AI Trained on a Diverse Dataset Perform Better Chest X-ray Analysis

November 29, 2021
[Katie Robinson](#)

Deep learning models trained on a dataset lacking racial diversity could hinder the detection of pathology in underrepresented minority patients.

A study presented at the [Radiological Society of North America (RSNA) 2021 Annual Meeting](#) demonstrates the importance of using racially diverse datasets while training artificial intelligence (AI) systems to ensure fair outcomes.

"As the rapid development of deep learning in medicine continues, there are concerns of potential bias when interpreting radiological images," the authors wrote. "As future medical AI systems are approved by regulators, it is crucial that model performance on different racial/ethnic groups is shared to ensure that safe and fair systems are being implemented."

The findings were presented by Brandon Price, a medical student at Florida State University College of Medicine in Tallahassee.

Many studies have shown that deep learning systems are subjective in their interpretation of data. Bias is often accidentally introduced into the training data, or a racial or ethnic group is under sampled causing susceptible models to develop bias. In this study, the researchers investigated how a deep learning model trained on a dataset lacking racial diversity could impede the detection of pathology in underrepresented minority patients.

The researchers used a dataset with over 300,000 chest X-ray images and 14 labeled findings. A low sample size of other races/ethnicities meant that only images of Black and White patients were included. One training dataset included only White patients and the other training dataset comprised 26% Black and 74% White patients. An equal distribution of labeled findings was shared between the datasets and a DenseNet model was trained on each dataset 25 times. The receiver operating characteristics (ROC) area under the curve (AUC) and sensitivity, with a specificity threshold of 0.75, were compared for each of the 14 labeled findings.

Compared with a model trained on only White patients, the model trained with a diverse dataset had a significantly better ROC-AUC performance at identifying six of the 14 labeled findings in a test dataset of only Black patients (P <0.05). Additionally, compared with a model trained on only White patients, the model trained with a diverse dataset found a significant increase in sensitivity performance for six of the 14 labeled findings on a test dataset of only Black patients (P <0.05).

"As more AI systems are developed, it is imperative that they are fair and perform equally well with groups that have been historically underserved," the authors wrote.

For more coverage of RSNA 2021, [click here](#).

Related Content:

[Conferences](#) | [X-Ray](#) | [Technology](#)

NSAID Use May Affect Axial Spondyloarthritis Diagnosis via MRI

Breast Ultrasound/MRI Fusion May Help Reduce MRI-Guided Biopsies

Cardiovascular Imaging Manifestations of COVID-19 to Watch For