

Bern, 15.04.2021

b
UNIVERSITÄT
BERN

[Bern Data Science Day 2021 - April 23](#)

Contribution - A Deep-Learning Diagnostic Support System for the Detection of COVID-19 Using Chest Radiographs: A Multireader Validation Study

DS applications and challenges in Medicine, Natural Sciences, and Engineering

Matthias Fontanellaz

From the ARTORG Center for Biomedical Engineering Research, University of Bern.

matthias.fontanellaz@artorg.unibe.ch

Lukas Ebner

Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital.

lukas.ebner@insel.ch

Adrian Huber

Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital.

adrian.huber@insel.ch

Alan Peters

Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital.

alan.peters@insel.ch

Laura Löbelenz

Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital.

Cynthia Hourscht

Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital.

Jeremias Klaus

Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital

Jaro Munz

Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital.

jaro.munt@szb-chb.ch

Thomas Ruder

Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital.

thomas.ruder@insel.ch

Dionysios Drakopoulos

Department of Radiology, Division City and County Hospitals, Inselgroup, Bern University Hospital, University of Bern, Bern, Switzerland.

dionysios.drakopoulos@insel.ch

Dominik Sieron

Department of Radiology, Division City and County Hospitals, Inselgroup, Bern University Hospital, University of Bern, Bern, Switzerland.

Elias Primetis

Department of Radiology, Division City and County Hospitals, Inselgroup, Bern University Hospital, University of Bern, Bern, Switzerland.

elias.primetis@insel.ch

Johannes T Heverhagen

Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital.

johannes.heverhagen@insel.ch

Stavroula Mouggiakakou

From the ARTORG Center for Biomedical Engineering Research, University of Bern.

Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital.

Stavroula.mouggiakakou@artorg.unibe.ch

Andreas Christe

Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital.

Department of Radiology, Division City and County Hospitals, Inselgroup, Bern University Hospital, University of Bern, Bern, Switzerland.

andreas.christe@insel.ch

Abstract

The aim of this study was to compare a diagnosis support system to detect COVID-19 pneumonia on chest radiographs (CXRs) against radiologists of various levels of expertise in chest imaging.

Materials and methods: Five publicly available databases comprising normal CXR, confirmed COVID-19 pneumonia cases, and other pneumonias were used. After the harmonization of the data, the training set included 7966 normal cases, 5451 with other pneumonia, and 258 CXRs with COVID-19 pneumonia, whereas in the testing data set, each category was represented by 100 cases. Eleven blinded radiologists with various levels of expertise independently read the testing data set. The data were analyzed separately with the newly proposed artificial intelligence-based system and by consultant radiologists and residents, with respect to positive predictive value (PPV), sensitivity, and F-score (harmonic mean for PPV and sensitivity). The χ test was used to compare the sensitivity, specificity, accuracy, PPV, and F-scores of the readers and the system.

Results: The proposed system achieved higher overall diagnostic accuracy (94.3%) than the radiologists (61.4% \pm 5.3%). The radiologists reached average sensitivities for normal CXR, other type of pneumonia, and COVID-19 pneumonia of 85.0% \pm 12.8%, 60.1% \pm 12.2%, and 53.2% \pm 11.2%, respectively, which were significantly lower than the results achieved by the algorithm (98.0%, 88.0%, and 97.0%; $P < 0.00032$). The mean PPVs for all 11 radiologists for the 3 categories were 82.4%, 59.0%, and 59.0% for the healthy, other pneumonia, and COVID-19 pneumonia, respectively, resulting in an F-score of 65.5% \pm 12.4%, which was significantly lower than the F-score of the algorithm (94.3% \pm 2.0%, $P < 0.00001$). When other pneumonia and COVID-19 pneumonia cases were pooled, the proposed system reached an accuracy of 95.7% for any pathology and the radiologists, 88.8%. The overall accuracy of consultants did not vary significantly compared with residents (65.0% \pm 5.8% vs 67.4% \pm 4.2%); however, consultants detected significantly more COVID-19 pneumonia cases ($P = 0.008$) and less healthy cases ($P < 0.00001$).

Conclusions: The system showed robust accuracy for COVID-19 pneumonia detection on CXR and surpassed radiologists at various training levels.