This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

Supplementary Material for: CI-MAS: Clinically Informed Restoration-Configuration Simulation for Dental CBCT Metal Artifact Reduction

Jiatao Wang et al.

Hangzhou Dianzi University, Hangzhou, China

1 Complete Restoration Configurations

Table S1 lists all 17 restoration configurations used in the feasibility study. The subject-disjoint partition was fixed before slice extraction: cases 2, 7, 9, 14, and 15 for training; case 18 for validation/model selection; and cases 6, 23, and 24 for testing.

Table S1. Complete CI-MAS restoration configurations. C denotes crown and F denotes filling.

Case	Variant	Restoration site(s)	Type
2	v1	C36 + F37	Mixed
2	v2	F16 + F17	Two fillings
6	v1	C37 + F36	Mixed
7	v1	C16	Single crown
7	v2	F36 + F37	Two fillings
9	v1	C36 + F37	Mixed
14	v1	C36	Single crown
14	v2	F16	Single filling
14	v3	F36 + F37	Two fillings
15	v1	C37	Single crown
15	v2	F36	Single filling
18	v1	C37	Single crown
18	v2	F16 + F36	Two fillings
23	v1	C36	Single crown
23	v2	F34 + F35	Two fillings
24	v1	C16	Single crown
24	v2	F36 + F37	Two fillings

2 Method S1. Scanner Geometry

All scans were acquired using a Jirox dental CBCT system. The source-to-detector distance was 1,700 mm and the source-to-axis distance was 1,500 mm.

Acquisition used 100 kV, 6 mA, a full 360-degree rotation, 1,600 views, a 768 × 768 detector matrix, 0.2-mm detector pitch, and an approximately 55-mm horizontal detector offset.

3 Method S2. Screening and Tooth-Semantic QC

This technical feasibility study used retrospective convenience sampling. An initial pool of 25 metal-free examinations was screened for complete projection sequences and FDK reconstructions, coverage of the target dental arch and posterior region, absence of obvious metal artifacts that would invalidate the metal-free baseline, and absence of severe motion, truncation, or reconstruction failure. Tooth-semantic post-processing merged pulp labels into the corresponding teeth, assigned crown-region labels to the nearest tooth, and refined per-tooth masks by morphological closing, hole filling, and largest-component extraction. Manual QC checked instance completeness, consistency between FDI identifiers and anatomy, and posterior-tooth continuity. Nine examinations met these criteria.

4 Method S3. Projection-Domain Degradation Details

After restoration instantiation, the metal-free volume V_{clean} and binary metal mask M were processed separately in the projection domain. The cone-beam operator A produced

$$P_{\text{clean}} = A(V_{\text{clean}}), \qquad P_{\text{metal}} = A(M). \tag{S1}$$

The simulation then applied four empirical components.

High attenuation.

$$P_{\text{sim}} = P_{\text{clean}} + \alpha\beta P_{\text{metal}}, \tag{S2}$$

with $\alpha = 1.35$ and $\beta = 1.12$.

Beam hardening. Let T denote the metal trace, defined by $P_{\text{metal}} > 0.1$ followed by a three-pixel dilation. With $s = \text{Percentile}_{99}(P_{\text{metal}}[T])$ and $t_n = \text{clip}(P_{\text{metal}}/s, 0, 4)$,

$$P_{\text{sim}} \leftarrow P_{\text{sim}} - \frac{\gamma s t_n^2}{1 + t_n}, \tag{S3}$$

where $\gamma = 0.42$.

Photon starvation. Signal-dependent Gaussian noise was amplified inside the metal trace:

$$P_{\text{sim}} \leftarrow P_{\text{sim}} + \sigma\sqrt{\max(\alpha, 10^{-6})}(1 + \delta t_n)\mathcal{N}(0, 1), \tag{S4}$$

with $\sigma = 0.009$ and $\delta = 3.6$.

Detector saturation. Projection values within T were clipped at the 99.4th percentile of the simulated projection. The degraded projection was reconstructed with the same FDK geometry used for the metal-free volume. These coefficients are proof-of-concept simulation parameters and were not calibrated as material-specific or scanner-specific physical constants.

Pseudo-code summary. The complete procedure for each subject and selected configuration is:

```
Input : metal-free V_clean, FDI label map L, configuration C
Output: triplet {V_artifact, V_clean, M_bin}

1. M      <- G(L, C)                      # instantiate metal mask
2. P_clean<- A(V_clean);  P_metal <- A(M)
3. T      <- dilate({P_metal > 0.1}, 3)  # metal trace
4. P      <- P_clean + alpha*beta*P_metal   # high attenuation
   P      <- P - gamma*s*t_n^2/(1+t_n)      # beam hardening
   P      <- P + sigma*sqrt(alpha)*(1+delta*t_n)*N(0,1)  # photon starvation
   P      <- clip(P, 99.4th percentile)     # detector saturation
5. V_artifact <- FDK(P)                 # reconstruct
6. save {V_artifact, V_clean, M_bin}
```

5 Method S4. RISE-MAR Retraining and Evaluation

The nine examinations were partitioned at the subject level before slice extraction. The resulting sets contained 724 training slices, 174 validation slices, and 392 test slices. No subject contributed slices to more than one partition. RiseMARNet was initialized from the official dental checkpoint and trained for 10 epochs with Adam, batch size 2, random 256×256 crops, random flips and rotations, and L1 supervision in non-metal regions. The best validation checkpoint was selected independently for each of three random seeds. This protocol replaces the slice-level 8:2 split used in the original submission and is the protocol reported in the camera-ready manuscript.

PSNR was calculated in non-metal regions using $I_{\max} = 3072$ HU and $I_{\min} = -1024$ HU; SSIM was computed slice-wise over the full slice and averaged over metal-containing slices. Test metrics were first averaged across configurations within each held-out subject and then summarized across the three held-out subjects; CI-MAS and RGMB results were additionally averaged over the three training seeds.

Table S2 summarizes the controlled comparison between CI-MAS and the matched Random Geometric Mask Baseline (RGMB). All confounded dimensions (subject split, training scale, restoration-component count, metal-volume range, projection simulator, and MAR training protocol) were matched; the two dimensions intentionally varied are the site semantics (FDI-guided vs. none) and the restoration geometry (tooth-local vs. random 3D).

Table S2. Controlled comparison between CI-MAS and the matched Random Geometric Mask Baseline (RGMB). "Matched" indicates the dimension was controlled to be comparable across the two simulation strategies; "identical" indicates the same protocol or implementation was used.

Aspect	CI-MAS	RGMB
Subject split	matched	matched
Training scale	matched	matched
Restoration-component count	matched	matched
Metal-volume range	matched	matched
Projection simulator	identical	identical
MAR training protocol	identical	identical
Site semantics	FDI-guided	none
Restoration geometry	tooth-local	random 3D

6 Overview of CI-MAS and Feasibility Demonstrations

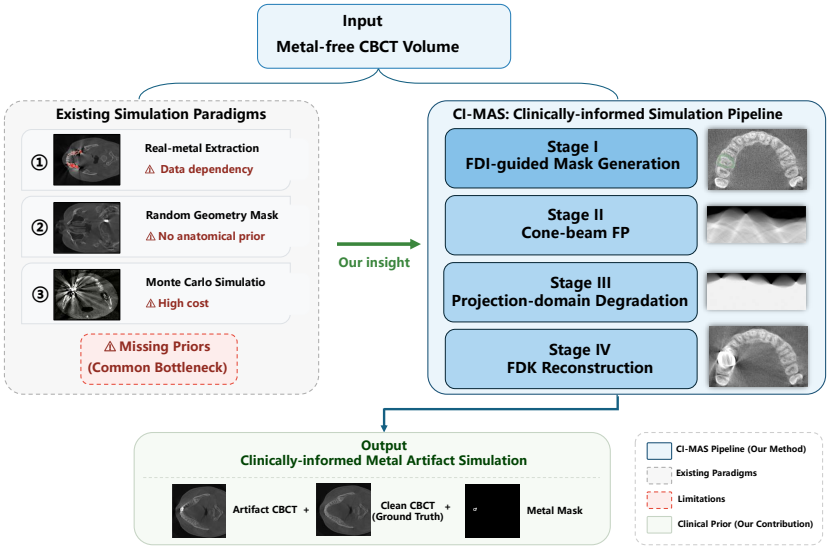


Fig. S1. Overview of CI-MAS in the context of existing simulation paradigms. Existing strategies (real-metal extraction, random-geometry masks, Monte Carlo simulation) address metal representation or artifact formation after a metal object has been specified, but share a common bottleneck: the lack of a clinical restoration-configuration prior. CI-MAS adds this decision layer before physical simulation.

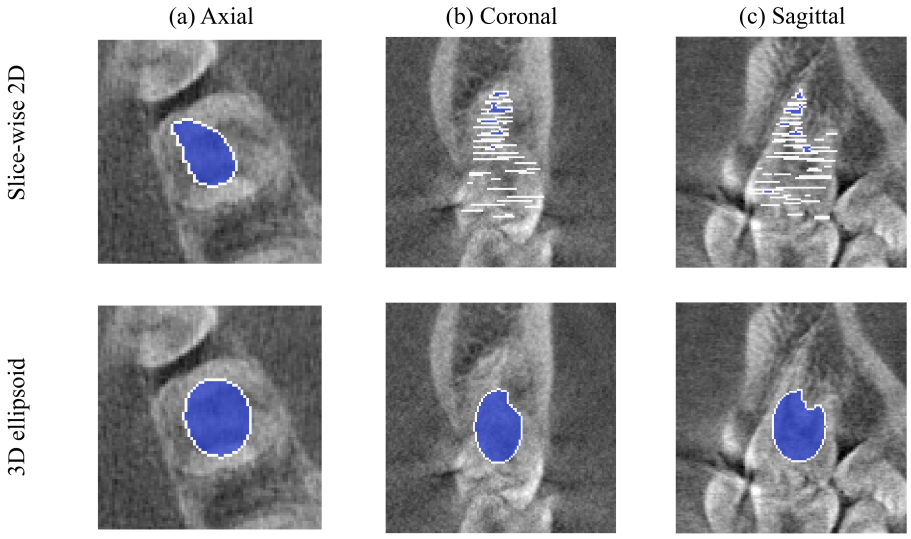


Fig. S2. Comparison of the current slice-wise filling mask and a direct 3D ellipsoidal alternative in axial, coronal, and sagittal views (Case 6, FDI 36). The 3D construction improves inter-slice continuity and is shown only as a feasibility illustration; it was not used in the reported CI-MAS experiments.

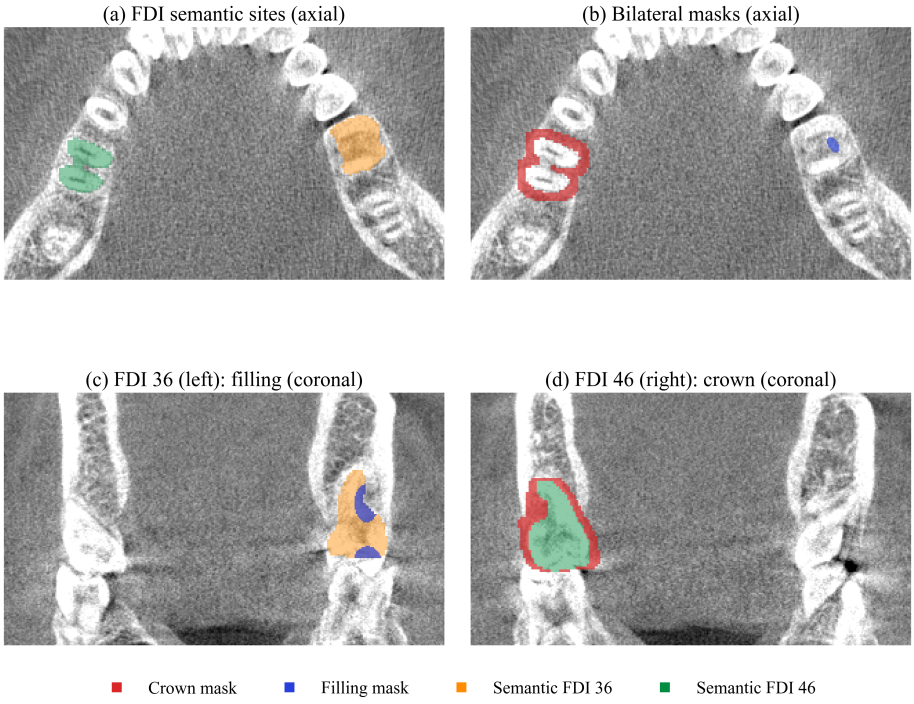


Fig. S3. Bilateral contralateral mask configuration. The same tooth-local instantiation operators were applied to a true contralateral pair (FDI 46 crown, FDI 36 filling) not included in the 17-configuration evaluation set. Top row: axial full-arch views show both FDI semantic sites and the generated bilateral masks; bottom row: coronal slices through each side show the filling (left) and crown (right), respectively. This demonstrates that the unilateral candidate set reflects the controlled scope of the proof-of-concept dataset rather than an algorithmic restriction. The additional configuration was not included in MAR training or quantitative evaluation.

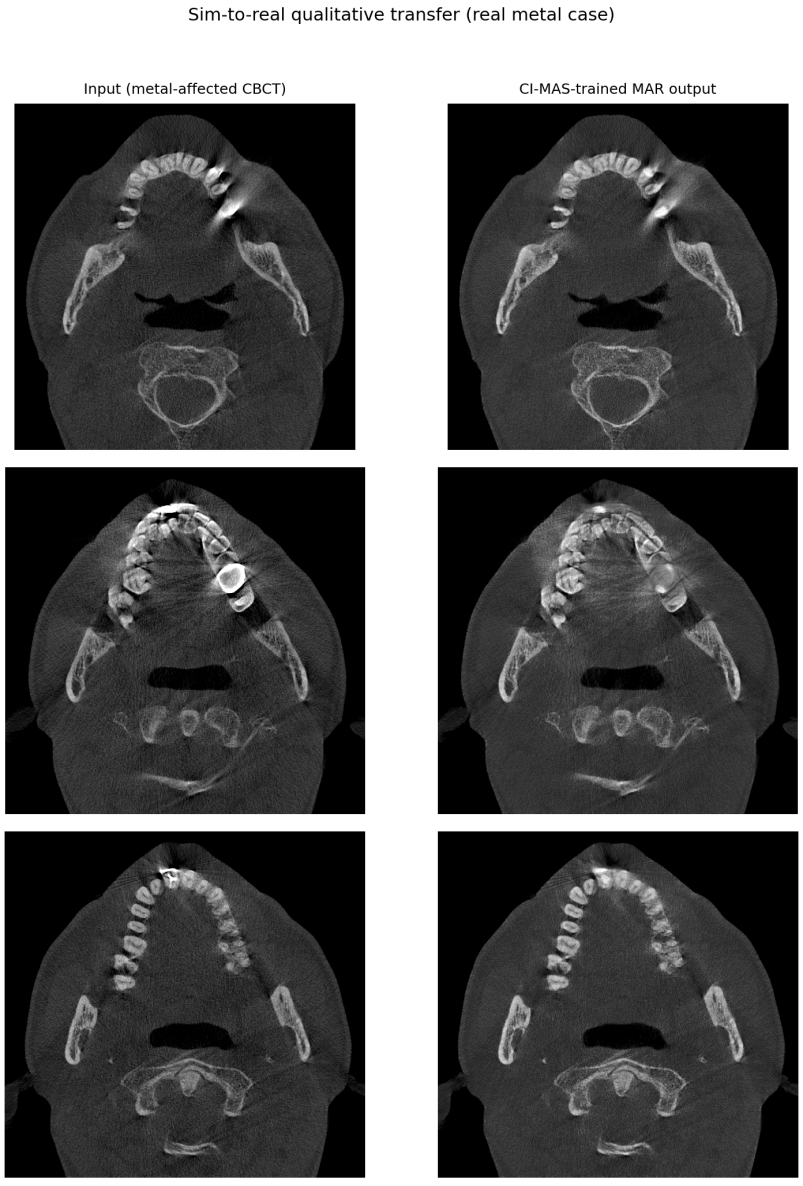


Fig. S4. Non-paired qualitative comparison with real artifact-affected CBCT examples. Clinical images are used only as visual context and are not paired ground truth or part of the quantitative evaluation.