

学 位 論 文 要 旨 (Summary of the Doctoral Dissertation)	
学位論文題目 (Dissertation Title)	放射線治療のためのリアルタイムカラーX線透視画像からのマーカーレス肺腫瘍追跡
氏 名(Name)	エン ヨウケン
放射線治療は肺がんの主な治療法の一つであるが、その有効性は呼吸による腫瘍運動に大きく影響される。画像誘導放射線治療（IGRT）は、腫瘍をリアルタイムで追跡することでこれを軽減することを目的とする。臨床的な金標準は、フィデューシャルマーカーを外科的に埋め込む方法であるが、これは侵襲的な手技であり、気胸や出血などの合併症のリスクを伴う。そのため、「マーカーレス」追跡法が強く求められているが、手技中のキロボルト（kV）蛍光透視画像の軟部組織コントラストの低さと高ノイズレベルにより、その開発が妨げられている。 深層学習（DL）モデルは、マーカーレスセグメンテーションに有望な解決策を提供するが、膨大な注釈付き訓練データが必要である。ノイズの多い蛍光透視シーケンスの手動注釈が現実的でないため、通常、計画 CT スキャンから生成された合成的なデジタル再構成放射線写真（DRRs）で訓練される。これにより、重要な「ドメインギャップ」が生じる：クリーンな合成 DRRs で訓練されたモデルは、ノイズが多くぼやけ、スタイルが異なる実際の臨床蛍光透視画像に直面すると一般化に失敗する可能性がある。 本論文は、このドメインギャップを克服し、患者固有の再訓練を必要とせずにクロス患者モデルを用いた正確でリアルタイムのマーカーレス肺腫瘍追跡を実現するための包括的なフレームワークを提示する。本研究の貢献は三つである：（1）現実的な訓練データ生成のための新しい物理情報に基づく劣化モデルの開発；（2）3D 腫瘍追跡のための完全な深層学習ワークフローの設計と検証；（3）研究および検証パイプライン全体を効率化する統合ソフトウェアプラットフォームの実装。 第2章で詳述される最初の主要貢献は、ドメインギャップを橋渡しするための新しい手法である。本研究では、単純なノイズシミュレーション（例：ガウスノイズ）が不十分であると主張する。SyncTraX 立体視システムで撮像されたカスタムゼラチンファントムを用いた実証的な分析を通じて、蛍光透視ノイズの統計的調査を行った。主要な発見は、ノイズがランダムではなく、ピクセル強度および材料密度と相関する内容依存性と、X線ビーム中心からの半径距離に関連する空間依存性を有することである。これらの発見に基づき、洗練された確率的劣化モデルを開発した。このモデルは、クリーンな DRR を入力とし、非一様な物理情報に基づくノイズと拡散の多段階プロセスを適用して、高度に現実的な GFI スタイルの画像を生成する。これにより、堅牢な画像復元ネットワークの訓練に不可欠な高忠実度のペアデータセット（現実的な GFI & クリーン DRR）が作成される。 第3章で提示される第二の主要貢献は、エンドツーエンド追跡ワークフローの開発と検証である。このパイプラインは、臨床的なカラーフロスコピー画像（CFI）を入力とする。まず、CorrC2G アルゴリズムが CFI を高コントラストのグレースケール蛍光透視画像（GFI）に変換する。ソリューションの核心は、第2章のペアデータセットで訓練された Restormer（Transformer ベース）ネットワークである。このネットワークは「ドメイン橋渡し」として機能し、ノイズの多い臨床 GFI を入力とし、クリーンな DRR スタイルの画像を出力する。この「クリーン化された」画像は、クリーン DRRs のみで訓練された DUCK-Net セグメンテーションモデルに供給される。得られたマスクの2D重心を抽出、低パスフィルタで軌跡を平滑化し、最終的な3D位置を立体視から再構成する。 このフレームワークは、7人の肺がん患者の臨床データセット（100シーケンス）で検証され、埋め込まれたフィデューシャルマーカーを地面実測値として用いた場合、完全パイプラインで中央値3Dユークリッド誤差（MEE）1.53 mm（IQR: 1.22 mm）を達成した。 重要なことに、アブレーション研究は精度-速度トレードオフのニュアンスある分析を提供した。Restormer を含む完全パイプライン	

（和文 2,000 字程度 / 英文 800 語程度）
（about 800 words）

様式 7 号 (第 12 条, 第 31 条関係)

(様式 7 号) (Format No.7) 日本語版

ンが最高の精度 (1.53 mm MEE) を達成した。しかし、計算コストの高いスタイル転送ステップ (78 ms 追加) を省略した簡略化ワークフローは、わずかに精度が低下するのみで 1.62 mm MEE を達成した。これは、堅牢なデータ拡張 (ノイズ、ぼかし) で強化された DUCK-Net モデルが、すでにドメインギャップに対して高度に耐性があったことを示す。この発見は、実用的でリアルタイムの実装では、簡略化ワークフローが中程度の中央値精度の 5.6%犠牲で処理速度の 76.6%大幅向上を提供する優位で実行可能な代替案であることを強調する。

第 4 章で詳述される第三の貢献は、研究ワークフローの実用的ボトルネックへの対処である。CT 処理、DRR 生成、AI モデル訓練、バッチ推論、軌跡評価の断片化したプロセスは時間消費が多くエラー prone である。これを解決するため、専用の一体型デスクトップアプリケーション「MedKit」を開発した。PyQt、Python、PyTorch で構築されたこのソフトウェアは、4D-CT 可視化と DRR 生成か

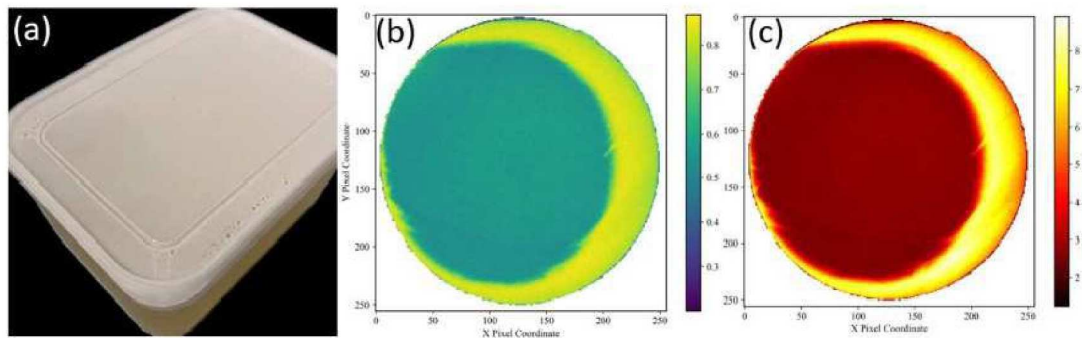


Figure 1. (a) ゼラチン充填容器。ゼラチン立方体の蛍光透視画像におけるノイズ特性 (背景除外)。 (b) 劣化誘発ノイズの確率および (c) 振幅のヒートマップで、1800 枚の減算画像 (フレームマイナス平均) からのピクセルごとの分散から導出。ノイズは中央領域 (Prob: ~ 0.56 , Amp: ~ 2.7) から右周辺弧 (Prob: ~ 0.83 , Amp: ~ 7.2) へ大幅に増加する。低いグレースケール値の領域 (例: 立方体エッジ) はノイズが低い。主要発見: ノイズ確率および振幅は X 線ビーム中心 (画像中心ではない) からの距離とともに増加する。このプロファイルは非単調で、ノイズは約 1 半径まで増加し、その後減少する。この減少は右側で顕著であり、ビーム中心のわずかな左方シフトに起因する。

ら Restormer/DUCK-Net による AI 駆動バッチ処理および自動 3D 評価まで、全体のパイプラインを統合する。このツールは、本論文の大規模実験を可能にする上で不可欠であった。結論として、本論文は実行可能で正確な非侵襲的マーカーレス肺腫瘍追跡フレームワークを実証した。主要な革新は、内容依存および空間依存の劣化モデルを開発し、最高の精度を実現したことである。しかし、主要な発見は、スタイル転送ステップを省略した簡略化ワークフローが、許容可能な精度 (1.62 mm MEE) とリアルタイム処理能力のより実用的バランスを提供することである。本研究は、物理モデリングから検証された AI ワークフローおよびそれをサポートするソフトウェアまで、完全なエンドツーエンドソリューションを提供し、より安全で正確な画像誘導放射線治療への道を拓く。

(和文 2,000 字程度 / 英文 800 語程度)
(about 800 words)

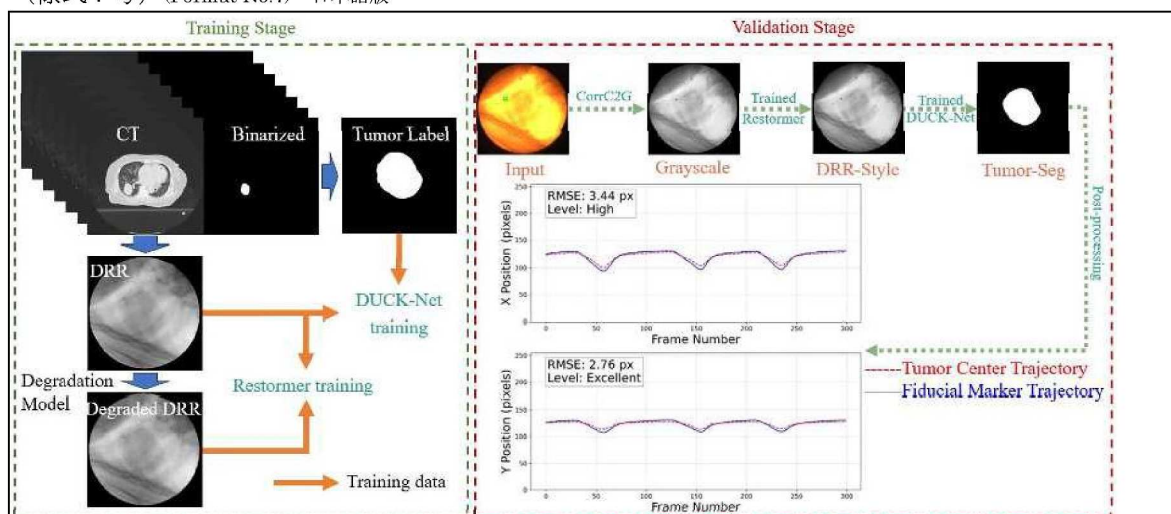


Figure 2. 提案パイプラインのワークフロー。入力 CFI の小さな緑の四角は、腫瘍位置を決定するための補助として埋め込まれた選択されたフィデュシャルマーカを表す。本研究では、光学フロー法を用いて選択されたフィデュシャルマーカの運動軌跡を取得し、それを予測腫瘍軌跡との比較に使用した。

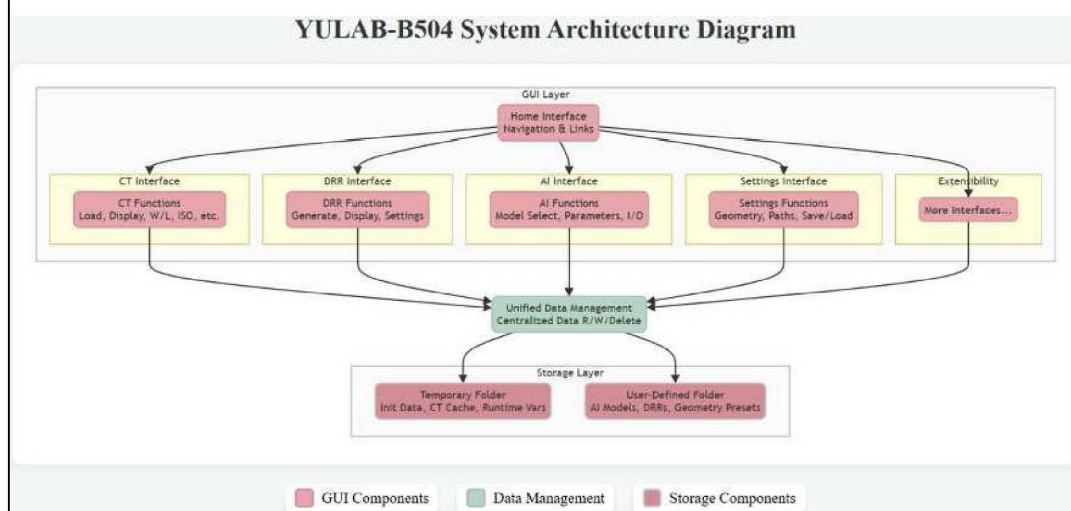


Figure 3. 開発されたアプリケーション YULAB-B504 のモジュラーシステムアーキテクチャ。

学 位 論 文 要 旨 (Summary of the Doctoral Dissertation)	
学位論文題目 (Dissertation Title)	Maker-less Lung Tumor Tracking from Real-time Color X-ray Fluoroscopic Images for Radiotherapy
氏 名 (Name)	Yan Yongxuan
Radiotherapy is one of the main treatments for lung cancer, but its efficacy is crucially influenced by respiratory-induced tumor motion. Image-Guided Radiation Therapy (IGRT) aims to mitigate this by tracking the tumor in real-time. The clinical gold standard involves surgically implanting fiducial markers, which is an invasive procedure associated with risks of complications such as pneumothorax and hemorrhage. Consequently, "marker-less" tracking methods are highly sought after, but their development is hindered by the poor soft-tissue contrast and high noise levels of intraoperative Kilovoltage (kV) fluoroscopic images. Deep learning (DL) models offer a promising solution for marker-less segmentation, but they require vast amounts of annotated training data. As manual annotation of noisy fluoroscopic sequences is infeasible, models are typically trained on synthetic Digitally Reconstructed Radiographs (DRRs) generated from planning CT scans. This creates a critical "domain gap": models trained on clean, synthetic DRRs may fail to generalize when confronted with the noisy, blurry, and stylistically different images captured during real clinical fluoroscopy. This dissertation presents a comprehensive framework aiming to overcome this domain gap to achieve accurate, real-time, marker-less lung tumor tracking using a cross-patient model, eliminating the need for patient-specific retraining. The contributions of this work are threefold: (1) the development of a novel, physics-informed degradation model to generate realistic training data; (2) the design and validation of a complete deep learning workflow for 3D tumor tracking; and (3) the implementation of an integrated software platform that streamlines the entire research and validation pipeline. The first major contribution, detailed in Chapter 2, is a novel methodology for bridging the domain gap. This work posits that simple noise simulation (e.g., Gaussian noise) is insufficient. Through empirical analysis of a custom gelatin phantom imaged by the SyncTraX stereoscopic system, we performed a statistical investigation of fluoroscopic noise. The key finding was that noise is not random: it is both content-dependent, correlating with pixel intensity and material density, and spatially-dependent, related to the radial distance from the X-ray beam center. Based on these findings, a sophisticated stochastic degradation model was developed. It takes a clean DRR and applies a multi-stage process of non-uniform, physically-informed noise and diffusion to generate a highly realistic, GFI-style image. This creates a high-fidelity, paired dataset (Realistic GFI & Clean DRR) essential for training a robust image restoration network. The second major contribution, presented in Chapter 3, is the development and validation of the end-to-end tracking workflow. The pipeline takes a clinical color fluoroscopic image (CFI) as input. First, the CorrC2G algorithm converts the CFI to a high-contrast grayscale fluoroscopic image (GFI). The core of the solution is a Restormer (Transformer-based) network, which is trained on the paired dataset from Chapter 2. This network functions as a "domain bridge," taking the noisy clinical GFI as input and outputting a clean, DRR-style image. This "cleaned" image is then fed into a DUCK-Net segmentation model, which is trained only on clean DRRs. The 2D centroid of the resulting mask is extracted, a low-pass filter smooths the trajectory, and the final 3D	

position is reconstructed from the stereo views.

This framework was validated on a clinical dataset from seven lung cancer patients (100 sequences), achieving a median 3D Euclidean Error (MEE) of 1.53 mm (IQR: 1.22 mm) with the full pipeline, using implanted fiducial markers as the ground truth.

Crucially, the ablation studies provided a nuanced analysis of the accuracy-speed trade-off. The full pipeline with Restormer achieved the highest accuracy (1.53 mm MEE). However, a streamlined workflow omitting the computationally expensive style-transfer step (which adds 78 ms) was only marginally less accurate, achieving a 1.62 mm MEE. This demonstrates that the

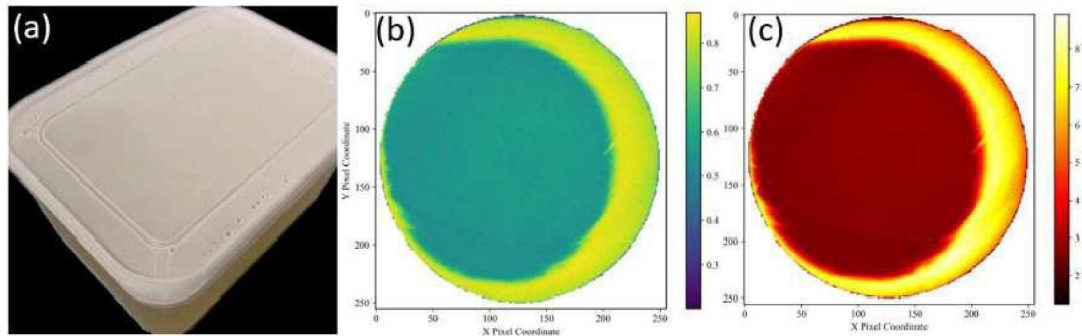


Figure 1. (a) Gelatin-filled container. Noise characteristics in a gelatin cube fluoroscopic image (background excluded). (b) Heatmap of degradation-induced noise probability and (c) amplitude, derived from pixel-wise variance across 1800 subtracted images (frame minus average). Noise increases significantly from the central region (Prob: ~ 0.56 , Amp: ~ 2.7) to the right peripheral arc (Prob: ~ 0.83 , Amp: ~ 7.2). Regions with lower grayscale values (e.g., cube edges) exhibit lower noise. Key Finding: Noise probability and amplitude increase with distance from the X-ray beam center (not image center). This profile is non-monotonic: noise increases up to approximately one radius, then declines. This decline is noticeable on the right side, attributed to a slight leftward shift of the beam center.

DUCK-Net model, fortified with robust data augmentation (noise, blur), was already highly resilient to the domain gap. This finding highlights that for practical, real-time implementation, the streamlined workflow is a superior and viable alternative, sacrificing a modest 5.6% in median accuracy for a substantial 76.6% gain in processing speed.

The third contribution, detailed in Chapter 4, addresses the practical bottlenecks of the research workflow itself. The fragmented process of CT handling, DRR generation, AI model training, batch inference, and trajectory evaluation is time-consuming and error-prone. To solve this, a dedicated, all-in-one desktop application, "MedKit," was developed. Built with PyQt, Python, and PyTorch, this software integrates the entire pipeline—from 4D-CT visualization and DRR generation to AI-driven batch processing with Restormer/DUCK-Net and automated 3D evaluation. This tool was instrumental in enabling the large-scale experiments of this dissertation. In conclusion, this dissertation successfully demonstrates a feasible, accurate, and non-invasive marker-less lung tumor tracking framework. The primary innovation is the development of a content- and spatially-dependent degradation model that effectively enabled the highest accuracy. However, the key finding is that a streamlined workflow, omitting the style transfer step, provides a more practical balance of acceptable accuracy (1.62 mm MEE) and real-time processing capability. This work provides a complete, end-to-end solution—from physical modeling to a validated AI workflow and the software to support it—paving the way for safer and more precise image-guided radiotherapy.

様式 7 号 (第 12 条, 第 31 条関係)
(様式 7 号) (Format No.7) 英語版

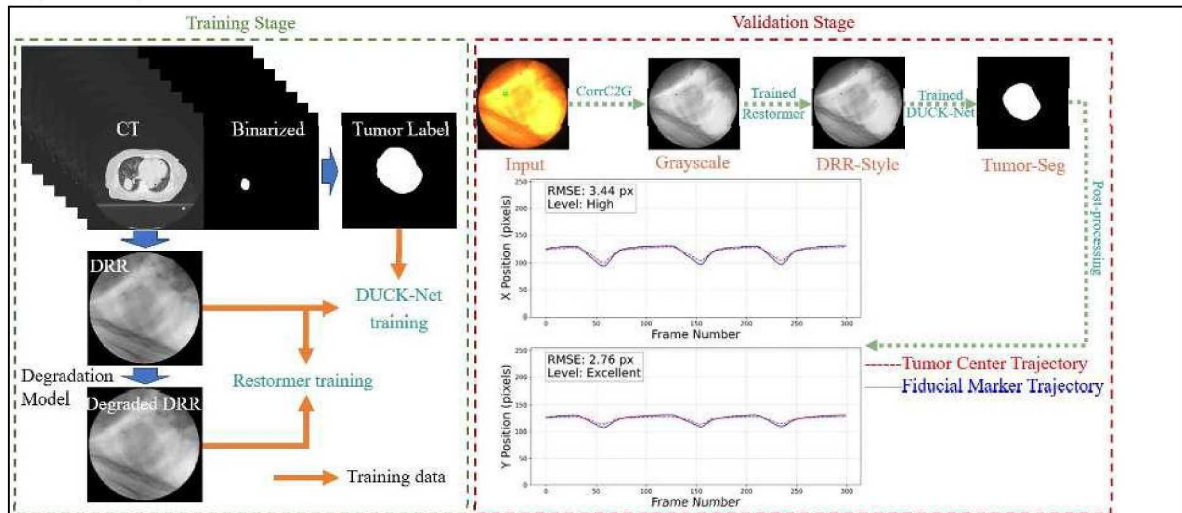


Figure 2. Workflow of the proposed pipeline. The small green square from the input CFI represents the selected fiducial marker; which was implanted around the tumor to assist in determining the tumor location. In this study, we used the optical flow method to obtain the motion trajectory of the selected fiducial marker, which was used for comparison against the predicted tumor trajectory.

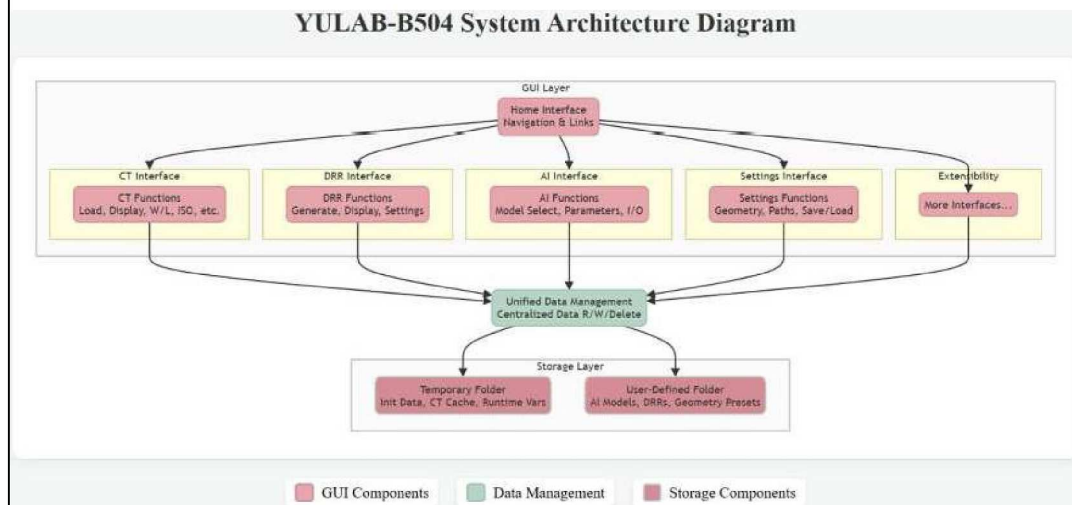


Figure 3. The modular system architecture of the developed application YULAB-B504.

(和文 2,000 字程度 / 英文 800 語程度)
(about 800 words)

(様式 9 号)

学位論文審査の結果及び最終試験の結果報告書

山口大学大学院創成科学研究科

氏 名	Yan Yongxuan
審 査 委 員	主 査： 藤井 文武
	副 査： 間普 真吾
	副 査： 大木 順司
	副 査： 山口 皓平
	副 査： 椎木 健裕
論 文 題 目	Marker-less Lung Tumor Tracking from Real-time Color X-ray Fluoroscopic Images for Radiotherapy (放射線治療のためのリアルタイムカラーX線透視画像からのマーカーレス肺腫瘍追跡)
【論文審査の結果及び最終試験の結果】	
悪性腫瘍の放射線治療では、正常な細胞への放射線投与を避けながら腫瘍に処方線量を集中させる精度管理が極めて重要である。しかしながら、肺腫瘍は患者の呼吸により体内で位置が変化する「呼吸性移動」を示すため、腫瘍位置変化を踏まえて放射線投与精度を確保する治療線量投与法の開発が臨床技術として重要な意味を持つ。放射線治療におけるこの技術を「動体追跡照射」と呼ぶ。 申請者は、山口大学医学部附属病院に設置され臨床運用されている動体追跡装置（商品名：SyncTraX、島津製作所）により 30Hz のレートで治療中に得られる kV のカラーステレオ透視像（CFI）を用いて、画像情報のみで実時間の腫瘍位置追跡を行う画像処理システムの構築に関する研究を行い、2つのシステム構成を提案した。 申請者の第一の成果は、画像セグメンテーションに利用される U-net を、①カラーからグレースケールへの変換、②腫瘍像と骨とのオーバーラップの影響低減、③腫瘍のセグメンテーション、の 3 つの目的に特化して学習させ、それらを順次利用することで、治療中に SyncTraX を用いて得られる CFI 上の腫瘍位置特定を行うシステムの提案である。 申請者の第二の成果は、CFI 取得時に印加されるノイズに注目し、腫瘍セグメンテーション精度の改善を図ろうとしたものである。腫瘍セグメンテーションを行うためのモデルを DUCK-net に変更、その学習データとしてゼラチンファントムを利用して同定したノイズモデルを用いて意図的に劣化させた DRR（デジタル再構成透視像）由来の画像を用いる等複数の改善を試み、その結果として、同定された腫瘍位置の誤差が大幅に低減されることを示した。また、対象とする患者の腫瘍近辺に埋め込まれたマーカの SyncTraX を用いた 3 次元位置測定結果と、提案した画像による肺腫瘍追跡結果との比較では、臨床適用可能なレベルまで追跡誤差が低減化されたことも示した。	

(様式 9 号)

なお、申請者は、研究に伴い開発したソフトウェアを公開し、当該領域の今後の研究進展のための貢献も行っていることも評価に値する。

公聴会は、2月7日に20名の参加を得て対面により行った。主な質問内容は、腫瘍位置特定に用いた DUCK-net モデルの原理とハイパーパラメータに関するもの、腫瘍が複雑な形状を示している場合の影響に関するもの、動体追跡装置を変更した場合の影響に関するもの、および 3 次元追跡誤差の時系列について、誤差が大きな領域と呼吸相との関係に関するもので、いずれの質問にも申請者からの的確な回答がなされた。

以上より本研究は、独創性、信頼性、有効性、実用性共に優れ、博士（工学）の論文に十分値するものと判断した。

論文内容及び審査会、公聴会での質問に対する応答などから、最終試験は合格とした。

なお、主要な論文の発表状況は下記のとおりである。（関連論文 計4編、参考論文 計1編）

- 1) Yongxuan Yan, Fumitake Fujii, Takehiro Shiinoki, and Shengping Liu, Markerless Lung Tumor Localization From Intraoperative Stereo Color Fluoroscopic Images for Radiotherapy, IEEE Access, Vol.12, pp. 40809-40826. 2024
- 2) Yongxuan Yan, Fumitake Fujii, and Takehiro Shiinoki, Marker-Less Lung Tumor Tracking from Real-Time Color X-Ray Fluoroscopic Images Using Cross-Patient Deep Learning Model, Bioengineering, Vol.12, No.11, 1197. 2025
- 3) Yongxuan Yan, Fumitake Fujii, Takehiro Shiinoki, Shengping Liu, Markerless Lung Tumor Detection from Color Fluoroscopic Images Via Cascade U-Net Based Patient-Specific Models, The 65th American Association of Physicists in Medicine Annual Meeting, Session: MO-345-IePD-F6 - IGRT Tracking, 2023