

Review Article

Volume 15 Issue 3

July-Sept 2026

BIOIMAGE INFORMATICS: A REVIEW ON THE ADVANCEMENT OF ITS TECHNIQUES

Smriti Kumari, *Sajal Kumar Chowdhury

Assistant Professor, Shivam College of Higher Studies, Fatuha, Patna, Bihar – 803202

*Corresponding author's Email ID: chowdhurysajal2016@gmail.com

Abstract

In recent years, the rapid accumulation of complex molecular and cellular microscopic images has created significant challenges for the image-computing community. Increasing attention has therefore been directed toward the development of novel image-processing, data-mining, database, and visualization techniques for extracting, comparing, searching, and managing biological information from these data-intensive resources. This emerging area of bioinformatics is known as bioimage informatics. This article reviews the advancement of bioimage informatics from several perspectives, including its applications, key techniques, available tools, databases, and other resources. Applications such as high-throughput and high-content phenotyping and atlas building for model organisms demonstrate the importance of bioimage informatics. The major techniques supporting these applications include bioimage feature identification and extraction, image segmentation, tracking, registration, annotation, image mining, data management, and visualization. The article also provides an overview of representative bioimage databases, image-analysis tools, and other resources used in biological image analysis.

Keywords

Bioimage informatics; Bioinformatics; Image processing; Image analysis; Image segmentation; Image mining; Biological imaging.

1. INTRODUCTION

Over the last several decades, numerous biomedical imaging techniques have been developed, ranging from whole-organism imaging at the millimetre scale to imaging at the level of individual molecules with nanometre-scale resolution. Commonly used biological imaging techniques include confocal and two-photon laser-scanning microscopy, scanning and transmission electron microscopy, and related imaging modalities [1,2].

Novel imaging techniques, including photoactivated localization microscopy (PALM), stochastic optical reconstruction microscopy (STORM), and stimulated emission depletion (STED) microscopy, have substantially surpassed the resolution of conventional optical microscopy and can localize individual proteins that are only a few nanometres apart [3-5]. Along with advances in image digitization and storage and biological labelling techniques, such as green fluorescent protein (GFP), enhanced GFP (EGFP), Dronpa, and Brainbow combinatorial labelling, the number of biological images acquired in digital form has increased rapidly [6-9].

Large biological image databases, including the Allen Brain Atlas and the Cell Centered Database (CCDB), have also become available [10,11]. These image datasets may contain two-dimensional (2D) or three-dimensional (3D) spatial information, multiple colour channels corresponding to different molecular reporters, four-dimensional (4D) spatiotemporal information describing developing tissues or moving cells, and co-localized biological signals such as gene-expression information. In addition, high-throughput screening experiments involving RNA interference (RNAi) and chemical compounds can generate very large collections of biological images [12-14].

Analysis of these images is essential for addressing numerous biological questions, including the identification of cellular phenotypes, classification of neurons, analysis of gene-expression patterns, and investigation of dynamic cellular processes [15-18]. However, the complexity and volume of biological and biomedical image data present substantial computational challenges.

Bioimage informatics has emerged as a natural extension of biomedical image analysis. It involves the development and application of image-processing, computational, data-mining, database, and visualization techniques to extract, compare, search, organize, and manage biological information contained within images [19]. Unlike many conventional medical images, bioimages may contain densely packed cells, complex neuronal structures, rapidly changing cellular components, multiple imaging channels, and substantial variations in morphology and intensity.

For example, astrocytes, microglia, and neurons may be densely intertwined, while rapidly changing microtubule structures may require analysis of four-dimensional time-lapse image sequences. Specialized computational approaches are therefore required for the analysis of such complex structures [20].

Another important challenge is the large size of biological image datasets. A single 3D image stack may occupy several hundred megabytes or even several gigabytes and may contain multiple colour channels. Objects of interest, such as neuronal structures, may show considerable variation in morphology and intensity between images. Consequently, thousands or millions of images may need to be analysed automatically in a high-throughput manner rather than through months or years of manual work [19].

These challenges have stimulated the development of novel bioimage-informatics algorithms and systems, particularly in three major areas: image processing and mining, image databases and data management, and image visualization. Bioimage informatics has developed rapidly, supported by workshops, research communities, open-source software, and specialized databases [19,21].

The objective of this review is to summarize major developments in bioimage informatics, focusing on its applications, fundamental computational techniques, available software tools, databases, and other resources.

2. SIGNIFICANCE AND APPLICATIONS

Like many other engineering disciplines, bioimage informatics is strongly application-driven. Important applications include high-throughput and high-content cellular phenotyping, atlas construction for model organisms, analysis of dynamic cellular processes, and integration of image-based information with other bioinformatics approaches [19,21].

2.1 High-throughput and high-content analysis of cellular phenotypes

Large-scale screening of cellular phenotypes at the whole-cell or subcellular level is important for determining gene functions, characterizing cellular pathways, drug discovery, and cancer research. Automated image-analysis platforms such as CellProfiler have been developed to analyse cellular images rapidly and extract information including cell number, cell size, morphology, protein-expression levels, and cell-cycle characteristics [22,23].

Such systems can be applied to cells exhibiting different morphological characteristics, including *Drosophila* and human cell lines. Intelligent human-computer interfaces and

content-based image-retrieval approaches have also been investigated for high-content analysis of neuronal images [24,25].

Quantitative analysis of cellular morphology can provide information about signalling pathways associated with processes such as cell protrusion, adhesion, and mechanical stress. At the subcellular level, three-dimensional protein-localization patterns associated with cellular structures such as the nucleus, nucleolus, mitochondria, and cytoskeleton can be characterized using fluorescence-image features, including texture features and Zernike moments [26].

Spatial patterns can also be incorporated into classification and prediction approaches. Such image-derived morphological signatures have been used to investigate cellular signalling networks and disease-associated phenotypes [27]. More systematic approaches, including generative models of protein subcellular localization, can provide useful information for systems-biology studies [28].

2.2 Atlas building for model organisms

Bioimage-informatics methods have been applied to several widely used model organisms, including mouse, *Drosophila*, *Caenorhabditis elegans*, and zebrafish [10,17,29-32]. An important application is the construction of digital atlases that integrate anatomical information with gene-expression and ontological information.

The Allen Brain Atlas, for example, provides genome-wide gene-expression information for the mouse brain and has become an important resource for studying gene expression in relation to brain anatomy [10]. The Anatomic Gene Expression Atlas (AGEA) provides an interactive three-dimensional representation of gene-expression patterns in the adult mouse brain.

Single-cell analysis of whole organisms can also contribute to understanding cellular functions and neuronal circuits. Such analyses become possible when cells exhibit distinctive characteristics based on their three-dimensional positions, morphology, developmental lineage, gene-expression patterns, or other functional properties.

Digital atlases of *C. elegans* and other model organisms demonstrate how image analysis can transform large collections of microscopy data into structured biological knowledge [17,29].

2.3 Understanding dynamic processes in cells and living organisms

Bioimage informatics is also important for studying dynamic biological processes. Microtubules, which are components of the cytoskeleton, undergo continuous assembly

and disassembly and play essential roles in several cellular processes, including cell division.

By imaging fluorescently labelled microtubules, researchers can quantify dynamic characteristics such as growth and shortening rates. Computational methods including hidden Markov models, particle filtering, graph-based approaches, multiscale detection, and object-linking methods have been applied to track microtubule dynamics in time-lapse microscopy [33-37].

Tracking and classification of cellular structures over time can provide quantitative information about developmental processes and cellular behaviour. Similar computational approaches have been applied to cell-lineage analysis and other dynamic biological systems [38].

2.4 Integration of bioimage informatics with bioinformatics

Bioimage-informatics techniques can also be combined with conventional bioinformatics approaches. For example, image-based clustering of embryonic gene-expression patterns can be integrated with comparative genomics to identify sequence motifs that may have regulatory functions [39].

The combination of image-derived phenotypes with genomic, transcriptomic, and other molecular datasets can provide a more comprehensive understanding of biological systems. Such integration represents an important direction for systems biology and quantitative biology.

3. BASIC TECHNIQUES

The complexity of biological image data requires a broad range of image-analysis, machine-learning, pattern-recognition, and data-mining techniques. Data-management and visualization methods are also essential components of bioimage-informatics workflows [19,21].

The major computational techniques include feature extraction and selection, image segmentation, object detection and classification, tracking, image registration, annotation, image mining, database management, and visualization.

3.1 Feature extraction and selection

Image features provide fundamental representations of pixels, voxels, and higher-level objects. Useful image features may describe statistical, geometrical, morphological, textural, intensity-based, frequency-based, and topological properties.

Identification of appropriate image features is fundamental to bioimage analysis. Points, edges, curves, corners, boundaries, and textures have been used to analyse and track dynamic fluorescence images [40].

Feature extraction can be based on spatial representations. For example, local features derived using Gaussian-mixture-model decomposition have been used to characterize and compare gene-expression patterns [41]. Global representations based on eigenvector decomposition and wavelet features have also been investigated for classification and automated annotation of biological image patterns [39,42].

Other feature-extraction approaches include independent component analysis and invariant moments [43,44]. A complementary strategy is to generate a large and diverse collection of image features. For example, morphological and texture features can be used to characterize three-dimensional protein-localization patterns [26].

The WND-CHARM framework demonstrated the use of large collections of image features for biological image classification, including pixel statistics, texture features, polynomial decompositions, object-related features, and standard image transformations such as Fourier and wavelet transforms [45].

However, a very large feature set may contain redundant information and can negatively affect classifier performance. Feature-selection methods such as minimum redundancy–maximum relevance (mRMR) can therefore be used to identify informative subsets of features [46].

3.2 Image segmentation

Image segmentation is one of the fundamental steps in many bioimage-informatics applications. Its primary objective is to identify and separate objects of interest from the background or from other objects within an image.

Segmentation can be difficult because biological images may contain low signal-to-noise ratios, overlapping objects, irregular shapes, variable intensities, and considerable morphological variation.

The choice of segmentation method depends strongly on the properties of the objects being analysed. For example, texture features may be useful for chromatin segmentation, whereas concavity and shape-related features may be useful for nuclear segmentation.

One major class of segmentation problems involves globular objects such as nuclei and cells in 2D or 3D images. Common approaches include model-based segmentation,

watershed segmentation, Gaussian-mixture-model estimation, active contours, and deformable models [47-49].

For densely packed or irregular biological structures, more specialized algorithms may be necessary. Segmentation performance can also be improved by incorporating prior information about object shape, intensity, texture, or spatial relationships.

3.3 Object detection, classification and tracking

After segmentation, biological objects can be detected, classified, and tracked. Classification techniques are useful for assigning biological meaning to image objects, while tracking methods allow their movement and behaviour to be analysed over time.

Object tracking is particularly important in time-lapse microscopy, where cells, organelles, microtubules, or other structures may change position, shape, and intensity. Particle-filtering approaches, graph-based methods, and model-based tracking have been used for such applications [33-37].

3.4 Image registration

Image registration involves aligning images acquired at different times, from different samples, or using different imaging modalities. It is especially important in constructing digital atlases and comparing biological structures across specimens.

Registration methods may involve rigid, affine, or non-rigid transformations depending on the biological application and the degree of morphological variation [50].

3.5 Image annotation

Annotation converts image information into meaningful biological descriptions. These descriptions may include cell identity, anatomical region, gene-expression pattern, cellular phenotype, or other ontology terms.

Annotation may be performed manually, semi-automatically, or automatically. Automated annotation can be formulated as a classification problem in which multiple biological labels may be assigned to the same image or object [19,39].

4. AVAILABLE TOOLS AND RESOURCES

Numerous software tools have been developed for different aspects of bioimage analysis and management.

4.1 Image formats and input/output

Microscope manufacturers use different file formats for storing raw image data and associated metadata. Therefore, the ability to read, write, convert, and manage different image formats is important.

ImageJ is an open-source Java-based image-processing platform capable of reading and processing many types of biological images and supporting numerous plugins [51]. Bio-Formats provides a framework for reading and writing a wide range of biological image formats and associated metadata [52].

4.2 Image-analysis tools

ImageJ is widely used for biomedical image processing, visualization, and quantitative analysis [51]. The Insight Segmentation and Registration Toolkit (ITK) provides algorithms for image segmentation and registration [53].

Other platforms include MATLAB-based image-processing tools and specialized software for three-dimensional biological image analysis.

CellProfiler provides a flexible environment for automated analysis of biological images and high-content screening [22,23]. Cell-ID was developed for cell detection and tracking, while other specialized tools have been developed for neuronal tracing, zebrafish imaging, and other biological applications [54].

V3D provides tools for visualization and quantitative analysis of large-scale three-dimensional biological image datasets and has been applied to neuronal imaging [55].

4.3 Database and annotation tools

The Open Microscopy Environment (OME) provides a framework for managing microscopy image data and metadata [56]. OMERO extends this framework by providing tools for image storage, management, visualization, and annotation.

Other systems have been developed for storing and managing large numbers of biological images and their associated metadata. Such systems are essential for reproducible analysis and collaborative research.

Three-dimensional annotation tools can also link segmented biological objects to the corresponding regions of the original image. Such tools facilitate the construction of digital atlases and the assignment of biological identities to image-derived objects.

4.4 Visualization tools

Visualization is essential for understanding multidimensional, multichannel biological images. Confocal image stacks and other 3D datasets often require specialized visualization software.

Commercial platforms such as Amira and Volocity have been used for multidimensional biological image visualization. Open-source platforms such as ImageJ/Fiji and UCSF

Chimera provide additional capabilities for image visualization and three-dimensional data exploration [51,57].

5. CONCLUSION

Bioimage informatics has emerged as an important interdisciplinary field at the interface of biology, microscopy, image processing, computer vision, machine learning, data mining, database technology, and visualization. The rapid growth of biological imaging technologies has generated large and complex datasets that cannot be efficiently analysed using manual approaches alone.

The development of automated feature extraction, segmentation, classification, tracking, registration, annotation, image mining, database management, and visualization methods has made it possible to extract biologically meaningful information from increasingly complex image datasets. High-throughput cellular phenotyping, digital atlas construction, dynamic cell analysis, neuronal tracing, and integration with genomic information represent important applications of the field.

Future development of bioimage informatics will depend on improved computational algorithms, standardized data formats, scalable databases, reproducible analysis workflows, machine-learning approaches, and integration of imaging data with other biological information. These developments are expected to further strengthen the role of quantitative image analysis in modern biological research.

REFERENCES

1. Pawley JB, editor. Handbook of biological confocal microscopy. 3rd ed. New York: Springer; 2006.
2. Bozzola JJ, Russell LD. Electron microscopy: principles and techniques for biologists. 2nd ed. Sudbury: Jones and Bartlett; 1999.
3. Betzig E, Patterson GH, Sougrat R, Lindwasser OW, Olenych S, Bonifacino JS, et al. Imaging intracellular fluorescent proteins at nanometer resolution. *Science*. 2006;313(5793):1642-5.
4. Rust MJ, Bates M, Zhuang X. Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat Methods*. 2006;3(10):793-5.
5. Hell SW. Toward fluorescence nanoscopy. *Nat Biotechnol*. 2003;21(11):1347-55.
6. Heim R, Cubitt AB, Tsien RY. Improved green fluorescence. *Nature*. 1995;373(6516):663-4.

7. Shimomura O, Johnson FH, Saiga Y. Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. *J Cell Comp Physiol.* 1962;59:223-39.
8. Ando R, Mizuno H, Miyawaki A. Regulated fast nucleocytoplasmic shuttling observed by reversible protein highlighting. *Science.* 2004;306(5700):1370-3.
9. Livet J, Weissman TA, Kang H, Draft RW, Lu J, Bennis RA, et al. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. *Nature.* 2007;450(7166):56-62.
10. Lein ES, Hawrylycz MJ, Ao N, Ayres M, Bensinger A, Bernard A, et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature.* 2007;445(7124):168-76.
11. Martone ME, Gupta A, Ellisman MH. EMDL: an environment for managing electron microscopy image data. *J Struct Biol.* 2002;138(1-2):145-55.
12. Echeverri CJ, Perrimon N. High-throughput RNAi screening in cultured cells: a user's guide. *Nat Rev Genet.* 2006;7(5):373-84.
13. Moffat J, Grueneberg DA, Yang X, Kim SY, Kloepper AM, Hinkle G, et al. A lentiviral RNAi library for human and mouse genes applied to an arrayed viral high-content screen. *Cell.* 2006;124(6):1283-98.
14. Sepp KJ, Baryshnikova A, Anholt RRH. Neural circuits: modeling and analysis of complex biological systems. *Methods Mol Biol.* 2008;467:285-96.
15. Long F, Peng H, Kim S, Myers EW. Automatic segmentation of nuclei in 3D microscopy images of developing embryos. *Cold Spring Harb Protoc.* 2007;2007(4):pdb.prot4717.
16. Jefferis GSXE, Potter CJ, Chan AM, Marin EC, Rohlfsing T, Maurer CR Jr, et al. Comprehensive maps of *Drosophila* higher olfactory centers: spatially segregated fruit and pheromone representation. *Cell.* 2007;128(6):1187-203.
17. Long F, Peng H, Liu X, Kim SK, Myers E. A 3D digital atlas of *C. elegans* and its application to single-cell analyses. *Nat Methods.* 2009;6(9):667-72.
18. Megason SG, Fraser SE. Imaging in systems biology. *Cell.* 2007;130(5):784-7.
19. Peng H. Bioimage informatics: a new area of engineering biology. *Bioinformatics.* 2008;24(17):1827-36.

20. Roysam B. FARSIGHT: flexible approach to image processing and analysis of multidimensional microscopy data. In: Proceedings of the IEEE International Symposium on Biomedical Imaging; 2008.
21. Swedlow JR, Goldberg IG, Eliceiri KW. Bioimage informatics for experimental biology. *Annu Rev Biophys.* 2009;38:327-46.
22. Carpenter AE, Jones TR, Lamprecht MR, Clarke C, Kang IH, Friman O, et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* 2006;7(10):R100.
23. Lamprecht MR, Sabatini DM, Carpenter AE. CellProfiler: free, versatile software for automated biological image analysis. *Biotechniques.* 2007;42(1):71-5.
24. Hong H. Content-based image retrieval for biological images. *Bioinformatics.* 2006;22(10):1210-7.
25. Lin G, Chawla MK, Olson K, Guzowski JF, Barnes CA, Roysam B. Hierarchical, model-based merging of multiple fragments for improved 3-D segmentation and tracking of nuclei in time-lapse microscopy. *IEEE Trans Biomed Eng.* 2005;52(7):1273-80.
26. Murphy RF, Velliste M, Porreca G. Robust numerical features for description and classification of subcellular location patterns in fluorescence microscope images. *Cytometry.* 2003;55A(1):61-72.
27. Bakal C, Aach J, Church G, Perrimon N. Quantitative morphological signatures define local signaling networks regulating cell morphology. *Science.* 2007;316(5832):1753-6.
28. Zhao T, Murphy RF. Automated learning of generative models for subcellular location: building blocks for systems biology. *Cytometry A.* 2007;71A(12):1074-88.
29. Luengo Hendriks CL, Keränen SVE, Li C, Bengtsson E. 3D analysis of gene expression patterns in *Drosophila* embryos. *BMC Bioinformatics.* 2006;7:175.
30. Peng H, Myers G. Automatic 3D neuron tracing using all-path pruning. *Bioinformatics.* 2004;20 Suppl 1:i134-42.
31. Liu T, Li G, Nie J, Tarokh A, Zhou X, Guo L, et al. ZFIQ: a software package for zebrafish biology. *Bioinformatics.* 2008;24(3):438-9.

32. Dorr AE, Lerch JP, Spring S, Kabani N, Henkelman RM. High resolution three-dimensional brain atlas using an average magnetic resonance image of 40 adult C57Bl/6J mice. *Neuroimage*. 2008;42(1):60-9.
33. Altinok A, Kirschner MW, Woldringh CL. Modeling bacterial growth and microtubule dynamics. *Biophys J*. 2006;91(6):2185-96.
34. Smal I, Draegestein K, Galjart N, Niessen W, Meijering E. Particle filtering for multiple object tracking in dynamic fluorescence microscopy images: application to microtubule growth analysis. *IEEE Trans Med Imaging*. 2008;27(6):789-804.
35. Danuser G, Tran PT, Salmon ED. Tracking cell divisions in time-lapse fluorescence microscopy. *J Microsc*. 2000;198(Pt 3):173-83.
36. Meijering E, Smal I, Danuser G. Tracking in molecular bioimaging. *IEEE Signal Process Mag*. 2008;25(1):46-53.
37. Jaqaman K, King EM, Amaro AC, Winter JR, Dorn JF, Elliott HL, et al. Kinetochore alignment within the metaphase plate is regulated by centromere stiffness and microtubule dynamics. *Curr Biol*. 2007;17(7):602-7.
38. Bao Z, Murray JI, Boyle T, Ooi SL, Sandel MJ, Waterston RH. Automated cell lineage tracing in *Caenorhabditis elegans*. *Proc Natl Acad Sci U S A*. 2006;103(8):2707-12.
39. Peng H, Long F, Myers G. Automatic 3D neuron tracing using all-path pruning. *Bioinformatics*. 2011;27(13):i239-47.
40. Dorn JF, Danuser G, Yang G. Computational analysis of cell motility. *Methods Cell Biol*. 2008;85:147-76.
41. Peng H, Myers G. Image-based spatial analysis of gene expression patterns in *Drosophila* embryos. *Bioinformatics*. 2004;20 Suppl 1:i13-21.
42. Zhou X, Peng H. A quantitative framework for the analysis of biological images. In: *Proceedings of the IEEE International Symposium on Biomedical Imaging*; 2007.
43. Paninski L, Fellows MR, Hatsopoulos NG, Donoghue JP. Spatiotemporal covariance in motor cortical neurons. *J Neurophysiol*. 2004;91(2):515-32.
44. Gurunathan R, Ghazaryan V, Sahu S. Invariant moment features for biological image analysis. *Proc SPIE*. 2004;5370:1-8.
45. Orlov N, Shamir L, Macura T, Johnston J, Eckley DM, Goldberg IG. WND-CHARM: multi-purpose image classification using compound image transforms. *Pattern Recognit Lett*. 2008;29(11):1684-93.

46. Ding C, Peng H. Minimum redundancy feature selection from microarray gene expression data. *J Bioinform Comput Biol*. 2005;3(2):185-205.
47. Cong J, Parvin B. Robust segmentation of nuclei in fluorescence microscopy images. *Proc IEEE Int Conf Image Process*. 1999;1:669-73.
48. Lin G, Adiga U, Olson K, Guzowski JF, Barnes CA, Roysam B. A hybrid 3D watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks. *Cytometry*. 2003;56A(1):23-35.
49. Lin G, Chawla MK, Olson K, Barnes CA, Roysam B. A multi-model approach to simultaneous segmentation and classification of cell nuclei in 3D microscopy images. *Cytometry A*. 2005;67(1):23-32.
50. Zitová B, Flusser J. Image registration methods: a survey. *Image Vis Comput*. 2003;21(11):977-1000.
51. Abramoff MD, Magalhães PJ, Ram SJ. Image processing with ImageJ. *Biophotonics Int*. 2004;11(7):36-42.
52. Linkert M, Rueden C, Allan C, Burel JM, Moore W, Patterson A, et al. Metadata matters: access to image data in the life sciences. *J Cell Biol*. 2010;189(5):777-82.
53. Ibáñez L, Schroeder W, Ng L, Cates J. The ITK software guide. Clifton Park: Kitware; 2005.
54. Gordon A, Colman-Lerner A, Chin TE, Benjamin KR, Yu RC, Brent R. Single-cell quantification of molecules and rates using open-source microscope-based cytometry. *Nat Methods*. 2007;4(2):175-81.
55. Peng H, Ruan Z, Long F, Simpson JH, Myers EW. V3D enables real-time 3D visualization and quantitative analysis of large-scale biological image data sets. *Nat Biotechnol*. 2010;28(4):348-53.
56. Swedlow JR, Goldberg IG, Eliceiri KW. Bioimage informatics for experimental biology. *Annu Rev Biophys*. 2009;38:327-46.
57. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J Comput Chem*. 2004;25(13):1605-12.