



The Next Generation of Biological Image Analysis: Machine Learning and Generative AI in Biotechnology

Masoud Arabfard ^{1*}, Ali Ghofrani ¹

¹ Artificial Intelligence in Health Research Center, Biomedicine Technologies Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran

Corresponding Author: Masoud Arabfard, Assistant Professor, Artificial Intelligence in Health Research Center, Biomedicine Technologies Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran. Tel: +989132649405, E-mail: arabfard@gmail.com

Received August 24, 2026; Accepted September 18, 2026; Online Published September 30, 2026

Biotechnology has long been a visually oriented discipline, evolving from early microscopic observations of cellular structures to contemporary high-content screening (HCS) platforms capable of producing millions of fluorescent and phase-contrast images each day.¹ Despite these technological advances, image analysis remains a major bottleneck.² Traditional approaches continue to rely substantially on handcrafted features, threshold-based segmentation, and supervised learning algorithms that depend on extensive manual annotation.³ Such limitations are particularly problematic because biological systems exhibit complex spatial organization, temporal dynamics, and multiscale characteristics that conventional computer vision pipelines often fail to adequately represent.⁴

Recent advances in deep learning (DL), particularly self-supervised vision transformers (ViTs) and generative diffusion models, have the potential to fundamentally reshape image-based biotechnology rather than merely provide incremental improvements to existing analytical workflows. By learning rich and transferable representations directly from large-scale unlabeled image datasets, these approaches can capture the complex visual patterns and latent biological features underlying cellular and multicellular systems.⁵ This capability may facilitate the interpretation of biological structures across multiple organizational scales, ranging from subcellular organelles and individual cells to organoids, tissues, and microbial colonies, thereby enabling biological insights that may remain inaccessible to conventional image-analysis strategies.

To date, biological image analysis has been predominantly driven by supervised convolutional neural networks (CNNs) applied to tasks such as cell segmentation and counting, organelle tracking, phenotypic classification, and microbial colony identification.⁶ Although these models have

demonstrated substantial performance across diverse applications, their effectiveness is strongly dependent on the availability of large, high-quality, manually annotated datasets. Generating hundreds of thousands of expert-labeled images is often labor-intensive, time-consuming, and economically prohibitive, particularly for biotechnology applications involving diverse experimental conditions and complex biological phenotypes. This annotation dependency represents a critical barrier to the scalability, generalizability, and broader adoption of current DL-based imaging approaches.

Recent studies have advocated a transition toward self-supervised learning (SSL), particularly approaches based on masked autoencoders (MAEs) and contrastive learning frameworks such as SimCLR.⁷ These methods can learn informative and transferable visual representations directly from large collections of unlabeled biological images, substantially reducing dependence on costly manual annotation. Emerging evidence further suggests that SSL-based models trained on cell-line imaging datasets can outperform conventional supervised CNNs in predicting drug responses from morphological phenotypes alone.⁸ This development highlights the potential of representation learning to enable scalable, label-free phenotypic screening, reducing the reliance on predefined molecular markers or extensive expert annotation. A persistent challenge in biological imaging, however, is the fundamental trade-off among spatial resolution, temporal resolution, and phototoxicity.^{9,10}

Generative DL, provides a powerful framework for overcoming several persistent limitations in biological imaging.¹¹ By learning complex mappings between different image representations, these models can reconstruct, synthesize, and predict biologically meaningful visual information that is difficult to obtain directly through conventional imaging approaches.

Modern biotechnology increasingly relies on the integration of complementary imaging modalities, each capturing distinct aspects of biological organization. Conventional multimodal integration strategies are often highly specialized, rule-based, and dependent on modality-specific processing pipelines, limiting their flexibility and scalability. To address these limitations, cross-modal transformer architectures could provide a promising unified framework for integrating heterogeneous biological imaging modalities. By learning shared latent representations across complementary imaging domains, these models can establish meaningful correspondences between distinct visual signals and facilitate the reconstruction or imputation of missing modalities.

A persistent concern surrounding the application of DL in biotechnology is its perceived “black-box” nature, which can limit biological interpretability and confidence in model-derived predictions. Nevertheless, recent advances in explainable artificial intelligence (XAI), including attention rollout, gradient-weighted class activation mapping (Grad-CAM), and feature-attribution methods, provide opportunities to move beyond prediction toward mechanistic insight and biological discovery.¹²

DL systems have the potential to move beyond prediction toward hypothesis generation, transforming biological image analysis into a more predictive and discovery-oriented discipline. However, realizing this potential requires overcoming the scarcity and fragmentation of large-scale, standardized, publicly accessible bioimaging datasets. Establishing coordinated bioimage repositories, building on resources such as EMPIAR, with standardized metadata and validated annotations, alongside federated learning frameworks that enable collaborative model development without sharing sensitive data, could substantially improve scalability and generalizability.¹³ In parallel, standardized community benchmarks and challenges could facilitate reproducible evaluation and fair comparison of models across tasks such as cell tracking, organoid assessment, phenotypic profiling, and microbial imaging, thereby accelerating the responsible integration of DL into biological research and biotechnology.

The proposed DL-based imaging paradigm has broad applicability across biotechnology, including cell and gene therapy, microbial biotechnology, tissue engineering, drug discovery, and synthetic biology. These applications illustrate the potential of advanced DL approaches to address diverse biological imaging challenges across multiple biotechnology domains.

In conclusion, Emerging advances in deep learning are reshaping biological imaging by enabling more efficient, integrative, and interpretable approaches to the analysis of complex biological systems. Rather than replacing human expertise, these technologies can augment researchers' ability to extract meaningful patterns, generate hypotheses,

and connect visual phenotypes with underlying biological processes. However, realizing this potential will require open, collaborative, and reproducible efforts, including standardized datasets, robust benchmarking, and transparent computational workflows across academia and biotechnology. Ultimately, biological images are not merely collections of pixels but multidimensional representations of the organization, dynamics, and complexity of life; unlocking their full scientific value will depend on computational approaches capable of capturing this biological depth.

Conflict of Interest Disclosures

The authors declare that they have no conflicts of interest.

References

1. Zhu Y, Ma Y. Biotechnology-Driven Innovations in Visual Communication: Harnessing Digital Imaging Technology and Biomicrographics for Commercial Design Solutions. *J Commer Biotechnol*. 2025;30(2):98-108. doi:10.5912/jcb2062
2. Jung SK. Image Processing and Analysis for Biotechnology and Bioprocess Engineering. *Appl Sci*. 2024;14(2):711. doi:10.3390/app14020711
3. Xu Y, Quan R, Xu W, Huang Y, Chen X, Liu F. Advances in medical image segmentation: A comprehensive review of traditional, deep learning and hybrid approaches. *Bioengineering*. 2024;11(10):1034. doi:10.3390/bioengineering11101034
4. Walpole J, Papin JA, Peirce SM. Multiscale computational models of complex biological systems. *Annu Rev Biomed Eng*. 2013;15(1):137-54. doi:10.1146/annurev-bioeng-071811-150104
5. Pastore VP, Ciranni M, Bianco S, Fung JC, Murino V, Odone F. Efficient unsupervised learning of biological images with compressed deep features. *Image Vision Comput*. 2023;137:104764. doi:10.1016/j.imavis.2023.104764
6. Singh H, Kaur H. A systematic survey on biological cell image segmentation and cell counting techniques in microscopic images using machine learning. *Wireless Pers Commun*. 2024;137(2):813-51. doi:10.1007/s11277-024-11379-7
7. Chen T, Kornblith S, Norouzi M, Hinton G. A simple framework for contrastive learning of visual representations. In: *International conference on machine learning*. PmlR. 2020. pp. 1597-1607.
8. Husain SS, Bober J, Irizar A, Bober M. Bridging self-supervision and mechanism of action discovery in morphological profiling. In: *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE. 2025. pp. 4278-4285.
9. Peng X, Huang X, Du K, Liu H, Chen L. High spatiotemporal resolution and low photo-toxicity fluorescence imaging in live cells and *in vivo*. *Biochem Soc Trans*. 2019;47(6):1635-50. doi:10.1042/BST20190020
10. Zhang M, Li R, Fu S, Kumar S, McGinty J, Qin Y, Chen L. Deep learning enhanced light sheet fluorescence microscopy for *in vivo* 4D imaging of zebrafish heart beating. *Light Sci Appl*. 2025;14(1):92. doi:10.1038/s41377-024-01710-z
11. Khazrak I, Takhirova S, Rezaee MM, Yadollahi M, Green II RC, Niu S. Addressing small and imbalanced medical image datasets using generative models. *Artifi Intell Appl*.

2025. doi:[10.47852/bonviewAIA52026661](https://doi.org/10.47852/bonviewAIA52026661)
12. Cheng Z, Wu Y, Li Y, Cai L, Ilnaini B. A comprehensive review of explainable artificial intelligence (XAI) in computer vision. *Sensors*. 2025;25(13):4166. doi:[10.3390/s25134166](https://doi.org/10.3390/s25134166)
 13. Sarkans U, Chiu W, Collinson L, Darrow MC, Ellenberg J, Grunwald D, et al. REMBI: Recommended Metadata for Biological Images—enabling reuse of microscopy data in biology. *Nat Methods*. 2021;18(12):1418-22. doi:[10.1038/s41592-021-01166-8](https://doi.org/10.1038/s41592-021-01166-8)