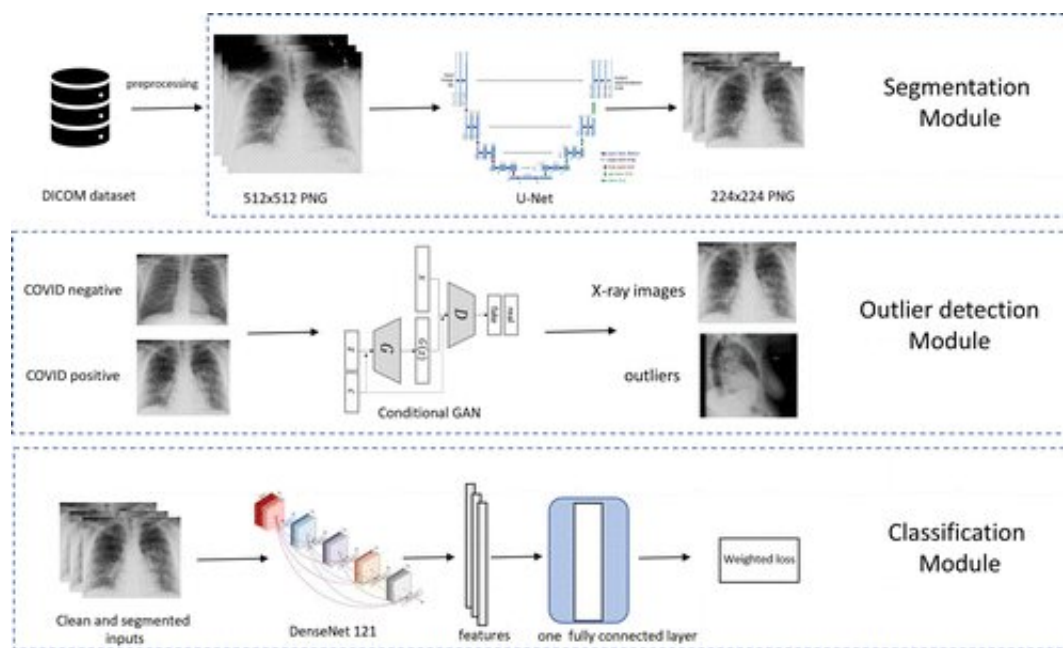# Researchers determine that AI-based tools have not yet reached full diagnostic potential in COVID-19

July 28 2022, by Jack Alexander

Overview of the COVID-19 diagnostic model pipeline shows segmentation module (top), outlier detection module (middle), and classification module (bottom). DICOM = Digital Imaging and Communications in Medicine, GAN = generative adversarial network, PNG = portable network graphics format. Credit: Ju Sun et al, *Radiology: Artificial Intelligence* (2022). DOI: 10.1148/ryai.210217

Published by the journal *Radiology: Artificial Intelligence*, a prospective observational study across 12 hospital systems from the University of

Minnesota Medical School has evaluated the real-time performance of an interpretable artificial intelligence (AI) model to detect COVID-19 from chest X-rays.

Participants with COVID-19 had a significantly higher COVID-19 diagnostic score than participants who did not have COVID-19. However, researchers found the real-time model performance was unchanged over the 19 weeks of implementation. The model sensitivity was significantly higher in men, while model specificity was significantly higher in women. Sensitivity was significantly higher for Asian and Black participants than for white participants. The COVID-19 AI [diagnostic system](link) had significantly worse accuracy than predictions made by radiologists.

"This study, which represents the first live investigation of an AI COVID-19 diagnostic model, highlights the potential benefits but also the limitations of AI," said Christopher Tignanelli, MD, MS, FACS, FAMIA, an associate professor of surgery at the University of Minnesota Medical School and general surgeon at M Health Fairview. "While promising, AI-based tools have not yet reached full diagnostic potential."

The research findings were informed by an AI algorithm developed by Ju Sun, an assistant professor at the U of M College of Science and Engineering, and his team in collaboration with M Health Fairview and Epic.

- COVID-19 diagnostic models perform well for participants with severe COVID-19 effects; however, they fail to differentiate participants with mild COVID-19 effects.
- Many of the early-pandemic AI models that were published boasted overly optimistic performance metrics using publicly available datasets.

- The AI model's diagnostic accuracy was inferior to the [predictions](link) made by board-certified radiologists.

"We saw the same overly optimistic performance in this study when we validated against two publicly available datasets; however, as we showed in our manuscript, this does not translate to the real world," Dr. Tignanelli said. "It is imperative moving forward that researchers and journals alike develop standards requiring external or real-time prospective validation for peer-reviewed AI manuscripts."

Researchers hope to develop a simpler diagnostic AI model by integrating data from more than 40 U.S. and European sites and multi-modal models that leverage structured data and clinical notes along with images.

More information: Ju Sun et al, Performance of a Chest Radiograph AI Diagnostic Tool for COVID-19: A Prospective Observational Study, *Radiology: Artificial Intelligence* (2022). [DOI: 10.1148/ryai.210217](link)

Provided by University of Minnesota Medical School