



STS 2026 CHALLENGE

Segmentation-Guided 3D Dental ROI Transformer for Structured Clinical Report Generation from CBCT

Team: SJTU_SEE_2426Lab

Presenter: Changkai Ji

Changkai Ji¹, Yusheng Liu¹, Yuxian Jiang¹, and Lisheng Wang^{1*}

Code: <https://github.com/duola-wa/MICCAI-2026-STS-Task-3>

¹School of Automation and Intelligent Sensing, Shanghai Jiao Tong University, Shanghai, China.



杭州电子科技大学
HANGZHOU DIANZI UNIVERSITY



上海交通大学
Shanghai Jiao Tong University



杭州钱塘牙科医院
HANGZHOU QIANTANG DENTAL HOSPITAL



杭州牙科医院
HANGZHOU DENTAL HOSPITAL



浙江大学医学院附属
邵逸夫醫院
SIR RUN RUN SHAW HOSPITAL
MEDICAL SCHOOL ZHEJIANG UNIVERSITY



DataCanvas
九章云极

Background

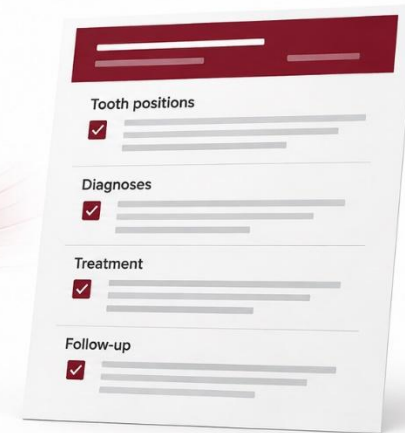


- Dental diagnostic reports translate complex CBCT findings into clinically meaningful information, supporting treatment planning, clinical communication, and follow-up.
- Preparing these reports requires clinicians to review numerous slices in a CBCT volume, which is a time-consuming and expertise-intensive process.
- Generating structured reports from CBCT connects volumetric image analysis with organized clinical information, extending dental AI beyond anatomical segmentation.



Dental CBCT

Conceptual illustration



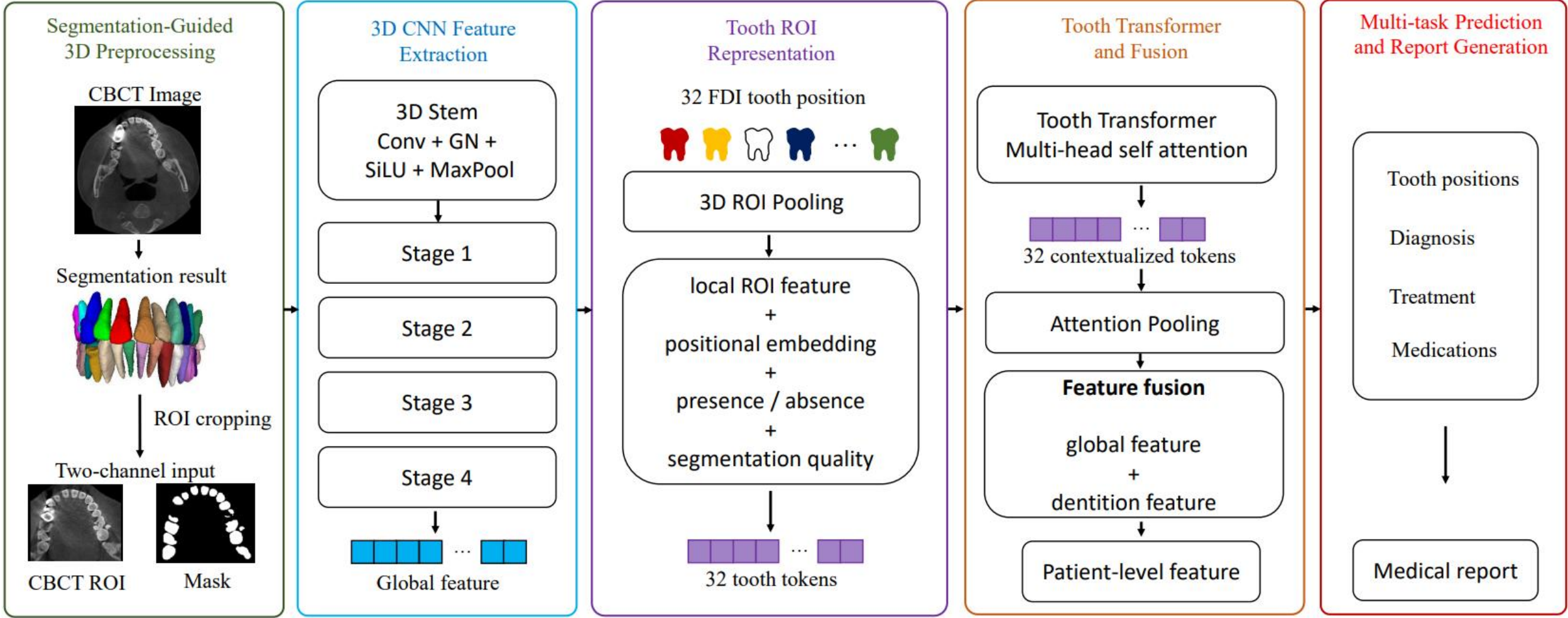
Structured clinical report

Motivation

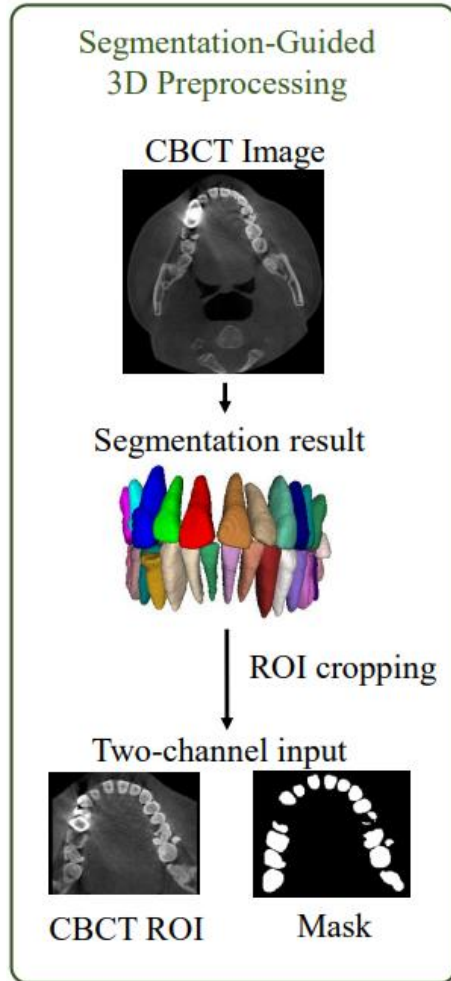


- Preserving 3D anatomical context: Clinically relevant findings may involve tooth roots, periapical regions, and surrounding bone. The model therefore needs to retain spatial continuity and information beyond tooth boundaries.
- Linking findings to specific teeth: Global image features alone may not sufficiently capture localized abnormalities. Combining global context with FDI-aligned tooth-level features can help associate clinical findings with the corresponding tooth positions.
- Learning from limited and imbalanced clinical data: Limited number of labeled CBCT cases are available, and diagnoses and treatments occur at different frequencies. Learning clinically relevant patterns under these conditions is challenging.

Methodology

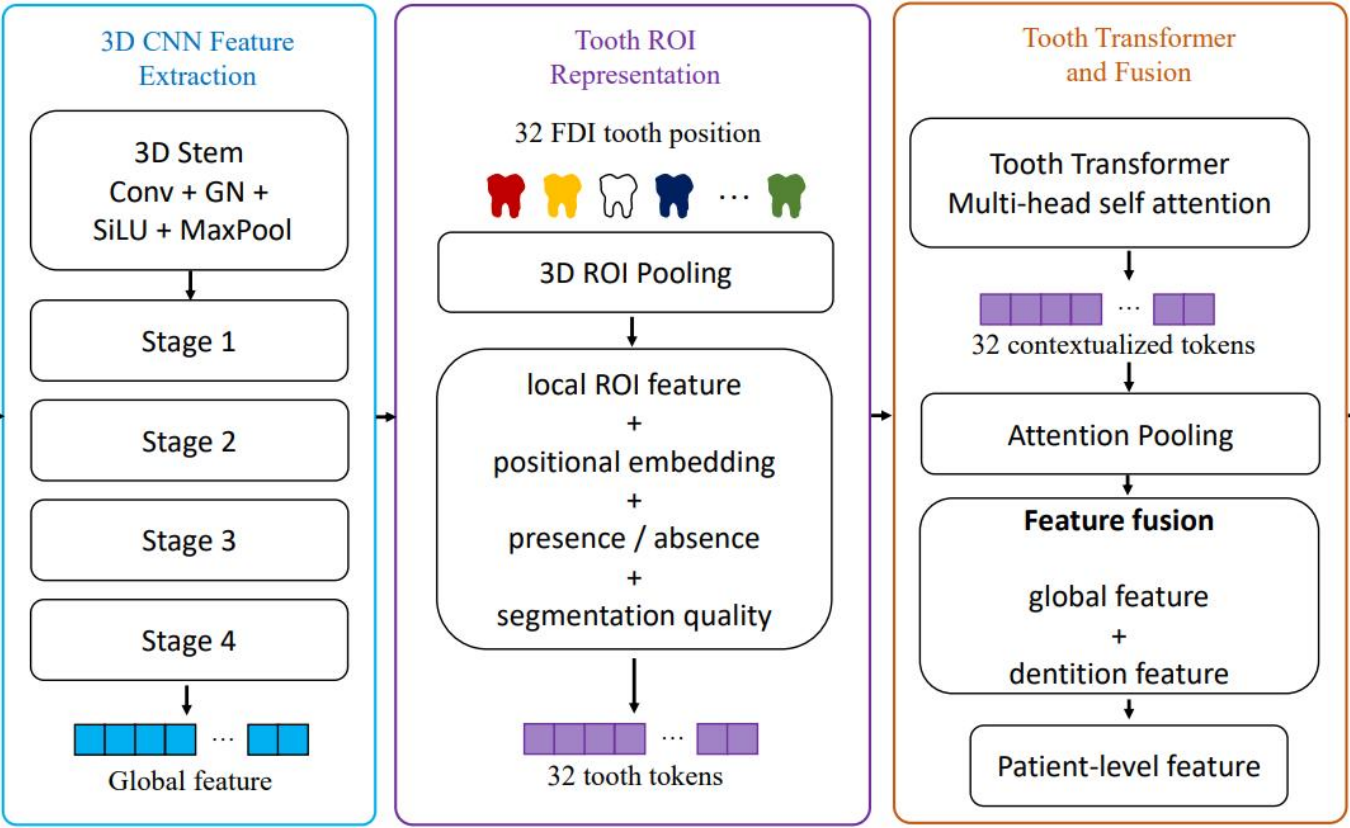


Overview of the proposed framework. Global CBCT features and FDI-aligned tooth-level features are jointly modeled for structured clinical prediction and report generation.

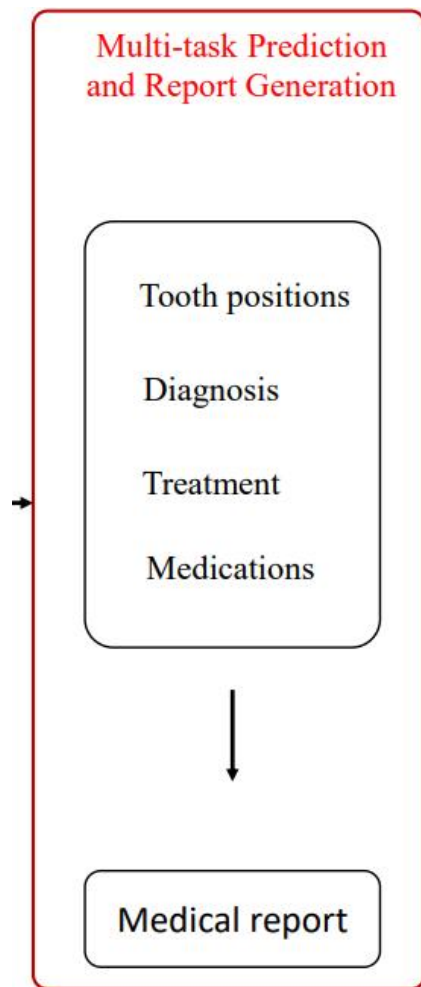


- Anatomical localization: Individual teeth were first segmented using an open-source, pretrained nnU-Net model [1], and the segmentation results were then mapped to the 32 tooth FDI positions to guide ROI construction.
- Context-preserving cropping: Expanded 3D regions retain the teeth and surrounding tissues, including periapical regions, periodontal structures, and alveolar bone.
- Two-channel input: The cropped CBCT and tooth mask are jointly provided to the network, preserving image intensities while adding anatomical guidance.

[1] Van Nistelrooij, Niels, et al. "ToothSeg: Robust tooth instance segmentation and numbering in CBCT using deep learning and self-correction." *IEEE Journal of Biomedical and Health Informatics* (2026).



- Global and local features: A shared 3D CNN extracts a global representation from deep features and tooth-level ROI features from an intermediate feature map.
- FDI-aligned tooth tokens: Each token combines local visual features, FDI positional information, tooth presence.
- Relationship modeling and fusion: A Tooth Transformer models interactions among tooth slots, followed by attention pooling and fusion with the global representation.



- Multi-task clinical prediction: Separate heads predict tooth positions, diagnoses, tooth–diagnosis associations, treatment actions, medications, and auxiliary clinical information.
- Class-balanced optimization: A weighted combination of class-balanced binary cross-entropy losses addresses label imbalance across prediction tasks.
- Structured report rendering: Predicted entities are mapped to predefined clinical phrases and organized into the seven required report fields.

Experiments: Training and Inference



Item	Setting
Supervised training	50 labeled CBCT cases
Internal evaluation	5-fold patient-level cross-validation
Optimization	AdamW and early stopping
Inference	Five-model logit averaging
Hardware	One NVIDIA RTX 4090

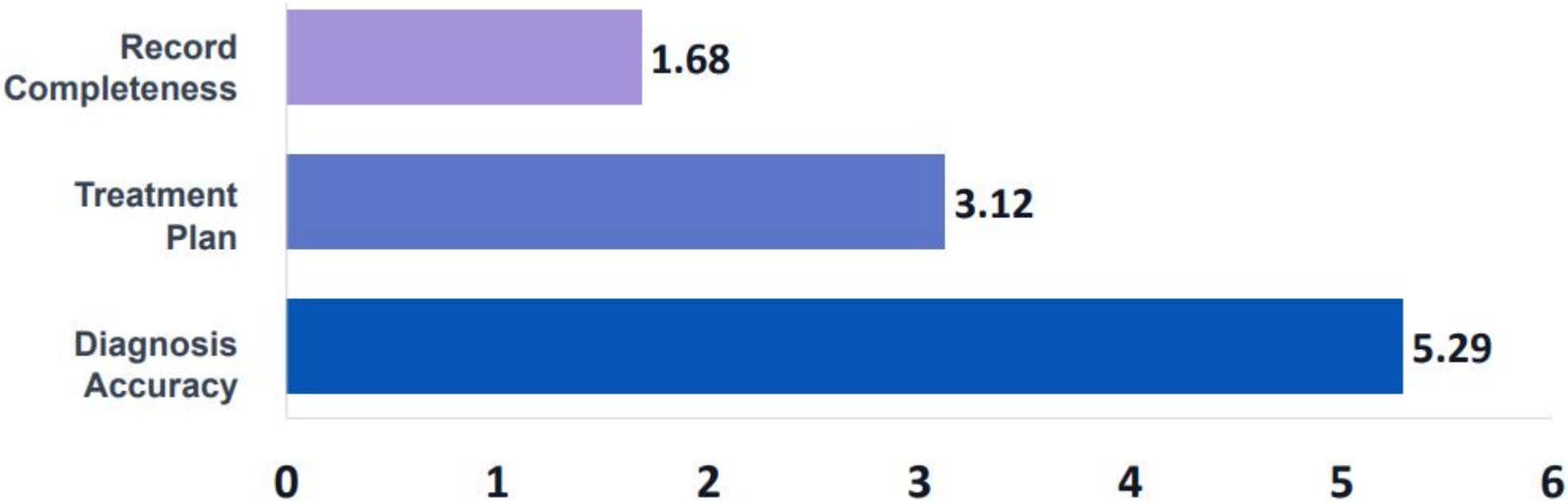
- Task-specific thresholds were selected using out-of-fold predictions from the training set.

Experiments and Results



Table 2. Validation performance on MICCAI STSR 2026 Challenge Task 3.

Diagnosis Accuracy	Treatment Plan	Record Completeness	Total Score
5.29	3.12	1.68	10.08



Visualization

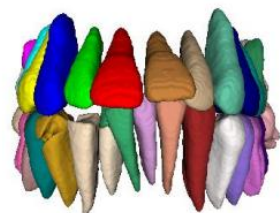
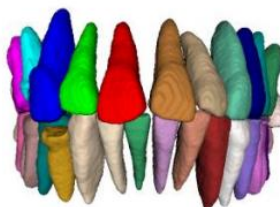


MICCAI2026

CBCT Image



Segmentation Result



Generated Report

```
"Main appeal": "Food trapping in the upper right posterior region.",  
"Present medical history": "Intermittent food trapping for several weeks, without  
"Oral Check": "Clinical examination identifies a cavitated lesion on tooth 16. Im  
"Diagnosis": "Dental caries of tooth 16.",  
"Treatment plan": "Remove carious tissue and restore tooth 16.",  
"Handle": "Findings and treatment options discussed; restorative treatment schedu  
"Doctor advices": "Maintain interdental cleaning and return for the scheduled rest
```

```
"Main appeal": "Request for review of a previously restored lower right premolar."  
"Present medical history": "Previous restorative treatment; no current spontaneous  
"Oral Check": "An existing restoration is noted on tooth 44. Its margins and the s  
"Diagnosis": "Previously restored tooth 44 requiring review.",  
"Treatment plan": "Assess restoration integrity and determine whether repair or re  
"Handle": "Initial assessment completed; no operative treatment performed.",  
"Doctor advices": "Continue routine oral hygiene and return if pain or restoration
```

```
"Main appeal": "Discomfort when chewing on an upper left posterior tooth.",  
"Present medical history": "Intermittent biting discomfort for several weeks.",  
"Oral Check": "A periapical radiolucency is described at tooth 26 in this case. Pu  
"Diagnosis": "Suspected apical periodontitis of tooth 26.",  
"Treatment plan": "Confirm the diagnosis and assess the need for root canal treatm  
"Handle": "Diagnostic findings discussed; endodontic assessment arranged.",  
"Doctor advices": "Avoid heavy chewing on the affected tooth and attend the planne
```

Acknowledgements



- Thanks to the organizers for hosting the competition and providing valuable data for this research.
- Appreciation for the organizers' support throughout the competition, especially for their patience in addressing various questions and issues.



Q & A

changkaiji@sjtu.edu.cn