SIIM CMIMI24
Oct 21-22 | Boston, MA
Conference On Machine Intelligence In Medical Imaging

From AI to Eye: Training the Radiologist with Deep Learning Interpretations in Sex Differentiation

Shahriar Faghani, MD, Mayo Clinic; Christin A. Tiegs-Heiden, MD; Mana Moassefi, MD; Garret Powell, MD; Micheal Ringler, MD; Bradley Erickson, MD, PhD, CIIP, FSIIM; Nicholas Rhodes, MD

Introduction/Background

To use deep learning (DL) as an educational and scientific discovery tool to improve the radiologist's ability to directly make subtle imaging findings without additional DL assistance.

Methods/Intervention

We present a DL model that can identify sex differences from frontal knee radiographs with high accuracy, then use the resultant occlusion interpretation maps(OIMs) to train human readers to improve their ability to perform this same task. Two groups, each of three human readers, were tasked to separate radiographs into male and female sex correctly. Both groups were informed of the patient's sex, while the first group was also given these radiographs OIMs. After two weeks, the group was retested with a new set of 50 radiographs. This group was compared to a second group trained without the OIMs.

Results/Outcome

The DL model separated sex with 0.96 accuracy. The average accuracy of the six human readers initially was 0.62(range:0.56-0.74). After the study, the average accuracy of the six human readers was 0.77(range:0.7-0.84). The improvement in accuracy of the six human readers was statistically significant (p=0.0364). The accuracy of the "heat map" group was 0.8, and the control group was 0.74. When pooled as a collective group, Group 1 again showed significant improvement from baseline(p=0.0058), whereas Group 2 did not(p=0.1245), though there was no statistical difference between the two groups at the end of the experiment(p=0.2380).

Conclusion

OIMs could not be shown to definitively account for the improved accuracy in our test, though the group provided those maps demonstrated a statistically significant improvement from baseline, while the group without these maps did not. Moreover, simply the high accuracy of the DL model in performing this task proved it was possible and motivated our human readers to learn to perform this task.

Statement of Impact

This initiative seeks to identify new imaging biomarkers, thereby improving the functionality of existing DL systems and enabling human-led scientific advancements beyond the reach of DL alone. Additionally, this strategy incorporates a layer of explainability to facilitate the monitoring and troubleshooting of DL models when errors occur, contributing to the development of DL systems with increased resilience to such errors.

Keywords

Deep learning; Interpretability; Classification; MSK radiology; Knee radiographs; Education