Beyond Human Perception: Sexual Dimorphism in Hand and Wrist Radiographs Is Discernible by a Deep Learning Model

Sehyo Yune,# Hyunkwang Lee,# Myeongchan Kim, Shahein H. Tajmir, Michael S. Gee, and Synho Do

Reflection by Prateek Naharia | nahariap@bu.edu | MSBA, Boston University.
Dt. 3rd October 2023

Problem Being Addressed

The paper addresses the lack of detailed understanding and recognition of sexual dimorphism (distinct differences between males and females) in hand and wrist radiographs. While sex and sex hormones significantly impact bone structure and density, radiologists have had limited success in distinguishing sexes based on hand radiographs.

Contribution Made

1. Deep Learning Model Development: A deep convolutional neural network (CNN) model was developed to predict sex based on hand radiographs of individuals aged 5 to 70 years. The model achieved a remarkable accuracy of 95.9% in predicting the correct sex from a total of 1531 radiographs tested.
2. Comparative Evaluation: Two human radiologists were only able to achieve 58% and 46% accuracy, which starkly contrasts with the model's performance.
3. Visualization and Understanding: Attempts were made to understand and visualize the features the model focused on to distinguish between sexes, using techniques like Class Activation Mapping (CAM) and t-distributed Stochastic Neighbor Embedding (t-SNE).
4. Potential Clinical Implications: The model's ability to distinguish sex based on radiographs was seen as a stepping stone towards exploring associations between bone sex and various clinical outcomes, potentially extending the utility of bone age radiographs.

Evaluation Used (or the Argument Made)

1. Model Evaluation: The model was evaluated based on its accuracy in predicting sex from hand radiographs, achieving a high accuracy of 95.9% and a Cohen's kappa coefficient (κ) of 0.918, indicating strong agreement between the model predictions and the actual sexes.
2. Radiologist Evaluation: Two radiologists were evaluated on their ability to predict sex from radiographs, both before and after reviewing the model's feature descriptions and attention maps. Their accuracies remained significantly lower than the model's, showcasing the model's superior ability to identify sexual dimorphism.

Opinion on Whether It's a Meaningful Contribution

This work demonstrates a meaningful contribution in leveraging deep learning to unearth sexual dimorphism in hand and wrist radiographs, a task that proved challenging for human experts. The high accuracy achieved by the model, coupled with the potential clinical implications, signifies a notable advancement in medical image analysis and sets a precedent for further exploration into the hidden information within radiographs.

Limitations About the Paper

1. Dataset Limitation: The study only included left-hand radiographs, which could limit the model's applicability in clinical practice or other research endeavors that may include right-hand images.

2. Interpretability Challenge: Despite the model's high accuracy, the paper noted a difficulty in identifying the sex-specific radiographic features, which is a common challenge with deep learning models known as the "black box" problem. This lack of interpretability could pose challenges in adopting such models in clinical settings where understanding the decision-making process is crucial.
3. Small Sample Size in Preliminary Test: The preliminary test on selected radiographs from individuals with certain conditions had a small sample size, which limits the conclusiveness of the findings regarding the clinical implication of bone sex.