A Hybrid Deep Learning and Rule-Based Approach for Automatic Cephalometric Landmark Detection in 2D X-ray Images

Elham Hekmatnia[1], Matthew S. Holden[2]
[1,2] School of Computer Science, Carleton University, Ottawa, ON, Canada

Introduction: Accurate landmark localization in Cephalometric analysis plays a crucial role in orthodontic care and maxillofacial surgery. To rule out any skeletal or dental malocclusion, specific anatomical landmarks on Cephalometric images should be analyzed. Variations in anatomy, poor image quality, or overlapping structures can make it difficult to accurately locate some landmarks manually. The purpose of this study is to automate localization.

Methods: A total of 300 digital lateral Cephalometric radiographs with varying resolutions, obtained from a private orthodontic clinic, were used to localize 19 anatomical landmarks per image (Fig 1). The dataset was divided into three subsets: 150 images for training, 75 for validation, and 75 for testing. Preprocessing included noise removal and cropping, image contrast enhancement, and intensity standardization using a linear rescaling procedure. First, an end-to-end deep learning model called FARNet [1] was employed because of its state-of-the-art performance on public datasets. It uses a backbone network of Dense-Net pre-trained on ImageNet as an encoder. Its decoder is composed of two key parts: (1) the multi-scale feature aggregation (MSFA) module, which integrates multi-scale features from the encoder, and (2) the feature refinement (FR) module, which fuses features from the original image with up-sampled features and heatmaps from the MSFA. Since some soft tissue landmarks are difficult to detect due to low image contrast, anatomical variability, and overlapping structures, the localization accuracy achieved by FARNet for three critical soft tissue landmarks was not satisfactory. This further improved by rule-based image processing techniques where a Canny edge detection was modified with optimized smoothing and thresholding, followed by edge linking and edge tracking (Fig 2).

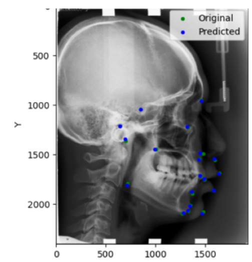

Fig1: Lateral Cephalogram with 19 landmarks

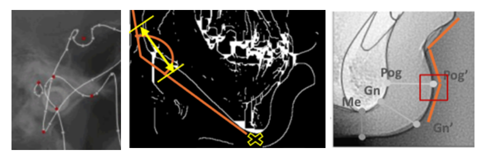

Fig2: Refinement of the location of critical landmarks (left: Porion, middle: Gonion, right: soft-tissue Pogonion)

Results: The performance of the proposed landmark detection method was quantitatively assessed using the Successful Detection Rate (SDR) at various precision thresholds and the Mean Radial Error (MRE). To assess the effectiveness of the proposed rule-based refinement methods which consider the expected shape or relative location of anatomical features, applied to FARNet outputs, we further evaluated the localization performance of the three critical soft tissue landmarks: Porion, Gonion, and soft-tissue Pogonion (Pog'). After applying the refinement, SDR for a 2 mm threshold improved from 48% to 57% for Porion, from 35% to 40% for Gonion, and from 66% to 69% for Pog' on our dataset. Likewise, performance improved after refinement on the public ISBI 2015 cephalometric X-ray dataset.

Table 1: Comparison of FARNet and FARNet + Rule-based refinement on public and private datasets

Methods	Dataset	MRE	SDR %			
			2mm	2.5mm	3mm	4mm
FARNet [1]	Public dataset (ISBI)	1.42	77.00	84.42	89.47	95.21
FARNet + Rule based		**1.36**	**79.44**	**84.63**	**91.40**	**96.31**
FARNet	Private dataset (ours)	1.64	77.08	83.64	87.57	92.56
FARNet + Rule based		**1.57**	**77.96**	**84.89**	**88.78**	**94.53**

Conclusions: Without applying any exclusion criteria, we used a dataset and compared our study with previous works that employed traditional methods, considered the same landmarks, and reported similar error metrics. Using FARNet, performance for three critical landmarks (i.e. Gonion, Porion and Pog') showed lower performance. By leveraging rule-based algorithms to identify landmarks located within well-defined anatomical shapes, lower MREs were achieved for these landmarks, particularly at clinically relevant thresholds of 2–4 mm for both our private dataset and a public dataset. The mean radial error (MRE) in this study was 1.57 mm, which falls within the acceptable range for cephalometric analysis according to previously reported clinical standards.

References: [1] Ao Y, Wu H. Feature aggregation and refinement network for 2D anatomical landmark detection. Journal of Digital Imaging. 2023 Apr;36(2):547-61.