PERSPECTIVE

AI analysis of super-resolution microscopy: Biological discovery in the absence of ground truth

Ivan R. Nabi[1,2] ![ORCID], Ben Cardoen[3] ![ORCID], Ismail M. Khater[3,4] ![ORCID], Guang Gao[1] ![ORCID], Timothy H. Wong[1] ![ORCID], and Ghassan Hamarneh[3] ![ORCID]

Super-resolution microscopy, or nanoscopy, enables the use of fluorescent-based molecular localization tools to study molecular structure at the nanoscale level in the intact cell, bridging the mesoscale gap to classical structural biology methodologies. Analysis of super-resolution data by artificial intelligence (AI), such as machine learning, offers tremendous potential for the discovery of new biology, that, by definition, is not known and lacks ground truth. Herein, we describe the application of weakly supervised paradigms to super-resolution microscopy and its potential to enable the accelerated exploration of the nanoscale architecture of subcellular macromolecules and organelles.

Introduction

Artificial intelligence (AI), "the capability of computer systems or algorithms to imitate intelligent human behavior" (Merriam-Webster.com, 2024), is increasingly present in our everyday lives. Recently, generative language and image AI models (ChatGPT, Sable Diffusion, Midjourney) created a storm of interest, with users challenged to differentiate between AI-generated and human-generated content. Users are both amazed by AI and disappointed by its occasional surprising failures and hallucinations. Like ChatGPT, current AI methods are incredible tools that are not always correct, necessitating expert "ground truth" validation.

"Ground truth is information that is known to be real or true, provided by direct observation and measurement (i.e., empirical evidence) as opposed to information provided by inference" (Wikipedia, 2024). For imaging-based AI applications, validation uses ground truth image annotations to test computer identification of images and image components, from distinguishing cats from dogs to real-world examples, such as medical imaging–based diagnosis or object recognition for self-driving cars. Machine learning models are trained and tested on acquired datasets. Supervised machine learning uses ground truth annotations to train new image prediction methods and validate predictions.

The most straightforward approach to train machine learning to identify objects within an image is to perform strong supervision. The machine learning model is trained on a curated dataset of images and their corresponding segmentation masks (dense, pixel/voxel-level annotations); class labels are assigned to every pixel in the image. However, when applying AI to novel bioimaging modalities, such as super-resolution microscopy

(SRM) that breaks Abbe's diffraction limit, even experts are challenged to define what is real within these images. As opposed to labeling street signs, vehicles, pedestrians, etc., for self-driving cars, for which we are all experts in principle, the time and financial cost of having expert biologists annotate images, at the pixel-, voxel-, or localization-level, can be astronomical. For SRM, strong supervision based on complete annotation is rarely feasible, with the noted exception of simulated data or phantom data such as DNA origami. While the application of machine learning to SRM has tremendous potential to address unanswered questions and discover novel biology, ground truth expert annotation of image content is often infeasible. Annotation also relies on the assumption that experts know all there is to know about the underlying biology that these images capture, an assumption that may not always hold true. This is particularly the case for novel imaging modalities (such as SRM) whose primary purpose is to expand the boundaries of our understanding of biology (Fig. 1 A).

The "strength" of supervision describes the detail of knowledge that is known about and provided with the data to provide the training corpus. For example, in the case of classifying objects in an image, absence of supervision would mean there is no information provided about the class of each image, or about the objects depicted by the image. Weak supervision would describe a situation where we present information that, say, a dog or cat, is present in an image, but not where. Strong supervision would be a case where the machine learning method is provided with the knowledge of the exact location and outline of each identified object. When the supervisory signal is created from the input image itself without any annotation burden, we have self-supervision. For example, providing rotated images along with

---

[1]Department of Cellular and Physiological Sciences, Life Sciences Institute, University of British Columbia, Vancouver, Canada; [2]School of Biomedical Engineering, University of British Columbia, Vancouver, Canada; [3]School of Computing Science, Simon Fraser University, Burnaby, Canada; [4]Department of Electrical and Computer Engineering, Faculty of Engineering and Technology, Birzeit University, Birzeit, Palestine.

Correspondence to Ivan R. Nabi: irnabi@mail.ubc.ca; Ghassan Hamarneh: hamarneh@sfu.ca.
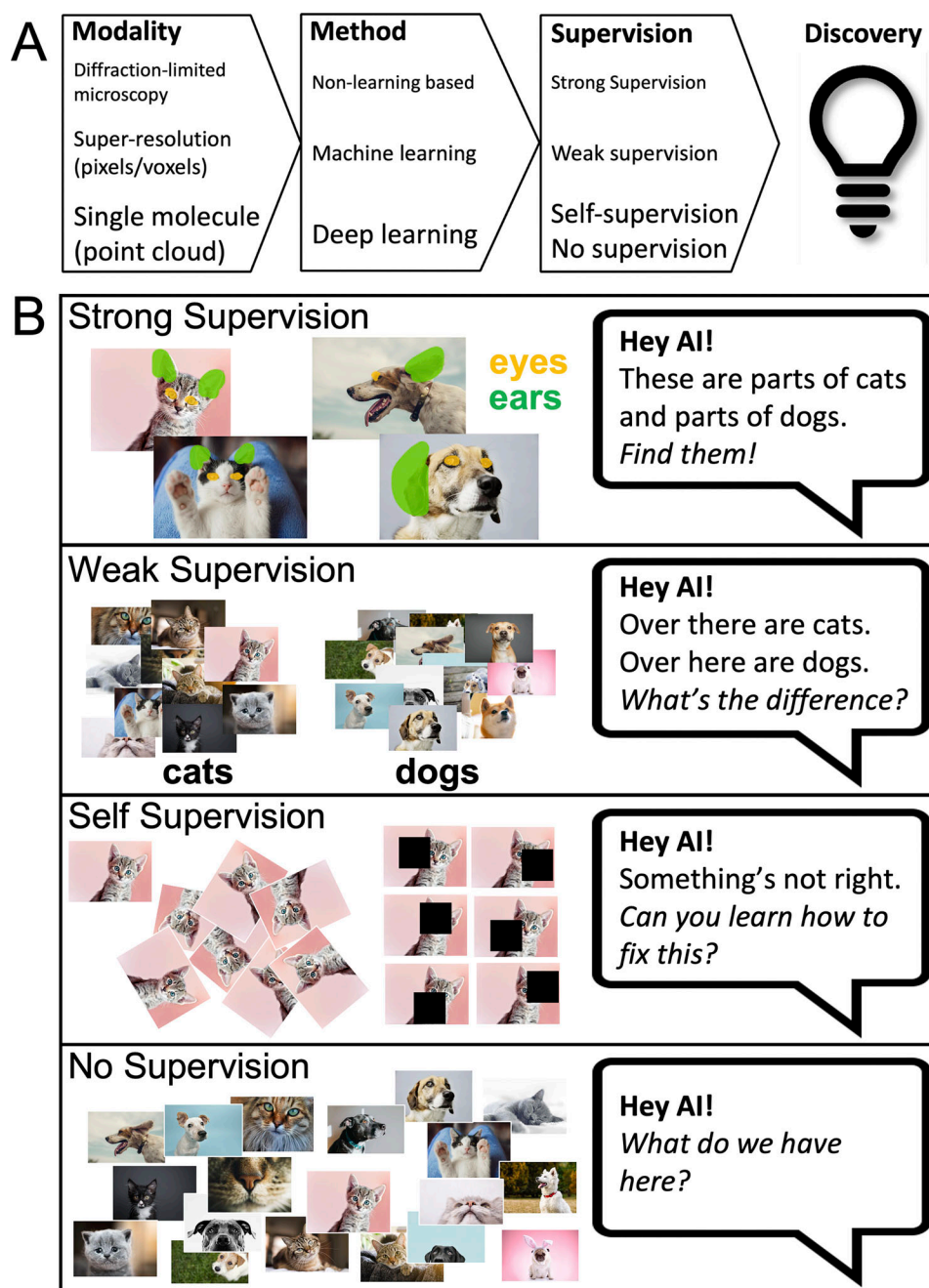
Figure 1. **AI and SRM. (A)** Continued and parallel improvement in SRM hardware and AI-based analysis will lead to new and better approaches to define the organization and dynamics of subcellular structures. For instance, analysis of higher resolution single-molecule approaches, such as MinFlux (2 nm lateral resolution) (Balzarotti et al., 2017) with deep learning and self-supervision modalities should lead to novel insights into molecular structure in whole cell analyses. **(B)** Different levels of supervision include strong supervision in which each object or even each pixel has annotated information that can be leveraged. This is rare in SRM as discovery implies absence of such data. Weak supervision occurs when partial information, often at the image level, is available, such as the presence or absence of a type of object, but not its location. In SRM use cases, this would be cell line, treatment, gene expression etc. Self-supervision is a hybrid form, where the model learns to operate on images using indirect information, such as rotations. This is then followed by a fine-tuning stage with small amounts of strongly supervised data. Unsupervised, i.e., no supervision, indicates a complete absence of such metadata.

an automatically generated annotation "angle of rotation" per image, or providing an image with added noise to it, along with the known original image, allows the machine learning method to learn how to cancel the rotation or remove the noise. By doing so it is possible to train a model to encode semantics of the objects in the images and leverage such knowledge to other tasks.

While some fine-tuning is invariably needed, a key advantage of self-supervised learning is that far less strongly supervised data is required for equal or better performance. The strength of supervision spans a spectrum, e.g., providing contours around each object is stronger than providing bounding boxes, which, in turn, is stronger than providing only the count of objects. It is

also possible to mix different forms of supervision, e.g., a machine learning method may leverage self-supervision, as well as some images with a variety of weak annotations, and other images with strong supervision. Unsupervised methods do not use any annotations (e.g., clustering). Semisupervised methods combine supervised and unsupervised datasets to train machine learning models (Fig. 1 B).

In contrast to the high-annotation burden of strong supervision, weak supervision assigns a single class label to the whole image rather than to every pixel (i.e., identifying a dog without specifying which pixels are part of the dog). An example of weak supervision in SRM is the provision of a training set of SRM images, along with a cell group or condition acting as the image class label. Dictating the image-level label does not involve annotating which pixels of the image are manifestations of that cell group or condition. Weakly supervised object detection and localization (i.e., training AI to find the object-specific locations by training on data labeled at the image level) has been popular for natural images (Zhou, 2017), and applied to biomedical images, such as MRI, histopathology, and confocal microscopy (D'Alonzo et al., 2021; Liu et al., 2022a; Xu et al., 2014). We suggest that this form of supervision is suited to SRM, given that the goal is to identify and characterize those subcellular structures that vary across experimental conditions (cell lines, gene expression, mutations, infection, and drug treatment).

For biological research exploring the subcellular space, including macromolecules, organelles, and cytoskeletal structures, ground truth has long been defined by high-resolution approaches, in particular electron microscopy (EM). EM provides exceptional resolution, <1 nm, and pioneering EM work from the 1950s revolutionized our understanding of cellular organelles, defining the morphological underpinnings of the cell that are now textbook cell biology. EM has provided unprecedented 3D views of the cell, such that we now understand that the cytoplasm is a complex and dense array of membrane-bound and nonmembrane-bound organelles organized amongst cytoskeletal elements. Indeed, Golgi's discovery of the Golgi apparatus in neurons of silver-stained nervous tissue in the late 1800s was challenged as an artifact until it was confirmed by EM in the 1950s (Bentivoglio and Mazzarello, 1998). This represents perhaps the first example of EM being used as ground truth validation.

Structural biology and SRM: Closing the mesoscale gap
Advances in structural biology technologies provide powerful tools to decipher molecular structures at the atomic level. X-ray crystallography and nuclear magnetic resonance (NMR) spectroscopy generated a wealth of molecular structural data that propelled the growth of the Protein Data Bank in the past 40 years. This enabled the development of coevolution and AI-based algorithms such as AlphaFold2 and RoseTTAFold that are capable of accurately predicting structures of proteins from amino acid sequences alone for many targets (Jumper et al., 2021). The recent resolution revolution in cryo-EM made it possible to use the single-particle approach or subtomogram-averaging approach to determine, at atomic resolution, the structures of a broad range of macromolecular complexes from

ribosomes to intact virions, filling the gap between subnanoscale and the mesoscale (Ke et al., 2020; Rozov et al., 2019; Zhang et al., 2010). However, with the exception of subtomogram averaging, these conventional structural biology approaches are primarily used for characterizing highly purified samples that have been removed from their native subcellular or cellular environment (Fig. 2). Further, EM is limited to fixed cell analysis, requires extensive sample preparation, and is time intensive. As such, while EM can provide ground truth validation, it cannot, however, provide the ground truth annotation required for strongly supervised learning training in super-resolution.

More rapid image acquisition, improved antibody labeling efficiency, and availability of fluorescent proteins for live cell analysis make fluorescent microscopy the method of choice for analyzing molecular distribution and dynamics in whole cell 3D volumes (Lippincott-Schwartz, 2011). Confocal microscopy improved axial resolution, to about 500 nm, and facilitated the use of dynamic photobleaching assays (i.e., fluorescence recovery after photobleaching [FRAP] and fluorescence loss in photobleaching [FLIP]) making it a routine tool to study subcellular structure and dynamics in many labs (Diekmann and Hoischen, 2014). While improving axial resolution, confocal microscopy does not address the diffraction barrier that limits the lateral resolution of fluorescence microscopy to about 200–250 nm. SRM, defined as microscopy approaches that break the diffraction limit of light, encompasses a number of distinct methodologies enabling nanoscale fluorescent microscopy and provides novel insight into subcellular structure and dynamics (Sahl et al., 2017; Sydor et al., 2015). With cryoEM now able to analyze whole viruses at subnanometre resolution (Zhang et al., 2010) and SRM approaches, such as MinFlux able to image down to 2 nm resolution (Balzarotti et al., 2017), the mesoscale gap is being bridged from both sides (Goodsell et al., 2020; Liu et al., 2022b) (Fig. 2). Indeed, recent live cell MinFlux imaging of kinesin stepping along microtubules highlights the potential of SRM for dynamic analysis of molecular structure (Deguchi et al., 2023).

The extensive application of AI to microscopy image acquisition and analysis, potentially leading to intelligent microscopes, has been extensively reviewed (Liu et al., 2021; Morgado et al., 2024; Pylvänäinen et al., 2023). Single-molecule reconstruction has been an active area of research and includes the application of a convolutional neural network to interpret single molecule localizations directly from SMLM images (DeepSTORM, DBlink) and implementation of anti-bunching to resolve closely spaced emitters, a major issue in single-molecule imaging approaches (Kudyshev et al., 2023; Nehme et al., 2018; Saguy et al., 2023; Yang et al., 2021). Recent application of AI to reconstruction methods has the potential to reduce noise, improve spatial and temporal resolution, and automate high-throughput super-resolution and live imaging (Fu et al., 2023; Priessner et al., 2024; Qiao et al., 2023). The application of AI and machine learning to SRM has been proposed to be the next step in advancing nanomedicine development, drug discovery, and antiviral research (Li et al., 2024a; Ortiz-Perez et al., 2024; Petkidis et al., 2023). In this Perspective, we focus on the specific
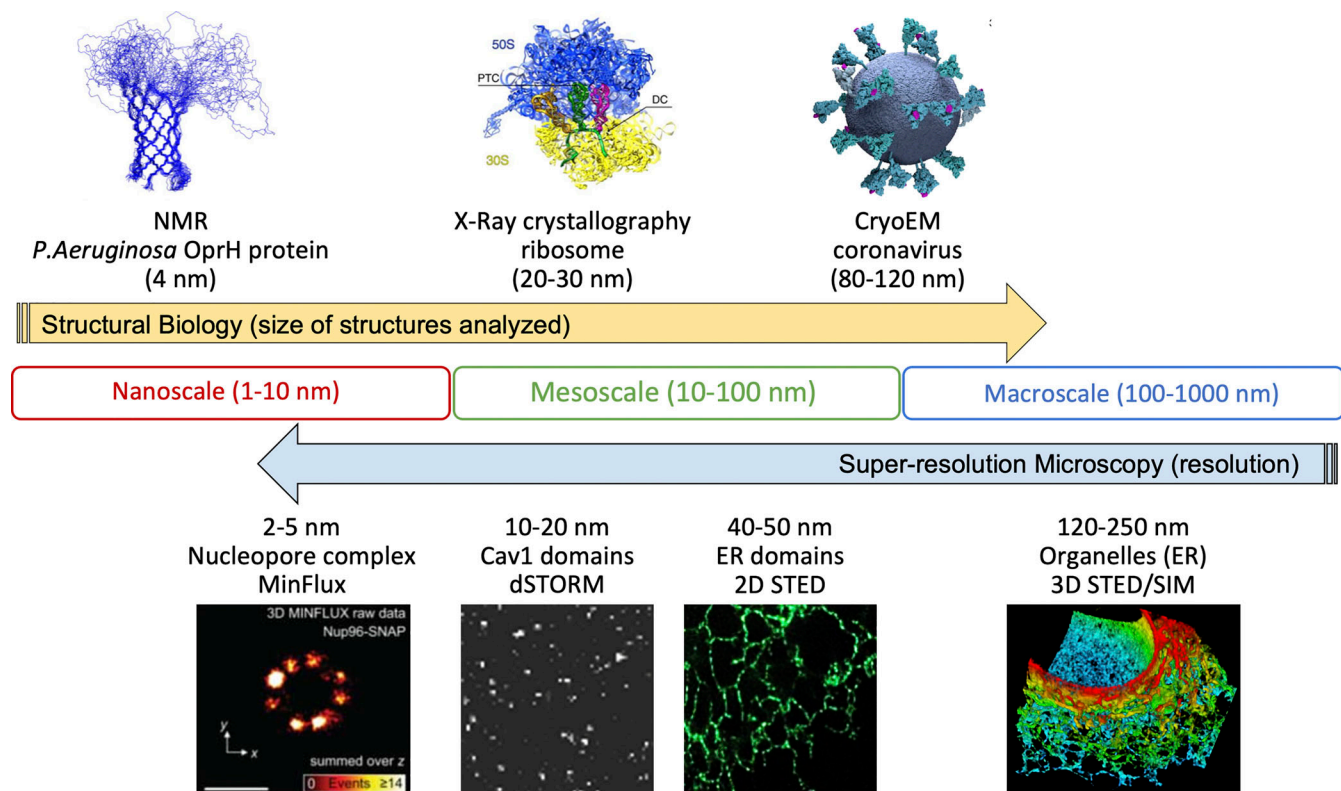
Figure 2. **Super-resolution microscopy and structural biology bridge the mesoscale domain.** Structural biology approaches (NMR, X-ray crystallography, and cryoEM) with angstrom level resolution analyze structures as large as the 80–120 nm coronavirus. Whole-cell super-resolution microscopy approaches such as structured illumination (SIM), stimulated emission depletion (STED), single-molecule localization microscopy (SMLM), and MinFlux (images from Edrington et al., 2011; Gao et al., 2019; Ke et al., 2020; Khater et al., 2018; Rozov et al., 2019; Schmidt et al., 2021) enable whole cell analysis with increasing resolution broaching the nanoscale.

application of AI to acquire semantic insight, new biology, from SRM.

Harnessing the power of AI for semantic insight from SRM

Pixel- and voxel-based SRM approaches, such as structured illumination (SIM) and stimulated emission depletion (STED) microscopy, provide increasingly detailed views of the subcellular space. However, they do not expose the underlying internal construction of subcellular structures. Single-molecule localization microscopy (SMLM) approaches, on the other hand, based on analysis of stochastic blinking of isolated fluorophores (Betzig et al., 2006), do not generate images but rather an event list of localizations, or a point cloud. Starting from a point cloud, one can create networks that in turn are well suited to graph-based analysis by machine and deep learning approaches. Despite this well-established network modeling approach, it has been common for many SMLM users to transform the point cloud data generated by SMLM into a pixelated image (Ruszczycki and Bernas, 2018). However, the end user should be aware of the pitfalls in these conversions (Ruszczycki and Bernas, 2018), whether for visualization or quantification.

Indeed, increasing the resolution of pixelated SRM images is akin to providing more enhanced detail, but only of the outer structure of a building (Fig. 3). Bridging the mesoscopic gap from diffraction-limited fluorescent microscopy to structural biology via SRM therefore requires visualizing inner structure or molecular architecture. However, in contrast to buildings where the design is generated before construction, the architecture of mesoscopic macromolecular structures and organelles is unknown, at best incomplete, and is what biological research endeavors to uncover. Fundamentally, novel biological discoveries lack ground truth (Fig. 3).

While SRM may have no definitive ground truth, the actual biology does not exist in isolation. SRM explores spatial and temporal data that, while yet to be described, occurs in the context of an existing accepted body of knowledge obtained by lower-resolution confocal microscopy, higher-resolution EM, as well as biochemical and structural analyses. SRM datasets, whether point cloud or pixel/voxel-based, generate large data sets including 3D information, time, and multiple channels at nanometer scales. Machine learning identification of patterns and differences from these large datasets can provide novel insight into subcellular structure and molecular architecture.

Machine learning approaches have been exploited to obtain biological insight from both pixel and point cloud super-resolution data. Machine Learning Structured Illumination Microscopy (MiLeSIM) used a supervised machine learning–based classifier to extract shape and size features in different strains of live attenuated influenza virus vaccines for high throughput imaging and assessment of viral production (Laine et al., 2018).

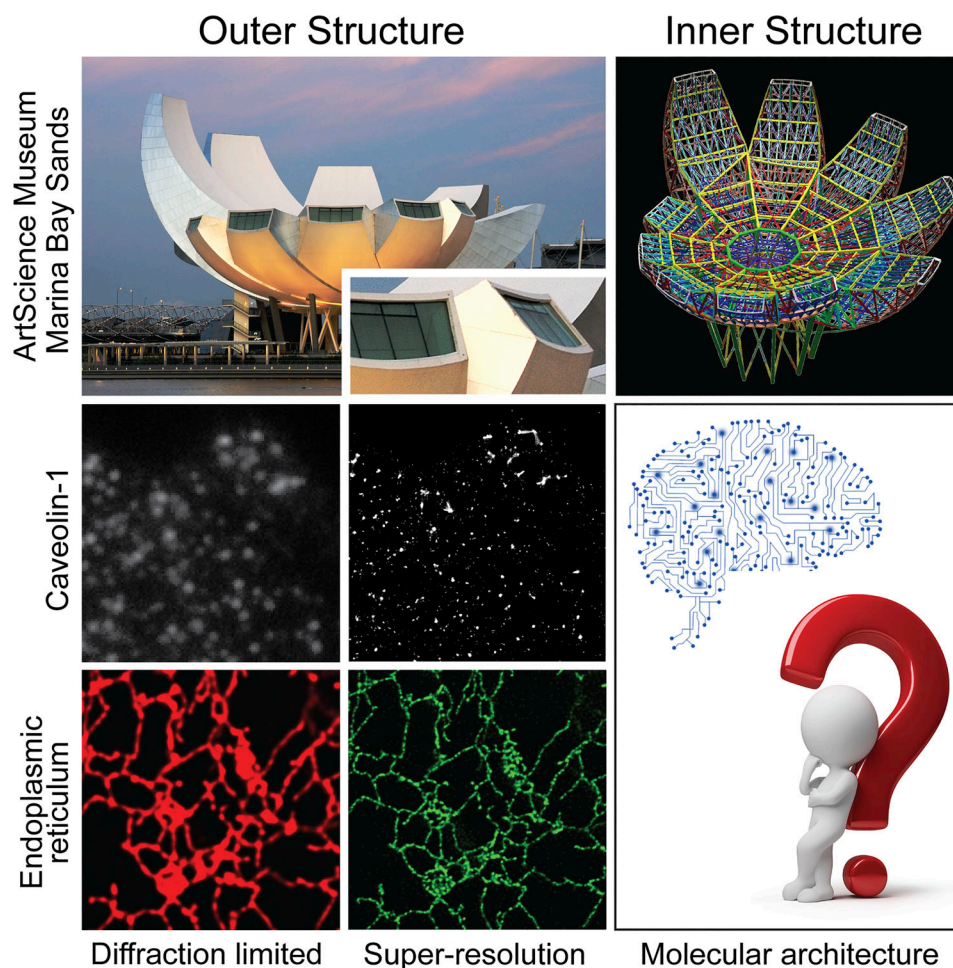## Outer Structure — Inner Structure

Figure 3. **Molecular architecture by SRM.** The top row shows outer (left) and inner (right) structure views of the ArtScience Museum at Marina Bay Sands. Below, pixel-based representation of SRM data provides unprecedented high-resolution views of subcellular structures, such as caveolin-1 domains, (wide-field on left versus SMLM on right) or the endoplasmic reticulum (confocal in red versus STED in green) but remain analogous to enlarged, more detailed views of the external face, or outer structure, of buildings (inset). Ongoing AI-based semantic analysis of SRM will provide the means to explore the design basis, or molecular architecture, of subcellular macromolecules and organelles.

SRM-based image-level class labels of endoplasmic reticulum (ER) were used to train a deep learning model to distinguish between Zika-infected and non-infected cells and showed that discriminating regions correspond to tubular matrix ER morphology (Long et al., 2020). Using a convolutional neural network model, SMLM image stacks are inputted to directly measure molecular diffusion in supported lipid bilayers (Park et al., 2023). Neural network extraction of features from nearest-neighbor distance-derived data identified cluster segmentation in SMLM point cloud space of C-terminal src kinase (CSK) or the associated PAG protein clustering on the cell membrane of T-cells (Williamson et al., 2020). Other advances involve the development of more generally applicable SMLM analysis tools, such as SuperResNET and SEMORE (SEgmentation and MORphological fingerprinting), to extract biological features from protein clusters imaged by SMLM (Bender et al., 2024; Khater et al., 2018; Li et al., 2024b, *Preprint*; Wong et al., 2024, *Preprint*).

Pairing machine learning with other biological information has also been applied to SMLM imaging to develop new insight.

LocoMoFit applies a maximum likelihood estimation that fits a provided model to a 3D point cloud structure (Wu et al., 2023). By fitting models, developed based on established structures, to individual coat structures obtained by SMLM, Locomofit supported the idea of a novel cooperative curvature model for clathrin endocytosis (Mund et al., 2023). Another approach used deep learning to pair single-molecule imaging of 3D chromatin structure using fluorescent oligo hybridization DNA probes with RNA expression measured by RNA FISH within the same cell and thereby predicted transcriptional states from the 3D structure of chromatin (Rajpurkar et al., 2021). A workflow developed for 3D reconstruction from 2D SMLM produced a 3D model of the centriole using iterative multi-reference refinement and dual color SMLM imaging between combinations of centriolar proteins allowed for novel insight into protein organization within the centriole (Sieben et al., 2018).

Applying machine learning to SRM requires approaches that do not depend on subcellular object-level ground truth for validation or algorithm training. Indeed, our AI-based image

analysis approach for biological discovery and validation has been to avoid pixel-/voxel-/localization-level annotation and use ground truth label annotations of groups of images of cells (e.g., wild-type versus mutant; over- versus under-expression; and infected versus uninfected cells) (Cardoen et al., 2022; Khater et al., 2018; Long et al., 2020; Saberian et al., 2022), each with prior knowledge of their biological features and functions. AI trained using reliable differential group labels is then used to predict pixel-/voxel-/localization labels. Key to weakly supervised approaches is the trustworthiness of labels, corresponding to phenotypic changes in the images confirmed by other modalities. Prediction of known biological features instills trust in AI for novel biological discovery (Fig. 4).

Fundamentally, within this context, weak supervision-inspired approaches for SRM could be based on three key principles:

1 Selection of trustworthy group labels grounded in biology;
2 AI-based identification of differential features across groups;
3 Use of a priori biological knowledge to corroborate AI-based discovery.

AI identification of novel biology

Current weakly supervised approaches are frequently characterized by modularly designed pipelines that, for example, require denoising, segmentation, and classification before reporting final results. Development of end-to-end approaches that take in the raw data and report the desired output without explicit intermediate stages can be beneficial when intermediate stages, such as segmentation, are either infeasible or too costly. However, end-to-end methods can require a final "fitting" or calibration stage, where the method's output is rescaled to correspond with validation data. Here, we present examples of weakly supervised approaches that incorporate modular design and end-to-end approaches.

Discovering a novel caveolin-1 scaffold domain

The protein caveolin-1 (CAV1) is the coat protein for 50–100 nm plasma membrane invaginations called caveolae. Caveolae formation requires the adaptor protein Cavin-1; in the absence of Cavin-1, CAV1 forms functional surface domains called scaffolds (Hill et al., 2008; Lajoie et al., 2009). While caveolae invaginations are morphologically distinct by EM, flat CAV1-positive domains are difficult to detect morphologically by EM and their identification by antibody labeling suffers from the poor antigenicity of most EM approaches. By diffraction-limited confocal microscopy, both caveolae and scaffolds present punctate surface labels; differentiating the two by the presence of Cavin-1 assumes Cavin-1 selectively associates with caveolae, which may not be correct (Khater et al., 2019a). Definitive identification of the two structures is problematic. The PC3 prostate cancer cell line, expressing elevated levels of CAV1 in the genetic absence of Cavin-1 and, therefore, of caveolae, together with PC3 cells transfected with Cavin-1 and expressing caveolae (Hill et al., 2008), provide cellular group labels for the identification of caveolae from SMLM data sets by machine learning.

A computational pipeline, SuperResNET (Fig. 4), inputs 3D single molecule point cloud data from dSTORM (direct stochastic optical reconstruction microscopy) labeling of CAV1, merges localizations below the resolution limit (20 nm), and filters localizations relative to randomized distributions. Feature-based cluster analysis of interactions below 80 nm identified two groups of blobs in PC3 cells and found two additional groups in Cavin-1 transfected PC3 cells, one of which was significantly larger than the PC3 groups, therefore corresponding to caveolae (Khater et al., 2018). Our validation of AI trained on discriminative group labels (cell lines), in this case, took the form of successfully identified clusters corresponding to the missing component (caveolae) in one of the groups (Khater et al., 2018, 2019b) (Fig. 4). Network analysis also identified three non-caveolar CAV1 scaffolds including small S1A scaffolds corresponding to 8S CAV1 complexes whose structure was recently reported by cryoEM (Han et al., 2020; Khater et al., 2018). Caveolae clusters consisted of 12-14 8S CAV1 complexes, matching the dodecahedral structure for caveolae reported by cryoEM (Khater et al., 2019b; Stoeber et al., 2016). The approach itself is therefore validated based on the identification of known structures: 8S CAV1 complexes and caveolae.

Beyond validating known biology (absence of caveolae in PC3 cells), network analysis of single-molecule SRM led to the identification of previously unidentified non-caveolar CAV1 scaffolds (Khater et al., 2018). Modularity analysis, in which clusters are broken down into smaller more closely associated clusters, showed that 8S CAV1 complexes combine to form caveolae as well as intermediate scaffold structures. These include previously undescribed 8S CAV1 complex dimers (S1B scaffolds) as well as larger hemispherical S2 scaffolds (Khater et al., 2019b) (Fig. 4). Intermediate scaffold structures are supported by the presence of a shoulder on the 8S CAV1 peak in fractionation studies (Hayer et al., 2010), by STED using belief-theory based weakly supervised object detection (Cardoen et al., 2022), and recent identification of CAV1 invaginations in the absence of cavin-1 called dolines (Lolo et al., 2023).

Identifying subpixel resolution riboMERCs

More recently, we developed a subpixel resolution approach to detect membrane contact sites, where the membrane of two organelles approach to within 10–30 nm (Helle et al., 2013). Mitochondria–ER contacts (MERCs) have been well-characterized by EM; however, their analysis by fluorescent microscopy is challenged by the fact that the distance between the two organelles (10–60 nm) is smaller than diffraction limited resolution and even that of 3D super-resolution approaches such as STED and SIM (Scorrano et al., 2019). To address this, we built on previous work (Cardoen et al., 2020, 2022) using Laplacian detection of local intensity changes to detect regions in 3D STED super-resolution images where the intensity of ER and mitochondria change in tandem. Independent of image segmentation, MCS-DETECT sensitively and robustly detects membrane contact sites independently of variations in local signal intensity or background (Fig. 4) (Cardoen et al., 2024).

However, be it diffraction-limited colocalization or subpixel MCS-DETECT detection of nanometer scale contact sites,
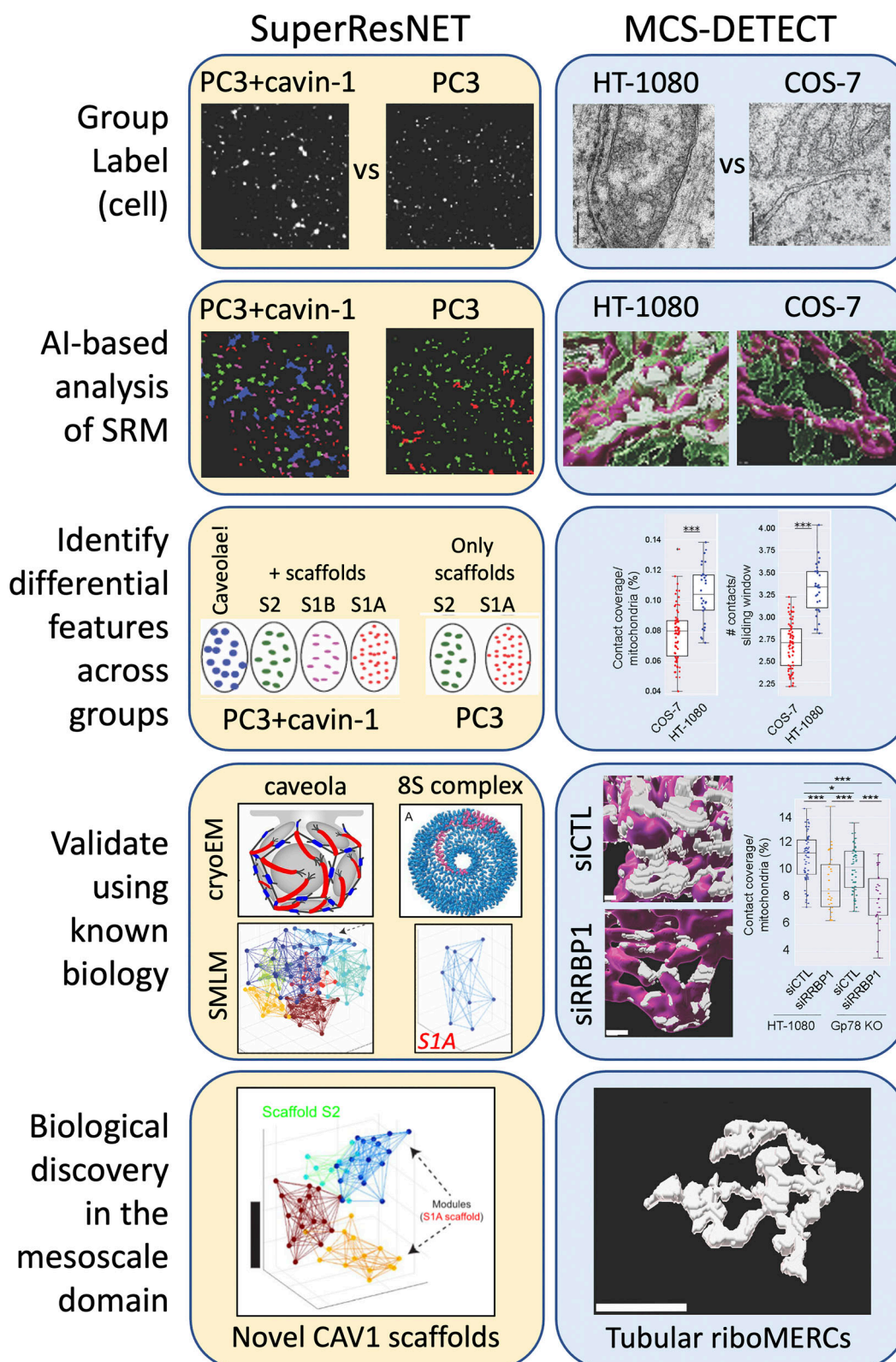
Figure 4. **The weakly supervised paradigm for AI-based semantic analysis of SRM datasets.** In the absence of pixel or object level ground truth (strong supervision), prior biological knowledge defines group labels for weakly supervised training of SRM datasets to identify group-specific protein structures. Left: Caveolae expression is known to require CAV1 and the adaptor protein cavin-1 (Anderson, 1998; Hill et al., 2008). PC3 cells lack cavin-1 and therefore caveolae; native and Cavin-1 transfected PC3 cells therefore provided group labels for weakly supervised network analysis: SuperResNET. Using a proximity threshold that most significantly distinguished point clouds of these two biologically distinct groups, we then segmented clusters of points into blobs (structures or objects), based on 28 features describing size, shape, topology and network measures. Four groups of blobs were found in Cavin-1 expressing PC3 cells, one of which was significantly larger than either of the two groups found in PC3 cells, therefore corresponding to caveolae. Structural correspondence of identified

structures to known biology, such as 8S complexes and caveolae whose structure has been determined by cryoEM (Han et al., 2020; Stoeber et al., 2016), validates the approach and enables identification of novel structures such as hemi-spherical CAV1 S2 scaffolds (Khater et al., 2018, 2019b). Right: HT-1080 and COS-7 cell lines differ in that only HT-1080 express elongated ribosome-studded mitochondria-ER contact sites (riboMERCS). Based on the differential expression of riboMERCs between these two cell lines, we developed a novel segmentation-free algorithm able to reconstruct MERCS, MCS-DETECT, that was optimized by maximizing the differential result between both cell lines (Cardoen et al., 2024). Feature analysis of MERCs matches differences between the two cell lines based on EM (i.e., ground truth), and the approach is validated by knockdown of a known riboMERC tether, RRBP1 (Hung et al., 2017). MCS-DETECT led to identification of extended tubular riboMERCs.

---

validation that contact sites are accurately detected remains challenging. As such, we tuned MCS-DETECT using a group label comparison of HT-1080 cells known to express a distinct type of ribosome-studded MERC (riboMERCs) (Wang et al., 2015) and COS-7 cells that do not express riboMERCs based on quantitative EM validation (Cardoen et al., 2024). This weakly supervised approach demonstrated that riboMERC expression is controlled by the RRBP1-SYNJ2BP tether (Hung et al., 2017) and that ribo-MERC size is regulated by the expression of the Gp78 ubiquitin ligase. MCS-DETECT further identified a convoluted tubular morphology for Gp78-dependent riboMERCs (Cardoen et al., 2024), similar to but less extended compared with wrappER riboMERCs described in the liver by 3D EM tomography (Anastasia et al., 2021). Using group labels based on known biology (i.e., EM), MCS-DETECT has defined a distinct morphology for riboMERCs and identified how their expression and size are controlled. MCS-DETECT, as a probabilistic reconstruction algorithm omitting segmentation, is an example of the move toward end-to-end approaches. However, it still requires specifically designed preprocessing filters to exclude false positives. In addition, the resulting reconstructed membrane contacts are not classified, so MCS-DETECT is not yet a complete end-to-end approach.

MERCs are functionally diverse, involved in ER–mitochondrial calcium exchange and regulation of mitochondrial metabolism, phospholipid, and sterol biosynthesis, as well as induction of apoptosis (Rowland and Voeltz, 2012). MERC formation is controlled by numerous tethers (Herrera-Cruz and Simmen, 2017), suggesting that there may very well be multiple subclasses of MERCs. RiboMERCs represent a subclass of morphologically distinct MERCs that provide useful parameters/group labels to define a novel MERC detection approach. Further definition of the other MCS-DETECT-detected MERCs may require additional information aside from the relatively limited size and shape features used to characterize riboMERCs (Cardoen et al., 2024). This would extend to defining MERCs based on their molecular composition or to live cell analyses to detect dynamics or functional outputs of individual MERCs using fluorescent reporters and high-speed SRM (Nixon-Abell et al., 2016). Deep learning could be applied to detect differences in MERC expression patterns due to the loss of specific tethers. A critical obstacle, however, is the extreme imbalance in whole-cell volumetric data, where larger 3D objects that are discriminative are by definition far less frequent, even though they are visually striking to a human expert. Whether current 3D spatial and temporal resolution limits for voxel-based imaging are sufficient to detect the more subtle distinctions amongst smaller MERCs remains to be determined, and one can only look forward to future hardware and software developments that improve our ability to see more clearly within the cell.

Weakly supervised approaches can exploit combinatorial information, which is often "nested" or "hierarchical." For example, a single cell can have multiple labels: cell line, expression, treatment, and so on. Furthermore, biological structures are often modular, with each modular part having a label that is hierarchically more refined with respect to its top label. The modular construction of caveolae is a great example of this. The use of weakly supervised learning paradigms to leverage such information, beyond the dual paradigms (i.e., PC3 versus PC3+cavin-1; HT-1080 versus COS-7) exploited here (Fig. 4), will necessarily lead to a more refined description and improved biological understanding of the studied structures.

Explainable AI in SRM
If our goal is to determine the inner structure or molecular architecture of biological structures (Fig. 3), then it is critical to understand the basis by which AI is making decisions. Raw image data (i.e., pixels, point clouds) are very large; training the model requires preparing the data in a format through which the machine learning algorithm can best identify biologically relevant differences. In the CAV1 point cloud data, 28 features were extracted from the segmented clusters and reported on the structural aspects of the clusters such as hollowness, topology, network interactions, size, and shape (Khater et al., 2018). This predefined feature approach reports on clearly understood structural aspects of the CAV1 clusters, showing that caveolae are large hollow blobs. Feature analysis of MERCs was limited to three features that could be validated by correspondence to EM (Cardoen et al., 2024).

Machine learning via handcrafted features is sometimes referred to as shallow learning since extraction of these features amounts to the execution of a particular known formula or short (shallow) recipe. This contrasts with deep learning, where the exact formula or recipe to extract the features is unknown beforehand; only its general form is known and takes the shape of a long (deep) sequence of operations whose exact equations are optimized to attain a certain objective, the accurate classification of images. Traditionally, deep models are based on convolutional neural networks (CNN) that implement feature extraction primarily via the application of many convolution operations in sequence (layers). The fact that the features are constructed from a deep sequence of operations and determined by a large number of parameters makes them hard to interpret and the decision process using these features (i.e., classification) difficult to understand—hence, the "black box" label associated with deep models. Over the last decade, the number of parameters of deep models has increased by several orders of magnitude, from

a few layers with millions of parameters to thousands of layers with trillions of parameters.

Deep learning is achieving state-of-the-art results on a wide array of prediction applications (e.g., classification and segmentation), surpassing their shallow machine learning counterparts often by big margins, and even meeting or surpassing human experts on biomedical image interpretation tasks (Fujisawa et al., 2019; Rajpurkar et al., 2017, Preprint; Zhang et al., 2019). While deep learning has found rapid adoption in SRM acquisition and image generation–related tasks (Hyun and Kim, 2023), discovery-oriented SRM data analysis is still limited (Khater et al., 2019a; Zehtabian et al., 2022). Explainable AI (XAI) is a fast-growing field aimed at improving our understanding of deep features and explaining deep model decision processes. XAI is moving toward standardizing the characteristics that an explainable or interpretable model should satisfy (Jin et al., 2023). Beyond natural image analysis, XAI has found rapid adoption in medical image analysis (Chaddad et al., 2023), a field that shares defining characteristics of SRM-based analysis: scarcity of ground truth, targeted at discovery, and high societal impact of findings. Extending deep learning models of SRM with XAI can enable the identification of novel subcellular structures and processes (Long et al., 2020; Nagao et al., 2020).

Region-based explanations, for example, gradient-weighted class activation mapping (Grad-CAM) and its many variants, were some of the earliest XAI approaches finding widespread adoption, as indicated by the citations of the introducing paper (Selvaraju et al., 2020). In contrast, Shapley-based approaches (Shapley, 1953) focus on establishing which features are strongly associated with the outcome of a model (van Zyl et al., 2024). Region- and feature-based explanations still require expert interpretation and may not always be robust. XAI methods can be difficult to use as evidence or validation for new discoveries in a context where validation is scarce. To ease this burden of uncertainty on expert users, recent advances focus on taxonomies and metrics that differentiate types of XAI (Ibrahim and Shafiq, 2023). More recently, XAI approaches are exploring large language models (LLM) based higher level explanations, such as providing language-based "concepts" (Maddigan et al., 2024, Preprint; Mavrepis et al., 2024, Preprint) alongside model performance. Here, the model would explain, using plain English, what concepts or domain-specific keywords are associated with the model's decision. It should be noted that using XAI is not necessarily guaranteed to augment AI-guided expert decisions (Jin et al., 2024). Finally, in discovery-oriented research, it is important to select methods that are causative and not just correlative (Vasudevan et al., 2021, Preprint), aligning with the current experiment design, where controlled experimental factors such as knockout can ensure the elimination of correlative effects in favor of the true causal information.

Beyond weak supervision

Beyond weakly supervised methods, self-supervision has emerged as a powerful paradigm to learn rich semantic features that can be specialized for a chosen end-task with a minimum of ground truth labels and have been successfully adapted to microscopy (Kobayashi et al., 2022; Wu et al., 2022). Unlike weak supervision, self-supervision does not require a group (e.g., cell) level label to learn an informative representation. However, it does not provide an end-task capable solution, as the final stage would be either unsupervised, supervised, or weakly supervised. The advantages, however, include requiring far fewer labels, higher performance, and the ability to reuse the learned encoding for multiple tasks.

Counterfactual learning simulates from existing data "what if" scenarios, causal relations between objects or conditions without the specific experiment taking place (Pearl, 2010). These then extend to generative models that learn to synthesize data from images, features, or even based on descriptions (language). While generative models do not specifically enforce a causal relation between their input and what they generate, the potential of such models in subcellular biology is largely untapped, but highly promising based on their interpretability and rapid adoption in medical imaging (Yi et al., 2019). These represent powerful paradigms for both medical and biology experiments, potentially enabling experiments otherwise not feasible due to ethical or resource constraints. Scalability is an issue, with SRM data often being orders of magnitude larger in dimension compared to datasets on which most deep learning models are being developed, whether it is for point cloud- or voxel-based data.

A recent perspective (Volpe et al., 2023, Preprint) argues that stability (i.e., robustness to confounding factors as well as reliability and performance on unseen datasets) of AI models is critical for the adoption of unsupervised learning for the reconstruction of SRM data. Stability of AI models is critical and can be addressed by the emerging work in continual (Parisi et al., 2019) and out-of-distribution learning, as well as by resolving "short-cuts" that degrade performance on new datasets (Robinson et al., 2021). We would argue that ongoing improvements in both SRM and AI offer tremendous potential not only for SRM data interpretation but also for semantic discovery leading to the accelerated exploration of nanoscale and mesoscale biology.

The ability of foundation and large language models (FLLM) offers great potential as these models are built on enormous datasets and then tuned to domain-specific tasks, as recently demonstrated in single-cell genomics (Luecken et al., 2022; Myers et al., 2024). However, the challenge in validation, interpretation, and reproduction of such models remains. The problem here is the rapid increase in the computing power and the associated carbon footprint in using, let alone training, such models (Bouza et al., 2023; Shah and Bhavsar, 2022). Validation and development of any novel technique are still constrained by access to representative, diverse open data that the community can review and challenge. In addition, translating results obtained on benchmark data to newly acquired data is sensitive to acquisition-specific signatures, such as PSF configuration, localization algorithm, and specimen-specific factors. For example, given that most SRM acquisitions are developed by closed-source vendors, it is often challenging to translate results across vendors. Recent efforts to standardize file formats and experimental design have been driven by the scientific community

rather than commercial needs and are a critical first step in resolving these challenges (Cardoen et al., 2023; Sarkans et al., 2021; Schmied et al., 2024). Community-driven efforts are underway to provide a single standardized format for critical metadata (Moore et al., 2023; Swedlow et al., 2021). Ignoring such confounding factors will only delay advances in AI-driven discovery and limit reproducibility. Open datasets, acquired and curated at great cost in expense, training, and expertise by publicly funded scientific centers, should be standardized and shared with a clear license to favor open science.

Be it determined by human endeavor or AI, newly discovered scientific knowledge needs to be understandable, reproducible, and trustworthy to be adopted by experts in the field. These challenges have been encountered by scientists introducing technological developments, going back to Golgi. The exponential changes that computer science and AI have brought to science, not to mention our everyday lives, means that validation will need to keep pace. Waiting 50 years for validation of scientific discovery, as for the Golgi apparatus, is not an option.

Acknowledgments

References

Anastasia, I., N. Ilacqua, A. Raimondi, P. Lemieux, R. Ghandehari-Alavijeh, G. Faure, S.L. Mekhedov, K.J. Williams, F. Caicci, G. Valle, et al. 2021. Mitochondria-rough-ER contacts in the liver regulate systemic lipid homeostasis. Cell Rep. 34:108873. https://doi.org/10.1016/j.celrep.2021.108873

Anderson, R.G. 1998. The caveolae membrane system. Annu. Rev. Biochem. 67: 199–225. https://doi.org/10.1146/annurev.biochem.67.1.199

Balzarotti, F., Y. Eilers, K.C. Gwosch, A.H. Gynnå, V. Westphal, F.D. Stefani, J. Elf, and S.W. Hell. 2017. Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. Science. 355: 606–612. https://doi.org/10.1126/science.aak9913

Bender, S.W.B., M.W. Dreisler, M. Zhang, J. Kæstel-Hansen, and N.S. Hatzakis. 2024. SEMORE: SEgmentation and MORphological fingErprinting by machine learning automates super-resolution data analysis. Nat. Commun. 15:1763. https://doi.org/10.1038/s41467-024-46106-0

Bentivoglio, M., and P. Mazzarello. 1998. One hundred years of the Golgi apparatus: History of a disputed cell organelle. Ital. J. Neurol. Sci. 19: 241–247. https://doi.org/10.1007/BF02427612

Betzig, E., G.H. Patterson, R. Sougrat, O.W. Lindwasser, S. Olenych, J.S. Bonifacino, M.W. Davidson, J. Lippincott-Schwartz, and H.F. Hess. 2006. Imaging intracellular fluorescent proteins at nanometer resolution. Science. 313:1642–1645. https://doi.org/10.1126/science.1127344

Bouza, L., A. Bugeau, and L. Lannelongue. 2023. How to estimate carbon footprint when training deep learning models? A guide and review. Environ. Res. Commun. 5:115014. https://doi.org/10.1088/2515-7620/acf81b

Cardoen, B., H. Ben Yedder, S. Lee, I.R. Nabi, and G. Hamarneh. 2023. DataCurator.jl: Efficient, portable and reproducible validation, curation and transformation of large heterogeneous datasets, using human-readable recipes compiled into machine-verifiable templates. Bioinform. Adv. 3:vbad068. https://doi.org/10.1093/bioadv/vbad068

Cardoen, B., K.R. Vandevoorde, G. Gao, P. Ortiz-Silva, P. Alan, W. Liu, E. Tiliakou, A.W. Vogl, G. Hamarneh, and I.R. Nabi. 2024. Membrane contact site detection (MCS-DETECT) reveals dual control of rough mitochondria-ER contacts. J. Cell Biol. 223:e202206109. https://doi.org/10.1083/jcb.202206109

Cardoen, B., T. Wong, P. Alan, S. Lee, J.A. Matsubara, I.R. Nabi, and G. Hamarneh. 2022. SPECHT: Self-tuning Plausibility based object detection Enables quantification of Conflict in Heterogeneous multi-scale microscopy. PLoS One. 17:e0276726. https://doi.org/10.1371/journal.pone.0276726

Cardoen, B., H.B. Yedder, A. Sharma, K.C. Chou, I.R. Nabi, and G. Hamarneh. 2020. ERGO: Efficient recurrent graph optimized emitter density estimation in single molecule localization microscopy. IEEE Trans. Med. Imaging. 39:1942–1956. https://doi.org/10.1109/TMI.2019.2962361

Chaddad, A., J. Peng, J. Xu, and A. Bouridane. 2023. Survey of explainable AI techniques in healthcare. Sensors. 23:634. https://doi.org/10.3390/s23020634

D'Alonzo, M., A. Bozkurt, C. Alessi-Fox, M. Gill, D.H. Brooks, M. Rajadhyaksha, K. Kose, and J.G. Dy. 2021. Semantic segmentation of reflectance confocal microscopy mosaics of pigmented lesions using weak labels. Sci. Rep. 11:3679. https://doi.org/10.1038/s41598-021-82969-9

Deguchi, T., M.K. Iwanski, E.M. Schentarra, C. Heidebrecht, L. Schmidt, J. Heck, T. Weihs, S. Schnorrenberg, P. Hoess, S. Liu, et al. 2023. Direct observation of motor protein stepping in living cells using MINFLUX. Science. 379:1010–1015. https://doi.org/10.1126/science.ade2676

Diekmann, S., and C. Hoischen. 2014. Biomolecular dynamics and binding studies in the living cell. Phys. Life Rev. 11:1–30. https://doi.org/10.1016/j.plrev.2013.11.011

Edrington, T.C., E. Kintz, J.B. Goldberg, and L.K. Tamm. 2011. Structural basis for the interaction of lipopolysaccharide with outer membrane protein H (OprH) from Pseudomonas aeruginosa. J. Biol. Chem. 286:39211–39223. https://doi.org/10.1074/jbc.M111.280933

Fu, S., W. Shi, T. Luo, Y. He, L. Zhou, J. Yang, Z. Yang, J. Liu, X. Liu, Z. Guo, et al. 2023. Field-dependent deep learning enables high-throughput whole-cell 3D super-resolution imaging. Nat. Methods. 20:459–468. https://doi.org/10.1038/s41592-023-01775-5

Fujisawa, Y., Y. Otomo, Y. Ogata, Y. Nakamura, R. Fujita, Y. Ishitsuka, R. Watanabe, N. Okiyama, K. Ohara, and M. Fujimoto. 2019. Deep-learning-based, computer-aided classifier developed with a small dataset of clinical images surpasses board-certified dermatologists in skin tumour diagnosis. *Br. J. Dermatol.* 180:373–381. https://doi.org/10.1111/bjd.16924

Gao, G., C. Zhu, E. Liu, and I.R. Nabi. 2019. Reticulon and CLIMP-63 regulate nanodomain organization of peripheral ER tubules. *PLoS Biol.* 17: e3000355. https://doi.org/10.1371/journal.pbio.3000355

Goodsell, D.S., A.J. Olson, and S. Forli. 2020. Art and science of the cellular mesoscale. *Trends Biochem. Sci.* 45:472–483. https://doi.org/10.1016/j.tibs.2020.02.010

Han, B., J.C. Porta, J.L. Hanks, Y. Peskova, E. Binshtein, K. Dryden, D.P. Claxton, H.S. Mchaourab, E. Karakas, M.D. Ohi, and A.K. Kenworthy. 2020. Structure and assembly of CAV1 8S complexes revealed by single particle electron microscopy. *Sci. Adv.* 6:eabc6185. https://doi.org/10.1126/sciadv.abc6185

Hayer, A., M. Stoeber, C. Bissig, and A. Helenius. 2010. Biogenesis of caveolae: Stepwise assembly of large caveolin and cavin complexes. *Traffic.* 11: 361–382. https://doi.org/10.1111/j.1600-0854.2009.01023.x

Helle, S.C., G. Kanfer, K. Kolar, A. Lang, A.H. Michel, and B. Kornmann. 2013. Organization and function of membrane contact sites. *Biochim. Biophys. Acta.* 1833:2526–2541. https://doi.org/10.1016/j.bbamcr.2013.01.028

Herrera-Cruz, M.S., and T. Simmen. 2017. Of yeast, mice and men: MAMs come in two flavors. *Biol. Direct.* 12:3. https://doi.org/10.1186/s13062-017-0174-5

Hill, M.M., M. Bastiani, R. Luetterforst, M. Kirkham, A. Kirkham, S.J. Nixon, P. Walser, D. Abankwa, V.M. Oorschot, S. Martin, et al. 2008. PTRF-Cavin, a conserved cytoplasmic protein required for caveola formation and function. *Cell.* 132:113–124. https://doi.org/10.1016/j.cell.2007.11.042

Hung, V., S.S. Lam, N.D. Udeshi, T. Svinkina, G. Guzman, V.K. Mootha, S.A. Carr, and A.Y. Ting. 2017. Proteomic mapping of cytosol-facing outer mitochondrial and ER membranes in living human cells by proximity biotinylation. *Elife.* 6:e24463. https://doi.org/10.7554/eLife.24463

Hyun, Y., and D. Kim. 2023. Recent development of computational cluster analysis methods for single-molecule localization microscopy images. *Comput. Struct. Biotechnol. J.* 21:879–888. https://doi.org/10.1016/j.csbj.2023.01.006

Ibrahim, R., and M.O. Shafiq. 2023. Explainable convolutional neural networks: A taxonomy, review, and future directions. *ACM Comput. Surv.* 55:206. https://doi.org/10.1145/3563691

Jin, W., M. Fatehi, R. Guo, and G. Hamarneh. 2024. Evaluating the clinical utility of artificial intelligence assistance and its explanation on the glioma grading task. *Artif. Intell. Med.* 148:102751. https://doi.org/10.1016/j.artmed.2023.102751

Jin, W., X. Li, M. Fatehi, and G. Hamarneh. 2023. Guidelines and evaluation of clinical explainable AI in medical image analysis. *Med. Image Anal.* 84: 102684. https://doi.org/10.1016/j.media.2022.102684

Jumper, J., R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature.* 596: 583–589. https://doi.org/10.1038/s41586-021-03819-2

Ke, Z., J. Oton, K. Qu, M. Cortese, V. Zila, L. McKeane, T. Nakane, J. Zivanov, C.J. Neufeldt, B. Cerikan, et al. 2020. Structures and distributions of SARS-CoV-2 spike proteins on intact virions. *Nature.* 588:498–502. https://doi.org/10.1038/s41586-020-2665-2

Khater, I.M., S.T. Aroca-Ouellette, F. Meng, I.R. Nabi, and G. Hamarneh. 2019a. Caveolae and scaffold detection from single molecule localization microscopy data using deep learning. *PLoS One.* 14:e0211659. https://doi.org/10.1371/journal.pone.0211659

Khater, I.M., Q. Liu, K.C. Chou, G. Hamarneh, and I.R. Nabi. 2019b. Super-resolution modularity analysis shows polyhedral caveolin-1 oligomers combine to form scaffolds and caveolae. *Sci. Rep.* 9:9888. https://doi.org/10.1038/s41598-019-46174-z

Khater, I.M., F. Meng, T.H. Wong, I.R. Nabi, and G. Hamarneh. 2018. Super resolution network analysis defines the molecular architecture of caveolae and caveolin-1 scaffolds. *Sci. Rep.* 8:9009. https://doi.org/10.1038/s41598-018-27216-4

Kobayashi, H., K.C. Cheveralls, M.D. Leonetti, and L.A. Royer. 2022. Self-supervised deep learning encodes high-resolution features of protein subcellular localization. *Nat. Methods.* 19:995–1003. https://doi.org/10.1038/s41592-022-01541-z

Kudyshev, Z.A., D. Sychev, Z. Martin, O. Yesilyurt, S.I. Bogdanov, X. Xu, P.G. Chen, A.V. Kildishev, A. Boltasseva, and V.M. Shalaev. 2023. Machine learning assisted quantum super-resolution microscopy. *Nat. Commun.* 14:4828. https://doi.org/10.1038/s41467-023-40506-4

Laine, R.F., G. Goodfellow, L.J. Young, J. Travers, D. Carroll, O. Dibben, H. Bright, and C.F. Kaminski. 2018. Structured illumination microscopy combined with machine learning enables the high throughput analysis and classification of virus structure. *Elife.* 7:e40183. https://doi.org/10.7554/eLife.40183

Lajoie, P., J.G. Goetz, J.W. Dennis, and I.R. Nabi. 2009. Lattices, rafts, and scaffolds: Domain regulation of receptor signaling at the plasma membrane. *J. Cell Biol.* 185:381–385. https://doi.org/10.1083/jcb.200811059

Li, H., X. Sun, W. Cui, M. Xu, J. Dong, B.E. Ekundayo, D. Ni, Z. Rao, L. Guo, H. Stahlberg, et al. 2024a. Computational drug development for membrane protein targets. *Nat. Biotechnol.* 42:229–242. https://doi.org/10.1038/s41587-023-01987-2

Li, Y.L., I.M. Khater, C. Hallgrimson, B. Cardoen, T.H. Wong, G. Hamarneh, and I.R. Nabi. 2024b. SuperResNET single molecule localization microscopy model-free network analysis achieves molecular resolution of Nup96 in preparation. *bioRxiv.* https://doi.org/10.1101/2024.03.12.584716 (Preprint posted March 13, 2024).

Lippincott-Schwartz, J. 2011. Emerging in vivo analyses of cell function using fluorescence imaging (*). *Annu. Rev. Biochem.* 80:327–332. https://doi.org/10.1146/annurev-biochem-121010-125553

Liu, M.Z., C. Swintelski, S. Sun, M. Siddique, E. Desperito, S. Jambawalikar, and R. Ha. 2022a. Weakly supervised deep learning approach to breast MRI assessment. *Acad. Radiol.* 29:S166–S172. https://doi.org/10.1016/j.acra.2021.03.032

Liu, S., P. Hoess, and J. Ries. 2022b. Super-resolution microscopy for structural cell biology. *Annu. Rev. Biophys.* 51:301–326. https://doi.org/10.1146/annurev-biophys-102521-112912

Liu, Z., L. Jin, J. Chen, Q. Fang, S. Ablameyko, Z. Yin, and Y. Xu. 2021. A survey on applications of deep learning in microscopy image analysis. *Comput. Biol. Med.* 134:104523. https://doi.org/10.1016/j.compbiomed.2021.104523

Lolo, F.N., N. Walani, E. Seemann, D. Zalvidea, D.M. Pavón, G. Cojoc, M. Zamai, C. Viaris de Lesegno, F. Martínez de Benito, M. Sánchez-Álvarez, et al. 2023. Caveolin-1 dolines form a distinct and rapid caveolae-independent mechanoadaptation system. *Nat. Cell Biol.* 25:120–133. https://doi.org/10.1038/s41556-022-01034-3

Long, R.K.M., K.P. Moriarty, B. Cardoen, G. Gao, A.W. Vogl, F. Jean, G. Hamarneh, and I.R. Nabi. 2020. Super resolution microscopy and deep learning identify Zika virus reorganization of the endoplasmic reticulum. *Sci. Rep.* 10:20937. https://doi.org/10.1038/s41598-020-77170-3

Luecken, M.D., M. Büttner, K. Chaichoompu, A. Danese, M. Interlandi, M.F. Mueller, D.C. Strobl, L. Zappia, M. Dugas, M. Colomé-Tatché, and F.J. Theis. 2022. Benchmarking atlas-level data integration in single-cell genomics. *Nat. Methods.* 19:41–50. https://doi.org/10.1038/s41592-021-01336-8

Maddigan, P., A. Lensen, and B. Xue. 2024. Explaining genetic programming trees using large language models. *arXiv.* https://doi.org/10.48550/arXiv.2403.03397 (Preprint posted March 6, 2024).

Mavrepis, P., G. Makridis, G. Fatouros, V. Koukos, M.M. Separdani, and D. Kyriazis. 2024. XAI for all: Can large language models simplify explainable AI? *arXiv.* https://doi.org/10.48550/arXiv.2401.13110 (Preprint posted January 23, 2024).

Merriam-Webster.com. 2024. Artificial intelligence. https://www.merriam-webster.com/dictionary/artificial%20intelligence (accessed April 1, 2024).

Moore, J., D. Basurto-Lozada, S. Besson, J. Bogovic, J. Bragantini, E.M. Brown, J.-M. Burel, X. Casas Moreno, G. de Medeiros, E.E. Diel, et al. 2023. OME-Zarr: A cloud-optimized bioimaging file format with international community support. *Histochem. Cell Biol.* 160:223–251. https://doi.org/10.1007/s00418-023-02209-1

Morgado, L., E. Gómez-de-Mariscal, H.S. Heil, and R. Henriques. 2024. The rise of data-driven microscopy powered by machine learning. *J. Microsc.* https://doi.org/10.1111/jmi.13282

Mund, M., A. Tschanz, Y.L. Wu, F. Frey, J.L. Mehl, M. Kaksonen, O. Avinoam, U.S. Schwarz, and J. Ries. 2023. Clathrin coats partially preassemble and subsequently bend during endocytosis. *J. Cell Biol.* 222:e202206038. https://doi.org/10.1083/jcb.202206038

Myers, D., R. Mohawesh, V.I. Chellaboina, A.L. Sathvik, P. Venkatesh, Y.-H. Ho, H. Henshaw, M. Alhawawreh, D. Berdik, and Y. Jararweh. 2024. Foundation and large language models: Fundamentals, challenges, opportunities, and social impacts. *Cluster Comput.* 27:1–26. https://doi.org/10.1007/s10586-023-04203-7

Nagao, Y., M. Sakamoto, T. Chinen, Y. Okada, and D. Takao. 2020. Robust classification of cell cycle phase and biological feature extraction by

image-based deep learning. *Mol. Biol. Cell.* 31:1346–1354. https://doi.org/10.1091/mbc.E20-03-0187

Nehme, E., L.E. Weiss, T. Michaeli, and Y. Shechtman. 2018. Deep-STORM: Super-resolution single-molecule microscopy by deep learning. *Optica.* 5:458–464. https://doi.org/10.1364/OPTICA.5.000458

Nixon-Abell, J., C.J. Obara, A.V. Weigel, D. Li, W.R. Legant, C.S. Xu, H.A. Pasolli, K. Harvey, H.F. Hess, E. Betzig, et al. 2016. Increased spatio-temporal resolution reveals highly dynamic dense tubular matrices in the peripheral ER. *Science.* 354:3928. https://doi.org/10.1126/science.aaf3928

Ortiz-Perez, A., M. Zhang, L.W. Fitzpatrick, C. Izquierdo-Lozano, and L. Albertazzi. 2024. Advanced optical imaging for the rational design of nanomedicines. *Adv. Drug Deliv. Rev.* 204:115138. https://doi.org/10.1016/j.addr.2023.115138

Parisi, G.I., R. Kemker, J.L. Part, C. Kanan, and S. Wermter. 2019. Continual lifelong learning with neural networks: A review. *Neural Netw.* 113:54–71. https://doi.org/10.1016/j.neunet.2019.01.012

Park, H.H., B. Wang, S. Moon, T. Jepson, and K. Xu. 2023. Machine-learning-powered extraction of molecular diffusivity from single-molecule images for super-resolution mapping. *Commun. Biol.* 6:336. https://doi.org/10.1038/s42003-023-04729-x

Pearl, J. 2010. An introduction to causal inference. *Int. J. Biostat.* 6:7. https://doi.org/10.2202/1557-4679.1203

Petkidis, A., V. Andriasyan, and U.F. Greber. 2023. Machine learning for cross-scale microscopy of viruses. *Cell Rep. Methods.* 3:100557. https://doi.org/10.1016/j.crmeth.2023.100557

Priessner, M., D.C.A. Gaboriau, A. Sheridan, T. Lenn, C. Garzon-Coral, A.R. Dunn, J.R. Chubb, A.M. Tousley, R.G. Majzner, U. Manor, et al. 2024. Content-aware frame interpolation (CAFI): Deep learning-based temporal super-resolution for fast bioimaging. *Nat. Methods.* 21:322–330. https://doi.org/10.1038/s41592-023-02138-w

Pylvänäinen, J.W., E. Gómez-de-Mariscal, R. Henriques, and G. Jacquemet. 2023. Live-cell imaging in the deep learning era. *Curr. Opin. Cell Biol.* 85:102271. https://doi.org/10.1016/j.ceb.2023.102271

Qiao, C., D. Li, Y. Liu, S. Zhang, K. Liu, C. Liu, Y. Guo, T. Jiang, C. Fang, N. Li, et al. 2023. Rationalized deep learning super-resolution microscopy for sustained live imaging of rapid subcellular processes. *Nat. Biotechnol.* 41:367–377. https://doi.org/10.1038/s41587-022-01471-3

Rajpurkar, A.R., L.J. Mateo, S.E. Murphy, and A.N. Boettiger. 2021. Deep learning connects DNA traces to transcription to reveal predictive features beyond enhancer-promoter contact. *Nat. Commun.* 12:3423. https://doi.org/10.1038/s41467-021-23831-4

Rajpurkar, P., J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, and K. Shpanskaya. 2017. Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv.* https://doi.org/10.48550/arXiv.1711.05225 (Preprint posted November 14, 2017).

Robinson, J., L. Sun, K. Yu, K. Batmanghelich, S. Jegelka, and S. Sra. 2021. Can contrastive learning avoid shortcut solutions? *Adv. Neural Inf. Process. Syst.* 34:4974–4986.

Rowland, A.A., and G.K. Voeltz. 2012. Endoplasmic reticulum-mitochondria contacts: Function of the junction. *Nat. Rev. Mol. Cell Biol.* 13:607–625. https://doi.org/10.1038/nrm3440

Rozov, A., I. Khusainov, K. El Omari, R. Duman, V. Mykhaylyk, M. Yusupov, E. Westhof, A. Wagner, and G. Yusupova. 2019. Importance of potassium ions for ribosome structure and function revealed by long-wavelength X-ray diffraction. *Nat. Commun.* 10:2519. https://doi.org/10.1038/s41467-019-10409-4

Ruszczycki, B., and T. Bernas. 2018. Quality of biological images, reconstructed using localization microscopy data. *Bioinformatics.* 34:845–852. https://doi.org/10.1093/bioinformatics/btx597

Saberian, M.S., K.P. Moriarty, A.D. Olmstead, C. Hallgrimson, F. Jean, I.R. Nabi, M.W. Libbrecht, and G. Hamarneh. 2022. DEEMD: Drug efficacy estimation against SARS-CoV-2 based on cell morphology with deep multiple instance learning. *IEEE Trans. Med. Imaging.* 41:3128–3145. https://doi.org/10.1109/TMI.2022.3178523

Saguy, A., O. Alalouf, N. Opatovski, S. Jang, M. Heilemann, and Y. Shechtman. 2023. DBlink: Dynamic localization microscopy in super spatiotemporal resolution via deep learning. *Nat. Methods.* 20:1939–1948. https://doi.org/10.1038/s41592-023-01966-0

Sahl, S.J., S.W. Hell, and S. Jakobs. 2017. Fluorescence nanoscopy in cell biology. *Nat. Rev. Mol. Cell Biol.* 18:685–701. https://doi.org/10.1038/nrm.2017.71

Sarkans, U., W. Chiu, L. Collinson, M.C. Darrow, J. Ellenberg, D. Grunwald, J.-K. Hériché, A. Iudin, G.G. Martins, T. Meehan, et al. 2021. REMBI: Recommended Metadata for Biological Images-enabling reuse of microscopy data in biology. *Nat. Methods.* 18:1418–1422. https://doi.org/10.1038/s41592-021-01166-8

Schmidt, R., T. Weihs, C.A. Wurm, I. Jansen, J. Rehman, S.J. Sahl, and S.W. Hell. 2021. MINFLUX nanometer-scale 3D imaging and microsecond-range tracking on a common fluorescence microscope. *Nat. Commun.* 12:1478. https://doi.org/10.1038/s41467-021-21652-z

Schmied, C., M.S. Nelson, S. Avilov, G.-J. Bakker, C. Bertocchi, J. Bischof, U. Boehm, J. Brocher, M.T. Carvalho, C. Chiritescu, et al. 2024. Community-developed checklists for publishing images and image analyses. *Nat. Methods.* 21:170–181. https://doi.org/10.1038/s41592-023-01987-9

Scorrano, L., M.A. De Matteis, S. Emr, F. Giordano, G. Hajnóczky, B. Kornmann, L.L. Lackner, T.P. Levine, L. Pellegrini, K. Reinisch, et al. 2019. Coming together to define membrane contact sites. *Nat. Commun.* 10:1287. https://doi.org/10.1038/s41467-019-09253-3

Selvaraju, R.R., M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra. 2020. Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Int. J. Comput. Vis.* 128:336–359. https://doi.org/10.1007/s11263-019-01228-7

Shah, B., and H. Bhavsar. 2022. Time complexity in deep learning models. *Procedia Comput. Sci.* 215:202–210. https://doi.org/10.1016/j.procs.2022.12.023

Shapley, L.S. 1953. A Value for n-Person Games. In Contributions to the Theory of Games (AM-28). Vol. 2. K. Harold William and T. Albert William, editors. Princeton University Press, Princeton, NJ. 307–318.

Sieben, C., N. Banterle, K.M. Douglass, P. Gönczy, and S. Manley. 2018. Multicolor single-particle reconstruction of protein complexes. *Nat. Methods.* 15:777–780. https://doi.org/10.1038/s41592-018-0140-x

Stoeber, M., P. Schellenberger, C.A. Siebert, C. Leyrat, A. Helenius, and K. Grünewald. 2016. Model for the architecture of caveolae based on a flexible, net-like assembly of Cavin1 and Caveolin discs. *Proc. Natl. Acad. Sci. USA.* 113:E8069–E8078. https://doi.org/10.1073/pnas.1616838113

Swedlow, J.R., P. Kankaanpää, U. Sarkans, W. Goscinski, G. Galloway, L. Malacrida, R.P. Sullivan, S. Härtel, C.M. Brown, C. Wood, et al. 2021. A global view of standards for open image data formats and repositories. *Nat. Methods.* 18:1440–1446. https://doi.org/10.1038/s41592-021-01113-7

Sydor, A.M., K.J. Czymmek, E.M. Puchner, and V. Mennella. 2015. Super-resolution microscopy: From single molecules to supramolecular assemblies. *Trends Cell Biol.* 25:730–748. https://doi.org/10.1016/j.tcb.2015.10.004

van Zyl, C., X. Ye, and R. Naidoo. 2024. Harnessing eXplainable artificial intelligence in time series energy forecasting: A comparative analysis of grad-CAM and shap. *Appl. Energy.* 353:122079. https://doi.org/10.1016/j.apenergy.2023.122079

Vasudevan, R.K., M. Ziatdinov, L. Vlcek, and S.V. Kalinin. 2021. Off-the-shelf deep learning is not enough, and requires parsimony, Bayesianity, and causality. *arXiv.* https://doi.org/10.48550/arXiv.2005.01557 (Preprint posted May 4, 2020).

Volpe, G., C. Wahlby, L. Tian, M. Hecht, A. Yakimovich, K. Monakhova, L. Waller, I.F. Sbalzarini, C.A. Metzler, M. Xie, et al. 2023. Roadmap on Deep Learning for Microscopy. *arXiv.* https://doi.org/10.48550/arXiv.2303.03793 (Preprint posted March 7, 2023).

Wang, P.T., P.O. Garcin, M. Fu, M. Masoudi, P. St-Pierre, N. Panté, and I.R. Nabi. 2015. Distinct mechanisms controlling rough and smooth endoplasmic reticulum contacts with mitochondria. *J. Cell Sci.* 128:2759–2765. https://doi.org/10.1242/jcs.171132

Wikipedia. 2024. Ground truth. Wikipedia, The Free Encyclopedia, https://en.wikipedia.org/w/index.php?title=Ground_truth&oldid=1195021762 (accessed April 2, 2024).

Williamson, D.J., G.L. Burn, S. Simoncelli, J. Griffié, R. Peters, D.M. Davis, and D.M. Owen. 2020. Machine learning for cluster analysis of localization microscopy data. *Nat. Commun.* 11:1493. https://doi.org/10.1038/s41467-020-15293-x

Wong, T.H., I.K. Khater, C. Halgrimsson, Y.L. Li, G. Hamarneh, and I.R. Nabi. 2024. Molecular architecture of clathrin pit formation and inhibition defined by super-resolution network analysis (SuperResNET). *bioRxiv.* https://doi.org/10.1101/2024.03.07.583946v1 (Preprint posted March 8, 2024).

Wu, Y.-L., P. Hoess, A. Tschanz, U. Matti, M. Mund, and J. Ries. 2023. Maximum-likelihood model fitting for quantitative analysis of SMLM data. *Nat. Methods.* 20:139–148. https://doi.org/10.1038/s41592-022-01676-z

Wu, Z., B.B. Chhun, G. Popova, S.-M. Guo, C.N. Kim, L.-H. Yeh, T. Nowakowski, J. Zou, and S.B. Mehta. 2022. DynaMorph: Self-supervised

learning of morphodynamic states of live cells. *Mol. Biol. Cell*. 33:ar59. https://doi.org/10.1091/mbc.E21-11-0561

Xu, Y., J.Y. Zhu, E.I. Chang, M. Lai, and Z. Tu. 2014. Weakly supervised histopathology cancer image segmentation and classification. *Med. Image Anal.* 18:591–604. https://doi.org/10.1016/j.media.2014.01.010

Yang, T., Y. Luo, W. Ji, and G. Yang. 2021. Advancing biological super-resolution microscopy through deep learning: A brief review. *Biophys. Rep.* 7:253–266. https://doi.org/10.52601/bpr.2021.210019

Yi, X., E. Walia, and P. Babyn. 2019. Generative adversarial network in medical imaging: A review. *Med. Image Anal.* 58:101552. https://doi.org/10.1016/j.media.2019.101552

Zehtabian, A., P.M. Müller, M. Goisser, L. Obendorf, L. Jänisch, N. Hümpfer, J. Rentsch, and H. Ewers. 2022. Precise measurement of nanoscopic septin ring structures with deep learning-assisted quantitative super-resolution microscopy. *Mol. Biol. Cell*. 33:ar76. https://doi.org/10.1091/mbc.E22-02-0039

Zhang, X., L. Jin, Q. Fang, W.H. Hui, and Z.H. Zhou. 2010. 3.3 A cryo-EM structure of a nonenveloped virus reveals a priming mechanism for cell entry. *Cell*. 141:472–482. https://doi.org/10.1016/j.cell.2010.03.041

Zhang, Z., P. Chen, M. McGough, F. Xing, C. Wang, M. Bui, Y. Xie, M. Sapkota, L. Cui, J. Dhillon, et al. 2019. Pathologist-level interpretable whole-slide cancer diagnosis with deep learning. *Nat. Mach. Intell.* 1:236–245. https://doi.org/10.1038/s42256-019-0052-1

Zhou, Z.-H. 2017. A brief introduction to weakly supervised learning. *Natl. Sci. Rev.* 5:44–53. https://doi.org/10.1093/nsr/nwx106