Received: 4 August 2023 | Revised: 10 January 2024 | Accepted: 19 February 2024

DOI: 10.1002/ima.23052

RESEARCH ARTICLE

WILEY

Improved senescent cell segmentation on bright-field microscopy images exploiting representation level contrastive learning

Fatma Çelebi[1,2] | **Dudu Boyvat**[3] | **Serife Ayaz-Guner**[4] | **Kasim Tasdemir**[5] | **Kutay Icoz**[1,2,6]

[1]BioMINDS (Bio Micro/Nano Devices and Sensors) Lab, Electrical and Electronics, Engineering Department, Abdullah Gül University, Kayseri, Turkey

[2]Computer Engineering Department, Abdullah Gül University, Kayseri, Turkey

[3]Moleculer Biology and Genetics Department, Abdullah Gül University, Kayseri, Turkey

[4]Department of Molecular Biology and Genetics, Izmir Institute of Technology, Izmir, Turkey

[5]Queen's University Belfast, UK

[6]Biomedical Engineering, University of Delaware, Newark, DE, USA

Correspondence
Fatma Çelebi and Kutay Icoz, BioMINDS (Bio Micro/Nano Devices and Sensors) Lab, Electrical and Electronics, Engineering Department, Abdullah Gül University, Kayseri, 38080, Turkey.
Email: fatma.celebi@agu.edu.tr and kutay.icoz@agu.edu.tr

Abstract

Mesenchymal stem cells (MSCs) are stromal cells which have multi-lineage differentiation and self-renewal potentials. Accurate estimation of total number of senescent cells in MSCs is crucial for clinical applications. Traditional manual cell counting using an optical bright-field microscope is time-consuming and needs an expert operator. In this study, the senescence cells were segmented and counted automatically by deep learning algorithms. However, well-performing deep learning algorithms require large numbers of labeled datasets. The manual labeling is time consuming and needs an expert. This makes deep learning-based automated counting process impractically expensive. To address this challenge, self-supervised learning based approach was implemented. The approach incorporates representation level contrastive learning component into the instance segmentation algorithm for efficient senescent cell segmentation with limited labeled data. Test results showed that the proposed model improves mean average precision and mean average recall of downstream segmentation task by 8.3% and 3.4% compared to original segmentation model.

KEYWORDS

cellular senescence, instance segmentation, mask R-CNN, microscopy images, self-supervised learning, SimCLR

1 | INTRODUCTION

Mesenchymal stem cells (MSCs) are adult stem cells that have a significant power of self-renewal and multi-differentiation. MSCs originate from a variety of sources, including cord blood, amniotic fluid, bone marrow, and adipose tissue; their extensive availability is a considerable trait for therapeutic and research purposes.[1] They are used in a variety of clinical applications, including regenerative medicine, immunological diseases, and cancer treatment.[2] MSCs exert their therapeutic effects largely through their differentiation capacity into multiple cell lineages and paracrine actions. MSCs gradually lose their potential therapeutic value with aging and cellular senescence. Cellular senescence is a complex state in which irreversible cell cycle arrest will limit their

ability to proliferate, as well as their ability to regenerate and homing capacity.[3,4] In addition, the senescence of the MSCs changes their secretions profile. Senescence-associated secretory phenotype (SASP) is a multicomponent phenotype released in large amounts by senescent MSCs that comprises pro-inflammatory chemicals, immune-modulators, and growth factors.[5–7] The secretions of these factors can cause adverse effects, such as damaging healthy tissues with inflammatory effects and inhibiting DNA repair.[8] Therefore, understanding the aging process of MSC is crucial for clinical applications.

Considering MSCs in clinical applications, it is crucial to detect and quantify senescent MSC in the population. One of the common techniques for the detection of cellular senescence is the beta-galactosidase assay (β-Gal), yet it has several drawbacks. Since assay outputs are evaluated visually, cell confluency, serum depletion, and operator bias can result in false positives, while some senescent cells may not show β-Gal activity, resulting in false negatives.[9] Besides, these cells must be manufactured and used in clinical settings in accordance with Good Manufacturing Practices (GMP). GMP requirements necessitate the organization of qualified staff, specialized and controlled equipment, and multiple quality controls.[10] The current protocols, such as β-Gal, to identify senescent cells are not easily adaptable to the GMP pipeline and are time-consuming considering the process. However, as previously stated, the appropriate use of MSCs for clinical applications demands the inspection of senescent cells prior to clinical use.

Senescent cells change their morphology, so image processing could read the visual cues emerging in the cell images to differentiate and count senescent cells. In the literature, image-based machine learning methods,[11,12] and conventional image processing techniques[13,14] were applied for various types of cell analysis. In addition, cell debris quantification to measure cellular apoptosis in bright-field microscopy images were achieved by using image processing techniques in Fiji open-source software.[15] However, the majority of those techniques depend on the use of extensive amount of labeled cell images. One prominent example of data hungry machine learning (ML) method is deep learning (DL) techniques which are promising for identification and quantification tasks. Kusumoto et al.[16] applied label-free senescent analysis for endothelial cells using DL. Since MSCs had unique cell shapes, there is no available image-based analysis method for the estimation of senescent cells.

Deep learning algorithms are widely applied for classification, segmentation, and object detection in medical images.[17] In Razzak et al.[18] usage of deep learning architectures was broadly reviewed. Hematological images were segmented into five leukocyte classes by SegNet, U-Net, and VGG-Net convolutional neural network architectures and results were compared.[19] Deep learning architectures were also implemented for cellular image analysis and details are extensively reviewed in Moen et al.[20] Microscopic cell images were segmented by using convolutional neural networks (CNN).[21–24] In microscopy images, cell abnormalities were segmented by classical LeNet architecture.[25] Two fully convolutional regression networks were developed and compared on fluorescent microscopy cell images for both detection and quantification in Xie et al.[26] Pre-trained single cell segmentation-based U-Net architectures were implemented on ImageJ software for non-machine-learning experts. The plugin allows users to analyze their data either on local computer or on cloud services.[27] Basic U-Net architecture which is mainly used to segment 2D biomedical images were modified to 3D to segment nuclei in fluorescence microscopy images.[28] Deep convolutional auto-encoders were proposed to count microscopic, stained cell images. However, DL techniques require a large number of well-labeled data, and the data labeling is time-consuming and requires an expert on that field. To overcome a small number of labeled dataset problems in DL, transfer learning techniques are commonly used. During the transfer learning, instead of training network from end-to-end to get weights, the weights of network are initialized by pre-trained network weights. There are some models that have pre-trained weights like Mask R-CNN, but those weights are obtained by training the models on natural images which is different than medical images. In addition, in Refs.,[29,30] it is shown that transfer learning from in-domain data improves performance compared to transfer learning from out-domain data. Since the original Mask R-CNN[31] is pre-trained on natural images, we aimed to use Mask R-CNN which is pre-trained on medical images. Nevertheless, even with transfer learning the number of required labeled samples is well high over the practical limits of the expert teams.

Therefore, to alleviate the problem of large number of labeled data requirement, we have resorted to self-supervised learning[32] on microscopic unlabeled cell images. As a result, this increased the mean average precision metric of Mask R-CNN based segmentation task. In this study, we showed that even if a small number of unlabeled images are available, self-supervised learning (SSL) on the in-domain dataset and fine-tuning for downstream labeled dataset training exceeds performance of the conventional instance segmentation algorithm. The main contributions of the proposed approach are concluded as follows.

1. Automated segmentation and counting of senescent cells: The study introduces a deep learning-based algorithm for the automated segmentation and counting of senescent cells in MSCs. This addresses the limitations of traditional manual cell counting methods, which are time-consuming and require expert operators.
2. Improvement by employing SSL: The proposed SSL-based approach demonstrates improved performance compared to the original segmentation model. The empirical test results show an increase in mean average precision (mAP) and mean average recall (mAR) of the downstream segmentation task by 8.3% and 3.4%, respectively.
3. Pre-training with unlabeled data: The proposed study demonstrates that leveraging a small number of unlabeled datasets during the SSL phase significantly enhances the performance of the downstream instance segmentation task on microscopy cell images.
4. Senescent cell quantification: The study demonstrates that SSL can be adopted to quantification task alongside the segmentation.

2 | RELATED WORK

Phase-contrast microscopy senescent cell images were identified by using morphology-based CNN architectures and then trained CNN architecture was used as a quantitative senescence score for the state of endothelial cells.[16] In Yan et al.,[34] to overcome large number of labeled dataset issue for DL techniques, unlabeled histopathology images were analyzed by deep learning based self-supervised contrastive learning method. Cellular senescence in cell culture was predicted by nuclear morphology based neural network.[35] Mask R-CNN is the conventional instance segmentation algorithm that is trained on natural images, and it is also applied to biomedical cell images. In Johnson,[36] Mask R-CNN was applied on microscopic images of cell nuclei. Instead of training the network end-to-end from the start, the weights are initialized with pre-trained MSCOCO[33] dataset weights. Average mask intersection over union (IoU) was reported as 70.54%. Pre-trained on natural images Mask R-CNN was used as instance segmentation in hematoxylin and eosin stained (H&E) microscopy cell nuclei images and average mask IoU was reported as 45.02% on the fivefold validation dataset.[37] Overlapping cells were segmented by Mask R-CNN based methods.[38] Infected and uninfected red blood-stained cell images segmentation was achieved by conventional Mask R-CNN by initializing weights from pre-trained model which is trained on the COCO dataset.[39]

3 | MATERIALS

For this study, 342 cell cultured images were used which was obtained by color camera and 10X objective (Leica DMI1, Germany). The size of the images was 2592 × 1944 pixels. Single image contains multiple young and senescence cells.

Proposed algorithms were developed with Python open-source libraries Keras 2.2.5 and Tensorflow 1.15.0.

3.1 | Mesenchymal stem cell culture

Adipose tissue-derived MSCs were obtained from the American Type Culture Collection (ATCC PCS-500-011). MSCs were grown in DMEM supplied with 10% FBS (Fetal Bovine Serum), 100 U/mL penicillin-streptomycin, 4 mM L-glutamine, and 5 ng/mL bFGF at 37°C in 5% CO_2 environment.

3.2 | Senescence induction

Senescence of the MSCs was induced by the addition of 300 μM H_2O_2 in phosphate buffered saline (PBS) for 30 min. Then, cells were washed three times with PBS and left in a complete growth medium.

3.3 | Microscopy

Visualization of the senescent and young stem cells is achieved after fixation of the cells. The cells were fixed for practical reasons to be able to collect large number of images, however, the method could be applied without fixation and staining of the cells. MSCs were washed with PBS three times. Then, cells were incubated in a 10 mL fixative solution (0.2% glutaraldehyde in PBS) for 15 min. Cells were washed with PBS to clean up the remaining glutaraldehyde. Fixed cells were stored with the addition of 50% glycerol.

4 | METHOD

4.1 | Instance segmentation

In this section, implementation of a conventional Computer Vision method segmenting the microscopic medical images is explained. The method's performance is further improved with the employment of SSL. It is explained in the next section.

Mask R-CNN is an instance segmentation network which was developed on top of Faster R-CNN. In an image, Mask R-CNN network classifies each pixel into a fixed set of categories (senescent cell, young cell, or background in our case), and at the same time the network precisely segments each instance. A general view of the proposed architecture and its output generated from our own dataset is given in Figure 1. It has two stages; the first stage consists of a backbone network (ResNet-101, VGG, Inception ...) for feature extraction and region proposal network. The second stage consists of returning confidence, bounding boxes, and binary masks based on region of interests. The first stage takes an input image and proposes bounding boxes of the candidate objects. During the second stage, a fully connected network classifies the proposed regions. The proposed region might be in an arbitrary size, but the fully connected classifier layer requires a fixed size image. To have the same size, the network uses RoIAlign methods. The output of RoIAlign layer goes to mask head layer, and the layer generates mask for each RoI with the predicted classes.[31]

In our approach, the bright-field microscopy images were trained then senescent and young cells were segmented. To be able to calculate density of each senescence and young cells number of objects is needed, instance segmentation Mask-RCNN algorithm was used rather than other semantic segmentation algorithms.

ResNet-101 and feature pyramid network (FPN) were used as the backbone. Matterport Inc. released under an MIT License implementation[40] which is based on Keras[41] and Tensorflow[42] libraries was used. Instead of training the network end-to-end from start, transfer learning approach was applied, and model weights were initialized by pre-trained MSCOCO dataset weights.

During the training, two approaches are implemented. The first approach is just training the network head layers, and the second approach is training all layers of the network. During the training data augmentation is applied to the training dataset. Dataset is expended with vertical and horizontal flipping, rotation, scaling, and adding blur augmentations.

4.2 | The proposed approach

Training deep learning networks on medical images needs vast amount of high-quality labeled data, and data labeling is a cumbersome process because it is time-consuming and needs an expert in medical field to label images correctly. Therefore, transfer learning is a popular remedy for the problem. With this approach, a model is pre-trained on labeled large natural image dataset (like ImageNet), and then learned generic representation is fine-tuned to medical dataset. Transfer learning from the same domain can improve the performance of the networks.[30,43] However, it still needs a large amount of labeled dataset which is a major problem in medical images requiring manual labeling by experts.

On natural images recognition tasks, recently another popular technique which is called self-supervised contrastive pre-training has drawn attention in the literature.[44] Despite the success of self-supervised contrastive learning on natural images,[32,45] application in senescent cell analysis is limited and needs more exploration. Self-supervised training technique is appealing because it allows pre-training with unlabeled domain-specific images to learn more pertinent features to that domain.[46]

Since we have a limited number of annotated dataset, on our unlabeled dataset we apply SSL approach by only

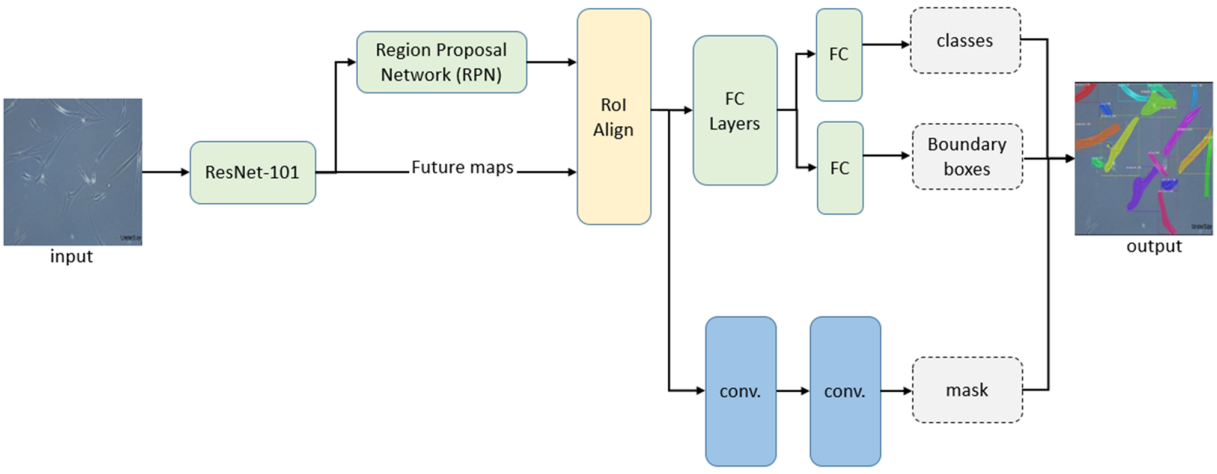

FIGURE 1 General view of Mask R-CNN architecture used on our dataset. Left to right; input image, architecture of network and output image. It takes an input image and returns output image with mask of each instance and their bounding boxes.

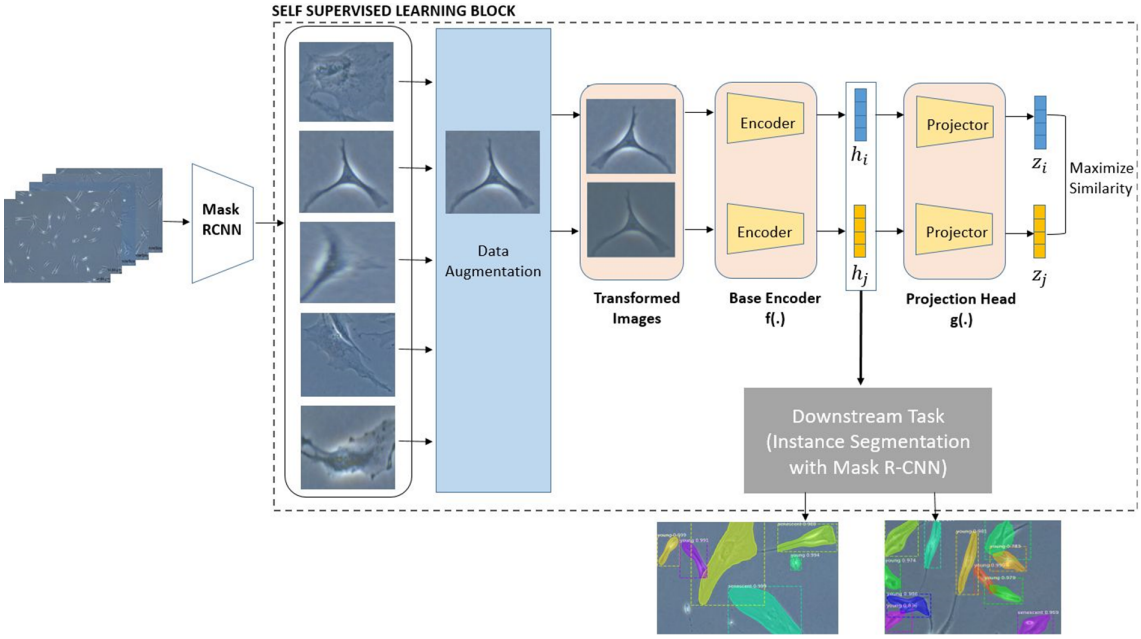

FIGURE 2 The proposed self-supervised instance segmentation on microscopic images workflow. It comprises of three steps. First, applying Mask R-CNN on unlabeled high-resolution images to get single object patches. Second, unsupervised learning to learn general representations of our microscopy image patches. Third, fine-tuning and training on labeled dataset, and then instance segmentation by using Mask R-CNN architectures.

training ResNet-101 network, which is the backbone network of Mask R-CNN.

Our self-supervised and instance segmentation workflow is given in Figure 2. Our pre-training includes two approaches: (1) training the network end-to-end from the start, (2) initializing the weights from ImageNet[47] and then training the network. Later, we used this pre-trained backbone network weights in Mask R-CNN architecture as initialization for downstream supervised training task on the labeled data. We wanted to increase the performance of Mask R-CNN by self-supervised pre-training on unlabeled medical images then fine-tuning on labeled data. We compared the results of both supervised pre-training on natural images and fine-tuning on labeled medical dataset, and self-supervised pre-training on in-domain medical images and fine-tuning on labeled medical dataset approaches.

As self-supervised pre-training there are various methods called SimCLR,[32] MoCo,[48] context restoration.[49] These methods are applied to medical images.[46,49,50] For our problem, we use SimCLR SSL algorithm which is a state-of-art method. It takes an image and by using basic augmentations (random cropping, random color distortion, and Gaussian Blur) it generates two transformed images. Method learns general representations by maximizing agreements between the same image transformations and minimizing agreements between different image transformations. We use sayakpaul/SimCLR-in-TensorFlow-2[51] implementations of SimCLR. As augmentation for SimCLR we use random flip, color distortion with strength 1.

In literature, it is stated that CNNs show high potential for identification task for biological data and cell images.[16] In Heckenbach et al.,[35] U-Net semantic segmentation architecture was used to segment nuclei of senescence cells. Also, in another study, breast cancer risk was associated with cellular senescence and senescence cells were segmented by basic U-Net architecture.[52] In addition, performance comparison of state-of-art cell nuclei segmentation was achieved by U-Net and DeepLabV3+ CNN architectures. In our study, to compare proposed methods with other segmentation techniques we implemented U-Net and DeepLabV3 CNN architectures.

5 | EXPERIMENTS AND RESULTS

5.1 | Dataset

We have a total of 342 microscopic images, each image has a size of 2592×1944. The number of images and sizes used for SSL and fine-tuning is given in Table 1. First, 14 high-resolution images were labeled by three medical experts. Then, the labeled images were trained with original Mask R-CNN by initializing weights from MSCOCO dataset. Then, unlabeled 242 high-resolution images were

TABLE 1 Number and properties of images used for self-supervised learning and fine-tuning.

	Total number of images	Image sizes	Number of objects	Labeled
Self-supervised learning Raw images	242	2592×1944	Multiple	No
Self-supervised learning Patches	8719	224×224	1	No
Fine-tuning	100	2592×1944	Multiple	Yes

used as test images. Since result of Mask R-CNN returns output image with mask of each instance and their bounding boxes, 242 test image results were used to extract unlabeled single cell images. After this step, total 8719 unlabeled single object images were obtained for self-supervised training. All 8719 images were resized to 224x224x3 to get the same size. Finally, to show performance of the proposed methods, unused 86 high-resolution images were also labeled by three medical experts for supervised training and all results were shown on these 100 labeled high-resolution images.

5.2 | Evaluation measures

To evaluate performance of Mask R-CNN on test dataset, mAP and mAR were calculated. Precision is the value of correctly predicted bounding boxes out of all bounding boxes found in image based on an IoU threshold, which is 0.5 for our case. While the mean of average precision across all images in the dataset is called mAP, average recall across all images in the dataset is called mAR. mAP is calculated as follows:

$$mAP = \frac{1}{N} \sum_{c \in N} \frac{TP_c}{TP_c + FP_c} \tag{1}$$

$$mAR = \frac{1}{N} \sum_{c \in N} \frac{TP_c}{TP_c + FN_c} \tag{2}$$

where N represents number of classes, TP represents true positives, FP represents false positives, FN represents false negatives, and classes represent senescent cells or young cells in our problem.

To measure quantification result of Mask R-CNN precision (PR), recall (RE) and dice similarity coefficient (DSC) metric scores were calculated, and they are described as follows:

$$PR = \frac{TP}{\mathrm{TP} + \mathrm{FP}} \tag{3}$$

$$RE = \frac{TP}{\mathrm{TP} + \mathrm{FN}} \tag{4}$$

$$DSC = \frac{2TP}{2\mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \tag{5}$$

In Equations (3)–(5), TP (true positive) represents the correctly quantified young or senescent cells, FP (false positive) represents the invalid quantified young or senescent cells, FN (false negative) represents missed quantified young or senescent cells.

To show performance of the instance segmentation algorithm, confusion matrix is used. TP is correctly predicted senescence cells, FP is number of incorrect predictions, FN is missed predictions, and true negative (TN) is correctly predicted young cells.

5.3 | Implementation details

5.3.1 | Mask R-CNN with transfer learning from MSCOCO dataset

In the first part of the proposed approach, Mask R-CNN weights were transferred from MSCOCO dataset that is pre-trained on labeled natural images, and Mask R-CNN was applied on our own labeled microscopic cell images. Labeled 100 high-resolution images were randomly divided into train, validation, and test dataset (50 images for training, 20 images for validation, and 30 images for testing). The labeled data were trained both with augmentation and without augmentation. Effects of data augmentation on our labeled dataset is given in Figure 3. It can be seen that the data augmentation technique dramatically affects the model's performance by decreasing validation loss.

While training, the initial parameters are set as follows: minimum confidence 0.9, steps per epoch 100, images per GPU 1, weight decay 0.0001, and learning rate 0.001. Maximum number of epoch was set to 100 and early stopping technique was used.[53]

After the training, the models were tested on the labeled test dataset. The resulted image has segmented objects masks, object labels with probability values, and bounding box coordinates for each object. The test images and their resulted images are given in Figure 4.

5.3.2 | Mask R-CNN with self-supervised learning and fine-tuning

In this part, implementation of self-supervised pre-training on unlabeled images is explained. For pre-training task, ResNet-101 architecture followed by a non-linear projection head were used. For pre-training two different scenarios were evaluated: (1) start training unlabeled images without initial weights, (2) start training unlabeled images with ImageNet weights. Self-supervised pre-training was completed with learning rate in {0.3, 0.1, 0.001, 0.0001}, batch size in {2, 4, 8, 16, 32, 64, 256}, and epoch in {50, 100, 150, 200, 250}. The best pre-training model was obtained with 256 batch size, 200 epoch, 0.1 learning rate, and 0.1 temp hyper-parameters. While pre-training, the normalized temperature-scaled cross entropy loss function (NT-Xent)[32] was used which computes the loss for positive pairs of two augmented images and contrasts with the distance to negative pair. Let sim $(\mathbf{u},\mathbf{v}) = \boldsymbol{u}^T\mathbf{v}/||\mathbf{u}||\ ||\mathbf{v}||$ denote the cosine similarity between two vectors **u** and **v,** and exp denote exponential function Then the loss function for a positive pair of examples $(\boldsymbol{i},\boldsymbol{j})$ formally expressed as:

$$l_{i,j} = -log\frac{\exp\left(sim\left(z_i,z_j\right)/\tau\right)}{\sum_{k=1}^{2N} 1_{[k\neq i]}\exp\left(sim\left(\left(z_i,z_j\right)/\tau\right)\right)} \quad (6)$$

where $1_{[k\neq i]} \in \{0,1\}$ is an indicator function evaluating to 1 if $k \neq i$, and τ is temperature parameter.

While training with larger batch sizes, LARS optimizer[54] was used to stabilize training. The best pre-training model was selected based on performance of network on the downstream fine-tuning and training task on labeled images.

After SSL on unlabeled data with ResNet-101 architecture, pre-trained ResNet-101 as backbone network without the nonlinear projection head was used for Mask R-CNN architecture in order to train and test labeled dataset. During the fine-tuning, the same labeled 100 high-resolution images were used. Images were

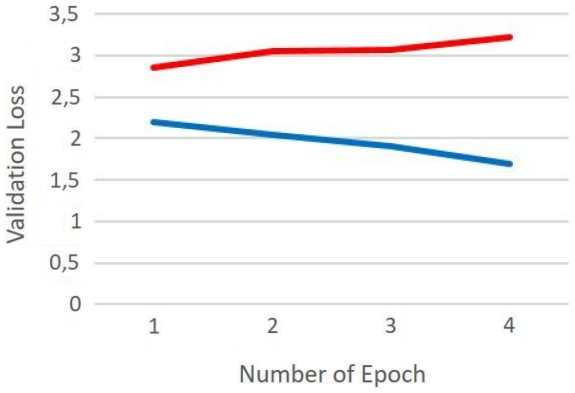

FIGURE 3 Effect of data augmentation on our dataset. Blue color indicates validation loss with augmentation, and red color indicates validation loss without augmentation.

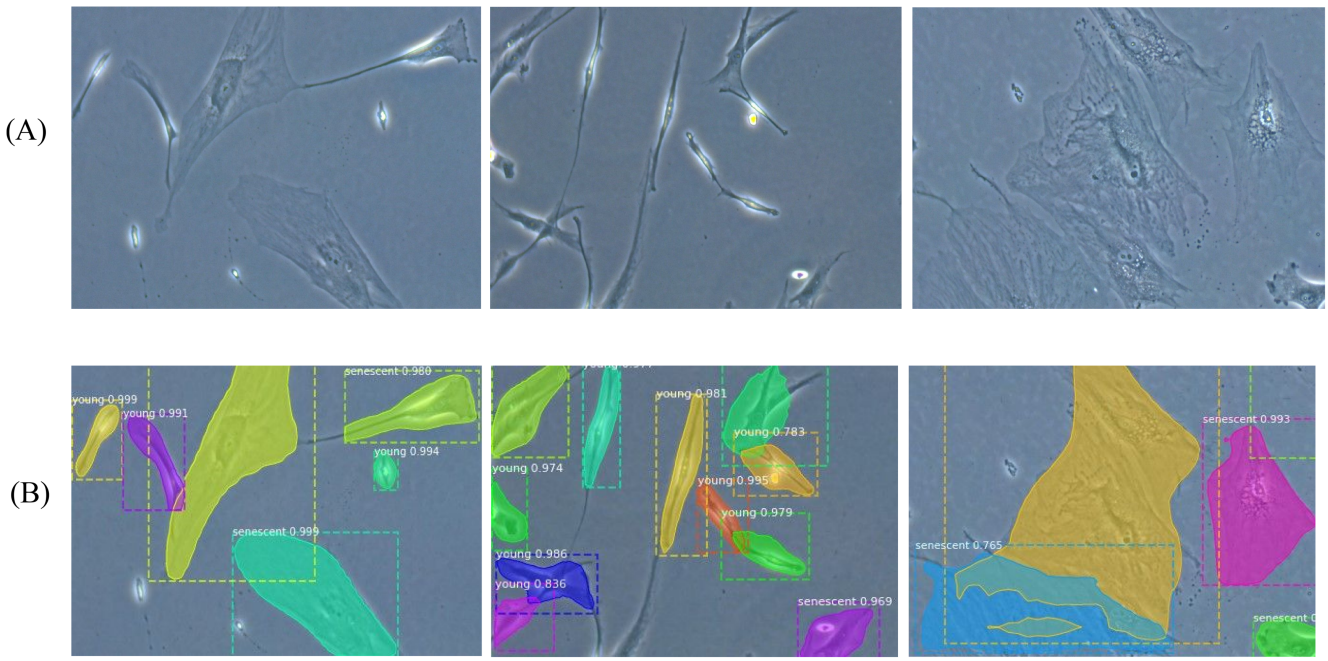

FIGURE 4 Test image samples and their Mask R-CNN without self-supervised learning segmentation results. Row (A) shows test images, and row (B) shows results of instance segmentation with masks, object labels, probability values, and bounding boxes.

TABLE 2 Evaluation results of Mask R-CNN on our microscopy test images with fine-tuning from MSCOCO dataset and self-supervised learning.

Backbone	Batch size of SSL	SSL from ImageNet weights	SSL end-to-end training from start	$mAP_{0.5}$	$mAR_{0.5}$	$mAP_{0.5-0.95}$	$mAR_{0.5-0.95}$
ResNet-101 + FPN				0.653	0.785	0.351	0.694
ResNet-101 + FPN	16	✓		0.683	0.790	0.341	0.697
ResNet-101 + FPN	32	✓		0.670	0.784	0.344	0.69
ResNet-101 + FPN	64	✓		0.670	0.792	0.331	0.648
ResNet-101 + FPN	256	✓		0.694	**0.821**	0.354	0.707
ResNet-101 + FPN	256		✓	**0.736**	0.819	**0.393**	**0.735**

Note: Maximum values for evaluation matrices were obtained by the proposed method and numbers are showed in bold.

randomly divided into training, validation, and test with number 50, 20, and 30 respectively. While training, initial parameters are set as follows: minimum confidence 0.9, steps per epoch 100, images per gpu 1, weight decay 0.0001, and learning rate 0.001. The maximum number of epoch was set to 800 and early stopping technique was used.

5.4 | Experimental results

Result of Mask R-CNN with transfer learning from MSCOCO dataset and transfer learning from SSL on in-domain dataset is given in Table 2. Mean average precision and mean average recall values were calculated at IoU threshold of 0.5 and IoU threshold of 0.5–0.95 range. It can be seen that in-domain data SSL improves performance of Mask R-CNN. The best result of fine-tuning was obtained by SSL with batch size 256. It can be concluded that SSL exploits larger batch sizes in spite of the small number of unlabeled dataset. Since larger batch sizes mean more diverse augmented version of an image, the network can learn more domain-specific features. The effect of SSL with different batch sizes and different approaches when fine-tuning with labeled dataset is given in Figure 5. The minimum loss was achieved on SSL with batch size 256 and without weight initialization while pre-training. In addition, even smaller batch size pre-training on unlabeled in-domain dataset gives better performance during fine-tuning compared to fine-tuning from MSCOCO dataset weights.

Besides, from Table 2, it can be concluded that self-supervised training end-to-end from start yields higher mAP metric for IoU 0.5 and IoU for range 0.5–0.95 than initializing self-supervised training weights from ImageNet. SSL end-to-end from start increased mAP of Mask R-CNN by %8.3 when IoU is 0.5, and %4.2 when IoU is in range 0.5–0.95. In addition, SSL end-to-end from start

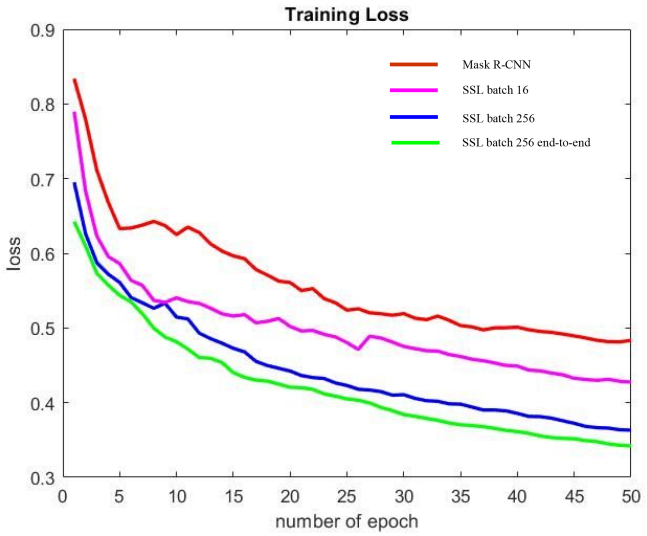

FIGURE 5 Training loss results on labeled dataset. Self-supervised learning (SSL) batch 16 indicates ImageNet initialized SSL with batch size 16 while pre-training, SSL batch 256 indicates ImageNet initialized SSL with batch size 256 while pre-training, SSL batch 256 end-to-end indicates SSL without any weights initialization in the proposed method. SSL improves instance segmentation performance.

increased mAR of Mask R-CNN by %3.4 when IoU is 0.5, and 4.1% when IoU is in the range of 0.5–0.95. To sum up, it can be seen from Table 2 that, both SSL approaches increased the evaluation metrics of original Mask R-CNN.

The self-supervised learning can learn finer details. In the images, there are young cells, senescent cells, and cells which just start aging and it is also taken as senescent cells. While original Mask R-CNN learn just started aging cells as young cell, SSL can learns those cells as senescent and so segment them correctly. Moreover, Mask R-CNN with SSL can detect more young cells correctly compared with original Mask R-CNN. In addition, SSL can find cell boundaries better when there are

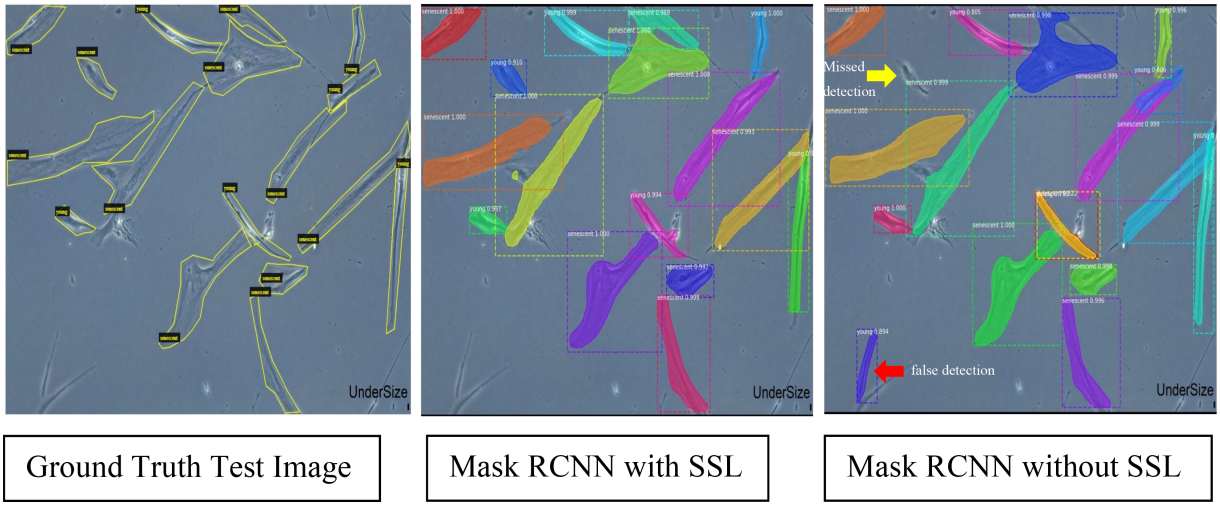

FIGURE 6 Test image result with self-supervised learning (SSL) and without SSL.

TABLE 3 Per-class performance of mask R-CNN and proposed method with IoU = 0.5.

	Mask RCNN		Mask RCNN + SSL	
Class	**mAP**	**mAR**	**mAP**	**mAR**
Senescent cell	0.742	0.832	0.819	0.928
Young cell	0.559	0.738	0.650	0.710

TABLE 4 Comparison of convolutional neural network segmentation models on our dataset.

Model	Backbone	$mAP_{0.5}$	$mAR_{0.5}$
DeepLabV3	ResNet-101	0.216	0.33
U-NET	—	0.648	0.508
Mask RCNN	ResNet-101	0.653	0.785
Mask RCNN + SSL	ResNet-101	**0.736**	**0.819**

Note: The proposed method yields the best results for mAP and mAR metrics and the results are shown in bold.

overlapping cells. Segmentation results on test image is shown in Figure 6. It can be concluded that SSL significantly improves Mask R-CNN segmentation result by detecting the cells more accurately and quantifying the number of cells more accurately.

Moreover, the per-class performance of the original model and the proposed model is shown in Table 3. While the proposed method increased mAP for both senescence and young cells, increased mAR for senescence cells and decreased mAR for young cells.

The study compares the segmentation results of Mask R-CNN, the proposed method, U-Net, and DeepLabV3 on test dataset, as shown in Table 4. The proposed method has the highest mean average precision and recall. DeepLabv3 does not perform well on our dataset. U-Net shows a comparable performance with original Mask R-CNN, but it falls short to segment young cells.

The total number of young and senescent cells in ground truth images, correctly found number of young and senescent cells in original Mask R-CNN, and the proposed method are given in Table 5. It can be seen from Table 5 that Mask R-CNN with SSL method finds more correct number of cells than original Mask R-CNN for both young and senescent cells.

To find TP, FP and FN, TN values, the quantitative results of instance segmentation algorithm were compared with manual counting of the experts. If an object in a high-resolution image is correctly classified by instance segmentation algorithm, it is counted as true positive. If it is not classified correctly, it is counted as a false negative. If an object is not found by algorithm, it is counted as missed detection (false negative). Confusion matric results are given in Tables 6 and 7 for the original and proposed models. Table 8 gives overall quantitative results of performance metrics for the proposed approach in microscopic cell images. It can be seen that Mask R-CNN with SSL is the highest performing method of all in all assessment metrics. Mask R-CNN with SSL increases mean precision, recall, dice similarity coefficient by 5.04%, 6.71%, and 5.98%, respectively.

6 | CONCLUSION AND DISCUSSION

The proposed study is aimed to develop an automated method to segment senescent and young cells in cell cultured microscopy images to estimate cellular senescence density while only limited number of labeled training

TABLE 5 Comparison of the methods by total number of correctly found cells.

	Ground truth		Prediction by Mask R-CNN		Prediction by SSL + Mask R-CNN	
	Senescent	Young	Senescent	Young	Senescent	Young
Total number of cells	334	339	277	239	304	262
Correct detection ratio (%)			83%	71%	91%	77%
False negative ratio (%)			17%	29%	9%	23%

TABLE 6 Confusion matrix results when Mask R-CNN + data augmentation is used.

		Ground truth	
		Senescence/pcs	Young/pcs
Predicted	Senescence	277 (TP)	56 (FP)
	Young	69 (FN)	239 (TN)

TABLE 7 Confusion matrice results when Mask R-CNN + data augmentation + SSL is used.

		Ground truth	
		Senescence/pcs	Young/pcs
Predicted	Senescence	304 (TP)	39 (FP)
	Young	47 (FN)	262 (TN)

samples are available. When the number of labeled dataset is limited, one approach is to use conventional image augmentation techniques.[55] To increase the number of labeled training dataset, image augmentation techniques, such as flipping, rotation, scaling, and adding blur, were applied during training. This approach increased the mAP performance of original Mask R-CNN from 0.585 to 0.653 on our test dataset. Therefore, we used data augmentation throughout the proposed work. Another approach is semi-supervised methods.[56–58] In Papandreou et al.,[57] it is stated that SSL performance is the highest when combining labeled data with large number of weakly annotated data. Since we do not have weakly annotated data, we did not use semi-supervised techniques. Another approach is to use transfer learning method which initializes network with pre-trained weights.[55]

When the labeled samples in training phase are scarce and the pre-training samples are from a different domain (natural images), the transfer learning methods fall short in microscopic image segmentation. This study showed that using transfer learning from the same domain thanks to incorporating SSL approach increased the performance of the model. Without cell staining, Mask R-CNN instance segmentation algorithm without SSL achieved 0.65 mAP and 0.78 mAR on test images with 0.5 IoU threshold. Mask R-CNN with SSL pre-training achieved 0.736 mAP and 0.819 mAR on test images with 0.5 IoU threshold, and 0.393 mAP and 0.735 mAR on test images with IoU threshold in range 0.5–0.95. These results show that benefiting from in-domain data SSL improved the performance of the instance segmentation algorithm. In addition, all performance metrics of quantification results are above 85% and for some images they are 100% as maximum. It means that the proposed method can guide experts about density of senescence cells in a given image.

Considering the promising and widespread use of MSCs, and the importance of detecting senescent cells in clinical practice, using the DL technique and automatically counting senescence cells, which can be easily adapted to GMP, is quite advantageous over time-consuming manual methods. A simple user interface can be developed to help experts to use the proposed study during senescence cell counting. Thanks to the assistance of interface, the expert can automatically segment senescent cells and calculate their total number by simply providing the microscopy image during their decisions about cellular senescence. The user interface will be designed to find and display senescence and young cells in the given image by running the DL model in the background.

We completed the first draft of the interface. Some screenshots of the interface are given in Supplementary Figure 1. To improve our model and user-interface it is needed to collect more dataset. Currently, we have collected data from a single research center. Later, we will need to collect data from different research centers. The developed method does not find all senescence and young cells 100% accurately, but it gives more than 85% precision and recall. In addition, the model and interface can determine whether there are more senescence cells or young cells in the given image.

Our dataset contains similar cell cultured microscopy images, as a result DL model makes an error and underperforms for other types of microscopy cell images. To overcome this problem, we will increase the diversity of the training dataset both in pre-training phase and

TABLE 8 Quantitative results of performance assessment procedures (mean, [min, max]) for proposed approach on microscopy cell images.

	Precision %	Recall %	Dice similarity coefficient %
Mask R-CNN	83.36 [64.70, 94.11]	79.96 [59.57, 91.66]	81.10 [63.63, 96.96]
Mask R-CNN + SSL	88.40 [73.33, 100]	86.67 [59.57, 100]	87.08 [73.91, 100]

during the training process. In this way, we will obtain a more general model and user interface that can detect cellular senescence in images of other cell types.

AUTHOR CONTRIBUTIONS

F.Ç., K.T., and K.I. developed the image analysis algorithms, D.B and S.A.G. prepared the cells and recorded the images. F.Ç. prepared the figures and tables. All authors reviewed the manuscript.

ACKNOWLEDGMENTS

The authors acknowledge Burak Acar and Servet Özcan for helping with cell cultures.

FUNDING INFORMATION

This research was not funded by any public or private entity.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Fatma Çelebi https://orcid.org/0000-0003-3157-6806
Kutay Icoz https://orcid.org/0000-0002-0947-6166

REFERENCES

1. Zhuang WZ, Lin YH, Su LJ, et al. Mesenchymal stem/stromal cell-based therapy: mechanism, systemic safety and biodistribution for precision clinical applications. *J Biomed Sci*. 2021;28(1):28. doi:10.1186/S12929-021-00725-7
2. Wei X, Yang X, Han ZP, Qu FF, Shao L, Shi YF. Mesenchymal stem cells: a new trend for cell therapy. *Acta Pharmacol Sin*. 2013;34(6):747-754. doi:10.1038/aps.2013.50
3. Alves-Paiva RM, Nascimento S, de Oliveira D, et al. Senescence state in mesenchymal stem cells at low passages: implications in clinical use. *Front Cell Dev Biol*. 2022;10:858996. doi:10.3389/fcell.2022.858996
4. Liu J, Ding Y, Liu Z, Liang X. Senescence in mesenchymal stem cells: functional alterations, molecular mechanisms, and rejuvenation strategies. *Front Cell Dev Biol*. 2020;8:258. doi:10.3389/FCELL.2020.00258/BIBTEX
5. Wiley CD, Campisi J. The metabolic roots of senescence: mechanisms and opportunities for intervention. *Nat Metab*. 2021;3(10):1290-1301. doi:10.1038/s42255-021-00483-8
6. Basisty N, Kale A, Jeon OH, et al. A proteomic atlas of senescence-associated secretomes for aging biomarker development. *PLoS Biol*. 2020;18(1):e3000599. doi:10.1371/JOURNAL.PBIO.3000599
7. Özcan S, Alessio N, Acar MB, et al. Unbiased analysis of senescence associated secretory phenotype (SASP) to identify common components following different genotoxic stresses. *Aging (Albany NY)*. 2016;8(7):1316-1329. doi:10.18632/AGING.100971
8. Galderisi U, Helmbold H, Squillaro T, et al. In vitro senescence of rat mesenchymal stem cells is accompanied by downregulation of stemness-related and DNA damage repair genes. *Stem Cells Dev*. 2009;18(7):1033-1042.
9. Pötzel T, Stoyanov J. Autofluorescence is a reliable in vitro marker of cellular senescence in human mesenchymal stromal cells. *Sci Rep*. 2019;9(1):1-15. doi:10.1038/s41598-019-38546-2
10. Lechanteur C, Briquet A, Bettonville V, Baudoux E, Beguin Y. Msc manufacturing for academic clinical trials: from a clinical-grade to a full gmp-compliant process. *Cells*. 2021;10(6):1320. doi:10.3390/cells10061320
11. Mualla F, Scholl S, Sommerfeldt B, Maier A, Hornegger J. Automatic cell detection in bright-field microscope images using sift, random forests, and hierarchical clustering. *IEEE Trans Med Imaging*. 2013;32(12):2274-2286. doi:10.1109/TMI.2013.2280380
12. Long X, Cleveland L, Lawrence Yao Y. Automatic detection of unstained viable cells in bright field images using a support vector machine with an improved training procedure. *Comput Biol Med*. 2006;36(4):339-362. doi:10.1016/j.compbiomed.2004.12.002
13. Uslu F, Icoz K, Tasdemir K, Doğan RS, Yilmaz B. Image-analysis based readout method for biochip: automated quantification of immunomagnetic beads, micropads and patient leukemia cell. *Micron*. 2020;133(January):102863. doi:10.1016/j.micron.2020.102863
14. Uslu F, Icoz K, Tasdemir K, Yilmaz B. Automated quantification of immunomagnetic beads and leukemia cells from optical microscope images. *Biomed Signal Process Control*. 2019;49:473-482. doi:10.1016/j.bspc.2019.01.002
15. Ölander M, Handin N, Artursson P. Image-based quantification of cell debris as a measure of apoptosis. *Anal Chem*. 2019;91(9):5548-5552. doi:10.1021/ACS.ANALCHEM.9B01243/SUPPL_FILE/AC9B01243_SI_002.ZIP
16. Kusumoto D, Seki T, Sawada H, et al. Anti-senescent drug screening by deep learning-based morphology senescence scoring. *Nat Commun*. 2021;12(1):257. doi:10.1038/s41467-020-20213-0
17. Hammad Saleem M, Khanchi S, Potgieter J, and Mahmood Arif K, "Image-based plant disease identification by deep

learning meta-architectures," *Plants* 2020, vol. 9, no. 11, p. 1451, doi: 10.3390/PLANTS9111451
18. Razzak MI, Naz S, Zaib A. Deep learning for medical image processing: overview, challenges and the future. *Lect Notes Comput Vis Biomech*. 2018;26:323-350. doi:10.1007/978-3-319-65981-7_12
19. Kadry S, Rajinikanth V, Taniar D, Damaševičius R, Valencia XPB. Automated segmentation of leukocyte from hematological images—a study using various CNN schemes. *J Supercomput*. 2022;78(5):6974-6994. doi:10.1007/s11227-021-04125-4
20. Moen E, Bannon D, Kudo T, Graf W, Covert M, Van Valen D. Deep learning for cellular image analysis. *Nat Methods*. 2019;16(12):1233-1246. doi:10.1038/s41592-019-0403-1
21. Gómez-de-Mariscal E, Maška M, Kotrbová A, Pospíchalová V, Matula P, Muñoz-Barrutia A. Deep-learning-based segmentation of small extracellular vesicles in transmission electron microscopy images. *Sci Rep*. 2019;9(1):1-10. doi:10.1038/s41598-019-49431-3
22. Al-Kofahi Y, Zaltsman A, Graves R, Marshall W, Rusu M. A deep learning-based algorithm for 2-D cell segmentation in microscopy images. *BMC Bioinformatics*. 2018;19(1):1-11. doi:10.1186/s12859-018-2375-z
23. Wang W, Taft DA, Chen YJ, et al. Learn to segment single cells with deep distance estimator and deep cell detector. *Comput Biol Med*. 2019;108(April):133-141. doi:10.1016/j.compbiomed.2019.04.006
24. Sadanandan SK, Ranefall P, Le Guyader S, Wählby C. Automated training of deep convolutional neural networks for cell segmentation. *Sci Rep*. 2017;7(1):1-7. doi:10.1038/s41598-017-07599-6
25. Araújo FHD, Silva RRV, Ushizima DM, et al. Deep learning for cell image segmentation and ranking. *Comput Med Imaging Graph*. 2019;72:13-21. doi:10.1016/j.compmedimag.2019.01.003
26. Xie W, Noble JA, Zisserman A. Microscopy cell counting and detection with fully convolutional regression networks. *Comput Methods Biomech Biomed Eng Imaging Vis*. 2018;6(3):283-292. doi:10.1080/21681163.2016.1149104
27. Falk T, Mai D, Bensch R, et al. U-Net: deep learning for cell counting, detection, and morphometry. *Nat Methods*. 2019;16(1):67-70. doi:10.1038/s41592-018-0261-2
28. Wu L, Chen A, Salama P, Winfree S, Dunn KW, Delp EJ. NISNet3D: three-dimensional nuclear synthesis and instance segmentation for fluorescence microscopy images. *Sci Rep*. 2023;13(1):1-19. doi:10.1038/s41598-023-36243-9
29. Heker M and Greenspan H. Joint liver lesion segmentation and classification via transfer learning, no. 1, pp. 1–5, 2020, [Online]. http://arxiv.org/abs/2004.12352
30. Liang G, Zheng L. A transfer learning method with deep residual network for pediatric pneumonia diagnosis. *Comput Methods Programs Biomed*. 2020;187:104964. doi:10.1016/j.cmpb.2019.06.023
31. He K, Gkioxari G, Dollár P, Girshick R. Mask R-CNN. *IEEE Trans Pattern Anal Mach Intell*. 2020;42(2):386-397. doi:10.1109/TPAMI.2018.2844175
32. Chen T, Kornblith S, Norouzi M, Hinton G. A simple framework for contrastive learning of visual representations. *37th Int. Conf. Mach. Learn. ICML 2020, vol. PartF16814, no. Figure 1*, PMLR; 2020:1575-1585.
33. Lin TY, Maire M, Belongie S, et al. "Microsoft COCO: common objects in context," Lect Notes Comput Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 8693 LNCS, no. PART 5, pp. 740–755, 2014, doi: 10.1007/978-3-319-10602-1_48
34. Yan J, Chen H, Li X, Yao J. Deep contrastive learning based tissue clustering for annotation-free histopathology image analysis. *Comput Med Imaging Graph*. 2022;97(December 2021):102053. doi:10.1016/j.compmedimag.2022.102053
35. Heckenbach I, Mkrtchyan GV, Ezra MB, et al. Nuclear morphology is a deep learning biomarker of cellular senescence. *Nat Aging*. 2022;2(8):742-755. doi:10.1038/s43587-022-00263-3
36. Johnson JW. Adapting mask-RCNN for automatic nucleus segmentation, pp. 1–7, 2018, doi: 10.1007/978-3-030-17798-0
37. Zhou S, Wang Q, Kumar N. Mask-RCNN for cell instance segmentation mask-RCNN for cell instance segmentation," no. April. 2020.
38. Chen J, Zhang B. Segmentation of overlapping cervical cells with mask region convolutional neural network. *Comput Math Methods Med*. 2021;2021:1-10. doi:10.1155/2021/3890988
39. Loh DR, Yong WX, Yapeter J, Subburaj K, Chandramohanadas R. A deep learning approach to the screening of malaria infection: automated and rapid cell counting, object detection and instance segmentation using mask R-CNN. *Comput Med Imaging Graph*. 2021;88(July 2020):101845. doi:10.1016/j.compmedimag.2020.101845
40. GitHub - matterport/Mask_RCNN: Mask R-CNN for object detection and instance segmentation on Keras and TensorFlow. https://github.com/matterport/Mask_RCNN (accessed Dec 09, 2021)
41. Keras: the Python deep learning API. https://keras.io/ (accessed Dec 09, 2021)
42. Abadi M, Agarwal A, Barham P, et al. TensorFlow: large-scale machine learning on heterogeneous distributed systems. 2016 [Online]. http://arxiv.org/abs/1603.04467
43. Chen S, Ma K, Zheng Y. Med3D: transfer learning for 3D medical image analysis," pp. 1–12, 2019, [Online]. http://arxiv.org/abs/1904.00625
44. Khosla P, Teterwak P, Wang C, et al. Supervised contrastive learning. *Adv Neural Inf Process Syst*. 2020;2020-Decem(NeurIPS):1-23.
45. He K, Fan H, Wu Y, Xie S, Girshick R. Momentum contrast for unsupervised visual representation learning. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recogn*. 2020;9726-9735. doi:10.1109/CVPR42600.2020.00975
46. Azizi S, Mustafa B, Ryan F, et al. Big self-supervised models advance medical image classification, no. 1. 2021 [Online]. http://arxiv.org/abs/2101.05224
47. Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. *ImageNet: a Large-Scale Hierarchical Image Database. IEEE*; 2009:248-255. doi:10.1109/cvprw.2009.5206848
48. He K, Fan H, Wu Y, Xie S, Girshick R. Momentum contrast for unsupervised visual representation learning. [Online]. https://github.com/facebookresearch/moco
49. Chen L, Bentley P, Mori K, Misawa K, Fujiwara M, Rueckert D. Self-supervised learning for medical image analysis using image context restoration. *Med Image Anal*. 2019;58:101539. doi:10.1016/j.media.2019.101539

50. Sowrirajan H, Yang J, Ng AY, and Rajpurkar P. MoCo-CXR: MoCo pretraining improves representation and transferability of chest X-ray models, pp. 1–17, 2020, [Online]. http://arxiv.org/abs/2010.05352
51. GitHub—sayakpaul/SimCLR-in-TensorFlow-2. (Minimally) implements SimCLR (https://arxiv.org/abs/2002.05709) in TensorFlow 2. https://github.com/sayakpaul/SimCLR-in-TensorFlow-2 (accessed Apr 04, 2022)
52. Heckenbach I, Powell M, Fuller S, et al. Breast cancer risk based on a deep learning predictor of senescent cells in normal tissue. *medRxiv*, p. 2023.05.22.23290327. 2023. doi:10.1101/2023.05.22.23290327v1.abstract
53. Liu L, Wu FX, Wang YP, Wang J. Multi-receptive-field CNN for semantic segmentation of medical images. *IEEE J Biomed Health Informatics*. 2020;24(11):3215-3225. doi:10.1109/JBHI.2020.3016306
54. You Y, Gitman I, Ginsburg B. Large batch training of convolutional networks. 2017 (accessed Apr 04, 2022). [Online]. https://github.com/BVLC/caffe/tree/master/models/bvlc_alexnet
55. Majurski M, Manescu P, Padi S, et al. Cell image segmentation using generative adversarial networks, transfer learning, and augmentations. *IEEE Comput Soc Conf Comput Vis Pattern Recogn Work*. 2019;2019-June:1114-1122. doi:10.1109/CVPRW.2019.00145
56. Khoreva A, Benenson R, Hosang J, Hein M, Schiele B. Simple does it: Weakly supervised instance and semantic segmentation. *Proc. - 30th IEEE Conf. Comput. Vis. Pattern Recognition, CVPR 2017, vol. 2017-January*. IEEE; 2017:1665-1674. doi:10.1109/CVPR.2017.181
57. Papandreou G, Chen LC, Murphy KP, Yuille AL. Weakly-and semi-supervised learning of a deep convolutional network for semantic image segmentation. *Proc. IEEE Int. Conf. Comput. Vis., vol. 2015 International Conference on Computer Vision, ICCV 2015*. IEEE; 2015:1742-1750. doi:10.1109/ICCV.2015.203
58. Xu K, Su H, Zhu J, Guan JS, Zhang B. Neuron Segmentation Based on CNN with Semi-Supervised Regularization. *IEEE Comput Soc Conf Comput Vis Pattern Recognit Work*. 2016; 1324-1332. doi:10.1109/CVPRW.2016.167

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Çelebi F, Boyvat D, Ayaz-Guner S, Tasdemir K, Icoz K. Improved senescent cell segmentation on bright-field microscopy images exploiting representation level contrastive learning. *Int J Imaging Syst Technol*. 2024;34(2):e23052. doi:10.1002/ima.23052