3D Tooth Movement Prediction

End-to-End ML Pipeline

1

Data Acquisition

iTero scanner captures 3D mesh (STL format) of upper/lower arches. Patient metadata: age, gender, compliance history. Historical treatment records from 10M+ cases.

↓

Raw Data

2

Data Preprocessing

Convert mesh to voxel grid (128^3). Segment individual teeth using U-Net. Normalize coordinates, align to standard reference frame. Extract 200+ features: tooth angles, distances, crowding index, occlusion metrics.

↓

Structured Features

3

Feature Extraction (CNN)

3D ResNet-50 processes voxelized scan. Outputs 96-dimensional feature vector per tooth capturing shape, position, and surrounding anatomy. Transfer learning from 5M pre-trained cases.

↓

Spatial Features

4

Sequential Modeling (RNN)

Bidirectional LSTM (3 layers, 512 units) models temporal dependencies. Inputs: CNN features + aligner specifications + treatment timeline. Predicts tooth position at each timestep (typically 30-50 stages).

↓

Movement Trajectory

5

Prediction & Optimization

Outputs: 3D coordinates per tooth per timestep, rotation matrices, confidence scores. Physics-based constraints ensure biologically feasible movements. Reinforcement learning optimizes aligner sequence to minimize treatment time.

↓

Treatment Plan

6

Clinical Integration

Interactive 3D visualization shows predicted outcome. Orthodontist reviews/modifies AI plan. System generates manufacturing specs for each aligner. Monitoring dashboard tracks actual vs predicted during treatment.

Performance

- Accuracy: 95%+ prediction
- Position error: <0.3mm
- Processing: <3 seconds
- Success rate: 88% first-time-right

Infrastructure

- Cloud: AWS/ Azure GPU clusters
- Training: 48 hours on 16x A100
- Storage: 2PB treatment data
- API: 99.9% uptime SLA

Business Impact

- 60% faster planning
- 40% fewer revisions
- 25% higher patient satisfaction
- $200M annual savings

Pipeline Summary

The end-to-end pipeline combines computer vision (CNN) for spatial understanding with sequential modeling (RNN) for temporal prediction. Data flows from iTero scanners through preprocessing and feature extraction, then through deep learning models trained on millions of historical cases. The hybrid CNN-RNN architecture captures both the complex 3D anatomy and the temporal dynamics of tooth movement. Physics-informed constraints ensure predictions are biologically feasible, while reinforcement learning optimizes the treatment sequence. The system processes a new case in under 3 seconds, achieving 95%+ accuracy in predicting final tooth positions 12-18 months in advance. Clinical integration allows orthodontists to review, modify, and approve AI-generated plans before manufacturing begins. Continuous learning from treatment outcomes improves the model over time, creating a virtuous cycle of better predictions and outcomes.