ARMADILLO: Augmented Reality Machine-Assisted Detection and Inference in Laparoscopic Liver Operations

Group Report for COMP5530

ABDUL KARIM ABBAS*, JAMES BORGARS*, AODHAN GALLAGHER*, AHMAD NAJMI MOHAMAD SHAHIR*, JIBRAN RAJA*, ABHINAV RAMAKRISHNAN*, and THEODORA VRAIMAKIS*, University of Leeds, United Kingdom
SHARIB ALI† and RAFAEL KUFFNER DOS ANJOS†‡, University of Leeds, United Kingdom

Laparoscopy is an approach to liver surgery which reduces complications and recovery time, and can also harness developments in machine-assisted surgery. In this report, we outline and implement the end-to-end process of performing 3D-2D registration using only a preoperative liver mesh and intraoperative laparoscopy footage, with no human involvement in respect to landmark annotation and alignment. We present novel research in the fields of 2D and 3D landmark segmentation, with best-in-class results for the dataset. We study iterative and deep learning approaches in the area of 3D-2D registration, with silhouette extrapolation implemented for improved results. Finally, we explore hardware implementation of the pipeline and data visualisation techniques using an Augmented Reality headset. Our results include a 30% relative increase in 2D segmentation precision, 36% improvement in 3D segmentation distance, and 31% improvement in reprojection error in registration compared to leading research.

CCS Concepts: • **Human-centered computing** → *Mixed / augmented reality*; *Human computer interaction (HCI)*; • **Computing methodologies** → **Computer vision**; **Image segmentation**; Tracking; *Object detection*; **Supervised learning**; **Neural networks**; *Point-based models*; • **Applied computing** → Life and medical sciences.

Additional Key Words and Phrases: segmentation, registration, deep learning, image-guided intervention, surgical data science, laparoscopy

1 INTRODUCTION

Laparoscopic liver surgery (also known as minimally-invasive liver surgery and keyhole surgery) is a surgical approach which minimises recovery time and the probability of complications [Slakey et al. 2013]. This surgical approach also facilitates developments in the area of machine-assisted surgery due to its use of a camera, such as 3D-2D registration of the liver, where a preoperative 3D mesh of the liver, including anatomical landmarks such as tumours and vessels, can be superimposed onto the liver in real-time during surgery.

In this report, we present an implemented end-to-end pipeline, automating the process of 3D-2D registration of the liver. This includes segmentation models of both preoperative 3D meshes, and intraoperative 2D laparoscopic images, which have both in of themselves warranted novel research currently under review, having outperformed prior research on the same dataset [Ali et al. 2025]. Model outputs are utilised in a registration pipeline that does not require manually annotated data. A visualisation implementation has also been completed to visualise segmentation predictions and create a model navigation environment. Our results have led to substantial improvement in all tasks of the pipeline.

2 BACKGROUND RESEARCH

2.1 2D Segmentation

Ronneberger et al. present the 'U-shaped architecture' for Fully Convolutional Networks (FCNs) in the form of UNet, proving that large datasets were not required for high accuracy in the field of biomedical segmentation [Ronneberger et al. 2015]. The UNet architecture consists of a contracting path, which has pooling layers, an expansive path with up-convolutions, with these two paths connected by a bottleneck and skip connections, achieving increased performance at reduced inference times compared to previous models [Ronneberger et al. 2015].

UNet++ by Zhou et al. builds upon the UNet architecture with a greater number of convolution blocks, dense skip connection pathways, and deep supervision [Zhou et al. 2020]. UNet3+ further develops upon the ideas of UNet++, proposing full-scale skip connections where each convolutional block in the contracting path connects to its opposing and below blocks in the expansive path; the bottleneck and expansive path is supervised by the ground truth and has skip connections to every block further up the path [Huang et al. 2020]. ResUNet is a deep residual UNet-based model, replacing the standard Convolution-ReLU block with residual convolution blocks utilising batch normalisation [Zhang et al. 2018]. Jha et al. propose ResUNet++, modifying the ResUNet architecture for medical image segmentation through the addition of squeeze-excitation blocks to dynamically weight convolutional channels, ASPP for increased context when classifying a pixel, and the introduction of attention for enhanced feature quality [Jha et al. 2019].

Chen et al. propose DeepLabV3+, building on top of the DeepLabV3 architecture with the addition of depth-wise separable convolution to both ASPP and decoder modules, resulting in improved performance, having been tested on non-medical benchmarks [Chen et al. 2018].

Various implementations of UNet have been thoroughly evaluated alongside ResUNet in the Liver Tumour Segmentation (LiTS) benchmark [Bilic et al. 2023], with UNet++ and UNet3+ both demonstrating their outperformance of UNet on the benchmark [Huang et al. 2020; Zhou et al. 2020], highlighting their relevance to the

*Sharing first authorship
†Also with the role of Supervisor.
‡Also with the role of Assessor.

Authors' addresses: Abdul Karim Abbas, sc21aka@leeds.ac.uk; James Borgars, sc20jdb@leeds.ac.uk; Aodhan Gallagher, sc20ag@leeds.ac.uk; Ahmad Najmi Mohamad Shahir, sc21anbm@leeds.ac.uk; Jibran Raja, ed20j3r@leeds.ac.uk; Abhinav Ramakrishnan, sc21a2r@leeds.ac.uk; Theodora Vraimakis, sc20tmv@leeds.ac.uk, University of Leeds, Woodhouse, Leeds, West Yorkshire, United Kingdom, LS2 9BW; Sharib Ali, s.s.ali@leeds.ac.uk; Rafael Kuffner dos Anjos, r.kuffnerdosanjos@leeds.ac.uk, University of Leeds, Woodhouse, Leeds, West Yorkshire, United Kingdom, LS2 9BW.