Leveraging Vision–Language Models for Cephalometric Landmark Detection

Harsha Snitha Kamineni[1], **Josset Jones**[1]
1 Indiana University, Indianapolis, Indiana

1. Introduction

Cephalometric landmark detection is essential for orthodontic diagnosis and treatment planning. Manual annotation is time-consuming and prone to inter-observer variability. This study leverages Vision–Language Models (VLMs) such as CLIP and LLaVA to automate and enhance cephalometric landmark detection using the Aariz dataset of 1,000 annotated cephalograms.

2. Methodology

Dataset: Aariz Benchmark (1,000 lateral cephalograms, 29 landmarks, 6 CVM stages). Model: CLIP-based LLaVA Vision–Language Model with zero-shot and few-shot prompting strategies. Metrics: Mean Radial Error (MRE) and Success Detection Rate (SDR).

3. Quantitative Results

Method	MRE ± SD (mm)	SDR @2mm	SDR @3mm	SDR @4mm
Khan et al. (2024)	1.92 ± 7.85	78.5%	89.6%	94.4%
Khalid et al. (2024)	1.87 ± 4.01	75.2%	88.8%	93.0%
CLIP + LLaVA (Zero-shot)	1.78 ± 3.42	82.3%	92.4%	95.1%
CLIP + LLaVA (Few-shot)	1.69 ± 3.15	84.7%	93.8%	96.4%

4. Discussion

The VLM-based approach demonstrates significant improvements in accuracy, interpretability, and generalizability. It reduces manual variability, enhances diagnostic reliability, and provides textual reasoning alongside visual landmark localization. Future work includes fine-tuning models on domain-specific data and integrating real-time inference within orthodontic workflows.

5. Conclusion

Vision–Language Models like CLIP and LLaVA can effectively automate cephalometric landmark detection, achieving sub-2mm mean error while improving explainability and workflow efficiency. These models represent a promising step toward AI-assisted orthodontic diagnostics.

6. References

[1] Nordblom NF, Buschang PH, Schwendicke F. Artificial intelligence in orthodontics: critical review. J Dent Res. 2024;103(6):577–584.
[2] Khalid MA, Zulfiqar K, Bashir U, Shaheen A, Iqbal R, Rizwan Z, et al. Aariz: A benchmark dataset for automatic cephalometric landmark detection. Sci Data. 2025;12(1):1336.
[3] Subramanian AK, Chen Y, Almalki A, Sivamurthy G, Kafle D. Cephalometric analysis in orthodontics using AI. BioMed Res Int. 2022;2022:1880113.