Autonomous Deep Learning for Pediatric Bone Age Estimation Using Historical Atlases

Authors: Samuel Lane, Carson Wood, Samuel Turner, Joshua Gullace, Rajendram Rajnarayanan

Background: Automated bone age assessment using deep learning has shown promise in pediatric radiology. This study investigates image CNN backbones for estimating skeletal maturity from hand radiographs and their relationship to the Greulich and Pyle methodology.

Objectives: To compare the performance of four CNN architectures (DenseNet121, EfficientNetB0, MobileNetV2, and ResNet50) trained with autonomous deep learning on the RSNA Pediatric Hand dataset, and to compare model performance between continuous ages and Greulich and Pyle intervals.

Methods: Four CNN architectures were trained using semi-supervised learning on the RSNA Pediatric Hand dataset to estimate skeletal age. All models were fine-tuned via transfer learning in a custom training pipeline that leveraged autonomous hyperparameter tuning. Performance was evaluated using accuracy, precision-recall, and ROC metrics. Models were compared between two approaches: regression on continuous ages and classification using Greulich and Pyle skeletal age intervals.

Results: Models trained with Greulich and Pyle intervals achieved 1.5 to 2 times higher validation and test accuracies compared to continuous age regression, with improved macro-averaged PR and ROC AUC scores. EfficientNetB0 and ResNet50 showed the greatest improvements.

Conclusion: CNN backbone architectures, particularly EfficientNetB0 and ResNet50, demonstrate significant performance gains when trained on pediatric skeletal age estimation tasks. Structuring training data using clinically established Greulich and Pyle intervals substantially enhances model accuracy. Future work should explore additional data augmentation, hyperparameter optimization, and physical optimizations such as region masking to further improve diagnostic performance across diverse architectural designs.

Submitted by Samuel Lane