Detecting the Sensing Area of A Laparoscopic Probe in Minimally Invasive Cancer Surgery

Design and Specification Proposal

COMP702 – M.Sc. project (2024/25)

Submitted by

GEO RAJU

(201855055)

under the supervision of

DR. BAORU HUANG

Department of Computer Science

University of Liverpool

1 Statement of ethical compliance: A0

Data Category: A (No sensitive personal data)

Participant Category: 0 (No use of human participants in any activity)

I confirm that I have read the ethical guidelines and will follow them during this project. I will use the Coffbee dataset from prior work [1] available at Imperial College London Box repository in CA1. The dataset contains stereo laparoscopic images with ground truth sensing area annotations obtained using laser-equipped gamma probes on artificial tissue phantoms rather than human tissue, eliminating ethical concerns while maintaining surgical realism. Human participants are not planned for requirements analysis or evaluation in CA3; the evaluation and testing are planned to be conducted using the existing dataset.

- CA1: Data Sources.
- CA3: Requirements Analysis; Evaluation.

2 Project description

In minimally invasive cancer surgery, surgeons rely on gamma probes such as the SENSEI probe [2] to detect radioactive tracers that help locate cancerous tissue and lymph nodes during procedures. While these probes can accurately detect gamma radiation emitted from injected radiotracers, they present a critical visualisation challenge: the probe is non-imaging and does not provide any visual indication of where the detected activity originates on the tissue surface.

This lack of spatial visualisation creates significant difficulties for surgeons who need to precisely identify the sensing area - the specific region on tissue where the probe is detecting gamma activity. Current pre-operative imaging systems like PET and CT scans, while useful for initial diagnosis, cannot provide the real-time, precise localisation needed during live surgical procedures.

The challenge is further complicated by the fact that the probe operates with an air gap from the tissue surface, making it difficult to determine the exact intersection point between the probe's sensing axis and the tissue. This uncertainty in sensing area localisation can impact surgical precision and decision-making during critical cancer resection procedures.

This project addresses the fundamental problem of accurately detecting the sensing area of gamma probes in real-time during laparoscopic cancer surgery, where precise tissue identification is essential for successful outcomes.

3 Aim and Requirements

3.1 The main aim

The main aim of the project is to propose a novel and effective method to detect the sensing area of a gamma probe ('SENSEI') [2] from the surgery scene in 3D and locate 'hot' tissue. The project will develop a deep learning-based approach that processes stereo laparoscopic images to accurately predict the probe axis-surface intersection points, providing real-time sensing area visualisation for surgeons during minimally invasive cancer surgery procedures.

3.2 Requirements

The following functional requirements define what the deep learning model must accomplish within the scope of this project:

- Identify the probe's sensing area in real time during minimally invasive cancer surgery

- Deliver accurate spatial localisation of the sensing area relative to tissue surface
- Handle varying probe orientations and positions within the surgical field of view
- Provide confidence indication for sensing area predictions

Based on these functional requirements, the following technical specifications must be met:

- Process exactly two stereo laparoscopic images (left and right) as input and output the 3D coordinates (x, y, z) of the probe axis–tissue surface intersection point in the stereo camera coordinate system.
- Achieve performance on par with recent benchmarks in sensing area detection: 2D error $\leq$ 40 pixels (mean) and 3D error $\leq$ 3.5 mm (mean Euclidean distance) on the test dataset.
- Maintain real-time inference capability with processing time $\leq$ 100 ms per stereo pair on standard GPU hardware (NVIDIA GTX 1080 or equivalent).
- Achieve generalized performance with R^2 score $\geq$ 0.75 for each coordinate (x, y, z) in 3D coordinate prediction on the test dataset.

If time permits, the following desirable features will be implemented:

- Explore different deep learning backbone architectures (e.g., ResNet, EfficientNet, Swin Transformer) to improve prediction performance.
- Provide per-coordinate uncertainty estimates (x, y, z) as well as overall confidence scores for each prediction.

4 Key literature and background reading

Deep learning has emerged as a transformative technology in the field of minimally invasive surgery (MIS), particularly with applications in surgical tool guidance and pose estimation [3], [4]. The integration of deep learning technologies in laparoscopic surgery is progressively reshaping minimally invasive procedures, with CNN architectures [3] coupled with tailored loss functions [5] and innovative training strategies [6] enhancing intraoperative capabilities while aligning visualisation with surgical demands.

The foundational work by Huang et al. [1] introduces the Coffbee dataset and establishes baseline methods for sensing area detection using laser-equipped gamma probes on artificial tissue phantoms. This work provides the primary benchmark for our project, demonstrating initial approaches to 2D and 3D localisation of probe sensing areas in laparoscopic environments using simple regression networks with ResNet backbones, achieving 2D errors of 52.9 pixels and 3D errors of 7.4mm. Building upon this foundation, Xu et al. [7] propose a Nested ResNet architecture with three-branch framework incorporating depth estimation and orientation guidance, demonstrating significant improvements with 22.10% reduction in 2D mean error and 41.67% reduction in 3D mean error over previous methods.

Recent methodological advances complement these foundational approaches. Zhou et al. [8] introduce a domain randomization approach for the pose estimation of surgical tools, enhancing model robustness by simulating varied environmental conditions across different surgical setups. Yin et al. [4] propose error analysis driven network modifications for surgical tool detection, introducing three-dimensional attention mechanisms and double-headed detection architectures that achieve improved localisation accuracy in challenging laparoscopic environments with insufficient illumination and extreme camera angles.

These advances collectively demonstrate the potential for sophisticated computer vision techniques to address the complex challenges of spatial localisation and instrument tracking in minimally invasive surgical environments, providing the foundation for accurate gamma probe sensing area detection in real-time surgical applications.

5 Development and implementation summary

The development of the sensing area detection system will follow a systematic approach divided into three main phases: data preparation and analysis, model development and training, and evaluation and optimisation.

- **Phase 1: Data preparation and analysis**

  The implementation will begin with comprehensive analysis of the Coffbee dataset from [1], which contains stereo laparoscopic images with ground truth sensing area annotations obtained using laser-equipped gamma probes on artificial tissue phantoms. Key activities include:

  - Dataset exploration and statistical analysis of sensing area distributions
  - Image preprocessing pipeline development for stereo pair normalisation
  - Train/validation/test split establishment following existing benchmark protocols

- **Phase 2: Model development and training**

  The core implementation phase will focus on developing a deep learning architecture with three-stage training capable of processing stereo image pairs to predict 3D intersection coordinates:

  - Implementation of a deep learning architecture with probe axis regression and intersection point prediction modules
  - Execution of three-stage training strategy: axis pretraining, intersection training, and end-to-end fine-tuning (as illustrated in Figure 1)
  - Development of multi-scale feature fusion using ResNet backbone with stereo image processing

- **Phase 3: Evaluation and optimisation**

  The final phase will focus on comprehensive model evaluation and performance optimisation:

  - Quantitative evaluation using 2D pixel error, 3D Euclidean distance error, and R^2 scores
  - Real-time performance testing on target GPU hardware (NVIDIA GTX 1080)
  - Comparative analysis against existing benchmarks from recent literature

The implementation will utilize PyTorch framework for model development, with standard computer vision libraries for image processing. Development will be conducted on GPU-enabled systems to ensure real-time performance capabilities suitable for surgical applications.

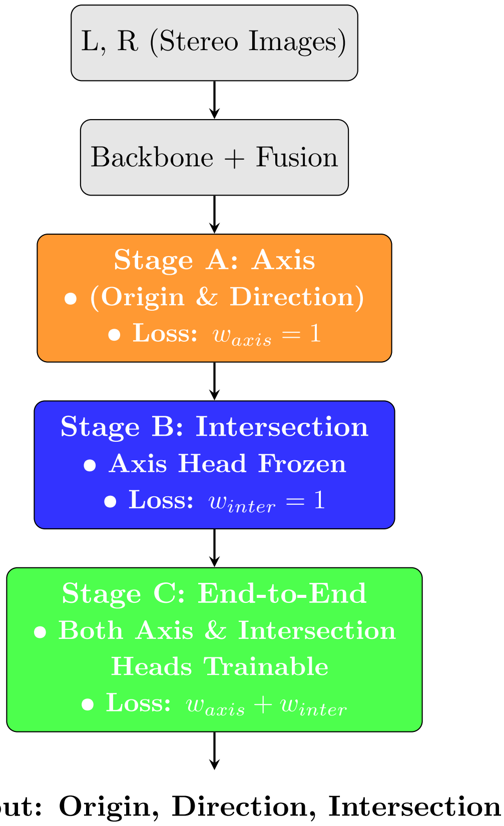

Figure 1: Three-stage training methodology for stereo vision-based gamma probe sensing area detection

6 Data source

This project will utilize the Coffbee dataset from [1], available at Imperial College London Box repository. The Coffbee dataset is specifically designed for sensing area detection research and contains stereo laparoscopic images with ground truth sensing area annotations obtained using laser-equipped gamma probes on artificial tissue phantoms. The dataset includes depth maps and probe axis information, enabling comprehensive evaluation of both 2D and 3D localisation accuracy.

7 Testing and evaluation

The system will undergo comprehensive evaluation using multiple metrics to assess both accuracy and performance capabilities required for surgical applications.

- **Accuracy evaluation**

  Quantitative assessment will be conducted using the following metrics aligned with the technical requirements:

  - **2D localisation accuracy**: Mean absolute error in pixel coordinates, targeting $\leq$ 40 pixels with additional reporting of median error and standard deviation
  - **3D localisation accuracy**: Mean Euclidean distance error in millimeters, targeting $\leq$ 3.5 mm between predicted and ground truth intersection points
  - **Per-coordinate R^2 score**: Individual coefficient of determination for x, y, and z coordinates, targeting $\geq$ 0.75 for each dimension to ensure balanced performance

- **Performance evaluation**

  Real-time capability assessment will include:

  – **Inference speed**: Processing time per stereo image pair on NVIDIA GTX 1080 hardware, targeting $\leq$ 100 ms for surgical viability

  – **Memory usage**: GPU memory consumption during inference to ensure deployability on standard surgical systems

  – **Throughput analysis**: Frames per second capability for continuous surgical video processing

- **Comparative analysis**

  The system will be benchmarked against existing methods from recent literature, particularly the baseline results reported in Huang et al. [1], to demonstrate performance improvements and validate the effectiveness of the proposed approach.

- **Evaluation protocol**

  Testing will follow established protocols using the reserved test set from the Coffbee dataset, ensuring no data leakage from training or validation phases. Results will be reported with confidence intervals and statistical significance testing where appropriate.

8 UI/UX mockup

While the primary focus of this project is the development and evaluation of the deep learning model for sensing area detection, a demonstration interface will be developed for static image processing and result visualisation.

The interface will be implemented using Python GUI frameworks (PyQt or Tkinter) with OpenCV for image display and matplotlib for result visualisation. This static processing approach ensures compatibility with the PyTorch-based model implementation while providing clear demonstration of sensing area detection capabilities, as illustrated in Figure 2.

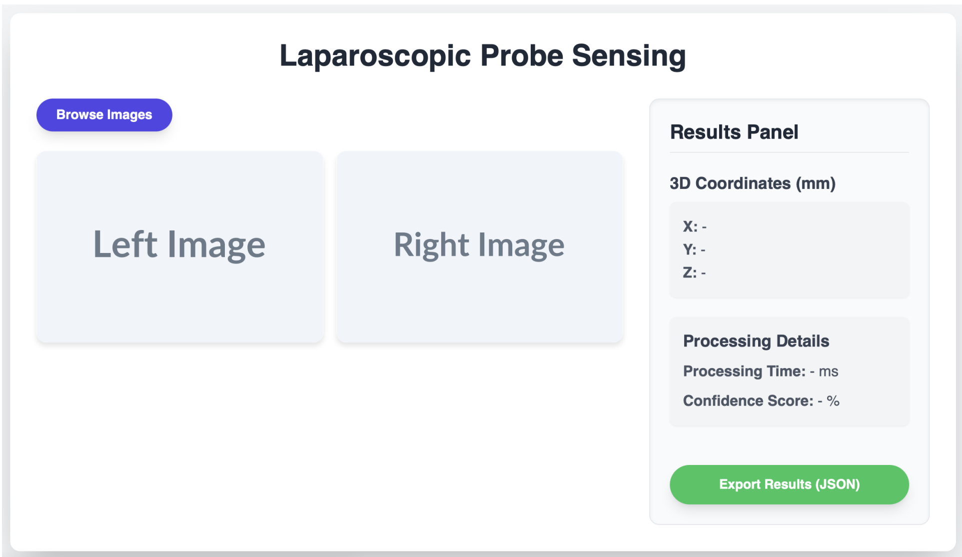

Figure 2: UI/UX mockup interface design

9 Project ethics and human participants

This project has been designed to comply with ethical guidelines and minimize any ethical concerns through careful data source selection and methodology design.

Data usage: The project exclusively utilizes the Coffbee dataset from [1], which contains stereo laparoscopic images with ground truth sensing area annotations obtained using laser-equipped gamma probes on artificial tissue phantoms rather than human tissue. This approach eliminates ethical concerns while maintaining surgical realism and relevance for the research objectives.

Human participants: No human participants are involved in any aspect of this research project. All data collection, model training, validation, and testing will be conducted using the existing synthetic dataset. No surveys, interviews, or human subject experiments are planned or required for the successful completion of this project.

Ethical compliance: The use of artificial tissue phantoms and synthetic data ensures full compliance with ethical guidelines while providing realistic surgical scenarios for model development and evaluation. This approach allows for comprehensive research into sensing area detection without any privacy, consent, or safety concerns associated with human subject research.

Data access and storage: The Coffbee dataset is publicly available through established academic channels, and all data handling will follow standard data management practices for research projects. No sensitive personal data will be collected, processed, or stored during the course of this project.

10 BCS project criteria

1. Application of practical and analytical skills

   This project will utilize practical and analytical skills from my Data Science and AI degree program. The development of a deep learning model for sensing area detection requires comprehensive application of machine learning techniques, computer vision algorithms, and software engineering practices. Key practical skills include PyTorch implementation, stereo image processing, statistical analysis for model evaluation, and real-time system optimisation. Analytical skills will be demonstrated through rigorous performance evaluation using multiple metrics (2D/3D error analysis, R^2 scores), comparative benchmarking against existing methods, and systematic analysis of model architecture choices including the three-stage training approach.

2. Innovation and creativity

   The project demonstrates innovation through the application of advanced deep learning techniques to a critical problem in minimally invasive cancer surgery. Creative elements include the development of a three-stage training strategy, integration of multi-scale feature fusion for stereo image processing, and the use of advanced loss functions for improved accuracy. The transformation of a geometric intersection problem into a high-dimensional feature learning problem represents a novel approach that leverages state-of-the-art computer vision techniques for surgical applications. The project creatively combines stereo vision, regression-based coordinate prediction, and real-time performance optimisation to address a previously unsolved visualisation challenge in surgical oncology.

11 Project plan

The project will be executed over a four-month period from August 4, 2025 to November 28, 2025, following the structured three-phase development approach outlined in the implementation summary, as shown in Figure 3.

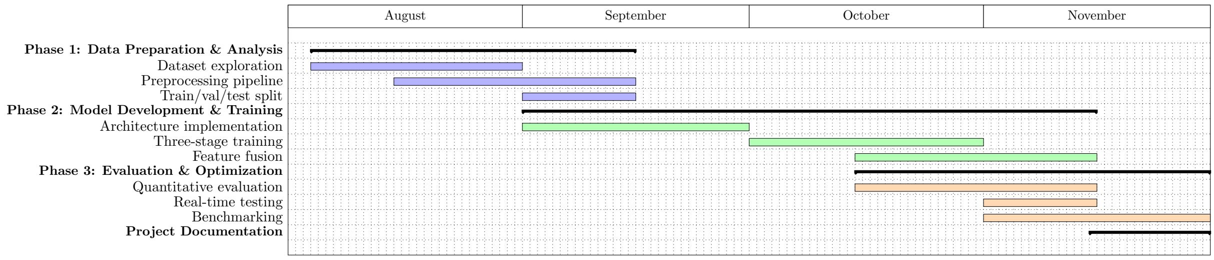

Figure 3: Project Timeline (August 4 - November 28, 2025)

12 Risks and contingency plans

Table 1 outlines the identified risks, their likelihood and impact assessments, and corresponding contingency plans.

Risk	**Contingency**	**Likelihood**	**Impact**
Model fails to achieve target accuracy metrics	Implement alternative architectures (EfficientNet, Swin Transformer)	Medium	High
Insufficient computational resources for training	Use cloud computing resources (Google Colab Pro, AWS)	Low	High
Dataset quality issues or missing annotations	Implement robust data preprocessing and validation	Low	Medium
Real-time performance requirements not met	Optimize model architecture and implement quantization	Medium	Medium
Training convergence difficulties	Adjust learning rates, batch sizes, and implement curriculum learning	Medium	Medium
Implementation delays affecting timeline	Reduce optional features scope and focus on core requirements	Low	Low

Table 1: Risk management plan

References

[1] B. Huang et al., "Sensing area detection for laparoscopic gamma probe using deep learning," in *Proc. Medical Imaging Conference*, 2023.

[2] Lightpoint Medical Ltd., "SENSEI product information," [Online]. Available: `https://img1.wsimg.com/blobby/go/5d369e4e-24bb-4f1f-b3be-ac252bedda10/downloads/SENSEI%20product%20flier%20ENGLISH%20(LPH040%20V.2).pdf?ver=1754269577667`. [Accessed: Dec. 10, 2024].

[3] H. ElMoaqet, H. Qaddoura, M. Ryalat, N. Almtireen, T. Alshirbaji, N. Jalal, and K. Möeller, "Deep learning architectures for single-label and multi-label surgical tool classification in minimally invasive surgeries," *Applied Sciences*, vol. 15, no. 11, p. 6121, 2025.

[4] B. Yin, S. Wang, S. Lu, G. Wang, and L. Dong, "Error analysis driven network modification for surgical tools detection in laparoscopic frames," *International Journal of Imaging Systems and Technology*, vol. 33, no. 1, pp. 192–203, 2022.

[5] A. Casella, S. Moccia, C. Carlini, E. Frontoni, E. Momi, and L. Mattos, ”Nephcnn: a deep-learning framework for vessel segmentation in nephrectomy laparoscopic videos,” in *Proc. IEEE International Conference on Pattern Recognition*, 2021, pp. 6144–6149.

[6] G. Loza, P. Valdastri, and S. Ali, ”Real-time surgical tool detection with multi-scale positional encoding and contrastive learning,” *Healthcare Technology Letters*, vol. 11, no. 2–3, pp. 48–58, 2023.

[7] L. Xu et al., ”Improved sensing area detection using nested ResNet architecture,” *Computer-Assisted Surgery Journal*, 2024.

[8] C. Zhou, L. Wang, B. Wu, and K. Xu, ”A markerless 3d tracking framework for continuum surgical tools using a surgical tool partial pose estimation network based on domain randomization,” *Advanced Intelligent Systems*, vol. 6, no. 4, 2024.