



OPEN

High throughput screening of mesenchymal stem cell lines using deep learning

Gyuwon Kim^{1,5}, Jung Ho Jeon^{2,3,5}, Keonhyeok Park¹, Sung Won Kim²✉ & Seungchul Lee^{1,4}✉

Mesenchymal stem cells (MSCs) are increasingly used as regenerative therapies for patients in the preclinical and clinical phases of various diseases. However, the main limitations of such therapies include functional heterogeneity and the lack of appropriate quality control (QC) methods for functional screening of MSC lines; thus, clinical outcomes are inconsistent. Recently, machine learning (ML)-based methods, in conjunction with single-cell morphological profiling, have been proposed as alternatives to conventional *in vitro/vivo* assays that evaluate MSC functions. Such methods perform *in silico* analyses of MSC functions by training ML algorithms to find highly nonlinear connections between MSC functions and morphology. Although such approaches are promising, they are limited in that extensive, high-content single-cell imaging is required; moreover, manually identified morphological features cannot be generalized to other experimental settings. To address these limitations, we propose an end-to-end deep learning (DL) framework for functional screening of MSC lines using live-cell microscopic images of MSC populations. We quantitatively evaluate various convolutional neural network (CNN) models and demonstrate that our method accurately classifies *in vitro* MSC lines to high/low multilineage differentiating stress-enduring (MUSE) cells markers from multiple donors. A total of 6,120 cell images were obtained from 8 MSC lines, and they were classified into two groups according to MUSE cell markers analyzed by immunofluorescence staining and FACS. The optimized DenseNet121 model showed area under the curve (AUC) 0.975, accuracy 0.922, F1 0.922, sensitivity 0.905, specificity 0.942, positive predictive value 0.940, and negative predictive value 0.908. Therefore, our DL-based framework is a convenient high-throughput method that could serve as an effective QC strategy in future clinical biomanufacturing processes.

Recent advances in regenerative medicine and tissue engineering have enabled stem cell therapies to progress to the preclinical and clinical phases in treating various degenerative diseases^{1,2}. In particular, mesenchymal stem cells (MSCs) have received considerable attention in tissue regeneration and the treatment of immune system-mediated diseases³. Although MSCs exhibit regenerative and immunomodulatory potentials, clinical trials have reported inconsistent therapeutic efficacies^{4,5}, likely attributable to functional heterogeneities among MSCs from different donors, variations in MSC production, treatment methods, and recipient conditions; all of these factors exhibit considerable variability^{4,6}. In addition, vast numbers of functional stem cells are required for tissue regeneration, in which mass production (using automated cell culture systems) is already commercially available^{7,8}. Therefore, effective quality control (QC) of MSC functions is needed for consistent, high-quality, large-scale biomanufacturing of MSCs and their products, as well as successful clinical translation^{9,10}. However, current assessment methods lack clinical relevance, adequate assay throughput, and robustness; improved techniques are necessary to better characterize MSC functions.

Recently, several studies have found connections between MSC functions and cell morphology by exploiting advances in high-content microscopic imaging¹¹. For example, MSC morphology has been correlated with differentiation capacity^{12,13}, motility¹⁴, and passage number^{15,16}. Another field of study seeks to predict *in vitro* MSC

¹Department of Mechanical Engineering, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea. ²Department of Otolaryngology-Head and Neck Surgery, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, 222 Banpo-Daero, Seocho-Gu, Seoul 06591, Republic of Korea. ³Department of Biomedicine and Health Science, College of Medicine, The Catholic University of Korea, 222 Banpo-Daero, Seocho-Gu, Seoul 06591, Republic of Korea. ⁴Graduate School of Artificial Intelligence, Pohang University of Science and Technology (POSTECH), Pohang 37673, Republic of Korea. ⁵These authors contributed equally: Gyuwon Kim and Jung Ho Jeon. ✉email: dohyuni9292@naver.com; seunglee@postech.ac.kr

functions via morphological profiling that employs machine learning (ML) algorithms, of which deep learning (DL) algorithms are optimal. Such algorithms have been used to predict MSC osteogenic potential^{17,18}, immunosuppressive capacity^{19,20}, and microenvironmental interactions^{21,22}. These works yield *in silico* predictions based solely on cell morphology, potentially eliminating the need for costly *in vitro/vivo* experiments. Candidate therapeutic stem cells can be then identified by selecting lines that are expected to be highly effective and then confirming such efficacy. If efficacy can be predicted using light microscope images periodically acquired during cell culture, the required candidates can be derived without compromising the culture flow.

The efficacy indicators of human mesenchymal tissues may vary depending on the disease that requires therapy. The multilineage differentiating stress-enduring (MUSE) marker evaluates stem cell capacity. MUSE cells are nontumorigenic pluripotent stem cells of human mesenchymal tissues expressing CD-105 and SSEA-3²³. MUSE cells readily home to damaged tissues and spontaneously differentiate into cells of such tissues, thus repairing the tissues and restoring function^{24,25}. Furthermore, MUSE cells exhibit immunomodulatory properties²⁶, and DNA repair capacity²⁴; these cells have been used to treat cardiovascular and neurological diseases^{27,28}. Therefore, it could be considered as an overall marker for evaluating the efficacy of stem cells.

In this study, human neural crest-derived nasal turbinate stem cells (hNTSCs) was used as the cell source for the analysis. In a previous study, after isolation of hNTSCs from inferior turbinate nasal tissue, MSC characteristics were investigated according to the statement of position of the International Society for Cell Therapy²⁹. Surface epitope analysis revealed that hNTSCs were negative for CD14, CD19, CD34 and HLA-DR and positive for CD29, CD73 and CD90, indicating a characteristic phenotype of MSCs. The plasticity of hNTSCs has been confirmed in cartilage, bone, adipose and neuronal differentiation conditions^{29–33}. hNTSCs have several advantages: donor morbidity is low, and the collection time is short (5–10 min); the cells are derived from the easily acquired turbinate mucosa of patients undergoing surgical treatment of chronic hypertrophic rhinitis (a common condition). The cell numbers are high; these cells exhibit good proliferation, multilineage differentiation potential, and immunomodulatory properties^{29,34,35}.

Therefore, we investigated whether high- and low-functioning hNTSC cell lines represented by MUSE markers could be distinguished from DLs through a simple live-cell microscopy imaging process.

Materials and methods

This section describes cell culture, *in vitro* assessments, and imaging methods and provides general explanations of DL-based methodologies and evaluation metrics.

Cell culture/expansion. We used human neural crest-derived nasal turbinate stem cells (hNTSCs); the work was approved by the Ethics Committee (approval no. KC08TISS0341) of Seoul St. Mary's Hospital of the Catholic University of Korea and all methods were performed in accordance with the relevant guidelines and regulations. Prior to surgery, participants provided written informed consent. hNTSCs were isolated from the turbinate tissue of a human who had undergone partial turbinate resection. The lower turbinate tissue was washed with saline and phosphate-buffered saline (Thermo Fisher Scientific, Waltham, MA, USA) containing an antibiotic-antibacterial solution (Thermo Fisher Scientific), then cut into 1-mm³ pieces. Subsequently, the dish was covered with a sterile glass slide. The growth medium was α -minimum essential medium (α -MEM, Thermo Fisher Scientific) with 1% (v/v) penicillin/streptomycin (Invitrogen) and 10% (v/v) fetal bovine serum (Thermo Fisher Scientific). Incubation was performed at 37 °C in a humidified atmosphere under 5% (v/v) CO₂. The medium was changed at 2–3-day intervals during the 3-week culture period. Finally, the glass cover slide was removed, and cells attached to the culture plate were harvested into a 0.25% (w/v) trypsin solution in 1 mM ethylenediaminetetraacetic acid. hNTSCs were expanded for use in experiments. To analyze MUSE marker expression and deep learning model training in the culture (passage 6), 1 × 10⁶ cells were seeded in 6-well culture plate and a training image was taken. 8 MSC lines were classified into two groups according to MUSE cell markers analyzed by immunofluorescence staining and FACS.

Flow cytometry. Single cell suspensions were prepared from hNTSCs. Cells were incubated for 30 min at 4 °C with stage-specific embryonic antigen-3 (SSEA-3) antibody (1:100, Abcam, Cambridge, UK, ab16286) followed by Alexa Fluor 633 anti-rat antibody (1:1000, Thermo Fisher Scientific, A21094). After incubation with the SSEA-3 antibody, the cells were incubated with CD105 antibody (1:100, PE-conjugated, BD Pharmingen; catalog no. 560839) for 30 min at 4 °C for double staining. The cells were re-suspended in DPBS (Gibco) and acquired through FACS Canto II (BD biosciences) with DIVA software.

Imaging and preprocessing. Live-cell microscopic imaging was performed using the Lionheart LX automated microscope (BioTek, Winooski, VT, USA) at 37 °C under 5% (v/v) CO₂. Images were acquired using phase objectives (40× and 100×). Data were acquired as 904 × 1224 8-bit grayscale images, which were preprocessed in terms of hue (H), saturation (S), and value (V) to reduce experimental variations. Images with mean V values > 230 were excluded because their excessive brightness impeded the distinction of cell morphology from the background. The remaining images were adjusted to achieve a mean V value of 130; the images were resized to 226 × 306 pixels using an interpolation method within Python OpenCV Toolbox software. Subsequently, the pixel values were normalized to [0, 1].

Immunofluorescence staining. The expression levels of SSEA-3 (derived using anti-SSEA-3; 1:300, Abcam, Cambridge, UK, ab16286) and CD-105 (derived using a PE mouse anti-human CD-105; 1:300, BD Pharmingen; catalog no. 560839) were determined via immunofluorescence staining. After 2 days of culture in the medium described above, hNTSCs were fixed in 2% (w/v) paraformaldehyde and washed with phos-

phate-buffered saline. The cells were then permeabilized with 0.3% (v/v) Triton X-100 (Sigma-Aldrich) and washed with phosphate-buffered saline. After cells had been blocked with 1% (v/v) normal goat serum (Jackson ImmunoResearch Laboratories Inc., West Grove, PA, USA), they were incubated with the primary antibodies mentioned above; they were then incubated with a goat anti-rat Alexa-Fluor 488 antibody (1:1000; Molecular Probes). The nuclei were labeled with 4',6-diamidino-2-phenylindole (DAPI; Sigma-Aldrich), and fluorescence was observed under a Zeiss LSM510 confocal microscope (Carl Zeiss).

Deep learning. We compared various convolutional neural network (CNN) models when selecting the appropriate architecture. Specifically, we reviewed VGG19³⁶, ResNet50V2³⁷, DenseNet121³⁸, InceptionV3³⁹, and Xception⁴⁰ (Supplementary Fig. 1). The final feature extraction layers of all networks were subjected to single-element averaging using global average pooling⁴¹, followed by the designation of surrogate classification layers as described in Supplementary Table 1. We leveraged various strategies to train our networks and optimize the trainable parameters. First, we employed transfer learning⁴², which effectively transfers knowledge among different domains and has been successfully employed in several biomedical imaging applications⁴³. Our transfer learning strategy featured three steps (Supplementary Fig. 2). In the first step, we obtained pre-trained networks on the ImageNet dataset⁴⁴, in which the feature extraction layers are pre-trained to identify low-level image features (Supplementary Fig. 2a). The trainable parameters of the feature extraction layers were frozen, and a new classification layer (trained using our cell image dataset) classified cell populations into high/low MUSE cells markers using the features learned in the previous step (Supplementary Fig. 2b). Finally, the entire network was fine-tuned at a reduced learning rate (Supplementary Fig. 2c). During training, trainable parameters were initialized using the normal He normalization method⁴⁵ and optimized using the Adam optimizer⁴⁶. Early stopping, L2 regularization⁴⁷, and dropout⁴⁸ were used to avoid overfitting. We then optimized each model's hyperparameters using Bayesian optimization (BO)⁴⁹ (Supplementary Fig. 3). Unlike systematic approaches such as the grid and random searches (Supplementary Fig. 3a,b respectively)⁵⁰, BO reduces the computational cost by considering all prior knowledge that facilitates the optimization process (Supplementary Fig. 3c). BO features simultaneous interactions between a surrogate model and an acquisition function. The surrogate model probabilistically models an unknown objective function that maps a set of hyperparameters to the evaluation results by considering previously evaluated samples. The acquisition function suggests a new set of hyperparameters based on exploration of (and exploitation within) the established surrogate model. A Gaussian process⁵¹ served as the surrogate model, while the expected improvement (EI) algorithm⁵² served as the acquisition function. Further details regarding model training are provided in Supplementary Table 1.

Evaluation metrics. To evaluate the results of our CNN models, we considered areas under the curves (AUCs), F1 scores, accuracies, sensitivities, specificities, positive predictive values (PPVs), and negative predictive values (NPVs) (Supplementary Table 2).

Statistical analysis. Statistical difference between two groups were assessed with the unpaired t test (two tailed). All results are presented as the mean \pm SD. A p-value of less than 0.05 was considered statistically significant.

Results

This section presents and discusses the experimental results of our DL-based MSC screening method. Previous studies found that the number of functional subpopulations was correlated with the overall therapeutic functionality of an MSC batch^{53–55}; moreover, subpopulations exhibited distinct morphological features^{20,56–58}. We thus hypothesized that microscopic images from functional cell lines contain a higher ratio of such subpopulation distributions; we also hypothesized that a CNN could classify the images based on the morphological characteristics. Therefore, all images were labeled as a whole; images of cell lines with high MUSE cell markers were positive, whereas images of cell lines with low MUSE cell markers were negative. A schematic is shown in Fig. 1. We first conducted multivariate in vitro assessments of eight cell lines from different donors. group with a high expression of MUSE cells (n = 4) and group with a low expression of MUSE cells (n = 4) were identified (Supplementary Fig. 4). Then (as shown in Fig. 1), the total dataset was split into training, validation, and test sets. Three sets of MSC images were obtained for training from six cell lines; each set contained 1,530 images from a positive cell line and 1530 images from a negative cell line. This dataset was used to conduct threefold cross-validation that quantitatively compared the classification performances of various CNN models by splitting the dataset into non-overlapping subsets. We identified the most promising model, then used this model to conduct a further quantitative and qualitative evaluation of an independent dataset with 6,120 cell images/class from test data.

Characteristics of the classified hNTSC group. Confocal microscopy images confirmed the difference in the MUSE cell markers SSEA-3 and CD-105 between the two groups (Supplementary Fig. 5). Supplementary Fig. 6 showed flow cytometry results of MUSE cells in the six passage showing of SSEA-3, CD105 expression. In FACS analyses, which were performed to confirm this result quantitatively, the results of comparing the MUSE cell markers of the four high groups and the four low groups showed that the MUSE cell marker expression was approximately twofold higher in the high group (3.5%) than in the low group (1.7%). It was confirmed that there was a significant difference between the forces (Supplementary Fig. 6b).

Bayesian optimization and model comparison. Using the search spaces of the hyperparameters listed in Supplementary Table 1, we subjected various models to BO (we recorded the mean threefold cross-validation

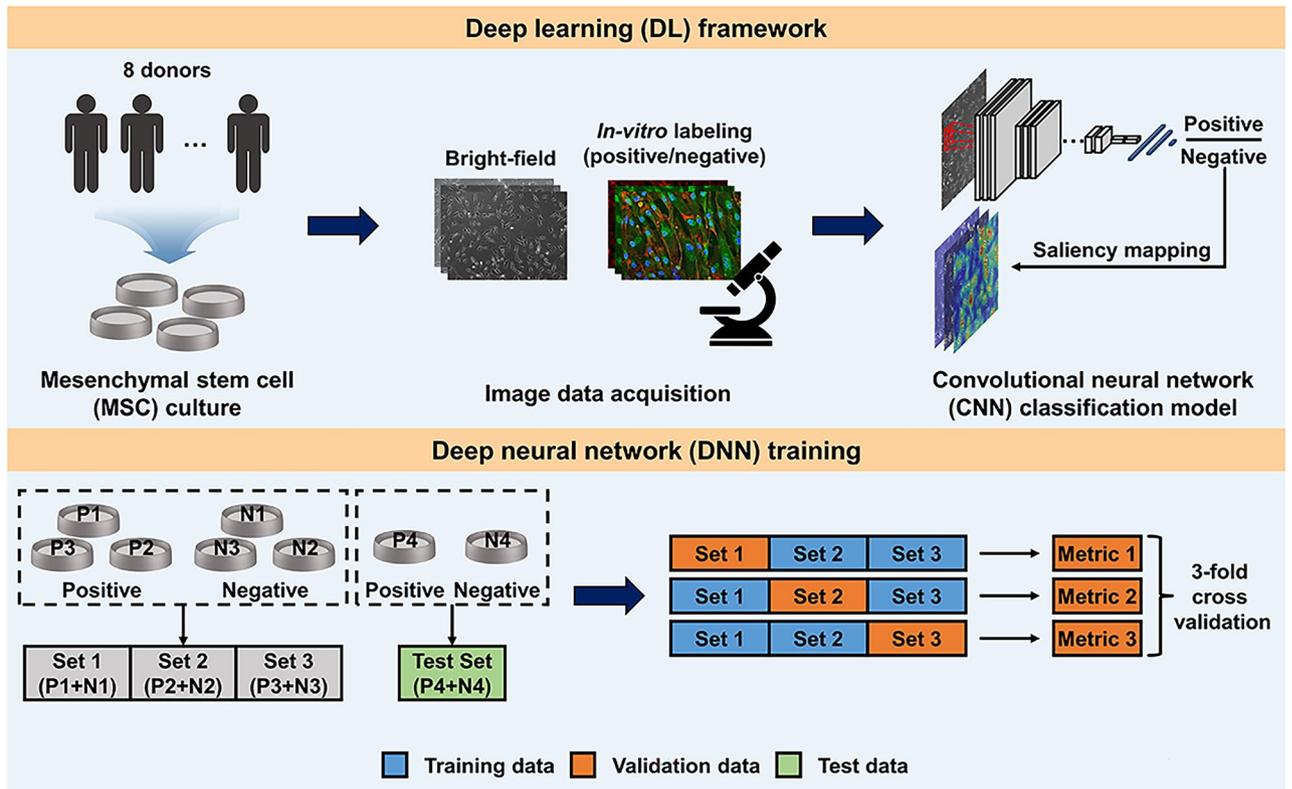


Figure 1. Schematic of the deep learning (DL) framework used to screen for functional mesenchymal stem cell (MSC) lines. Mesenchymal stem cell (MSC) cultures were obtained from eight cell lines from different donors for image data acquisition. To compare the classification performances of various convolutional neural network (CNN) models, threefold cross-validation was conducted by splitting the dataset into non-overlapping subsets.

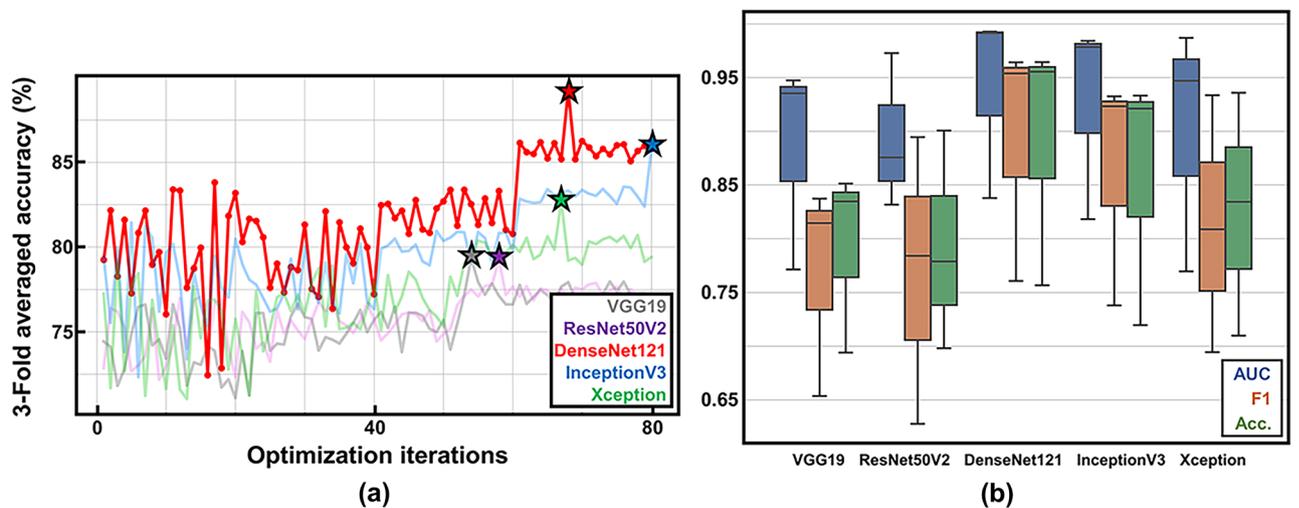


Figure 2. Hyperparameter optimization of comparative convolutional neural network (CNN) models during 80 Bayesian optimization (BO) iterations revealed by mean threefold cross-validation accuracies. (a) A progressive plot illustrating the BO process (stars indicate points of optimal performance and DenseNet121 was considered the best model) and (b) the threefold cross-validation metrics for each model after BO. AUC, the area under the curve; and Acc., accuracy.

accuracies); the results are shown in Fig. 2 and Table 1 while the confusion matrices are shown in Supplementary Fig. 7. BO utilizes prior knowledge concerning the optimization process, then searches for new hyperparameter sets (Fig. 2a). The threefold cross-validation accuracy gradually increased with increasing optimization; early, accuracy was low, and variance was high; later, accuracy was high and appeared to converge for all models. The optimal hyperparameter sets are starred; the values are listed in Supplementary Table 1. As shown in Fig. 2b, DenseNet121 outperformed the other methods in most evaluation metrics. When analyzed with the

	AUC	F1	Accuracy	Sensitivity	Specificity	PPV	NPV	FPS
VGG19	0.885±0.098	0.768±0.1	0.793±0.086	0.689±0.099	0.897±0.075	0.869±0.01	0.744±0.076	964±48
ResNet50V2	0.893±0.072	0.769±0.134	0.792±0.102	0.718±0.182	0.867±0.103	0.846±0.097	0.765±0.11	458±61
DenseNet121	0.941±0.089	0.893±0.115	0.892±0.117	0.885±0.099	0.899±0.138	0.902±0.133	0.884±0.104	378±25
InceptionV3	0.927±0.094	0.865±0.11	0.858±0.12	0.886±0.084	0.83±0.159	0.846±0.135	0.874±0.103	519±30
Xception	0.901±0.116	0.812±0.12	0.827±0.113	0.754±0.13	0.9±0.121	0.886±0.133	0.787±0.11	364±34

Table 1. Threefold cross-validation metrics of all convolutional neural networks (CNNs) after Bayesian optimization (BO). AUC, the area under the curve; PPV, positive predictive value; NPV, negative predictive value; FPS, frames per second. Significant values are given in bold.

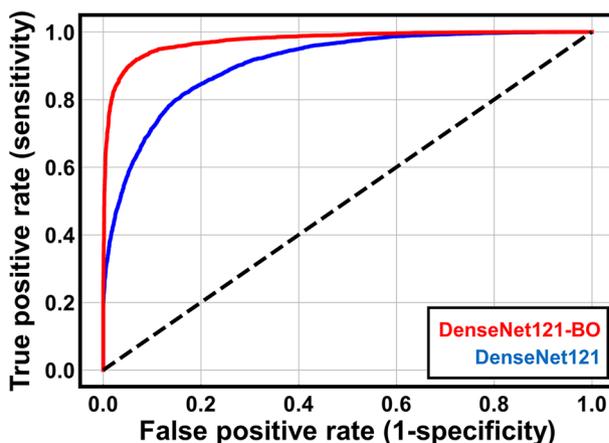


Figure 3. Receiver operating characteristic (ROC) curves of DenseNet121 operating on the independent test dataset before and after Bayesian optimization (BO). The prediction performance of DenseNet121 was improved after the BO process.

	AUC	F1	Accuracy	Sensitivity	Specificity	PPV	NPV	FPS
DenseNet121-BO	0.975	0.922	0.923	0.905	0.942	0.940	0.908	324
DenseNet121	0.908	0.820	0.825	0.800	0.849	0.842	0.810	389

Table 2. Quantitative metrics of DenseNet121 derived using an independent test dataset before and after Bayesian optimization (BO). AUC, the area under the curve; PPV, positive predictive value; NPV, negative predictive value; FPS, frames per second. Significant values are given in bold.

DenseNet121 model, AUC 0.941 ± 0.089 , F1 0.893 ± 0.115 , accuracy 0.892 ± 0.117 , sensitivity 0.885 ± 0.099 , specificity 0.899 ± 0.138 , PPV 0.902 ± 0.133 , and NPV 0.884 ± 0.104 were obtained. Comparing the best-performing CNN with others, DenseNet121 had an AUC average score of 0.941, while VGG19 had an AUC score of 0.885; VGG19 exhibited the worst performance, reflecting the consensus that deeper networks benefit from their ability to model mappings of higher complexity³⁶. InceptionV3 and Xception models performed poorly compared to DenseNet121; these findings suggested that their multi-scale convolution operations overexpressed our cell image dataset, thus leading to overfitting. Although VGG19 exhibits the fastest inference speed due to its relatively simple architecture [964 frames per second (FPS)], all models perform above 300 FPS, which is sufficient for real-time implementations. The threefold cross-validation results indicate that our method robustly managed the heterogeneities of cell lines. Below, we focus exclusively on DenseNet121.

Independent test set evaluation. All images from the threefold dataset were merged and used to retrain DenseNet121 with the hyperparameter set suggested by BO. We evaluated the performances of the optimized DenseNet121 (termed DenseNet121-BO) and the pre-optimized DenseNet121 (trained using the initial hyperparameters of Supplementary Table 1) using an independent test dataset derived from the remaining cell lines (Fig. 3 and Table 2; the confusion matrices are shown in Supplementary Fig. 8). Quantitatively, all metrics (including receiver operating characteristic curves) improved after BO. The optimized DenseNet121 model was improved to AUC 0.975, accuracy 0.922, F1 0.922, sensitivity 0.905, specificity 0.942, PPV 0.940, and NPV 0.908.

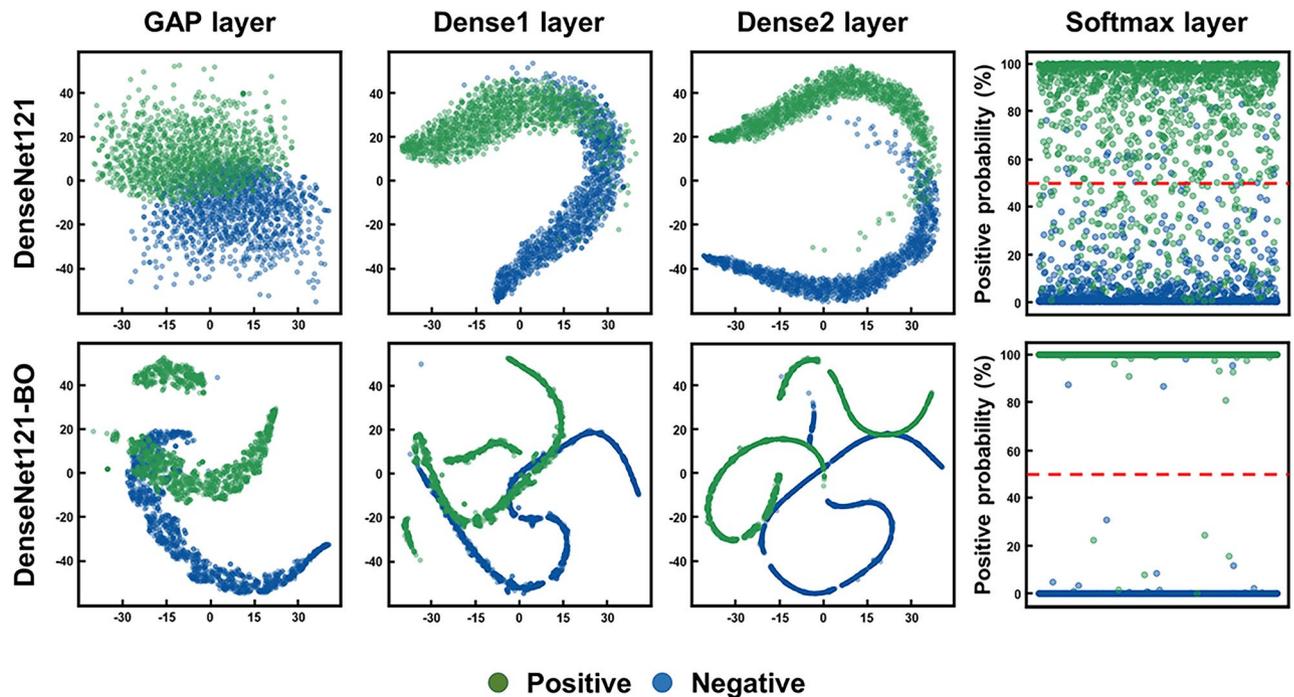


Figure 4. t-stochastic neighborhood embedding (t-SNE) visualizations of the feature outputs of different classification layers and the distributions of softmax outputs of the final classification layers before and after Bayesian optimization (BO). As the layers progressed within the models and after the BO process, the distributions of the two classes gradually became disentangled. GAP, global average pooling.

The FPS deteriorates slightly since the number of dense neurons in the classification layers increases after BO (Supplementary Table 1). However, FPS values above 300 are sufficient for real-time screening. The prediction performance of DenseNet121 after BO process would be expressed with an AUC score of 0.975. To further analyze the effects of BO, we examined the feature representations of all classification layers before and after BO using the t-stochastic neighborhood embedding algorithm⁵⁹. Figure 4 shows the embedded feature distributions of the test dataset in the global average pooling and dense layers of our model; the distributions of the two classes gradually became disentangled as they progressed within the layers. Notably, the layer distributions became more distinct after BO because BO optimizes the dense layer structures (Supplementary Table 1). Overall, BO improved the generalization afforded by our DenseNet121 model. Finally, we used several saliency mapping algorithms to highlight regions of interest (ROIs) that substantially contributed to the classification results. For this purpose, we used gradient-based class activation mapping (Grad-CAM++)⁶⁰, layer-wise relevance propagation (LRP)⁶¹, and visual back-propagation (VisualBackProp)⁶². Sample true-positive (TP) and true-negative (TN) images (with ROIs) are shown in Fig. 5. All three algorithms highlighted similar ROIs within the images; the TP images featured large ROIs (upper two rows), and the TN images featured small ROIs (bottom two rows). We calculated the normalized pixel scores of all images via saliency mapping by thresholding each pixel value below 0.5 to 0, then averaging the remaining values. The normalized pixel score distributions revealed clear differences between TP and TN images. The Grad-CAM++ distribution illustrates overlaps between the two classes due to the algorithm's tendency to highlight relatively large ROIs, resulting in large normalized pixel scores. The opposite was true of VisualBackProp and LRP; the ROIs of TP images were small, and the ROIs of TN images were nearly absent, which led to distinct distributions and small normalized pixel scores.

Discussion

In this study, we divided stem cell lines with different MUSE cell markers into two groups (high vs. low MUSE expression) when establishing our DL model that was trained using simple light microscopic images. Evaluation of hNTSCs with distinct patterns of MUSE cell marker expression enabled the separation of hNTSCs donor variation. It was verified that SSEA-3 and CD105 of the high MUSE expression group were higher than those of the low MUSE expression group by using immunofluorescence staining, and it was confirmed more clearly by FACS analysis. The optimized DenseNet121 DL model showed AUC 0.975, accuracy 0.922, F1 0.922, sensitivity 0.905, specificity 0.942, PPV 0.940, NPV 0.908 results, therefore, it is possible to discriminate groups with high or low expression of MUSE cell markers from cell images with very high accuracy.

Although ML-based methods effectively predict in vitro MSC function, several aspects require investigation; these motivated our work. First, DL-based principles should be used to create an end-to-end framework that optimizes feature extraction⁴⁷. ML algorithms require manual identification of morphological features in single-cell images; thus, high-content imaging followed by extensive image-processing is necessary⁶³. This approach limits screening throughput and hinders the complete exploitation of high-dimensional morphological information inherent in the image data. Additionally, manually identified features are experiment-dependent; they are not

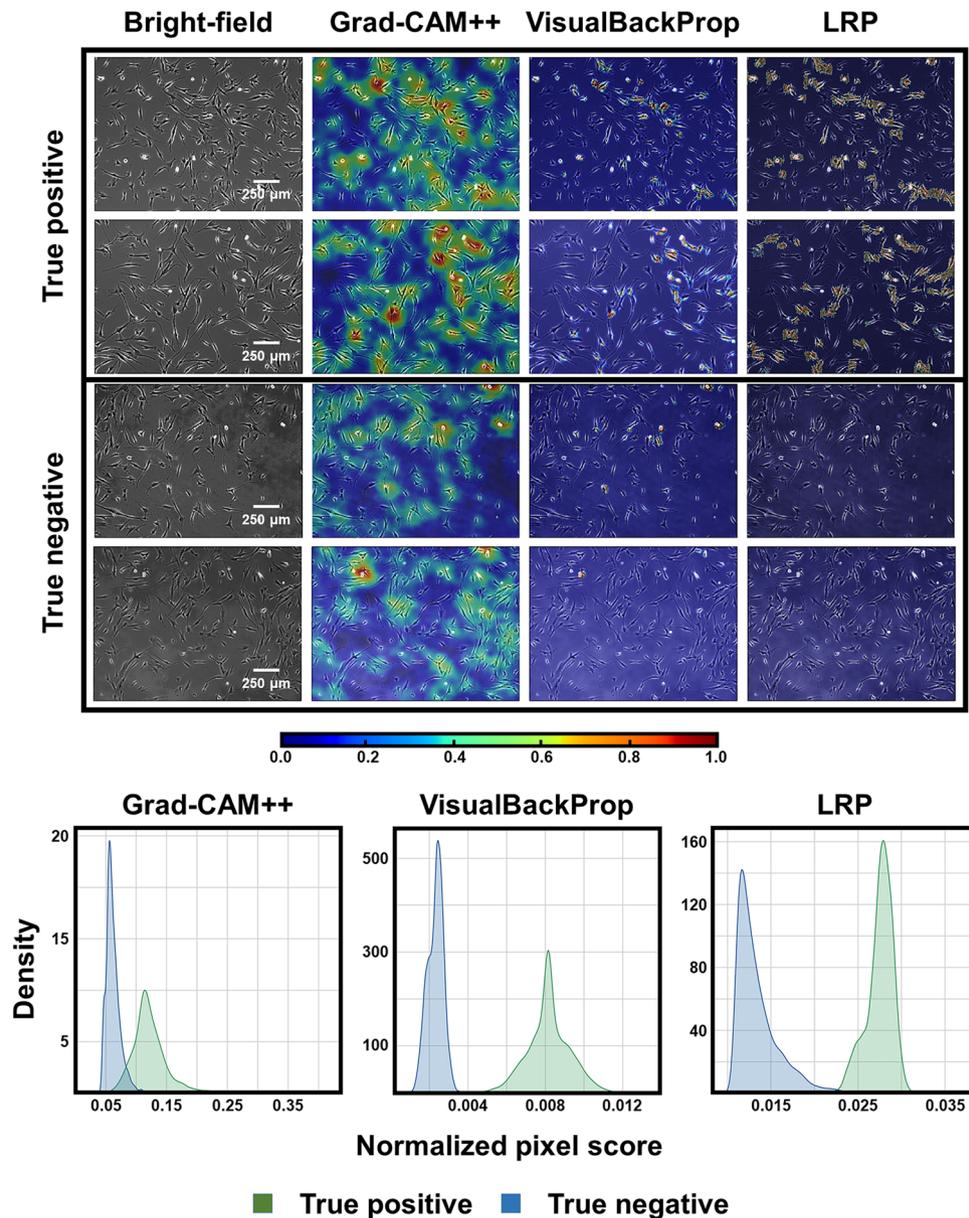


Figure 5. Interpretation of convolutional neural network (CNN) analyses using three saliency mapping algorithms: gradient-based class activation mapping (Grad-CAM++), visual back-propagation (VisualBackProp), and layer-wise relevance propagation (LRP). The Grad-CAM++ exhibited overlaps between the two classes due to the algorithms' propensity to highlight large ROIs, while VisualBackProp and LRP provided distinct distributions. Normalized pixel score distributions are shown for each algorithm.

generalizable to cells produced under different conditions¹⁹. Second, an optimal method must be generalizable to MSCs from various donors evaluated under different experimental conditions. Although a previous study⁶⁴ successfully utilized CNNs that evaluated cell morphology to predict the immunofluorescence levels of nine different surface markers, the proof-of-concept study reported results from only a single donor culture after only a few passages at a low seeding density. To address the above limitations, we proposed a novel DL-based method to identify MSC lines to high/low MUSE cells markers using live microscopic images. Previous research^{20,56–58} showed that the proportions of morphological subpopulations reflect MSC therapeutic functions. Also, the number of functional subpopulations was correlated with the overall therapeutic functionality of an MSC batch^{53–55}. We thus hypothesized that CNNs could be used to examine whole-labeled, live-cell microscopic images of MSC populations, then identify the *in vitro* characteristics of those populations. By leveraging various DL principles, we circumvented the extensive imaging procedures that were required by previous ML approaches; our MSC screening strategy is much faster performing above 300 FPS. We initially established and optimized several CNN models, using BO; DenseNet121 was considered the best model (Fig. 2, Table 1). Our model demonstrated AUCs > 0.94 during validation involving MSC images from eight different cell lines. We then explored the interpretations of the CNN using various saliency mapping algorithms; we sought to ensure that our method was

reliable. When used with single-cell morphological profiling⁶³, we provided insights that could serve as potential biomarker candidates for identifying desired subpopulations. The visualization results strongly supported our initial hypothesis: CNNs can classify whole-labeled images into high/low MUSE cells markers by examining the morphological features of subpopulations, as well as the proportions of such subpopulations. Our DL-based method is generalizable and offers high throughput. Identifying surface markers unique to these subpopulations⁶⁵ would enable their enrichment using methods such as fluorescence-activated cell sorting (FACS)⁵³. Finally, our approach will facilitate inter- and intra-laboratory research; morphological profiling is more cost-effective and less sensitive to experimental conditions compared with conventional immunohistochemistry- and flow cytometry-based methods⁶⁶.

Conclusions

We developed a DL-based *in silico* method to identify MSC lines to high/low MUSE cells markers; this method facilitates high-throughput QC during biomanufacturing. Through this process, function screening, replacing some of the *in vivo* and *in vitro* tests performed to confirm the function of each cell stem line can be accomplished. Therefore, our non-invasive, automated *in silico* system, which uses simple microscopy alone, will aid cell biomanufacturing and translation to cell therapies.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 17 May 2022; Accepted: 29 September 2022

Published online: 20 October 2022

References

- Margiana, R. *et al.* Clinical application of mesenchymal stem cell in regenerative medicine: A narrative review. *Stem Cell Res. Ther.* **13**, 366. <https://doi.org/10.1186/s13287-022-03054-0> (2022).
- Kouchakian, M. R. *et al.* The clinical trials of mesenchymal stromal cells therapy. *Stem Cells Int.* **2021**, 1634782. <https://doi.org/10.1155/2021/1634782> (2021).
- Shandil, R. K., Dhup, S. & Narayanan, S. Evaluation of the therapeutic potential of mesenchymal stem cells (MSCs) in preclinical models of autoimmune diseases. *Stem Cells Int.* **2022**, 6379161. <https://doi.org/10.1155/2022/6379161> (2022).
- Sharma, R. R., Pollock, K., Hubel, A. & McKenna, D. Mesenchymal stem or stromal cells: A review of clinical applications and manufacturing practices. *Transfusion* **54**, 1418–1437 (2014).
- Griffin, M. D. *et al.* Concise review: Adult mesenchymal stromal cell therapy for inflammatory diseases: How well are we joining the dots?. *Stem Cells* **31**, 2033–2041 (2013).
- Rizk, M. *et al.* Heterogeneity in studies of mesenchymal stromal cells to treat or prevent graft-versus-host disease: A scoping review of the evidence. *Biol. Blood Marrow Transplant.* **22**, 1416–1423 (2016).
- Thomas, R. J. *et al.* Manufacture of a human mesenchymal stem cell population using an automated cell culture platform. *Cytotechnology* **55**, 31–39. <https://doi.org/10.1007/s10616-007-9091-2> (2007).
- Yoshimoto, N. *et al.* An automated system for high-throughput single cell-based breeding. *Sci. Rep.* **3**, 1191. <https://doi.org/10.1038/srep01191> (2013).
- Dwarshuis, N. J., Parratt, K., Santiago-Miranda, A. & Roy, K. Cells as advanced therapeutics: State-of-the-art, challenges, and opportunities in large scale biomanufacturing of high-quality cells for adoptive immunotherapies. *Adv. Drug Deliv. Rev.* **114**, 222–239 (2017).
- Aijaz, A. *et al.* Biomanufacturing for clinically advanced cell therapies. *Nat. Biomed. Eng.* **2**, 362–376 (2018).
- Singh, S., Carpenter, A. E. & Genovesio, A. Increasing the content of high-content screening: An overview. *J. Biomol. Screen.* **19**, 640–650 (2014).
- Kilian, K. A., Bugarija, B., Lahn, B. T. & Mrksich, M. Geometric cues for directing the differentiation of mesenchymal stem cells. *Proc. Natl. Acad. Sci.* **107**, 4872–4877 (2010).
- Surdo, J. L. L., Millis, B. A. & Bauer, S. R. Automated microscopy as a quantitative method to measure differences in adipogenic differentiation in preparations of human mesenchymal stromal cells. *Cytotherapy* **15**, 1527–1540 (2013).
- Bertolo, A. *et al.* *In vitro* cell motility as a potential mesenchymal stem cell marker for multipotency. *Stem Cells Transl. Med.* **4**, 84–90 (2015).
- Lee, W. C. *et al.* Multivariate biophysical markers predictive of mesenchymal stromal cell multipotency. *Proc. Natl. Acad. Sci.* **111**, E4409–E4418 (2014).
- Lo Surdo, J. & Bauer, S. R. Quantitative approaches to detect donor and passage differences in adipogenic potential and clonogenicity in human bone marrow-derived mesenchymal stem cells. *Tissue Eng. Part C Methods* **18**, 877–889 (2012).
- Marklein, R. A. *et al.* High content imaging of early morphological signatures predicts long term mineralization capacity of human mesenchymal stem cells upon osteogenic induction. *Stem Cells* **34**, 935–947 (2016).
- Matsuoka, F. *et al.* Morphology-based prediction of osteogenic differentiation potential of human mesenchymal stem cells. *PLoS ONE* **8**, e55082 (2013).
- Klinker, M. W., Marklein, R. A., Surdo, J. L. L., Wei, C.-H. & Bauer, S. R. Morphological features of IFN- γ -stimulated mesenchymal stromal cells predict overall immunosuppressive capacity. *Proc. Natl. Acad. Sci.* **114**, E2598–E2607 (2017).
- Marklein, R. A. *et al.* Morphological profiling using machine learning reveals emergent subpopulations of interferon- γ -stimulated mesenchymal stromal cells that predict immunosuppression. *Cytotherapy* **21**, 17–31 (2019).
- Chen, D. *et al.* Machine learning based methodology to identify cell shape phenotypes associated with microenvironmental cues. *Biomaterials* **104**, 104–118 (2016).
- Vega, S. L. *et al.* High-content imaging-based screening of microenvironment-induced changes to stem cells. *J. Biomol. Screen.* **17**, 1151–1162 (2012).
- Kuroda, Y. *et al.* Unique multipotent cells in adult human mesenchymal cell populations. *Proc. Natl. Acad. Sci. USA* **107**, 8639–8643. <https://doi.org/10.1073/pnas.0911647107> (2010).
- Fisch, S. C. *et al.* Pluripotent nontumorigenic multilineage differentiating stress enduring cells (Muse cells): A seven-year retrospective. *Stem Cell Res. Ther.* **8**, 227. <https://doi.org/10.1186/s13287-017-0674-3> (2017).
- Simerman, A. A., Phan, J. D., Dumesic, D. A. & Chazenbalk, G. D. Muse cells: Nontumorigenic pluripotent stem cells present in adult tissues—a paradigm shift in tissue regeneration and evolution. *Stem Cells Int.* **2016**, 1463258. <https://doi.org/10.1155/2016/1463258> (2016).

26. Alessio, N. *et al.* The secretome of MUSE cells contains factors that may play a role in regulation of stemness, apoptosis and immunomodulation. *Cell Cycle* **16**, 33–44. <https://doi.org/10.1080/15384101.2016.1211215> (2017).
27. Park, Y. J., Niizuma, K., Mokin, M., Dezawa, M. & Borlongan, C. V. Cell-based therapy for stroke: Musing with muse cells. *Stroke* **51**, 2854–2862. <https://doi.org/10.1161/strokeaha.120.030618> (2020).
28. Yamashita, T., Kushida, Y., Abe, K. & Dezawa, M. Non-tumorigenic pluripotent reparative muse cells provide a new therapeutic approach for neurologic diseases. *Cells* <https://doi.org/10.3390/cells10040961> (2021).
29. Hwang, S. H. *et al.* Human inferior turbinate: An alternative tissue source of multipotent mesenchymal stromal cells. *Otolaryngol. Head Neck Surg.* **147**, 568–574. <https://doi.org/10.1177/0194599812447172> (2012).
30. Kwon, J. S. *et al.* In vivo osteogenic differentiation of human turbinate mesenchymal stem cells in an injectable in situ-forming hydrogel. *Biomaterials* **35**, 5337–5346. <https://doi.org/10.1016/j.biomaterials.2014.03.045> (2014).
31. Park, S. H. *et al.* Therapeutic potential of human nasal inferior turbinate-derived stem cells: Microarray analysis of multilineage differentiation. *ORL J. Otorhinolaryngol. Relat. Spec.* **84**, 153–166. <https://doi.org/10.1159/000516016> (2022).
32. Yun, B. G. *et al.* Accelerated bone regeneration via three-dimensional cell-printed constructs containing human nasal turbinate-derived stem cells as a clinically applicable therapy. *ACS Biomater. Sci. Eng.* **5**, 6171–6185. <https://doi.org/10.1021/acsbomaterials.9b01356> (2019).
33. Lim, J. Y. *et al.* Potential application of human neural crest-derived nasal turbinate stem cells for the treatment of neuropathology and impaired cognition in models of Alzheimer's disease. *Stem Cell Res. Ther.* **12**, 402. <https://doi.org/10.1186/s13287-021-02489-1> (2021).
34. Hwang, S. H. *et al.* Evaluation of characteristic of human turbinate derived mesenchymal stem cells cultured in the serum free media. *PLoS ONE* **12**, e0186249. <https://doi.org/10.1371/journal.pone.0186249> (2017).
35. Schürmann, M. *et al.* Identification of a novel high yielding source of multipotent adult human neural crest-derived stem cells. *Stem Cell Rev. Rep.* **14**, 277–285. <https://doi.org/10.1007/s12015-017-9797-2> (2018).
36. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).
37. He, K., Zhang, X., Ren, S. & Sun, J. Identity mappings in deep residual networks. In *European Conference on Computer Vision* 630–645 (Springer, 2016).
38. Huang, G., Liu, Z., Van Der Maaten, L. & Weinberger, K. Q. Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 4700–4708 (2017).
39. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 2818–2826 (2016).
40. Chollet, F. Xception: deep learning with depthwise separable convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* 1251–1258 (2017).
41. Lin, M., Chen, Q. & Yan, S. Network in network. *arXiv preprint arXiv:1312.4400* (2013).
42. Pan, S. J. & Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **22**, 1345–1359 (2009).
43. Raghun, M., Zhang, C., Kleinberg, J. & Bengio, S. Transfusion: understanding transfer learning for medical imaging. *arXiv preprint arXiv:1902.07208* (2019).
44. Deng, J. *et al.* Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* 248–255 (IEEE, 2009).
45. He, K., Zhang, X., Ren, S. & Sun, J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE International Conference on Computer Vision* 1026–1034 (2015).
46. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).
47. Goodfellow, I., Bengio, Y., Courville, A. & Bengio, Y. *Deep Learning* (MIT Press, 2016).
48. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **15**, 1929–1958 (2014).
49. Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & De Freitas, N. Taking the human out of the loop: A review of Bayesian optimization. *Proc. IEEE* **104**, 148–175 (2015).
50. Bergstra, J. & Bengio, Y. Random search for hyper-parameter optimization. *J. Mach. Learn. Res.* **13** (2012).
51. Rasmussen, C. E. Gaussian processes in machine learning. In *Summer School on Machine Learning* 63–71 (Springer, 2003).
52. Wilson, J. T., Hutter, F. & Deisenroth, M. P. Maximizing acquisition functions for Bayesian optimization. *arXiv preprint arXiv:1805.10196* (2018).
53. Rennert, R. C. *et al.* Microfluidic single-cell transcriptional analysis rationally identifies novel surface marker profiles to enhance cell-based therapies. *Nat. Commun.* **7**, 1–9 (2016).
54. Brett, E. *et al.* Isolation of CD248-expressing stromal vascular fraction for targeted improvement of wound healing. *Wound Repair Regen.* **25**, 414–422 (2017).
55. Yang, Z. X. *et al.* CD106 identifies a subpopulation of mesenchymal stem cells with unique immunomodulatory properties. *PLoS ONE* **8**, e59354 (2013).
56. Colter, D. C., Sekiya, I. & Prockop, D. J. Identification of a subpopulation of rapidly self-renewing and multipotential adult stem cells in colonies of human marrow stromal cells. *Proc. Natl. Acad. Sci.* **98**, 7841–7845 (2001).
57. Prockop, D., Sekiya, I. & Colter, D. Isolation and characterization of rapidly self-renewing stem cells from cultures of human marrow stromal cells. *Cytotherapy* **3**, 393–396 (2001).
58. Sekiya, I. *et al.* Expansion of human adult stem cells from bone marrow stroma: Conditions that maximize the yields of early progenitors and evaluate their quality. *Stem Cells* **20**, 530–541 (2002).
59. Van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9** (2008).
60. Chattopadhyay, A., Sarkar, A., Howlader, P. & Balasubramanian, V. N. Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks. In *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)* 839–847 (IEEE, 2018).
61. Bach, S. *et al.* On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS ONE* **10**, e0130140 (2015).
62. Bojarski, M. *et al.* Visualbackprop: Efficient visualization of cnns for autonomous driving. In *2018 IEEE International Conference on Robotics and Automation (ICRA)* 4701–4708 (IEEE, 2018).
63. Carpenter, A. E. *et al.* Cell Profiler: Image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* **7**, 1–11 (2006).
64. Imboden, S. *et al.* Investigating heterogeneities of live mesenchymal stromal cells using AI-based label-free imaging. *Sci. Rep.* **11**, 6728. <https://doi.org/10.1038/s41598-021-85905-z> (2021).
65. Lv, F.-J., Tuan, R. S., Cheung, K. M. & Leung, V. Y. Concise review: The surface markers and identity of human mesenchymal stem cells. *Stem Cells* **32**, 1408–1419 (2014).
66. Atha, D. H. *et al.* Standards for immunohistochemical imaging: A protein reference device for biomarker quantitation. *J. Histochem. Cytochem.* **58**, 1005–1014 (2010).

Acknowledgements

This research was supported by the National Research Foundation (NRF) (2020R1I1A1A01051844, 2020R1A6A1A03047902) and Ministry of Trade, Industry and Energy (MOTIE) (Development of Meta Soft Organ Module Manufacturing Technology without Immunity Rejection and Module Assembly Robot System, 20012378), Korea. The sponsors had no role in the study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Author contributions

Conceptualization, D.H.K., S.W.K., and S.L.; methodology, G.K., D.H.K., and S.L.; investigation, G.K., J.H.J., and K.P.; visualization, G.K. and J.H.J.; supervision, D.H.K., S.W.K., and S.L.; writing—original draft, G.K., D.H.K., and J.H.J.; writing—review and editing: G.K., D.H.K., J.H.J., K.P., S.W.K., and S.L.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-022-21653-y>.

Correspondence and requests for materials should be addressed to D.H.K. or S.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022