Editorial

Image Processing and Analysis for Biotechnology and Bioprocess Engineering

Sang-Kyu Jung

Department of Biological and Chemical Engineering, Hongik University, Sejong 2639, Republic of Korea; skjung@hongik.ac.kr; Tel.: +82-44-860-2127

1. Introduction

The development of high-performance computing hardware and image processing algorithms has led to the widespread application of image analysis in various fields [1,2]. In particular, image analysis is frequently applied in the fields of biotechnology and bioprocess engineering, where various imaging devices are used for research and development [3]. Compared to manual methods performed by humans, automated image analysis allows for fast, accurate, and reliable quantitative analysis. The classical approach to image analysis involves the sequential application of multiple image processing algorithms at the pixel level. The color value of a specific part of the final image is then analyzed, or an object is detected, and morphology such as coordinates, area, shape, and number is analyzed [4,5]. One of the important applications of image analysis in the field of biotechnology is probably high-throughput research that requires analyzing large amounts of images and videos, which is difficult for humans to perform [6]. For example, while exposing zebrafish to a large number of different chemicals, videos are recorded and analyzed to quantify the animal's growth, movement, organ morphology, social behavior, or feeding habits. Interpreting these results can reveal the function of the compound molecule and the efficacy and toxicity of the chemical or drug [7–9].

Recent advancements in machine learning techniques have made it possible to better utilize existing image analysis results. Rather than just using the area or color values of objects as endpoints, these image analysis results can now be used to classify objects according to their characteristics and predict important phenomena. For example, when obtaining multiple image analysis endpoints through image analysis, interpretation can become challenging. In such cases, applying dimensionality reduction techniques such as PCA (Principal Component Analysis), LDA (Linear Discriminant Analysis), T-SNE (t-distributed Stochastic Neighbor Embedding), and UMAP (Uniform Manifold Approximation and Projection) can significantly reduce the number of variables to be analyzed. This makes classification or prediction more feasible and enables visualization on a two-dimensional chart for easier human comprehension [10,11]. Moreover, tools like artificial neural networks (ANN), decision trees, and SVM (Support Vector Machine) have been employed for classification using multiple image analysis parameters [12–14].

Above all, the development of various deep learning architectures and training techniques suitable for image processing, along with the emergence of high-performance GPUs (Graphics Processing Units) capable of handling them, has led to the application of more sophisticated image processing technologies. Convolutional Neural Networks (CNN) and Transformers have shown excellent performance in object detection and classification. For instance, using CNN, it was revealed that female flies (*Drosophila*) were more attracted to sucrose during egg-laying [15], and CNN was employed for the detection, segmentation, and analysis of plants or plant stress [16]. This analysis could not be carried out with pixel-based classical image processing.

Meanwhile, ordinary biologists familiar with wet lab experiments often lack computer knowledge for conducting image analysis. Therefore, there is a need for the development

of user-friendly image analysis software. Numerous GUI (Graphical User Interface) applications utilizing image processing tools such as ImageJ [17], MATLAB (The MathWorks, Inc., Natick, MA, USA), openCV [18], PlantCV [19], and scikit-image [20] have been developed to cater to their needs. As documented in review papers, various software has been developed and used to analyze the locomotion of animals like zebrafish (*Danio rerio*) and *C. elegans*, as well as mice [21,22]. In addition, there is a variety of software that can detect plants, classify phenotypes [23,24], and classify cell types [25].

In this Special Issue, we introduce recent research articles using image analysis in biotechnology. It provides an overview not only of the applications of image analysis but also covers imaging devices, image processing algorithms, software development, and problem-solving processes.

2. An Overview of Published Articles

The article by Boruczkowski et al. (Contribution 1) analyzed the starch content of six potato cultivars through image analysis, presenting a case study where image analysis could be employed in food processing. For analysis, the authors sliced the potatoes to a thickness of 3 mm and stained them by soaking them in a potassium iodide (KI) solution. Higher starch content resulted in dark staining with iodine, while lower starch content exhibited a lighter color. Instead of using a digital camera, the authors utilized a scanner equipped with LED lights on top. The obtained images were then converted into brightness-corrected gray images. Although the average gray value of the potato cross-section could be used for starch content analysis, the authors uniquely positioned four cross-section lines at $45°$ intervals on the potato. The brightness distribution of pixels along these lines was then analyzed. In addition to the mean value, the median, range, skewness, and coefficient of variation (CV) of the gray pixels located on these measurement lines were computed. For that purpose, ImageJ software was employed for this image processing. The analysis revealed a good negative linear correlation between the mean gray value of cultivars and their actual starch content. Furthermore, while the actual starch content of the six cultivars was somewhat similar, making cultivar differentiation challenging, the image analysis statistical values mentioned above allowed for better discrimination. This image analysis method also proved useful in distinguishing potatoes stored at different temperatures (4 °C vs. −15 °C), demonstrating its potential as a practical on-site technology in potato-processing plants.

The study by Cha et al. (Contribution 2) utilized image analysis techniques to elucidate the correlation between genetic variants (SNPs: Single-Nucleotide Polymorphisms) and facial skin phenotypes. In this study, five facial skin phenotypes, such as wrinkles, pigmentation, moisture content, oil content, and sensitivity, were analyzed. Among these, wrinkles and pigmentation require analysis using special imaging equipment. First, the wrinkle phenotype was measured by calculating the average roughness (Ra) and maximum depth (Rmax) of the uneven skin surface. For this purpose, optical three-dimensional (3D) imaging was performed using the PRIMOS (Phaseshift Rapid In-vivo Measurement Of Skin) CR device (Canfield Scientific, Parsippany, NJ, USA). This equipment projects fringes made of parallel lines onto the skin surface and then photographs them with a digital camera pointed at an angle. When photographed at an angle, the fringe lines are observed to be curved depending on the depth of the surface. By interpreting this, a 3D-depth image can be obtained [26,27]. The pigmentation phenotype was analyzed for skin brightness and melanin content. For melanin measurement, the Mexameter MX 18 (Courage and Khazaka Electronic GmbH, Köln, Germany) was used. This equipment uses spectrophotometer technology to measure light in three wavelength bands (green at 568 nm, red at 660 nm, and infrared at 870 nm). Melanin content was quantified by measuring reflection and absorption in the skin, and skin brightness was assessed using a portable spectrophotometer, the CM-700d (Konica Minolta Inc., Tokyo, Japan), capable of full spectrum scanning from 400 nm to 700 nm. The L* value in the CIE L*a*b* color space was measured to serve as an

indicator of skin brightness. Finally, the measured skin phenotype values were divided into tertile groups and successfully used to discover their relationship with genetic variation.

The article by Hwang and Shin (Contribution 3) studied the fluorescence decay characteristics of photosynthetic phycobiliprotein complexes extracted from *Spirulina maxima*. Phycobiliproteins, including C-phycocyanin (C-PC), have a variety of applications, including the food industry, cosmetics, medical products, sensors, and photo harvesting. Understanding their optical properties is crucial for the successful utilization of these substances. In nature, C-PC exists in complex formations with various substances, and the aggregation of these complexes can impact their optical properties. Therefore, the authors explored the optical properties of C-PC aggregates. While most previous studies focused on the absorbance spectrum, the authors aimed to interpret fluorescence decay using mathematical models. For this purpose, C-PC aggregate was separated from *S. maxima* by precipitation and ion exchange, and then was followed by gel chromatography. The gel electrophoresis image confirmed the molecular weight of C-PC aggregates to be between approximately 17 and 20 kDa. The authors employed the Varioskan LUX multimode microplate reader (Thermo Scientific, Waltham, MA, USA) to study fluorescence decay and continuously imaged the fluorescence intensity emitted during excitation at 609 nm for 100 s. According to the analysis results, the C-PC aggregates exhibited a non-linear decay in fluorescence intensity over time. To mathematically simulate this decay trend, the authors devised two differential equation models based on mass balance. These mathematical models successfully regressed the experimental results of fluorescence intensity decay over time for standard C-PC and three C-PC aggregates.

The article by Wüstefeld et al. (Contribution 4) presents the development of deep neural network (DNN)-based segmentation algorithms for the automatic annotation of blob objects in images. Modern machine learning-based object detection and image classification usually require the preparation of training data and manual work by humans to label (or annotate) objects present in the image. When there are a large number of images or objects to be annotated, performing this task manually can be time-consuming, prone to inconsistencies, and susceptible to human errors. Especially in images where there is little difference in brightness and contrast between objects and the background, applying classical image processing algorithms based on pixel-by-pixel processing, such as adaptive thresholding or watershed, can be challenging. The images considered in this paper were captured using plasmon-assisted microscopy, utilizing the surface plasmon resonance effect (SPR), to image viruses and virus-like particles. These images fall into a category where accurate segmentation is difficult using classical segmentation algorithms. The authors of this article developed a technology using DNN for the automatic segmentation of such images. An interesting aspect is that, during training, instead of relying on individual annotation information for valid objects present in the images, they depended on the count of objects within the images. To achieve this, they connected a counting head, specialized in object counting, to the existing DNN, making the counting head responsible for object counting during training. In other words, the counting head was integrated into the network, and training was conducted using counting loss to ensure that the number of objects in the segmented result matches the manually provided count. However, relying solely on counting losses for training does not guarantee that the segmented images will necessarily exhibit a valid blob-like object. To address this, two approaches were devised. The first method involves utilizing class activation mapping (CAM) along with counting losses. The second method, instead of using CAM, adds contrast loss and morphological loss to counting loss, defining a total loss for training and segmentation. The authors conducted segmentation training using these two methods, using synthetically generated SPR images along with real SPR images. Both methods exhibited high segmentation accuracy, as measured by precision, recall, and F1-score, demonstrating the effective functioning of these weakly supervised automatic segmentation techniques. This paper serves as an example of how deep learning can be effectively applied in situations where classical segmentation techniques struggle to handle complex image processing tasks.

The article authored by Jung (Contribution 5) focuses on the development of software for analyzing the movement of small animals in a multi-well plate. Small animals ranging from a few millimeters to centimeters in size are widely utilized for studying the effects of genes or the environment on animals, as well as testing the toxicity of chemical substances and the efficacy of potential drugs. Although the movement of animals is a critical phenotype in such research, analyzing the rapid movements of animals through human observation is not feasible. The AniWellTracker software, introduced in this paper, is optimized for analyzing the movement of small animals in multi-well plates, which is essential for high-throughput research. It comes with a GUI designed for ease of use by researchers in the field of bioengineering. To briefly explain the image processing, the image captured in time-lapse is converted to a binary image through adaptive thresholding, which can be applied even in the case of nonuniform brightness, and then blob detection is performed. If the area and bounding box size of the detected object are appropriate, it is determined to be a valid animal. It then constructs locomotion paths by connecting the centroids of valid animals in each well of time-lapse images. AniWellTracker not only computes speed, travel distance, and location angle from the well center but also visualizes these results through colorful histograms. Particularly noteworthy is its capability to represent the distribution of animal positions through contoured heatmaps. The analyzed data can be conveniently stored in CSV (Comma Separated Values) file format, facilitating processing with spreadsheet programs like Microsoft Excel. The software is developed with native code, allowing it to be executed as a standalone application without relying on commercial software such as MATLAB or LabView (National Instruments, Austin, TX, USA) or external libraries such as ImageJ or OpenCV. Moreover, its open-source code enables advanced users to modify and adapt it to their needs, making it not only versatile but also cost-effective. It is anticipated that AniWellTracker will prove to be a valuable tool for future biological researchers because of its user-friendly and freely accessible nature.

3. Conclusions

This compilation of articles is dedicated to image analysis applications in the field of biotechnology. In terms of imaging hardware, various imaging devices are increasingly being utilized, as evidenced by Boruczkowski et al.'s scanner imaging (Contribution 1), Cha et al.'s PRIMOS 3D imaging and spectrophotometer imaging (Contribution 2), Hwang and Shin's fluorescence imaging (Contribution 3), and Wüstefeld et al.'s SPR imaging (Contribution 4), in addition to conventional digital cameras. Computing color intensity and distribution from acquired images has enabled the prediction of potato starch content and the differentiation of various cultivars (Contribution 1). Analyzing the average roughness, maximum depth, brightness, and melanin content of human skin has revealed SNPs influencing skin phenotypes (Contribution 2). Moreover, the fluorescence decay characteristics of photosynthetic phycobiliprotein complexes could be quantified (Contribution 3). For SPR images where accurate segmentation is challenging using traditional pixel-to-pixel-based image processing, DNN-based image segmentation algorithms were developed by incorporating CAM and counting, contrast, and morphological loss, achieving high-precision segmentation (Contribution 4). No matter how advanced image processing algorithms are, their utility is limited, if not easily accessible, to researchers. Therefore, the development of user-friendly and affordable image analysis software is crucial. Jung developed image analysis software using multi-well plates for high-throughput study of the locomotion of small animals (Contribution 5). The software was developed based on a GUI and is not only easy to use but can also calculate various image analysis features and has built-in visualization functions.

With recent remarkable developments in machine learning technology, it is anticipated that the current level of image analysis technology in the field of biotechnology will further elevate. In addition, hybrid image analysis techniques that combine machine learning algorithms with classic image processing algorithms are expected to be applied more widely.

Conflicts of Interest: The author declares no conflicts of interest.

List of Contributions

1. Boruczkowski, T.; Boruczkowska, H.; Drożdż, W.; Raszewski, B. Application of digital image analysis for assessment of starch content and distribution in potatoes. *Appl. Sci.* **2022**, *12*, 12988. https://doi.org/10.3390/app122412988.
2. Cha, M.-Y.; Choi, J.-E.; Lee, D.-S.; Lee, S.-R.; Lee, S.-I.; Park, J.-H.; Shin, J.-H.; Suh, I.S.; Kim, B.H.; Hong, K.-W. Novel genetic associations for skin aging phenotypes and validation of previously reported skin GWAS results. *Appl. Sci.* **2022**, *12*, 11422. https://doi.org/10.3390/app122211422.
3. Hwang, J.; Shin, A.H. A mathematical modeling and statistical analysis of phycobiliprotein fluorescence decay under exposure to excitation light. *Appl. Sci.* **2022**, *12*, 7469. https://doi.org/10.3390/app12157469.
4. Wüstefeld, K.; Ebbinghaus, R.; Weichert, F. Learning to segment blob-like objects by image-level counting. *Appl. Sci.* **2023**, *13*, 12219. https://doi.org/10.3390/app132212219.
5. Jung, S.-K. AniWellTracker: Image analysis of small animal locomotion in multiwell plates. *Appl. Sci.* **2023**, *13*, 2274. https://doi.org/10.3390/app13042274.

References

1. Kasturi, R. *Image Analysis Applications*; CRC Press: Boca Raton, FL, USA, 1990; Volume 24.
2. Jähne, B. *Digital Image Processing Concepts, Algorithms, and Scientific Applications*; Springer: Berlin/Heidelberg, Germany, 1995. [CrossRef]
3. Jung, S.-K. A review of image analysis in biochemical engineering. *Biotechnol. Bioprocess Eng.* **2019**, *24*, 65–75. [CrossRef]
4. Sonka, M.; Hlavac, V.; Boyle, R. *Image Processing, Analysis and Machine Vision*; Springer: Berlin/Heidelberg, Germany, 2013.
5. Shariff, A.; Kangas, J.; Coelho, L.P.; Quinn, S.; Murphy, R.F. Automated image analysis for high-content screening and analysis. *J. Biomol. Screen.* **2010**, *15*, 726–734. [CrossRef]
6. Das Choudhury, S.; Samal, A.; Awada, T. Leveraging image analysis for high-throughput plant phenotyping. *Front. Plant. Sci.* **2019**, *10*, 508. [CrossRef] [PubMed]
7. Rihel, J.; Prober, D.A.; Arvanites, A.; Lam, K.; Zimmerman, S.; Jang, S.; Haggarty, S.J.; Kokel, D.; Rubin, L.L.; Peterson, R.T.; et al. Zebrafish behavioral profiling links drugs to biological targets and rest/wake regulation. *Science* **2010**, *327*, 348–351. [CrossRef] [PubMed]
8. Miyawaki, I. Application of zebrafish to safety evaluation in drug discovery. *J. Toxicol. Pathol.* **2020**, *33*, 197–210. [CrossRef] [PubMed]
9. Dash, S.N.; Patnaik, L. Flight for fish in drug discovery: A review of zebrafish-based screening of molecules. *Biol. Lett.* **2023**, *19*, 20220541. [CrossRef]
10. Valletta, J.J.; Torney, C.; Kings, M.; Thornton, A.; Madden, J. Applications of machine learning in animal behaviour studies. *Anim. Behav.* **2017**, *124*, 203–220. [CrossRef]
11. Kraus, O.Z.; Frey, B.J. Computer vision for high content screening. *Crit. Rev. Biochem. Mol. Biol.* **2016**, *51*, 102–109. [CrossRef]
12. Mohammed, E.A.; Mohamed, M.M.; Far, B.H.; Naugler, C. Peripheral blood smear image analysis: A comprehensive review. *J. Pathol. Inform.* **2014**, *5*, 9. [CrossRef]
13. Zhang, J.; Li, C.; Rahaman, M.M.; Yao, Y.; Ma, P.; Zhang, J.; Zhao, X.; Jiang, T.; Grzegorzek, M. A comprehensive review of image analysis methods for microorganism counting: From classical image processing to deep learning approaches. *Artif. Intell. Rev.* **2022**, *55*, 2875–2944. [CrossRef]
14. Yan, J.; Wang, X. Unsupervised and semi-supervised learning: The next frontier in machine learning for plant systems biology. *Plant J.* **2022**, *111*, 1527–1538. [CrossRef] [PubMed]
15. Stern, U.; He, R.; Yang, C.-H. Analyzing animal behavior via classifying each video frame using convolutional neural networks. *Sci. Rep.* **2015**, *5*, 14351. [CrossRef] [PubMed]
16. Jiang, Y.; Li, C. Convolutional neural networks for image-based high-throughput plant phenotyping: A review. *Plant Phenomics* **2020**, 4152816. [CrossRef]
17. Schneider, C.A.; Rasband, W.S.; Eliceiri, K.W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **2012**, *9*, 671. [CrossRef]
18. Bradski, G. The openCV library. *Dr. Dobb's J. Softw. Tools Prof. Program.* **2000**, *25*, 120–123.
19. Gehan, M.A.; Fahlgren, N.; Abbasi, A.; Berry, J.C.; Callen, S.T.; Chavez, L.; Doust, A.N.; Feldman, M.J.; Gilbert, K.B.; Hodge, J.G. PlantCV v2: Image analysis software for high-throughput plant phenotyping. *PeerJ* **2017**, *5*, e4088. [CrossRef]
20. Van der Walt, S.; Schönberger, J.L.; Nunez-Iglesias, J.; Boulogne, F.; Warner, J.D.; Yager, N.; Gouillart, E.; Yu, T. scikit-image: Image processing in Python. *PeerJ* **2014**, *2*, e453. [CrossRef] [PubMed]

21. Franco-Restrepo, J.E.; Forero, D.A.; Vargas, R.A. A review of freely available, open-source software for the automated analysis of the behavior of adult zebrafish. *Zebrafish* **2019**, *16*, 223–232. [CrossRef]

22. Panadeiro, V.; Rodriguez, A.; Henry, J.; Wlodkowic, D.; Andersson, M. A review of 28 free animal-tracking software applications: Current features and limitations. *Lab Anim.* **2021**, *50*, 246–254. [CrossRef]

23. Lobet, G.; Draye, X.; Périlleux, C. An online database for plant image analysis software tools. *Plant Methods* **2013**, *9*, 38. [CrossRef]

24. Rahman, H.; Ramanathan, V.; Jagadeeshselvam, N.; Ramasamy, S.; Rajendran, S.; Ramachandran, M.; Sudheer, P.D.; Chauhan, S.; Natesan, S.; Muthurajan, R. Phenomics: Technologies and applications in plant and agriculture. In *PlantOmics: The Omics of Plant Science*; Springer: New Delhi, India, 2015; pp. 385–411.

25. Smith, K.; Piccinini, F.; Balassa, T.; Koos, K.; Danka, T.; Azizpour, H.; Horvath, P. Phenotypic image analysis software tools for exploring and understanding big image data from cell-based assays. *Cell Syst.* **2018**, *6*, 636–653. [CrossRef]

26. Benderoth, C.; Hainich, R. Optical 3D in-vivo skin imaging for topographical quantitative assessment of cosmetic and medical treatments. In Proceedings of the International Conference on 3D Body Scanning Technologies, Lugano, Switzerland, 19–20 October 2010; pp. 42–51.

27. Gevaux, L. 3D-Hyperspectral Imaging and Optical Analysis of Skin for the Human Face. Doctoral Dissertation, Université de Lyon, Lyon, France, 2019.