PROMPT-BASED VERTEBRAL SEGMENTATION USING A GENERATIVE AI APPROACH IN OVCF SPINAL RADIOGRAPHS

Po-Kai Su[1], Pei-Rong Jiang[1], Kai-Xuan Xu[2], Meng-Lei Su[3,4], Jiann-Her Lin[3,4], Hsin-Han Chiang[5], Hsiao-Chi Li[2,*]

[1] Master of Science in Artificial Intelligence Technology, National Taipei University of Technology, Taipei, Taiwan
[2] Department of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan
[3] Department of Neurosurgery, Taipei Medical University Hospital, Taipei, Taiwan
[4] School of Medicine, College of Medicine, Taipei Medical University, Taipei, Taiwan
[5] Department of Vehicle Engineering, National Taipei University of Technology, Taipei, Taiwan

ABSTRACT

(OVCFs) are prevalent among elderly patients, with X-ray imaging serving as the primary diagnostic tool. However, challenges such as organ obstruction and poor contrast after vertebroplasty procedures complicate vertebral segmentation in spinal X-rays.

This research introduces an innovative generative AI framework for vertebral segmentation in spinal X-ray images, combining YOLO-based detection with prompt-driven segmentation inspired by the Segment Anything Model (SAM). The system generates bounding boxes around vertebrae as segmentation prompts and employs an interpolation strategy to address potentially missed compressed vertebrae. By incorporating domain-specific knowledge of vertebral anatomy via the interpolation strategy, the framework enables accurate delineation of vertebral structures in cases of compression fractures.

The model achieves a Dice coefficient of 0.9389 ± 0.0026, an IoU of 0.8854 ± 0.0045, and a sensitivity of 0.9436 ± 0.0062. This generative AI application effectively addresses clinical challenges in vertebral segmentation for OVCF patients, potentially enhancing the accuracy of diagnoses and treatment planning.

BACKGROUND

Most OVCF studies focus on classification tasks rather than segmentation. Precise postoperative vertebrae segmentation remains underexplored despite its importance for treatment monitoring and providing biomechanics and radiomics a firm foundation for further research.

An X-ray image after vertebroplasty (VP) surgery with surgical implants, including cages, screws, and other hardware are hard to read, and the results are easily falsely cohesion and lead results reliable. And existing method relies on handcrafter parameter tuning leads to week generalizability cross dataset.

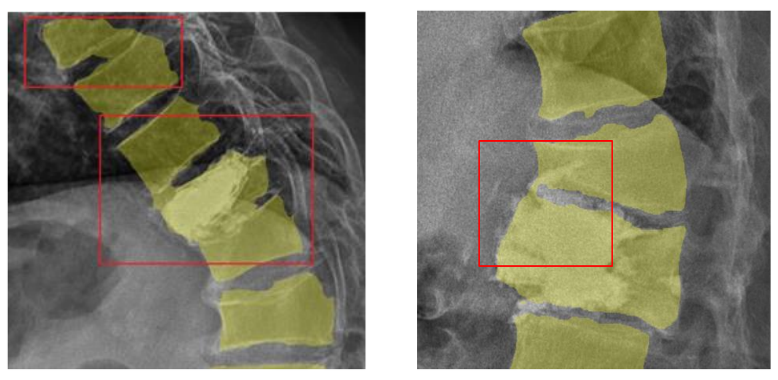


Figure 1. Example X-ray images of improperly fused vertebrae after vertebroplasty surgery

Proposed Architecture

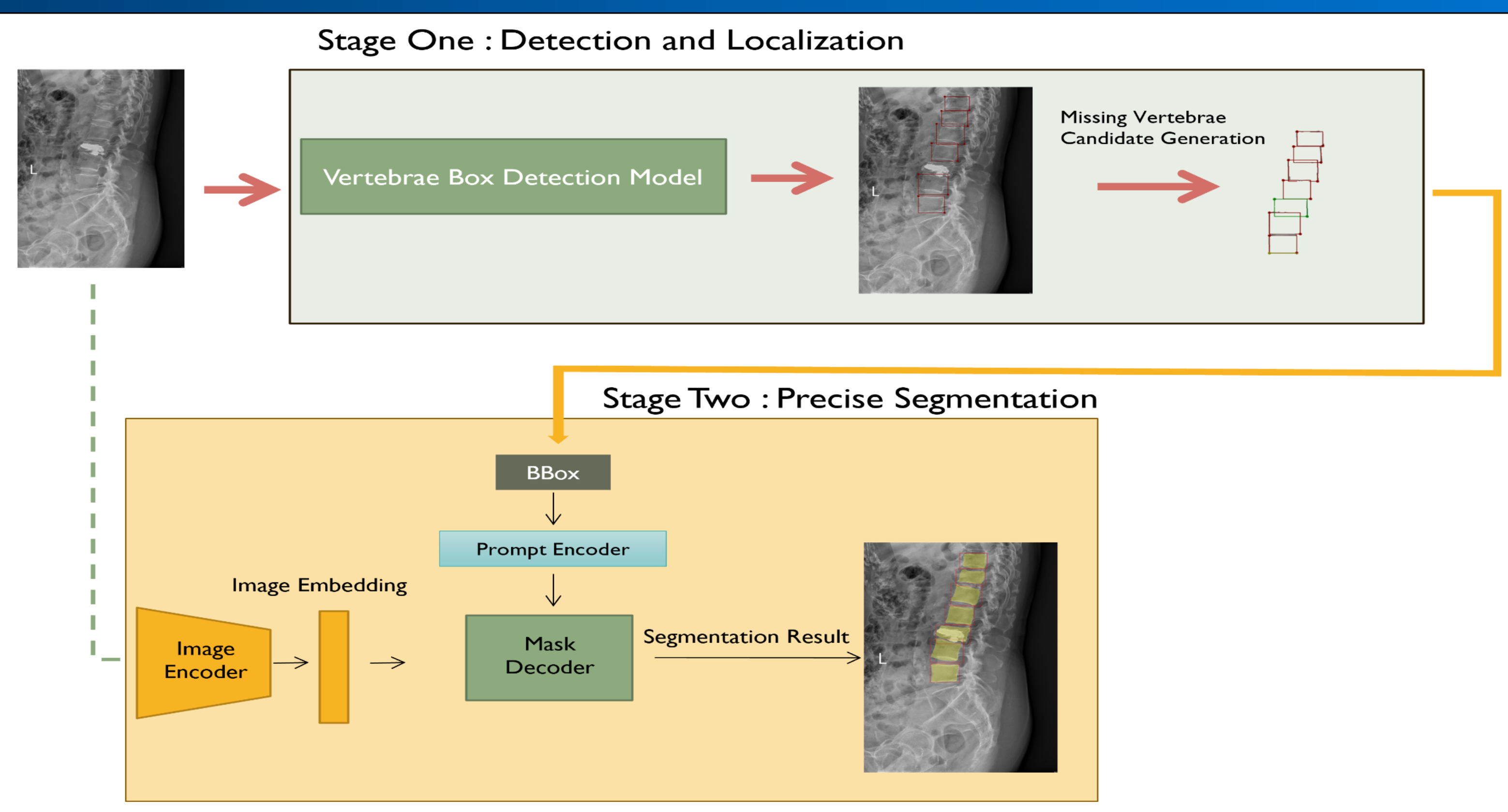


Figure 2. Overview of the proposed iterative vertebrae segmentation pipeline

Building on prior work showing that instance segmentation outperforms semantic segmentation for overlapping vertebrae, our framework (Fig. 2) follows a two-stage pipeline

Stage 1: Detection & Localization

YOLOv8 is used to detect vertebral bounding boxes. This helps isolate individual vertebrae. In difficult cases, some vertebrae may be missed. To address this, we introduce a missing vertebrae candidate generation strategy that leverages the spine's natural curvature and spatial regularity to infer missing detections.

Stage 2: Precise Segmentation

Instead of cropping, we retain the full context of the image by generating embeddings with ViT-B and using the bounding boxes as prompts for the poposed SAM-based model. This approach ensures precise, iterative segmentation at the vertebral level, particularly in cases like OVCF.

To prove a model's capability to accurately delineate individual vertebrae, we propose a *modified Detection Rate* metric. A vertebra (V_i) is considered detected if its IoU with groundtruth exceeds the threshold τ.

$$Detection\ Rate(\tau) = \frac{1}{N}\sum_{i=1}^{N}\begin{cases}1, & if\ IoU\left(V_i^{pred}, V_i^{GT}\right) > \tau \\ 0, & otherwise\end{cases}$$

CLINICAL DATASET

OVCF images are collected in Taipei Medical Hospital from 2014 to 2024

Phase I: Training Set

- Consists of 531 BMP-format radiographic images
- Collects from 164 patients (2014-2019) covering various clinical stages

Phase II: Testing Set

- Consists of 282 DICOM-format radiographic images
- 141 patients (2019-2024) who underwent VP

Morphometric Subset

- A random subset of 30 patients from Phase II
- 236 vertebrae manually annotated by expert clinicians
- Extract for Anterior Body Height (ABH), Middle Body Height (MBH), and Posterior Body Height (PBH)

Table 1. Summary of dataset COMPOSITION

Phase	Images	Annotated Vertebrae	Vertebrae with VP
Training (Phase I)	531	4759	651
Testing (Phase II)	282	2537	342
Morphometric Subset	30	236	35

EXPERIMENTAL RESULTS

We compare our method against traditional one-stage segmentation approaches and other two-stage detect–crop–segment–merge process. Our results show that two-stage models significantly outperform one-stage models across all evaluation metrics. Furthermore, our prompt-based segmentation achieves superior performance on most of these metrics (Table 2).

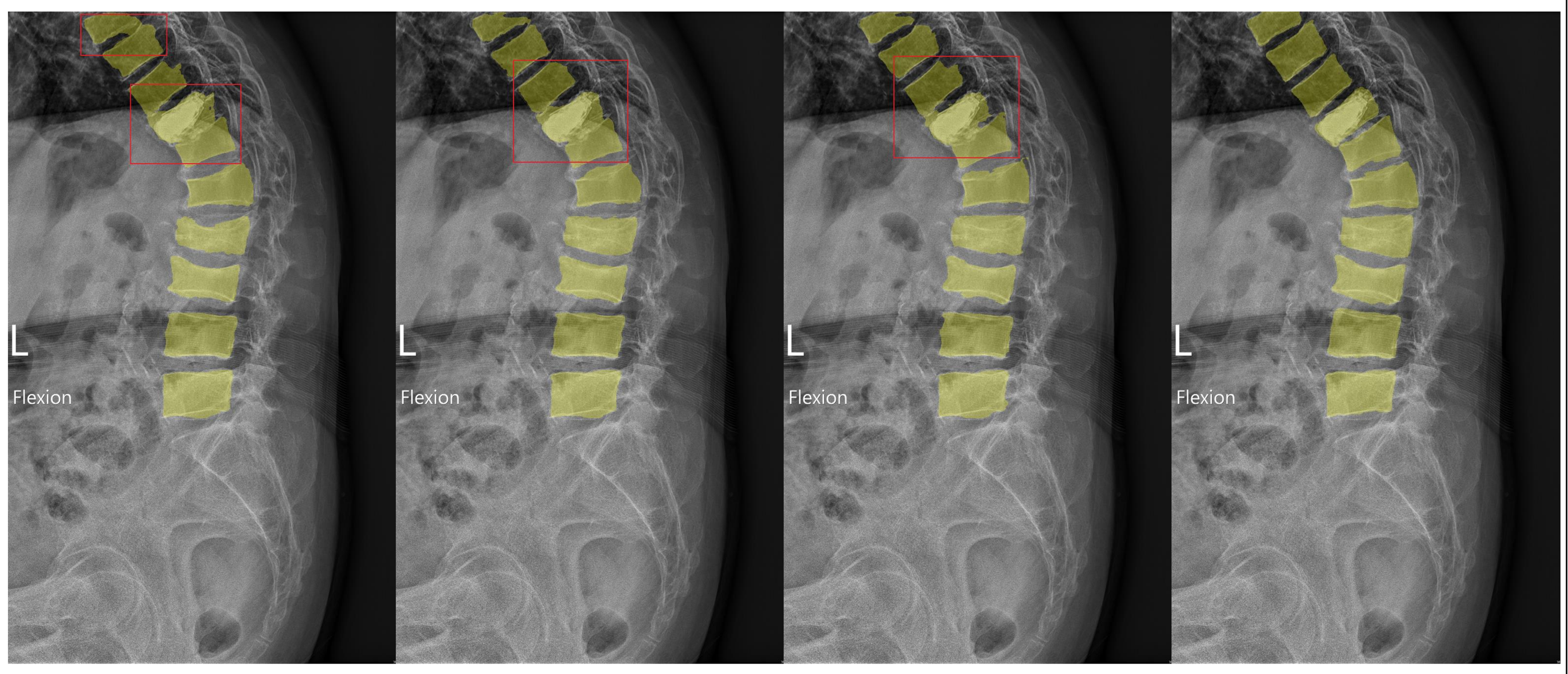
Table 2. Performance comparison among various segmentation models

	Method	Dice(↑)	IoU(↑)	ACC(↑)	SEN(↑)	DTR(↑)
One Stage	U-net	0.9113 ± 0.0027	0.8390 ± 0.0044	0.9860 ± 0.0004	0.9004 ± 0.0090	88.56 ± 0.845
	Res U-net	0.9054 ± 0.0036	0.8286 ± 0.0050	0.9846 ± 0.0007	0.9205 ± 0.0060	84.89 ± 3.13
	DeepLabV3+	0.8969 ± 0.0014	0.8150 ± 0.0022	0.9831 ± 0.0003	0.9195± 0.0047	67.89 ± 5.62
	YOLO v8	0.9025 ± 0.0018	0.8232 ± 0.0030	0.9831 ± 0.0004	**0.9675 ± 0.0022**	86.98 ± 2.55
Two Stages	U-net	0.9222 ± 0.0017	0.8566 ± 0.0029	0.9877 ± 0.0002	0.9082 ± 0.0079	94.53 ± 1.24
	Res U-net	0.9222 ± 0.0019	0.8566 ± 0.0033	0.9877 ± 0.0003	0.9094 ± 0.0055	94.54 ± 0.97
	DeepLabV3+	0.9226 ± 00032	0.8573 ± 0.0055	0.9878 ± 0.0004	0.9046 ± 0.0107	94.66 ± 0.77
	Our Method	**0.9389 ± 0.0026**	**0.8854 ± 0.0045**	**0.9901 ± 0.0004**	0.9436 ± 0.0062	**99.01 ± 0.24**

We further compared our biomechanical measurements with those of a physician expert and achieved the highest agreement (Table 3).

Table 3. Comparison of automatic and expert measurements in Heights

Method	Height	MAE (↓)	ICC(↑)	Pearson R(↑)	Pearson P(↑)
U-Net	ABH	2.88 ± 0.62	0.579 ± 0.093	0.665 ± 0.080	<0.0001
	MBH	2.53 ± 0.37	0.588 ± 0.063	0.646 ± 0.038	<0.0001
	PBH	4.09 ± 0.60	0.412 ± 0.085	0.638 ± 0.082	<0.0001
Res U-Net	ABH	2.63 ± 0.24	0.587 ± 0.013	0.670 ± 0.012	<0.0001
	MBH	2.61 ± 0.15	0.562 ± 0.058	0.638 ± 0.038	<0.0001
	PBH	4.12 ± 0.19	0.396 ± 0.023	0.629 ± 0.017	<0.0001
DeepLabV3+	ABH	2.53 ± 0.04	0.630 ± 0.034	0.694 ± 0.033	<0.0001
	MBH	2.53 ± 0.14	0.531 ± 0.039	0.579 ± 0.035	<0.0001
	PBH	3.94 ± 0.05	0.422 ± 0.025	0.628 ± 0.035	<0.0001
Our Method	ABH	**1.42 ± 0.15**	**0.896 ± 0.029**	**0.909 ± 0.025**	<0.0001
	MBH	**1.85 ± 0.08**	**0.798 ± 0.013**	**0.811 ± 0.016**	<0.0001
	PBH	**2.34 ± 0.28**	**0.745 ± 0.042**	**0.862 ± 0.017**	<0.0001



U-Net DeepLabV3+ Res U-Net Our Method

Figure 4. Visual comparison of two-stage segmentation results for vertebral segmentation

CONCLUSIONS

We propose a novel generative AI framework for vertebral segmentation in OVCF X-ray images, addressing key challenges in clinical spinal imaging. The method combines YOLO-based detection with prompt-driven segmentation inspired by the Segment Anything Model. It enables accurate segmentation even in challenging post-vertebroplasty cases where traditional methods often fail, providing a pathway toward robust and generalizable solutions for OVCF management.

ACKNOWLEDGEMENT

This work was supported by the National Science and Technology Council, Taiwan, under grant No. NSTC 113-2221-E-027-094.

TAIPEI TECH 國立臺北科技大學 National Taipei University of Technology HIBIO

hcli@mail.ntut.edu.tw