nature methods

Perspective

https://doi.org/10.1038/s41592-026-03202-x

From pixels to patterns: the AI revolution in stem cell-derived models

Received: 30 May 2025

Accepted: 21 June 2026

Published online: 14 August 2026

Check for updates

Luca Deininger [1,2,3,5] ✉, Paolo Caldarelli [4,5] ✉, Magdalena Zernicka-Goetz [4,6] ✉ & Ralf Mikut [1,6] ✉

Artificial intelligence (AI) is rapidly transforming stem cell and developmental biology, offering new strategies to analyze, interpret and optimize complex, dynamic systems such as organoids and stem cell-derived embryo models. In this Perspective, we chart the integration of AI into image-based analysis of stem cell systems, highlighting how deep learning, convolutional neural networks and emerging foundation models enable automated classification, segmentation and phenotyping at increasing scale and precision. We showcase applications in phenotyping, drug screening and mechanistic discovery, including real-time fate prediction and the identification of hidden morphological signatures linked to differentiation and disease. Practical challenges, including limited annotated data, model interpretability and live imaging constraints, are examined alongside future opportunities, such as multimodal integration, real-time experimental steering and protocol optimization. Altogether, we argue that AI is not merely an analytical tool, but a discovery engine that enhances reproducibility, accelerates insight and brings us closer to a mechanistic understanding of self-organization in complex stem cell-derived systems.

Stem cell-derived models have advanced both biomedical and fundamental research, offering physiologically relevant systems to study disease mechanisms, support drug discovery, advance regenerative medicine and elucidate fundamental principles of embryonic development. Leveraging the intrinsic self-organization of pluripotent stem cells, these models give rise to complex three-dimensional (3D) structures, including organoids that capture aspects of organ architecture and embryo-like models that recapitulate early developmental stages. Unlike traditional two-dimensional (2D) cell cultures, stem cell-derived 3D systems offer unique opportunities to study cellular behaviors, differentiation processes and therapeutic responses within a context that closely mirrors in vivo conditions. However, their inherent complexity and biological variability pose considerable challenges to standardized analysis and reproducibility.

As these systems increase in complexity and scale, their analysis has become a major bottleneck. Manual inspection and conventional image analysis struggle to capture subtle morphological features, temporal dynamics and rare developmental outcomes, limiting throughput and reproducibility. This gap has driven the adoption of AI-based image-analysis approaches.

AI-based and deep-learning-based image analysis are broadly used across biomedical imaging. In stem cell-derived systems, standard model architectures, combined with domain-specific adaptations, enable segmentation, detection and representation learning for feature extraction and clustering. By replacing manual analysis with standardized computational pipelines, these approaches support robust quality control, phenotyping, trajectory prediction, high-throughput drug screening and disease modeling, while also yielding mechanistic

[1]Group for Automated Image and Data Analysis, Institute for Automation and Applied Informatics, Karlsruhe Institute of Technology, Eggenstein-Leopoldshafen, Germany. [2]Division of Pediatric Neurology and Metabolic Medicine, Department I, Center for Pediatric and Adolescent Medicine, Medical Faculty Heidelberg, Heidelberg University, Heidelberg, Germany. [3]Department of Biosystems Science and Engineering, ETH Zürich, Basel, Switzerland. [4]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA. [5]These authors contributed equally: Luca Deininger, Paolo Caldarelli. [6]These authors jointly supervised this work: Magdalena Zernicka-Goetz, Ralf Mikut. ✉e-mail: ldeininger@ethz.ch; pcaldare@caltech.edu; magdaz@caltech.edu; ralf.mikut@kit.edu

insight. Collectively, they shift analysis from qualitative inspection to quantitative, reproducible measurement at scale.

In this Perspective, we review AI-based image-analysis methods applied to stem cell-derived systems, highlighting common methods and architectures alongside domain-specific adaptations, addressing current challenges and exploring emerging opportunities. We discuss how AI reshapes experimental workflows, improving analytical precision and enhancing reproducibility in fundamental and biomedical stem cell research.

Core concepts in AI-based image analysis

AI has swiftly emerged as a transformative force, reshaping how scientists analyze complex, high-dimensional datasets in biology. Many of these advances are driven by machine learning, a central branch of AI that enables models to classify samples and make predictions without explicit, hand-crafted rules. For a glossary of terms, see Box 1.

Classical machine-learning approaches remain useful in stem cell biology when datasets are modest, labels are scarce or interpretability is required. Support vector machines (SVMs)[1] and logistic regression[2] rely on previously extracted features to enable quality control, stage differentiation and lineage classification. Meanwhile, *k*-means[3] and related clustering can delineate morphological phenotypes of organoids and embryo-like models. Despite their simplicity, these approaches offer robustness, transparency and computational efficiency, which continue to make them effective in diverse biological applications[4,5].

A major shift occurred with the advent of neural networks, particularly convolutional neural networks (CNNs), which replaced manual feature engineering with automated, hierarchical feature learning directly from images. CNNs are built from convolutional layers (Fig. 1a), which apply learnable filters to identify patterns such as edges, textures and shapes. As information propagates through successive layers, these local features are progressively combined into more abstract representations, enabling the capture of complex biological structures.

Following convolution, pooling layers (Fig. 1b) reduce the spatial resolution of feature maps while retaining the most informative signals. By summarizing local neighborhoods through max or average pooling, they improve computational efficiency, increase robustness to small positional shifts, and help reduce overfitting by limiting model complexity.

To stabilize and accelerate training, batch normalization layers (Fig. 1c) are commonly added after convolutions. These normalize intermediate outputs and thereby improve model generalization. Convolutions, normalization and pooling are typically combined into convolutional blocks, the building units of deep CNNs. By stacking these blocks, CNNs can model increasingly complex image features.

At the network output, fully connected layers, like those used in multilayer perceptrons (MLPs), aggregate learned features for high-level interpretation (Fig. 1d), and a softmax layer converts predictions into class probabilities (Fig. 1e).

Beyond image-level classification, CNN-based architectures can be extended to tasks that require the localization of structures within an image. In this context, region proposal mechanisms (Fig. 1f) identify candidate regions that are likely to contain objects of interest. Region-based pooling operations then extract feature representations from these regions. Finally, bounding-box regression layers refine the spatial coordinates of predicted regions, allowing accurate localization of biological structures such as colonies or embryonic compartments (Fig. 1g). These components form the conceptual basis for CNN-based detection models discussed later in the Perspective.

More recently, transformer-based models have been introduced to image analysis by replacing convolutional operations with self-attention mechanisms[6]. This mechanism guides the transformer to prioritize the most relevant parts of the same image. Although conceptually powerful, their limited inductive bias and high data requirements currently constrain their adoption in stem cell imaging. As a result, CNNs remain the most practical and commonly used architectures for image-based analysis in this field.

BOX 1

Glossary

Activation function	Nonlinear mathematical function enabling neural networks to learn complex patterns. A common example is the ReLU.
Batch normalization layer	Layer that standardizes activations within a network, improving training speed and stability.
Convolutional block	Modular CNN unit typically consisting of convolution, batch normalization and pooling layers for hierarchical feature extraction.
Encoder–decoder	Architecture in which the encoder compresses input into a latent representation and the decoder reconstructs or transforms it. Common in segmentation and autoencoders.
Hierarchical feature learning	Extraction of features with increasing abstraction, from simple edges in early layers to complex structures in deeper layers.
MLP	Fully connected neural network with input, hidden and output layers, commonly used for classification.
Neural network	Machine-learning model composed of interconnected layers of artificial neurons that process data through weighted connections and nonlinear activations.
Overfitting	When a model learns training data too closely, including noise, reducing its ability to generalize to new data.
Pooling layer	Downsampling layer that reduces spatial resolution and computation (for example, maximum or average pooling).
Residual connections	Shortcut connections that add layer inputs to outputs, enabling stable training of very deep networks.
RNN	Neural network architecture for sequential data that maintains internal memory to capture temporal patterns.
Self-attention	Mechanism that weights relationships between elements within an input, capturing long-range dependencies. Core component of transformer models.
Softmax layer	Output layer that converts prediction scores into class probabilities.
Supervised learning	Training approach where models learn from labeled data to map inputs to known outputs.
Unsupervised learning	Learning approach that discovers patterns or structure in unlabeled data, such as clustering or dimensionality reduction.
Vanishing gradient problem	Training issue where gradients shrink during backpropagation, preventing effective learning in early network layers.

AI approaches for stem cell image analysis

CNN-based models form the backbone of most image-analysis pipelines in stem cell biology, enabling tasks ranging from image-level classification to segmentation, object detection and feature extraction. This section introduces these core task categories at a conceptual level, followed by domain-specific adaptations, while detailed and up-to-date applications are discussed in the next section.

https://doi.org/10.1038/s41592-026-03202-x

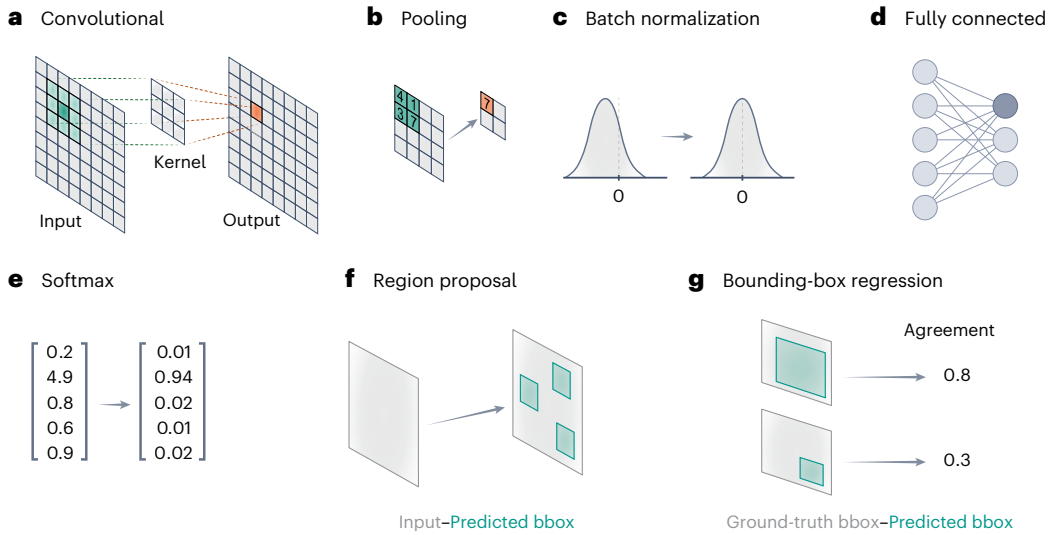

Fig. 1 | Core AI components used in image-based analysis in stem cell biology. Illustration of essential building blocks of deep-learning models, including convolution, pooling, batch normalization and region proposal layers. These components are commonly combined to enable feature extraction, classification and localization of biological structures in complex imaging datasets. **a**, Convolutional layers apply learnable filters to input images to extract spatial features such as edges, textures and patterns. **b**, Pooling layers reduce spatial resolution by summarizing local regions, enhancing computational efficiency and translation invariance. **c**, Batch normalization improves training stability and speed by normalizing activations across mini-batches. **d**, Fully connected layers integrate features extracted in previous layers to make final classification or regression decisions. **e**, Softmax layers convert raw network outputs into probability distributions for classification tasks. **f**, Region proposal mechanisms identify candidate regions within an image likely to contain objects of interest, feeding into subsequent classification and refinement stages. **g**, Bounding-box (bbox) regression quantifies the agreement between predicted and ground-truth bounding boxes in object detection.

Classification

Image classification assigns labels to entire images, enabling discrimination of stem cell types, differentiation states, culture conditions or model systems. A range of standard CNN architectures have been applied depending on dataset size, image complexity and the biological context. Early studies used LeNet[7,8], AlexNet[9,10] and VGG16 (ref. 11) for smaller datasets and simpler tasks, while GoogLeNet[12] and Xception[13,14] were applied to more complex or multi-scale morphologies. ResNet[15] (Fig. 2a) enabled training of deeper models for fine-grained differentiation state classification[16,17], while DenseNet[18] proved effective in data-limited settings due to efficient feature reuse[19,20]. Together, these studies illustrate pragmatic CNN architecture selection driven by experimental constraints and classification goals.

For time-series classification, recurrent neural networks (RNNs) and CNN–RNN hybrids capture temporal dynamics, such as developmental progression in time-lapse microscopy datasets[21]. However, due to the instability and vanishing gradient problems of standard RNNs, more advanced variants like long short-term memory networks or transformer architectures are now used[22,23]. These architectures are better suited for learning long-range dependencies in biological sequences and dynamic imaging data. In addition, a recent study demonstrated that training a CNN on single time-point images and then aggregating predictions across multiple time points outperformed both single-frame CNNs and end-to-end video classification models (Fig. 2b)[5].

Beyond conventional approaches, relative representation learning provides an alternative strategy by training models to assess similarity between samples. Rather than assigning labels based on predefined visual features, the model learns representations that place biologically and morphologically similar stem cell colonies close together, while separating dissimilar ones. Triplet-net approaches, for example, compare anchor, positive and negative samples to improve separation of visually similar classes of stem cell colonies (Fig. 2c)[24], complementing standard CNN-based classification.

Segmentation

To quantify stem cells and stem cell-derived models in microscopy images, researchers often need to determine cell locations and shapes. Image segmentation addresses this problem by assigning a label to every pixel in an image, thereby separating biological structures from the background.

Two main types of segmentation are commonly used in biological imaging: semantic and instance segmentation. Semantic segmentation labels all pixels of the same class together, for example, marking all nuclei in an image without distinguishing individual cells. In contrast, instance segmentation assigns a separate label to each object, enabling the delineation of individual cells or organoids.

Early deep-learning approaches for segmentation were based on fully convolutional networks (FCNs), which preserved spatial information across the image[25,26]. Building on this idea, U-Net[27,28] has become a standard architecture for segmentation in stem cell biology. Its encoder extracts image features while reducing resolution, and its decoder reconstructs a full-resolution segmentation mask (Fig. 2d). Skip connections between encoder and decoder layers help preserve fine spatial details such as thin cell boundaries.

Some biological objects have characteristic shapes that can be exploited explicitly. StarDist[29], for example, represents each object as a star-shaped polygon defined by radial distances from its center[30] (Fig. 2e). This approach is especially effective for round or mildly elongated structures, such as nuclei, single stem cells or spheroids, and often yields more accurate separation of touching objects than standard pixel-based methods.

Segmentation models have also been adapted to handle complex imaging conditions, such as dense organoid cultures or high-resolution screening data. For example, the residual double dynamic convolution attention U-Net (RDAU-Net) extends the standard U-Net architecture by incorporating attention mechanisms and dynamic convolutions[31] (Fig. 2f). This design helps the model focus on biologically relevant regions while suppressing background noise, leading to improved segmentation accuracy in challenging settings such as organoid-based drug screening[31].

https://doi.org/10.1038/s41592-026-03202-x

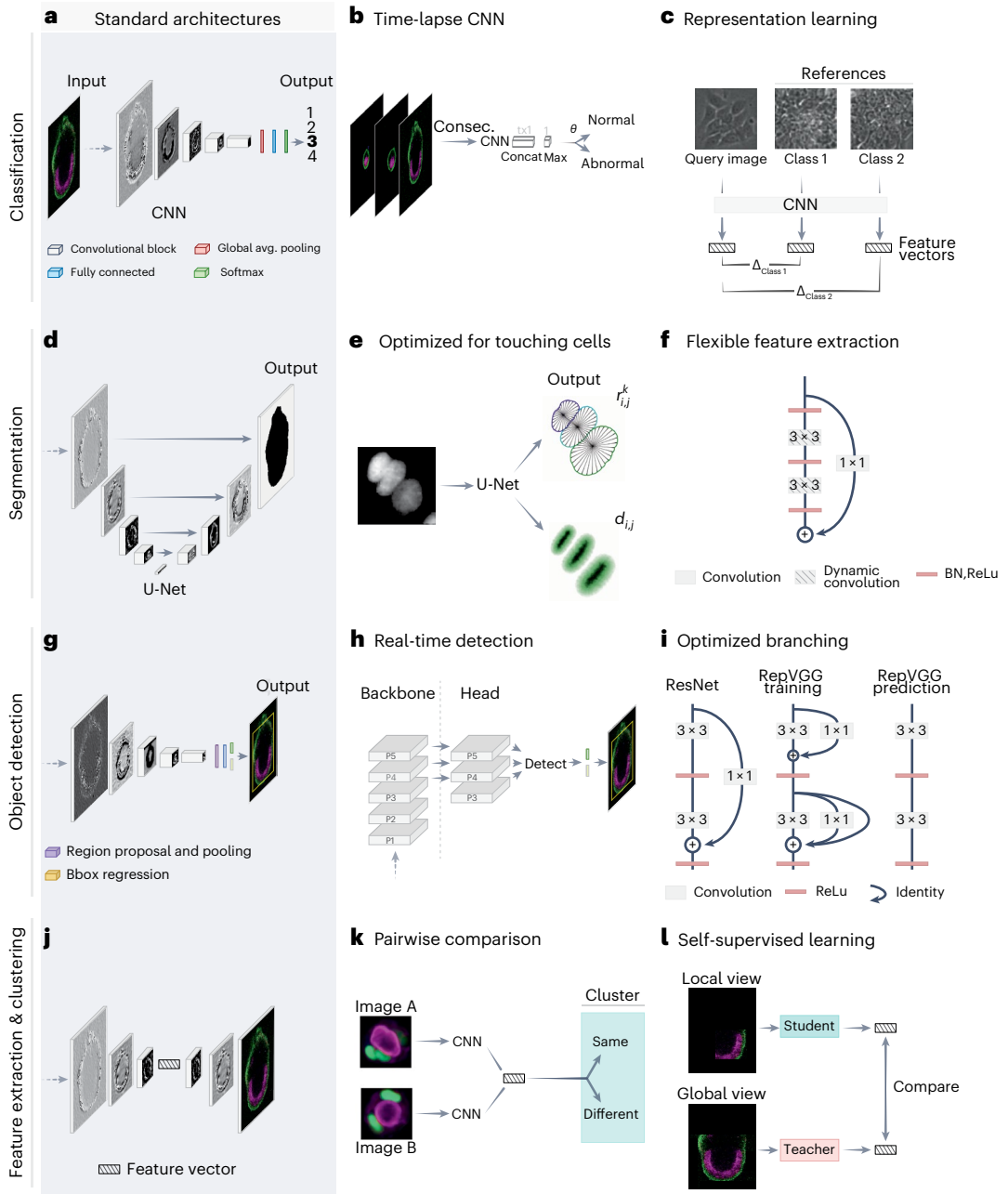

Fig. 2 | AI architectures and modifications for tasks in stem cell biology. **a**–**l**, Methods are grouped by task: classification (**a**–**c**), segmentation (**d**–**f**), object detection (**g**–**i**) and feature extraction with clustering (**j**–**l**). **a**, Standard classification networks are based on CNNs composed of sequential convolutional blocks and global average pooling, followed by fully connected and softmax layers to derive class probabilities. These architectures are commonly adapted or retrained for stem cell applications. **b**, A classification network for time-series data processes consecutive time points with a CNN, and the resulting predictions are combined to yield the final classification. **c**, In some cases, relative representation learning compares query image patches to reference patches[24], assigning the query to the class of the most similar reference based on feature representations. **d**, Segmentation networks, such as U-Net[27], use symmetric encoder–decoder architectures with skip connections and are widely applied in biomedical imaging, including stem cell systems. **e**, Specialized networks such as StarDist[29] are optimized for separating densely packed objects by jointly predicting an object probability (p) and a set of radial distances (r) from each pixel to the object boundary along predefined directions, enabling segmentation of cells inside organoids[69]. **f**, RDAU-Net[31] integrates dynamic convolution to expand the receptive field, which has been shown to be valuable for high-resolution organoid images. **g**, Object detection models, such as Faster R-CNN, combine a convolutional backbone with a region proposal network to predict bounding boxes and their corresponding class labels, valuable for localizing and counting organoids[37]. **h**, More recent architectures, such as YOLOv8 (ref. 42), predict bounding boxes and class labels in a single step, offering faster inference with high accuracy for large-scale organoid detection[72]. **i**, Further optimization includes replacing the YOLO backbone with RepVGG, achieving faster inference and improved performance in organoid detection[43]. RepVGG uses a multi-branch structure during training but uses a single path for inference. **j**, Convolutional autoencoders compress input images into a lower-dimensional latent space (feature vector), and then reconstruct the original image from this representation. This feature vector is used, for example, for unsupervised feature extraction in organoid drug screening[74]. **k**, Siamese networks encode image pairs into a feature vector and map them to similarity scores, assigning positive scores for images from the same cluster and negative scores for images from different clusters[46,50]. **l**, DINO[47] learns representations from differently augmented views of the same image without reconstruction or paired inputs, producing embeddings suitable for clustering, such as grouping stem cell-derived embryos[5]. ReLU, rectified linear unit; BN, batch normalization. Images reproduced with permission from the following: **c**, ref. 24, IEEE/ACM; **e**, ref. 29, Springer; **i**, ref. 43, Elsevier © 2024; **k**, ref. 50, CC BY 4.0.

More recently, transformer-based models have been introduced for image segmentation. These architectures capture long-range relationships across an image, which can be beneficial when structures span large spatial scales. Examples include SegFormer[32,33] and foundation-model approaches such as Segment Anything (SAM)[34,35]. Extensions such as micro-SAM[36] further tailor these models for interactive use in biological imaging, enabling researchers to refine segmentations with minimal manual input.

Object detection

In images involving overlapping stem cells or stem cell-derived models, pixel-level segmentation is infeasible and can also struggle with real-time image processing due to its computational complexity. Object detection offers a more scalable alternative by localizing objects via bounding boxes, an approach well suited for tasks like counting, tracking or density estimation. Bounding-box annotations also simplify the ground-truth generation compared with delineating precise contours for pixel-level segmentation, facilitating the creation of larger training datasets.

Object detection models differ mainly in how candidate object regions are generated and classified. Two-stage detectors, such as faster region-based CNN (Faster R-CNN), first extract shared feature representations using a CNN backbone and then perform region proposal and object classification in separate steps (Fig. 2g). Although this modular design has laid the foundation for object detection in stem cell biology[37], Faster R-CNN is increasingly being replaced by faster alternatives such as Mask R-CNN or YOLO. Mask R-CNN extends Faster R-CNN by adding a branch that predicts a pixel-level segmentation mask for each detected object, which can be useful when distinguishing between overlapping stem cells of the same class[38]. Cascade R-CNN improves detection accuracy through multiple refinement stages that enable more precise recognition of complex structures[39,40].

In contrast, one-stage models like YOLO have become widely popular due to their ability to perform detection and classification in a single step. Recent versions such as YOLOX[41] and YOLOv8 (ref. 42; Fig. 2h) strike a compelling balance between speed and accuracy, and have already been applied for organoid detection. Building on this, another study further optimized the branching of YOLO to improve performance in organoid detection from brightfield images (Fig. 2i)[43]. Although not yet commonly used in stem cell biology, transformer-based detectors are rapidly gaining traction as powerful alternatives for object detection across various domains[44].

Feature extraction and clustering

Beyond localization and classification, deep learning can also learn informative image representations and enable unsupervised comparisons. This capability is particularly valuable in stem cell biology, where labeled data are often scarce and manual feature engineering can be impractical. Models such as convolutional autoencoders[45], Siamese networks[46] and DINO[47–49] exemplify how deep learning can extract abstract features, facilitate anomaly detection and enhance comparative analysis without relying on predefined labels.

Convolutional autoencoders consist of an encoder that compresses input images into a lower-dimensional latent space and a decoder that reconstructs the original image from this representation (Fig. 2j). They are widely applied in anomaly detection, image denoising and feature extraction when labels are limited. Siamese networks, which use two branches with shared weights, learn similarity relationships between image pairs and are useful for tasks like stage estimation and temporal alignment without requiring explicit class labels (Fig. 2k)[46,50]. DINO, by contrast, does not rely on reconstruction or paired inputs, and instead learns from differently augmented views of the same image (Fig. 2l)[47,48] and is well suited to large unannotated datasets for clustering[5]. Crucially, tracking these learned representations over successive time points, whether interpretable morphological descriptors or latent features, can quantify developmental dynamics and trajectories of stem cell-derived models.

Domain-specific AI modifications

Deep-learning models developed for natural image analysis are increasingly applied to stem cell imaging, but differences in image structure, channel composition and semantic meaning often necessitate targeted domain-specific modifications. In addition to these architecture-level adaptations, training strategies such as data augmentation and transfer learning are commonly used to improve robustness in data-limited microscopy settings, as discussed in the next section.

Most pretrained deep-learning architectures are designed for three-channel (RGB) images, reflecting natural image datasets. This creates a mismatch with stem cell imaging, which typically involves grayscale brightfield images or multichannel fluorescence data capturing distinct molecular markers. For grayscale brightfield images, a common strategy is to replicate the single channel across all three input channels, enabling the reuse of pretrained RGB networks without architectural changes. In contrast, a widely used workaround is to assign specific biological channels (for example, nuclear, cytoskeletal and membrane stains) to the red, green and blue dimensions of the model input. While this enables the reuse of existing architectures with minimal modification, it can obscure the semantic independence of each channel.

To more faithfully represent biological data, CNNs can be adapted to accept arbitrary numbers of input channels[5]. This modification allows for direct modeling of high-dimensional fluorescence stacks and other imaging modalities, without the need to collapse biologically distinct signals. These domain-specific adjustments substantially extend the utility of deep-learning models, ensuring compatibility with the complex and varied image formats encountered in stem cell research.

AI applications in stem cell biology

Stem cell research presents a uniquely complex challenge: stem cell-derived systems are dynamic, heterogeneous and difficult to analyze manually, especially when scaled for high-throughput experimentation. Subtle morphological cues, asynchronous differentiation events, and culture-to-culture variability create considerable barriers to reproducibility and strain traditional image-analysis methods. In this context, AI offers scalable, automated and sensitive approaches for monitoring and dissecting stem cell systems in vitro. From quality control and phenotyping to drug screening and mechanistic discovery, AI-driven image analysis is reshaping stem cell research.

Quality control and phenotyping

Stem cells in culture. The promise of pluripotent stem cells lies in their unparalleled ability to self-renew and to differentiate into virtually any cell type. This unique biology has made them foundational tools for developmental biology, regenerative medicine and disease modeling. Most commonly, these cells are harvested from embryonic or adult tissues or are generated via reprogramming of somatic cells into induced pluripotent stem (iPS) cells[51]. Once established, cultures must be maintained under tightly controlled conditions to preserve their pluripotency and expansion potential. Yet even under optimal conditions, stem cell cultures are prone to heterogeneity, senescence and unwanted differentiation, which can jeopardize downstream applications. This creates a need for automated, reliable and noninvasive approaches to monitor stem cell quality, a demand increasingly addressed by AI-powered image analysis.

Recent studies have demonstrated the value of deep learning in the routine monitoring of stem cell cultures based solely on label-free light microscopy. For example, VGG16-based models have accurately distinguished between high-quality and low-quality human embryonic

https://doi.org/10.1038/s41592-026-03202-x

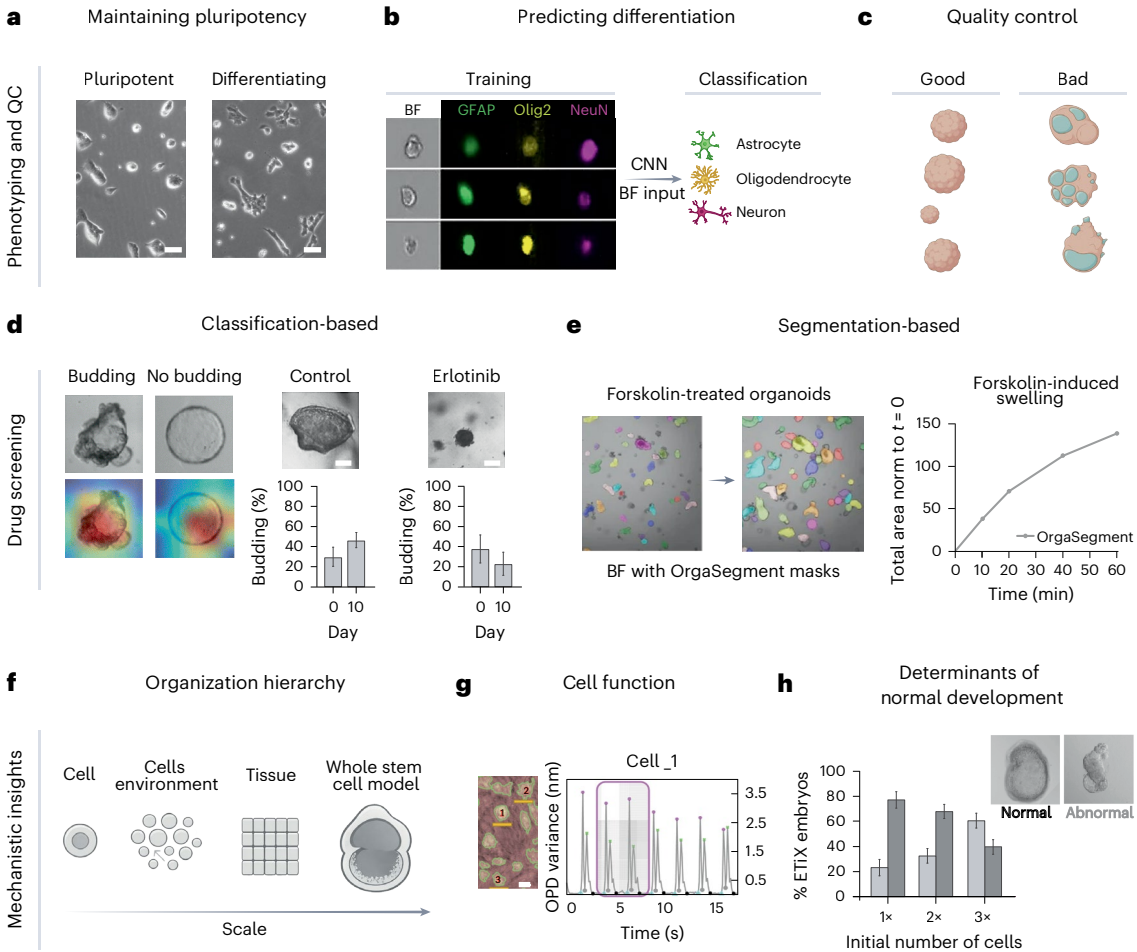

Fig. 3 | Representative AI-enabled image-based analysis in stem cell biology. AI enables scalable, reproducible image-based classification, segmentation and prediction in stem cell systems. **a**–**c**, In the context of phenotyping and quality control (QC), AI distinguishes pluripotent from differentiating stem cells based on subtle morphological cues (**a**)[17], predicts lineage outcomes from brightfield (BF) images combined with marker expression (**b**)[62] and classifies stem cell-derived structures by developmental quality, supporting objective scoring across experiments (**c**)[64]. **d**,**e**, In drug screening and disease modeling, AI classifies phenotypic outcomes such as organoid budding behavior and opacity under compound exposure (**d**)[14], while segmentation-based analysis quantifies treatment responses by tracking morphological changes, including lumen expansion, over time (**e**)[75]. **f**–**h**, For mechanistic insight, AI enables multi-scale feature extraction, from cellular microenvironments to tissue organization (**f**)[28], detects functional behaviors such as cardiomyocyte beating from video data (**g**)[26] and identifies early predictors of successful development, including initial cell number, supporting refinement of differentiation protocols (**h**)[5]. Scale bars, 50 µm for **a**, 100 µm for **d** and **h** and 10 µm for **g**. Images reproduced with permission from the following: **d**, ref. 14, Royal Society of Chemistry; **g**, ref. 26, © Optical Society of America. Other images reproduced under a Creative Commons license CC BY 4.0: **a**, ref. 17; **b**, ref. 62; **c**, ref. 64; **e**, ref. 75; **h**, ref. 5.

stem cell colonies[52], as well as between functional and nonfunctional iPS cell-derived cardiomyocytes[53]. Likewise, DenseNet-based networks have been applied to predict the multipotency of nasal turbinate-derived adult stem cells solely from morphology, eliminating the need for invasive biochemical assays[20]. Cascade R-CNN models have shown to be adept at detecting senescent mesenchymal stem cells (MSCs), identifying hallmarks of cellular aging that compromise regenerative potential[40]. Together, these models enable early exclusion of nonfunctional cells, thus improving the consistency of downstream differentiation protocols.

A central parameter for assessing stem cell health, particularly in iPS cell systems, is colony morphology. iPS cells, which are reprogrammed from somatic cells into a pluripotent state, form distinct colonies that reflect their pluripotent status, and alterations in shape, size or texture often precede changes in cell fate. These colonies serve as a key indicator of pluripotency, making their monitoring essential for ensuring reproducibility and consistency in stem cell cultures. CNNs such as AlexNet have been used to detect and score these colonies based on brightfield imaging, offering a label-free, scalable alternative to conventional staining or reporter-based methods[10].

Beyond snapshot-based quality assessment, AI has also enabled dynamic analysis of cellular behavior over time. For instance, VGG16-based models integrated with object detection and tracking pipelines have been used to analyze keratinocyte stem cell motion, providing kinetic information that correlates with cellular potency and viability[54]. This temporal dimension enables detection of gradual phenotypic drift that may precede overt differentiation.

Stem cell cultures are rarely synchronized. At any given time, cells in the same dish may occupy different stages along a differentiation continuum. This heterogeneity complicates analysis but also presents an opportunity for deeper insight if it can be properly resolved. AI has shown considerable power in this area. ResNet, known for its ability to train very deep networks via residual connections, has been applied to distinguish undifferentiated mouse embryonic stem cells from epiblast-like cells and, similarly, to separate iPS cells from their mesodermal derivatives (Fig. 3a)[17].

Custom CNNs trained on phase-contrast microscopy have effectively discriminated between undifferentiated iPS cells and iPS cell-derived hepatocytes[55], highlighting how tailored architectures can capture lineage-specific morphology. Similarly, AlexNet applied

to phase-contrast images has accurately identified iPS cell-derived endothelial cells within mixed cultures[8], supporting fast and robust classification.

In MSCs, where morphological variation is subtle, ResNet has enabled fine-grained phenotyping of subpopulations in imaging flow cytometry[56]. CNN-based approaches have also classified MSCs in label-free light microscopy, undergoing adipogenic[57], tenogenic[58] and osteogenic differentiation[59], enabling quantitative assessments of lineage commitment and differentiation potential.

Importantly, several studies have extended beyond binary classification to model gradual cell-state transitions in label-free light microscopy. A CNN-based pipeline classified C2C12 myoblasts into three distinct differentiation stages[60], while another identified six reprogramming stages in iPS cell cultures[61]. These multiclass models demonstrate deep learning's ability to capture continuous cell-fate trajectories rather than discrete end states.

While distinguishing stem cells at different differentiation stages is important, selecting cells before overt differentiation allows detection of earlier and more subtle fate commitments. AI-driven approaches are particularly well suited to this regime. For example, Xception-based models trained on imaging flow cytometry data, which combine brightfield and fluorescence imaging, have predicted whether neural stem cells would become neurons, astrocytes or oligodendrocytes, providing a window into fate commitment at the earliest morphological stages (Fig. 3b)[62]. In hematopoietic systems, a hybrid model combining LeNet with RNNs inferred lineage choices in brightfield microscopy several generations before traditional markers emerged, enabling the rapid assessment of hematopoietic differentiation[21].

Taken together, these advances highlight the enormous potential of AI in stem cell culture monitoring. By enabling real-time, noninvasive and highly accurate assessments of stem cell identity, quality and fate potential, AI is progressively influencing how researchers select, cultivate and characterize stem cells.

Stem cell-derived models. Beyond simple stem cell cultures, the formation of complex 3D structures such as organoids and stem cell-derived embryo models has revolutionized in vitro modeling of development, disease and tissue regeneration. These self-organizing systems mimic key aspects of organ architecture or early pre-implantation and post-implantation development. However, their inherent complexity and variability introduce substantial challenges for standardization, reproducibility and phenotypic consistency. AI-driven image analysis provides a scalable solution for objective quality control, morphological quantification and high-throughput phenotyping across model systems.

Similarly to traditional stem cell cultures, organoid differentiation trajectories vary widely, and subjective manual selection often introduces observer bias. To address this, deep-learning models have been used to predict differentiation status directly from brightfield images. For instance, DenseNet architectures trained on human iPS cell-derived kidney organoids achieved accurate, noninvasive assessments of differentiation potential, enabling earlier and more reliable selection of well-formed structures[19]. Likewise, a ResNet model outperformed expert annotations in classifying retinal differentiation from mouse embryonic stem cell-derived retinal organoids, allowing objective identification of developmental progress without requiring fluorescent reporters or destructive sampling[16].

Beyond early-stage classification, AI has proven instrumental in monitoring the structural quality and consistency of developing organoids. In cerebral organoids, U-Net-based pipelines applied to magnetic resonance imaging enable longitudinal tracking of size and shape[63], while *k*-means clustering has identified key morphological indicators such as Feret diameter as proxies for developmental robustness (Fig. 3c)[64]. In stem cell-derived embryo models, a ResNet-based CNN distinguished high-quality and low-quality mouse post-implantation models (ETiXs) at both early and advanced stages, enhancing reproducibility and enabling objective selection[5]. Related approaches applied to in vivo embryos have identified early morphological descriptors predictive of later morphogenetic outcomes[65], conceptually aligning with AI-based strategies that link early image-derived features to downstream developmental trajectories. Similarly, AI classified human blastoids into five distinct quality groups from brightfield images, demonstrating high-throughput phenotyping[66]. In liver and alveolar organoid cultures, ResNet architectures have been used to track morphogenesis, growth rates and symmetry features, allowing for automated quality monitoring[67]. AI also improves the robustness of the analysis pipeline itself, with SVM models used to detect and remove segmentation artifacts in intestinal organoids, thereby refining quantitative outputs and downstream analyses[4].

AI-driven phenotyping extends beyond quality control to detailed morphometric analysis and case–control comparisons. Advanced pipelines now combine segmentation, object detection and clustering to extract and interpret phenotypic features across organoid types. In brain organoids, a U-Net-based pipeline applied to light-sheet imaging enabled quantification of rosette-like structures, facilitating assessment of neuronal development[28]. In parallel, brightfield pipelines such as MOrgAna integrate MLP models with logistic regression to analyze shape and growth dynamics in large organoid populations[68]. Additional approaches use custom CNN architectures[58], the SegFormer model[33], Segment Anything[35] and fluorescence-based methods such as StarDist to segment cells and structures within brain organoids[69]. Instance segmentation models, which are particularly important when organoids overlap or form in clusters, have become a major focus. For example, Faster R-CNN is used in the OrgaQuant pipeline to localize and count intestinal organoids from brightfield images[37]. U-Net and Mask R-CNN have both been applied to densely seeded kidney organoid cultures[70]. By combining object detection with pixel-level segmentation, Mask R-CNN enables accurate delineation of individual organoids in crowded, high-throughput settings.

Object detection-based pipelines are particularly useful for real-time monitoring and for identifying overlapping stem cell-derived models within a culture dish. Deep-Orga, based on the YOLOX architecture, enables rapid detection and tracking of densely packed intestinal organoids under brightfield microscopy[43]. While ResNet is traditionally used for classification tasks, it can also serve as a robust feature extractor in detection and tracking frameworks. In hepatic and pulmonary organoid systems, paired ResNet modules have been used to automate long-term monitoring of organoid growth and morphology by capturing fine-grained features across time points[67].

Beyond segmentation and object detection, some AI-driven phenotyping pipelines assess organoid similarity through clustering, deep learning or integrating multiple computer vision tasks. Clustering has been used to characterize and differentiate morphological development in brain organoids across different stages based on segmented brightfield images and to systematically classify phenotypic subtypes in intestinal organoids[71]. To improve organoid reproducibility, ResNet-based Siamese networks have been utilized to identify morphologically similar organoids, facilitating standardized assessments[50]. Additionally, a multilevel pipeline integrating DenseNet for classification, U-Net for segmentation and YOLOv8 for object detection has been designed to analyze brain organoids using both brightfield and fluorescence microscopy, enabling multilevel phenotypic analysis[72].

As stem cell-derived models increase in complexity, the need for scalable and unbiased phenotyping grows. AI provides an integrated toolkit spanning early differentiation prediction and longitudinal monitoring, supporting standardization and high-throughput analysis.

Disease modeling and drug screening

A central goal of stem cell research is to uncover the mechanisms underlying human disease and to accelerate the development of new

therapies. iPS cells have become a cornerstone of disease modeling, offering patient-specific platforms that preserve genetic backgrounds and mimic in vivo biology more faithfully than traditional cell lines. These systems allow for the generation of organoids or differentiated cell types that recapitulate disease-specific phenotypes, providing a unique window into pathophysiology. Yet, many disease-related alterations are subtle, heterogeneous and dynamically regulated, making them difficult to detect using conventional imaging or manual inspection. Deep learning enables sensitive and scalable identification of such phenotypes, which is particularly important for rare diseases and personalized medicine.

In this context, CNNs have demonstrated an exceptional capacity to recognize disease-specific morphological patterns. For example, a custom CNN architecture trained on iPS cell-derived colonies from individuals with Huntington's disease successfully identified subtle morphological shifts that were not apparent to human experts[24]. This illustrates AI's capacity to uncover cryptic phenotypic signatures and stratify patient-derived lines.

AI's utility extends even further when applied to drug screening, an essential phase in the therapeutic development pipeline. Organoids derived from patient iPS cells offer a physiologically relevant setting for testing compound efficacy and toxicity, but analyzing responses across hundreds or thousands of samples quickly becomes infeasible through manual means. Deep-learning-based pipelines allow for rapid, standardized evaluation of treatment outcomes, greatly accelerating discovery and reducing human bias. These models can quantify treatment-induced changes in morphology, viability or marker expression, even when effects are subtle or heterogeneous across the organoid population.

Building on these capabilities, several AI tools automate the classification of organoid responses to drug treatment. For instance, D-CryptO, leveraging an Xception-based architecture, was applied to colon organoids to predict both structural maturity and drug responsiveness with high fidelity (Fig. 3d)[14]. This enabled rapid and noninvasive evaluation of treatment outcomes. Likewise, OrgaNet, a CNN-based model, classified organoid viability across diverse conditions, surpassing traditional ATP bioluminescence-based viability assays in sensitivity and reproducibility[73]. These tools exemplify how CNNs can replace or complement biochemical readouts with image-based assessments that preserve spatial context and allow real-time tracking.

In more complex scenarios, AI enables integrative analysis that goes beyond classification. One study combined a ResNet model for disease detection with a convolutional autoencoder for unsupervised feature extraction, revealing latent phenotypic clusters that correspond to drug-response profiles in Huntington's disease organoids[74]. This dual strategy improved discrimination and enabled an unbiased, data-driven stratification of compounds, thereby facilitating more refined drug-response profiling. Such approaches hold promise for drug-repurposing screens, where phenotypic outcomes may not follow predefined categories.

Segmentation models are also central to drug screening, as they allow precise measurement of organoid shape and behavior in response to treatment. For example, a U-Net-based model was applied to brightfield images of bladder cancer organoids. This approach, termed OrgaSegment, tracked drug responses through changes in organoid area without the need for staining[75] (Fig. 3e). In the specific context of cystic fibrosis, where swelling responses of intestinal organoids serve as functional readouts, the pipeline automatically quantified CFTR-mediated organoid swelling under different drug treatments. Other pipelines combine brightfield and fluorescence imaging. In one case, an FCN (the predecessor of U-Net) was used to extract morphology from brightfield images while detecting apoptosis markers in fluorescence channels. This enabled detailed time-lapse analysis of drug-induced cell death in patient-derived organoids[76].

Taken together, these advances highlight the transformative potential of AI in disease modeling and therapeutic screening. Deep learning enables robust and scalable analysis of organoid phenotypes, revealing subtle disease signatures and drug responses that conventional approaches often miss. By integrating classification, segmentation and clustering, unified AI pipelines enable systematic characterization of biological variability and treatment effects, advancing precision medicine. While stem cell-derived models still require in vivo validation, their integration with AI improves experimental rigor, reduces bias and boosts throughput, strengthening the link between discovery and clinical application.

Mechanistic insights

Beyond their established roles in quality control, phenotyping and drug screening, AI models are increasingly being leveraged to investigate the underlying biological mechanisms that govern the development of stem cell-derived models, stem cell differentiation and disease progression. By detecting patterns that elude human perception, these models enable researchers to probe questions related to morphogenesis, cell–cell communication and spatial organization, which are core aspects of developmental biology that have historically been challenging to quantify at scale.

Segmentation-based approaches deepen our mechanistic understanding by extracting features at single-cell resolution. StarDist-based 3D cell segmentation has enabled high-throughput characterization of organoid topology, facilitating the quantification of cell packing, spatial heterogeneity and neighborhood relationships across thousands of structures[30,77]. This allows researchers to bridge scales, from subcellular structures to tissue-level architecture (Fig. 3f). These tools could be particularly powerful when applied to perturbed environments, where conventional section-based analyses provide limited three-dimensional information. For example, studies of organoids cultured in microgravity have revealed how altered mechanical forces reshape epithelial architecture and lumen formation, offering clues about the physical constraints guiding morphogenesis[78]. Automated 3D segmentation could quantify these effects throughout whole organoids. In human iPS cell-derived cardiomyocytes, FCNs have been used to segment nuclei and quantify contractile behavior, offering a dynamic readout of electrophysiological maturation and rhythmicity (Fig. 3g)[26]. This enables investigation of how genetic or pharmacological perturbations affect coordinated beating.

Integrating deep learning with multimodal data enables functional insights into developmental processes. In one study, the MOrgAna pipeline combined deep-learning-based segmentation of stem cell-derived embryo models with transcriptomic data to link early metabolic states and morphological outcomes[79]. High glycolytic activity predicted successful trunk-like cell formation, while lower glycolysis and relatively increased oxidative phosphorylation led to disrupted morphogenesis and a bias toward neural-like cells. In another study, FCNs were used to extract morphological and fluorescence features from thousands of 3D gastruloid images, creating an AI-driven phenotypic landscape[80]. Integration with single-cell transcriptomic and epigenomic data then revealed regulatory modules of symmetry breaking and informed a dual Wnt-modulation strategy that enhanced anterior patterning.

Mechanistic insights further emerge from classification models trained on high-dimensional image datasets. In stem cell-derived embryo models, for example, CNNs such as ResNet, along with classical machine-learning algorithms like SVMs, have been applied to predict developmental success based on early morphological features. These models uncovered that a higher number of initial cells significantly increases the likelihood of successful embryoid formation (Fig. 3h)[5], illustrating how AI can connect early conditions to emergent tissue organization.

At the same time, while stem cell-derived systems provide experimentally tractable platforms to study developmental processes, their

https://doi.org/10.1038/s41592-026-03202-x

Challenges	Future
Data • Batch variability • Availability and annotation • Biases	• Open imaging datasets • AI foundation models • Generative AI forecasting • AI-integrated experimental workflow • AI-driven multimodal data integration • Personalized medicine
AI limitations • Model generalizability • Data complexity and infrastructure • Model interpretability and trust	
External factors • Regulatory and ethical considerations	

Fig. 4 | Challenges and future directions for integrating AI in stem cell biology. The adoption of AI in stem cell research is shaped by constraints and opportunities across data generation, computational modeling and regulatory frameworks. Data-related challenges include batch-to-batch variability, limited access to well-annotated training datasets and biases that reduce model generalizability. Computational barriers include high resource demands for training on large image datasets, as well as limited interpretability and user trust, because models often function as black boxes that are difficult to validate or integrate into decision-making. External constraints include evolving regulations and ethical concerns, particularly for stem cell-derived embryo models and clinical AI applications. Future directions focus on improved experimental and computational integration. Large open-source imaging datasets will support the development of both specialized models and general-purpose foundation models. Foundation models are flexible, pretrained architectures capable of tasks such as segmentation, classification and detection with minimal additional training, and may generalize across systems to streamline phenotyping, organoid classification and adaptive experimentation. Generative models enable synthetic data generation and may forecast experimental outcomes, including developmental trajectories, supporting in silico experimentation. Real-time AI-integrated workflows can link image acquisition with on-the-fly analysis and intervention, enabling closed-loop experimental design. Integration of multimodal data, including imaging, transcriptomics, proteomics and metadata, will improve modeling of cell state and behavior by linking morphology with molecular profiles. Patient-specific AI models integrating genetic, imaging and treatment data may enable optimized differentiation protocols, individualized organoids and tailored therapies, advancing personalized medicine through improved disease modeling and drug testing in stem cell-derived systems.

ability to recapitulate in vivo embryogenesis is inherently partial. Differences in geometry, mechanical constraints, signaling environments and extraembryonic interactions can limit direct generalization. These conceptual and technical gaps must be considered when transferring AI-derived insights from organoids or embryo-like models to in vivo embryos.

In summary, AI is increasingly enabling mechanistic insights into stem cell and developmental biology by bridging imaging data with biological interpretation. Combined with high-resolution microscopy such as confocal and light-sheet imaging, AI models can extract spatial and topological features, including curvature, connectivity and tissue geometry, that reveal how local microenvironments influence cell fate, pattern formation and morphogenetic processes. By correlating these features with molecular readouts, mechanical stress and dynamic behaviors, AI helps explain not only what occurs during development, but also how and why it unfolds. AI-driven analyses further reveal how genetic, chemical or physical perturbations disrupt normal self-organization. Together, these advances mark a shift from descriptive to explanatory models of development, positioning AI as a central tool for uncovering principles of tissue patterning and morphogenesis.

Challenges and practical insights

Despite the growing impact of AI in stem cell research, several technical, practical and ethical challenges continue to limit its broader adoption and generalizability. A persistent obstacle is the scarcity of high-quality, annotated datasets, particularly for specialized tasks such as classifying intermediate cell states, segmenting organoid substructures or predicting rare developmental outcomes (Fig. 4). Unlike general computer vision datasets, stem cell imaging requires expert annotations, which are time-consuming, subjective and often inconsistent across laboratories. To alleviate this bottleneck, researchers commonly turn to data augmentation techniques to artificially increase the training dataset diversity, as well as transfer learning. Models pretrained on large-scale datasets such as ImageNet can be fine-tuned on smaller biological datasets, improving performance despite the differences between natural images and microscopy data[19].

Dataset complexity also presents trade-offs. Low-complexity datasets, such as 2D static images of well-characterized cell lines, are easier to process and share but may overlook biologically relevant heterogeneity. In contrast, high-complexity datasets, including 3D time-lapse imaging of organoids or embryo-like structures, capture richer biological dynamics but require substantial computational resources (Fig. 4), including multi-terabyte storage, high-throughput pipelines and scalable infrastructures such as GPU clusters or cloud-based environments. Live imaging introduces additional optical constraints. Researchers must balance temporal resolution against phototoxicity and photobleaching, particularly in sensitive mammalian cells, while accounting for limited optical transparency and increased light scattering in thick 3D samples. Widely accessible modalities such as confocal or multiphoton microscopy are, therefore, often restricted in their ability to image large or optically complex structures, as their coupled illumination and detection paths limit penetration depth and multiview acquisition. These limitations can reduce signal quality and accuracy, particularly during long-term imaging of heterogeneous systems such as organoids.

Light-sheet fluorescence microscopy partially overcomes these limitations by selectively illuminating thin planes of the sample, reducing light exposure and enabling long-term volumetric imaging[81]. However, light-sheet fluorescence microscopy setups are technically demanding, require careful optical alignment and sample mounting, and generate large datasets that must be accommodated by downstream analysis pipelines. Imaging requirements therefore vary substantially across model systems: early-stage stem cell-derived embryo models can often be imaged effectively using standard confocal approaches, whereas larger and more complex organoids frequently require multiview or light-sheet imaging. The lack of standardized imaging platforms across laboratories remains a major challenge for reproducibility and large-scale AI deployment.

Data bias is another critical limitation (Fig. 4). Overrepresentation of specific experimental conditions, model systems or imaging modalities in the training data can reduce model performance on new datasets or across laboratories. Robust validation and benchmarking practices, such as predefined data splits, cross-validation, cross-instrument or batch testing, simple negative controls (for example, channel swaps), calibration or normalization checks, and standardized performance metrics, are essential for identifying and mitigating such biases. Just as crucial is the need for interdisciplinary collaboration: biologists can provide context for ambiguous annotations, highlight biologically relevant features and define meaningful endpoints, while computational scientists contribute algorithmic expertise and technical robustness. This dialogue is key to ensuring that AI models remain both scientifically insightful and technically sound.

Beyond data bias, a broader limitation lies in the generalization behavior of neural networks themselves. These models typically perform well on data distributions similar to those seen during training and may behave unpredictably when applied to out-of-distribution inputs[82]. Furthermore, deep-learning models often lack explicit physical constraints and can struggle with tasks that require geometric consistency[83].

Increasing imaging resolution and temporal depth also creates practical challenges. Large datasets require substantial storage and computational resources (Fig. 4), while training advanced models

often demands GPUs or tensor processing units that may be inaccessible to smaller laboratories. Efficient data management strategies, such as intelligent compression, distributed storage and automated metadata tracking, are therefore essential. To support reproducibility and cross-study comparability, the systematic reporting of imaging and sample metadata is increasingly important. A minimal microscopy–AI reporting checklist should include microscope configuration, objective parameters, voxel size, imaging settings, environmental conditions, sample passage number, acquisition software and version, file format and the underlying metadata schema. The use of electronic lab notebooks and standardized, machine-readable metadata formats is therefore becoming a necessary requirement for reproducible and scalable AI workflows[84].

Within this practical context, empirical studies in microscopy suggest that the performance gains from increasing dataset size are strongly task and model dependent. Classical convolutional architectures such as U-Net or ResNet often approach performance saturation with dataset sizes of approximately 10^2–10^3 annotated images, after which additional data tend to yield diminishing returns in raw accuracy but can still improve robustness and cross-dataset generalization[27]. In contrast, foundation models typically derive their advantage from large-scale pretraining on 10^5–10^8 images and are most beneficial in settings involving heterogeneous imaging modalities, substantial domain shifts or extremely sparse labeled data[6,48]. For well-defined microscopy tasks, classical architectures therefore often provide sufficient performance at lower computational cost.

At the same time, these models can fail under domain shift due to sensitivity to imaging conditions (for example, illumination, contrast, magnification, voxel size) and optical artifacts such as scattering or out-of-focus blur. In multi-laboratory datasets, models may also learn instrument-specific signatures rather than biologically meaningful morphology. In practice, lightweight adaptation strategies (for example, linear probes or adapters[85]) are often sufficient when the target data closely resemble the pretraining domain and the task aligns with learned primitives such as generic object boundaries. Full fine-tuning is typically required when domain shifts are substantial, new modalities or resolution regimes are involved or the task demands novel semantic distinctions.

Beyond these model-level considerations, usability represents a major barrier to adoption. Many stem cell and developmental biology laboratories lack dedicated computational expertise, and deploying advanced AI models often requires substantial technical overhead. Wider impact therefore depends on accessible tools, including graphical user interfaces, reproducible pipelines and integration with widely used bioimage platforms such as Fiji/ImageJ and Napari. Tools such as DeepImageJ[86] and ZeroCostDL4Mic[87] help lower these barriers by enabling nonspecialists to apply deep-learning models with minimal setup and thereby facilitate broader dissemination of AI-based workflows.

Interpretability is another major concern, especially in translational and clinical contexts. Most deep-learning models, especially convolutional architectures, function as black boxes, which makes it difficult to understand which features drive predictions or how errors arise (Fig. 4). This opacity poses challenges when AI tools are used to inform critical decisions, such as selecting organoids for drug testing or identifying patient-specific treatment responses. In such settings, explainable AI tools (for example, saliency maps[88] or Grad-CAM[89]; Fig. 3d) are increasingly being used to trace back decisions and enhance transparency. Importantly, these approaches explain model behavior, meaning which input regions or features influence a prediction, rather than directly revealing the underlying biological mechanisms driving the observed phenotype. To mitigate potential overinterpretation, complementary strategies such as counterfactual analyses[90] or systematic input-ablation experiments[91] can be used to test the robustness and causal relevance of highlighted features. In parallel, classical machine-learning models, such as SVMs or logistic regression, may also serve as interpretable baselines when datasets are small or model transparency is required[5].

One of the most important unmet challenges involves the disconnect between AI classification and real-time experimental control. While current models can accurately predict outcomes or classify developmental stages from recorded image data, they are typically applied only retrospectively, after an experiment is complete. No widely available technology allows real-time intervention based on AI predictions, such as selecting a normally developing organoid for further culture or targeting an abnormal structure for removal or rescue. Bridging this gap will require not only low-latency inference pipelines but also the development of responsive technologies, such as photoactivatable dyes, optogenetic tools, microfluidic environments or robotic micromanipulators that can interact with the sample in real time based on model output.

Finally, as AI systems move closer to clinical application, ethical and regulatory considerations become increasingly relevant (Fig. 4). Questions surrounding data privacy, informed consent and compliance with biomedical standards must be addressed, especially in personalized medicine or patient-derived organoid research. In this context, stem cell-derived systems are increasingly discussed as part of new approach methodologies[92] aimed at supporting the 3R principle (replace, reduce, refine) in preclinical research. Standardized image-based readouts supported by AI could contribute quantitative endpoints for applications such as developmental toxicity assessment. However, any regulatory uptake by agencies such as the US Food and Drug Administration and the European Medicines Agency would require rigorous validation, standardized imaging and analysis protocols, and transparent reporting of model performance and uncertainty. AI-enabled human stem cell-derived embryos also raise profound ethical questions. As algorithms improve and enable the in vitro development of these models beyond the 14-day guideline imposed on natural human embryos, new debates may emerge around human embryo identity, developmental potential and moral status[93]. Addressing these issues will require parallel progress in bioethical frameworks to ensure thoughtful governance and public engagement.

Given these constraints, selecting an appropriate AI approach remains nontrivial and depends on a combination of imaging modality, object geometry, annotation effort and throughput requirements. To provide practical guidance, Table 1 summarizes common experimental scenarios in stem cell imaging and maps them to suitable model families and training strategies, highlighting trade-offs between accuracy, data requirements and computational cost.

Conclusions and future directions

As the intersection of artificial intelligence and stem cell biology continues to mature, open-source biomedical imaging repositories are becoming increasingly essential for democratizing AI development (Fig. 4). Initiatives that aggregate large-scale, annotated datasets across stem cell-derived model types, differentiation protocols and experimental conditions will provide critical training resources and promote reproducibility across the community. These datasets will also enable benchmarking, stimulate competition and help identify areas where AI tools remain underdeveloped or biased.

Alongside data sharing initiatives, analytical frameworks are rapidly evolving. One promising direction is the emergence of foundation models, which are highly flexible, pretrained architectures capable of performing a wide range of tasks with minimal task-specific tuning (Fig. 4). Tools like Segment Anything, originally developed for natural images, are now being adapted to biomedical datasets[36,94], offering generalization capacity across imaging techniques and biological contexts. These models mark a shift away from narrowly specialized pipelines toward universal frameworks that can perform segmentation, classification and object detection simultaneously. These models may

https://doi.org/10.1038/s41592-026-03202-x

Table 1 | Decision framework for selecting AI models in stem cell imaging based on imaging modality, data complexity and experimental constraints

Imaging context	Key challenge	Recommended models	Training strategy
Brightfield, 2D, sparse objects	Low contrast, weak boundaries	CNNs (ResNet, DenseNet); U-Net	Supervised training with limited labels; strong augmentation
High-quality 2D brightfield or low-channel fluorescence, sparse objects	Minimal structural ambiguity	Foundation models	Zero-shot inference or light prompt-based adaptation
Brightfield, overlapping objects	Object separation	Object detection (YOLO, Faster R-CNN); instance segmentation (Mask R-CNN)	Bounding-box annotation to reduce labeling effort
Fluorescence, multichannel	Channel specificity, noise	U-Net, StarDist, multi-input CNNs	Channel-aware architectures; normalization
3D volumes	Compute and memory limits	3D U-Net; slice-wise 2D CNNs	Patch-based training; hybrid 2D/3D approaches
Time-lapse imaging	Temporal resolution versus phototoxicity	CNN + temporal aggregation; CNN–RNN hybrids	Frame-wise training with temporal aggregation
High geometry variability	Shape diversity	StarDist; transformer-based segmentation	Geometry-aware representations
Low label budget	Limited annotations	Transfer learning; self-supervised models (for example, DINO); foundation models	Fine-tuning pretrained representations
High throughput/real-time	Latency constraints	Lightweight CNNs; YOLO-family detectors	Model compression; inference-optimized architectures
Cross-lab generalization	Domain shift, reproducibility	CNNs with strong augmentation; foundation models	External benchmarking; domain-aware validation

accelerate phenotypic screening, streamline organoid classification and support adaptive experimentation during live imaging.

Simultaneously, generative AI has demonstrated remarkable capabilities in language, code and increasingly, in scientific applications. In stem cell research, it is beginning to be used for tasks such as generating synthetic training data[95]. Beyond data augmentation, large language models and agentic AI systems hold the potential to actively assist experimental workflows by autonomously generating analysis pipelines, selecting model architectures, tuning hyperparameters or integrating prior biological knowledge from the literature. Looking ahead, generative models may forecast experimental outcomes (Fig. 4), predicting developmental trajectories or system-level responses under different perturbations. These tools may eventually allow researchers to replace certain experimental steps with in silico simulations, reducing time, cost and resource consumption. Combined with explainable AI, such models could identify early predictors of morphogenesis, guide targeted data acquisition and enable more efficient experimental design. Integration of agentic large language models with image-based AI systems may further enable closed-loop scientific assistants that connect imaging data, experimental metadata and literature knowledge to support automated hypothesis generation and experimental planning[96,97].

Alongside advances in generative and foundation models, increasing attention is being given to integrating physical principles into AI frameworks. Physics-based models, long used to study tissue mechanics, are now combined with deep learning to create hybrid approaches that are both data driven and grounded in biophysical laws[98,99]. In particular, physics-informed neural networks offer a compelling strategy for inferring mechanical properties such as stress and tension directly from live imaging of stem cell-derived systems. By embedding force-balance equations into the training process, these models enable noninvasive, high-throughput mapping of biomechanical states during development. This map helps link mechanical cues to differentiation outcomes, optimizing culture conditions, and even steering morphogenetic processes in real time through feedback-informed experimentation.

Another emerging direction is the integration of AI-based analysis with automated experimental control (Fig. 4). Real-time classification systems could identify specific morphologies or developmental states as they arise, allowing dynamic intervention through robotic micromanipulators, microfluidic platforms or optogenetic tools. Such closed-loop systems would offer a powerful framework for dissecting the mechanisms underlying stem cell self-organization, including symmetry breaking, axis formation and tissue patterning. At the same time, they could be harnessed to improve culture outcomes by enriching for correctly assembled models and eliminating aberrant ones, thereby enhancing both the scientific insights and the experimental efficiency of stem cell-derived systems. In particular, the integration of such tools may lower the expertise required to operate advanced analysis pipelines, broadening accessibility and accelerating adoption in experimental laboratories.

Furthermore, AI has the potential to optimize experimental protocols themselves. By systematically analyzing large datasets spanning culture conditions, molecular perturbations and phenotypic outcomes, AI models can help identify optimal combinations of media, growth factors and physical environments to guide differentiation toward specific fates. Such data-driven protocol refinement could reduce variability, improve efficiency and lower the barrier to reproducing complex tissue models across laboratories.

A particularly promising frontier lies in the integration of multimodal data (Fig. 4). Combining live imaging with transcriptomics, proteomics or spatial data will allow AI models to go beyond morphology and infer causal links between structure, gene expression and cell fate[100]. For instance, models trained on paired image and single-cell RNA-sequencing data may eventually predict molecular signatures directly from morphology, offering noninvasive proxies for lineage specification or disease state. This would greatly enhance the interpretability and utility of stem cell-derived systems for mechanistic studies and therapeutic screening.

Another exciting application of AI in stem cell biology is its potential to advance personalized medicine (Fig. 4). By learning from patient-specific datasets, including genetic profiles, imaging data and treatment responses, AI can help tailor stem cell differentiation protocols to individual needs. For example, models trained on diverse patient-derived iPS cell lines could predict optimal conditions for generating personalized organoids or cell therapies. Such individualized

modeling could improve the fidelity of disease models and enable more precise drug testing, bringing stem cell technologies closer to clinical translation.

Looking further ahead, the convergence of foundation models, agentic AI, robotics and real-time imaging suggests a future in which AI systems function not only as analytical tools but also as autonomous experimental partners. As models grow more generalizable and tools become more accessible, AI is poised to move from a downstream analytic tool to a central component of experimental design, capable of steering experiments in real time, identifying new biological principles, and ultimately enhancing our capacity to model and manipulate human development and disease.

References

1. Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297 (1995).
2. Lever, J., Krzywinski, M. & Altman, N. Logistic regression. *Nat. Methods* **13**, 541–542 (2016).
3. Hartigan, J. A. & Wong, M. A. Algorithm AS 136: a *K*-means clustering algorithm. *Appl. Stat.* **28**, 100–108 (1979).
4. Lukonin, I. et al. Phenotypic landscape of intestinal organoid regeneration. *Nature* **586**, 275–280 (2020).
5. Caldarelli, P. et al. AI-based approach to dissect the variability of mouse stem cell-derived embryo models. *Nat. Commun.* **16**, 1772 (2025).
   This paper presented one of the first applications of AI-based image analysis to systematically characterize variability in stem cell-derived embryo models.
6. Dosovitskiy, A. et al. An image is worth 16 × 16 words: transformers for image recognition at scale. In *Proc. 9th International Conference on Learning Representations* (ICLR, 2021).
7. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2324 (1998).
8. Kusumoto, D. et al. Automated deep learning-based system to identify endothelial cells derived from induced pluripotent stem cells. *Stem Cell Rep.* **10**, 1687–1695 (2018).
9. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **60**, 84–90 (2017).
10. Fan, K. et al. A Machine learning assisted, label-free, non-invasive approach for somatic reprogramming in induced pluripotent stem cell colony formation detection and prediction. *Sci. Rep.* **7**, 13496 (2017).
11. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. Preprint at https://doi.org/10.48550/arXiv.1409.1556 (2015).
12. Szegedy, C. et al. Going deeper with convolutions. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 1–9 (2015).
13. Chollet, F. Xception: deep learning with depthwise separable convolutions. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 1800–1807 (2017).
14. Abdul, L. et al. D-CryptO: deep learning-based analysis of colon organoid morphology from brightfield images. *Lab Chip* **22**, 4118–4128 (2022).
15. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 770–778 (2016).
16. Kegeles, E., Naumov, A., Karpulevich, E. A., Volchkov, P. & Baranov, P. Convolutional neural networks can predict retinal differentiation in retinal organoids. *Front. Cell. Neurosci.* **14**, 171 (2020).
17. Waisman, A. et al. Deep learning neural networks highly predict very early onset of pluripotent stem cell differentiation. *Stem Cell Rep.* **12**, 845–859 (2019).
18. Huang, G., Liu, Z., van der Maaten, L. & Weinberger, K. Q. Densely connected convolutional networks. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 2261–2269 (2017).
19. Park, K. et al. Deep learning predicts the differentiation of kidney organoids derived from human induced pluripotent stem cells. *Kidney Res. Clin. Pract.* **42**, 75–85 (2023).
20. Kim, H. et al. Predicting multipotency of human adult stem cells derived from various donors through deep learning. *Sci. Rep.* **12**, 21614 (2022).
21. Buggenthin, F. et al. Prospective identification of hematopoietic lineage choice by deep learning. *Nat. Methods* **14**, 403–406 (2017).
    This paper demonstrated that deep learning can predict lineage decisions from time-lapse microscopy images.
22. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997).
23. Vaswani, A. et al. Attention is all you need. In *Adv. Neural Inf. Process. Syst.* **30**, 6000–6010 (2017).
24. Witmer, A., Theagarajan, R. & Bhanu, B. Triplet-net classification of contiguous stem cell microscopy images. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **20**, 2314–2327 (2023).
25. Long, J., Shelhamer, E. & Darrell, T. Fully convolutional networks for semantic segmentation. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 3431–3440 (2015).
26. Ahmadzadeh, E., Jaferzadeh, K., Shin, S. & Moon, I. Automated single cardiomyocyte characterization by nucleus extraction from dynamic holographic images using a fully convolutional neural network. *Biomed. Opt. Express* **11**, 1501–1516 (2020).
27. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. In *Lecture Notes in Computer Science* 234–241 https://doi.org/10.1007/978-3-319-24574-4_28 (2015).
28. Albanese, A. et al. Multiscale 3D phenotyping of human cerebral organoids. *Sci. Rep.* **10**, 21487 (2020).
29. Schmidt, U., Weigert, M., Broaddus, C. & Myers, G. Cell detection with star-convex polygons. In *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018* (eds. Frangi, A. F. et al.) Vol. 11071, 265–273 (Springer International, 2018).
30. Ong, H. T. et al. Digitalized organoids: integrated pipeline for high-speed 3D analysis of organoid structures using multilevel segmentation and cellular topology. *Nat. Methods* https://doi.org/10.1038/s41592-025-02685-4 (2025).
    This research enabled scalable, quantitative analysis of 3D organoid architecture at cellular resolution to advance organoid research from qualitative imaging toward systematic, data-driven tissue phenotyping.
31. Wang, X. et al. A novel deep learning segmentation model for organoid-based drug screening. *Front. Pharmacol.* **13**, 1080273 (2022).
32. Xie, E. et al. SegFormer: simple and efficient design for semantic segmentation with transformers. In *Advances in Neural Information Processing Systems* 34 (NeurIPS, 2021).
33. Schröter, J. et al. A large and diverse brain organoid dataset of 1,400 cross-laboratory images of 64 trackable brain organoids. *Sci. Data* **11**, 514 (2024).
34. Kirillov, A. et al. Segment Anything. In *Proc. IEEE/CVF Int. Conf. Computer Vision* 4015–4026 (IEEE, 2023).
35. Xing, X., Tang, C., Guo, Y., Kurniawan, N. & Yang, G. SegmentAnything helps microscopy images based on automatic and quantitative organoid detection and analysis. In *Medical Imaging 2024: Clinical and Biomedical Imaging* 58 (2024)
36. Archit, A. et al. Segment Anything for microscopy. *Nat. Methods* **22**, 579–591 (2025).
    This paper adapted the Segment Anything foundation model for microscopy, enabling general-purpose segmentation across diverse imaging modalities with minimal task-specific training.

37. Kassis, T., Hernandez-Gordillo, V., Langer, R. & Griffith, L. G. OrgaQuant: human intestinal organoid localization and quantification using deep convolutional neural networks. *Sci Rep.* **9**, 12479 (2019).
38. He, K., Gkioxari, G., Dollár, P. & Girshick, R. Mask R-CNN. In *Proc. IEEE International Conference on Computer Vision* 2961–2969 (IEEE, 2017).
39. Cai, Z. & Vasconcelos, N. Cascade R-CNN: delving into high quality object detection. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition* 6154–6162 (IEEE, 2018).
40. He, L. et al. Morphology-based deep learning enables accurate detection of senescence in mesenchymal stem cell cultures. *BMC Biol.* **22**, 1 (2024).
41. Ge, Z., Liu, S., Wang, F., Li, Z. & Sun, J. YOLOX: exceeding YOLO series in 2021. Preprint at https://doi.org/10.48550/arXiv.2107.08430 (2021).
42. Varghese, R. & M., S. YOLOv8: a novel object detection algorithm with enhanced performance and robustness. In *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)* 1–6 https://doi.org/10.1109/ADICS58448.2024.10533619 (IEEE, 2024).
43. Leng, B., Jiang, H., Wang, B., Wang, J. & Luo, G. Deep-Orga: an improved deep learning-based lightweight model for intestinal organoid detection. *Comput. Biol. Med.* **169**, 107847 (2024).
44. Li, Y., Mao, H., Girshick, R. & He, K. Exploring plain vision transformer backbones for object detection. In *Computer Vision – ECCV 2022* 280–296 (Springer, 2022).
45. Masci, J., Meier, U., Cireşan, D. & Schmidhuber, J. Stacked convolutional auto-encoders for hierarchical feature extraction. In *Artificial Neural Networks and Machine Learning – ICANN 2011* (eds. Honkela, T. et al.) 6791, 52–59 (Springer, 2011).
46. Koch, G., Zemel, R. & Salakhutdinov, R. Siamese neural networks for one-shot image recognition. In *Proc. 32nd Int. Conf. Mach. Learn. ICML Deep Learn. Workshop* (2015).
47. Caron, M. et al. Emerging properties in self-supervised vision transformers. In *Proc. IEEE/CVF International Conference on Computer Vision* 9630–9640 (IEEE, 2021).
48. Oquab, M. et al. DINOv2: learning robust visual features without supervision. Preprint at https://doi.org/10.48550/arXiv.2304.07193 (2023).
49. Moutakanni, T. et al. Cell-DINO: Self-supervised image-based embeddings for cell fluorescent microscopy. *PLoS Comput. Biol.* **21**, e1013828 (2025).
50. Presberger, J. et al. Correlation clustering of organoid images. Preprint at https://doi.org/10.48550/arXiv.2403.13376 (2024).
51. Poliwoda, S. et al. Stem cells: a comprehensive review of origins and emerging clinical roles in medical practice. *Orthop. Rev.* **14**, 37498 (2022).
52. Mamaeva, A. et al. Quality control of human pluripotent stem cell colonies by computational image analysis using convolutional neural networks. *Int. J. Mol. Sci.* **24**, 140 (2023).
53. Orita, K., Sawada, K., Koyama, R. & Ikegaya, Y. Deep learning-based quality control of cultured human-induced pluripotent stem cell-derived cardiomyocytes. *J. Pharmacol. Sci.* **140**, 313–316 (2019).
54. Hirose, T., Kotoku, J., Toki, F., Nishimura, E. K. & Nanba, D. Label-free quality control and identification of human keratinocyte stem cells by deep learning-based automated cell tracking. *Stem Cells* **39**, 1091–1100 (2021).
55. Marzec-Schmidt, K. et al. Artificial intelligence supports automated characterization of differentiated human pluripotent stem cells. *Stem Cells* **41**, 850–861 (2023).
56. Mukhopadhyay, R., Chandel, P., Prasad, K. & Chakraborty, U. Machine learning aided single cell image analysis improves understanding of morphometric heterogeneity of human mesenchymal stem cells. *Methods* **225**, 62–73 (2024).
57. Mai, M. et al. Morphology-based deep learning approach for predicting adipogenic and osteogenic differentiation of human mesenchymal stem cells (hMSCs). *Front. Cell Dev. Biol.* **11**, 1329840 (2023).
58. Dursun, G. et al. Development of convolutional neural networks for recognition of tenogenic differentiation based on cellular morphology. *Comput. Methods Programs Biomed.* **208**, 106279 (2021).
59. Lan, Y. et al. Morphology-based deep learning approach for predicting osteogenic differentiation. *Front. Bioeng. Biotechnol.* **9**, 1329840 (2022).
60. Niioka, H. et al. Classification of C2C12 cells at differentiation by convolutional neural network of deep learning using phase contrast images. *Hum. Cell* **31**, 87–93 (2018).
61. Chang, Y. -H. et al. Human induced pluripotent stem cell region recognition in microscopy images using Convolutional Neural Networks. In *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* 4058–4061 https://doi.org/10.1109/EMBC.2017.8037747 (2017).
62. Zhu, Y. et al. Deep learning-based predictive identification of neural stem cell differentiation. *Nat. Commun.* **12**, 2614 (2021). **Showed that label-free brightfield images can be used to predict neural stem cell fate at an early stage.**
63. Deininger, L. et al. An AI-based segmentation and analysis pipeline for high-field MR monitoring of cerebral organoids. *Sci Rep.* **13**, 21231 (2023).
64. Kachkin, D. et al. Deciphering brain organoids heterogeneity by identifying key quality determinants. Preprint at *bioRxiv* https://doi.org/10.1101/2025.01.13.632763 (2025).
65. Čapek, D. et al. EmbryoNet: using deep learning to link embryonic phenotypes to signaling pathways. *Nat. Methods* **20**, 815–823 (2023).
66. Fan, Z. et al. deepBlastoid: a deep learning model for automated and efficient evaluation of human blastoids. *Life Med.* **4**, lnaf026 (2025).
67. Bian, X. et al. A deep learning model for detection and tracking in high-throughput images of organoid. *Comput. Biol. Med.* **134**, 104490 (2021).
68. Gritti, N. et al. MOrgAna: accessible quantitative analysis of organoids with machine learning. *Development* **148**, dev199611 (2021).
69. Shi, J. Automatic quantitative analysis of brain organoids via deep learning. Preprint at https://doi.org/10.48550/arXiv.2211.00750 (2022).
70. MacDonald, M. et al. Improved automated segmentation of human kidney organoids using deep convolutional neural networks. In *Medical Imaging 2020: Image Processing* 11313, 832–839 (SPIE, 2020).
71. Brémond-Martin, C. et al. TDA-clustering strategies for the characterization of brain organoids. In *Workshop on the Ethical and Philosophical Issues in Medical Imaging* 13755, 113–122 (Springer Nature, 2022).
72. Kahveci, B. et al. BrAIn: a comprehensive artificial intelligence-based morphology analysis system for brain organoids and neuroscience. *Bioeng. Transl. Med.* https://doi.org/10.1002/btm2.70123 (2025).
73. Bian, X. et al. OrgaNet: a deep learning approach for automated evaluation of organoids viability in drug screening. In *Bioinformatics Research and Applications* (eds. Wei, Y. et al.) 411–423 https://doi.org/10.1007/978-3-030-91415-8_35 (Springer International, 2021).
74. Metzger, J. J. et al. Deep-learning analysis of micropattern-based organoids enables high-throughput drug screening of Huntington's disease models. *Cell Rep. Methods* **2**, 100297 (2022).

75. Lefferts, J. W. et al. OrgaSegment: deep-learning based organoid segmentation to quantify CFTR dependent fluid secretion. *Commun. Biol.* 7, 319 (2024).
76. Deben, C. et al. OrBITS: label-free and time-lapse monitoring of patient derived organoids for advanced drug screening. *Cell. Oncol.* **46**, 299–314 (2023).
77. Beghin, A. et al. Automated high-speed 3D imaging of organoid cultures with multi-scale phenotypic quantification. *Nat. Methods* **19**, 881–892 (2022).
This paper introduced an automated light-sheet imaging platform enabling high-throughput 3D imaging and deep learning-based multi-scale phenotypic analysis of large organoid populations.
78. Guo, D. et al. The critical role of YAP/BMP/ID1 axis on simulated microgravity-induced neural tube defects in human brain organoids. *Adv. Sci.* **12**, 2410188 (2025).
79. Villaronga-Luque, A. et al. Integrated molecular-phenotypic profiling reveals metabolic control of morphological variation in a stem-cell-based embryo model. *Cell Stem Cell* https://doi.org/10.1016/j.stem.2025.03.012 (2025).
80. Suppinger, S. et al. Multimodal characterization of murine gastruloid development. *Cell Stem Cell* **30**, 867–884 (2023).
81. Corsetti, S., Gunn-Moore, F. & Dholakia, K. Light sheet fluorescence microscopy for neuroscience. *J. Neurosci. Methods* **319**, 16–27 (2019).
82. Ovadia, Y. et al. Can you trust your model's uncertainty? Evaluating predictive uncertainty under dataset shift. In *Advances in Neural Information Processing Systems* 32 (NeurIPS, 2019).
83. Bronstein, M. M., Bruna, J., Cohen, T. & Veličković, P. Geometric deep learning: grids, groups, graphs, geodesics, and gauges. Preprint at https://doi.org/10.48550/arXiv.2104.13478 (2021).
84. Leipzig, J., Nüst, D., Hoyt, C. T., Ram, K. & Greenberg, J. The role of metadata in reproducible computational research. *Patterns* **2**, 100322 (2021).
85. Zanella, M. & Ben Ayed, I. Low-rank few-shot adaptation of vision-language models. In *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition Workshops* 1593–1603 (2024).
86. Gómez-de-Mariscal, E. et al. DeepImageJ: a user-friendly environment to run deep learning models in ImageJ. *Nat. Methods* **18**, 1192–1195 (2021).
This study introduced a plugin that enables running trained deep learning models directly within the ImageJ/Fiji platform.
87. Von Chamier, L. et al. Democratising deep learning for microscopy with ZeroCostDL4Mic. *Nat. Commun.* **12**, 2276 (2021).
88. Simonyan, K., Vedaldi, A. & Zisserman, A. Deep inside convolutional networks: visualising image classification models and saliency maps. Preprint at https://doi.org/10.48550/arXiv.1312.6034 (2013).
89. Selvaraju, R. R. et al. Grad-CAM: visual explanations from deep networks via gradient-based localization. https://doi.org/10.48550/arXiv.1610.02391 (2016).
90. Verma, S. et al. Counterfactual explanations and algorithmic recourses for machine learning: a review. *ACM Comput. Surv.* **56**, 1–42 (2024).
91. Sun, J. et al. Improving explainable AI with patch perturbation-based evaluation pipeline: a COVID-19 X-ray image analysis case study. *Sci. Rep.* **13**, 19488 (2023).
92. European Medicines Agency. Regulatory acceptance of new approach methodologies (NAMs) to reduce animal use testing. https://www.ema.europa.eu/en/human-regulatory-overview/research-development/ethical-use-animals-medicine-testing/regulatory-acceptance-new-approach-methodologies-nams-reduce-animal-use-testing (2025).
93. Gyngell, C., Lynch, F., Sawai, T. & Savulescu, J. Stem cell-derived embryo models: moral advance or moral obfuscation? *J. Med. Ethics* https://doi.org/10.1136/jme-2023-109605 (2024).
94. Pachitariu, M., Rariden, M. & Stringer, C. Cellpose-SAM: superhuman generalization for cellular segmentation. Preprint at *bioRxiv* https://doi.org/10.1101/2025.04.28.651001 (2025).
95. Wu, Z., Park, J., Steiner, P. R., Zhu, B. & Zhang, J. X. J. Generative adversarial network model to classify human induced pluripotent stem cell-cardiomyocytes based on maturation level. *Sci. Rep.* **14**, 27016 (2024).
96. Lei, W. et al. IO Chatbot: a community-driven AI assistant for integrative computational bioimaging. *Nat. Methods* **21**, 1368–1370 (2024).
97. Royer, L. A. Omega — harnessing the power of large language models for bioimage analysis. *Nat. Methods* **21**, 1371–1373 (2024).
98. Ichbiah, S., Delbary, F., McDougall, A., Dumollard, R. & Turlier, H. Embryo mechanics cartography: inference of 3D force atlases from fluorescence microscopy. *Nat. Methods* **20**, 1989–1999 (2023).
99. Schmitt, M. S. et al. Machine learning interpretable models of cell mechanics from protein images. *Cell* **187**, 481–494 (2024).
100. Hu, J. et al. SpaGCN: integrating gene expression, spatial location and histology to identify spatial domains and spatially variable genes by graph convolutional network. *Nat. Methods* **18**, 1342–1351 (2021).

Acknowledgements

This work was supported by the Helmholtz Association under the joint research school 'HIDSS4Health' – Helmholtz Information and Data Science School for Health to L.D., and the Helmholtz program NACIP to R.M. and L.D. This work was supported by the NIH through the Director's Pioneer Award (HD104575A to M.Z.-G.) and R01HD101489 (GRANT13035316 to M.Z.-G.), the NOMIS Foundation (12540449 to M.Z.-G.), the Wellcome Trust (207415/Z/17/Z to M.Z.-G.) and the Open Philanthropy Project (to M.Z.-G.).

Author contributions

Conceptualization, figure preparation and writing: L.D. and P.C.; supervision: R.M. and M.Z.-G.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence should be addressed to Luca Deininger, Paolo Caldarelli, Magdalena Zernicka-Goetz or Ralf Mikut.

Peer review information *Nature Methods* thanks Patrick Müller and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Madhura Mukhopadhyay, in collaboration with the *Nature Methods* team.