AI for radiographic COVID-19 detection selects shortcuts over signal: Online Figures

Alex J. DeGrave[1,2,3], Joseph D. Janizek[1,2,3], and Su-In Lee[1]

[1]Paul G. Allen School of Computer Science and Engineering, University of Washington
[2]Medical Scientist Training Program, University of Washington
[3]These authors contributed equally and are ordered alphabetically: Alex J. DeGrave, Joseph D. Janizek

References

1. Bustos, A., Pertusa, A., Salinas, J.-M. & de la Iglesia-Vayá, M. Padchest: A large chest x-ray image dataset with multi-label annotated reports. *Medical Image Analysis* **66,** 101797 (2020).
2. Vayá, M. d. l. I. *et al.* BIMCV COVID-19+: A large annotated dataset of RX and CT images from COVID-19 patients. *arXiv:2006.01174* (2020).
3. Wang, X. *et al. ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases* in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017), 2097–2106.
4. Cohen, J. P., Morrison, P. & Dao, L. COVID-19 image data collection. *arXiv 2003.11597.* `https://github.com/ieee8023/covid-chestxray-dataset` (2020).

Online Supplementary Fig. 7 | Saliency maps for 15 radiographs from each of the four data sources. Across all four data sources, saliency maps highlight text tokens and laterality markers (e.g., the first radiograph-saliency map pair in the first row of the PadChest[1] examples, the second-to-last and last radiograph-saliency map pairs in the third row of the PadChest examples, the first four radiograph-saliency map pairs in the second row of the BIMCV-COVID-19+[2] examples, all five radiograph-saliency map pairs in the third row of the ChestX-ray14[3] examples, and the second-to-last radiograph-saliency map pair in the last row of the GitHub-COVID[4] examples). Note that the GitHub-COVID images inncluded were selected from the subset that was suitable for reproduction (*i.e.*, licensed as CC BY, CC BY-SA, or CC BY-NC-SA) and therefore underrepresents images that originate from academic publications, which may be more likely to feature annotations.

Online Supplementary Fig. 7 | (continued) The online version of this figure is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 (CC BY-NC-SA 4.0) license. This figure is composed from pairs of images: the left image of each pair is an original radiograph converted to grayscale, but otherwise unaltered; the right image is the corresponding expected gradients saliency map. Below we list the attributions for the creators of the original images in the Github-COVID section, which were compiled from a variety of online sources as part of the "GitHub-COVID" dataset. Image numbering is left to right, top to bottom.

1. Robert Vollmann. CC BY-NC-SA 4.0. https://www.eurorad.org/case/16747
2. Hinrich B. Winther, Hans Laser, Svetlana Gerbel, Sabine K. Maschke, Jan B. Hinrichs, Jens Vogel-Claussen, Frank K. Wacker, Marius M. Höper, Bernhard C. Meyer. CC BY 4.0. doi.org/10.6084/m9.figshare.12275009
3. Domenico Nicoletti. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-35?lang=us
4. Julia Tran, Justin Glavis-Bloom, Theodore Bryan, Karen Tran Harding, Chantal Chahine, Roozbeh Houshyar. CC BY-NC-SA 4.0. https://www.eurorad.org/case/16654
5. Hinrich B. Winther, Hans Laser, Svetlana Gerbel, Sabine K. Maschke, Jan B. Hinrichs, Jens Vogel-Claussen, Frank K. Wacker, Marius M. Höper, Bernhard C. Meyer. CC BY 4.0. doi.org/10.6084/m9.figshare.12275009
6. Ian Bickle. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-progression-and-regression
7. Roma Patel. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-42
8. Muhammad Imran Khan. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/post-intubation-pneumomediastinum-and-pneumothorax-background-covid-19-pneumonia?lang=us
9. Edgar Lorente. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-rapidly-progressive
10. Domenico Nicoletti. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-38
11. Roma Patel. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-41
12. Julia Tran, Justin Glavis-Bloom, Theodore Bryan, Karen Tran Harding, Chantal Chahine, Roozbeh Houshyar. CC BY-NC-SA 4.0. https://www.eurorad.org/case/16654
13. Soon Ho Yoon, Kyung Hee Lee, Jin Yong Kim, Young Kyung Lee, Hongseok Ko, Ki Hwan Kim, Chang Min Park, Yun-Hyeon Kim. CC BY-NC 4.0. https://www.kjronline.org/DOIx.php?id=10.3348/kjr.2020.0132
14. Ian Bickle. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-progression-and-regression
15. Alexander Gross, Martin Schwarz, Thomas Albrecht. CC BY-NC-SA 4.0. https://www.eurorad.org/case/16663

Online Supplementary Fig. 8 | Examples images generated by a CycleGAN that was trained to alter COVID-19 negative images from the ChestX-ray14 dataset[3] to appear like COVID-19 positive images from the GitHub-COVID[4] dataset and vice versa. Red boxes in lower panel mark annotations that were removed by the CycleGAN, as referenced in the main text. Images from the GitHub-COVID repository may contain annotations, as many were scraped from figures in academic publications. Note that the COVID-19+ → COVID-19− images were selected from the subset of the GitHub-COVID dataset that was suitable for reproduction (*i.e.*, licensed as CC BY, CC BY-SA, or CC BY-NC-SA) and therefore underrepresents images that originate from academic publications, which may be more likely to feature annotations.

Online Supplementary Fig. 8 | (continued) The online version of this figure is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 (CC BY-NC-SA 4.0) license. This figure is composed from pairs of images: the left image of each pair is cropped from the original and converted to grayscale, but otherwise unaltered; the right image was transformed by a CycleGAN. Below we list the attributions for the creators of the original images in the COVID-19+ → COVID-19− section, which were compiled from a variety of online sources as part of the "GitHub-COVID" dataset; the COVID-19− → COVID-19+ images originate from the NIH ChestX-ray14 repository. Image numbering is left to right, top to bottom.

1. Robert Vollmann. CC BY-NC-SA 4.0. https://www.eurorad.org/case/16747
2. Hinrich B. Winther, Hans Laser, Svetlana Gerbel, Sabine K. Maschke, Jan B. Hinrichs, Jens Vogel-Claussen, Frank K. Wacker, Marius M. Höper, Bernhard C. Meyer. CC BY 4.0. doi.org/10.6084/m9.figshare.12275009
3. Domenico Nicoletti. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-35?lang=us
4. Julia Tran, Justin Glavis-Bloom, Theodore Bryan, Karen Tran Harding, Chantal Chahine, Roozbeh Houshyar. CC BY-NC-SA 4.0. https://www.eurorad.org/case/16654
5. Hinrich B. Winther, Hans Laser, Svetlana Gerbel, Sabine K. Maschke, Jan B. Hinrichs, Jens Vogel-Claussen, Frank K. Wacker, Marius M. Höper, Bernhard C. Meyer. CC BY 4.0. doi.org/10.6084/m9.figshare.12275009
6. Ian Bickle. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-progression-and-regression
7. Roma Patel. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-42
8. Muhammad Imran Khan. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/post-intubation-pneumomediastinum-and-pneumothorax-background-covid-19-pneumonia?lang=us
9. Edgar Lorente. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-rapidly-progressive
10. Domenico Nicoletti. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-38
11. Roma Patel. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-41
12. Julia Tran, Justin Glavis-Bloom, Theodore Bryan, Karen Tran Harding, Chantal Chahine, Roozbeh Houshyar. CC BY-NC-SA 4.0. https://www.eurorad.org/case/16654
13. Soon Ho Yoon, Kyung Hee Lee, Jin Yong Kim, Young Kyung Lee, Hongseok Ko, Ki Hwan Kim, Chang Min Park, Yun-Hyeon Kim. CC BY-NC 4.0. https://www.kjronline.org/DOIx.php?id=10.3348/kjr.2020.0132
14. Ian Bickle. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-progression-and-regression
15. Alexander Gross, Martin Schwarz, Thomas Albrecht. CC BY-NC-SA 4.0. https://www.eurorad.org/case/16663
16. Tiffany Chao, Justin Glavis-Bloom, Louis Fanucci, Theodore Bryan, Roozbeh Houshyar. CC BY-NC-SA 4.0. https://www.eurorad.org/case/16669
17. Edna M. Edrada, Edmundo B. Lopez, Jose Benito Villarama, Eumelia P. Salva Villarama, Bren F. Dagoc, Chris Smith, Ana Ria Sayo, Jeffrey A. Verona, Jamie Trifalgar-Arches, Jezreel Lazaro, Ellen Grace M. Balinas, Elizabeth Freda O. Telan, Lynsil Roy, Myvie Galon, Carl Hill N. Florida, Tatsuya Ukawa, Annavi Marie G. Villanueva, Nobuo Saito, Jean Raphael Nepomuceno, Koya Ariyoshi, Celia Carlos, Amalea Dulcene Nicolasora, Rontgene M. Solante. "First COVID-19 infections in the Philippines: a case report" *Tropical Medicine and Health* **48**(21), 2020. CC BY 4.0. https://doi.org/10.1186/s41182-020-00203-0
18. Mohammad Al-Tibi. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-49
19. Edgar Lorente. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-28
20. Aveek Mukherjee, Mudassar Ahmad, Douglas Frenia. "A coronavirus disease 2019 (COVID-19) patient with multifocal pneumonia treated with hydroxychloroquine". *Cureus* **12**(3): e7473. CC BY 4.0. doi.org/10.7759/cureus.7473
21. Hinrich B. Winther, Hans Laser, Svetlana Gerbel, Sabine K. Maschke, Jan B. Hinrichs, Jens Vogel-Claussen, Frank K. Wacker, Marius M. Höper, Bernhard C. Meyer. CC BY 4.0. doi.org/10.6084/m9.figshare.12275009
22. Edgar Lorente. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-rapidly-progressive-acute-respiratory-distress-syndrome-ards
23. Hinrich B. Winther, Hans Laser, Svetlana Gerbel, Sabine K. Maschke, Jan B. Hinrichs, Jens Vogel-Claussen, Frank K. Wacker, Marius M. Höper, Bernhard C. Meyer. CC BY 4.0. doi.org/10.6084/m9.figshare.12275009
24. Christina Aslanidi, Sofia Athanasiou, Eleni Lazaridou, Anastasia Kotanidou, Demetrios Exarhos. CC BY-NC-SA 4.0. https://www.eurorad.org/case/16744
25. Anton Aubanell Creus. CC BY-NC-SA 3.0. https://radiopaedia.org/cases/covid-19-pneumonia-22
26. Edna M. Edrada, Edmundo B. Lopez, Jose Benito Villarama, Eumelia P. Salva Villarama, Bren F. Dagoc, Chris Smith, Ana Ria Sayo, Jeffrey A. Verona, Jamie Trifalgar-Arches, Jezreel Lazaro, Ellen Grace M. Balinas, Elizabeth Freda O. Telan, Lynsil Roy, Myvie Galon, Carl Hill N. Florida, Tatsuya Ukawa, Annavi Marie G. Villanueva, Nobuo Saito, Jean Raphael Nepomuceno, Koya Ariyoshi, Celia Carlos, Amalea Dulcene Nicolasora, Rontgene M. Solante. "First COVID-19 infections in the Philippines: a case report" *Tropical Medicine and Health* **48**(21), 2020. CC BY 4.0. https://doi.org/10.1186/s41182-020-00203-0
27. Hinrich B. Winther, Hans Laser, Svetlana Gerbel, Sabine K. Maschke, Jan B. Hinrichs, Jens Vogel-Claussen, Frank K. Wacker, Marius M. Höper, Bernhard C. Meyer. CC BY 4.0. doi.org/10.6084/m9.figshare.12275009
28. Hinrich B. Winther, Hans Laser, Svetlana Gerbel, Sabine K. Maschke, Jan B. Hinrichs, Jens Vogel-Claussen, Frank K. Wacker, Marius M. Höper, Bernhard C. Meyer. CC BY 4.0. doi.org/10.6084/m9.figshare.12275009
29. Hinrich B. Winther, Hans Laser, Svetlana Gerbel, Sabine K. Maschke, Jan B. Hinrichs, Jens Vogel-Claussen, Frank K. Wacker, Marius M. Höper, Bernhard C. Meyer. CC BY 4.0. doi.org/10.6084/m9.figshare.12275009
30. Bruno Giesteira, Joana Pinto, Jessica Sousa, Elisa Melo Abreu. CC BY-NC-SA 4.0. https://www.eurorad.org/case/16706