XAI for image based COVID-19 detection

Weronika Hryniewska

Warsaw Univeristy of Technology

Do not repeat these mistakes -- a critical appraisal of applications of explainable artificial intelligence for image based COVID-19 detection

Weronika Hryniewska, Przemysław Bombiński, Patryk Szatkowski, Paulina Tomaszewska, Artur Przelaskowski, Przemysław Biecek

The sudden outbreak and uncontrolled spread of COVID-19 disease is one of the most important global problems today. In a short period of time, it has led to the development of many deep neural network models for COVID-19 detection with modules for explainability. In this work, we carry out a systematic analysis of various aspects of proposed models. Our analysis revealed numerous mistakes made at different stages of data acquisition, model development, and explanation construction. In this work, we overview the approaches proposed in the surveyed ML articles and indicate typical errors emerging from the lack of deep understanding of the radiography domain. We present the perspective of both: experts in the field - radiologists, and deep learning engineers dealing with model explanations. The final result is a proposed a checklist with the minimum conditions to be met by a reliable COVID-19 diagnostic model.

Subjects: **Image and Video Processing (eess.IV)**; Computer Vision and Pattern Recognition (cs.CV); Machine Learning (cs.LG)

Cite as: arXiv:2012.08333 **[eess.IV]**

Table of contents

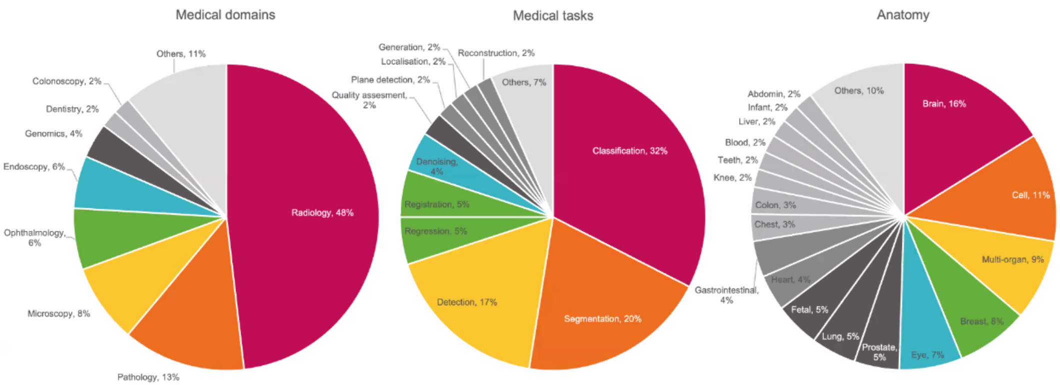


Figure: Artificial intelligence in modern medicine [1]

[1] https://ahmedhosny.github.io/aimindex/#

Methods

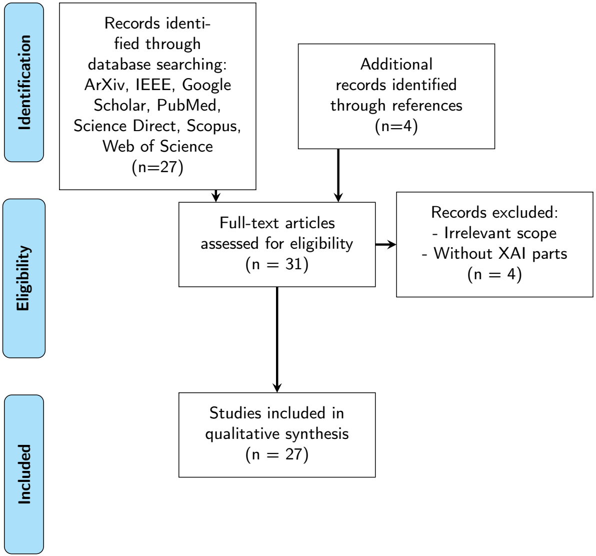


Figure: PRISMA Flow Diagram

("XAI covid-19"
OR "explainable artificial intelligence covid-19"
OR "explainable covid-19"
OR "explanations covid-19"
OR "interpretable covid-19"
OR "interpretations covid-19"
OR "transparent covid-19")

AND

("X-ray"
OR "CT"
OR "computed tomography"
OR "radiography")

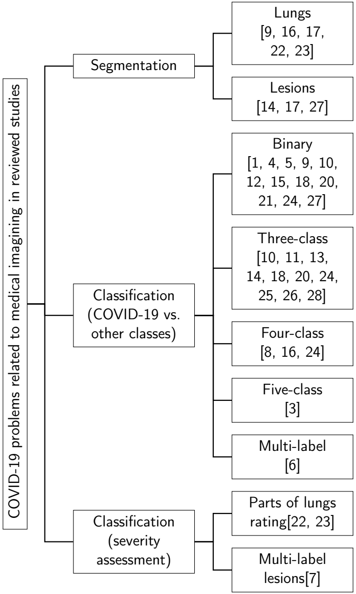


Figure: Categorisation of AI applications in 27 reviewed studies

Model training

Databases

Link to dataset	Number of COVID-19 cases	Used in paper
github.com/lindawangg/COVID-Net	917	[14, 20, 22, 26]
github.com/ieee8023/covid-chestxray-dataset	584	[3, 5, 6, 7, 8, 12, 13, 15, 16, 18, 20, 22, 24, 25, 28]
kaggle.com/tawsifurrahman/covid19-radiography-database	219	[13]
github.com/agchung/Actualmed-COVID-chestxray-dataset	58	[3]
github.com/agchung/Figure1-COVID-chestxray-dataset	56	[13, 20]
github.com/ari-dasci/OD-covidgr	426	[23]
github.com/UCSD-AI4H/COVID-CT	349	[1, 4, 12, 15]
github.com/muhammedtalo/COVID-19	125	[5, 13]
sirm.org/category/senza-categoria/covid-19	111	[21]
kaggle.com/nabeelsajid917/covid-19-x-ray-10000-images	70	[13]
kaggle.com/praveengovi/coronahack-chest-xraydataset	58	

Link to dataset	Used in article	Number of COVID-19 cases	Presence of children in database	Format
github.com/lindawangg/COVID-Net	[14, 20, 22, 26]			
github.com/ieee8023/covid-chestxray-dataset	[3, 5, 6, 7, 8, 12, 13, 15, 16, 18, 20, 22, 24, 25, 28]	584	no kids	JPG & PNG
github.com/agchung/Figure1-COVID-chestxray-dataset	[13, 20]	56	no kids	JPG & PNG
github.com/agchung/Actualmed-COVID-chestxray-dataset		58	no info	PNG
kaggle.com/tawsifurrahman/covid19-radiography-database	[13]	219	no info	PNG
kaggle.com/c/rsna-pneumonia-detection-challenge	[7, 13, 20, 22]	0	present <10	DICOM
kaggle.com/paultimothymooney/chest-xray-pneumonia	[6, 8, 16, 24, 28]	0	no info	JPG
github.com/UCSD-AI4H/COVID-CT	[1, 4, 12, 15]	349	10 cases	JPG & PNG
nihcc.app.box.com/v/ChestXray-NIHCC	[7, 18, 20]	0	present <10	PNG
github.com/muhammedtalo/COVID-19	[5, 13]	125	no info	JPG & PNG
openi.nlm.nih.gov/	[7, 16]	0	no info	DICOM & PNG
stanfordmlgroup.github.io/competitions/chexpert	[7, 22]	0	no kids	JPG
bimcv.cipf.es/bimcv-projects/padchest/	[7]	0	1535 cases <10	PNG
data.mendeley.com/datasets/rscbjbr9sj/2	[25]	0	no info	JPG
github.com/ari-dasci/OD-covidgr	[23]	426	no info	JPG
kaggle.com/nabeelsajid917/covid-19-x-ray-10000-images	[13]	70	present <10	JPG & PNG
kaggle.com/nih-chest-xrays/sample	[5]	0	present <10	PNG
kaggle.com/praveengovi/coronahack-chest-xraydataset	[3]	58	no info	JPG & PNG
physionet.org/content/mimic-cxr/2.0.0	[7]	0	no info	DICOM
radiopaedia.org	[21]	0	present <10	JPG
sirm.org/category/senza-categoria/covid-19	[21]	111	no info	JPG & PNG
wiki.cancerimagingarchive.net/display/Public/LIDC-IDRI	[9]	0	no info	DICOM
offline database or from hospital	[9, 10, 17, 27]			
lack of information	[11, 19]			

DICOM

For medical imaging, standard extension is Digital Imaging and Communications in Medicine (DICOM).

Pixel intensity for DICOM is between:
$2^{12} = 4096$ and $2^{16} = 65536$

Pixel intensity for JPG and PNG:
$2^8 = 256$

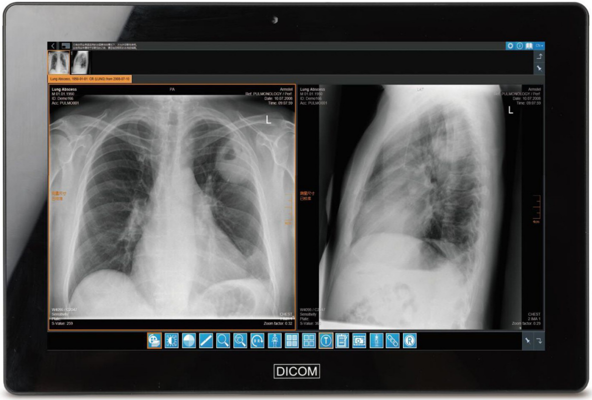


Figure: Medical display[2]

[2] www.medicalexpo.com/prod/estone-group/product-119365-950540.html

Children are not small adults

Figure: Baby scan tube [3]

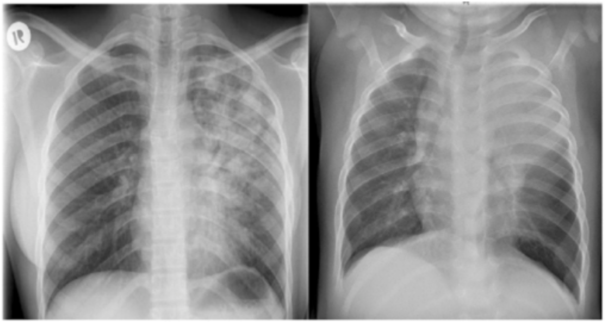
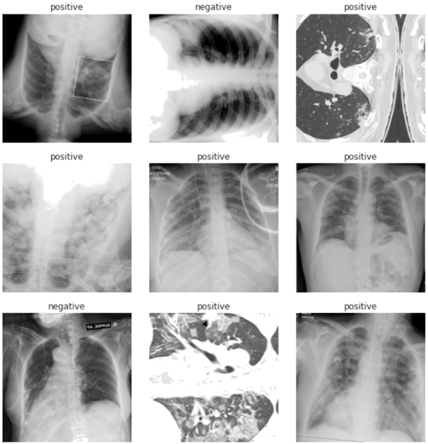


Figure: Chest X-ray image of an adult and a child with tuberculosis [4]

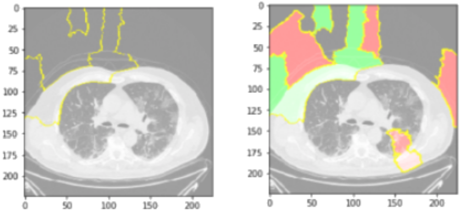
[3] www.dailystar.co.uk/news/latest-news/confusing-picture-baby-stuck-tube-17367514

[4] https://bmcinfectdis.biomedcentral.com/articles/10.1186/s12879-020-05381-0

No mixing CT with X-ray



COVID-19 Chest X-ray Sample



Model Interpretability - COVID-19 Chest Radiographs Dataset

Figure: CT images are mixed with X-rays [12]

Lung windows

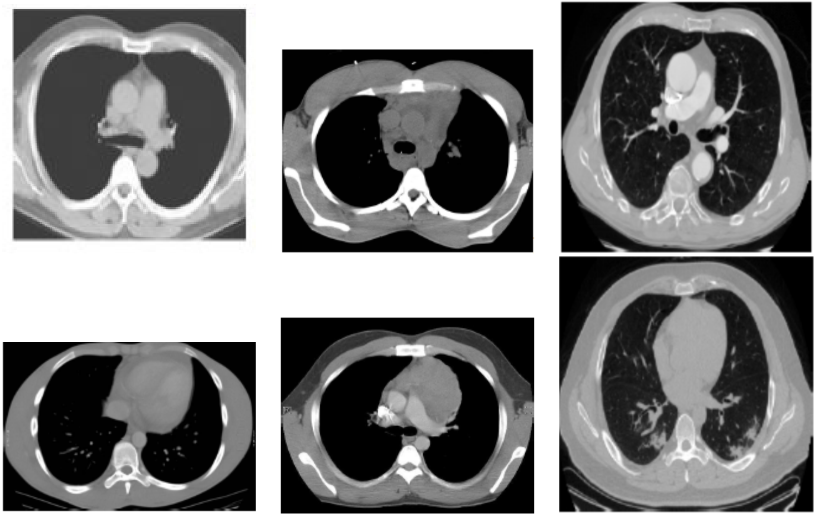
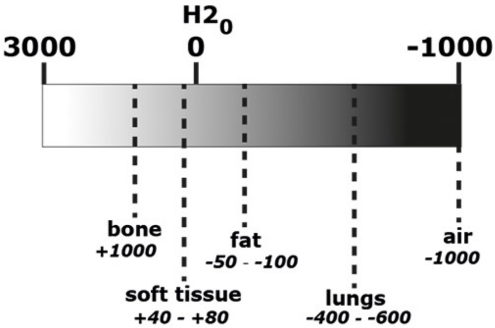
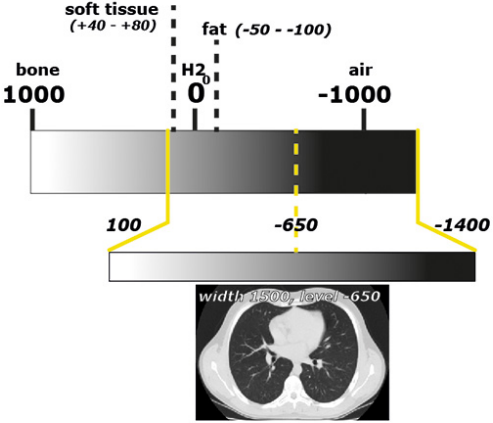


Figure: These are not lung windows, but tissue windows. The lungs cannot be examined. [1, 4, 10]

Hounsfield Units

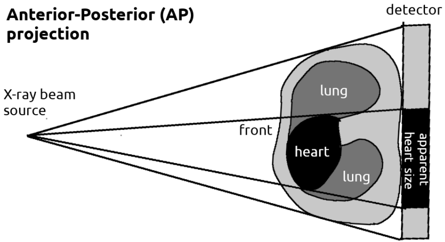


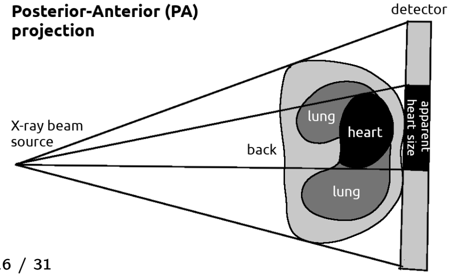


[5]

[5] www.startradiology.com/the-basics/x-rayct-technique/index.html

PA vs AP projection





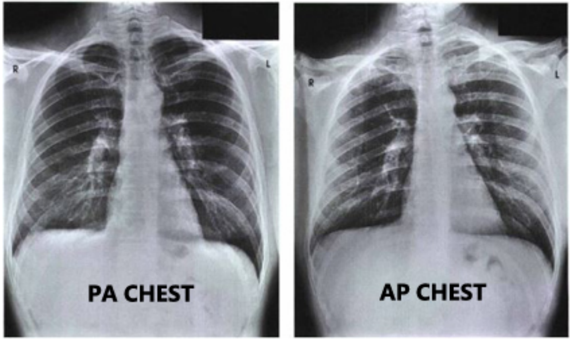


Figure: PA and AP chest radiograph comparison[a]

[a] www.radtechonduty.com/2015/02/chest-xray

Image preprocessing

- cropping, changing colour space, proportionally resizing, or zooming
- normalizing pixel intensity or equalizing histogram

Preprocessing technique	Reference
Resize to the same size	[3, 4, 6, 7, 8, 12, 13, 15, 16, 17, 21, 24, 25, 28]
Normalize pixel intensity	[6, 7, 8, 13, 21, 22, 24]
Change color space	[12, 13, 21, 25]
Eliminate noise	[13, 15, 22, 23]
Use Perona-Malik filter	[13]
Limit image intensity	[6, 22]
Equalize histogram	[3, 13, 16, 22]
Perform image enhancement	[13, 15, 16, 17]
Crop image	[3, 6, 23]
Cast data type	[16, 22]
Zoom image / augmentation	[21, 4]
Add pixels	[23]
Feature encoding	[28]
Feature extraction	[14, 28]
Rotate image	[4]
Use 2D wavelet transform	[15]
Lack or lack of description	[1, 5, 9, 10, 11, 14, 18, 19, 20, 26, 27]

Lung segmentation

- without a distorted lung border
- without liver, bowel, etc.

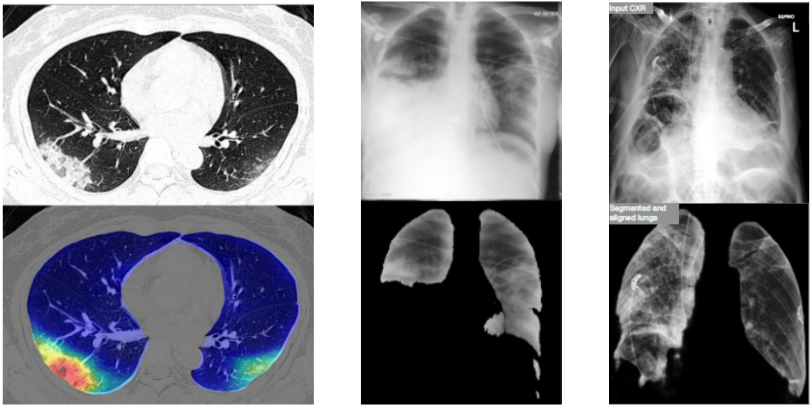


Figure: Examples of lung segmentation [9, 16, 22]

Data augmentation

- rotation
- cropping
- proportionally scalling/zooming up
- brightness, and contrast

Data augmentation technique	Values and studies
Affine transformations	[10]:
Rotation	[9, 20, 24, 26], 5°[3], 15°[5, 13, 21], 20°[8], 25°[22], 5-35°[11]
Scaling / Zooming	[8, 20, 26], 10% [3, 12, 22], 20% [24]
Horizontally, vertically	20% [11]
Flip	[8, 12, 14]
Horizontal	[3, 9, 20, 26, 27]
Vertical	
Shifting / Translation	[8, 20, 26], height 5% [3], 10% [22]
Shearing	[8, 24, 25]
Crop	[9, 27]
Color jitter (brightness, contrast, saturation, and hue)	[10]
Brightness change	[20, 26] +/-30 [25], 10% [3]
Gaussian noise	[25]
ZCA whitening transformation	[8]
Elastic transformation	α=60, σ=12 [22]
Grid distortion	steps=5, limit=0.3 [22]
Optical distortion	distort=0.2, shift=0.05 [22]
Warping	10% [12]
Multiple patches from each image	[16]
Class-inherent transformations Network [a]	[23]
Lack	[1, 6, 7, 15, 17, 18, 19, 28]
Not specified	[4]

[a] inspired by Generative Adversarial Networks

Model architecture

Classification

Model architecture	Reference
ResNet	[13]
ResNet18	[6, 14, 16, 22]
ResNet34	[6, 12, 17]
ResNet50	[1, 3, 9, 11, 12, 15]
DenseNet	[7, 13, 20, 22]
DenseNet121	[3, 21]
DenseNet-161	[6]
DenseNet-201	[1, 12, 21]
VGG	[13, 22]
VGG-16	[1, 5, 21, 28]
VGG-19	[1, 12, 21]
Inception	[22]
InceptionV3	[6, 11, 24]
InceptionResNetV2	[1, 21]
MobileNetV2	[1, 21]
NasNetMobile	[1, 21]
EfficientNet-B0	[12]
Efficient TBCNN	[3]
MobileNet	[21]
NasNetLarge	[21]
Res2Net	[27]
Residual Attention Net	[21]
ResNet15V2	[1]
ResNet50V2	[8]
ResNeXt	[6]
WideResNet	[6]
Xception	[21]
own model	[4, 10, 12, 18, 22, 23, 25, 26]

Simple classifier or ensembles	Reference
AdaBoost	[4]
Deep tree	[11]
KNN	[4]
Naive Bayes	[4]
SVM	[4]

Segmentation

Model architecture	Reference
U-net	[9, 22, 23]
AutoEncoder	[14]
VGG-16 backbone + enhanced feature module	[27]
(FC)-DenseNet-103	[16]
Nested version of Unet (Unet++)	[22]
VB-Net	[17]

Transfer learning

Transfer learning on the dataset consists of chest medical imagining (not on ImageNet [13])

Figure: Examples of images in ImageNet database[6]

X-ray with DICOM lung images:
7 470 images[7], 29 684 images[8] , 3 209 images[9], 377 110 images[10].

CT with DICOM lung images:
15 419 images[11], 260 826 images[12], 21 082 502 images[13].

[6] https://i0.wp.com/syncedreview.com/wp-content/uploads/2020/06/Imagenet.jpg
[7] pubmed.ncbi.nlm.nih.gov/25525580
[8] kaggle.com/c/rsna-pneumonia-detection-challenge
[9] kaggle.com/c/siim-acr-pneumothorax-segmentation
[10] physionet.org/content/mimic-cxr/2.0.0
[11] wiki.cancerimagingarchive.net/display/Public/RIDER+Lung+CT
[12] wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=70224216
[13] wiki.cancerimagingarchive.net/display/NLST/National+Lung+Screening+Trial

Validation

In external validation type called cross-database validation the models are trained on one database (source database) and tested on another database (target database).

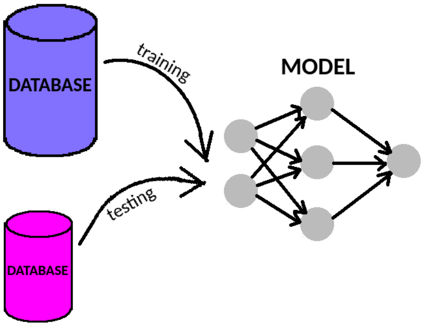


Figure: Cross-database validation

Explainable artificial intelligence

XAI methods used in considered studies

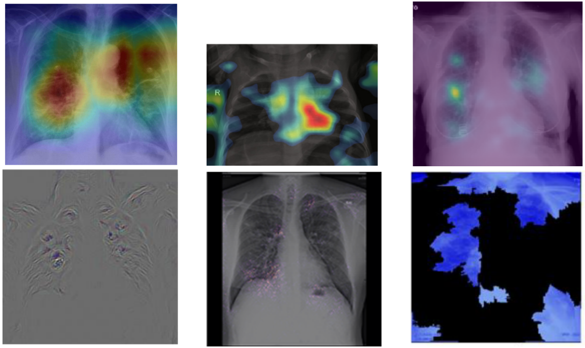


Figure: Examples of XAI visualizations from studies: [23, 25, 7, 8, 6, 1]. Following explanations are used (in a reading direction): Grad-CAM, CAM, saliency, guided backpropagation, integrated gradients, LIME. Such explanations can be divided into 4 types: heat maps (image from 1 to 3), contour lines (image 4), points (image 5), and image pieces (image 6).

- Grad-CAM (gradient-weighted class activation mapping)
- LIME (local interpretable model-agnostic explanations)
- CAM (class activation mapping)
- Saliency (saliency map)
- Guided Backpropagation
- LRP (layer-wise relevance propagation)
- Occlusion (occlusion sensitivity)
- AM (activation mapping)
- Attribution maps
- DeepLIFT
- Feature maps
- Grad-CAM++
- Guided Grad-CAM
- GSInquire
- Input X Gradient
- Integrated Gradients

Expert based evaluation of XAI methods

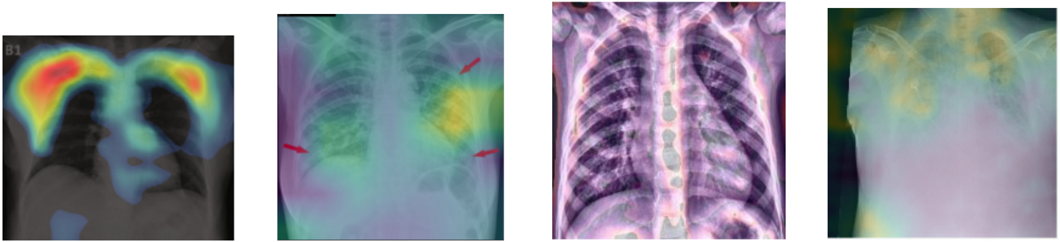


Figure: Examples of biased model explanations [25, 5, 3, 14] Red arrows in the second photo are marked by a radiologist to help localize the lesions. They were not present during model training.

The checklist for responsible analysis of lung images with deep learning models

- Data resources
  - D Does the data format provide diagnostic quality? (DICOM is recommended)
  - R Are the low quality images (i.e., blurred, too dark, or too bright) rejected?
  - D Is the dataset balanced in terms of sex and age?
  - R Does the dataset contain one type of images (CT or X-ray)?
  - R Are the lung structures visible ("lung" window) on CT images?
  - D Are children and adult images not mixed within the dataset?
  - R Are images correctly categorized in relation to class of pathology?
  - D Are AP/PA projections described for every X-ray image?
- Image preprocessing
  - D Is the data preprocessing described?
  - D Are artefacts (like captions) removed?
- Data augmentation (if needed)
  - D Are lungs fully present after transformations?
  - R Are lung structures visible after brightness or contrast transformations?
- Transfer learning (if used)
  - D Is the transfer learning procedure described?
- Model performance
  - D At least a few metrics out of the proposed in [2] are used?
  - D Is the model validated on a different database than used for training?
- Domain quality of model explanations
  - R Are other structures (i.e., bowel loops) not interpreted as lungs?
  - R Are areas marked as explanation inside the chest?
  - R Are artifacts (cables, breathing tubes, image compression, embedded markup symbols) not identified as part of the explanations?
  - R Are areas pointed out as explanations consistent with opinions of radiologists?
  - R Are explanations sharply indicate lesions?

Bibliography I

[1] Ahsan, M. M., Gupta, K. D., Islam, M. M., Sen, S., Rahman, M. L., and Hossain, M. S.
Study of Different Deep Learning Approach with Explainable AI for Screening Patients with COVID-19 Symptoms: Using CT Scan and Chest X-ray Image Dataset.

[2] Albahri, O. S., Zaidan, A. A., Albahri, A. S., Zaidan, B. B., Abdulkareem, K. H., Alamoodi, A. H., Aleesa, A. M., Chyad, M. A., Alesa, R. M., Kim, L. C., Lakulu, M. M., Ibrahim, A. B., and Rashid, N. A.
Systematic Review of Artificial Intelligence Techniques in the Detection and Classification of COVID-19 Medical Images in Terms of Evaluation and Benchmarking: Taxonomy Analysis, Challenges, Future Solutions and Methodological Aspects.
Journal of Infection and Public Health, June (2020).

[3] Albert, N.
Evaluation of Contemporary Convolutional Neural Network Architectures for Detecting COVID-19 from Chest Radiographs.
arXiv.org (2020).

[4] Angelov, P. P., and Soares, E. A.
Explainable-By-Design Approach For Covid-19 Classification Via CT-Scan.

[5] Brunese, L., Mercaldo, F., Reginelli, A., and Santone, A.
Explainable Deep Learning for Pulmonary Disease and Coronavirus COVID-19 Detection from X-rays.
Computer Methods and Programs in Biomedicine 196 (2020), 105608.

[6] Chatterjee, S., Saad, F., Sarasaen, C., Ghosh, S., Khatun, R., Radeva, P., Rose, G., Stober, S., Speck, O., and Nürnberger, A.
Exploration of Interpretability Techniques for Deep COVID-19 Classification using Chest X-ray Images.

[7] Cohen, J. P., Dao, L., Morrison, P., Roth, K., Bengio, Y., Shen, B., Abbasi, A., Hoshmand-Kochi, M., Ghassemi, M., Li, H., and Duong, T. Q.
Predicting COVID-19 Pneumonia Severity on Chest X-ray with Deep Learning.

[8] Ghoshal, B., and Tucker, A.
Estimating Uncertainty and Interpretability in Deep Learning for Coronavirus (COVID-19) Detection.
1–14.

Bibliography II

[9] Gozes, O., Frid, M., Greenspan, H., and Patrick, D.
Rapid AI Development Cycle for the Coronavirus (COVID-19) Pandemic : Initial Results for Automated Detection & Patient Monitoring using Deep Learning CT Image Analysis.
arXiv.org (2020).

[10] Han, Z., Wei, B., Hong, Y., Li, T., Cong, J., Zhu, X., Wei, H., and Zhang, W.
Accurate Screening of COVID-19 using Attention Based Deep 3D Multiple Instance Learning.
IEEE Transactions on Medical Imaging XX, XX (2020), 1–1.

[11] Hossain, M. S., Muhammad, G., and Guizani, N.
Explainable AI and Mass Surveillance System-Based Healthcare Framework to Combat COVID-I9 Like Pandemics.
IEEE Network 34, 4 (2020), 126–132.

[12] Jaiswal, A. K., Tiwari, P., Rathi, V. K., Qian, J., Pandey, H. M., and Albuquerque, V. H. C.
COVIDPEN: A Novel COVID-19 Detection Model using Chest X-Rays and CT Scans.
medRxiv (2020), 2020.07.08.20149161.

[13] Karim, M. R., Döhmen, T., Rebholz-Schuhmann, D., Decker, S., Cochez, M., and Beyan, O.
DeepCOVIDExplainer: Explainable COVID-19 Diagnosis Based on Chest X-ray Images.

[14] Khobahi, S., Agarwal, C., and Soltanalian, M.
CoroNet: A Deep Network Architecture for Semi-Supervised Task-Based Identification of COVID-19 from Chest X-ray Images.
medRxiv (2020), 2020.04.14.20065722.

[15] Matsuyama, E.
A Deep Learning Interpretable Model for Novel Coronavirus Disease (COVID-19) Screening with Chest CT Images.
Journal of Biomedical Science and Engineering 13, 7 (2020), 140–152.

Bibliography III

[16] Oh, Y., Park, S., and Ye, J. C.
Deep Learning COVID-19 Features on CXR using Limited Training Data Sets.
IEEE Transactions on Medical Imaging 0062, c (2020), 1–1.

[17] Ouyang, X., Huo, J., Xia, L., Shan, F., Liu, J., Mo, Z., Yan, F., Ding, Z., Yang, Q., Song, B., Shi, F., Yuan, H., Wei, Y., Cao, X., Gao, Y., Wu, D., Wang, Q., and Shen, D.
Dual-Sampling Attention Network for Diagnosis of COVID-19 from Community Acquired Pneumonia.
IEEE Transactions on Medical Imaging 39, XX (2020), 1–1.

[18] Ozturk, T., Talo, M., Yildirim, E. A., Baloglu, U. B., Yildirim, O., and Rajendra Acharya, U.
Automated detection of COVID-19 cases using deep neural networks with X-ray images.
Computers in Biology and Medicine 121, April (2020), 103792.

[19] Rahman, M. A., Hossain, M. S., Alrajeh, N. A., and Guizani, N.
B5G and Explainable Deep Learning Assisted Healthcare Vertical at the Edge: COVID-I9 Perspective.
IEEE Network 34, 4 (2020), 98–105.

[20] Sarker, L., Islam, M., Hannan, T., Zakaria, A., Ahmed, Z., and Zakaria, A.
COVID-DenseNet: A Deep Learning Architecture to Detect COVID-19 from Chest Radiology Images.
Preprints, May (2020).

[21] Sharma, V., and Dyreson, C.
COVID-19 detection using Residual Attention Network an Artificial Intelligence approach.

[22] Signoroni, A., Savardi, M., Benini, S., Adami, N., Leonardi, R., Gibellini, P., Vaccher, F., Ravanelli, M., Borghesi, A., Maroldi, R., and Farina, D.
End-to-end learning for semiquantitative rating of COVID-19 severity on Chest X-rays.
1–22.

Bibliography IV

[23] Tabik, S., Gómez-Ríos, A., Martín-Rodríguez, J. L., Sevillano-García, I., Rey-Area, M., Charte, D., Guirado, E., Suárez, J. L., Luengo, J., Valero-González, M. A., García-Villanova, P., Olmedo-Sánchez, E., and Herrera, F.
COVIDGR dataset and COVID-SDNet methodology for predicting COVID-19 based on Chest X-Ray images.

[24] Tsiknakis, N., Trivizakis, E., Vassalou, E., Papadakis, G., Spandidos, D., Tsatsakis, A., Sánchez-García, J., López-González, R., Papanikolaou, N., Karantanas, A., and Marias, K.
Interpretable artificial intelligence framework for COVID-19 screening on chest X-rays.
Experimental and Therapeutic Medicine (2020), 727–735.

[25] Ucar, F., and Korkmaz, D.
COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images.
Medical Hypotheses 140, April (2020), 109761.

[26] Wang, L., and Wong, A.
COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images.
1–12.

[27] Wu, Y.-H., Gao, S.-H., Mei, J., Xu, J., Fan, D.-P., Zhao, C.-W., and Cheng, M.-M.
JCS: An Explainable COVID-19 Diagnosis System by Joint Classification and Segmentation.
1–11.

[28] Zokaeinikoo, M., Mitra, P., Kumara, S., and Kazemian, P.
AIDCOV: An Interpretable Artificial Intelligence Model for Detection of COVID-19 from Chest Radiography Images.
medRxiv (2020), 2020.05.24.20111922.

Thank you for your attention!