Interdisciplinary Journal of Health, Environment and Computation
ISSN Online: 3108-1444 Vol.: 1, No.: 1 (July – September) 2025. PP. : 1-8

ARTIFICIAL INTELLIGENCE FOR DETECTING CORONAVIRUS (COVID-19): AN EXTENDED REVIEW

Mohammed Hussein Khalil[1], Ank Gabor[2]

[1]*University of Information Technology and Communications, Baghdad, Iraq.*

[2]*Dicle University, Turkey.*

Email : Ank.gabor@yahoo.com

https://doi.org/10.65983/ijhec.2025.01.0001

Abstract:

Rapid, accurate detection of coronavirus disease 2019 (COVID-19) has been essential for epidemic control. While reverse-transcription polymerase chain reaction (RT-PCR) remains the reference standard, constraints in turnaround time, access, and sensitivity under certain conditions have motivated the use of Artificial Intelligence (AI) to assist screening and diagnosis from complementary data sources: medical images (chest radiographs and computed tomography), respiratory audio (cough/voice/breath), wearable and consumer-device signals, and clinical/electronic health record (EHR) data. This review synthesizes key datasets, model families, validation strategies, and performance trends, and highlights persistent pitfalls including dataset bias, information leakage, lack of external validation, and explainability gaps. We conclude with a roadmap for robust, clinically useful AI systems: rigorous study design, standardized reporting, multicenter external testing, prospective evaluation, human factors integration, and governance for safety, privacy, and equity.

Keywords - COVID-19; SARS-CoV-2; artificial intelligence; machine learning; deep learning; chest X-ray; CT; cough; voice; wearables; EHR; diagnosis; screening; triage.

Received: July 12, 2025; Revised: August 24, 2025; Accepted: September 04, 2025

1. Introduction

The COVID-19 pandemic spurred an unprecedented wave of AI research aimed at assisting detection, triage, and prognosis. Early work focused on distinguishing COVID-19 from other causes of pneumonia on chest imaging; soon after, researchers explored respiratory audio captured by smartphones, passive sensing from wearables, and structured clinical data. Despite thousands of publications, translation to routine care has been uneven. This review concentrates on AI for *detection*—identifying current infection or COVID-compatible pneumonia—rather than prognosis, resource allocation, or epidemiologic modeling. We survey data modalities, representative datasets, model architectures, validation practices, and performance, then discuss common pitfalls and pathways to trustworthy deployment.

2. Data Modalities and Typical AI Pipelines

AI detection systems follow a common pipeline: (i) data acquisition; (ii) quality control and preprocessing (normalization, augmentation, segmentation); (iii) model training with appropriate cross-validation; (iv) evaluation on internal and external test sets; and (v) explainability and uncertainty quantification. Below we outline strengths/limitations by modality.

2.1 Chest X-ray (CXR)

Advantages: ubiquitous, low cost, portable. Challenges: subtle findings, confounding from devices/positioning, and label noise due to imperfect RT-PCR ground truth.

2.2 Computed Tomography (CT)

Advantages: higher sensitivity for pulmonary opacities and extent quantification. Challenges: resource intensity, radiation, and variable protocols across scanners and sites.

2.3 Lung Ultrasound (LUS)

Advantages: bedside, no radiation; promising for triage. Challenges: operator dependence and limited large public datasets.

2.4 Respiratory Audio (Cough/Voice/Breath)

Advantages: scalable, non-invasive, amenable to telehealth. Challenges: heterogeneous devices/environments, label quality, and the need for symptom controls to avoid confounding.

2.5 Wearables and Mobile Sensing

Advantages: continuous passive monitoring (heart rate, step count, sleep, skin temperature). Challenges: population selection bias, temporal confounding, and personalization requirements.

2.6 Clinical/EHR Data

Advantages: rich multimodal features (vitals, labs, symptoms). Challenges: missingness, shift over time, and transportability.

3. Representative Public and Consortial Datasets

Table 1 summarizes widely referenced COVID-19 detection datasets. Availability and licensing are subject to change; researchers should confirm governance and intended use.

Dataset / Source	Modality	Approx. Size (period)	Notes	Citation
COVID-19 Image Data Collection	CXR/CT	Thousands (2020–2022)	Curated from publications; heterogeneous quality	Cohen et al., 2020
COVID-Net CXR datasets	CXR	Tens of thousands (2020–2021)	Benchmarks for COVID-Net models	Wang & Wong, 2020
BIMCV-COVID19+	CXR/CT	Thousands (2020–2021)	Spanish cohort with metadata	de la Iglesia Vayá et al., 2020
UK Biobank COVID-19 Imaging	CXR/CT/MRI	Thousands (2020–2021)	Linked clinical/longitudinal data	UK Biobank Consortium
MosMedData	CT	1,100+ (2020)	Severity-graded CT scans from Moscow	Morozov et al., 2020
RSNA International COVID-19 Open Radiology Database (RICORD)	CXR/CT	Thousands (2020–2021)	Expert-labeled radiology data	RSNA/ACR, 2021
Coswara / Cambridge COVID-19 Sounds / COUGHVID	Audio	Tens of thousands (2020–2022)	Crowdsourced cough/voice/breath	Brown et al., 2021; Orlandic et al., 2021
Scripps DETECT / RADAR-Base	Wearables	Tens of thousands (2020–2021)	Consumer wearable + symptom data	Radin et al., 2020; Quer et al., 2021
Israeli MOH Symptom Survey	Clinical/Symptom	Millions (2020)	Self-reported symptoms linked to tests	Zoabi et al., 2021

4. Model Families and Training Practices

Imaging models predominantly use transfer learning with ResNet, DenseNet, EfficientNet, and custom architectures such as COVID-Net. Segmentation (e.g., U-Net) aids lesion localization and severity scoring. Audio models often use CNNs on spectrograms, sometimes with RNN/Transformer back-ends. Wearable/EHR models employ gradient boosting, random forests, or temporal deep networks. To reduce overfitting and leakage, robust design choices include patient-level splits, site-held-out external tests, time-based validation, augmentation, calibration, and pre-registration of analysis plans.

5. Survey of Results

Aggregated performance varies by modality, dataset quality, and evaluation design. Tables 2–4 summarize representative peer-reviewed results. Values are indicative, emphasizing study design and validation strategy over single-number leaderboards.

Table 2. Selected Imaging-Based Detection Studies

Study	**Modality**	**Model**	**Data/Validation**	**Reported Metric(s)**	**Notes**
Li et al., 2020 (Radiology)	CT	DL classifier + lesion analysis	1.3k pts; internal & external test	AUC ~0.96 (COVID vs CAP)	Multicenter; early CT focus
Harmon et al., 2020 (Nat Commun)	CT	3D CNN	Multinational; external tests	AUC >0.90 (site-pooled)	Showed generalization across scanners
Bai et al., 2020 (Radiology)	CT	AI + radiologist comparison	Multicenter; reader study	AI AUC ~0.90	Augmented radiologist accuracy
Wang & Wong, 2020 (COVID-Net)	CXR	Tailored CNN (COVID-Net)	Public CXR sets; external test	Accuracy/AUC vary by split	Open-source model/datasets
Ozturk et al., 2020 (CBM)	CXR	DarkCovidNet	Multi-source CXR; holdout test	Acc. ~98% (binary)†	High internal score; caution on leakage
RSNA RICORD Benchmarks, 2021	CXR/CT	ResNet/EfficientNet	Expert labels; external eval	AUC ~0.80–0.92	More conservative, standardized

† High internal accuracies often drop on external datasets; see Discussion on generalization.

Table 3. Selected Audio-Based Detection Studies

Study	**Signal**	**Model**	**Data/Validation**	**Reported Metric(s)**	**Notes**
Laguarta et al., 2020 (IEEE OJEMB)	Cough/Voice	CNN + biomarker features	4k+ samples; held-out test	Acc./AUC up to ~0.97 (subset)	Smartphone-only feasibility study
Brown et al., 2021 (IEEE	Cough/Voice	CNN on log-mel	Crowdsourced; cross-val &	AUC ~0.80–	Importance of

T-ASLP)		spectrograms	external	0.90	symptom-matched controls
Orlandic et al., 2021 (COUGHVID)	Cough	Classical ML + CNN	Crowdsourced; demographic strat.	AUC ~0.70–0.85	Label noise and device variability

Table 4. Selected Wearables and Clinical Data Studies

Study	Data	Model	Design	Reported Metric(s)	Notes
Mishra et al., 2020 (Nat Biomed Eng)	Smartwatch HR/steps/sleep	Personalized anomaly detection	Prospective case series	Detection before symptoms in many cases	Early warning feasibility
Quer et al., 2021 (Nat Med)	Wearables + symptoms	Gradient boosting	Prospective app cohort	AUC ~0.80	Improved with symptom data
Zoabi et al., 2021 (Sci Rep)	Symptoms + demographics	Gradient boosting	National survey linked to PCR	AUC ~0.90	Population-scale screening

6. Explainability, Uncertainty, and Human Factors

Explainability tools (Grad-CAM, occlusion, SHAP/LIME) can localize salient regions or features, supporting clinician trust. Calibrated probabilities and uncertainty estimation (e.g., temperature scaling, MC-Dropout) are essential for safe triage decisions. Human-centered design—clear user interfaces, failure mode communication, and alignment with clinical pathways—improves adoption.

7. Common Pitfalls and Reporting Checklist

Table 5 lists frequent issues and mitigation strategies for COVID-19 detection models.

Pitfall	Why It Matters	Mitigation
Information leakage (image-level splits)	Inflated performance estimates	Patient-level or site-level splits
Confounders (site marks, devices)	Spurious learning	Harmonization, segmentation, adversarial debiasing
Imbalanced datasets	Biased PPV/NPV; instability	Stratified sampling, class-balanced losses, calibration
Poor ground truth	Noisy labels	Confirmatory testing, radiology consensus
No external validation	Poor generalization	Hold-out institutions, time-shift tests, prospective trials
Opaque models	Low trust and safety	XAI, uncertainty quantification, decision support not replacement

8. Regulatory, Privacy, and Ethics

Detection tools that inform clinical decision-making may fall under medical-device regulations. Developers should document intended use, performance on target populations, risk controls, and post-market surveillance. Privacy-preserving techniques such as federated learning and differential privacy can reduce data sharing risks. Equity assessments must examine performance across demographic and clinical subgroups to avoid exacerbating disparities.

9. Future Directions

Promising directions include: (i) larger, harmonized, and prospectively curated datasets; (ii) multimodal fusion of imaging, audio, and clinical context; (iii) self-supervised and foundation models adapted to limited labels; (iv) causal inference and counterfactual explanations for robustness; (v) prospective, randomized impact evaluations; and (vi) lifecycle monitoring with dataset and model cards.

10. Discussion

Across modalities, internal test performance often appears high, yet decreases under external validation, reflecting distribution shifts and confounding. Imaging models are strongest for identifying moderate-to-severe pneumonia but cannot replace virologic testing for early or asymptomatic infection. Audio and wearable approaches offer scalable screening but require careful personalization and context. Clinical/EHR models can aid pre-test risk estimation when combined with local prevalence and symptom profiles. A pragmatic approach is human-AI teaming: algorithms prioritize cases or flag anomalies, while clinicians integrate history, examination, and laboratory testing. Well-designed systems document uncertainty, provide interpretable cues, and undergo continuous quality improvement.

Abbreviations

AI: Artificial Intelligence; ML: Machine Learning; DL: Deep Learning; CNN: Convolutional Neural Network; RNN: Recurrent Neural Network; LSTM: Long Short-Term Memory; CXR: Chest X-ray; CT: Computed Tomography; AUC: Area Under the Receiver Operating Characteristic Curve; AUCPR: Area Under the Precision-Recall Curve; PPV: Positive Predictive Value; NPV: Negative Predictive Value; RT-PCR: Reverse Transcription Polymerase Chain Reaction; EHR: Electronic Health Record; XAI: Explainable AI; Grad-CAM: Gradient-weighted Class Activation Mapping.

References

1. Bai, H. X., et al. "Artificial Intelligence Augmentation of Radiologist Performance in Distinguishing COVID-19 from Pneumonia on Chest CT." Radiology, vol. 296, no. 3, 2020, pp. E156–65. 10.1148/radiol.2020201491.
2. Brown, C., J. Chauhan, A. Grammenos, et al. "Exploring Automatic Diagnosis of COVID-19 from Crowdsourced Respiratory Sound Data." IEEE Trans Audio Speech Lang Process., vol. 29, 2021, pp. 2010–23. 10.48550/arXiv.2006.05919
3. Cohen, J. P., P. Morrison, and L. Dao. "COVID-19 Image Data Collection." arXiv, 2020, arXiv:2003.11597.
4. de la Iglesia Vayá, M., et al. "BIMCV COVID-19+: A Large Annotated Dataset of Chest Radiograph Images and Clinical Data." medRxiv, 2020. 10.48550/arXiv.2006.01174
5. de Souza, J., et al. "Lung Ultrasound in COVID-19: Insights and Prospects for AI." Ultrasound in Medicine and Biology, vol. 47, no. 6, 2021, pp. 1465–80.
6. Harmon, S. A., et al. "Artificial Intelligence for the Detection of COVID-19 Pneumonia on Chest CT Using Multinational Datasets." Nat Commun, vol. 11, 2020, p. 4080. doi:10.1038/s41467-020-17971-2.
7. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep Residual Learning for Image Recognition.", 2016. 10.1109/CVPR.2016.90
8. Laguarta, Jordi, Ferran Hueto, and Brian Subirana. "COVID-19 Artificial Intelligence Diagnosis Using Only Cough Recordings." IEEE Open Journal of Engineering in Medicine and Biology, vol. 1, 2020, pp. 275–81. 10.1109/OJEMB.2020.3026928
9. Li, L., et al. "Artificial Intelligence Distinguishes COVID-19 from Community-Acquired Pneumonia on Chest CT." Radiology, vol. 296, no. 2, 2020, pp. E65–71. 10.1148/radiol.2020200905.
10. Lundberg, Scott M., and Su-In Lee. "A Unified Approach to Interpreting Model Predictions." Advances in Neural Information Processing Systems (NeurIPS), 2017. https://proceedings.neurips.cc/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf
11. Mishra, T., et al. "Pre-Symptomatic Detection of COVID-19 from Smartwatch Data." Nature Biomedical Engineering, vol. 4, 2020, pp. 1208–10. 10.1038/s41551-020-00640-6
12. Morozov, S. P., et al. "MosMedData: Chest CT Scans with COVID-19 Related Findings Dataset." medRxiv, 2020. 10.1101/2020.05.20.20100362
13. Ozturk, T., et al. "Automated Detection of COVID-19 Cases Using Deep Neural Networks with X-Ray Images." Computers in Biology and Medicine, vol. 121, 2020, p. 103792. 10.1016/j.compbiomed.2020.103792
14. Orlandic, L., T. Teijeiro, and D. Atienza. "The COUGHVID Crowdsourcing Dataset." Scientific Data, vol. 8, 2021, p. 156. 10.1038/s41597-021-00937-4

15. Quer, G., et al. “Wearable Sensor Data and Self-Reported Symptoms for COVID-19 Detection.” Nature Medicine, vol. 27, 2021, pp. 73–77. 10.1038/s41591-020-1123-x
16. Radin, J. M., et al. “Harnessing Wearable Device Data to Improve COVID-19 Detection.” Scientific Reports, vol. 10, 2020, p. 13773.
17. Rajpurkar, P., et al. “CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning.” arXiv, 2017, arXiv:1711.05225.
18. Roberts, M., et al. “Common Pitfalls and Recommendations for Using ML to Detect and Prognosticate for COVID-19 Using Chest Radiographs and CT Scans.” Nature Machine Intelligence, vol. 3, 2021, pp. 199–217. 10.1038/s42256-021-00307-0
19. RSNA/ACR. RICORD – The RSNA International COVID-19 Open Radiology Database, 2021. 10.1148/radiol.2021203957
20. Selvaraju, R. R., et al. “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization.” Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017. 10.1109/ICCV.2017.74
21. Shorten, C., and T. M. Khoshgoftaar. “A Survey on Image Data Augmentation for Deep Learning.” Journal of Big Data, vol. 6, 2019, p. 60. 10.1186/s40537-019-0197-0
22. UK Biobank. COVID-19 Imaging Data. (Consortium Resources).
23. Wang, L., Z. Q. Lin, and A. Wong. “COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images.” arXiv, 2020, arXiv:2003.09871.
24. Wynants, L., et al. “Prediction Models for Diagnosis and Prognosis of COVID-19: Systematic Review and Critical Appraisal.” BMJ, vol. 369, 2020, m1328. 10.1136/bmj.m1328
25. Yan, L., et al. “An Interpretable Mortality Prediction Model for COVID-19 Patients.” Nature Machine Intelligence, vol. 2, 2020, pp. 283–88. 10.1038/s42256-020-0180-7