

TECHNICAL REPORT

Application of Deep Neural Networks in the Manufacturing Process of Mesenchymal Stem Cells Therapeutics

Dat Ngo^{1,*}, Jeongmin Lee^{2,3,*}, Sun Jae Kwon², Jin Hun Park^{4,5}, Baek Hwan Cho^{5,6}, Jong Wook Chang^{2,3,7}

¹Department of Computer Engineering, Korea National University of Transportation, Chungju, Korea

²CDMO Technology Institute, ENCell Co., Ltd., Seoul, Korea

³Cell and Gene Therapy Institute, Samsung Medical Center, Seoul, Korea

⁴Department of Media and Communication, College of Future Convergence Division of Healthcare Sciences, CHA University, Seongnam, Korea

⁵Department of Biomedical Informatics, School of Medicine, CHA University, Seongnam, Korea

⁶Institute of Biomedical Informatics, School of Medicine, CHA University, Seongnam, Korea

⁷Department of Health Sciences and Technology, Samsung Advanced Institute for Health Sciences & Technology, Sungkyunkwan University, Seoul, Korea

Current image-based analysis methods for monitoring cell confluency and status depend on individual interpretations, which can lead to wide variations in the quality of cell therapeutics. To overcome these limitations, images of mesenchymal stem cells cultured adherently in various types of culture vessels were captured and analyzed using a deep neural network. Among the various deep learning methods, a classification and detection algorithm was selected to verify cell confluency and status. We confirmed that the image classification algorithm demonstrates significant accuracy for both single- and multistack images. Abnormal cells could be detected exclusively in single-stack images, as multistack culture was performed only when abnormal cells were absent in the single-stack culture. This study is the first to analyze cell images based on a deep learning method that directly impacts yield and quality, which are important product parameters in stem cell therapeutics.

Keywords: Mesenchymal stem cells, Deep learning, Digital image processing, Cell proliferation, Neural network models

Received: June 14, 2024, Revised: August 26, 2024, Accepted: September 2, 2024, Published online: September 26, 2024

Correspondence to **Jong Wook Chang**

Department of Health Sciences and Technology, Samsung Advanced Institute for Health Sciences & Technology, Sungkyunkwan University, 115 Irwon-ro, Gangnam-gu, Seoul 06355, Korea

E-mail: jongwook.chang@samsung.com

Co-Correspondence to **Baek Hwan Cho**

Department of Biomedical Informatics, School of Medicine, CHA University, 335 Pangyo-ro, Bundang-gu, Seongnam 13488, Korea

E-mail: baekhwan.cho@cha.ac.kr

Co-Correspondence to **Jin Hun Park**

Department of Media and Communication, College of Future Convergence Division of Healthcare Sciences, CHA University, 335 Pangyo-ro, Bundang-gu, Seongnam 13488, Korea

E-mail: parkjinhun@cha.ac.kr

*These authors contributed equally to this work.

© This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Copyright © 2024 by the Korean Society for Stem Cell Research

Introduction

Since their discovery by Friedenstein et al. (1) in the 1960s and the 1970s, mesenchymal stem cells (MSCs) have demonstrated therapeutic potential against various diseases (1, 2). Beyond their regenerative potential, the main therapeutic mechanism of MSCs is mediated through paracrine actions, such as immunomodulatory, trophic, and protective effects in various diseases (3, 4). The criteria for defining cells as MSCs are characterized by the International Society for Cell & Gene Therapy (5). MSCs have been found in a variety of adult and neonatal tissue sources (6), but an essential requirement is that they need to be cultured under plastic-adherent conditions (7).

The yield of adherent cells is assessed by monitoring their confluence relative to the surface of the culture vessel. Cell confluency is crucial in the manufacture of cell therapy products because it determines the subculture and harvest stages. Microscopy is commonly used to measure cell confluency, as it allows for nondestructive, label-free analysis, which is critical for maintaining product quality. In addition, visual measurements enable the detection of cells with an abnormal status, as indicated by an enlarged and flattened shape.

Nevertheless, the reliance on manual assessment of microscope images to evaluate MSCs growth and determine the optimal harvest time introduces subjectivity and significant time investment. In the field of therapeutic production, these variations have a significant impact (8). Different interpretations of cell images by individual technicians can lead to large variations in key process steps, resulting in quality deviations such as decreased therapeutic efficacy and production yield. Therefore, developing an objective and automated protocol for monitoring MSCs cultures is crucial, as this area remains underexplored.

Efforts to automate the examination of cell proliferation have encountered specific challenges, especially when dealing with MSCs at the production scale. To maximize their productivity, MSCs are cultured in multilayer flasks (9). The multilevel culture vessels enable savings in cost, space, time, and labor, but also present limitations in the resolution of the images owing to the different refractions of light under overlapped layers. Furthermore, distinguishing cells from the background becomes increasingly subtle as cell confluence increases. This inherent complexity underscores the difficulty of accurately assessing cell growth based on microscope images and highlights the need for monitoring algorithms that meet rigorous requirements in terms of accuracy and generalizability.

Despite the scarcity of studies on automated monitoring

of MSCs cultures, related research on cell segmentation (10) and cell confluence estimation (11) offers valuable insights. These studies have employed various computer vision techniques with different levels of algorithmic complexity. Heuristic methods, including image thresholding, edge detection, and clustering, lack generalizability because they rely on predefined rules and assumptions. In contrast, deep convolutional neural networks (CNNs) offer a data-driven approach that captures hidden regularities within data. CNNs have achieved remarkable performance in medical applications such as skin cancer classification (12). Unlike heuristic methods, CNNs learn intricate patterns and features directly from data, making them particularly promising for MSCs culture monitoring.

In this study, we present a novel approach based on CNNs for classifying cell confluence and detecting abnormal cells in microscope images. Our objective was to accurately categorize the entire image into different confluence levels while identifying and localizing the presence of abnormal cells.

Materials and Methods

Isolation and cultivation of human Wharton's jelly-derived MSCs

This study was approved by the Institutional Review Board of the Samsung Medical Center (IRB# SMC-2016-07-102-037, renewal date August 19th, 2022). Umbilical cords were collected with informed consent from pregnant mothers. Wharton's jelly-derived mesenchymal stem cells (WJ-MSCs) were isolated from umbilical cord tissues, as previously described (13). WJ-MSCs were cultured in single- and multilayer flasks according to the standard operating procedures of the Good Manufacturing Practice facility at ENCell Co., Ltd. Images of the cells were taken daily at identical locations at the four corners and center of the culture vessels. Cell confluence was categorized into four levels: class 1, seeding point (0%~40%); class 2, media change decision point (40%~60%); class 3, subculture decision point (60%~80%); and class 4, overgrown decision point (80%~100%; Fig. 1).

Dataset exploration

Our dataset comprised 720 single-stack and 413 multistack images, each of which was associated with one of four different confluence levels: 1, 2, 3, and 4. Furthermore, a subset of the single-stack images (165 of 720) contained bounding box annotations indicating the presence of abnormal cells.

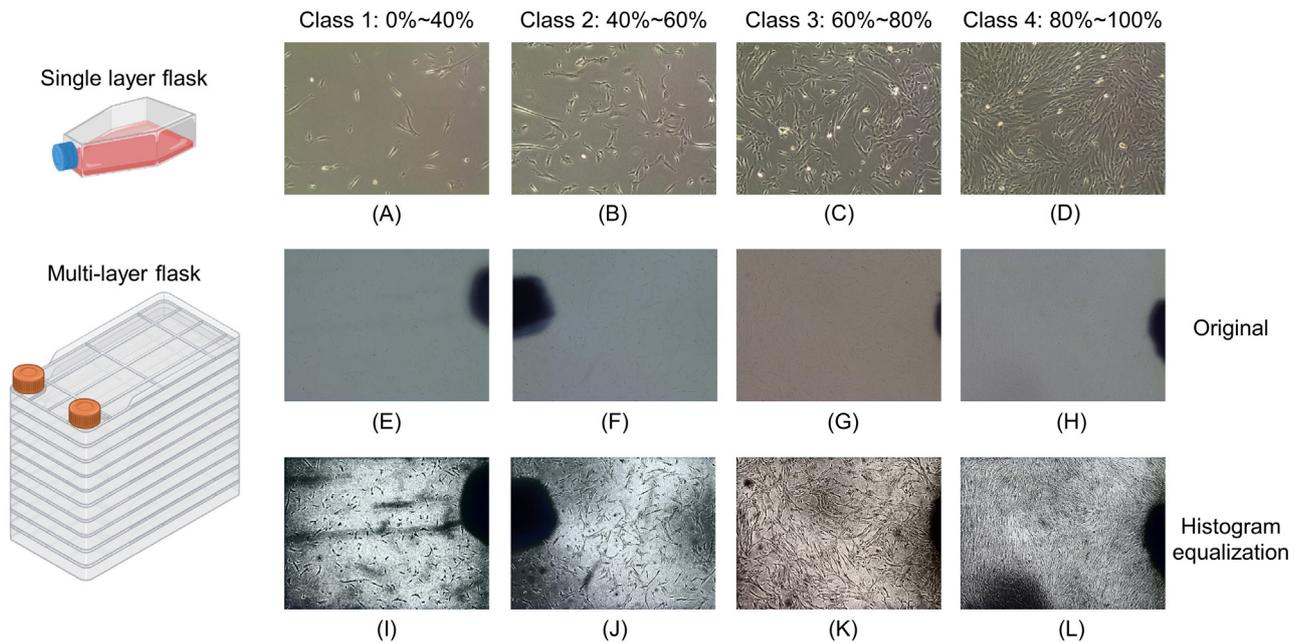


Fig. 1. Microscope images of Wharton's jelly-derived mesenchymal stem cells in single-stack and multistack configurations at various sub-classes: (A-D) show classification images from single layer flasks, (E-H) depict original classification images from multi-layer flasks, and (I-L) display histogram-equalized classification images from multi-layer flasks. The images of the flasks were created using BioRender.com

Augmentation methods

Given the limited number of images in the dataset, we applied seven augmentation methods to mitigate overfitting: random rotation, horizontal flip, vertical flip, color jitter, Gaussian blur, sharpness enhancement, and contrast enhancement. The first three methods improved the model's robustness against geometric rotations, while the last four address variations in the microscopic imaging conditions (such as lighting and contrast).

The rotation operator adjusted each pixel location in an image by rotating it through a specified angle about the image center. Horizontal and vertical flips are self-explanatory. Color jitter randomly adjusted the image brightness and hue, increasing the model's robustness by introducing variations in the color properties. Gaussian blurring involved convolving an image with a low-pass filter whose coefficients were drawn from a normal distribution, thereby smoothing the image, and reducing noise.

Sharpness enhancement was implemented by adding a scaled edge map to a smoothed version of the input image. An edge map was obtained by subtracting the smoothed image from the original image. This technique enhanced the clarity and definition of image edges, promoting clearer object boundaries. Contrast enhancement was achieved through min-max image intensity stretching, thereby expanding the dynamic range of the pixel intensities, and

increasing the contrast. In this study, we applied sharpness and contrast enhancements with equal probability.

Deep learning model

Fig. 2 presents an overview of the proposed model, which comprises the following modules: feature extraction, classification, and abnormal cell detection. We assessed the performance of six state-of-the-art CNN architectures—AlexNet, InceptionV3, ResNet50 (14), ShuffleNetV2 (15), MobileNetV3 (16), and Vision Transformer (17)—to identify the optimal feature extractor. We subsequently adapted RetinaNet (18), a widely used object detection framework, to perform image classification and abnormal cell detection.

The source code and dataset used in this study are publicly available at <https://github.com/MIHlabCHA/Mesenchymal-Stem-Cells-Classification-and-Detection>

Feature extraction: In classification tasks, obtaining a highly abstract task-specific feature map is crucial. Although heuristic methods are effective at capturing primitive features relevant to a specific problem, their limited generalizability stems from their lack of feature abstraction capability. By contrast, CNNs offer a powerful framework for extracting and abstracting features from images. CNNs consist of sequentially arranged convolutional layers, where each layer typically performs convolution, nonlinear mapping, and pooling. Convolution uses learnable filters to de-

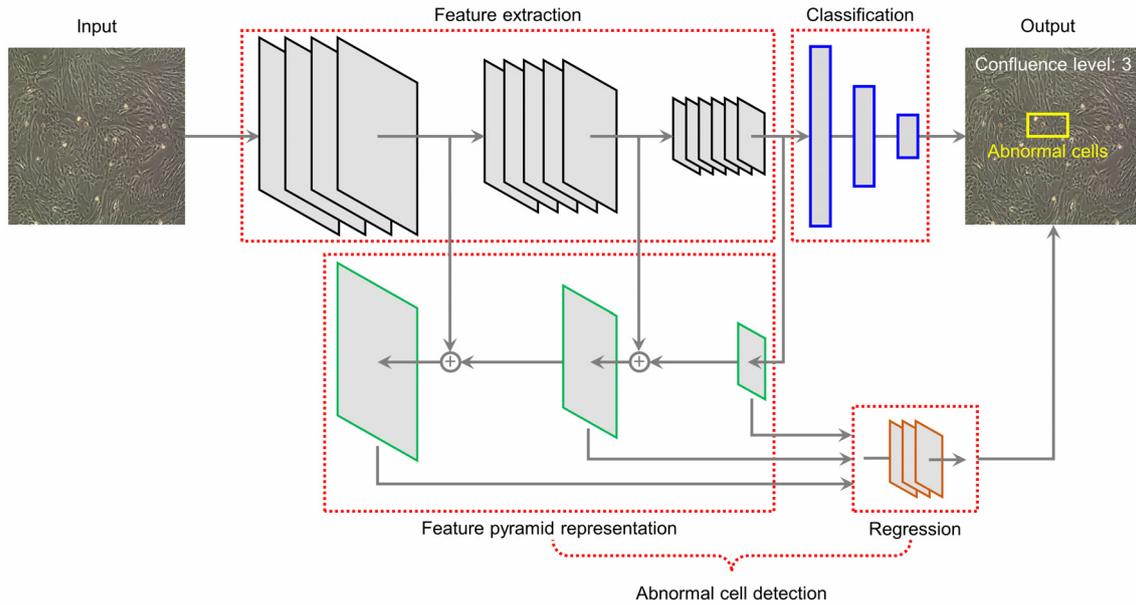


Fig. 2. Overview of the proposed deep learning model for mesenchymal stem cells image classification and abnormal cell detection. The red-dotted boxes correspond to the explanation in the deep learning model subsection.

detect various visual patterns in an image, producing feature maps that highlight the presence of the detected patterns. Nonlinear mapping applies an element-wise function to feature maps, enabling CNNs to discover complex relationships. Pooling summarizes local information in feature maps, reducing spatial dimensions while preserving essential patterns, which helps CNNs remain invariant to small spatial translations.

By stacking multiple convolutional layers, CNNs progressively learn abstract and higher-level features. The lower layers capture primitive features, such as edges and textures, while the higher layers capture more complex patterns and semantic information. In this study, we used adaptive average pooling at the final convolutional layer to aggregate and flatten the feature maps into a feature vector of length 2048, thereby facilitating the subsequent classification process.

Classification: Given the feature vector, we employed two fully connected layers (FCLs), followed by a softmax layer for classification. Each FCL, which is characterized by a weight matrix and a bias vector, mapped the input feature vector to an output of the desired length. In this study, the two FCLs reduced the feature vector length from 2,048 to 1,024, and then to four, corresponding to the number of classes. The softmax layer normalized the output vector into a probability distribution, with each element representing the probability of the corresponding class.

To train the model, we used the cross-entropy loss func-

tion, which measured the dissimilarity between the predicted and true probability distributions and served as an effective objective for model training.

Abnormal cell detection: To detect abnormal cells, we constructed a feature pyramid to handle different object scales, thereby enhancing the scale invariance of our model. The feature pyramid was built top-down, starting with the final feature map from the feature extractor. For each convolutional layer, the corresponding final feature map was upsampled to match the size of the preceding layer's feature map. The two maps were combined to form a pyramidal feature map.

These pyramid feature maps were then input into independent regression networks for abnormal cell detection. Each map could be viewed as a set of grids, where each grid covered a specific region in the input image and contained predefined anchor boxes of various sizes and aspect ratios. The network identified the presence and shape of abnormal cells within each anchor box.

To ensure accurate bounding box predictions and minimize the impact of outliers, we employed a smooth L1 function for model training. This function enabled our model to learn precise bounding-box predictions while being less sensitive to extreme values.

Results

Model evaluation using cross-validation

We performed five-fold cross-validation to assess the efficacy of our proposed deep learning solution. More specifically, we split the dataset into five equally sized groups, reserving one group for validation while using the remaining four for model training. This process was repeated five times, with each of the five groups serving as the validation dataset exactly once. During model training, we utilized the AMSGrad variant of the Adam optimization algorithm, incorporating L2 regularization and a learning rate uniformly sampled from the range $[10^{-2}, 10^{-8}]$. We set the total number of epochs to 1,000 and allowed the training to terminate if there was no improvement after 100 epochs.

We trained and validated our model on a deep learning server equipped with two 8-core Intel(R) Xeon(R) Gold 5315Y CPUs (each supporting hyper-threading with two threads), 256 GB RAM, and four NVIDIA RTX A2000 12 GB GPUs.

Performance metrics and image classification

To evaluate WJ-MSCs image classification, we used the area under the receiver operating characteristic curve (ROCAUC), accuracy, and the area under the precision-recall curve (PRAUC) as assessment metrics. Because our task involved multiclass classification, we calculated the ROCAUC and PRAUC using the micro-average one-versus-all approach. We investigated six different backbone CNNs, namely AlexNet, InceptionV3, ResNet50, ShuffleNetV2, MobileNetV3, and Vision Transformer, to determine the optimal feature extractor. The experimental results are summarized in Table 1 and the best results are denoted by the superscript ‘a’.

Table 1 shows that all six CNNs are highly appropriate

for single-stack WJ-MSCs image classification. InceptionV3 achieved the highest mean ROCAUC score of 0.958 ± 0.017 , the highest mean PRAUC score of 0.923 ± 0.030 , and the highest accuracy, with a mean score of 0.928 ± 0.029 . In classifying images based on cell confluence levels, images obtained from single-flask cultivation were visually distinguishable, albeit with some confusion between classes 2 and 3 (Fig. 1).

Challenges in multistack image classification

Nevertheless, this task became more challenging when dealing with images from multi-flask cultivation, which exhibited greater complexity. Consequently, certain CNNs, including AlexNet, ShuffleNetV2, MobileNetV3, and Vision Transformer, struggled to extract informative high-level features, as evidenced by their relatively low scores. In contrast, InceptionV3 and ResNet50, with architectural designs that facilitate the stacking of multiple convolutional layers, yielded promising results. Notably, ResNet50 emerged as the most suitable architecture for feature extraction, yielding the highest scores for ROCAUC, accuracy, and PRAUC. The inclusion of auxiliary classifiers in InceptionV3 helps mitigate the problem of gradient vanishing in very deep neural networks, whereas ResNet50

Table 2. Mean hit rate and false discovery rate scores for abnormal cell detection

	Hit rate	False discovery rate
Fold 1	0.879	0.000
Fold 2	0.968	0.000
Fold 3	1.000	0.000
Fold 4	0.923	0.000
Fold 5	1.000	0.000
Average	0.954 ± 0.047	0.000 ± 0.000

Values are presented as mean \pm SD.

Table 1. Mean ROCAUC, accuracy, and PRAUC scores for mesenchymal stem cells image classification

Backbone CNN	Single-stack			Multistack		
	ROCAUC	Accuracy	PRAUC	ROCAUC	Accuracy	PRAUC
AlexNet	0.949 ± 0.008	0.863 ± 0.035	0.879 ± 0.026	0.751 ± 0.051	0.525 ± 0.055	0.494 ± 0.080
InceptionV3	0.958 ± 0.017^a	0.928 ± 0.029^a	0.923 ± 0.030^a	0.937 ± 0.033	0.850 ± 0.074	0.839 ± 0.082
ResNet50	0.941 ± 0.009	0.863 ± 0.038	0.857 ± 0.032	0.958 ± 0.015^a	0.875 ± 0.045^a	0.897 ± 0.034^a
ShuffleNetV2	0.915 ± 0.023	0.766 ± 0.041	0.816 ± 0.052	0.892 ± 0.033	0.765 ± 0.068	0.751 ± 0.067
MobileNetV3	0.940 ± 0.014	0.816 ± 0.108	0.861 ± 0.041	0.894 ± 0.034	0.750 ± 0.098	0.727 ± 0.084
Vision Transformer	0.948 ± 0.013	0.809 ± 0.065	0.867 ± 0.036	0.616 ± 0.023	0.380 ± 0.053	0.341 ± 0.045

Values are presented as mean \pm SD.

ROCAUC: receiver operating characteristic curve, PRAUC: precision-recall curve, CNN: convolutional neural network.

^aBest results.

addresses this issue using skip connections. These architectural choices contributed significantly to the improved performance of InceptionV3 and ResNet50 in our experiments.

Abnormal cell detection efficacy

Based on the experimental results of the WJ-MSCs image classification, it is evident that both InceptionV3 and ResNet50 excel as feature extractors regardless of whether the input images are single-stack or multistack images. This observation, coupled with the favorable characteristics of RetinaNet in terms of implementation convenience, led us to select ResNet50 as the backbone feature extractor for abnormal cell detection.

To evaluate the performance of abnormal cell detection, we utilized two metrics: hit and false discovery rates. We also performed five-fold cross-validation and the results are presented in Table 2. Notably, our experiments for detecting abnormal cells were conducted exclusively with single-stack images due to the absence of bounding box annotations for multistack images, as indicated in Table 1. The analysis demonstrates the outstanding performance of the proposed model, particularly in achieving zero false alarms across all five folds. Notably, the third and fifth folds displayed ideal performance, while the remaining folds experienced a few instances of misclassification, where the model incorrectly indicated the absence of abnormal cells. Overall, the experimental results confirm the suitability of our model for accurate abnormal cell detection.

Detection analysis of the multi-layer flask image set was not possible as these images were taken at later passages after abnormal conditions had already been verified.

Discussion

MSCs are promising therapeutic candidates for various diseases, as demonstrated by numerous clinical trials and approvals worldwide (8). Their intrinsic features allow them to naturally migrate, localize, and respond to disease environments, providing significant benefits with minimal immune reactions, which enables the use of allogenic products (6, 7). However, challenges remain, as MSCs from various sources present diversity, complicating the manufacturing process. From the perspective of the manufacturing process alone, adherent culture remains the key process for MSCs production, and methods such as total cell count, cell image-based analysis, and metabolic analysis are being employed to enable a more efficient scale-up (19). Analyzing the total number of cells requires detachment, which risks cell destruction. Analyzing the metabolic changes can confirm overall cell growth patterns but fails

to independently detect abnormal cell status. Assessing MSC confluency is the simplest method, but is prone to subjective interpretation.

Despite its recognized significance, cell confluence has received limited attention in the literature. Related studies focused on the classification and segmentation of cell phenotypes. Traditional approaches rely on heuristic methods for image preprocessing, followed by handcrafted feature extraction and classification using models, such as SVM and kNN (10). However, these methods often suffer from limited generalizability because of manual parameter tuning. Recent efforts by Bai et al. (11) introduced deep-learning-based approaches to address these challenges. However, these methods often use standard architectures that may lack the flexibility to capture complex relationships in cellular imagery.

Our study explored various deep learning models, such as AlexNet, InceptionV3, ResNet50, ShuffleNetV2, MobileNetV3, and Vision Transformer. We leveraged RetinaNet for rapid and accurate detection of abnormal cells, which is a critical aspect of the manufacturing process. We validated our model across different image types, including single-stack and multistack images, demonstrating its robustness and potential in WJ-MSC applications.

While InceptionV3 demonstrated outstanding performance in classifying cell confluence in single-stack images, the other CNN architectures also yielded acceptable scores. This can be attributed to the distinct features of single-stack images, which allow CNNs to capture high-level indicators of cell confluence.

The performance of CNN architectures varied substantially when analyzing multistack images. While the best ROCAUC using ResNet50 (0.958 ± 0.015) was comparable to that for single-stack images, accuracy (0.875 ± 0.045) and PRAUC (0.897 ± 0.034) were notably lower. This diminished performance is likely due to cell clumping, represented as black spots in Fig. 1, which distracts the CNN models from discerning meaningful patterns within the training data. Further studies should include comprehensive experiments to clarify the impact of cell clumping on model performance.

Furthermore, we present the experimental findings on the detection of abnormal cells using single-stack images. This task could not be performed in the multistack scenario because of the absence of bounding box annotations. The detection results were impressive, with a mean false discovery rate of zero and a mean hit rate of 0.954 ± 0.047 . This performance is attributed to anchor boxes of varying sizes and aspect ratios, which provided prior information about expected object dimensions and geometries. The

network learned to classify abnormal cells and predicted their bounding boxes relative to these anchors, thereby improving localization accuracy.

However, two major limitations of this study must be acknowledged. Firstly, the absence of an external test set collected from different institutes under diverse cultivation settings hindered the assessment of the generalizability of the proposed model. Secondly, our dataset included only a single bounding box annotation per image, whereas abnormal cells could occur at multiple locations within an image. Future efforts should focus on acquiring additional data to create an external test set and address cases where abnormal cells are dispersed across the image.

Furthermore, it is imperative to investigate whether flattened or enlarged abnormal cells in multilayer cultivation impact classification. One approach involves training and validating two separate models on variants of the same dataset: one with cell clumps present and the other with cell clumps masked out. A comparative analysis of these models is likely to yield valuable insights into the impact of cell clumps on the model's ability to discern cell confluence levels, which is a major goal of our future research.

This study aimed to reduce subjectivity and increase practicality by using a simple index. We demonstrated that the noninvasive classification of multilayer flask images can be effectively applied to the GMP manufacturing process, potentially enhancing the efficiency of WJ-MSCs therapeutics for various diseases.

ORCID

Dat Ngo, <https://orcid.org/0000-0002-3896-648X>
 Jeongmin Lee, <https://orcid.org/0000-0003-3983-3950>
 Sun Jae Kwon, <https://orcid.org/0009-0008-7966-9706>
 Jin Hun Park, <https://orcid.org/0009-0009-7633-4545>
 Baek Hwan Cho, <https://orcid.org/0000-0001-7722-5660>
 Jong Wook Chang, <https://orcid.org/0000-0001-9335-5510>

Funding

This research was partly supported by grants from the Korea Health Technology R&D Project through the Korea Health Industry Development Institute, funded by the Ministry of Health & Welfare, Republic of Korea (grant numbers: HR14C0008 and HR22C1363), and a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. NRF-2023R1A2C2003577).

Potential Conflict of Interest

There is no potential conflict of interest to declare.

Authors' Contribution

Conceptualization: DN, JL, BHC, JWC. Data curation: DN, JL. Formal analysis: DN, BHC. Funding acquisition: JHP, BHC, JWC. Investigation: DN, JL. Methodology: DN, JL. Project administration: JHP, BHC, JWC. Resources: SJK. Software: DN, BHC. Supervision: JHP, BHC, JWC. Visualization: DN, JL. Writing – original draft: DN, JL, BHC. Writing – review and editing: JHP, BHC, JWC.

References

1. Friedenstejn AJ, Piatetzky-Shapiro II, Petrakova KV. Osteogenesis in transplants of bone marrow cells. *J Embryol Exp Morphol* 1966;16:381-390
2. Caplan AI, Correa D. The MSC: an injury drugstore. *Cell Stem Cell* 2011;9:11-15
3. Margiana R, Markov A, Zekiy AO, et al. Clinical application of mesenchymal stem cell in regenerative medicine: a narrative review. *Stem Cell Res Ther* 2022;13:366
4. Jovic D, Yu Y, Wang D, et al. A brief overview of global trends in MSC-based cell therapy. *Stem Cell Rev Rep* 2022;18:1525-1545
5. Dominici M, Le Blanc K, Mueller I, et al. Minimal criteria for defining multipotent mesenchymal stromal cells. The International Society for Cellular Therapy position statement. *Cytotherapy* 2006;8:315-317
6. Berebichez-Fridman R, Montero-Olvera PR. Sources and clinical applications of mesenchymal stem cells: state-of-the-art review. *Sultan Qaboos Univ Med J* 2018;18:e264-e277
7. Fernández-Santos ME, Garcia-Arranz M, Andreu EJ, et al. Optimization of mesenchymal stromal cell (MSC) manufacturing processes for a better therapeutic outcome. *Front Immunol* 2022;13:918565
8. Fernández-Garza LE, Barrera-Barrera SA, Barrera-Saldaña HA. Mesenchymal stem cell therapies approved by regulatory agencies around the world. *Pharmaceuticals (Basel)* 2023;16:1334
9. Hassan MNFB, Yazid MD, Yunus MHM, et al. Large-scale expansion of human mesenchymal stem cells. *Stem Cells Int* 2020;2020:9529465
10. Mota SM, Rogers RE, Haskell AW, et al. Automated mesenchymal stem cell segmentation and machine learning-based phenotype classification using morphometric and textural analysis. *J Med Imaging (Bellingham)* 2021;8:014503
11. Bai H, Lu C, Ma M, Yan S, Zhang J, Han Z. An improved U-Net for cell confluence estimation. *Optoelectron Lett* 2022;18:378-384
12. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 2017;542:115-118
13. Kwon S, Ki SM, Park SE, et al. Anti-apoptotic effects of human Wharton's jelly-derived mesenchymal stem cells on skeletal muscle cells mediated via secretion of XCL1. *Mol*

- Ther 2016;24:1550-1560
14. Alzubaidi L, Zhang J, Humaidi AJ, et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J Big Data* 2021;8:53
 15. Ma N, Zhang X, Zheng HT, Sun J. ShuffleNet V2: practical guidelines for efficient CNN architecture design. In: Ferrari V, Hebert M, Sminchisescu C, Weiss Y, editors. *Computer Vision – ECCV 2018*. Springer, Cham; 2018. 122-138
 16. Howard A, Sandler M, Chen B, et al. In: *IEEE International Conference on Computer Vision, IEEE Computer Society, editors. Searching for MobileNetV3. 2019 IEEE/CVF International Conference on Computer Vision (ICCV)*. IEEE; 2019. 1314-1324
 17. Dosovitskiy A, Beyer L, Kolesnikov A. An image is worth 16x16 words: transformers for image recognition at scale. *arXiv:2010.11929* [Preprint]. 2020 [cited 2023 Mar 10]. Available from: <https://doi.org/10.48550/arXiv.2010.11929>
 18. Lin TY, Goyal P, Girshick R, He K, Dollar P. Focal loss for dense object detection. *IEEE Trans Pattern Anal Mach Intell* 2020;42:318-327
 19. Yin JQ, Zhu J, Ankrum JA. Manufacturing of primed mesenchymal stromal cells for therapy. *Nat Biomed Eng* 2019;3:90-104