



Faculty of Dentistry, Oral & Craniofacial Sciences, King's College London

EviDent-CBCT

Evidence-Bottlenecked Report Generation from Dental CBCT

Ruiyang Hao

Project supervised by Dr. Yunpeng Li and Prof. Owen Addison



ODIN 2026

ToothFairy4 Challenge

27 September 2026

CBCT interpretation – Clinically Important and Challenging

Diagnosis and surgical planning depend on a careful reading of the full 3D scan.

Specialist training

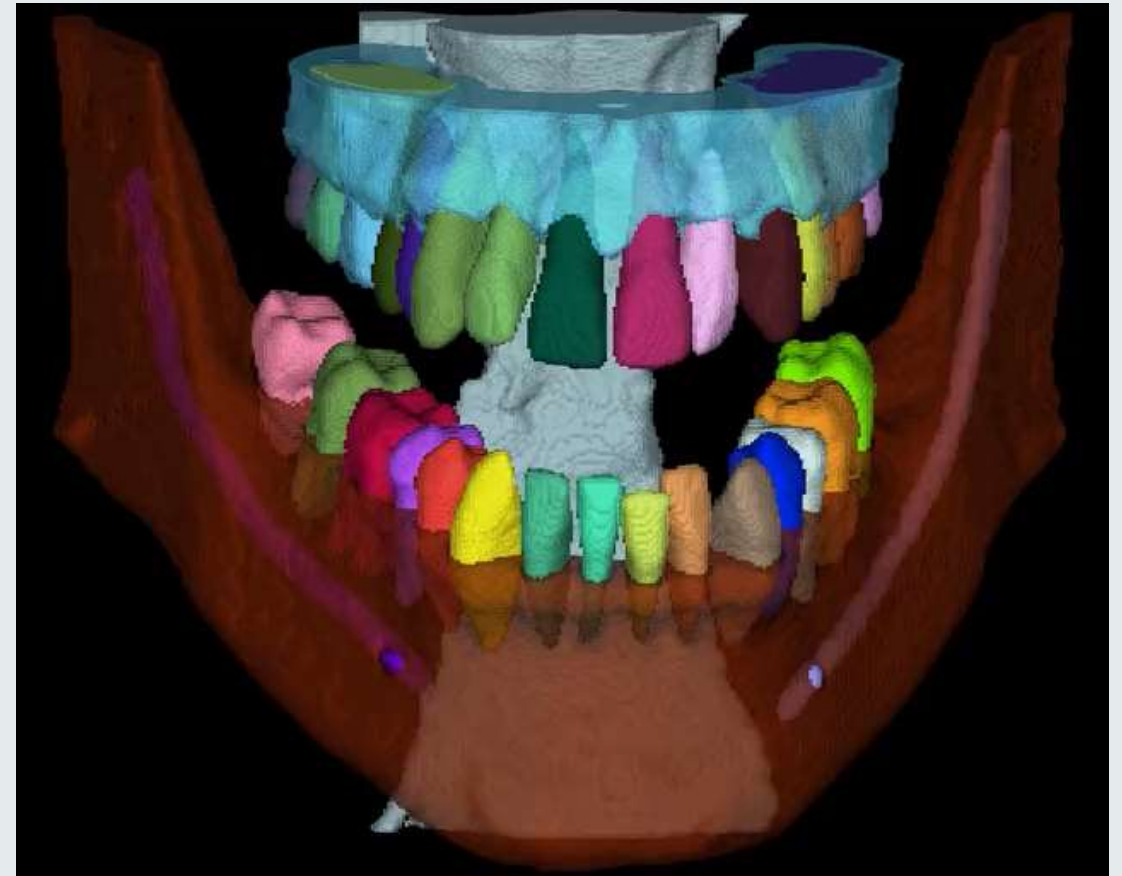
3D anatomy, pathology and radiological judgement

Substantial workload

Systematic volume review and report writing

Variation in practice

Clinical experience, report detail and descriptive terminology



Public CBCT datasets make these tasks possible

Comparing with medical CT/MRI 3D data, high-quality annotations and reports remain scarce in Dental CBCT

Thank you to the ToothFairy organisers, clinicians and annotators

2023

ToothFairy

Inferior alveolar
canal segmentation

2024

ToothFairy2

Broader anatomical
segmentation

2025

ToothFairy3

Expanded anatomy
and efficient
segmentation

2026

ToothFairy4

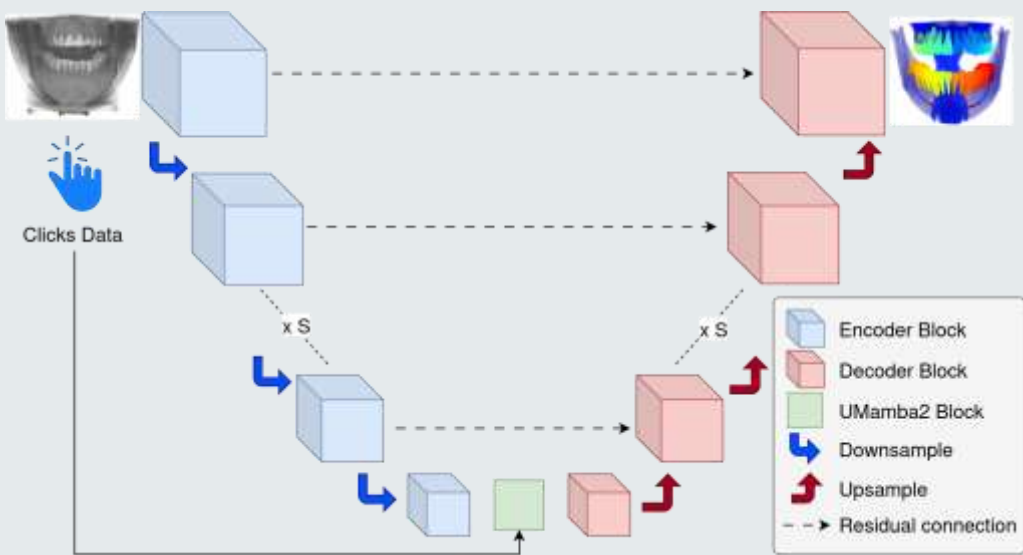
CBCT report
generation

From anatomical segmentation to clinical interpretation

Our second year in the ToothFairy community

2025

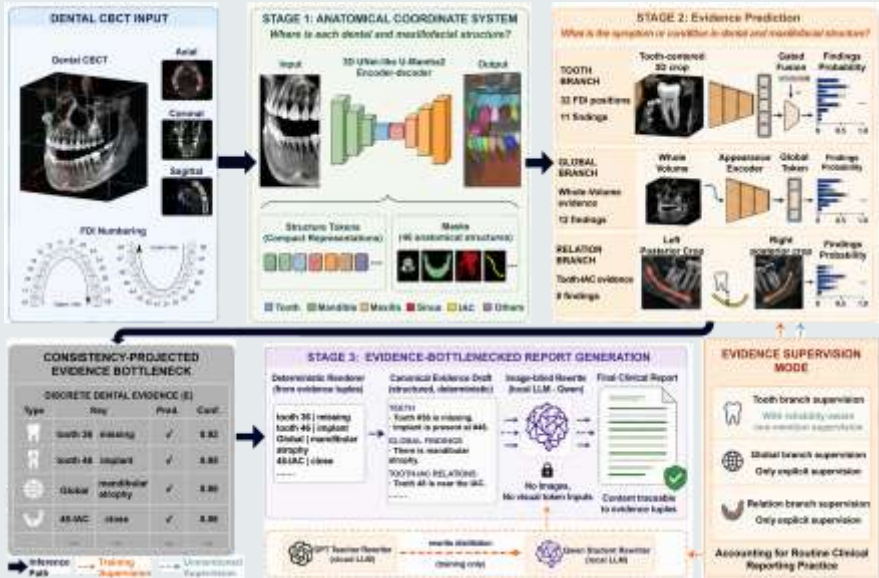
ToothFairy3: Anatomy Segmentation



U-Mamba2^[1]

2026

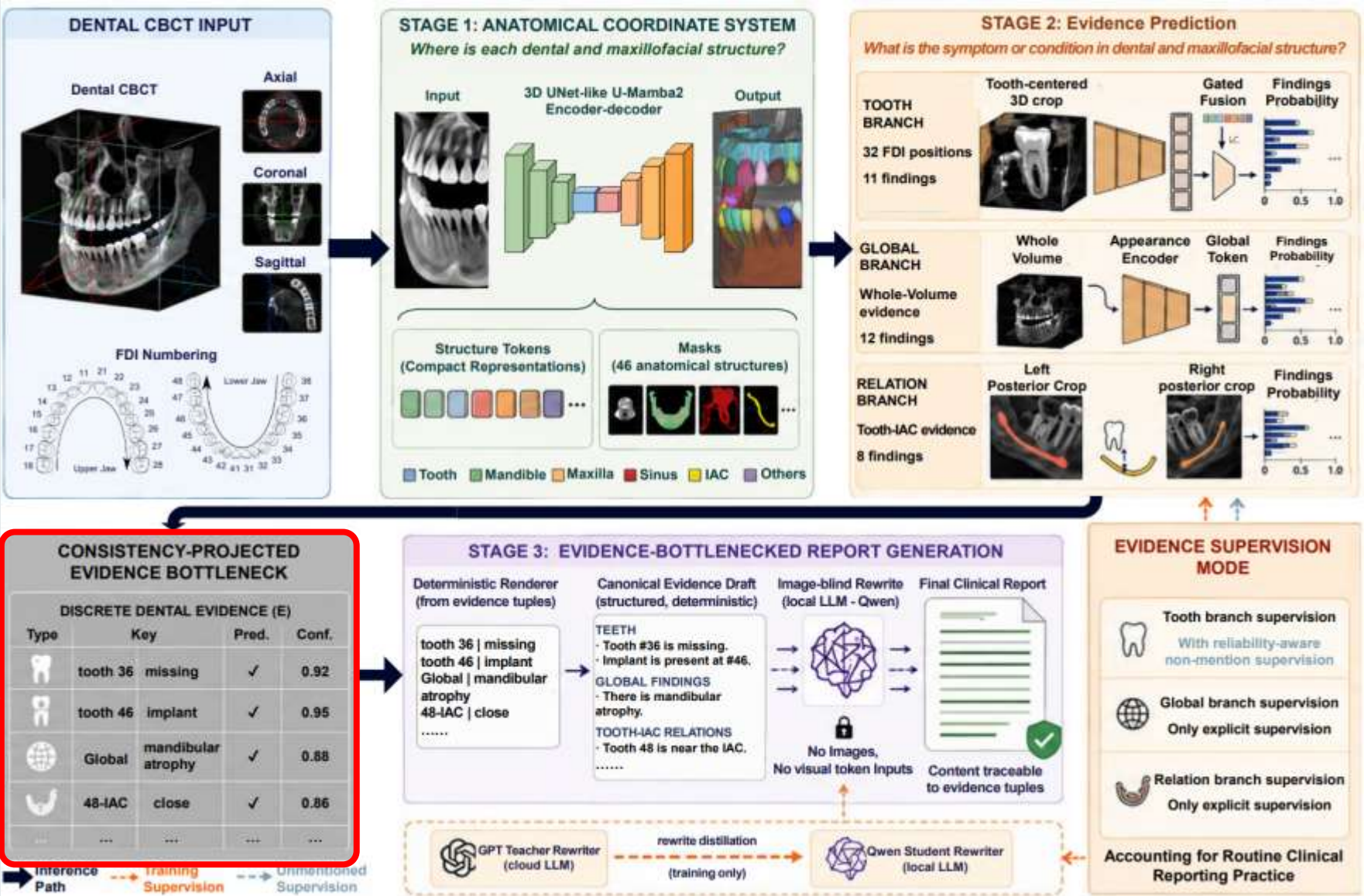
ToothFairy4: Report Generation



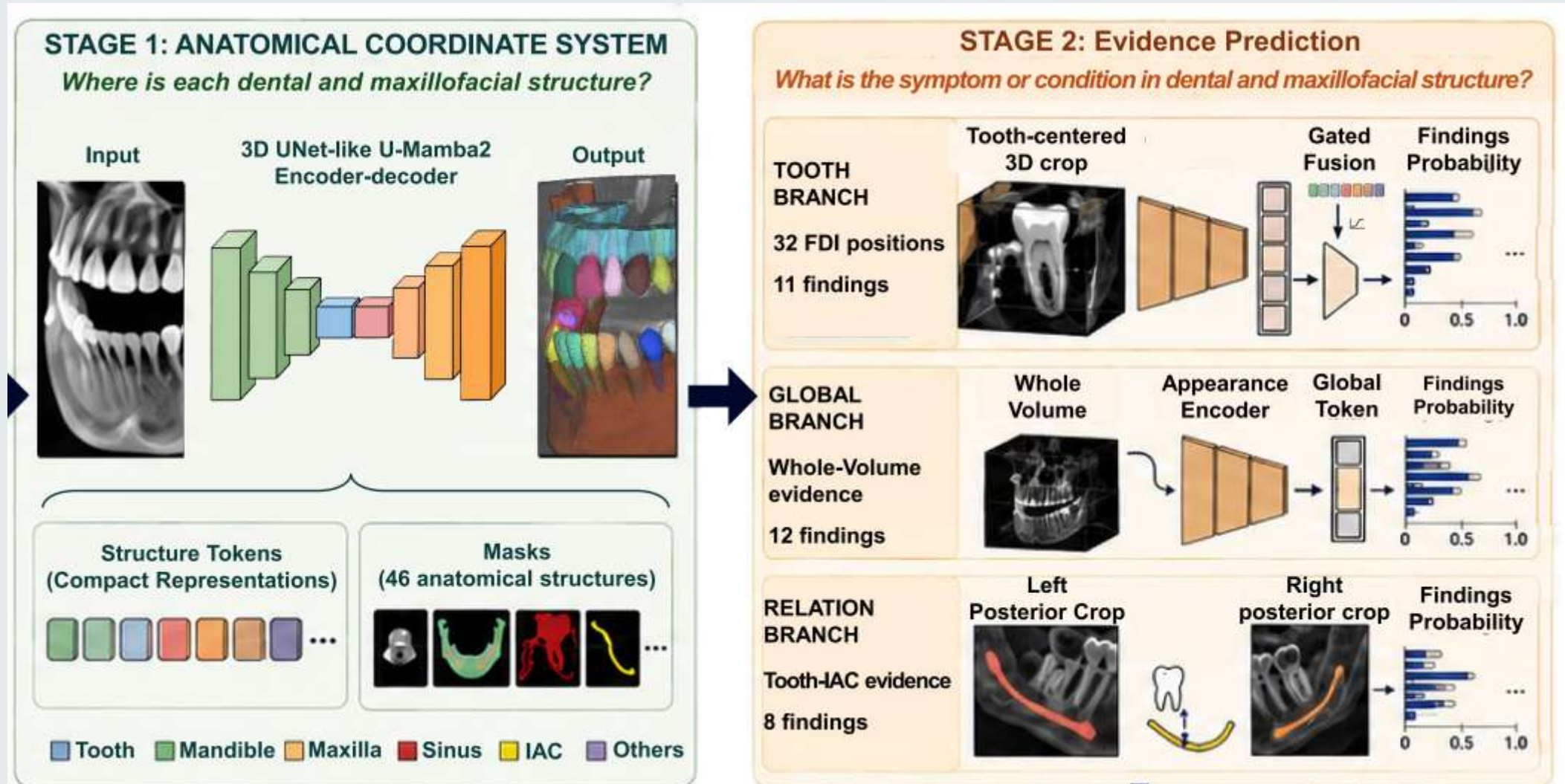
EviDent-CBCT

EviDent-CBCT makes the dental evidence explicit

Can a
structured
evidence record
improve CBCT
report
generation?



Anatomy aligns evidence with the right structures



Label Extraction for Finding Prediction

From recurring report findings to structured training labels

Select recurring findings

Review report coverage.
Merge similar expressions.



Compact finding schema

11

Tooth findings

12

Global findings

8

Tooth–canal relations



Clinical report

Tooth 36 has a crown.

Tooth 48 is erupted.

Other findings are not documented.



GPT-assisted extraction

Map each report to the finding schema.

Findings and tooth / region
Explicit positives and negatives



Structured finding labels

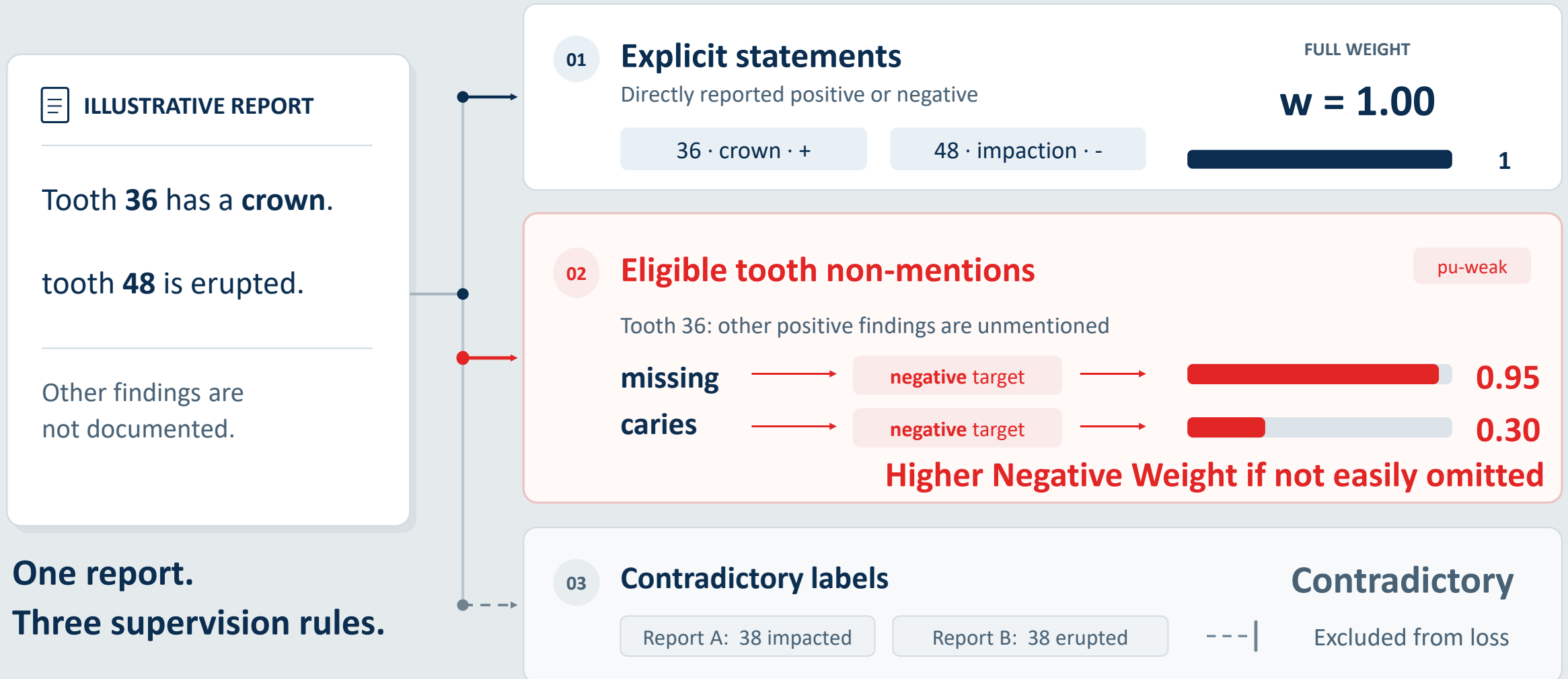
Tooth	Finding	Status
36	crown	Positive
48	impaction	Negative
36	caries	Unmentioned

Illustrative tooth-level mapping

Unmentioned ≠ confirmed negative

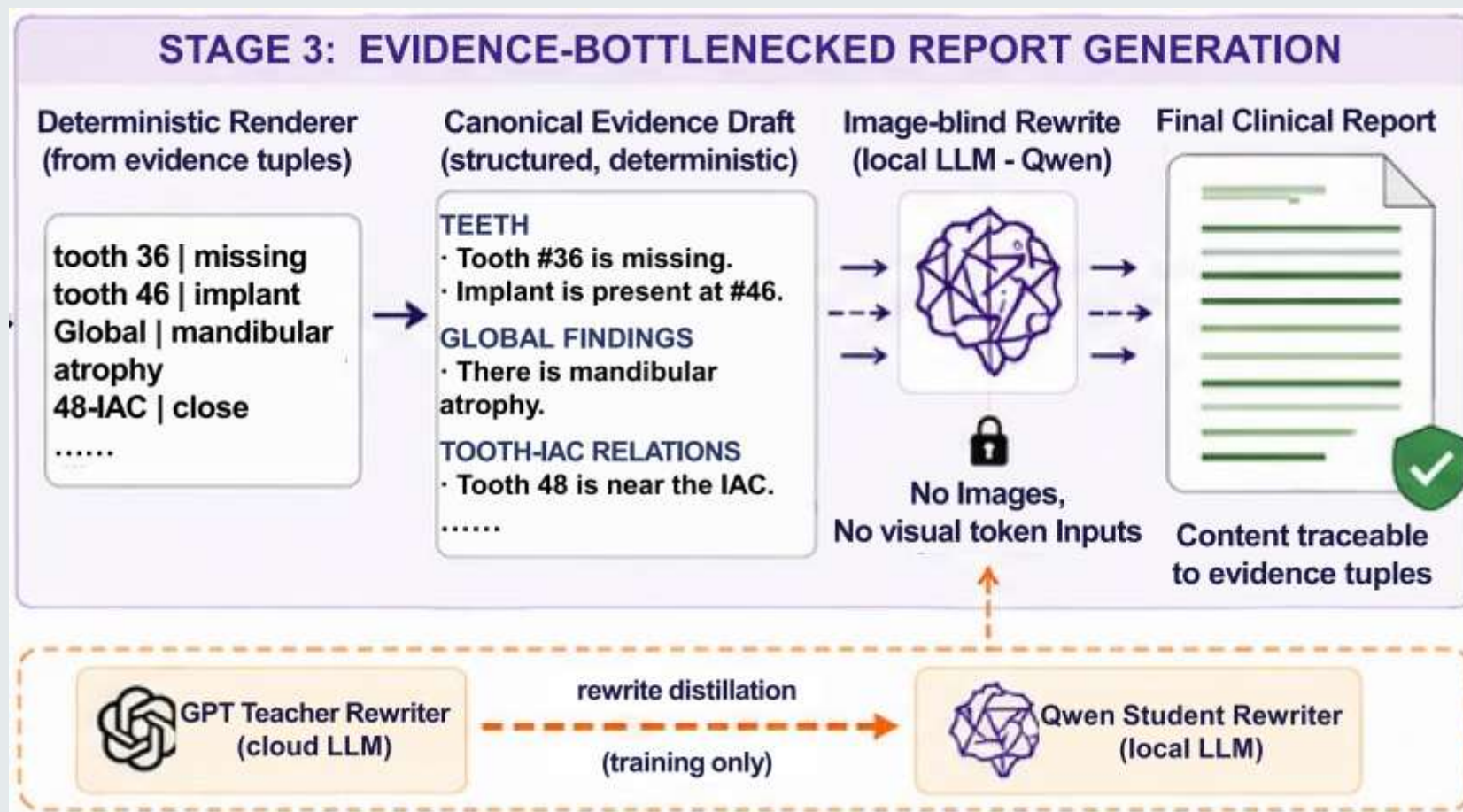
Next: reliability-aware supervision

Not mentioned ≠ confirmed normal: Reliability-aware Supervision from Reports



This label assign design is inspired by positive unlabeled (PU) learning^[1], We call this supervision as **PU-Weak**

The local rewriter receives only the evidence draft



Before writing

Resolve missing-tooth and canal-proximity conflicts

Choose between conflicting canal courses by confidence

This module adds no positive finding

Method Comparison: Whether evidence bottleneck actually help?

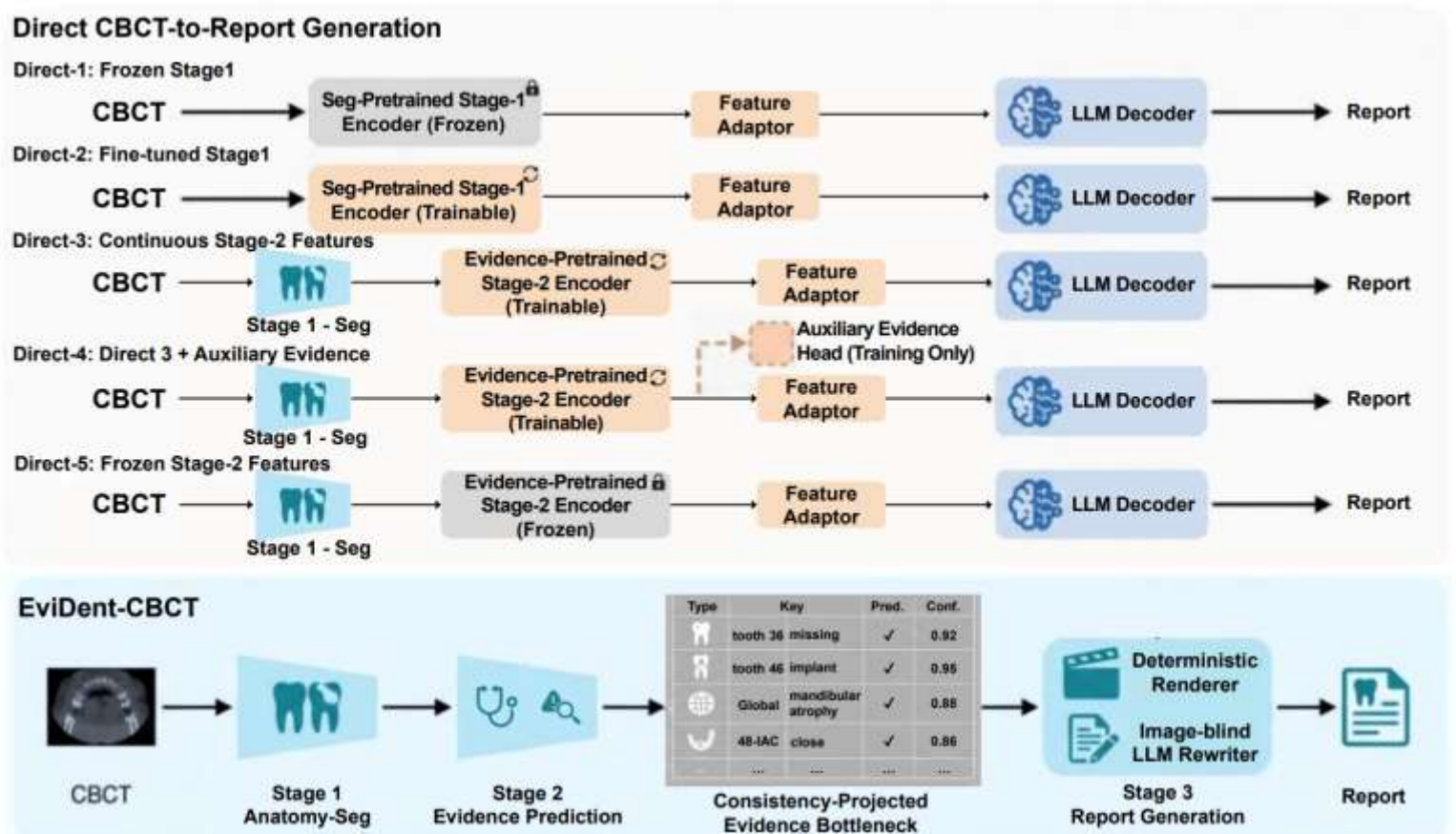
624 scans

1,001 English reports

561 training scans
63 validation scans

Controlled direct generation

Five visual interfaces; Qwen3-8B and Med42-8B; Raw and canonical targets



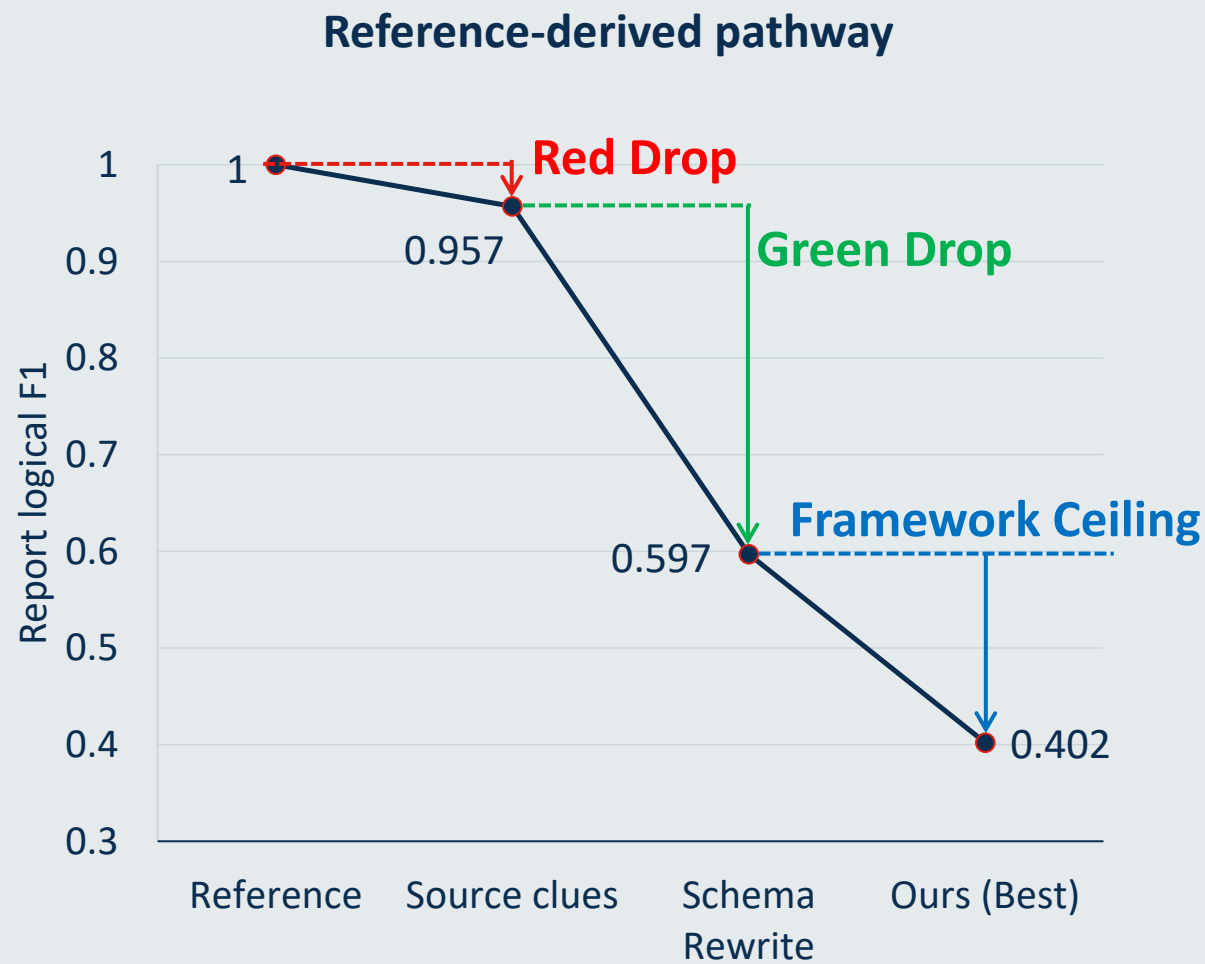
The evidence interface improves report logical F1

Ours 0.402 vs. Best Direct 0.371

Family	Visual/report interface	Report target	Logical P	Logical R	Logical F1
Direct Med42-8B	Direct 1: frozen Stage 1 tokens	Raw	.227 $\pm$.016	.264 $\pm$.015	.243 $\pm$.005
	Direct 1: frozen Stage 1 tokens	Canonical	.356 $\pm$.020	.334 $\pm$.014	.344 $\pm$.012
	Direct 2: fine-tuned Stage 1 encoder	Raw	.266 $\pm$.031	.293 $\pm$.016	.278 $\pm$.024
	Direct 2: fine-tuned Stage 1 encoder	Canonical	.388 $\pm$.019	.354 $\pm$.017	.371 $\pm$.018
	Direct 3: continuous Stage 2 features	Raw	.291 $\pm$.003	.272 $\pm$.036	.280 $\pm$.018
	Direct 3: continuous Stage 2 features	Canonical	.366 $\pm$.005	.359 $\pm$.011	.362 $\pm$.006
	Direct 4: Direct 3 + auxiliary evidence	Raw	.273 $\pm$.020	.272 $\pm$.019	.271 $\pm$.002
	Direct 4: Direct 3 + auxiliary evidence	Canonical	.374 $\pm$.008	.356 $\pm$.010	.365 $\pm$.007
	Direct 5: frozen Stage 2 encoder	Raw	.252 $\pm$.004	.274 $\pm$.020	.262 $\pm$.010
	Direct 5: frozen Stage 2 encoder	Canonical	.380 $\pm$.024	.342 $\pm$.006	.360 $\pm$.013
EviDent-CBCT	Consistency-projected evidence record	Deterministic template	.372 $\pm$.002	.428 $\pm$.008	.398 $\pm$.004
	Consistency-projected evidence record	Local distilled rewrite	.384 $\pm$.002	.423 $\pm$.004	.402 $\pm$.003

Notes: Logical metrics in this page are based on our improved RadFact-Lite-Dental module, slightly different from official metrics. Details in paper.

The compact schema accounts for the largest information loss



Red Drop – Rare Content Drop: A small amount of rare report content is not retained during evidence extraction.

Green Drop – Schema Description-Omitting Drop: Our Evidence Schema represents recurrent categorical findings, locations, and relations, but only limited descriptive qualifier.

Promising direction – how to robustly output the detailed information in CBCT report?

Notes: Logical metrics in this page are based on our improved RadFact-Lite-Dental module, slightly different from official metrics. Details in paper.

Weak negatives improve tooth-evidence F1

Ablation on How we treat non-mentioned labels:

Eligible tooth non-mentions

pu-weak

Tooth 36: other positive findings are unmentioned

missing

negative target

0.95

caries

negative target

0.30

Hard negatives: Non-mention as Full-weight negative

Explicit only: Non-mention labels are excluded in supervision

Base (pu-weak)

0.480 ± 0.007

Hard negatives

0.467 ± 0.017

Explicit only

0.359 ± 0.018

0.3

0.35

0.4

0.45

0.5

0.55

Case Illustration: complete original comparison

Green: matched Light green: partial Blue: unsupported / contradictory Red: unrecovered

EviDent-CBCT Prediction

The CT scan encompasses the mandible, excluding the condylar and coronoid processes bilaterally.

In the mandibular region, the right and left mandibular canals exhibit a regular course, predominantly lingual. Notably, the mandibular canal is in close proximity to teeth 38 and 48. There is evidence of mandibular atrophy in the alveolar processes of the edentulous areas. Teeth 38 and 48 are impacted. Periodontal disease is present in the existing teeth, and areas of osteolysis are observed.

The maxilla is not included in the scan volume.

The mandibular condyles are excluded from the acquisition and are not visible.

Reference ToothFairy4 Report

Mandible: tooth 3.8 and 4.8 are in complete bony impaction. The mandibular canal has a regular course and is in close relationship with tooth 3.8 and 4.8. The maxilla is not included in the acquisition and is not assessable. The mandibular condyles are not included in the scan.

Best Direct Model Prediction

The CT scan encompasses the mandible, excluding the condylar and coronoid processes bilaterally.

In the mandibular region, the right and left mandibular canals exhibit a regular course. Notably, the mandibulargia is in close proximity to tooth 38. Tooth 48 is absent, while teeth 38 and 43 are impacted. Tooth 38 shows evidence of an periapical lesion, and areas of osteolysis are observed.

The maxilla is not included in the scan volume.

The mandibular condyles are excluded from the acquisition and are not visible.

Reference ToothFairy4 Report

Mandible: tooth 3.8 and 4.8 are in complete bony impaction. The mandibular canal has a regular course and is in close relationship with tooth 3.8 and 4.8. The maxilla is not included in the acquisition and is not assessable. The mandibular condyles are not included in the scan.



Thank you

Explicit dental evidence before report generation

Code: <https://github.com/ryhnhao/EviDent-CBCT>

Paper: <http://arxiv.org/abs/2610.02375>

If you find our methods or results useful, please consider citing us as follows:

```
@misc{hao2026evidentcbct,
  title = {EviDent-CBCT: Evidence-Bottlenecked Report Generation from Dental CBCT under Non-Exhaustive Report Supervision},
  author = {Ruiyang Hao and Zhi Qin Tan and Yulan He and Owen Addison and Yunpeng Li},
  year = {2026},
  eprint = {2610.02375},
  archivePrefix = {arXiv},
  primaryClass = {cs.CV},
  note = {Preprint},
  url = {https://arxiv.org/abs/2610.02375}
}
```

Ruiyang Hao

ruiyang.hao@kcl.ac.uk

