

A PACS-Integrated Artificial Intelligence Platform for Visualizing Radiolucent Foreign Body Aspiration

Xiaofan Liu

Central Hospital of Wuhan

Zhe Chen

University of Southampton

Jixiang Ni

Central Hospital of Wuhan

Zhiyong Tang

University of Southampton

Yan Jiang

Central Hospital of Wuhan

Kai Wei

Central Hospital of Wuhan

Qilin Chang

Jiangnan University

Jingwen Wang

Wuhan University of Science and Technology

Liangchao Wang

Central Hospital of Wuhan

Fang Ni

Central Hospital of Wuhan

Qiong Qian

Central Hospital of Wuhan

Si Fang

Central Hospital of Wuhan

Junjie Xu

Central Hospital of Wuhan

Yin Wang

Central Hospital of Wuhan

Manman Hu

Central Hospital of Wuhan

Dian Quan

Central Hospital of Wuhan

Lei Li

Central Hospital of Wuhan

Rob M. Ewing

University of Southampton

Zehor Belkhatir

University of Southampton

Guqin Zhang

Zhongnan Hospital of Wuhan University

Hanxiang Nie

Renmin Hospital of Wuhan University

Yi Hu

Central Hospital of Wuhan

Xun Yang

Central Hospital of Wuhan

Yihua Wang

yihua.wang@soton.ac.uk

University of Southampton

Research Article

Keywords: Foreign body aspiration (FBA) Deep learning, Self-supervised anomaly detection, Medical image segmentation, Computer-aided diagnosis (CAD)

Posted Date: October 14th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10991925/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

A PACS-Integrated Artificial Intelligence Platform for Visualizing Radiolucent Foreign Body Aspiration

Xiaofan Liu^{1,*}, Zhe Chen^{2,*}, Jixiang Ni^{1,*}, Zhiyong Tang², Yan Jiang³, Kai Wei⁴, Qilin Chang⁵, Jingwen Wang⁶, Liangchao Wang¹, Fang Ni¹, Qiong Qian¹, Si Fang¹, Junjie Xu¹, Yin Wang¹, Manman Hu¹, Dian Quan¹, Lei Li¹, Rob M. Ewing^{2,7}, Zehor Belkhatir^{8,#}, Guqin Zhang^{9,#}, Hanxiang Nie^{10,#}, Yi Hu^{1,#}, Xun Yang^{1,#}, Yihua Wang^{2,7,11,#†}

¹Department of Pulmonary and Critical Care Medicine, The Central Hospital of Wuhan, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei 430014, China; ²Biological Sciences, Faculty of Environmental and Life Sciences, University of Southampton, Southampton SO17 1BJ, UK; ³Department of Radiology, The Central Hospital of Wuhan, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei 430014, China; ⁴Department of Gastrointestinal Surgery, The Central Hospital of Wuhan, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei 430014, China. ⁵School of Medicine, Jiangnan University, Wuhan, Hubei 430056, China; ⁶School of Medicine, Wuhan University of Science and Technology, Wuhan, Hubei 430081, China. ⁷Institute for Life Sciences, University of Southampton, Southampton SO17 1BJ, UK; ⁸School of Electronics and Computer Science, Digital Health & Biomedical Engineering Group, University of Southampton, Southampton SO17 1BJ, UK; ⁹Department of Respiratory and Critical Care Medicine, Zhongnan Hospital of Wuhan University, Wuhan, Hubei 430071, China; ¹⁰Department of Respiratory and Critical Medicine, Renmin Hospital of Wuhan University, Wuhan, Hubei 430060, China; ¹¹NIHR Southampton Biomedical Research Centre, University Hospital Southampton, Southampton SO16 6YD, UK.

*These authors contributed equally.

#For correspondence: Yihua Wang: yihua.wang@soton.ac.uk; Yi Hu: huyi@zxhospital.com; Hanxiang Nie: nhxiang@whu.edu.cn; Guqin Zhang: zgqxiaoxiao@126.com; Xun Yang: 78719976@qq.com; Zehor Belkhatir: z.belkhatir@soton.ac.uk

†Lead contact

ABSTRACT:

Diagnosis of radiolucent foreign body aspiration (FBA) in adults is a major clinical challenge. While previous deep learning models have successfully achieved patient-level diagnostic screening, they remain fundamentally limited as "black-box" classifiers, failing to provide the precise 3D spatial coordinates required for surgical intervention. Furthermore, the extreme scarcity of labelled FBA cases forces data-starved models to overfit, hindering broader clinical deployment. To overcome this and bridge the critical gap between radiological suspicion and surgical execution, we developed AI-PACS, a self-supervised anomaly detection and 3D visualization platform. Driven by a Lean Joint Embedding Predictive Architecture (LeJEPA) and a Sparse-Inferred Gaussian Regularization (SIGReg) module, the system is pre-trained on 804 unlabelled CT scans to establish a standard Gaussian manifold of "non-FBA" airway topology. This allows AI-PACS to identify FBA as high-norm statistical outliers with "pathological invariance," effectively filtering out confounding structural noise. Validated across multi-center internal (n=268), external (n=103), and independent (n=70) cohorts, AI-PACS demonstrated robust generalizability with diagnostic accuracies of 75.7%–82.5%. In a head-to-head benchmark against expert radiologists, AI-PACS significantly improved detection sensitivity from 18.8% to 56.3% ($P < 0.05$) and reduced the average negative exclusion screening time by 75%. By shifting the paradigm from pure diagnostic classification to automated 3D anatomical localization (S1-S10 mapping), AI-PACS acts as a digital rehearsal tool that translates invisible structural outliers into actionable interventional guidance.

Clinical trial number Not applicable.

Keywords: Foreign body aspiration (FBA) Deep learning, Self-supervised anomaly detection, Medical image segmentation, Computer-aided diagnosis (CAD)

Introduction

Foreign body aspiration (FBA) constitutes a critical diagnostic challenge in pulmonary medicine, carrying substantial risks of chronic lung injury, recurrent pneumonia, and acute asphyxiation if not promptly identified [1, 2]. While radiopaque objects are readily discernible on standard imaging, the diagnostic landscape for radiolucent FBA is far more complex in the adult population [3-5]. Our previous 10-year retrospective cohort study revealed that approximately 75% of FBA cases in adults involve radiolucent objects, such as bone fragments, food particles, or plastic materials [6]. On chest Computed Tomography (CT), these objects possess densities remarkably similar to soft tissue or mucus (approximately 30–60 HU), rendering them virtually indistinguishable to the naked eye [7]. In current clinical practice, this creates a "needle in a haystack" scenario where radiologists must manually scrutinize thousands of slices. The difficulty is further complicated by vague and nonspecific clinical presentations; patients often lack a witnessed choking history, leading to frequent misdiagnoses of pneumonia or malignancy [8].

To address these diagnostic gaps, artificial intelligence (AI) has emerged as a promising solution, initially gaining attention in paediatric imaging using chest X-rays. Truong *et al.* (2024) demonstrated the potential of machine learning to augment human performance in paediatric FBA. In their study of 566 patients, standard radiologist interpretation achieved a sensitivity of only 50.6%, whereas their machine learning algorithm improved this to 66.7%, highlighting the capacity of AI to detect subtle patterns in single-view radiographs [9]. More recently, Çoşkun *et al.* (2025) applied Convolutional Neural Networks (CNN) to reduce the rate of unnecessary invasive procedures. Their deep learning model achieved a remarkable sensitivity of 94.1% and specificity of 97.8% in predicting bronchoscopy necessity, significantly outperforming multiple logistic regression models [10]. These studies collectively suggest the feasibility of AI in FBA screening. However, they predominantly focus on paediatric populations and 2D planar imaging, which are insufficient for resolving the complex, overlapping anatomy of adult airways in 3D. Building on these foundations, our recent study pioneered a supervised deep learning framework specifically for adult CT scans. By integrating high-precision

airway segmentation (MedpSeg) with a multi-view convolutional classifier (ResNet-18), it was successfully demonstrated that AI could significantly improve detection sensitivity from 35.7% to 71.4% compared to expert radiologists, while maintaining a robust specificity of 94.6%[11]. This work marked a milestone in proving that deep neural network can perceive subtle morphological signatures of radiolucent objects in 3D space that typically escape human vision.

Building on these foundations, our recent study pioneered a deep learning framework specifically for adult CT scans. By utilizing a convolutional classifier (2D ResNet-18), it successfully demonstrated that AI could achieve high accuracy in patient-level FBA screening, significantly improving detection sensitivity compared to expert radiologists¹¹. While this model showed immense promise as a first-line diagnostic alert, its clinical utility for downstream surgical intervention remains fundamentally limited. Traditional classifiers operate largely as "black boxes"; by compressing complex 3D airway topology into global probability scores, they inherently lose the spatial coordinates necessary for pre-operative bronchoscopy planning. Clinical interventionalists require more than just a binary "yes/no" diagnosis—they need precise anatomical localization within the tracheobronchial tree (S1-S10 segments) to safely navigate and retrieve the aspirated object. Furthermore, achieving robust spatial localization is severely bottlenecked by the "unbalanced reality" of FBA data—an extreme scarcity of labelled cases that forces data-starved supervised classifiers to overfit to confounding structural noise.

To bridge the critical gap between radiological screening and surgical intervention, we propose a definitive shift from pure diagnostic classification to precise 3D visualization and lesion-level localization. To overcome the data starvation that plagues this localization task, we shifted to a self-supervised anomaly detection paradigm. Our central hypothesis is that the anatomical features of the "non-FBA" airway maintain a consistent topological manifold. Rather than relying on scarce labelled cases, we conceptualize a model that mathematically defines this baseline "normality" by pre-training on a large cohort of unlabelled CT scans. Under this framework, the system learns the standard geometry of the healthy airway and subsequently identifies foreign bodies as statistical outliers that deviate from this learned norm. This approach uniquely allows the model to achieve "pathological invariance," filtering

102 out confounding structural noise from chronic diseases that typically triggers false positives in
103 conventional classifiers. Furthermore, we developed AI-PACS, an automated clinical visualization
104 platform. By translating these "invisible" structural outliers into high-intensity heatmaps mapped to the
105 standardized segmental nomenclature (S1–S10) of the tracheobronchial tree, the AI-PACS platform acts
106 as a digital rehearsal tool. It not only serves as a high-throughput diagnostic safeguard but also provides
107 actionable spatial coordinates, potentially redefining the standard of care from initial suspicion to
108 successful surgical retrieval.

109

Study Design and Methods

Study Design and Ethical Approval

This retrospective, multi-center diagnostic accuracy study was conducted in strict adherence to the Standards for Reporting Diagnostic Accuracy Studies (STARD) guidelines [12]. The research protocol and patient characteristics for internal, external and independent cohorts utilized in this work were reported earlier [11]. Full ethical approval was sustained by the Institutional Review Boards (IRBs) of all participating clinical centers: The Central Hospital of Wuhan (Approval Nos. WHZXKYL2024-108 and WHZXKYL2025-069), The Renmin Hospital of Wuhan University (Approval No. WDRY2025-K083), and The Zhongnan Hospital of Wuhan University (Approval No. 2025086K). Consistent with our previous methodology [11], the requirement for written informed consent was formally waived by the respective ethics committees due to the retrospective nature of the study and the rigorous anonymization of all patient imaging data prior to analysis.

Study Cohorts and Data Acquisition

To address the inherent data scarcity associated with foreign body aspiration, we designed a comprehensive dual-phase data acquisition strategy that decouples representation learning from clinical validation.

For the self-supervised pre-training phase (Phase I), we constructed a large-scale "normality" manifold by aggregating 804 unlabeled chest CT scans from three distinct sources. Rather than treating this as a monolithic dataset, each source was strategically selected to impart a specific dimension of generalizability to the model: anatomical, clinical, and technical. First, to establish an anatomical baseline, we utilized 254 scans from the healthy control subset of the public MosMedData repository ^[13], teaching the model the standard, disease-free geometry of the tracheobronchial tree. Second, to cultivate clinical adaptability (achieving "pathological invariance"), we complemented this with a private cohort of 250 randomized, non-FBA patients from The Central Hospital of Wuhan. This cohort provided a wide spectrum of non-specific respiratory presentations, forcing the model to generalize across physiological noise by learning to distinguish common chronic airway variations (e.g., age-related remodelling or

inflammation) from acute, out-of-distribution anomalies. Finally, to ensure technical invariance against varying imaging parameters, we integrated 300 multi-center scans from the ATM22 challenge dataset [14-18]. Given that the ATM22 corpus was curated primarily for airway structural modeling from non-emergency cohorts (e.g., routine screening and chronic lung diseases), these scans were fundamentally devoid of acute foreign body obstructions, ensuring the purity of the pre-training manifold across diverse CT scanners and acquisition protocols. Crucially, all scans in this pre-training corpus were paired with high-fidelity, voxel-level ground-truth airway masks, serving as a strict anatomical prior that confined the self-supervised learning process exclusively to verified airway regions without introducing segmentation noise.

Following pre-training, the model's clinical performance was rigorously evaluated on a separate validation dataset comprising 441 patients organized into three retrospective cohorts, summarized in Table 1, to test real-world generalizability. The primary evaluation was conducted on an Internal Cohort (n=268) from The Central Hospital of Wuhan, consisting of 41 confirmed radiolucent FBA cases and 227 non-FBA (NFBA) controls. This was followed by validation on an External Cohort (n=103) from The Renmin Hospital of Wuhan University, comprising 21 FBA cases and 82 NFBA controls, to assess cross-center robustness. Finally, a strictly held-out Independent Cohort (n=70) from The Zhongnan Hospital of Wuhan University, containing 14 FBA cases and 56 NFBA controls, was utilized to evaluate the platform's cross-center generalizability and conduct comprehensive ablation analyses within a heterogeneous clinical environment. Patients characteristics for internal, external and independent cohorts were reported earlier [11]. All validation images were acquired using standard helical CT protocols with a slice thickness of < 1.25 mm. Flexible bronchoscopy reports served as the definitive ground truth for confirming the presence, type, and precise anatomical location of foreign bodies for all diagnostic and localization metrics. Collectively, these datasets encompass diverse patient populations from three independent tertiary centers, enhancing the translational relevance of the findings and ensuring the platform's utility across varied institutional workflows.

To rigorously benchmark the AI-PACS platform against standard clinical practice (Phase III), a dedicated human-AI comparative evaluation subset was established through random sampling from the

internal validation cohort at The Central Hospital of Wuhan. This subset comprised 35 patients, specifically partitioned into a confirmed FBA group (n=16) and a non-FBA (NFBA) control group (n=19). For this cohort, detailed baseline clinical and radiological characteristics were systematically extracted to assess cohort comparability and task difficulty, encompassing demographic data (age, sex, BMI), clinical history (length of disease course, hospital and ICU stay, chronic respiratory or neurological diseases, and smoking history), and specific morphological presentations on multi-detector computed tomography (MDCT). This specialized subset served as the basis for a head-to-head performance comparison between the algorithmic pipeline and an expert radiologist panel. Consistent with the protocols of all study phases, flexible bronchoscopy reports remained the definitive ground truth for confirming the presence, morphological type, and precise anatomical location of foreign bodies.

System Overview and End-to-End Pipeline

The AI-PACS platform operates as a fully automated, end-to-end diagnostic pipeline designed to transition from raw volumetric imaging to clinical decision support. As illustrated in Fig. 1, the workflow comprises three primary functional stages: (1) Anatomical Prior Integration, where a dedicated 3D U-Net extracts the tracheobronchial tree to mask out extra-pulmonary noise (Fig. 1c); (2) Feature Encoding and Anomaly Scoring, where the pre-trained, frozen LeJEPa encoder projects airway patches into the learned normality manifold to identify statistical outliers (Fig. 1b); and (3) Spatial Localization and Visualization, where anomaly scores are back-projected onto the 3D airway mesh and registered to the S1–S10 segmental nomenclature to guide bronchoscopic intervention (Fig. 1d). This integrated approach ensures that the detection process is both mathematically rigorous and anatomically grounded.

Anatomical Prior and Mask-Guided Pre-processing

To mitigate scanner-specific variations across heterogeneous data sources, all raw volumetric data underwent a rigorous standardization pipeline prior to feature encoding. All CT sequences were first resampled to a uniform isotropic resolution of $1.0 \times 1.0 \times 1.0 \text{ mm}^3$ using tricubic interpolation to preserve the integrity of fine bronchial structures. Voxel intensities were subsequently normalized via a

static "Lung Window" clipping strategy (range: -1000 HU to +400 HU) and linearly scaled to the [0, 1] interval, thereby enhancing the contrast between the radiolucent airway lumen and the surrounding bronchial walls. To enforce the "Anatomical Prior," we implemented a spatial filtering mechanism designed to constrain the model's receptive field exclusively to the tracheobronchial tree. Mathematically, this was achieved via the element-wise multiplication (Hadamard product) of the normalized CT volume (X) and a binary airway mask (M), yielding a masked volume $X_{masked} = X \odot M$. This operation effectively zeroes out high-variance background signals—including pulmonary parenchyma, vascular trees, and bony thorax artifacts—ensuring that the subsequent LeJEPA encoder focuses solely on the relevant topological features of the airway manifold. Crucially, the source of the mask (M) differed by experimental phase to ensure rigor: during the self-supervised pre-training (Phase I), we utilized the high-fidelity ground-truth masks provided by the MosMedData and ATM22 datasets to generate a noise-free training signal. Conversely, during the clinical inference phase (Phase II) where ground truth is unavailable, the mask was generated dynamically by a specialized U-Net architecture implemented within the MedpSeg framework, as previously described [11]. This module was specifically optimized for high-precision airway extraction to provide the necessary anatomical prior for FBA detection.

LeJEPA Architecture and Manifold Contraction

The core anomaly detection engine is underpinned by a Lean Joint Embedding Predictive Architecture (LeJEPA), instantiated with a Vision Transformer (ViT-B/16) backbone [19]. Unlike generative approaches that expend capacity reconstructing high-frequency pixel noise, LeJEPA operates exclusively within the abstract latent space to capture semantic topology. The technical mechanism for this detection capability is detailed in Fig. 2. The network processes the masked airway volume as a sequence of non-overlapping patches, projecting them via dual encoders (Context and Target) into a high-dimensional embedding space (Fig. 2a). We designed the learning objective around a "context-target" paradigm: the encoder processes a subset of visible "context" patches to predict the latent representations of masked "target" regions.

To formally implement the "Manifold Contraction" hypothesis, we introduced a regularization objective termed SIGReg [20]. While the predictive loss minimizes the distance between the predicted and actual embeddings, the SIGReg component explicitly constrains the geometry of the latent space. Mathematically, it enforces two critical conditions: (1) Variance regularization, which prevents the hypersphere from collapsing to a single point by maintaining amplitude on the diagonal; and (2) Covariance regularization, which minimizes off-diagonal correlations to zero. As illustrated in Fig. 3b, this dual mechanism exerts opposing geometric forces—expanding useful variance while compressing redundancy—forcing the latent distribution of all "normal" airway features to collapse into a standard isotropic Gaussian hypersphere $N(0, I)$, with zero mean and unit variance. By compacting the "normality" manifold into this predictable statistical structure, the network maximizes the entropic distance to out-of-distribution events. Consequently, radiolucent foreign bodies resist this compression and manifest as distinct outliers with high embedding norms (Fig. 2b, Stage 3).

Inference and Automated Localization Pipeline

In the clinical inference phase (Phase II), the system operates in a fully automated, end-to-end manner designed for high-throughput triage without manual intervention. For a query scan, airway extraction is performed using the modified 3D U-Net architecture (MedpSeg) that was explicitly developed and validated in our previous study [11]. While our prior framework incorporated a human-in-the-loop refinement step, the AI-PACS platform deploys this validated network in a strictly automated mode. This lightweight, supervision-guided module generates a binary airway mask, isolating the tracheobronchial tree from the pulmonary parenchyma to ensure analysis is confined to the region of interest. The masked volume is then tokenized into non-overlapping patches and processed by the frozen LeJEPA encoder, which preserves the "normality manifold" established during Phase I.

Anomaly Scoring and Statistical Thresholding

The transition from abstract latent embeddings to a binary clinical diagnosis is governed by a precise geometric and statistical pipeline. In the inference phase, the frozen LeJEPA encoder projects tokenized

airway patches x_i into the high-dimensional latent space. Because the SIGReg loss explicitly constrains all "normal" anatomical features to collapse into a standard isotropic Gaussian hypersphere $N(0, I)$ centered at the origin, the degree of abnormality is represented by the Euclidean distance from this origin. The Anomaly Score $A(x_i)$ is calculated as the L_2 norm of the latent embedding vector z_i :

$$A(x_i) = \|z_i\|_2 = \sqrt{\sum_{j=1}^D z_{i,j}^2},$$

where D represents the dimensionality of the embedding space. Under this manifold contraction paradigm, healthy airway tissues yield embeddings with consistently low norms clustered near the origin, whereas radiolucent foreign bodies—representing out-of-distribution (OOD) morphological structures—resist compression and manifest as high-magnitude outliers.

To define a clinically actionable decision boundary, a global diagnostic threshold (τ) was established. This "threshold line" was optimized exclusively using the Internal Validation Cohort ($n=268$) to prevent data leakage and preserve the blind integrity of the external and independent sets. We utilized a Receiver Operating Characteristic (ROC) analysis to identify the τ that maximizes Youden's J statistic:

$$J(\tau) = \text{Sensitivity}(\tau) + \text{Specificity}(\tau) - 1.$$

This optimization ensures a mathematically balanced trade-off between the sensitivity required to detect occult objects and the specificity needed to minimize false-positive surgical explorations. Once τ was fixed based on internal data, it was applied as a rigid criterion across the External and Independent cohorts. A detection is flagged if $A(x_i) > \tau$. To translate these scalar scores into 3D heatmaps, a Gaussian smoothing kernel G_σ was applied to the back-projected scores to generate a continuous volumetric intensity field H :

$$H = G_\sigma * [A(x_i) \cdot 1_{(A(x_i) > \tau)}],$$

where $1_{(A(x_i) > \tau)}$ serves as a spatial binary indicator function—evaluating to 1 if the localized anomaly score exceeds the threshold τ , and 0 otherwise—to strictly suppress physiological background noise. The

operator $\cdot$ denotes element-wise (point-wise) multiplication, retaining only the magnitude of significant outliers. Where $*$ denotes the 3D spatial convolution operator.

G_σ defines the isotropic 3D Gaussian smoothing kernel, which is mathematically formulated with respect to the standard deviation σ :

$$G_\sigma(\mathbf{v}) = \frac{1}{(2\pi\sigma^2)^{\frac{3}{2}}} \exp\left(-\frac{\|\mathbf{v}\|_2^2}{2\sigma^2}\right).$$

Here, $\mathbf{v}$ represents the 3D spatial displacement vector from the kernel center. The parameter σ , empirically set to 1.5 mm, dictates the spatial spread of the distribution. Applying this convolution effectively diffuses the sparse, discrete high-scoring patches into a smooth, continuous volumetric intensity field, providing an intuitive visual gradient for clinical interpretation. This heatmap is registered to the bronchial tree via centerline extraction, enabling automated localization from the trachea down to the segmental bronchi (S1–S10). To establish a standardized spatial reference, the system maps these detections directly onto the anatomical segmentation nomenclature illustrated in Fig.4. By utilizing this comprehensive topological map, AI-PACS translates mathematical outlier scores into precise anatomical coordinates. Finally, this localization is rendered as a color-coded overlay on multi-planar reconstruction (MPR) views to provide intuitive, actionable guidance for pre-operative bronchoscopy planning.

Baseline Model Implementation

To rigorously evaluate the AI-PACS platform, our comparative framework was strategically designed to address two distinct methodological frontiers: transitioning from our previously proposed supervised paradigm to a label-free approach, and selecting the optimal self-supervised geometry. First, to benchmark the paradigm shift from our prior work, we implemented a standard 3D ResNet-18 as the supervised baseline. Second, and more importantly, it is critical to compare our system against established unsupervised and self-supervised anomaly detection paradigms. This ensures that the

observed performance gains are specifically attributable to our LeJEPA manifold contraction rather than merely the transition to a label-free framework. To achieve this, we implemented two representative unsupervised generative baselines (a 3D Denoising Autoencoder [DAE] and a 3D Variational Autoencoder [VAE]). Crucially, these unsupervised baselines did not replace the entire AI-PACS pipeline. Instead, they were integrated strictly as alternative detection engines at Stage 2 (Feature Encoding and Anomaly Scoring). To guarantee a strictly controlled comparative analysis, all generative baselines processed the exact same U-Net-generated anatomical masks from Stage 1 and operated under identical data augmentation constraints. Regarding the specific architectures, the DAE was constructed using a symmetric 3D Convolutional architecture comprising four down-sampling and four up-sampling blocks [21]. To enforce robust feature learning rather than simple identity mapping, additive Gaussian noise ($\mu = 0, \sigma_{noise} = 0.1$) was applied to the input volumes, and the model was optimized via Mean Squared Error (MSE) loss. The anomaly score for the DAE was defined as the voxel-wise L_2 reconstruction error between the input and the output volumes. Similarly, a Variational Autoencoder (VAE) was implemented with a commensurate backbone but incorporated the reparameterization trick to enforce a continuous latent space aligned with a standard Gaussian prior. Its anomaly score was computed as the Evidence Lower Bound (ELBO), aggregating both the reconstruction residual and the Kullback-Leibler divergence [22]. For a fair and direct comparison, the bottleneck dimension for both generative baselines was set to $d=256$, explicitly matching the projection capacity of our LeJEPA target encoder. All baseline models were trained for 300 epochs on the 804-scan "normality" cohort using the Adam optimizer ($lr = 10^{-4}$) on two high-performance NVIDIA A5000 GPUs.

Ablation Study Design

To dissect the individual contributions of our specific pipeline engineering, we systematically degraded the AI-PACS architecture through an ablation study. First, to assess the critical necessity of the anatomical prior, we removed the spatial filtering mechanism (w/o Anatomical Mask). This forced the

encoder to process the entire raw thoracic volume, exposing the model to high-variance background noise from pulmonary parenchyma and bones. Second, we evaluated the necessity of manifold contraction by disabling the Variance and Covariance Regularization components of the SIGReg objective (w/o SIGReg Loss). Under this degraded configuration, the dual encoders learned predictive features, but the latent space was not explicitly constrained to an isotropic Gaussian manifold $N(0, I)$. These degraded configurations were subsequently compared against the full AI-PACS pipeline integrating both the anatomical prior and manifold contraction.

Comparative Localization Study: Radiologist vs. AI Model

To rigorously benchmark the anatomical localization precision of the LeJEPa framework against clinical standards, a comparative evaluation was performed using the dedicated human-AI comparative subset (n=35), which comprised 16 confirmed FBA cases and 19 NFBA controls. A panel of three board-certified radiologists, each possessing over 10 years of clinical experience, was tasked with an independent review of this randomized subset. To ensure direct metric parity with the model's probabilistic output, the readers were instructed to record a ranked differential diagnosis for each case, identifying the three most probable anatomical segments (primary, secondary, and tertiary suspects) likely to harbor a foreign body. This human benchmark was compared against the standalone AI model's output, where the "Top 3 prediction" was algorithmically defined as the three bronchial segments (S1–S10) exhibiting the highest aggregated anomaly scores. For both the radiologist panel and the AI platform, a localization was deemed successful if the bronchoscopy-confirmed target segment was encompassed within their respective top three ranked candidates. This metric specifically quantifies the clinical utility of each modality in narrowing the effective search volume for pre-operative bronchoscopy planning.

Statistical Analysis

Evaluation metrics were computed at both the patient level (for binary diagnostic classification) and the lesion level (for anatomical localization). The primary endpoints for diagnostic performance included

Area Under the Receiver Operating Characteristic curve (AUROC), sensitivity (recall), specificity, accuracy, positive predictive value (PPV), and negative predictive value (NPV). To determine the optimal operating point on the ROC curve, we utilized Youden's J statistic derived exclusively from the internal validation cohort; this threshold was then frozen and applied blindly to external and independent test sets.

Localization Performance was evaluated using two rigorous standards against the bronchoscopy gold standard. First, we calculated Strict Top 1 Accuracy, recording a "hit" only if the centroid of the predicted heatmap fell exactly within the ground-truth bronchial segment (S1–S10). Second, recognizing the anatomical contiguity and potential registration variances at segmental boundaries, we computed Top 3 Accuracy. This metric considers a localization successful if the ground-truth segment is contained within the three segments exhibiting the highest aggregated anomaly scores. This broader metric reflects clinical utility, as narrowing the search space to a "neighborhood" of three adjacent segments is often sufficient to guide successful bronchoscopy navigation.

All computational experiments were conducted on a Linux-based workstation (Ubuntu 20.04 LTS) equipped with two NVIDIA RTX A5000 GPUs (48GB RAM) to accommodate the memory-intensive requirements of high-resolution 3D volumetric processing. The deep learning pipeline was implemented in Python (v3.8) using the PyTorch framework (v1.18.1) with CUDA (v11.3) support for hardware acceleration.

Results

Quantitative Validation of the Self-Supervised Normality Manifold

To establish a robust foundation for anomaly detection, we first quantified the representation learning performance of the LeJEPA engine and the SIGReg module during Phase I (Fig. 1). Using an aggregated normality cohort of 804 unlabelled CT scans (Table 1 - Phase I; Fig. 1a), the system was trained to define the standard topological manifold of the non-FBA airway. This diverse pre-training cohort was strategically curated to impart "pathological invariance" and structural robustness through 3 distinct and complementary data sources. To provide a baseline of anatomical integrity, 254 healthy control scans were utilized from the public MosMedData repository, ensuring the model captured standard physiological airway geometry without underlying disease. This was augmented by a private clinical cohort of 250 randomized patients from The Central Hospital of Wuhan, which was essential for teaching the model to distinguish acute anomalies from common chronic pathological changes. Unlike the healthy controls, this specific subset encompassed a wide spectrum of non-specific respiratory presentations and chronic airway variations—such as non-NFBA inflammation or age-related structural remodelling, thereby enabling the platform to achieve "pathological invariance" by filtering out confounding structural noise. Furthermore, the dataset was supported by 300 multi-center scans from the ATM22 challenge, which provided high-fidelity airway trees from multiple clinical sites and domains. The inclusion of ATM22 data, characterized by diverse scanner protocols and complex bronchial branching patterns, ensured that the learned manifold remained robust across heterogeneous institutional workflows. The stability of the self-supervised pre-training phase is quantitatively illustrated in Fig. 3. As shown in Fig. 3a, the joint optimization of the total loss (L_{total}) and the SIGReg regularization (L_{SIGReg}) demonstrates a rapid initial decay. This is followed by transient fluctuations between epochs 300 and 500, which reflect the model's adaptation to the structural diversity inherent in multi-center data sources, such as the MosMedData and ATM22 repositories. The eventual stabilization of both losses confirms that the system successfully "sculpted" the latent space into a predictable geometric structure. This process is further validated by the Manifold Stability Chart (Fig. 3d), where the progressive reduction and eventual plateau

of the mean L_2 norm for normal airway patches confirm the successful execution of the Manifold Contraction process."

To visualize this learned manifold, we applied Principal Component Analysis (PCA) to the high-dimensional latent embeddings, revealing that AI-PACS effectively compressed the majority of normal airway features into a compact, isotropic Gaussian hypersphere. As shown in Fig. 3b, while the core normality is tightly clustered, the visualization also captures the inherent complexity of clinical imaging, where approximately 10% of normal samples—representing physiological variants like focal inflammation or bronchial remodelling—form a "long-tail" distribution. Notably, while most radiolucent objects were projected as distinct high-norm outliers (marked as 'x'), a critical subset of "occult" FBA cases was observed within the peripheral boundaries of the normal cluster. These overlapping cases correspond to minute fragments or low-density organic materials that closely mimic healthy anatomical signatures, representing the primary diagnostic frontier addressed by our platform. While aggressive hyperparameter tuning of the SIGReg manifold constraints could mathematically force the separation of these minute objects, it would inevitably compromise the model's learned "pathological invariance," leading to an unacceptable increase in false-positive detections for patients with chronic airway variations. Thus, the current manifold reflects the optimal clinical trade-off between sensitivity and specificity.

The transition from abstract latent topology to clinical decision-making is governed by the distribution of L_2 norms, which serve as the basis for our anomaly scoring. The scores for the normal airway were primarily concentrated near the origin, whereas FBA cases demonstrated significantly higher magnitudes. Consistent with the PCA findings, the histogram in Fig. 3c reveals a "Detection Gap" or gray zone where the scores of occult FBAs overlap with the right-tail variants of the normal distribution. By implementing a diagnostic threshold (τ) at the 95th percentile of the internal validation scores, the system provides a mathematically grounded safeguard. This threshold ensures that AI-PACS can maintain a high sensitivity for emergency triage while minimizing the risk of false-positive surgical explorations in patients with complex pre-existing pulmonary conditions.

Furthermore, the stability of this learned representation was verified by the Manifold Stability Chart (Fig. 3d). As the self-supervised training progressed, the mean L_2 distance of normal airway patches from

the latent origin exhibited a consistent decline, eventually plateauing at a high level of compactness. This convergent trend, synchronized with the loss stabilization observed in Fig. 3a, provides direct evidence of the Manifold Contraction process. By forcing the diverse features of the 804-scan pre-training corpus into this predictable geometric structure, AI-PACS established a robust anatomical prior that maximizes the entropic distance to pathological outliers, ensuring that any deviation—regardless of the object's specific density—manifests as a detectable statistical anomaly.

Deployment of the AI-PACS Visualization Ecosystem

To translate our algorithmic findings into clinical reality, we deployed the LeJEPA-based detection engine within a fully integrated AI-PACS visualization ecosystem (Figs. 1d and 5). This framework establishes the "normality manifold" by pre-training on a diverse cohort of 804 unlabelled CT scans, projecting complex "non-FBA" airway features into a standard isotropic Gaussian distribution (Figs. 1a and b). Within this structured latent space, normal airway anatomy forms a compact cluster. In contrast, foreign bodies act as out-of-distribution (OOD) samples that resist compression, appearing as distinct statistical outliers with high embedding norms—calculated as the L_2 norm of the latent vector—which allows the system to identify FBA by detecting deviations from the learned manifold (Fig. 1c).

The system was engineered for high-throughput clinical environments, achieving an average end-to-end processing latency of approximately 60 seconds per scan (from DICOM ingestion to heatmap generation) on standard hospital-grade workstations. This rapid turnaround facilitates "near real-time" analysis, seamlessly integrating into the radiological workflow without introducing computational bottlenecks. Comparative workflow analysis indicates that this automated triage system reduces the average screen-time for negative exclusion by approximately 75% relative to traditional slice-by-slice manual inspection, which typically requires 3–5 minutes per case. This approach effectively transforms a tedious "needle-in-a-haystack" search into a targeted confirmation task.

The centrepiece of this ecosystem is the "Interactive Diagnostic Cockpit," designed to bridge the gap between abstract anomaly scores and anatomical intuition (Fig. 5a). The interface visualizes the global lung segmentation and strictly registers anomaly detections to the extracted centreline of the bronchial

tree. This allows the system to generate an automated localization report referencing standard segmental nomenclature (S1–S10). The clinical utility of this visualization is most profound in cases of radiolucent organic aspiration (*e.g.*, peanuts or beans). These objects are often indistinguishable from mucus or soft tissue on standard lung windows (Window Width:1600, Window Level: -600). For granular diagnostic review, the Automated Localization module (Fig. 5b) highlights suspected foreign bodies using red regions of interest (ROIs) across axial, coronal, and sagittal planes. Crucially, the "Topology Location" map projects these detections onto the bronchial centreline, identifying the specific anatomical segment (represented by a red star) according to the S1–S10 nomenclature. This mapping relies on the standardized anatomical segmentation of the tracheobronchial tree illustrated in Fig. 4, which provides the topological coordinates necessary for precise pre-operative bronchoscopy planning.

Beyond diagnosis, the platform extends capabilities into pre-operative surgical planning via a "Virtual Bronchoscopy" module. By correlating the 3D coordinates of the predicted anomaly with the airway topology, the system provides an intuitive spatial reference from the trachea to the target lesion. In a pilot user-experience evaluation involving interventional pulmonologists, this automated localization feature was reported to significantly enhance spatial awareness, allowing operators to visualize the precise depth, orientation, and occlusion status of the foreign body prior to sedation (Supplementary Video 1). This "digital rehearsal" capability is instrumental in assisting physicians to determine the approach, minimizing exploration-related trauma and reducing total procedure time, validating the platform as a comprehensive solution for interventional guidance.

Baseline Comparisons and Ablation Study

The architectural superiority of the full AI-PACS platform is systematically benchmarked in Fig 6. Performance metrics (Accuracy, Sensitivity, Specificity, and F1-Score) across the Internal (Fig. 6a), External (Fig. 6b), and Independent (Fig. 6c) cohorts (Table 1 - Phase II) demonstrate that the LeJEP-driven system consistently outperforms supervised (3D ResNet-18), generative (3D DAE/VAE), and contrastive (SimCLR) baselines. Furthermore, the performance degradation observed in the ablated variants (AI-PACS w/o Anatomical Mask and w/o SIGReg Loss) confirms that both spatial filtering of

anatomical noise and explicit geometric regularization are indispensable for resolving radiolucent aspiration. Patients characteristics for internal, external and independent cohorts were reported earlier [11].

To rigorously isolate the impact of the underlying learning paradigms, the supervised baseline in this experiment was implemented as a standard 3D ResNet-18, architecturally commensurate with our LeJEPa encoder. In contrast to our previous 2D framework¹¹, which relied on complex multi-view ensembles and extensive heuristic engineering to compensate for data scarcity—this native 3D configuration was deliberately selected to facilitate a controlled comparative analysis. By eliminating the confounding variables introduced by multi-view heuristics, this approach ensures an equitable benchmark between supervised cross-entropy classification and self-supervised anomaly detection under identical 3D volumetric constraints. Devoid of these specialized heuristics, the fully supervised 3D baseline exhibited substantial performance degradation, yielding a sensitivity of merely 47.3% and an F1-Score of 44.4% on the internal cohort. This precipitous decline underscores the inherent vulnerability of conventional supervised classifiers when confronted with the statistical rarity of FBA cases. Similarly, the generative baselines (DAE and VAE) exhibited the characteristic limitations associated with pixel-level reconstruction. Although the DAE achieved a relatively high sensitivity in the independent cohort (61.4%), it incurred an unacceptably high false-positive rate, compromising specificity (72.1%). Given that the radiodensity of radiolucent foreign objects is virtually indistinguishable from surrounding soft tissue and mucus, the reconstruction errors leveraged by these generative models failed to reliably differentiate true morphological anomalies from physiological noise, resulting in significant over-prediction. Furthermore, while the SimCLR contrastive framework enhanced overall feature representation—achieving F1-Scores of 47.3% and 49.4% in the internal and independent cohorts, respectively—it remained constrained by the absence of explicit geometric regularization. It ultimately struggled to compress the normal anatomical manifold into a uniform distribution, limiting its diagnostic precision.

Furthermore, the component ablation study confirmed that both anatomical isolation and geometric regularization are strict prerequisites for robust anomaly detection, revealing a critical sensitivity-

specificity trade-off. Removing the anatomical prior (AI-PACS w/o Mask) exposed the network to massive background variance. Interestingly, while this configuration maintained a high sensitivity (*e.g.*, achieving a sensitivity of 64.1% in the independent cohort), it caused the model to heavily conflate complex parenchymal textures and vascular trees with bronchial anomalies. This anatomical noise drastically reduced Specificity (79.2%) and yielded a suboptimal F1-Score (48.7%).

Conversely, ablating the manifold contraction mechanism (AI-PACS w/o SIGReg) weakened the anomaly scoring logic. Without the explicit geometric forces compressing normal features into a standard Gaussian hypersphere $N(0, I)$, the system struggled to define a crisp decision boundary, resulting in internal and independent F1-Scores of 48.9% and 50.2%, respectively. In contrast, the predictive LeJEPa architecture of the full AI-PACS successfully integrated both the mask-guided spatial focus and the geometric manifold constraints. By bypassing low-level intensity confounders and filtering out background noise, the full pipeline achieved the optimal balance between sensitivity and specificity, yielding the highest overall F1-Scores across the Internal (49.5%), External (50.0%), and Independent (51.4%) cohorts.

Evaluating the Performance of AI-PACS vs. Expert Radiologists

To rigorously quantify the clinical utility of the AI-PACS platform, we conducted a head-to-head performance comparison against manual radiological review on the dedicated human-AI comparative evaluation cohort ($n=35$, comprising 16 FBA and 19 NFBA cases) (Table 1 - Phase III). This cohort was randomly selected from the internal cohort (The Central Hospital of Wuhan).

The baseline demographic and clinical characteristics of the 35 patients in the independent human-AI comparative cohort (16 FBA and 19 NFBA) are summarized in Table 2. The two groups were statistically comparable regarding baseline demographics, including age ($P = 0.5620$), sex ($P = 0.5001$), and BMI ($P = 0.1236$). There were also no significant differences in the history of chronic respiratory diseases, such as COPD ($P = 0.7003$) or asthma ($P = 1.0000$). Notably, the length of the disease course was significantly prolonged in the FBA group compared to the NFBA group (median 20 days vs. 10 days, $P = 0.0238$). Furthermore, the MDCT presentations of the NFBA control group exhibited a significantly higher

prevalence of confounding structural abnormalities. Specifically, patients in the NFBA group had higher rates of pulmonary consolidation (52.6% vs. 12.5%, $P = 0.0298$), pulmonary emphysema (36.8% vs. 6.2%, $P = 0.0472$), obsolete pulmonary tuberculosis (73.7% vs. 6.2%, $P < 0.001$), lung nodules (73.7% vs. 18.8%, $P = 0.0020$), and interstitial changes of the lung (68.4% vs. 0.0%, $P < 0.001$). These complex presentations confirm that the comparative cohort provided a demanding clinical scenario characterized by substantial background anatomical noise, mirroring the "needle-in-a-haystack" challenge of real-world FBA diagnosis.

As demonstrated in the baseline characteristics, this specific subset presented a highly demanding "needle-in-a-haystack" scenario laden with confounding anatomical noise. Standard manual slice-by-slice inspection by experienced radiologists yielded a sensitivity (Recall) of merely 18.8% (3/16), underscoring the extreme visual difficulty of diagnosing radiolucent objects unaided. In stark contrast, the AI-PACS platform successfully identified 9 out of 16 positive cases, elevating the Sensitivity to 56.3%. This represents a statistically significant absolute improvement of +37.5% ($P < 0.05$), effectively cutting the missed diagnosis (false negative) rate by more than half. Furthermore, the model maintained a robust capability to reject false alarms; it outperformed human readers in Specificity (78.9% vs. 63.2%), although this specific difference did not reach statistical significance ($P = 0.28$). Driven by the substantial surge in sensitivity and the positive trend in specificity, the Overall Accuracy saw a highly significant enhancement from 42.9% (Manual) to 68.6% (AI-PACS, $P < 0.05$). These precise metrics, detailed in Table 3, firmly establish that the self-supervised manifold contraction approach perceives subtle 3D morphological deviations that consistently escape human vision, validating the platform's immense potential as an automated clinical safeguard against diagnostic delays.

Discussion

The diagnostic management of FBA in adults represents one of the most persistent "blind spots" in pulmonary imaging [23]. Our study introduces a paradigm shift by deploying AI-PACS, a self-supervised anomaly detection platform that bypasses the traditional limitations of data scarcity and human visual thresholds (Fig. 1c). By learning the compact topological manifold of the "non-FBA" airway, the system perceives radiolucent obstructions as statistical outliers, achieving a diagnostic sensitivity that significantly surpasses experienced radiologists in head-to-head clinical testing.

The primary clinical dilemma in adult FBA is the radiographic radiolucency of aspirated objects [24]. As our data confirms, radiolucent materials such as peanuts, fish bones, or plastic fragments possess Hounsfield Unit (HU) values (30–60 HU) that are virtually indistinguishable from surrounding soft tissue or mucus. This physical reality leads to a high false-negative rate in manual reviews; in our comparative cohort, senior radiologists missed 13 out of 16 positive cases, yielding a baseline sensitivity of only 18.8%. The AI-PACS platform effectively overcomes this visual bottleneck. Unlike the human eye, which relies on contrast and intensity, the LeJEPa engine utilizes manifold contraction (Fig. 2) to identify subtle 3D morphological deviations. By forcing all normal airway features into a standard Gaussian hypersphere $N(0, I)$, as visually demonstrated by the high degree of clustering in our latent manifold projection (Fig. 3b), any foreign object manifests as a high-norm outlier regardless of its density (Fig. 1c). This approach resulted in an absolute sensitivity boost of +37.5%, effectively cutting the rate of missed diagnoses by more than half while maintaining a robust specificity of 78.9%.

Although AI-driven FBA screening has shown promise in recent literature, existing solutions are largely constrained to pediatric populations and 2D planar imaging modalities. For instance, machine learning applications for pediatric chest X-rays have demonstrated the ability to augment human sensitivity, yet these 2D-based approaches are fundamentally limited by anatomical summation [9]. This visual overlapping is particularly obstructive in the adult tracheobronchial tree, where 3D volumetric CT remains the diagnostic gold standard for resolving complex airway structures. Unlike these 2D paradigms, AI-PACS operates within a 3D latent space, enabling the system to perceive subtle morphological

signatures and topological deviations that are inherently lost during 2D projection. By bridging this dimensional gap, our platform provides a necessary diagnostic safety net tailored for the intricate anatomy of the adult airway.

A core methodological advantage of the AI-PACS platform is its ability to circumvent the "data starvation" that plagues traditional deep learning models in rare disease diagnostics. The stark performance contrast between our self-supervised LeJEPA engine and the supervised 3D ResNet-18 baseline (Recall: 56.1% vs. 47.3% in internal validation; Fig. 6a) underscores the inherent limitations of the "classification" paradigm. Traditional supervised models operate under a "closed-world" assumption; they are fundamentally data-starved in FBA scenarios and struggle to define a discriminative boundary for anomalies they rarely encounter during training¹⁴. This forces the classifier to overfit to a limited set of morphological features, leaving it vulnerable to out-of-distribution variations in real-world clinical settings.

Our strategy bypasses this bottleneck by operating under an "open-world" zero-shot assumption. By leveraging 804 unlabeled, accessible scans (Fig. 1a), we define anatomical "normality" with high mathematical precision. As shown in the Manifold Stability Chart (Fig. 3d), the consistent contraction of the latent representation during pre-training ensures that the model achieves "pathological invariance". This explicit geometric constraint effectively filters out confounding structural noise—such as inflammation or chronic remodeling—that typically triggers false positives in pixel-level reconstruction models. While a "gray zone" of statistical overlap exists for certain occult fragments (Fig. 3c), the overall convergence of the training dynamics (Fig. 3a) provides a stable foundation for the automated diagnostic safeguard. By abandoning the reliance on scarce positive labels, AI-PACS sacrifices artificial, overfitted classification metrics in exchange for a highly robust, generalized anomaly detection mechanism.

Beyond algorithmic accuracy, the true clinical advantage of AI-PACS lies in its seamless integration. Operating within a 3D latent space allows the system to perceive topological deviations that are inherently lost during 2D projection. In the clinical diagnostic pathway, the platform acts as a critical

"safety net" for patients presenting with high clinical suspicion but negative or ambiguous standard CT findings. By deploying the Interactive Diagnostic Cockpit (Fig. 5a), AI-PACS reduces negative exclusion screen time by 75%, transforming a tedious slice-by-slice manual search into a rapid triage confirmation. Furthermore, the Automated Localization module correlates anomalies directly with the standardized anatomical segmentation of the tracheobronchial tree (Fig. 4). By rendering these detections as topology location maps (Fig. 5b), the platform bridges the gap between diagnosis and intervention, enhancing the spatial awareness of interventional pulmonologists for precise pre-operative bronchoscopy planning.

Despite the robust performance across multi-center cohorts, several limitations warrant careful consideration. First, the retrospective nature of this study, while a necessary strategy to overcome the "data starvation" inherent to rare clinical events like adult FBA, limits our ability to evaluate the platform's impact on long-term patient outcomes in real-time clinical environments. While we have demonstrated significant gains in diagnostic sensitivity and screening efficiency, prospective clinical trials are essential to determine if these algorithmic advantages translate directly into clinical endpoints, such as reduced hospital stays, lower rates of obstructive complications, and improved survival in acute emergency scenarios [25]. Second, the morphological detection thresholds of the AI-PACS platform require further characterization. Although the model demonstrated high sensitivity across a range of organic and inorganic materials, its performance on extremely small fragments (< 2 mm) in the distal sub-segmental bronchi remains a challenge. Such minute objects may fall below the effective spatial resolution of standard clinical CT protocols or the specific patch-tokenization range of the LeJEPa encoder, potentially leading to occult obstructions in the extreme periphery. Future iterations could integrate higher-resolution reconstruction algorithms or adaptive multi-scale patch processing to better capture these subtle distal signals. Finally, while the integration of the SIGReg loss and anatomical priors significantly mitigated the false-positive burden, certain physiological confounders persist. High-density impacted mucus or focal endobronchial inflammatory masses can exhibit MDCT densities and topological deviations that mimic radiolucent foreign bodies. While our "pathological invariance" training on non-FBA cohorts mitigated many such errors, the current system is primarily optimized for the first six to seven generations of the tracheobronchial tree (S1–S10). Extending precise detection to

the extreme peripheral airways—beyond the reach of standard centerline extraction—remains a technical frontier requiring further optimization of automated segmentation and 3D back-projection algorithms.

The AI-PACS platform represents a significant advancement in the automated detection of radiographically occult foreign body aspiration in adults. By shifting the diagnostic paradigm from traditional supervised classification to self-supervised anomaly detection, our system effectively circumvents the "data scarcity" that has historically limited the utility of deep learning in rare disease diagnostics. Through the application of LeJEPa and SIGReg manifold contraction, the platform establishes a high-precision anatomical prior that perceives subtle 3D morphological deviations as statistical outliers, effectively identifying "invisible" radiolucent objects that consistently escape human vision. Our multi-center validation and head-to-head benchmarking demonstrate that AI-PACS provides a robust clinical "safety net," significantly increasing diagnostic sensitivity by +37.5% and reducing the time required for negative exclusion by 75%. Beyond simple detection, the integrated Virtual Bronchoscopy module offers precise anatomical localization referenced to standard nomenclature, bridging the gap between radiological suspicion and successful surgical intervention. While further prospective studies are needed to fully characterize its impact on long-term clinical outcomes, the AI-PACS ecosystem stands as a powerful, data-efficient tool for emergency and respiratory settings. By automating the tedious "needle-in-a-haystack" search for radiolucent obstructions, this platform has the potential to minimize diagnostic delays, reduce unnecessary patient morbidity, and establish a new standard of care for suspected foreign body aspiration.

Acknowledgments

Not applicable.

Author contributions

Y.W., W.W., H.N., Y.H., Z.B., G.Z., R.M.E., and X.L. conceptualized and designed the study. All authors had full access to the data and contributed to the decision to submit the manuscript for publication. Y.W., W.W., H.N., Y.H., Z.B., R.M.E., X.L., Z.C., Z.T., and X.Y. contributed to manuscript drafting. Z.C., X.L., Y.J., and Z.T. performed the data analyses. All authors reviewed the manuscript critically for important intellectual content and approved the final version. Data collection was performed by X.L., X.Y., D.Z., F.N., S.G., Q.Q., Y.H., J.X., YinW., M.Z., G.Z., H.N., and X.W. We thank Dr. Lan Lan and Dr. Yiyuan Cao for their assistance in CT image review and clinical interpretation.

Funding

XL was supported by the Beijing Medical and Health Public Welfare Foundation (YWJKJHHKYJJ-KXCRHX004) and Wuhan Natural Science Foundation Exploration Program -Morning Light Program (2025020701020246). JX was funded by the Wuhan Central Hospital's Institutional Discipline Fund (2021XK104). XY was funded by Science and Technology Plan Project of Wuhan Municipal Health Commission WX20B01. HN was supported by the National Natural Science Foundation of China (82370031). YH received funding from the National Natural Science Foundation of China (62073149) and the Wuhan Municipal Health Commission (WX23A81). ZC and ZT were supported by the China Scholarship Council (202308060201; 202308060211). YW was supported by the UK Medical Research Council (MR/S025480/1).

Availability of data and materials

The data supporting the findings of this study are available upon reasonable request from the corresponding author with approval from the corresponding hospital. The source code for this implementation is publicly accessible at the following website: Code is available at https://github.com/ZheChen1999/LeJEP3D_CT.

Declarations

The authors declare that they have no conflicts of interest regarding the publication of this work.

Ethics approval

As part of routine clinical care, patients provided general informed consent for the use of their imaging data. For this retrospective, multicenter study, the ethics committees of the participating institutions—

As part of routine clinical care, patients provided general informed consent for the use of their imaging data. For this retrospective, multicenter study, the ethics committees of the participating institutions: The Central Hospital of Wuhan (Approval Nos. WHZXKYL2024-108 and WHZXKYL2025-069), The Renmin Hospital of Wuhan University (Approval No. WDRY2025-K083), and The Zhongnan Hospital of Wuhan University (Approval No. 2025086K), approved the use of de-identified imaging data and waived the requirement for additional study-specific informed consent. The study was conducted in accordance with the Declaration of Helsinki.

Consent to participate

Not applicable. The requirement for written informed consent to participate was waived by the institutional review boards due to the retrospective and observational design of the study and the use of fully de-identified imaging data.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no conflicts of interests.

References

- [1] Bajaj D, Sachdeva A, Deepak D. Foreign body aspiration. *Journal of Thoracic Disease*. 2021;13:5159.
- [2] Wang W-y, Wang S-y, Zhao Q-j, Li W-y, Zhang T, Wang Y-j. Analysis of risk factors for complications of foreign body aspiration in children. *Respiratory Medicine*. 2025:108363.
- [3] Morris P, Perkins A. Diagnostic imaging. *The Lancet*. 2012;379:1525-33.
- [4] Takeyama R, Yamauchi Y, Nishida T, Saito Y, Saito K, Sasajima Y, et al. A Retained Childhood Toy Mimicking Chronic Bronchial Obstruction in Adulthood: A Case Report. *Respirology Case Reports*. 2026;14:e70505.
- [5] Milivojević I, Kontić-Jovanović M. Foreign body aspiration misdiagnosed as asthma: A case report and literature review. *Srpski arhiv za celokupno lekarstvo*. 2026:13-.
- [6] Liu X, Ni F, Guo T, Jiang F, Jiang Y, Song C, et al. Risk factors associated with radiolucent foreign body inhalation in adults: a 10-year retrospective cohort study. *Respiratory research*. 2022;23:238.
- [7] Flak B. *Computed Tomography of the Body*. *Computed Tomography-E-Book: Computed Tomography-E-Book*. 2022:393.
- [8] Kumar S, Kumar H, Kumar G, Singh SP, Bijalwan A, Diwakar M. A methodical exploration of imaging modalities from dataset to detection through machine learning paradigms in prominent lung disease diagnosis: a review. *BMC Medical Imaging*. 2024;24:30.
- [9] Truong B, Zapala M, Kammen B, Luu K. Automated Detection of Pediatric Foreign Body Aspiration from Chest X-rays Using Machine Learning. *The Laryngoscope*. 2024;134:3807-14.
- [10] Çoşkun N, Yalçınkaya M, Demir E. The impact of Deep Learning on determining the necessity of Bronchoscopy in Pediatric Foreign Body Aspiration: can negative bronchoscopy Rates be reduced? *Journal of Pediatric Surgery*. 2025;60:162014.
- [11] Liu X, Chen Z, Tang Z, Yang X, Jiang Y, Zheng D, et al. Automated detection of radiolucent foreign body aspiration on chest CT using deep learning. *npj Digital Medicine*. 2025;8:647.
- [12] Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. *BMJ Open*. 2016;6:e012799.
- [13] Morozov SP, Andreychenko AE, Blokhin IA, Gelezhe PB, Gonchar AP, Nikolaev AE, et al. MosMedData: data set of 1110 chest CT scans performed during the COVID-19 epidemic. *Digital Diagnostics*. 2020;1:49-59.
- [14] Zhang M, Wu Y, Zhang H, Qin Y, Zheng H, Tang W, et al. Multi-site, multi-domain airway tree modeling. *Medical image analysis*. 2023;90:102957.
- [15] Qin Y, Chen M, Zheng H, Gu Y, Shen M, Yang J, et al. Airwaynet: a voxel-connectivity aware approach for accurate airway segmentation using convolutional neural networks. *International conference on medical image computing and computer-assisted intervention: Springer*; 2019. p. 212-20.
- [16] Yu W, Zheng H, Zhang M, Zhang H, Sun J, Yang J. Break: Bronchi reconstruction by geodesic transformation and skeleton embedding. *2022 IEEE 19th international symposium on biomedical imaging (ISBI): IEEE*; 2022. p. 1-5.
- [17] Zheng H, Qin Y, Gu Y, Xie F, Yang J, Sun J, et al. Alleviating class-wise gradient imbalance for pulmonary airway segmentation. *IEEE transactions on medical imaging*. 2021;40:2452-62.

- [18] Zhang M, Zhang H, Yang G-Z, Gu Y. CFDA: Collaborative feature disentanglement and augmentation for pulmonary airway tree modeling of COVID-19 CTs. *International conference on medical image computing and computer-assisted intervention*: Springer; 2022. p. 506-16.
- [19] Balestrierio R, LeCun Y. Lejepa: Provable and scalable self-supervised learning without the heuristics. *arXiv preprint arXiv:251108544*. 2025.
- [20] Saheb A, Obando-Ceron J, Courville A, Bashivan P, Castro PS. Stable Deep Reinforcement Learning via Isotropic Gaussian Representations. *arXiv preprint arXiv:260219373*. 2026.
- [21] Zhu G, Yu Y, Li Z, Dai Y. GMDDPM-CT: Advancing network intrusion detection through diffusion-based sample synthesis and deep learning. *ICT Express*. 2026.
- [22] Moayedikia A, Jensen R, Fin S. Feature selection for high-dimensional imbalanced class datasets using Harmony Search and Kullback–Leibler divergence. *Data Mining and Knowledge Discovery*. 2026;40:6.
- [23] Daher LA, Guerra-Londono CE, Debiante LG. Rigid Bronchoscopy and Bronchial Interventions. *Anesthesia for Interventional Pulmonary Procedures*: Springer; 2026. p. 61-81.
- [24] Meyers ML, Moore MM, Baker JB, Clemenshaw MN, Cooper ML, Hammer MR, et al. ACR Appropriateness Criteria® Ingested or Aspirated Foreign Body-Child. *Journal of the American College of Radiology*. 2026.
- [25] Ramanath Mahale AA, Shirodkar D, Prabhu AS. Occult foreign body airway in children—Case series with review of literature. 2026.

Figures

Fig. 1. Overview of the AI-PACS platform for radiolucent FBA detection. (a) Data Curation (Phase I). An aggregated pre-training cohort of 804 scans was constructed from public (MosMedData, ATM22) and private sources to capture diverse airway norms. Crucially, ground-truth masks were used as anatomical priors to isolate pure airway volumes, ensuring noise-free training. (b) Manifold Contraction. The LeJEPA encoder utilizes SIGReg loss to compress normal airway features into a standard Gaussian manifold ($N(0, I)$). (c) Inference (Phase II). For new patients, an automated U-Net segmenter extracts the airway. The frozen encoder computes anomaly scores ($\|z\|_2$), identifying foreign bodies as statistical outliers that deviate from the learned manifold. (d) Clinical Visualization. The AI-PACS interface renders high-scoring regions as 3D volumetric heatmaps, providing precise localization mapped to bronchial segments (S1–S10).

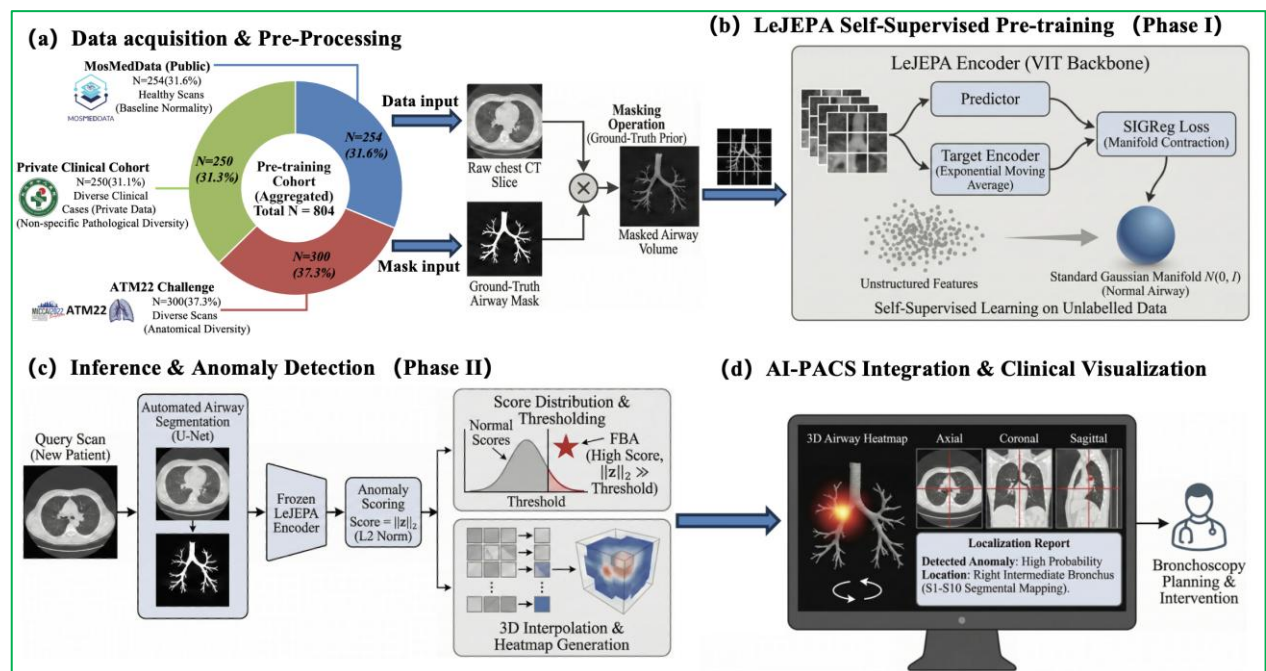


Fig. 2. LeJEPA architecture and manifold contraction via SIGReg loss. (left) LeJEPA Engine. The framework processes masked airway volumes using asymmetric dual encoders. A predictor reconstructs the latent features of masked regions from visible context patches, driving self-supervised representation learning. (right) Manifold Contraction. Visualization of the SIGReg mechanism. Starting from an unstructured distribution (Stage 1), the loss applies Variance Regularization (orange arrows) to prevent collapse and Covariance Regularization (blue arrows) to enforce orthogonality (Stage 2). This "sculpting" process compacts normal airway features into a standard Gaussian manifold ($\mathcal{N}(0, I)$), forcing foreign bodies to manifest as detectable high-norm outliers ($\|z\|_2$) (Stage 3).

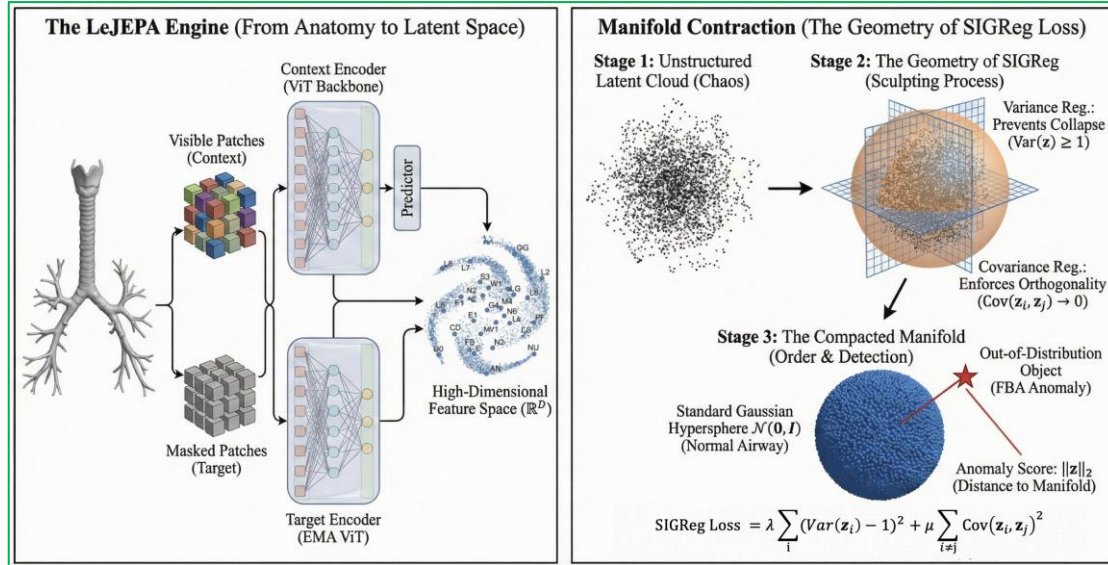


Fig. 3. Quantitative evaluation of representation learning and manifold stability during Phase I pre-training. (a) Training dynamics. Convergence of the joint loss (L_{total}) and SIGReg regularization (L_{SIGReg}) across 1,000 epochs. Stochastic fluctuations during the intermediate phase (epochs 300–500) indicate the model's adaptation to multi-center structural heterogeneity. (b) Latent manifold projection. PCA visualization demonstrates the successful compression of normal airway features into a compact Gaussian cluster. Red markers illustrate how radiolucent foreign bodies (FBA) manifest as high-norm outliers, including "occult" cases that reside at the manifold boundary. (c) Anomaly score distribution. Histogram analysis showing the statistical separation between healthy anatomy and radiographically occult obstructions. The dashed line defines the operational 95% specificity threshold. (d) Manifold stability chart. The progressive reduction in the mean L_2 norm of normal airway patches over training epochs confirms the "Manifold Contraction" process, proving that the system successfully established a high-precision and stable anatomical prior of normality.

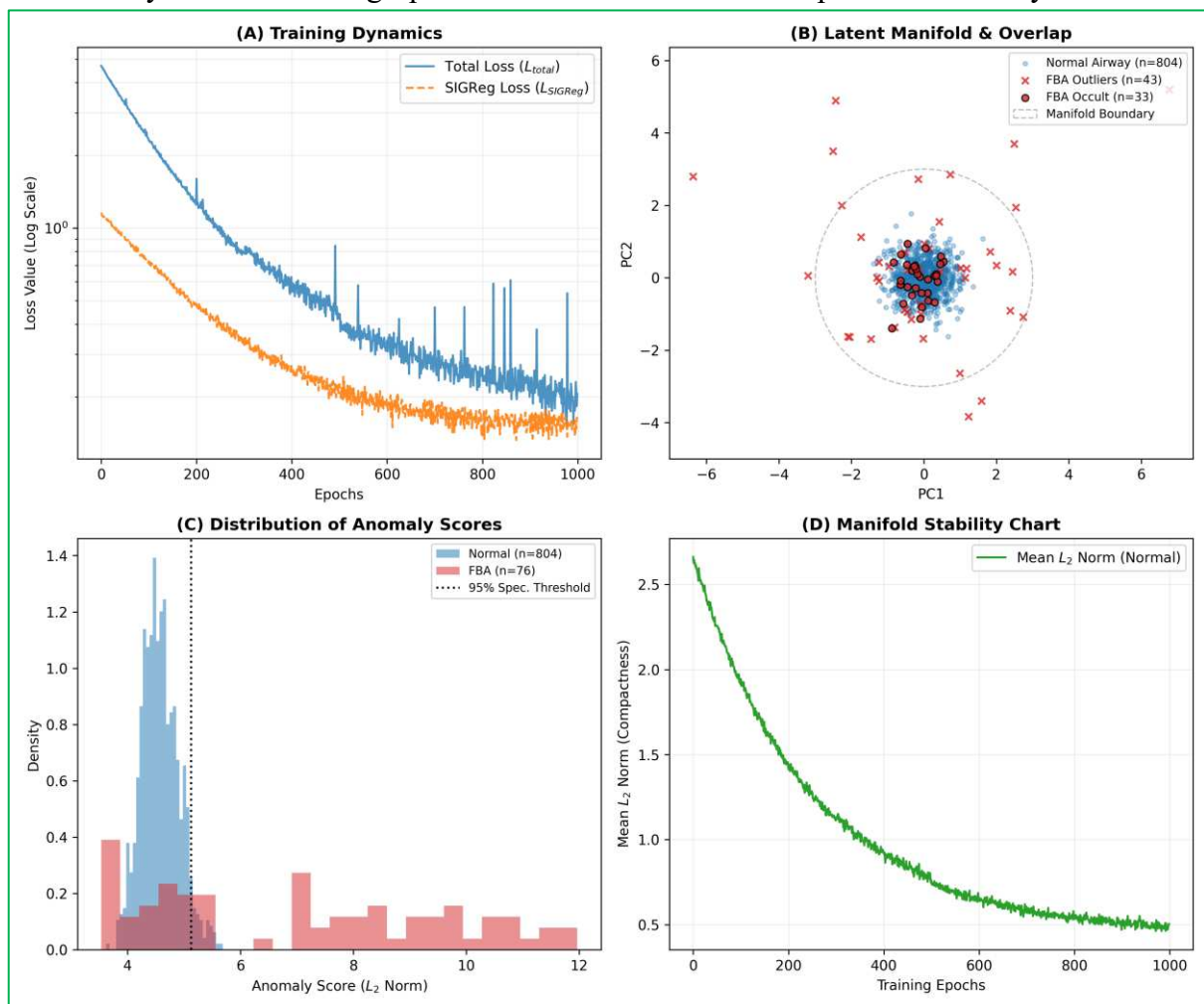


Fig. 4. Anatomical segmentation of the tracheobronchial tree. Schematic diagram illustrating the standard nomenclature for the lobar and segmental bronchi (S1–S10). The AI-PACS platform utilizes this topological map to provide precise localization of detected anomalies, providing anatomical guidance for bronchoscopy planning.

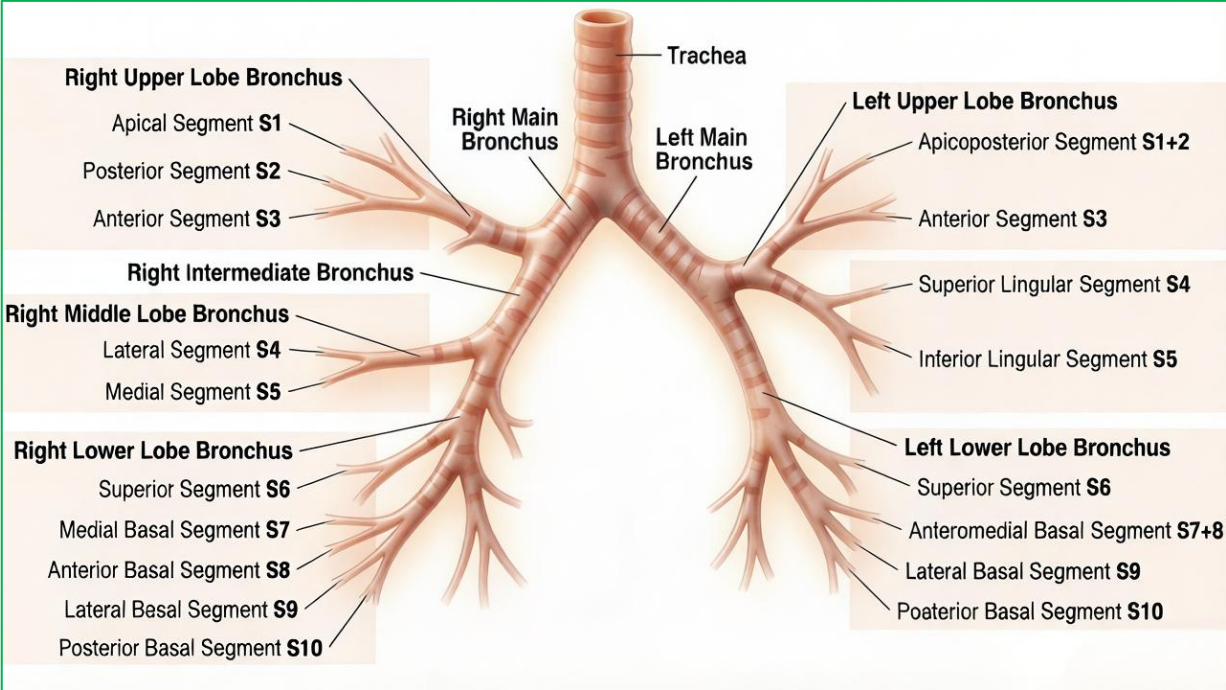


Fig. 5. AI-PACS user interface and automated anomaly localization. (a) AI-PACS Clinical Dashboard. The integrated workstation interface displays the global lung segmentation (green overlay) and multi-planar reconstruction (MPR) views. The dashboard incorporates a dedicated 3D volume rendering (inset) to visualize the spatial relationship between the bronchial tree (blue) and the detected foreign body (gold), alongside a ranked list of detection scores. (b) Automated Localization. Detailed visualization of the top-ranked anomaly. The system automatically highlights the suspected foreign body with red regions of interest (ROIs) across axial, coronal, and sagittal planes. The "Topology Location" map (bottom left) projects the detection onto the bronchial centerline, identifying the specific anatomical segment (red star) to provide an anatomical reference for bronchoscopy navigation.

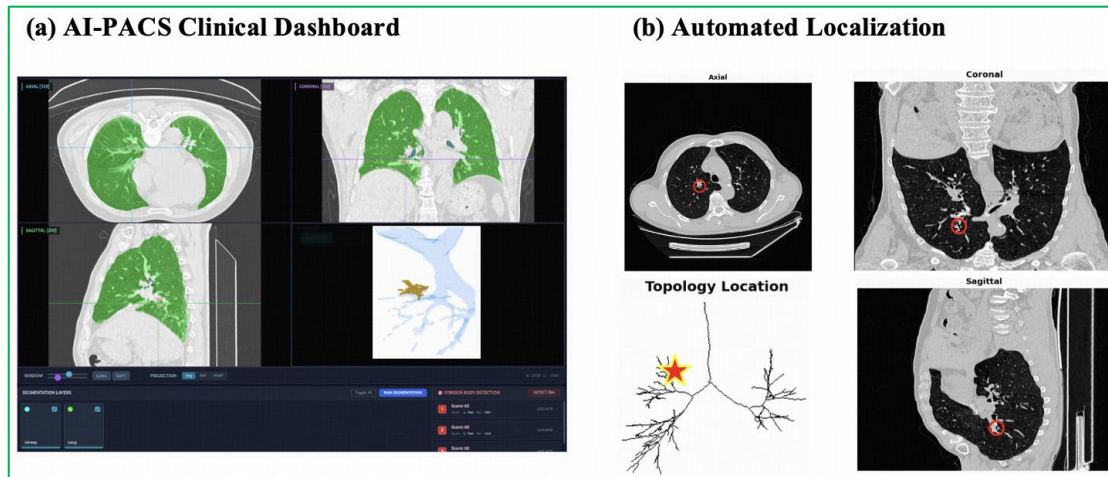
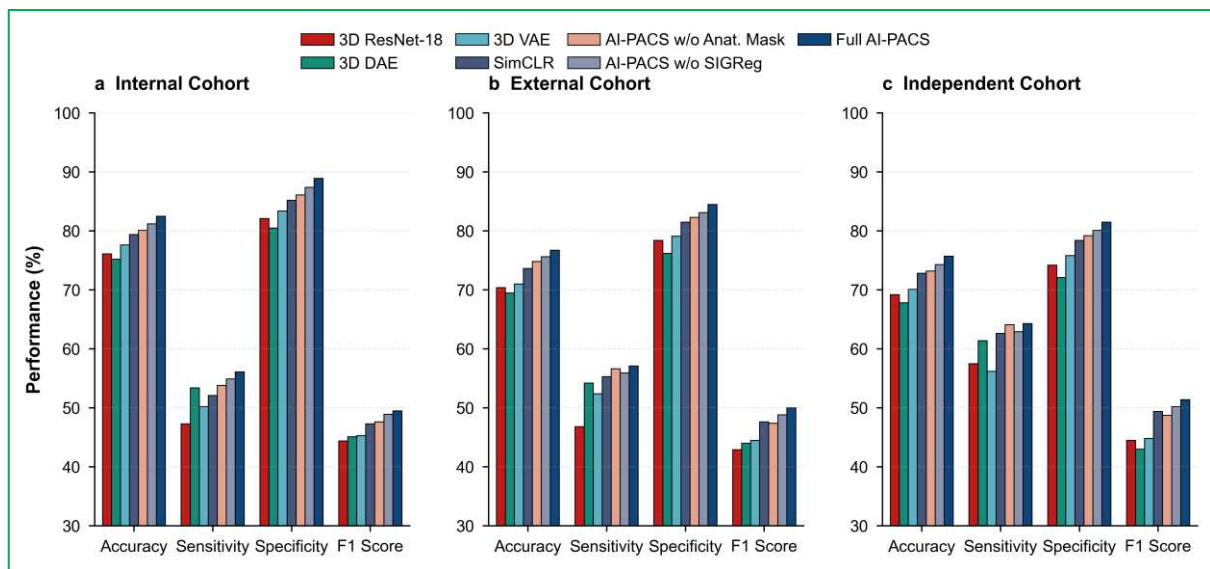


Fig. 6. Comprehensive baseline comparison and ablation study across multi-center cohorts. Bar plots evaluating the diagnostic performance (Accuracy, Precision, Recall, and F1 Score) of the proposed Full AI-PACS platform against established representation learning paradigms and ablated configurations on the (a) Internal Cohort, (b) External Cohort, and (c) Independent Cohort. The compared models include a supervised baseline (3D ResNet-18), generative baselines (3D DAE, 3D VAE), a contrastive baseline (SimCLR), and two architecturally degraded variants (AI-PACS w/o Anatomical Mask, AI-PACS w/o SIGReg Loss). Across all three clinical centers, the complete LeJEPA-driven AI-PACS consistently achieves the highest overall performance, demonstrating robust generalization. The performance drops observed in the ablated variants explicitly validate that both the spatial filtering of anatomical noise (Anatomical Mask) and explicit geometric regularization (SIGReg) are indispensable prerequisites for resolving radiolucent foreign body aspiration. Abbreviations: DAE, denoising autoencoder; VAE, variational autoencoder; SimCLR, simple framework for contrastive learning of visual representations; SIGReg, variance-covariance regularization.



Tables

Table 1

Summary of Patient Cohorts and Dataset Distribution Across Study Phases.

Study Phase	Cohort/Source	FBA (Radiolucent)	Non-FBA (Control)	Total (n)
Phase I: Pre-training	Aggregated Normality Corpus	0	804	804
	MosMedData (Healthy)	0	254	254
	Private Clinical (Randomized)	0	250	250
	ATM22 Challenge (Multi-center)	0	300	300
Phase II: Validation	Multi-Center Aggregate	76	365	441
	Internal (The Central Hospital of Wuhan)	41	227	268
	External (The Renmin Hospital of Wuhan University)	21	82	103
	Independent (The Zhongnan Hospital of Wuhan University)	14	56	70
Phase III: Comparison	Human-AI Benchmarking Subset*	16	19	35

Note: The FBA group specifically focuses on radiolucent cases to evaluate the platform’s performance in clinically occult scenarios.

Table 2

Baseline demographic, clinical, and radiological characteristics of the independent human-AI comparative evaluation cohort. Data are presented as median (interquartile range) for continuous variables and as number (percentage) for categorical variables. Statistical comparisons between the FBA and NFBA groups were performed using the Mann-Whitney U test for continuous variables and Fisher's exact test for categorical variables.

Feature	FBA Group (n=16)	NFBA Group (n=19)	P-value
Age, years	67.5 (57.5-76.0)	63.0 (49.5-77.0)	0.5620
Sex			0.5001
Female	6 (37.5%)	10 (52.6%)	
Male	10 (62.5%)	9 (47.4%)	
BMI, kg/m²	24.3 (21.6-26.2)	22.8 (19.0-24.3)	0.1236
Length of disease course, days	20 (7-150)	10 (2.5-15)	0.0238*
Hospital stays, days	10 (9-12)	8 (5-12)	0.2516
Experience in ICU			1.0000
Yes	2 (12.5%)	3 (15.8%)	
No	14 (87.5%)	16 (84.2%)	
Chronic respiratory disease			
COPD	3 (18.8%)	5 (26.3%)	0.7003
Asthma	1 (6.2%)	2 (10.5%)	1.0000
Bronchiectasis (History)	1 (6.2%)	3 (15.8%)	0.6081
Lung cancer	0 (0.0%)	0 (0.0%)	1.0000
Interstitial lung disease	0 (0.0%)	1 (5.3%)	1.0000
Nervous system disease	2 (12.5%)	2 (10.5%)	1.0000
Smoking history	5 (31.2%)	5 (26.3%)	1.0000
MDCT presentations			
Pneumonic patch	13 (81.2%)	14 (73.7%)	0.7003
Airway stenosis	4 (25.0%)	1 (5.3%)	0.1558
Atelectasis	5 (31.2%)	4 (21.1%)	0.7003
Bronchiectasis	1 (6.2%)	6 (31.6%)	0.0964
Pleural effusion	0 (0.0%)	4 (21.1%)	0.1088
Consolidation	2 (12.5%)	10 (52.6%)	0.0298*
Pulmonary emphysema	1 (6.2%)	7 (36.8%)	0.0472*
Obsolete Pulmonary Tuberculosis	1 (6.2%)	14 (73.7%)	<0.001*
Lung nodule	3 (18.8%)	14 (73.7%)	0.0020*
Interstitial changes of the lung	0 (0.0%)	13 (68.4%)	<0.001*

Statistically significant difference ($P < 0.05$). Abbreviations: FBA, foreign body aspiration; NFBA, non-foreign body aspiration; BMI, body mass index; ICU, intensive care unit; COPD, chronic obstructive pulmonary disease; MDCT, multi-detector computed tomography.

849 **Table 3**

850 Head-to-Head Diagnostic Performance Comparison Between AI-PACS and Expert Radiologist
851 Panel (n=35).

Diagnostic Modality	TP	FP	TN	FN	Sensitivity (%)	Specificity (%)	Accuracy (%)
Manual Radiologist Review	3	7	12	13	18.8	63.2	42.9
AI-PACS Platform	9	4	15	7	56.3	78.9	68.6
Absolute Improvement	+6	-3	+3	-6	+37.5	+15.7	+25.7
P-value	-	-	-	-	< 0.05	0.28	< 0.05

852 Data are n (%). FBA: Foreign Body Aspiration; TP: True Positive; FP: False Positive; TN: True Negative; FN:
853 False Negative.

854 Note: Values in parentheses represent the percentage relative to the full evaluation cohort (N = 35; 16 FBA and
855 19 NFBA). Percentages for TP and FN are calculated with denominator = 16 (FBA cases), and for TN and FP
856 with denominator = 19 (NFBA cases).

857 P values with McNemar's test.