Caries Detection Performance between A Deep-Learning Model and Human Subjects

Maryam Baldawi[1], Tanjida Kabir[2], MS, Sung K Kim [1], DDS, Ana Neumann[1], DDS, MPH, PhD, Jiang, Xiaoqian[2], PhD, Chun-Teh Lee[1], DDS, MS, DMSc, Shayan Shams [2,3] PhD, [1]University of Texas Health Science Center at Houston, School of Dentistry, [2]University of Texas Health Science Center at Houston School of Biomedical Informatics, Houston, Texas, USA, [3]Department of Applied Data Science, San Jose State University, San Jose, California

Objectives: Application of deep learning to assist clinicians to detect diagnose caries radiographically is growing in dentistry. Convolutional neural network (CNN), a deep learning algorithm inspired by the human brain, can learn hierarchical features through multiple layers to make classification or segmentation tasks trivial. This study aimed to compare varying levels of human expertise performance in radiographic caries lesion detection to the performance of the deep learning CNN model on radiographic images.

Methods: A CNN model was developed to detect caries on the intraoral radiographic images, as demonstrated in figure 1. Among 362 periapical and bitewing radiographic images, 80% of images were used for training, 10% for model validation and 10% for testing. U-Net with attention was used to program the model to overcome irrelevant features transfer. The performance of caries detection by the model was evaluated by the positive predictive value (PPV), negative predictable value (NPV) and F1-score. Then, 43 human subjects of varying levels of experience (12 dental students without clinical experience (group DN), 20 dental students with clinical experience (group DE), 8 residents (group ER), 3 experienced dentists (group ED)) annotated 10 radiographic images and the subject's predictions were compared to that of the model's using dice similarity coefficients (DSC) score. Total numbers of identified radiographic caries were compared by one-way ANOVA (analysis of variance).

Results: For the model's performance, the PPV was 0.998, NPV was 0.947 and F1-score was 0.999. The mean total numbers of identified radiographic caries in the four groups were 31.58±9.29 (DN), 38.60±11.36 (DE), 40.13±7.68 (ER) and 41.33±9.02 (ED) respectively. There was no significant difference between groups. The agreement of caries detection between the model and human subjects will be analyzed in the future.

Conclusion: The deep learning model can help dental education and clinical practice with real-time and reliable radiographic caries detection.