016/j.chest.2023.08.015 2

[5] I. Isikay, E. Cekic, B. Baylarov, O. Tunc, and S. Hanalioglu. Narrative review of patient-specific 3D visualization and reality technologies in skull base neurosurgery. *Frontiers in Surgery*, 11, 2024. doi: 10.3389/fsurg.2024.1427844 2, 3

[6] A. Kildahl-Andersen et al. Bronchoscopy using a HM MR device—a phantom study and a first in-patient user experience. *Frontiers in Virtual Reality*, 4, 2023. doi: 10.3389/frvir.2023.940536 2, 3

[7] V. Mehta. Artificial intelligence augmentation raises questions about the future of bronchoscopy. *ERJ Open Research*, 11, 2025. doi: 10.1183/23120541.00931-2024 2

[8] C. A.-B. Nielsen, L. Lönn, L. Konge, and M. Taudorf. Simulation-Based Virtual-Reality Patient-Specific Rehearsal Prior to Endovascular Procedures: A Systematic Review. *Diagnostics*, 10, 2020. doi: 10.3390/diagnostics10070500 2

[9] S. Okachi et al. Clinical Applicability of Three-Dimensional Holographic Virtual Bronchoscopy with Mixed Reality. *Annals of the American Thoracic Society*, 22, 2025. doi: 10.1513/annalsats.202407-729cc 1, 2, 3

[10] A. Sachdeva and S. Sethi. Motivation and Learning. *CHEST*, 165, 2024. doi: 10.1016/j.chest.2023.09.018 3

[11] A. H. Sadeghi et al. Virtual reality and artificial intelligence for 3-dimensional planning of lung segmentectomies. *JTCVS Techniques*, 7, 2021. doi: 10.1016/j.xjtc.2021.03.016 2, 3

[12] B. Tian et al. Bridging traditional gaps in medical training with mixed reality: insights from bronchoscopy education—a narrative review. *Journal of Thoracic Disease*, 17, 2025. doi: 10.21037/jtd-2025-565 2, 3

[13] X. Wang, J. Yang, B. Zhou, L. Tang, and Y. Liang. Integrating mixed reality, augmented reality, and artificial intelligence in complex liver surgeries: Enhancing precision, safety, and outcomes. *iLIVER*, 4, 2025. doi: 10.1016/j.iliver.2025.100167 2

[14] J. e. a. Wasserthal. TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images. *Radiology: Artificial Intelligence*, 2023. doi: 10.1148/ryai.230024 2

[15] S. Winiarski et al. Integrating Artificial Intelligence in Bronchoscopy and Endobronchial Ultrasound (EBUS) for Lung Cancer Diagnosis and Staging. *Cancers*, 17, 2025. doi: 10.3390/cancers17172835 2, 3