

# Deep and shallow data science for multi-scale optical neuroscience

Gal Mishne<sup>a</sup> and Adam Charles<sup>b</sup>

<sup>a</sup>Halicioğlu Data Science Institute, Department of Electrical and Computer Engineering and the Neurosciences Graduate Program, UC San Diego, 9500 Gilman Drive, La Jolla, CA 92093 USA

<sup>b</sup>Department of Biomedical Engineering, Kavli Neuroscience Discovery Institute, Center for Imaging Science, Department of Neuroscience, and Mathematical Institute for Data Science, Johns Hopkins University, Baltimore, MD 21287 USA

## ABSTRACT

Optical imaging of the brain has expanded dramatically in the past two decades. New optics, indicators, and experimental paradigms are now enabling in-vivo imaging from the synaptic to the cortex-wide scales. To match the resulting flood of data across scales, computational methods are continuously being developed to meet the need of extracting biologically relevant information. In this pursuit challenges arise in some domains (e.g., SNR and resolution limits in micron-scale data) that require specialized algorithms. These algorithms can, for example, make use of state-of-the-art machine learning to maximally learn the details of a given scale to optimize the processing pipeline. In contrast, other methods, however, such as graph signal processing, seek to abstract away from some of the details that are scale-specific to provide solution to specific sub-problems common across scales of neuroimaging. Here we discuss limitations and tradeoffs in algorithmic design with the goal of identifying how data quality and variability can hamper algorithm use and dissemination.

**Keywords:** fluorescence microscopy, calcium imaging, functional imaging, data analysis

## 1. INTRODUCTION

Data science, in the form of statistical modeling and image processing, has become a central topic in extracting the most out of complex imaging data, such as taken by the latest in optical designs.<sup>1</sup> Computational approaches in particular have blossomed as imaging methods have extended across the micro- and macro- scales that aim to glean information about how the brain processes information at all scales.<sup>2</sup> These methods fundamentally take into account—implicitly or explicitly—properties of the data imbued by the imaging target and the imaging device. Colloquially we refer to these as the likelihood and the prior, after the successful terminology from the Bayesian inference and estimation literature.

Due to the enormous range of targets (different brain areas across species) at different scales (cell bodies, widefield, dendrites, axons, vasculature) that exhibit different temporal dynamics (different indicators of calcium and voltage, hemodynamics, long-term AMPA receptor changes, etc.), fundamentally different tools have been developed by the community. Some of these tools have aimed to solve core processing challenges, for example, denoising or identifying functional components in videos, across many different imaging contexts. Other tools instead have focused on concentrating effort toward solving a very specific set of challenges in a single context. As an extreme example of the latter, computational imaging techniques use specialized algorithms to process and interpret data from one specific instance of an imaging device.

Inherent to these methods are thus a set of tradeoffs surrounding the ability of these algorithms to generalize across contexts. These tradeoffs can impact how new methods can be robustly assessed and shared. We discuss here a number of considerations one should consider both in the development of such algorithms and in the deployment and dissemination thereof.

---

Further author information: (Send correspondence to G.M. and A.S.C.)

G.M.: E-mail: gmishne@ucsd.edu

A.S.C.: E-mail: adamsc@jhu.edu

## 2. THE PROMISE OF DEEP LEARNING?

One increasingly popular group of approaches in the analysis of functional optical imaging is to leverage advances in deep learning, where deep learning has been integrated into various points of the processing pipeline for functional imaging, from preprocessing, through inference to postprocessing, which we briefly review below.

One challenge deep learning methods have sought to address is denoising of the acquired video. Functional imaging suffers from imaging-dependent noise, where signal-to-noise (SNR) can vary across different neuronal components due to differences in the expression of the indicator. Identifying neural components, e.g., cell bodies, is enabled because the location of cells is fixed in the field-of-view and their activity results in fluorescence persisting over multiple frames. Extraction of spatial components and their corresponding time traces can be improved by incorporating a denoising step as part of preprocessing (following motion correction). Recently, general advances from deep learning<sup>3-5</sup> have been proposed for denoising calcium imaging. One class of methods, including DeepInterpolation,<sup>3</sup> SUPPORT<sup>6</sup> and DeepVID,<sup>7,8</sup> leverage the spatio-temporal consistency of the fluorescence signal to separate it from the noise. In essence, these methods train a neural network to predict a movie frame (for DeepInterpolation) or pixel (for SUPPORT) based on previous and past frames or the surrounding pixels in space and time, respectively. As independent noise cannot be predicted from the spatio-temporal neighborhood, it is thus effectively removed. A broader class of image-restoration methods are based on more traditional static architectures. CARE,<sup>9</sup> and the more recent XTC that specializes for synaptic imaging,<sup>10</sup> are based on a U-net architecture trained on high- and low-resolution images. These images can be obtained by more expensive measurements in vivo, or as in XTC, by creating training data ex-vivo using super-resolution microscopy.

The fundamental step for the analysis of functional imaging data is identifying spatial regions of interest (ROIs) and their corresponding time traces. While earlier methods mostly relied on regularized matrix decomposition,<sup>11-15</sup> recently, general deep learning methods for image segmentation have been adapted to the ROI extraction problem.<sup>16-22</sup> These are supervised approaches that require labeled training data in order to train a network to identify ROIs in new datasets.

Some networks tackle only the problem of identifying spatial components and ignore the temporal dimension. UNet2DS<sup>16</sup> is based on the fully convolutional UNet<sup>17</sup> model, which was developed for biomedical image segmentation. UNet2DS takes the mean image as an input and outputs the probability of a pixel belonging to a cell or background. UNet2DS, however, ignores important temporal information, leading to a difficulty in separating overlapping neurons. Furthermore, it is dependent on the summary image that may be biased towards cells that fire more often, or have high expression despite exhibiting no activity. Similar to UNet2DS, Cellpose<sup>22</sup> is similarly based on a UNet architecture, with an added global “style” designator to help differentiate different image classes, and so suffers the same drawbacks.

Other networks have added temporal statistics or activity to their design.<sup>18-20</sup> A (2+1)D convolutional neural network<sup>18</sup> was trained on spatiotemporal sliding windows and the output represented the probability of a pixel belonging to an ROI centroid. This approach better suppressed noisy detections compared to a 2D network that only took as input the time-averaged image, yet still only provided spatial components as outputs.

STNeuroNet<sup>19</sup> is a 3D convolutional neural network for neuron segmentation, trained on overlapping spatiotemporal blocks and outputs a binarized probability map. Shallow U-Net Neuron Segmentation (SUNS)<sup>20</sup> simplifies STNeuroNet’s architecture for speed and incorporates temporal information by applying a temporal matched filter and whitening to the input video. Both methods post-process the output probability maps by clustering to detect individual cells and aggregating detected cells across the maps. The design of these networks incorporates temporal information while efficiently reducing it in order to obtain fast processing speeds and low-latency inference in closed-loop experiments. Finally, DISCo (Deep learning, Instance Segmentation, and Correlations)<sup>21</sup> uses a combination of time-trace correlation-based pixel segmentation on a graph and a convolutional neural network to identify individual spatial profiles. While these methods incorporate temporal information in different ways, they only identify cells and do not address the issue of estimating time traces.

Finally, deep learning has also served in postprocessing to classify identified regions-of-interest (ROIs) as either true or false, based on human-annotated training data, to improve the precision of calcium imaging analysis methods, e.g., in CaImAn.<sup>12</sup>

On the practical side, training these networks might involve a lot of training data and time. The gain is the fast processing speed of these networks at test time. Critical to their success is 1) the availability of training data, 2) sufficient computing resources, and 3) the generalization of the trained system to new datasets. These impact the performance of deep learning models when deployed to unseen data in a new lab not encountered in training. As with general deep learning models, the exact expected biases of the black box system are unknown.

### 3. 10 LABS, 11 OPINIONS: THE OUT-OF-DISTRIBUTION PROBLEM.

One of the most frustrating situations is the inability of an algorithm that worked fantastically on one dataset but failed utterly on another. In neuroimaging—wherein surgical protocols, custom microscopes, and variability in environment can impact signal statistics from lab-to-lab—this effect is further frustrated when a method thoroughly validated on many datasets from one lab do not perform to the same degree in another lab. One reason can be traced back to an insidious type of overfitting to the distribution of one lab’s data, while even a similar lab might have data that may as well be *out of distribution (OOD)*.

OOD is a central challenge in machine learning more generally.<sup>23,24</sup> In particular, the fact that much of modern-day machine learning focuses on deriving much of what can be derived from the data itself, a central axiom is that the full dataset follows a static distribution. When, of the full distribution (say all videos of calcium imaging of GCaMP-6f transgenic mice in Hippocampus area CA1), only a biased subset is used to train the ML algorithm (say all videos of calcium imaging of GCaMP-6f transgenic mice in Hippocampus area CA1 taken at Johns Hopkins University) then the extra assumption is needed that the subset of the data follows the same distribution of the full data. If it does not then a different subset (say all videos of calcium imaging of GCaMP-6f transgenic mice in Hippocampus area CA1 taken at UC San Diego) will fail to leverage what the ML algorithm learned from the first subset. This effect is being continuously studied in more complex, modern, ML systems, however will continue to have an impact in how we can expect ML solutions to generalize across users.

One direct consequence of the “every lab is OOD” problem is that algorithms that learn too much from data can suffer drawbacks in dissemination. An obvious case is the use of deep learning. Deep learning fits highly over-parameterized models and is known to be able to interpolate within a distribution extremely well, while extrapolate very poorly to new domains. Thus, despite the fact that pre-trained networks are fast to run, OOD create a high level of uncertainty in using pre-trained networks accurately. In developing methods for functional imaging analysis, and especially for training deep learning models during method development, most use NeuroFinder<sup>25</sup> and/or the Allen Institute Mouse Brain Observatory.<sup>26</sup> However, both datasets suffer from both mislabeled data and missing labels, or erroneous time-trace estimates.<sup>27</sup> Thus limited and mislabeled training data creates a limited and potentially distorted *in distribution* that the deep learning model learns, which can affect performance in unexpected ways. Not all is lost, however, as fine-tuning models on new datasets can be a work-around for new data sets, albeit at the cost of further annotation or data collection. In fact, many labs have manually labeled data acquired from their own optical systems and in the brain regions they typically study, this barrier can be somewhat mitigated provided easy processes by which to use this data to fine-tune the model.

Thus, even though pre-trained networks are fast to run, these challenges create a high level of uncertainty in using pre-trained networks accurately. Individual labs can instead choose to annotate their own data to train networks from scratch, such that they are optimized for the imaging used locally. While this solves some of the aforementioned challenges (assuming care is taken in annotation), the training procedure can be computationally intensive relative to a lab’s computational capabilities.

Some methods explicitly seek to transcend such challenges by providing models that learn from data in an instance-specific setting. For example, matrix factorization and other such methods for source identification use more general assumptions of low-rank and non-negativity rather than learning full feature spaces as in deep learning. Danger, however, lies here as well. These methods also have multiple (numerous) parameters that are often fit in certain domains, or have regularizations that are fixed to work well for certain imaging contexts. For example, algorithms that constrain components to be contained in a small area in the image work excellently for somatic imaging across brain areas and species but fail for different imaging scales and non-compact or non-continuous morphologies (e.g., dendrites or widefield).

## 4. ONE PIPELINE DOES NOT FIT ALL.

The generalization constraints discussed create a fundamental foil in the field’s goal of creating highly disseminable and reproducible data pipelines that can work “out of the box”. Such a goal is highly appealing, as we should seek as much as possible to leverage the advances of others to optimize our data analysis. To realize this goal, however, two challenges must be met: robust assessment, and adaptive retraining.

The first of these challenges must face the fact that OOD means that the assessment performed of the pipeline and its constituent parts might not be correct for all labs/users. To enable new users to correctly choose an optimal computational approach, new approaches to validation must be developed. These validation approaches must promote data diversity and scale, rather than dissecting singular datasets. Promising avenues in this direction include the growing repositories of data in, e.g., the Dandi archive, and new simulation-based data generators.<sup>28</sup>

The second challenge lies in having the pipeline meet the users. Currently, the goal of robust dissemination has been to make the code as immutable as possible.<sup>29</sup> In some dimensions, this is the correct requirement, as one does not want the code itself to change. In others, the pipeline must be adaptive to identify when the approach is failing at processing a dataset and which modification must be taken to improve the approach. Simple versions of such methods would be simple parameter re-fitting (e.g., using BayesOpt) or fine-tuning of deep learning models.<sup>30</sup> More complex solutions should be able to swap and reorder stages in the pipeline. None of these is currently “out of the box” for users and require significant manual intervention to adapt pipelines.

## 5. CONCLUSIONS

The pace of neuroimaging analysis is rapid and unrelenting. Progress is often achieved within the context of specific labs tackling problems which, while critical to their work, would be invaluable to the community at large. Robust dissemination is thus key and must account for the OOD problem.

New efforts offer promise. Efforts at the Allen Institute, the International Brain Lab, DataJoint, and NeuroCAAS, amongst others, are striving to create stable pipelines may lead to easier access. These efforts can form central points for tools that assess and correct OOD problems as new users build up their neuroimaging needs.

## ACKNOWLEDGMENTS

G.M. was partially funded by a Kavli Institute for Brain and Mind (KIBM) Innovative Research Grant award. We would like to thank Michael Xie for helpful feedback.

## REFERENCES

- [1] Benisty, H., Song, A., Mishne, G., and Charles, A. S., “Review of data processing of functional optical microscopy for neuroscience,” *Neurophotonics* **9**(4), 041402–041402 (2022).
- [2] Charles, A. S., Falk, B., Turner, N., Pereira, T. D., Tward, D., Pedigo, B. D., Chung, J., Burns, R., Ghosh, S. S., Kebschull, J. M., et al., “Toward community-driven big open brain science: open big data and tools for structure, function, and genetics,” *Annual review of neuroscience* **43**, 441–464 (2020).
- [3] Lecoq, J., Oliver, M., Siegle, J. H., Orlova, N., Ledochowitsch, P., and Koch, C., “Removing independent noise in systems neuroscience data using deepinterpolation,” *Nature Methods* **18**(11), 1401–1408 (2021).
- [4] Li, X., Zhang, G., Wu, J., Zhang, Y., Zhao, Z., Lin, X., Qiao, H., Xie, H., Wang, H., Fang, L., et al., “Reinforcing neuron extraction and spike inference in calcium imaging using deep self-supervised denoising,” *Nature Methods* **18**(11), 1395–1400 (2021).
- [5] Weigert, M., Schmidt, U., Boothe, T., Müller, A., Dibrov, A., Jain, A., Wilhelm, B., Schmidt, D., Broaddus, C., Culley, S., et al., “Content-aware image restoration: pushing the limits of fluorescence microscopy,” *Nature methods* **15**(12), 1090–1097 (2018).
- [6] Eom, M., Han, S., Park, P., Kim, G., Cho, E.-S., Sim, J., Lee, K.-H., Kim, S., Tian, H., Böhm, U. L., et al., “Statistically unbiased prediction enables accurate denoising of voltage imaging data,” *Nature Methods* **20**(10), 1581–1592 (2023).

- [7] Liu, C., Platisa, J., Ye, X., Ahrens, A. M., Chen, I. A., Davison, I. G., Pieribone, V. A., Chen, J. L., and Tian, L., “Two-photon voltage imaging denoising by self-supervised learning,” in [*Neural Imaging and Sensing 2023*], **12365**, 13–14, SPIE (2023).
- [8] Platisa, J., Ye, X., Ahrens, A. M., Liu, C., Chen, I. A., Davison, I. G., Tian, L., Pieribone, V. A., and Chen, J. L., “High-speed low-light in vivo two-photon voltage imaging of large neuronal populations,” *Nature Methods*, 1–9 (2023).
- [9] Wei, Z., Lin, B.-J., Chen, T.-W., Daie, K., Svoboda, K., and Druckmann, S., “A comparison of neuronal population dynamics measured with calcium imaging and electrophysiology,” *PLoS computational biology* **16**(9), e1008198 (2020).
- [10] Xu, Y. K. T., Graves, A. R., Coste, G. I., Hugarir, R. L., Bergles, D. E., Charles, A. S., and Sulam, J., “Cross-modality supervised image restoration enables nanoscale tracking of synaptic plasticity in living mice,” *Nature Methods*, 1–10 (2023).
- [11] Pachitariu, M., Packer, A. M., Pettit, N., Dalglish, H., Hausser, M., and Sahani, M., “Suite2p: beyond 10,000 neurons with standard two-photon microscopy,” *bioRxiv* (2016).
- [12] Giovannucci, A., Friedrich, J., Gunn, P., Kalfon, J., Brown, B. L., Koay, S. A., Taxidis, J., Najafi, F., Gauthier, J. L., Zhou, P., Khakh, B. S., Tank, D. W., Chklovskii, D. B., and Pnevmatikakis, E. A., “CaImAn an open source tool for scalable calcium imaging data analysis,” *Elife* **8**, e38173 (2019).
- [13] Mishne, G. and Charles, A. S., “Learning spatially-correlated temporal dictionaries for calcium imaging,” in [*IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP-2019)*], 1065–1069, IEEE (2019).
- [14] Charles, A. S., Cermak, N., Affan, R. O., Scott, B. B., Schiller, J., and Mishne, G., “Graft: Graph filtered temporal dictionary learning for functional neural imaging,” *IEEE Transactions on Image Processing* **31**, 3509–3524 (2022).
- [15] Mukamel, E. A., Nimmerjahn, A., and Schnitzer, M. J., “Automated analysis of cellular signals from large-scale calcium imaging data,” *Neuron* **63**(6), 747–760 (2009).
- [16] Klibisz, A., Rose, D., Eicholtz, M., Blundon, J., and Zakharenko, S., “Fast, simple calcium imaging segmentation with fully convolutional networks,” in [*Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*], 285–293, Springer (2017).
- [17] Ronneberger, O., Fischer, P., and Brox, T., “U-net: Convolutional networks for biomedical image segmentation,” in [*International Conference on Medical image computing and computer-assisted intervention*], 234–241, Springer (2015).
- [18] Apthorpe, N., Riordan, A., Aguilar, R., Homann, J., Gu, Y., Tank, D., and Seung, H. S., “Automatic neuron detection in calcium imaging data using convolutional networks,” *Advances in neural information processing systems* **29**, 3270–3278 (2016).
- [19] Soltanian-Zadeh, S., Sahingur, K., Blau, S., Gong, Y., and Farsiu, S., “Fast and robust active neuron segmentation in two-photon calcium imaging using spatiotemporal deep learning,” *Proceedings of the National Academy of Sciences* **116**(17), 8554–8563 (2019).
- [20] Bao, Y., Soltanian-Zadeh, S., Farsiu, S., and Gong, Y., “Segmentation of neurons from fluorescence calcium recordings beyond real time,” *Nature Machine Intelligence*, 1–11 (2021).
- [21] Kirschbaum, E., Bailoni, A., and Hamprecht, F. A., “Disco: Deep learning, instance segmentation, and correlations for cell segmentation in calcium imaging,” in [*International Conference on Medical Image Computing and Computer-Assisted Intervention*], 151–162, Springer (2020).
- [22] Stringer, C., Wang, T., Michaelos, M., and Pachitariu, M., “Cellpose: a generalist algorithm for cellular segmentation,” *Nature methods* **18**(1), 100–106 (2021).
- [23] Geisa, A., Mehta, R., Helm, H. S., Dey, J., Eaton, E., Dick, J., Priebe, C. E., and Vogelstein, J. T., “Towards a theory of out-of-distribution learning,” *arXiv preprint arXiv:2109.14501* (2021).
- [24] Fort, S., Ren, J., and Lakshminarayanan, B., “Exploring the limits of out-of-distribution detection,” *Advances in Neural Information Processing Systems* **34**, 7068–7081 (2021).
- [25] Berens, P., Theis, L., Stone, J. S., Sofroniew, N., Tolias, A., Bethge, M., and Freeman, J., “Standardizing and benchmarking data analysis for calcium imaging,” in [*COSYNE*], (2017).

- [26] “Allen institute brain observatory whitepaper.” <https://help.brain-map.org/display/observatory/Documentation>. Accessed: 2019-09-27.
- [27] Gauthier, J. L., Koay, S. A., Nieh, E. H., Tank, D. W., Pillow, J. W., and Charles, A. S., “Detecting and correcting false transients in calcium imaging,” *bioRxiv* , 473470 (2018).
- [28] Song, A., Gauthier, J. L., Pillow, J. W., Tank, D. W., and Charles, A. S., “Neural anatomy and optical microscopy (naomi) simulation for evaluating calcium imaging methods,” *Journal of Neuroscience Methods* **358**, 109173 (2021).
- [29] Abe, T., Kinsella, I., Saxena, S., Buchanan, E. K., Couto, J., Briggs, J., Kitt, S. L., Glassman, R., Zhou, J., Paninski, L., et al., “Neuroscience cloud analysis as a service: An open-source platform for scalable, reproducible data analysis,” *Neuron* **110**(17), 2771–2789 (2022).
- [30] Pachitariu, M. and Stringer, C., “Cellpose 2.0: how to train your own model,” *Nature methods* **19**(12), 1634–1641 (2022).