



Open-Source Biomedical Image Analysis Models: A Meta-Analysis and Continuous Survey

Rui Li^{1†}, Vaibhav Sharma^{1†}, Subasini Thangamani^{1†} and Artur Yakimovich^{1,2,3,4*}

¹Center for Advanced Systems Understanding (CASUS), Helmholtz-Zentrum Dresden-Rossendorf e. V. (HZDR), Görlitz, Germany, ²Bladder Infection and Immunity Group (BIIG), Department of Renal Medicine, Division of Medicine, University College London, Royal Free Hospital Campus, London, United Kingdom, ³Artificial Intelligence for Life Sciences CIC, Dorset, United Kingdom, ⁴Roche Pharma International Informatics, Roche Diagnostics GmbH, Mannheim, Germany

Open-source research software has proven indispensable in modern biomedical image analysis. A multitude of open-source platforms drive image analysis pipelines and help disseminate novel analytical approaches and algorithms. Recent advances in machine learning allow for unprecedented improvement in these approaches. However, these novel algorithms come with new requirements in order to remain open source. To understand how these requirements are met, we have collected 50 biomedical image analysis models and performed a meta-analysis of their respective papers, source code, dataset, and trained model parameters. We concluded that while there are many positive trends in openness, only a fraction of all publications makes all necessary elements available to the research community.

Keywords: machine learning, deep learning, open source, bioimaging, image analysis, medical imaging

INTRODUCTION

The source code of data analysis algorithms made freely available for possible redistribution and modification (i.e. open source) has been beyond any doubt driving the ongoing revolution in Data Science (DS), Machine Learning (ML), and Artificial Intelligence (AI) (Sonnenburg et al., 2007; Landset et al., 2015; Abadi et al., 2016; Paszke et al., 2019). Encouraging open collaboration, the open-source model of code redistribution allows researchers to build upon their peers' work on a global scale fueling the rapid iterative improvement in the respective fields (Sonnenburg et al., 2007). Conversely, "closed-source" publications not only hamper the development of the field but also make it hard for the researchers to reproduce the results disseminated in the research articles. While *de jure* all published work resides in the public domain, reverse engineering of an advanced algorithm implementation may often take weeks or months, making such works hard to reproduce.

Needless to say, open source comes in a great variety of shapes and kinds. Remarkably, just making the source code of your research software available publicly or upon request does not *per se* make it open source. Usage and redistribution of any original creation, be it a research article or source code, lies within the legal boundaries of copyright laws, which differ significantly from country to country. Therefore, for example, publicly available code without an explicit attribution of a respective open-source license cannot be counted or treated as open source. Due to the sheer diversity, it may be difficult to judge which specific license is right for one's project. Yet the choice of the license must always be dictated by the project and the intent of its authors. Consulting the licenses list approved by the Open Source Initiative is generally considered to be a good starting point.

OPEN ACCESS

Edited by:

Jan Eglinger,
Friedrich Miescher Institute for
Biomedical Research (FMI),
Switzerland

Reviewed by:

Jean-Karim Hériché,
European Molecular Biology
Laboratory Heidelberg, Germany
Christopher Schmied,
Fondazione Human Technopole, Italy

*Correspondence:

Artur Yakimovich
a.yakimovich@hzdr.de

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Computational BiolMaging,
a section of the journal
Frontiers in Bioinformatics

Received: 04 April 2022

Accepted: 13 June 2022

Published: 05 July 2022

Citation:

Li R, Sharma V, Thangamani S and
Yakimovich A (2022) Open-Source
Biomedical Image Analysis Models: A
Meta-Analysis and Continuous Survey.
Front. Bioinform. 2:912809.
doi: 10.3389/fbinf.2022.912809

The importance of open source software for computational biomedical image analysis has become self-evident in the past 3 decades. Packages like ImageJ/Fiji (Schindelin et al., 2012; Schneider et al., 2012), CellProfiler (Carpenter et al., 2006), KNIME (Tiwari and Sekhar, 2007), and Icy (de Chaumont et al., 2011) not only perform the bulk of quantification tasks in the wetlabs but also serve as platforms for distribution of modules containing cutting-edge algorithms. The ability to install and use these modules and algorithms by researchers from various fields *via* a point-and-click interface made it possible for the research groups without image analysis specialists to obtain a qualitatively new level of biomedical insights from their data. Yet, as we transition into the data-driven and representation learning paradigm of biomedical image analysis, the availability of datasets and trained model parameters becomes as important as the open-source code.

The ability to download training parameters may allow researchers to skip the initial model training and focus on gradual model improvement through a technique known as transfer learning (West et al., 2007; Pan and Yang, 2010). Transfer learning has proven effective in Computer Vision (Deng et al., 2009) and Natural Language Processing (Wolf et al., 2020) domains (further reviewed in (Yakimovich et al., 2021)). However, the complexity of sharing the trained parameters of a model differs significantly between ML algorithms. For example, while model parameters of a conventional ML algorithm like linear regression may be conveniently shared in the text of the article, this is impossible for DL models with millions of parameters. This, in turn, requires rethinking conventional approaches to ML/DL models sharing under an open-source license.

In this review, we collate ML models for biomedical image analysis recently published in the peer-reviewed literature and available as open-source. We describe open-source licenses used, code availability, data availability, biomedical and ML tasks, as well as the availability of model parameters. We make the collated collection of the open-source model available *via* a GitHub repository and call on the research community to contribute their models to it *via* pull requests. Furthermore, we provide descriptive statistics of our observations and discuss the pros and cons of the status quo in the field of biomedical image analysis as well as perspectives in the general DS context. Several efforts to create biomedical ML model repositories or so-called “zoos” (e.g. bioimage.io) and web-based task consolidators (Hollandi et al., 2020; Stringer et al., 2021) have been undertaken. Here, rather than proposing a competing effort, we propose a continuous survey of the field “as is”. We achieve this through collating metadata of published papers and their respective source code, data, and model parameters (also known as weights and checkpoints).

Continuous Biomedical Image Analysis Model Survey

To understand the availability, reproducibility, and accessibility of published biomedical image analysis models we have collected a survey meta-dataset of 50 model articles and preprints published within the last 10 years. During our collection effort, we have prioritized publications with accompanying source code freely available online. In an

attempt to minimize bias, we made sure that no individual medical imaging modality or biomedical task represents more than 25% of our dataset. Additionally, we have attempted to sample models published by both the biomedical community (e.g. Nature group journals), engineering community (IEEE group journals and conferences), as well as models published as preprints. For each publication we have noted the biomedical imaging modality, biomedical task (e.g. cancer), the open-source license used, reported model performance with respective metric, whether the model is dealing with the supervised task, whether the model parameters can be downloaded (as well as the respective link), links to code and dataset. Noteworthy, performance reporting is highly dependent on a dataset or benchmark. Therefore, to avoid confusion or bias we have recorded the best-reported performance for illustrative purposes only. Identical performance on a different dataset should not be expected. For the purpose of this review, we have split this meta-dataset into three tables according to the ML task of the models. The full dataset is available on GitHub (<https://github.com/casus/bim>). To ensure the completeness and correctness of this meta-dataset we invite the research community to contribute their additions and corrections to our survey meta-dataset.

First display table obtained from our meta-dataset contains 14 models aimed at biomedical image classification (**Table 1**). The most prevalent imaging modalities for this ML task are computed tomography (CT) and digital pathology—both highly clinically relevant modalities. We noted that most publications had an open-source license clearly defined in their repositories. The consensus between the choices of metric is rather low, making it difficult to compare one model to the other. Although most models had both source code and datasets available, only 4 out of 14 models had trained model parameters available for download.

The second display table contains 25 models (**Table 2**) aimed at biomedical image segmentation—a task relevant for obtaining quantitative insights from the biomedical images (e.g. size of the tumor). Similarly, to the models for biomedical image classification, the vast majority of the segmentation models have a well-defined open-source license with only a few exceptions. Again, similarly to the classification models, the consensus between performance metric choices is rather low, although Dice score reports clearly dominated. Conversely, the percentage of models with pre-trained parameters available for download is slightly higher than in the case of the classification models (36% vs 29%). However, over half of the models do not provide pre-trained parameters for the download for both segmentation and classification tasks.

Finally, we have also examined biomedical image analysis models aimed at less popular ML tasks including data generation, object detection or reconstruction (**Table 3**). Apart from digital pathology, CT scans this group of models also contains light and electron microscopy. Remarkably, only 19% of models in this group had downloadable model parameters. At the same time, almost

TABLE 1 | Biomedical Image Classification Models. Here, AUC is Area under curve, CT is computed tomography.

Imaging Modality	Biomed Task	License	Reported Performance	Parameters Download	References
CT	Lung tumor	Apache-2.0	0.93 Accuracy	No	LaLonde et al. (2020)
CT	Lung tumor	MIT	0.76 AUC	No	Guo et al. (2020)
CT	Pulmonary nodule	GPL-3.0	0.90 Accuracy	No	Zhu et al. (2018a)
CT	Pulmonary nodule	MIT	0.96 AUC	No	Al-Shabi et al. (2019)
CT	Pulmonary nodule	MIT	0.95 AUC	No	Dey et al. (2018)
Dermatoscopy	Skin tumor	N/a	0.93 Accuracy	No	Datta et al. (2021)
Dermatoscopy	Skin tumor	MIT	0.81 AUC	Yes	Zunair and Ben Hamza, (2020)
Mammography	Breast tumor	CC BY-NC-ND 4.0	0.93 AUC	Yes	Shen et al. (2021)
Digital Pathology	Breast tumor	CC BY-NC-ND 4.0	0.63 F1	No	Pati et al. (2022)
Mammography	Breast tumor	CC BY-NC-SA 4.0	0.84 Accuracy	Yes	Shen et al. (2019)
Digital Pathology	Breast tumor	MIT	0.93 Accuracy	Yes	Rakhlin et al. (2018)
Digital Pathology	Lung tumor	GPL-3.0	0.53 Kappa	No	Wei et al. (2019)
Digital Pathology	Lung tumor	MIT	0.97 AUC	No	Coudray et al. (2018)
Fluorescence microscopy	Host-pathogen interactions	N/a	0.92 Accuracy	No	Fisch et al. (2019)

TABLE 2 | Biomedical Image Segmentation Models. Here, CT is computed tomography, DSC is Dice similarity coefficient, AP is Average Precision, IoU is Intersection over Union, DOF is Depth of field, AUC is Area under curve, SHG is Second harmonic generation microscopy.

Imaging Modality	Biomed Task	License	Reported Performance	Parameters Download	References
3D microscopy	Nuclei detection	MIT	0.937 AP	No	Hirsch and Kainmueller, (2020)
CT	Kidney tumor	GPL-3.0	0.95 Dice	No	Müller and Kramer, (2021)
CT	Pulmonary nodule	BSD-3-Clause	N/a	No	Hancock and Magnan, (2019)
CT	Pulmonary nodule	CC BY-NC-SA 4.0	0.55 IoU	Yes	Aresta et al. (2019)
CT	Pulmonary nodule	MIT	0.83 DSC	No	Keetha et al. (2020)
CT	Pancreas & Brain tumor	MIT	0.84 Dice	No	Oktay et al. (2018)
CT, Dermatoscopy	Lung tumor and Skin tumor	N/a	0.9965 Jaccard	No	Kaul et al. (2019)
CT	Brain tumor	Apache 2.0	0.89 Dice	No	Isensee et al. (2018)
MRI	Brain tumor	Apache 2.0	0.79 Dice	No	Wang et al. (2021)
MRI	Brain tumor	CC BY-NC-ND 4.0	0.76 Dice	No	Baek et al. (2019)
Digital Pathology	Breast tumor	CC BY-NC-ND 4.0	0.893 F1	Yes	Le et al. (2020)
Digital Pathology	Lung tumor	CC-BY	0.83 Accuracy	No	Tomita et al. (2019)
Digital Pathology	Multiple pathologies	MIT	N/a	No	Khened et al. (2021)
Electron microscopy	Multiple pathologies	MIT	0.5 VI	Yes	Lee et al. (2017)
Fluorescence microscopy	Cellular structures reconstruction	N/a	20 x Enhancement in DOF	Yes	Wu et al. (2019)
Fluorescence microscopy	Nuclei detection	BSD-3-Clause	0.94 Accuracy	Yes	Weigert et al. (2020)
Microscopy	Cellular reconstruction	N/a	0.69 AP	No	Hirsch et al. (2020)
MRI	Brain tumor	BSD-3-Clause	0.87 Dice	Yes	Wang et al. (2018)
MRI	Brain tumor	MIT	0.85 Dice	Yes	Havaei et al. (2017)
MRI	Brain tumor	MIT	0.90 Dice	No	Isensee et al. (2018)
MRI	Brain tumor	MIT	0.91 Dice	No	Myronenko, (2019)
SHG	Bone disease	GPL-3.0	0.78 Accuracy	No	Schmarje et al. (2019)
Time-lapse microscopy	Nuclei detection	N/a	0.92 Accuracy	Yes	Shailja et al. (2021)
Ultrasound imaging	Intraventricular hemorrhage	MIT	0.89 Dice	No	Valanarasu et al. (2020)
MRI	Brain tumor	N/a	0.81 Dice	Yes	Larrazabal et al. (2021)

all the models in this group had well attributed open-source licenses. This may suggest that parameter sharing is not very common in highly specialized fields like microscopy. Interestingly, for this and other groups of ML tasks, we have found that parameter sharing was more common in models submitted as a part of a data challenge. This may be simply a result of data challenge participation conditions.

Trends Meta-Analysis in Biomedical Image Analysis Model

To understand general trends in the collection of our open-source models we have computed respective fractions of each descriptive category we have assigned to each work. The assignment was performed through careful analysis of the respective research article, code repository, dataset repository, and the availability of

TABLE 3 | Other Biomedical Image Models. Here, CT is computed tomography.

Imaging Modality	Biomed Task	ML Task	License	Parameters Download	References
Mammography	Breast tumor	Classification & Detection	N/a	Yes	Ribli et al. (2018)
Fluorescence microscopy	Cellular structures reconstruction	Data generation	Apache-2.0	No	Eschweiler et al. (2021)
CT	Pulmonary nodule	Detection	Apache-2.0	No	Zhu et al. (2018b)
CT	Pulmonary nodule	Detection	MIT	No	Li and Fan, (2020)
Digital Pathology	Multiple pathologies	Graph embedding	AGPL 3.0	No	Jaume et al. (2021)
Mammography	Breast tumor	Image Inpainting & Data generation	CC BY-NC-ND 4.0	Yes	Wu et al. (2018)
Confocal microscopy	Cellular structures reconstruction	Reconstruction	Apache-2.0	No	Vizcaíno et al. (2021)
Cryo-electron microscopy	Cellular structures reconstruction	Reconstruction	GPL-3.0	No	Zhong et al. (2019)
Cryo-electron microscopy	Protein structures reconstruction	Reconstruction	GPL-3.0	No	Ullrich et al. (2019)
Electron microscopy	Cellular structures reconstruction	Reconstruction	N/a	No	Guay et al. (2021)
3D microscopy	Image acquisition	Reconstruction	BSD-3-Clause	No	Saha et al. (2020)

the trained model parameters (**Figure 1**). While admittedly 50 papers constitute a relatively small sample size, we have made the best reasonable effort to ensure the sampling was unbiased. Specifically, the set of models we have reviewed addresses the following biomedical tasks (from most to least frequent): pulmonary nodule, brain tumor, breast tumor, cellular structures reconstruction, lung tumor, cell nuclei detection, multiple pathologies, skin tumor, protein structures reconstruction, kidney tumor, pancreas and brain tumor, lung tumor and skin tumor, host-pathogen interactions, bone disease, image acquisition, intraventricular hemorrhage (**Figure 1A**).

From the perspective of imaging modalities, the models we reviewed span the following: computed tomography (CT), digital pathology, magnetic resonance imaging (MRI), mammography, fluorescence microscopy, 3D microscopy, cryo-electron microscopy, dermatoscopy, electron microscopy, confocal microscopy, CT and dermatoscopy, light and electron microscopy, second harmonic generation microscopy, time-lapse microscopy, ultrasound imaging (**Figure 1B**). From the perspective of ML tasks these models covered the following: segmentation, reconstruction classification, object detection, image inpainting and data generation, graph embedding, classification, and detection (**Figure 1C**). 86% of the models we have reviewed were addressing supervised tasks and 14% unsupervised tasks (**Figure 1D**).

Within our collection of open-source models, we have noted that 32% of the authors have selected the MIT license, 18% have selected Apache-2.0, 12%—GPL-3.0, 10%—BSD-3-Clause license, 8%—CC BY-NC-SA 4.0 license. Remarkably, another 8% have published their code without license attribution, arguably making it harder for the field to understand the freedom to operate with the code made available with the paper (**Figure 1E**). Within these papers, 84% of the authors made the dataset used to train the model available and clearly indicated within the paper or the code repository (**Figure 1F**). Overall, this amounted to the vast majority of the works which we have selected to have a clear open-source license designation, as well as a dataset available.

Remarkably, while providing the model's source code, as well as, in most cases, the model's dataset, an impressive 68% of the contributions we have reviewed did not provide trained model parameters (**Figure 1G**). Breaking down by the publishers or repositories, 43% and 31% of papers published by Nature group and Springer respectively provided model parameters. However, only 25% of IEEE papers and 14% of arXiv preprints provided parameters. Altogether, the low percentage of shared parameters are suggesting that the efforts to reproduce these papers came with the caveat of provisioning a hardware setup capable of wielding the computational load required by the respective model. In some cases that requires access to the high-capacity computing. Furthermore, this way, instead of simply building upon the models trained, the efforts of the authors would have to be first reproduced. Needless to say, should any of the papers become seminal these high-performance computations would have to be repeated time and time again, possibly taking days of GPU computation.

Interestingly, of the authors who have chosen to make the trained parameters available to the readers around 25% have chosen to deposit the parameters on GitHub, while 19% and 6% have opted for Google drive and Dropbox services respectively. The rest deposited their parameters on the proprietary and other services (**Figure 1H**).

DISCUSSION

The advent of ML and specifically representation learning is opening a new horizon for biomedical image analysis. Yet, the success of these new advanced ML approaches brings about new requirements and standards to ensure quality and reproducibility (Hernandez-Boussard et al., 2020; Mongan et al., 2020; Norgeot et al., 2020; Heil et al., 2021; Laine et al., 2021). Several minimalistic quality standards applicable to the clinical setting have been proposed (Hernandez-Boussard et al., 2020; Mongan et al., 2020; Norgeot et al., 2020), and while coming from slightly

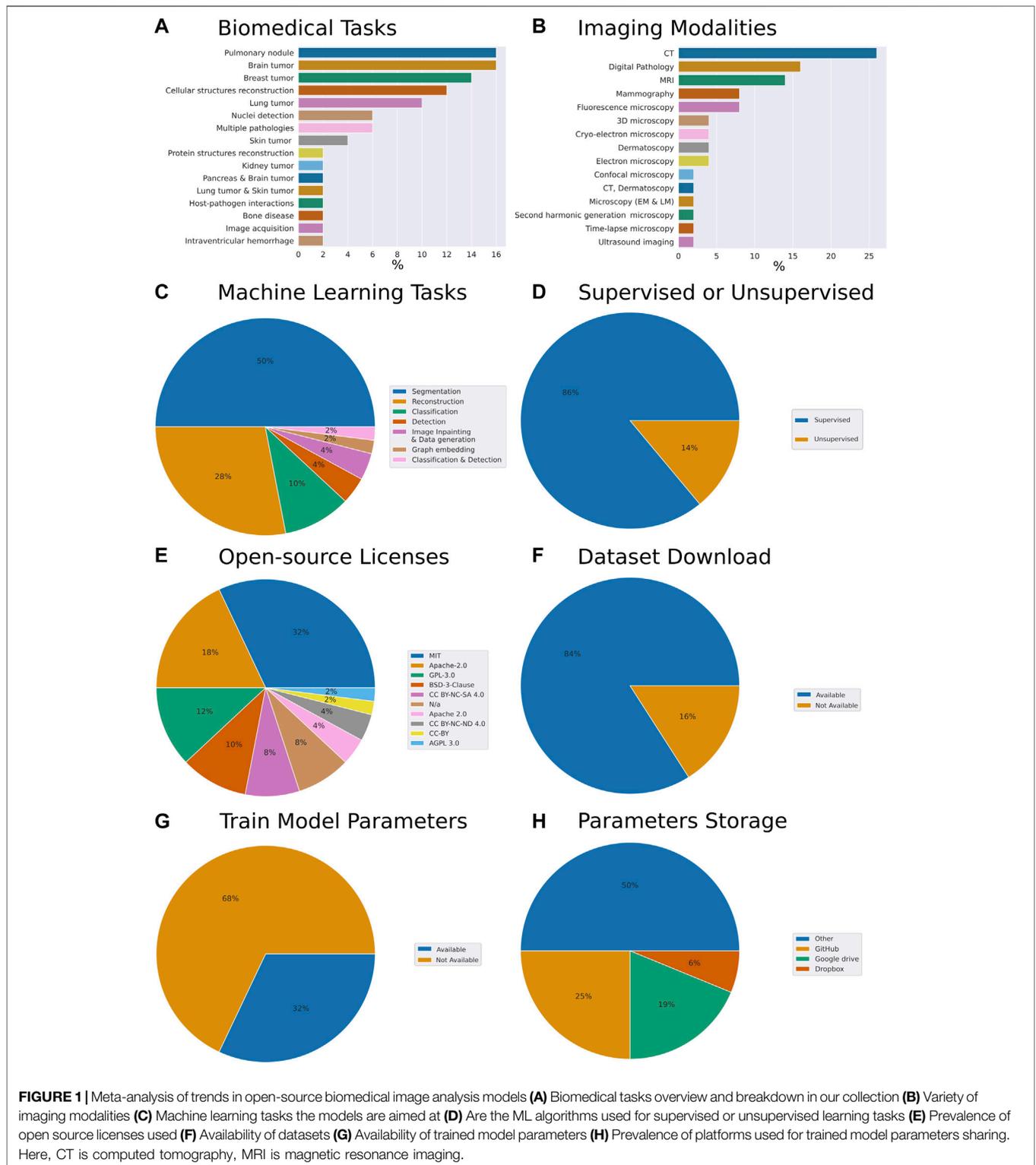


FIGURE 1 | Meta-analysis of trends in open-source biomedical image analysis models **(A)** Biomedical tasks overview and breakdown in our collection **(B)** Variety of imaging modalities **(C)** Machine learning tasks the models are aimed at **(D)** Are the ML algorithms used for supervised or unsupervised learning tasks **(E)** Prevalence of open source licenses used **(F)** Availability of datasets **(G)** Availability of trained model parameters **(H)** Prevalence of platforms used for trained model parameters sharing. Here, CT is computed tomography, MRI is magnetic resonance imaging.

different perspectives they demonstrate an overlap on essential topics like the dataset description, comparison to baseline and hyperparameters sharing. For example, CLAIM (Mongan et al.,

2020) and MINIMAR (Hernandez-Boussard et al., 2020) approaches aim to adhere to a clinical tradition. Authors define a checklist including a structure of an academic

biomedical paper, requiring either a lengthy biomedical problem description (CLAIM) or descriptive statistics of the dataset's internal structure (MINIMAR). At the same time, MI-CLAIM (Norgeot et al., 2020) aims to adhere to the Data Science tradition, focusing specifically on data preprocessing and baseline comparison. Remarkably, even though item 24 of the CLAIM checklist explicitly mentions the importance of specifying the source of the starting weights (parameters) if transfer learning is employed, all three approaches fail to explicitly encourage sharing of the trained model parameters. Instead of proposing yet another checklist, the current survey aims to understand to extend to which the model parameters are shared in the biomedical image analysis field and emphasize the importance of parameters sharing to foster reproducibility in the field.

The past 3 decades have successfully demonstrated the viability of the open-source model for the research software in this field, as well as the role of open-source software in fostering scientific progress. However, the change of modeling paradigm to DL requires new checks and balances to ensure the results are reproducible and the efforts are not doubled. Furthermore, major computational efforts inevitably come with an environmental footprint (Strubell et al., 2020). Making parameters of the trained models available to the research community not only could minimize this footprint, but also open new prospects for the researcher wishing to fine-tune the pre-trained models to their task of choice. Such an approach proved incredibly fruitful in the field of natural language processing (Zhang et al., 2020).

Remarkably, in the current survey, we have found that only 32% of the biomedical models we have reviewed made the train model parameters available for download. On one hand, such a low number of trained models available for download may be explained by the fact that many journals and conferences do not require trained models to warrant publication. On another hand, with parameters of some models requiring hundreds of megabytes of storage, there are not many opportunities to share these files. Interestingly, while some researchers shared their trained model parameters *via* platforms like GitHub, Google drive, and Dropbox, the vast majority opted for often proprietary sites to share these parameters (Figure 1H). In our opinion, this indicates the necessity of hubs and platforms for sharing trained biomedical image analysis models.

It is worth noting that most cloud storage services like Google drive or Dropbox are more suited for instant file sharing rather than archival deposition of model parameters. These storage solutions don't offer data immutability or digital object identifiers attached to them, and hence can simply be overwritten or disappear leaving crucial content inaccessible. Authors opting for self-hosting of model parameters also likely underestimate the workload of the long-term serving of archival data. Instead of the aforementioned approaches to model sharing, one should take advantage of efforts like BioImage.io, Tensorflow Hub (Paper, 2021), PyTorch Hub, DLHub (Chard et al., 2019), or similar in order to foster consistency and reproducibility of their

results. Arguably, one of the most intuitive experiences of model parameters sharing for the end-users is currently offered by the HuggingFace platform in the domain of natural language processing. This has largely been possible through the platform's own ML library allowing for improved compatibility (Wolf et al., 2020).

Interestingly, the vast majority of authors have chosen MIT and Apache-2.0 as their open-source licenses. Both Apache-2.0 and MIT are known for being permissive, rather than copyleft licenses. Furthermore, both licenses are very clearly formulated and easy to use. It is tempting to speculate that their popularity is a result of the simplicity and openness that these licenses offer.

However, noteworthy, our survey is limited to the papers we reviewed. To improve the representativeness of our meta-analysis, as well as encourage the dissemination of the open-source models in biomedical image analysis we call on our peers to contribute to our collection *via* the GitHub repository. Specifically, we invite the researchers to fork our repository, make additions to the content of the list following the contribution guidelines and merge them in *via* pull request. This way we hope to not only obtain an up-to-date state of the field but also ensure the code, datasets and trained model parameters are easier to find.

AUTHOR CONTRIBUTIONS

AY conceived the idea. AY, ST, VS, and RL reviewed the published works and collated the data. AY, ST, VS, and RL wrote the manuscript.

FUNDING

This work was partially funded by the Center for Advanced Systems Understanding (CASUS) which is financed by Germany's Federal Ministry of Education and Research (BMBF) and by the Saxon Ministry for Science, Culture and Tourism (SMWK) with tax funds on the basis of the budget approved by the Saxon State Parliament. This work has been partially funded by OPTIMA. OPTIMA is funded through the IMI2 Joint Undertaking and is listed under grant agreement No. 101034347. IMI2 receives support from the European Union's Horizon 2020 research and innovation programme and the European Federation of Pharmaceutical Industries and Associations (EFPIA). IMI supports collaborative research projects and builds networks of industrial and academic experts in order to boost pharmaceutical innovation in Europe. The views communicated within are those of OPTIMA. Neither the IMI nor the European Union, EFPIA, or any Associated Partners are responsible for any use that may be made of the information contained herein.

REFERENCES

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., et al. (2016). "TensorFlow: A System for Large-Scale Machine Learning," in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)* (usenix.org), 265
- Al-Shabi, M., Lan, B. L., Chan, W. Y., Ng, K. H., and Tan, M. (2019). Lung Nodule Classification Using Deep Local-Global Networks. *Int. J. Comput. Assist. Radiol. Surg.* 14, 1815–1819. doi:10.1007/s11548-019-01981-7
- Aresta, G., Jacobs, C., Araújo, T., Cunha, A., Ramos, I., van Ginneken, B., et al. (2019). iW-Net: an Automatic and Minimalistic Interactive Lung Nodule Segmentation Deep Network. *Sci. Rep.* 9, 11591. doi:10.1038/s41598-019-48004-8
- Baek, S., He, Y., Allen, B. G., Buatti, J. M., Smith, B. J., Tong, L., et al. (2019). Deep Segmentation Networks Predict Survival of Non-small Cell Lung Cancer. *Sci. Rep.* 9, 17286. doi:10.1038/s41598-019-53461-2
- Carpenter, A. E., Jones, T. R., Lamprecht, M. R., Clarke, C., Kang, I. H., Friman, O., et al. (2006). CellProfiler: Image Analysis Software for Identifying and Quantifying Cell Phenotypes. *Genome Biol.* 7, R100. doi:10.1186/gb-2006-7-10-r100
- Chard, R., Li, Z., Chard, K., Ward, L., Babuji, Y., Woodard, A., et al. (2019). "DLHub: Model and Data Serving for Science," in 2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS) (ieeexplore.ieee.org), 283–292. doi:10.1109/ipdps.2019.00038
- Coudray, N., Ocampo, P. S., Sakellaropoulos, T., and Narula, N. (2018). "Classification and Mutation Prediction from Non-small Cell Lung Cancer Histopathology Images Using Deep Learning," *Nat. Med.* Available at: <https://www.nature.com/articles/s41591-018-0177-5?sf197831152=1>.
- Datta, S. K., Shaikh, M. A., Srihari, S. N., and Gao, M. (2021). "Soft Attention Improves Skin Cancer Classification Performance," in *Interpretability of Machine Intelligence in Medical Image Computing, and Topological Data Analysis and its Applications for Medical Data* (Springer International Publishing), 13–23. doi:10.1007/978-3-030-87444-5_2
- de Chaumont, F., Dallongeville, S., and Olivo-Marin, J.-C. (2011). "ICY: A New Open-Source Community Image Processing Software," in 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (ieeexplore.ieee.org), 234–237. doi:10.1109/isbi.2011.5872395
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). "ImageNet: A Large-Scale Hierarchical Image Database," in 2009 IEEE Conference on Computer Vision and Pattern Recognition, 248–255. doi:10.1109/cvpr.2009.5206848
- Dey, R., Lu, Z., and Hong, Y. (2018). "Diagnostic Classification of Lung Nodules Using 3D Neural Networks," in 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018) (ieeexplore.ieee.org), 774–778. doi:10.1109/isbi.2018.8363687
- Eschweiler, D., Rethwisch, M., Jarchow, M., Koppers, S., and Stegmaier, J. (2021). 3D Fluorescence Microscopy Data Synthesis for Segmentation and Benchmarking. *PLoS One* 16, e0260509. doi:10.1371/journal.pone.0260509
- Fisch, D., Yakimovich, A., Clough, B., Wright, J., Bunyan, M., Howell, M., et al. (2019). Defining Host-Pathogen Interactions Employing an Artificial Intelligence Workflow. *Elife* 8, e40560. doi:10.7554/eLife.40560
- Guay, M. D., Emam, Z. A. S., Anderson, A. B., Aronova, M. A., Pokrovskaya, I. D., Storrie, B., et al. (2021). Dense Cellular Segmentation for EM Using 2D-3D Neural Network Ensembles. *Sci. Rep.* 11, 2561–2611. doi:10.1038/s41598-021-81590-0
- Guo, H., Kruger, U., Wang, G., Kalra, M. K., and Yan, P. (2020). Knowledge-Based Analysis for Mortality Prediction from CT Images. *IEEE J. Biomed. Health Inf.* 24, 457–464. doi:10.1109/JBHI.2019.2946066
- Hancock, M. C., and Magnan, J. F. (2019). Level Set Image Segmentation with Velocity Term Learned from Data with Applications to Lung Nodule Segmentation. *arXiv [eess.IV]*. Available at: <http://arxiv.org/abs/1910.03191> (Accessed March 31, 2022).
- Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., et al. (2017). Brain Tumor Segmentation with Deep Neural Networks. *Med. Image Anal.* 35, 18–31. doi:10.1016/j.media.2016.05.004
- Heil, B. J., Hoffman, M. M., Markowetz, F., Lee, S. I., Greene, C. S., and Hicks, S. C. (2021). Reproducibility Standards for Machine Learning in the Life Sciences. *Nat. Methods* 18, 1132–1135. doi:10.1038/s41592-021-01256-7
- Hernandez-Boussard, T., Bozkurt, S., Ioannidis, J. P. A., and Shah, N. H. (2020). MINIMAR (MINimum Information for Medical AI Reporting): Developing Reporting Standards for Artificial Intelligence in Health Care. *J. Am. Med. Inf. Assoc.* 27, 2011–2015. doi:10.1093/jamia/ocaa088
- Hirsch, P., and Kainmueller, D. (2020). "An Auxiliary Task for Learning Nuclei Segmentation in 3D Microscopy Images," in Proceedings Of the Third Conference On Medical Imaging With Deep Learning *Proceedings of Machine Learning Research*. Editors T. Arbel, I. Ben Ayed, M. de Bruijne, M. Descoteaux, H. Lombaert, and C. Pal (Montreal, QC, Canada: PML), 304
- Hirsch, P., Mais, L., and Kainmueller, D. (2020). PatchPerPix for Instance Segmentation. *arXiv [cs.CV]*. Available at: <http://arxiv.org/abs/2001.07626> (Accessed March 30, 2022).
- Hollandi, R., Szkalitsy, A., Toth, T., Tasnadi, E., Molnar, C., Mathe, B., et al. (2020). nucleAIzer: A Parameter-free Deep Learning Framework for Nucleus Segmentation Using Image Style Transfer. *Cell. Syst.* 10, 453–e6. doi:10.1016/j.cels.2020.04.003
- Isensee, F., Kickingereder, P., Wick, W., Bendszus, M., and Maier-Hein, K. H. (2018). "Brain Tumor Segmentation and Radiomics Survival Prediction: Contribution to the BRATS 2017 Challenge," in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Quebec City, Canada: Springer International Publishing), 287–297. doi:10.1007/978-3-319-75238-9_25
- Jaume, G., Pati, P., Anklin, V., Foncubierta, A., and Gabrani, M. (2021). "HistoCartography: A Toolkit for Graph Analytics in Digital Pathology," in Proceedings Of the MICCAI Workshop On Computational Pathology *Proceedings of Machine Learning Research*. Editors M. Atzori, N. Burlutskiy, F. Ciompi, Z. Li, F. Minhas, and H. Müller (PMLR), 117
- Kaul, C., Manandhar, S., and Pears, N. (2019). "Focusnet: An Attention-Based Fully Convolutional Network for Medical Image Segmentation," in 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019) (ieeexplore.ieee.org), 455–458. doi:10.1109/isbi.2019.8759477
- Keetha, N. V., Samson, A. B. P., and Annavarapu, C. S. R. (2020). U-det: A Modified U-Net Architecture with Bidirectional Feature Network for Lung Nodule Segmentation. *arXiv [eess.IV]*. Available at: <http://arxiv.org/abs/2003.09293> (Accessed March 31, 2022).
- Khened, M., Kori, A., Rajkumar, H., Krishnamurthi, G., and Srinivasan, B. (2021). A Generalized Deep Learning Framework for Whole-Slide Image Segmentation and Analysis. *Sci. Rep.* 11, 11579. doi:10.1038/s41598-021-90444-8
- Laine, R. F., Arganda-Carreras, I., Henriques, R., and Jacquemet, G. (2021). Avoiding a Replication Crisis in Deep-Learning-Based Bioimage Analysis. *Nat. Methods* 18, 1136–1144. doi:10.1038/s41592-021-01284-3
- LaLonde, R., Torigian, D., and Bagci, U. (2020). "Encoding Visual Attributes in Capsules for Explainable Medical Diagnoses," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020* (Springer International Publishing), 294–304. doi:10.1007/978-3-030-59710-8_29
- Landset, S., Khoshgoftaar, T. M., Richter, A. N., and Hasanin, T. (2015). A Survey of Open Source Tools for Machine Learning with Big Data in the Hadoop Ecosystem. *J. Big Data* 2, 1–36. doi:10.1186/s40537-015-0032-1
- Larrazabal, A. J., Martínez, C., Dolz, J., and Ferrante, E. (2021). "Orthogonal Ensemble Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021* (Strasbourg, France: Springer International Publishing), 594–603. doi:10.1007/978-3-030-87199-4_56
- Le, H., Gupta, R., Hou, L., Abousamra, S., Fassler, D., Torre-Healy, L., et al. (2020). Utilizing Automated Breast Cancer Detection to Identify Spatial Distributions of Tumor-Infiltrating Lymphocytes in Invasive Breast Cancer. *Am. J. Pathol.* 190, 1491–1504. doi:10.1016/j.ajpath.2020.03.012
- Lee, K., Zung, J., Li, P., Jain, V., and Sebastian Seung, H. (2017). Superhuman Accuracy on the SNEMI3D Connectomics Challenge. *arXiv [cs.CV]*. Available at: <http://arxiv.org/abs/1706.00120>.
- Li, Y., and Fan, Y. (2020). DeepSEED: 3D Squeeze-And-Excitation Encoder-Decoder Convolutional Neural Networks for Pulmonary Nodule Detection. *Proc. IEEE Int. Symp. Biomed. Imaging* 2020, 1866–1869. doi:10.1109/ISBI45749.2020.9098317

- Mongan, J., Moy, L., and Kahn, C. E., Jr (2020). Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers. *Radiol. Artif. Intell.* 2, e200029. doi:10.1148/ryai.2020200029
- Müller, D., and Kramer, F. (2021). MIScnn: a Framework for Medical Image Segmentation with Convolutional Neural Networks and Deep Learning. *BMC Med. Imaging* 21, 12. doi:10.1186/s12880-020-00543-7
- Myronenko, A. (2019). “3D MRI Brain Tumor Segmentation Using Autoencoder Regularization,” in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Granada, Spain: Springer International Publishing), 311–320. doi:10.1007/978-3-030-11726-9_28
- Norgeot, B., Quer, G., Beaulieu-Jones, B. K., Torkamani, A., Dias, R., Gianfrancesco, M., et al. (2020). Minimum Information about Clinical Artificial Intelligence Modeling: the MI-CLAIM Checklist. *Nat. Med.* 26, 1320–1324. doi:10.1038/s41591-020-1041-y
- Oktay, O., Schlemper, J., Le Folgoc, L., Lee, M., Heinrich, M., Misawa, K., et al. (2018). Attention U-Net: Learning where to Look for the Pancreas. *arXiv [cs.CV]*. Available at: <http://arxiv.org/abs/1804.03999> (Accessed March 31, 2022).
- Pan, S., and Yang, Q. (2010). A Survey on Transfer Learning. *IEEE Transaction Knowl. Discov. Data Eng.* 22 (10), 191. doi:10.1109/tkde.2009.191
- Paper, D. (2021). “Simple Transfer Learning with TensorFlow Hub,” in *State-of-the-Art Deep Learning Models in TensorFlow: Modern Machine Learning in the Google Colab Ecosystem*. Editor D. Paper (Berkeley, CA: Apress), 153–169. doi:10.1007/978-1-4842-7341-8_6
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. *Adv. Neural Inf. Process. Syst.* 32. Available at: <https://proceedings.neurips.cc/paper/2019/hash/bdbca288fee7f92f2bfa9f7012727740-Abstract.html> (Accessed March 24, 2022).
- Pati, P., Jaume, G., Foncubierta-Rodríguez, A., Feroce, F., Anniciello, A. M., Scognamiglio, G., et al. (2022). Hierarchical Graph Representations in Digital Pathology. *Med. Image Anal.* 75, 102264. doi:10.1016/j.media.2021.102264
- Rakhlin, A., Shvets, A., Igloukov, V., and Kalinin, A. A. (2018). “Deep Convolutional Neural Networks for Breast Cancer Histology Image Analysis,” in *Image Analysis and Recognition* (Póvoa de Varzim, Portugal: Springer International Publishing), 737–744. doi:10.1007/978-3-319-93000-8_83
- Ribli, D., Horváth, A., Unger, Z., Pollner, P., and Csabai, I. (2018). Detecting and Classifying Lesions in Mammograms with Deep Learning. *Sci. Rep.* 8, 4165. doi:10.1038/s41598-018-22437-z
- Saha, D., Schmidt, U., Zhang, Q., Barbotin, A., Hu, Q., Ji, N., et al. (2020). Practical Sensorless Aberration Estimation for 3D Microscopy with Deep Learning. *Opt. Express* 28, 29044–29053. doi:10.1364/OE.401933
- Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an Open-Source Platform for Biological-Image Analysis. *Nat. Methods* 9, 676–682. doi:10.1038/nmeth.2019
- Schmarje, L., Zelenka, C., Geisen, U., Glüer, C.-C., and Koch, R. (2019). “2D and 3D Segmentation of Uncertain Local Collagen Fiber Orientations in SHG Microscopy,” in *Pattern Recognition* (Dortmund, Germany: Springer International Publishing), 374–386. doi:10.1007/978-3-030-33676-9_26
- Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 Years of Image Analysis. *Nat. Methods* 9, 671–675. doi:10.1038/nmeth.2089
- Shailja, S., Jiang, J., and Manjunath, B. S. (2021). “Semi Supervised Segmentation and Graph-Based Tracking of 3D Nuclei in Time-Lapse Microscopy,” in *2021 IEEE 18th International Symposium on Biomedical Imaging* (Nice, France: ISBI IEEE), 385–389. doi:10.1109/isbi48211.2021.9433831
- Shen, L., Margolies, L. R., Rothstein, J. H., Fluder, E., McBride, R., and Sieh, W. (2019). Deep Learning to Improve Breast Cancer Detection on Screening Mammography. *Sci. Rep.* 9, 12495. doi:10.1038/s41598-019-48995-4
- Shen, Y., Wu, N., Phang, J., Park, J., Liu, K., Tyagi, S., et al. (2021). An Interpretable Classifier for High-Resolution Breast Cancer Screening Images Utilizing Weakly Supervised Localization. *Med. Image Anal.* 68, 101908. doi:10.1016/j.media.2020.101908
- Sonnenburg, S., Braun, M. L., Ong, C. S., Bengio, S., Bottou, L., Holmes, G., et al. (2007). The Need for Open Source Software in Machine Learning. Available at: <https://www.jmlr.org/papers/volume8/sonnenburg07a/sonnenburg07a.pdf> (Accessed March 24, 2022).
- Stringer, C., Wang, T., Michaelos, M., and Pachitariu, M. (2021). Cellpose: a Generalist Algorithm for Cellular Segmentation. *Nat. Methods* 18, 100–106. doi:10.1038/s41592-020-01018-x
- Strubell, E., Ganesh, A., and McCallum, A. (2020). Energy and Policy Considerations for Modern Deep Learning Research. *AAAI* 34, 13693–13696. doi:10.1609/aaai.v34i09.7123
- Tiwari, A., and Sekhar, A. K. (2007). Workflow Based Framework for Life Science Informatics. *Comput. Biol. Chem.* 31, 305–319. doi:10.1016/j.compbiolchem.2007.08.009
- Tomita, N., Abdollahi, B., Wei, J., Ren, B., Suriawinata, A., and Hassanpour, S. (2019). Attention-Based Deep Neural Networks for Detection of Cancerous and Precancerous Esophagus Tissue on Histopathological Slides. *JAMA Netw. Open* 2, e1914645. doi:10.1001/jamanetworkopen.2019.14645
- Ulrich, K., van den Berg, R., Brubaker, M., Fleet, D., and Welling, M. (2019). Differentiable Probabilistic Models of Scientific Imaging with the Fourier Slice Theorem. *arXiv [cs.LG]*. Available at: <http://arxiv.org/abs/1906.07582> (Accessed March 31, 2022).
- Valanarasu, J. M. J., Sindagi, V. A., Hacihaliloglu, I., and Patel, V. M. (2020). “KiU-Net: Towards Accurate Segmentation of Biomedical Images Using Over-complete Representations,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020* (Lima, Peru: Springer International Publishing), 363–373. doi:10.1007/978-3-030-59719-1_36
- Vizcaino, J. P., Saltarin, F., Belyaev, Y., Lyck, R., Lasser, T., Favaro, P., et al. (2021). Learning to Reconstruct Confocal Microscopy Stacks from Single Light Field Images. *IEEE Trans. Comput. Imaging* 7, 775. doi:10.1109/TCI.2021.3097611
- Wang, G., Li, W., Ourselin, S., and Vercauteren, T. (2018). “Automatic Brain Tumor Segmentation Using Cascaded Anisotropic Convolutional Neural Networks,” in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* (Quebec City, Canada: Springer International Publishing), 178–190. doi:10.1007/978-3-319-75238-9_16
- Wang, W., Chen, C., Ding, M., Yu, H., Zha, S., and Li, J. (2021). “TransBTS: Multimodal Brain Tumor Segmentation Using Transformer,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, Strasbourg, France, September 27–October 1, 2021 (Springer International Publishing), 109–119. doi:10.1007/978-3-030-87193-2_11
- Wei, J. W., Tafe, L. J., Linnik, Y. A., Vaickus, L. J., Tomita, N., and Hassanpour, S. (2019). Pathologist-level Classification of Histologic Patterns on Resected Lung Adenocarcinoma Slides with Deep Neural Networks. *Sci. Rep.* 9, 3358. doi:10.1038/s41598-019-40041-7
- Weigert, M., Schmidt, U., Haase, R., Sugawara, K., and Myers, G. (2020). “Star-convex Polyhedra for 3d Object Detection and Segmentation in Microscopy,” in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Snowmass village, CO, United States, March 2–5, 2020* (openaccess.thecvf.com), 3666–3673. doi:10.1109/wacv45572.2020.9093435
- West, J., Ventura, D., and Warnick, S. (2007). *Spring Research Presentation: A Theoretical Foundation for Inductive Transfer*, 1. Provo, UT, United States: Brigham Young University, College of Physical and Mathematical Sciences.
- Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., et al. (2020). “Transformers: State-Of-The-Art Natural Language Processing,” in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, Online, November 16–20, 2020 (Online: Association for Computational Linguistics). doi:10.18653/v1/2020.emnlp-demos.6
- Wu, E., Wu, K., Cox, D., and Lotter, W. (2018). “Conditional Infilling GANs for Data Augmentation in Mammogram Classification,” in *Image Analysis for Moving Organ, Breast, and Thoracic Images* (Granada, Spain: Springer International Publishing), 98–106. doi:10.1007/978-3-030-00946-5_11
- Wu, Y., Rivenson, Y., Wang, H., Luo, Y., Ben-David, E., Bentolila, L. A., et al. (2019). Three-dimensional Virtual Refocusing of Fluorescence Microscopy Images Using Deep Learning. *Nat. Methods* 16, 1323–1331. doi:10.1038/s41592-019-0622-5
- Yakimovich, A., Beaugnon, A., Huang, Y., and Ozkirimli, E. (2021). Labels in a Haystack: Approaches beyond Supervised Learning in Biomedical Applications. *Patterns* 2, 100383. doi:10.1016/j.patter.2021.100383
- Zhang, N., Li, L., Deng, S., Yu, H., Cheng, X., Zhang, W., et al. (2020). Can Fine-Tuning Pre-trained Models Lead to Perfect Nlp? a Study of the Generalizability

- of Relation Extraction. Available at: <https://openreview.net/forum?id=3yzlj2-eBbZ> (Accessed April 4, 2022).
- Zhong, E. D., Bepler, T., Davis, J. H., and Berger, B. (2019). Reconstructing Continuous Distributions of 3D Protein Structure from Cryo-EM Images. *arXiv [q-bio.QM]*. Available at: <http://arxiv.org/abs/1909.05215> (Accessed March 30, 2022).
- Zhu, W., Liu, C., Fan, W., and Xie, X. (2018a). “DeepLung: Deep 3D Dual Path Nets for Automated Pulmonary Nodule Detection and Classification,” in 2018 IEEE Winter Conference on Applications of Computer Vision (WACV) (ieeexplore.ieee.org), Granada, Spain, March 12–15, 2018, 673–681. doi:10.1109/wacv.2018.00079
- Zhu, W., Vang, Y. S., Huang, Y., and Xie, X. (2018b). “DeepEM: Deep 3D ConvNets with EM for Weakly Supervised Pulmonary Nodule Detection,” in Medical Image Computing and Computer Assisted Intervention – MICCAI 2018 (Springer International Publishing), Lake Tahoe, NV, United States, March 12–15, 2018, 812–820. doi:10.1007/978-3-030-00934-2_90
- Zunair, H., and Ben Hamza, A. (2020). Melanoma Detection Using Adversarial Training and Deep Transfer Learning. *Phys. Med. Biol.* 65, 135005. doi:10.1088/1361-6560/ab86d3

Conflict of Interest: AY was employed by Roche Pharma International Informatics, Roche Diagnostics GmbH, Mannheim, Germany

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Li, Sharma, Thangamani and Yakimovich. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.