Accurate Cephalometric Landmark Detection Based on Knowledge Distillation

Yu Song[1], Xu Qiao[2], Yinhao Li[1], Rahul Kumar Jain[1] and Yen-Wei Chen[1]

[1]Graduate School of Engineering, Ritsumeikan University, Japan
[2] Department of Biomedical Engineering, Shandong University, Jinan, China
E-mail: gr0398ep@ed.ritsumei.ac.jp

Cephalometric analysis is widely utilized by dental professionals and orthodontists to examine the role of the human skull in diagnosing conditions and devising treatment strategies. To conduct such an analysis, the manual annotation of cephalometric landmarks is a prerequisite. In recent times, various automatic cephalometric landmark detection methods have been developed, employing deep learning techniques, which have demonstrated promising outcomes. Among these methods, regression based on convolutional neural networks (CNN) is prevalent, wherein the ground-truth (GT) information primarily informs the calculation of the loss function, thereby facilitating the refinement of predicted landmark locations through backpropagation. However, owing to the scarcity of annotated cephalometric data, there is potential for enhanced performance through improved utilization of ground-truth information.

In this study, we propose a knowledge distillation approach employing GT images to achieve precise cephalometric detection. Initially, we employ images composed of GT landmarks as input to train a detection model, which serves as the teacher model. Subsequently, the teacher model imparts guidance to a student model, trained on original images, through the distillation of pertinent features. We posit that the features shared between GT images and original images exhibit comparable domain distribution, as they both represent the same anatomical structure. Our proposed method is rigorously evaluated using a public grand challenge dataset, and it demonstrates superior performance when compared to state-of-the-art techniques.

Furthermore, we enhance our knowledge distillation approach by incorporating focal and global distillation during the distillation process. Specifically, we introduce a masking mechanism to differentiate between the foreground and background areas of landmarks, facilitating separate focal distillation. Additionally, we implement a feature similarity penalty to mediate between the teacher and student models, which acts as the global distillation process. The efficacy of our refined approach is thoroughly evaluated on an in-house dataset.