Development and Validation of an AI-Based System for Classification, Detection and Tracking of Bleeding in Laparoscopic Surgery

Xin Wang，MD[1]，Bing Peng，MD[1]，Ang Li，MD[1]，Jie Liu，BS[2]，Dian Qin，PhD[2]，Yuxian Wang，BS[2].

1. Department of General Surgery, Division of Pancreatic Surgery, West China Hospital of Sichuan University；2. ChengDu Withai Innovations Technology Company, Chengdu, China

- **Introduction：** The advancement of minimally invasive technologies has expanded the use of laparoscopic techniques in various surgical fields, significantly improving patient outcomes by reducing recovery times and minimizing trauma. Despite this, intraoperative bleeding remains a prevalent issue, accounting for 80% to 90% of all adverse events and complicating surgery. Excessive bleeding not only necessitates surgical conversions but also elevates the risk of postoperative complications, adversely affecting long-term outcomes for cancer patients. This study develops an AI-based algorithm to classify, detect, and track intraoperative bleeding, and preliminarily validates its clinical feasibility and efficacy.

  Methods: In this study conducted from 2023 to 2024, 257 surgical videos including hepatobiliary, pancreatic, gastrointestinal, and gynecology surgery were collected from 25 hospitals. The data, categorized into three bleeding categories—none, mild oozing, and significantly vascular bleeding—were used to train，validation and test a deep learning model for bleeding classification, detection and tracking in surgeries. We developed a multi-task hierarchical deep learning framework that includes a 3D Patch Embedding module for structured data representation, progressive 'Hiera stage' modules for feature extraction, and specialized output heads for precise task performance (Figure 1). The model's effectiveness is demonstrated through metrics such as precision, recall, F1-score for bleeding classification, and mean distance between predicted and ground truth for bleeding sites detection and tracking. Finally, we deployed the algorithm on the previously developed SurgSmart intelligent workstation, running it in real time during surgeries to verify its feasibility and effectiveness.

- **Results:** As illustrated in Figure 2, 9,499/5,662 clips and 12,436/162,421 frames of mild/significant vascular bleeding were annotated for the training, validation, and testing of the model. The model demonstrated robust performance across various tasks, achieving a precision of 0.830, a recall of 0.776, and an F1-score of 0.799 in bleeding classification, along with a mean distance error of 0.087 in bleeding point detection and tracking. When deployed on the SurgSmart workstation and used in the operating room (Figure 3), over 80% of surgeons reported that the algorithm exceeded surgeons in both accuracy and speed, indicating its strong potential to improve hemostasis in future procedures.

- **Conclusions：** This study developed an AI algorithm that effectively classifies, detects, and tracks bleeding in real-time across diverse procedures, demonstrating excellent accuracy and significant potential for clinical application.

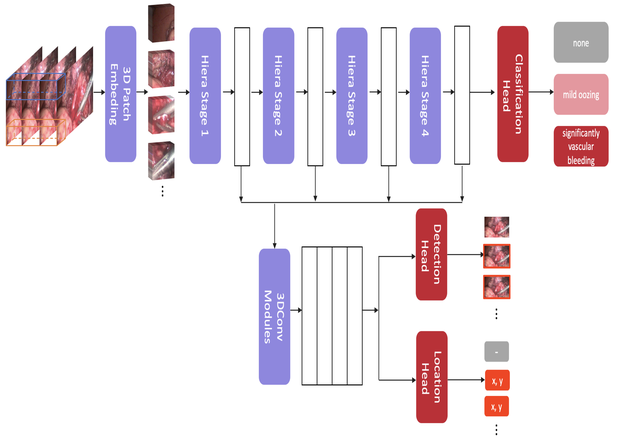


Figure 1 Algorithm development



Figure 3 SurgSmart M3 workstation

Surgical Dataset									
	None		Mild Ooze			Significantly vascular bleeding			
	Cases	Clips	Cases	Bleeding Clips	Frames of bleeding site detection	Cases	Bleeding Clips	Frames of bleeding site detection	
Train	145	13280	138	4942	6607	127	3062	90730	
Validation	55	4949	55	2647	3018	50	1494	41483	
Test	57	5256	56	1910	2811	54	1106	30208	
Total	257	23485	249	9499	12436	231	5662	162421	

Algorithm performance				
	Metrics			Mean
Precision	0.894 (95% CI: 0.886-0.902)	0.772 (95% CI: 0.751-0.791)	0.823 (95% CI: 0.797-0.846)	0.830 (95% CI: 0.811-0.847)
Recall	0.964 (95% CI: 0.959-0.969)	0.685 (95% CI: 0.664-0.706)	0.677 (95% CI: 0.649-0.704)	0.776 (95% CI: 0.757-0.793)
F1-score	0.928 (95% CI: 0.921-0.934)	0.726 (95% CI: 0.705-0.746)	0.743 (95% CI: 0.715-0.769)	0.799 (95% CI: 0.780-0.816)
Mean distance error	0.087 (95% CI: 0.086-0.088)			

Figure 2 Datasets and Algorithm Performance