An Intelligent Care System for Postoperative Wound Status Assessment Following Total Knee Replacement
人工膝關節置換術之術後傷口狀態評估智慧照護系統

張品堯 王麒安 張正璽

國立成功大學 資訊工程學系

詹慧伶 教授｜指導教授

嘉義長庚紀念醫院 骨科部｜合作單位

彭國狄 部主任、陳紀樺 個管師｜合作夥伴

Introduction

Background

- Exceed 60,000 Total Knee Replacement (TKR) are performed annually in Taiwan, requiring postoperative monitoring.
- 0.5%~1% postoperative wound complication are identified after hospital discharge.
- Current follow-up relies on manual review of wound image via messaging application.
- Existing AI wound assessment studies rarely focus on TKR postoperative care.

Objective

- Develop an automated wound status assessment system for early complication after TKR, improving the efficiency of remote follow-up and reducing nurses' workload.

Proposed System

- We proposed a framework for TKR patients based on segmentation and classification.
- The proposed system integrate with LINE and has been deployed.

Method

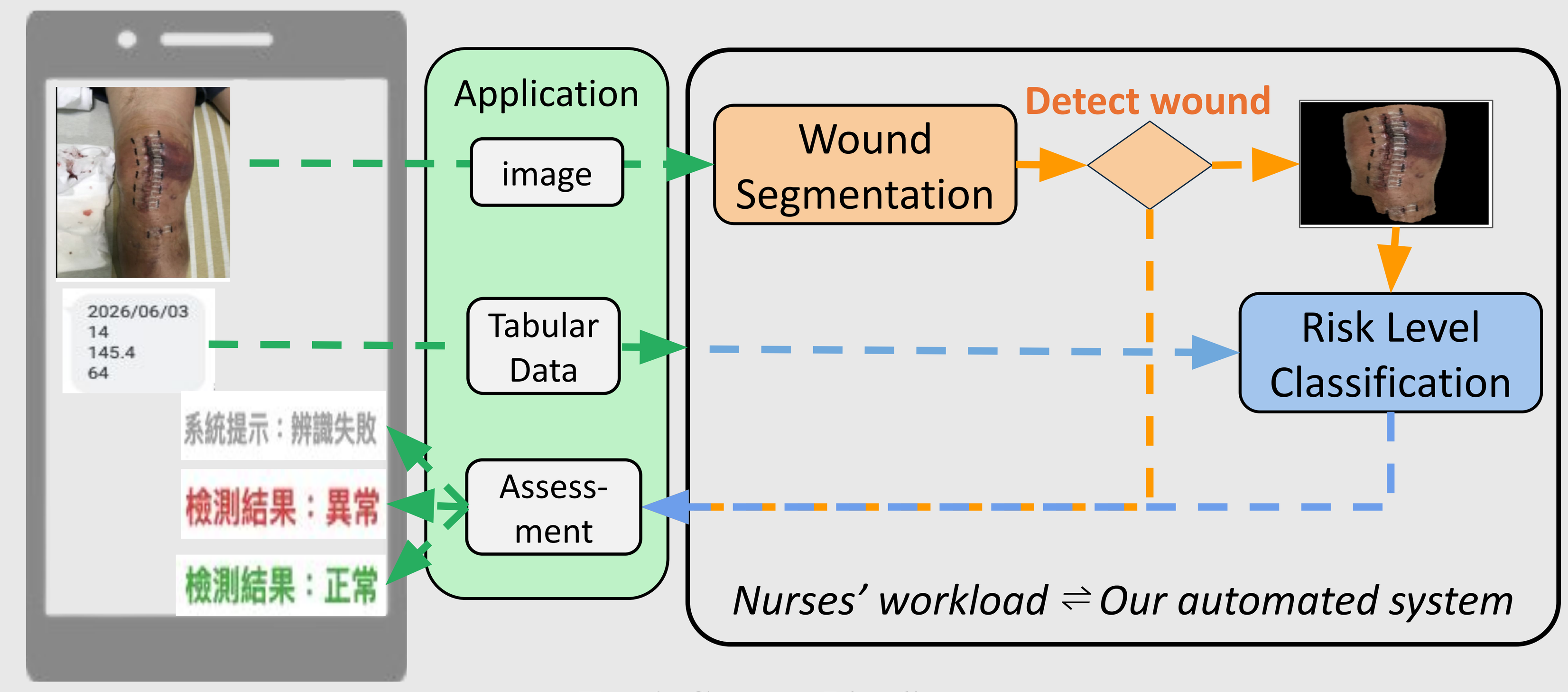


*Fig 1. **System Pipeline***

TKR Dataset Introduction

- Real-world clinical image collected by nurses during routine patient care.
- Professional label by nurses.
- Small dataset size.
- Class imbalance.

*Table I. **TKR Dataset.** The table shows the number and proportion of images.*

Normal	Abnormal	Total
1142	192	1334
85.6%	14.4%	100%

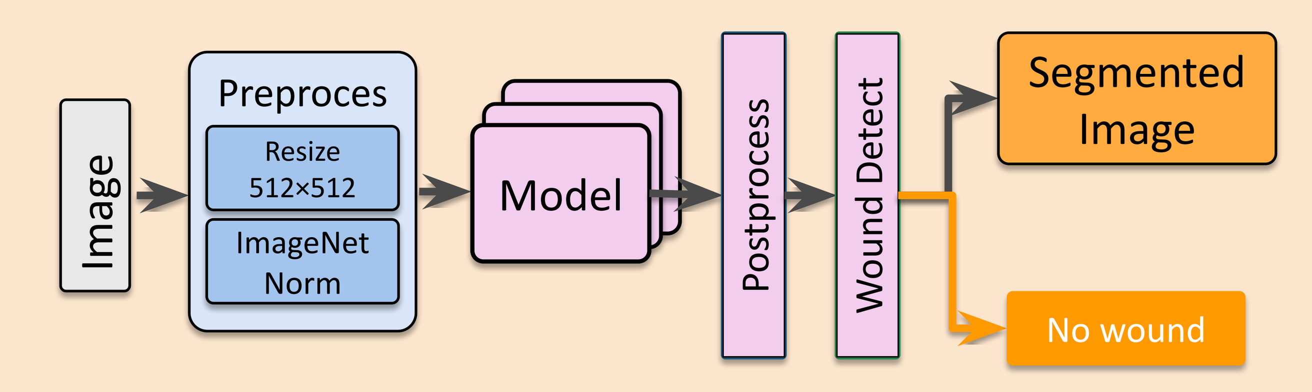


*Fig 2. **Wound Segmentation Inference Pipeline.***

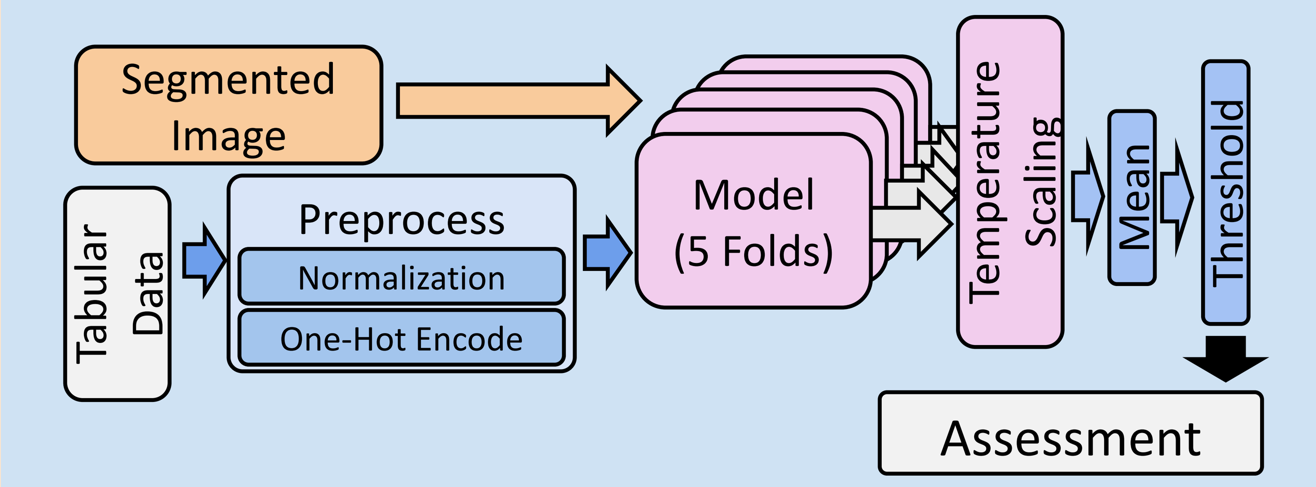


*Fig 3. **Risk Level Classification Inference Pipeline.***

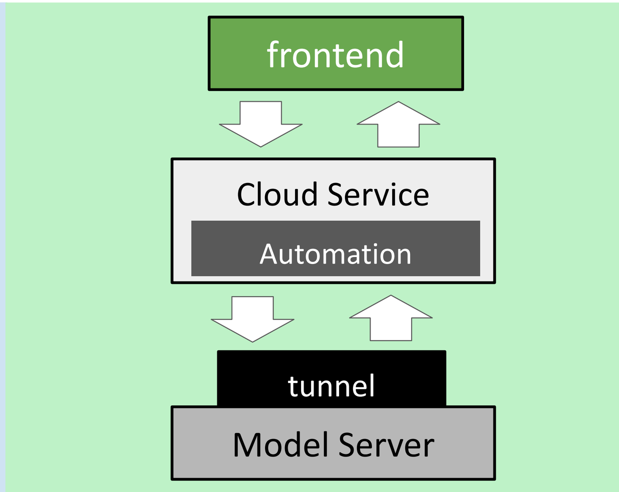


*Fig 4. **Application Pipeline***

Experiments & Results

A. Why Segmentation

*Table II. **From Original Image to ROI Cropping and Precise Segmentation***

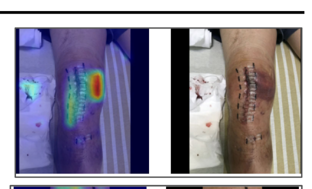
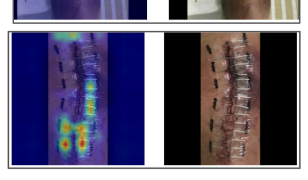
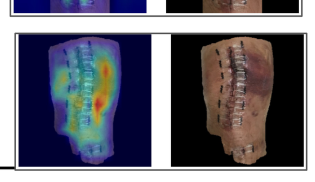
	Background	Surroundings	
Raw Image	✗	✗	
YOLO	✔	✗	
Segme-ntation	✔	✔	

*Table IV. **Inference Results.** We apply ensemble to improve prediction result and get further improvement by calibration and threshold.*

Calib-rated	Thre-shold	Abnormal Recall	Abnormal F_2-score	Macro F_1-score
✗	0.5	82.86	68.08	69.67
✓	**0.5**	**85.71**	**72.82**	**73.93**

B. 3 Metrics: Abnormal Recall, Abnormal F2-score, and Macro F1-score.

- Underestimation carries high clinical risk.
- Abnormal classes show imbalanced distribution.

*Table III. **Training Strategies Ablation Study**, We present key comparisons among different training strategies, and our best training strategy achieve highest overall score.*

Data Augmen-tation	Image Preprocess	Loss Function	Model	Fine-tune	Training Data	Abnormal Recall	Abnormal F_2-score	Macro F_1-score
Basic & Mixup	Segmen-tation	Balanced Cross Entropy	DINOv2 -small	LoRA (Rank 64)	Multi-modal	76.57 ± 5.11	**64.97 ± 4.11**	69.23 ± 7.26
Basic & CutMix	–	–	–	–	–	**77.14 ± 14.98**	60.41 ± 3.90	62.87 ± 7.59
–	YOLO	–	–	–	–	58.86 ± 7.45	54.03 ± 5.38	68.21 ± 3.73
–	–	Cross Entropy	–	–	–	60.57 ± 11.14	56.92 ± 6.41	71.05 ± 3.69
–	–	–	ViT-s/16	–	–	62.86 ± 5.71	59.80 ± 2.81	**73.38 ± 6.89**
–	–	–	–	Full Fine-tune	–	69.71 ± 14.23	51.32 ± 7.00	54.37 ± 4.37
–	–	–	–	–	Image Only	70.29 ± 18.02	59.52 ± 8.79	67.24 ± 11.63

Clinical Testing

- The system has been deployed and is currently under clinical testing at CGMH.

*Table V. **Testing Results.** We collect 36 cases and evaluate with 3 metrics*

Abnormal Recall	Abnormal F_2-score	Macro F_1-score
90.00	77.59	70.24

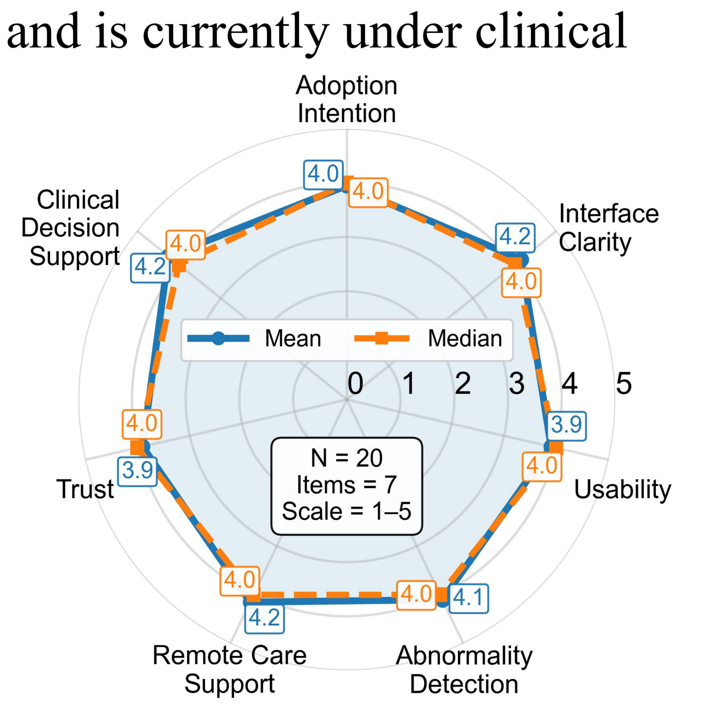


*Fig 6 . **Testing Feedback.** All 20 nurses give positive responses during the system testing.*

Conclusion

- The Proposed segmentation and classification framework for TKR wound image assessment has good performance and is is currently under clinical testing.
- The LINE-based system has been deployed and has received positive feedback from nurses.

Future Work

- Enhance the system with visual explanations, VLM-generated descriptions, confidence scores, and expanded clinical data.
- Integrate the system into hospital platforms through the established collaboration with the SafeSay team at CGMH.
- Extend the segmentation classification framework to broader wound image assessment scenarios.