QUANTITATIVE IMAGE-BASED ANALYSIS OF MULTIDIMENSIONAL MESENCHYMAL STROMAL CELL CULTURES VIA COMPUTER VISION AND MACHINE LEARNING

A Dissertation

by

SAKINA MOHAMMED MOTA

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Kristen C. Maitland
Committee Members,	Carl A. Gregory
	Roland Kaunas
	Wonmuk Hwang
Head of Department,	Michael J. McShane

May 2022

Major Subject: Biomedical Engineering

ABSTRACT

Chronic disorders are leading causes of disability and death, and their high pervasiveness calls for effective treatments leading to a long-term cure. Studies show that stem cell-based therapies can provide successful solutions to chronic diseases. Mesenchymal stromal cells (MSCs), a multipotent group of stem cells, are extensively used for cell-based therapies owing to their immunomodulation, ex vivo proliferation, and clinically significant effects. MSC-based solutions require manufacturing large volumes of viable cells that is highly dependent on reliable methods to characterize cell mechanisms. Previous findings have revealed that cell morphology can serve as a critical quality attribute to predict the therapeutic potency of MSCs. Current standards to estimate MSC effectiveness based on their morphological phenotype are subjective, destructive, or time-consuming. Therefore, an objective method for morphological screening is needed to analyze the viability of live MSC cultures non-invasively. Computer-aided image analysis is an excellent tool to extract relevant cellular features rapidly and efficiently. This dissertation aims to facilitate large-scale cell growth strategies by providing automated technologies for evaluating images of cells grown in monolayer and three-dimensional environments.

We developed an analysis scheme based on conventional image processing techniques combined with machine learning to examine monolayer cultured MSCs. This methodology proved the applicability of image analysis as a robust tool for quantifiable culture monitoring. To address the challenges of standard image analysis approaches, we explored deep learning to better identify individual cells. It showed improved results for cell localization and also optimized overall MSC assessment. Lastly, we extrapolated deep learning-based analysis to study three-dimensional cultures used to produce MSCs in commercially viable quantities. This work shows great promise to fulfill the unmet need for cytomorphological analysis and downstream representation of MSCs adhered to spherical microcarriers. The algorithms were validated using visual inspection by a biologist with 15+ years of experience with MSCs. Thus, our research exhibits valuable potential to quantitatively analyze MSC efficacy and functionality, enabling the advancement of cytotherapies.

ACKNOWLEDGMENTS

First and foremost, I am very thankful to my committee chair, Dr. Kristen Maitland, for her invaluable advice, endless support, and patience during my Ph.D. study. Her immense knowledge and experience have always encouraged me in my academic research and daily life. She has always been there for me and gone above and beyond to help me grow professionally and personally. For this, I feel very privileged and extend my gratitude. I would also like to thank our collaborators and my committee members Dr. Carl Gregory and Dr. Roland Kaunas, for their guidance throughout this research. Their feedback and willingness to brainstorm research ideas were very influential in shaping my experimental methods and pushed me to further my technical expertise. I also appreciate my other committee member, Dr. Wonmuk Hwang, for his help and valuable time. I wish to express my special gratitude to Dr. Maryellen Giger for her mentorship and willingness to make time for me over the last four years. She was instrumental in providing me with the tools needed to complete my dissertation and build a strong foundation in my research area.

I am truly honored to have been introduced to Mr. and Mrs. Worsham and especially thank them for awarding me the Hagler fellowship. I wish to grow up to be like them and give back to society in the best way possible. I would also like to acknowledge my colleagues Robert Rogers, Andrew Haskell, and Berkley White for aiding my experiments and teaching me topics related to this research. A special shout out to my lab mates Cyril Soliman, Blanche Ter Hofstede, Oscar Benavides, Latifah Maasarani, and Elizabeth Bullard for their help with my writing and mainly many memorable discussions that broadened my understanding of optics, biomedical sciences, and various social issues. I am thankful to all the student organizations that I had the opportunity to serve. These organizations not only honed my interpersonal skills but gave me a wonderful group of friends. It has been my pleasure to know leading women scientists who set strong examples through mentoring and professional pursuits. Here, I have to mention Dr. Kristen Maitland, Dr. Maryellen Giger, and Dr. Madeleine Durkee. I would also like to thank the faculty, academic advisors, and staff of the Biomedical Engineering department for always assisting me and making

me a part of a big research-driven community.

I would also like to express my gratitude to my father and mother, Mohammed and Munira Mota, for their wise counsel, a sympathetic ear, and unconditional love. I would not have accomplished this without their encouragement and backing at every stage of my life. Whatever I am today, I owe it to them. I would also like to thank all my friends, especially Meet Sanghvi, for constantly cheering me, motivating me, and being a fun distraction to rest my mind outside research. Lastly, I would like to acknowledge my pillar of strength, my partner Nimish Ojale. Thank you for pushing me to become a better version of myself daily. You have made my doctoral journey incredible.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

The work in this dissertation was supported by a committee consisting of Dr. Kristen C. Maitland, Dr. Wonmuk Hwang, and Dr. Roland Kaunas from the Department of Biomedical Engineering at the Texas A&M University and Dr. Carl A. Gregory from the Department of Molecular and Cellular Medicine at the Texas A&M Health Science Center. Additionally, the dissertation was also supported by Dr. Maryellen L. Giger from the Department of Radiology at the University of Chicago.

Funding sources

Funding for the work related to this dissertation was provided by the President's Excellence Fund X-Grant. Funding was also provided in part by the Texas A&M University Hagler Institute for Advanced Study and the Graduate HEEP Fellowship by Mr. and Mrs. Bradley L. "Brad" Worsham.

NOMENCLATURE

α-MEM	α-Minimum Essential Media
2D	Two-dimensional
2U-Net	Double U-Net
3D	Three-dimensional
AF	Area Fraction
APS	Average Phenotype Score
AUC	Area under the ROC Curve
BF	Brightfield
CCM	Complete Culture Medium
CMFDA	5-Chloromethylfluorescein Diacetate
CQA	Critical Quality Attributes
CTG	CellTracker Green
EDTA	Ethylenediaminetetraacetic Acid
FBS	Fetal Bovine Serum
FN	False Negatives
FP	False Positives
FPF	False Positive Fraction
gelMA	Gelatin Methacryloyl
KNN	K-nearest Neighbors
LDA	Linear Discriminant Analysis
LR	Logistic Regression
LSFM	Light-sheet Fluorescence Microscopy

LSVM	Linear Kernel Support Vector Machine
MIP	Maximum Intensity Projection
MSC	Mesenchymal Stromal Cell
P	Precision
PBS	Phosphate-buffered Saline
PC	Pixel-wise Confluency
PFA	Paraformaldehyde
RMSE	Root Mean Square Error
ROC	Receiver Operating Characteristic
RoI	Regions of Interest
RS	Rapidly Self-renewing
RSVM	Radial-basis Kernel Support Vector Machine
RWV	Rotating Wall Vessel
S	Sensitivity
s.d.	Standard Deviation
s.e.	Standard Error
SNR	Signal-to-noise Ratio
SR	Slowly Replicating
TP	True Positives
TPF	True Positive Fraction

TABLE OF CONTENTS

Page

ABSTRACT ii

ACKNOWLEDGMENTS iii

CONTRIBUTORS AND FUNDING SOURCES v

NOMENCLATURE vi

TABLE OF CONTENTS viii

LIST OF FIGURES xi

LIST OF TABLES xv

1. INTRODUCTION AND MOTIVATION 1

    1.1 Chronic diseases and conditions 1
    1.2 Mesenchymal stromal cell-based therapies to address chronic diseases 1
    1.3 Current MSC evaluation strategies and their shortcomings 3
    1.4 Exploiting computer vision to quantitatively assess MSCs 4
    1.5 Proposed algorithms to analyze monolayer and three-dimensional MSC cultures 4

2. TECHNICAL BACKGROUND 7

    2.1 Cytotherapy: a brief overview 7
    2.2 Mesenchymal stromal cells: current understanding, therapeutic potential, and clinical challenges 8
    2.3 Culture protocol, imaging, and dataset for MSCs grown on monolayer 10
        2.3.1 Protocol for monolayer MSC cultures 10
        2.3.2 Imaging of monolayer MSC cultures 11
        2.3.3 Monolayer MSC culture dataset 12
    2.4 Culture protocol, imaging, and dataset for MSCs grown on microcarriers 14
        2.4.1 Protocol for three-dimensional MSC cultures 14
        2.4.2 Imaging of three-dimensional MSC cultures 14
        2.4.3 Three-dimensional MSC culture dataset 17
    2.5 Quantitative assessment of monolayer and 3D MSC cultures using image-based analysis 19
        2.5.1 MSC segmentation from their monolayer culture images using conventional processing techniques 19

2.5.2 MSC segmentation from their monolayer and 3D culture images using deep learning 21
2.5.3 Extraction of human-engineered features for monolayer and 3D cultured MSCs 23
2.5.4 Classification of monolayer cultured MSCs based on morphological phenotype 24

3. PRELIMINARY STUDY OF MORPHOLOGICAL CELL IMAGE ANALYSIS FOR MONITORING MONOLAYER CULTURED MESENCHYMAL STROMAL CELLS ... 25

3.1 Introduction 25
3.2 Methods 25
3.2.1 Image pre-processing and cell segmentation 25
3.2.2 Feature extraction 28
3.2.3 Classification of mesenchymal stromal cells based on their morphological phenotype 28
3.3 Results 29
3.4 Discussion 31

4. AUTOMATED MESENCHYMAL STEM CELL SEGMENTATION AND MACHINE LEARNING-BASED PHENOTYPE CLASSIFICATION USING MORPHOMETRIC AND TEXTURAL ANALYSIS 33

4.1 Introduction 33
4.2 Methods 34
4.2.1 Overview 34
4.2.2 Image pre-processing and area fraction estimation 34
4.2.3 Candidate cell region detection 37
4.2.4 Candidate marker detection 39
4.2.5 Cell segmentation and validation 41
4.2.6 Feature extraction 43
4.2.7 Cell classification and validation 44
4.3 Results 45
4.3.1 Cell detection and segmentation 45
4.3.2 Cell phenotype classification 48
4.3.3 Generalizability of phenotype classification 51
4.4 Discussion 54

5. U-NET-BASED IMAGE SEGMENTATION OF MESENCHYMAL STROMAL CELLS . 58

5.1 Introduction 58
5.2 Methods 58
5.2.1 Pre-processing and augmentation 59
5.2.2 U-Net model training 60
5.2.3 Post-processing and performance evaluation 60
5.3 Results 61

5.4 Discussion 65

6. 2U-NET: DOUBLE U-NET SEGMENTATION TO OPTIMIZE MACHINE LEARNING-DRIVEN MORPHOLOGICAL PROFILING OF MESENCHYMAL STROMAL CELLS 67

6.1 Introduction 67
6.2 Methods 68
6.2.1 Pre-processing, annotation, and augmentation 69
6.2.2 Training procedure 69
6.2.3 Post-processing and cell segmentation 71
6.2.4 Feature extraction and phenotype classification 72
6.2.5 Performance evaluation 73
6.3 Results 75
6.3.1 Detection and segmentation of MSCs using 2U-Net 75
6.3.2 Classification of MSC phenotype using machine learning 77
6.3.3 Generalization of MSC phenotype prediction and assessment of culture attributes 80
6.4 Discussion 83

7. QUANTITATIVE ASSESSMENT OF THREE-DIMENSIONAL MESENCHYMAL STROMAL CELL CULTURE ATTRIBUTES USING A COMBINATION OF DEEP LEARNING AND CONVENTIONAL IMAGE ANALYSIS 86

7.1 Introduction 86
7.2 Methods 88
7.2.1 Overview 88
7.2.2 Implementation of 2D U-Net and 3D U-Net deep learning models 89
7.2.3 Segmentation of cell regions from CellTracker Green channel image volumes 91
7.2.4 Segmentation of cell nuclei from DRAQ5 channel image volumes 92
7.2.5 Computation of human-engineered features and culture attributes at the single-cell level 93
7.3 Results 97
7.3.1 Cell region and nucleus segmentation from image volumes of microcarrier-based 3D MSC cultures 97
7.3.2 Cell count and confluency estimation for microcarrier-based 3D MSC cultures 101
7.4 Discussion 103

8. CONCLUSION AND FUTURE SCOPE 106

REFERENCES 108

APPENDIX A. 2U-NET TRAINING PROCEDURE 133

A.1 Determination of loss function for U-Net 1 133
A.2 Determination of annotation method and loss function for U-Net 2 134

LIST OF FIGURES

FIGURE Page

2.1 Example images showing distinct morphology of RS and SR MSCs. 10

2.2 Basic schematic of a phase-contrast microscope adapted from [1]. 12

2.3 Example phase-contrast micrograph from the acquired monolayer MSC culture dataset. 13

2.4 Basic optical configuration of a light-sheet fluorescence microscope adapted from [2]. 16

2.5 Example light-sheet fluorescence volumetric image channels from the acquired three-dimensional MSC culture dataset. a) Brightfield channel used to detect microcarrier locations in the sample mount and center them on the camera's field of view. b) CellTracker green channel showing fluorescently labeled cell bodies. c) DRAQ5 channel expressing fluorescence for nuclei. 18

2.6 Layout of conventional image processing-based segmentation steps for monolayer cultured MSCs. 21

2.7 U-Net model architecture adapted from [3]. Each blue and white box denotes a multi-channel feature map and a copied feature map, respectively. The top of the box corresponds to the number of channels, and box side edges correspond to layer size. Arrows represent different operations listed in the figure. 23

3.1 Flow diagram for the image analysis algorithm. 26

3.2 Figure showing pipeline of the image analysis algorithm. a) Input phase-contrast microscope image. b) Input image pre-processing by contrast adjustment and edge sharpening to make cell bodies more detectable. c) Cell edge detection using their gradient. d) Cell region prediction within the image using the detected edges. e) Seed (candidates cell markers) identification to distinguish each region proposal as either an individual cell or a cluster of cells. f) Clustered cell delineation using the marker-based watershed technique. g) Individual cell segmentation with the edge-based method. h) Classification of MSCs as either RS or SR phenotype using features extracted from segmented cells for logistic regression method. 27

3.3 Performance of the logistic regression model for classifying morphological phenotype of MSCs was validated using AUC $\pm$ s.e. calculated as 0.98 ± 0.02 and 0.87 ± 0.08 for training and testing data sets, respectively, from the displayed ROC curves. 31

4.1 Pipeline of algorithm to classify mesenchymal stem cells in phase-contrast micrographs. Example images illustrate (a) low and (b) moderate density. Cell regions of interest (RoI) in image (b) include (c) individual cells and (d) clusters of cells, differentiated by the number of candidate markers (blue) inside RoIs (white). 35

4.2 Example image of MSCs shows the steps involved in detecting initial cell regions. a) The input phase-contrast micrograph was converted to grayscale and b) pre-processed using contrast adjustment, sharpening, and anisotropic filtering. c) Edges were detected using Sobel operator. d) Edges were connected and filled using dilation, closing, and flood-fill operation to obtain initial cell regions ($I_{\text{initial-region}}$). 36

4.3 Illustration of the steps involved in identifying candidate cell regions. a) $I_{\text{initial-region}}$ obtained after edge detection and morphological operations were used as the input. b) Thresholding using size, intensity, and shape criteria removed detected objects that were not cells. c) Opening and erosion optimized the shape of cell areas. d) Clearing the image border removed incomplete cells resulting in the final candidate cell regions. (Borders of cell regions are highlighted in blue.) 39

4.4 Pipeline for segmentation of markers inside the RoIs (shown in yellow) from $I_{\text{cell-region}}$. a) $I_{\text{pre-processed}}$ was processed to get an image ($I_{\text{marker-processed}}$) suitable for detecting markers. b) Markers were obtained from $I_{\text{marker-processed}}$ using H-minima transform with higher threshold A. c) Markers were obtained from $I_{\text{marker-processed}}$ using H-minima transform with lower threshold B. d) Markers from both the thresholds were merged for RoIs satisfying the perimeter threshold criterion and also for RoIs with zero markers from threshold A. 40

4.5 Representation of the cell segmentation steps. a) $I_{\text{cell-region}}$ was integrated with the results from marker detection (Markers shown in blue). b) RoIs with single markers were identified as individual cells, and c) RoIs with more than one marker were identified as clusters of cells. d) Marker-controlled watershed was carried out to segment individual cells within clusters (Watershed ridge lines shown in red). e) Results from the steps (b) and (d) were combined to get the segmentation output. 42

4.6 Sample images of MSCs from the independent test dataset comparing the ground truth cell outlines (top row) with the segmentation results of the algorithm (bottom row). Columns a and b are examples of low density images whose AF values were estimated as 0.04 and 0.08, respectively by the algorithm. Similarly, examples of moderate density images with AF values of 0.15 and 0.21 are shown in column c and d, respectively. The ground truth images also illustrate that the truncated cells on image borders were not considered for analysis. 49

4.7 Diagonal classifier agreement plots between a) logistic regression classifier (x-axis) and radial kernel support vector machine (y-axis) for day 2 culture and b) K-nearest neighbor (x-axis) and linear discriminant analysis (y-axis) for day 4 culture. Each point represents a cell whose phenotype was predicted as RS (green) or SR (red) by each classifier. 51

4.8 Fitted binormal ROC curves illustrating the performances of day 2 and day 4 ensemble classifiers for the task of distinguishing between RS and SR cells during independent testing. The dashed blue line represents the "RSVM + LR" day 2 classifier and the solid orange line represents the "LDA + KNN" day 4 classifier. 52

5.1 Pipeline of the proposed U-Net based mesenchymal stem cell segmentation method. 59

6.1 Flowchart illustrating the presented 2U-Net algorithm for image-based analysis of MSC cultures. a) U-Net 1 was trained with pre-processed, annotated, and augmented phase-contrast micrographs of MSCs. It was then used to predict cell regions that were further post-processed. Similarly, b) U-Net 2 was modeled to locate algorithm-defined cell markers in the images. These predictions were followed by c) identification of individual cells (blue) and cell clusters (brown); clusters were subsequently segmented using a marker-controlled watershed. After segmentation, object features were extracted and utilized by a machine learning classifier to categorize cells as either RS (green) or SR (red) phenotype. Quantitative metrics viz. culture confluency, cell count, and efficacy ratio were obtained using segmentation and classification results. d) Lastly, the algorithm's potential to estimate culture quality was verified using standard visual inspection as ground truth. 70

6.2 Example images (a – d) of MSCs from independent testing data comparing the ground truth cell outlines (first row) with segmentation results of previous morphological (second row), standard U-Net (third row), post-processed U-Net (fourth row), and reported 2U-Net (fifth row) methods. Cells in ground truth images are outlined in black and they also show that the truncated cells on borders are omitted for analysis. For all the algorithm outputs, true positives (correctly detected cells) are outlined in blue, false negatives (cells that were not detected) are outlined and shaded in blue, while false positives (objects incorrectly detected as cells) are outlined and shaded in green. 78

6.3 Fitted binormal ROC curves illustrating the performance of machine learning models to estimate cell phenotype as RS or SR during independent testing. a) Classifiers trained with features from MSCs segmented by morphological (dashed orange) and 2U-Net (solid blue) algorithms at low density. b) Classifiers trained with features from MSCs segmented by morphological (dashed red) and 2U-Net (solid green) algorithms at medium density. 81

7.1 Schema describing the proposed image analysis framework. 90

7.2 Illustration of steps involved in segmenting single microcarriers. a) Brightfield (BF) channel volume used as input. b) Reduction of noise and intensity enhancement for BF volume through median filtering and standard contrast adjustment. c) Microcarrier edge detection using Sobel filter. d) Removal of background noise and false positive objects through volume-based thresholding. e) Connection and refinement of detected microcarrier edges via morphological erosion followed by closing. f) Generation of microcarrier sphere using bounding box (shown in yellow) dimensions from the previous step. 95

7.3 Representation of segmentation steps for multiple microcarriers. a) DRAQ5 channel volumetric image used as input. b) Improvement of intensity for microcarrier voxels applying limit-controlled contrast adjustment. c) Binarization using Otsu's method with algorithm-defined intensity threshold. d) Connection and filling of binarized voxels via morphological closing. e) Removal of background noise and false positive objects through volume-based thresholding. f) Fine-tuning segmented border of multiple microcarriers utilizing morphological opening. 96

7.4 Sample image volumes of 3D cultured MSCs stained with CTG (first column) from validation data comparing the ground truth cell regions (second column) with the segmentation results of 3D conventional (third column), 2D U-Net (fourth column), and 3d U-Net (fifth column) methods. Each row is an example of a volumetric image from days 3 and 7 of culture expansion consisting of single and multiple microcarriers. DICE score (D) is shown in yellow for each segmentation result image on the top left corner. 99

7.5 Sample image volumes of 3D cultured MSCs stained with DRAQ5 (first column) from validation data comparing the ground truth nuclei (second column) with the segmentation results of 3D conventional (third column), 2D U-Net (fourth column), and 3d U-Net (fifth column) methods. Each row is an example of a volumetric image from days 3 and 7 of culture expansion consisting of single and multiple microcarriers. DICE score (D) is shown on the top left corner, and sensitivity (S) and precision (P) are shown on the bottom left corner for each segmentation result image in yellow. 102

7.6 Scatter plots of observed (Imaris) versus predicted (algorithm) culture confluency values for training (N = 23) and validation (N=21) dataset. The solid black lines indicate computed linear regression lines, and the region shaded in gray color is a 95% confidence interval (CI) for the slope of the regression lines. Root mean square error (RMSE) and adjusted R^2 values are displayed on the top left corner of the plots. a) Plot showing a high linear correlation and close fit of the algorithm with the ground truth for training data. b) Validation data plot confirming high correlation and good fit of the presented algorithm in estimating confluency. 104

LIST OF TABLES

TABLE Page

2.1 Monolayer MSC culture dataset used for the development and validation of image analysis algorithms. 13

2.2 Three-dimensional MSC culture dataset used for the development and validation of image analysis algorithms. 17

3.1 Cell detection accuracy and average DICE score for cell segmentation. 29

3.2 Segmentation error rates. 30

3.3 Performance metrics of the algorithm for classification of MSCs based on their morphological phenotype. 30

4.1 Mesenchymal stem cell culture dataset used for feature extraction and classification. 43

4.2 Human-engineered features extracted for the segmented MSCs. 44

4.3 Cell detection sensitivity of the algorithm for training and independent testing. 46

4.4 Cell detection precision of the algorithm for training and independent testing. 46

4.5 Cell segmentation DICE score of the algorithm for training and independent testing. 47

4.6 Cell detection and segmentation performance of the algorithm per image from training and testing datasets. Mean and standard deviations were calculated for all cells over all images. 48

4.7 5-fold cross-validation to compare performance of models for distinguishing between RS and SR cells. AUC values are in bold for Day 2 and Day 4 classifiers that performed best during cross-validation. 50

4.8 Performance of the algorithm for classifying MSCs as RS or SR for independent testing. 51

4.9 Classification sensitivity of the algorithm for the 5 cell phenotype categories. 53

5.1 MSC detection and segmentation performance of morphological, standard U-Net, and presented U-Net algorithms for the training dataset. 62

5.2 MSC detection and segmentation performance of morphological, standard U-Net, and presented U-Net algorithms for the independent test dataset. 63

5.3 P-values from two-sample t-tests for statistical comparison between morphological, standard U-Net, and presented U-Net algorithms. 64

5.4 10-fold cross-validation AUC of top-performing features extracted from cells correctly segmented using morphological image analysis and presented U-Net for predicting phenotype of MSCs as viable or non-viable. 65

6.1 Dataset used for feature extraction and classification of MSC morphological phenotype. 73

6.2 Object-based cell detection sensitivity and precision of morphological, standard U-Net, post-processed U-Net, and 2U-Net algorithms for training and independent testing. 76

6.3 Image-based cell detection and segmentation performance of morphological, standard U-Net, post-processed U-Net, and 2U-Net algorithms for training and independent testing. Mean and standard deviations (s.d.) were calculated across all images. Statistical comparison of 2U-Net with the other three methods illustrates its significant improvement across all metrics. P-values were adjusted for multiple comparisons using Bonferroni–Holm correction. 77

6.4 Comparison of machine learning models to distinguish between RS and SR phenotype of MSCs. AUC values of low and medium cell density classifiers that performed best during five-fold cross-validation are presented in bold. 79

6.5 Classification performance of morphological and 2U-Net algorithms for independent testing. 80

6.6 Classification sensitivity of morphological and 2U-Net algorithms for five cell phenotype categories determined by average phenotype scores (APS) obtained from 20 readers for training and independent testing data. "Truth" here corresponds to labeling done by the single expert with 15+ years of experience in working with MSCs. 82

6.7 Comparison of culture attributes computed using 2U-Net algorithm and truth for training and independent testing. 83

7.1 Performance comparison of 3D conventional, 2D U-Net, and 3D U-Net methods for cell region and nucleus segmentation from image volumes of microcarrier cultured MSCs. The image analysis approach that performed best during training and validation is presented in bold for each metric. 98

7.2 Statistical comparisons between performance metrics of 3D conventional, 2D U-Net, and 3D U-Net methods for cell region and nucleus segmentation using Imaris ground truth (N = 21). P-values were adjusted for multiple comparisons using Bonferroni-Holm correction, and statistically significant results are indicated in bold. "$>$" in the table stands for "outperformed." 101

A.1 Comparison of loss functions for training U-Net 1 model to segment cell regions. The metrics for the loss function that performed best during five-fold cross-validation are presented in bold. 133

A.2 Comparison of marker annotation methods and loss functions for training U-Net 2 model to locate cell markers. The metrics for the combination of marker annotation method and loss function that performed best during five-fold cross-validation are presented in bold. 135

1. INTRODUCTION AND MOTIVATION

1.1 Chronic diseases and conditions

Chronic diseases are persistent conditions that last more than three months, as defined by the U.S. National Center for Health Statistics [4]. They require ongoing medical attention and have a huge impact on quality of life as well as hinder daily activities. Other traits of chronic diseases include complex causalities, prolonged development periods without symptoms, and associated functional impairments that trigger further health complications. These ailments become more common with age and can neither be prevented by vaccines, cured by medication nor do they spontaneously heal. Chronic diseases such as cancer, diabetes, stroke, lung disease, and arthritis are some of the leading causes of disability and death, affecting nearly half (approximately 45%, or 133 million) of all Americans [5]. According to the National Center for Chronic Disease Prevention and Health Promotion, six in ten adults in the U.S. have a chronic disease and four in ten adults have two or more [6]. Seven out of every 10 deaths in the U.S. are because of chronic illnesses [7]. Additionally, they are also a primary driver of the nation's medical costs, accounting for 90% of the 3.8 trillion dollars expenditure annually [8]. The high prevalence, economic burden, and complicated needs of patients with chronic diseases make them one of the greatest problems faced by the healthcare system, making it crucial to have efficacious therapies and treatments that can provide a long-term cure.

1.2 Mesenchymal stromal cell-based therapies to address chronic diseases

Studies have proven that the approach of cell-based therapeutics has valuable potential to treat a variety of widespread chronic diseases [9]. As an emerging tool for regenerative medicine and tissue engineering, cell-based therapeutics is considered one of the most promising fields of modern medicine. Over the past few years, there has been exponential growth in research and clinical trials employing stem cells for therapies, leading to a significant impact on several disorders [10, 11, 12]. Stem cell-based therapies primarily exploit the unique property of stem cells to self-renew and dif-

ferentiate into the specific cell types required to repair damaged tissues or organs. However, due to the complexity of stem cell-based products, it is important to carefully select a stable, safe, and easily available stem cell source that has the ability to differentiate into several lineages [13].

Mesenchymal stromal cells, one of the most prominent stem cells, has been extensively used in more than 950 registered clinical trials listed with the FDA because of their potential therapeutic effects [14]. They are characterized by fast in vitro proliferation, self-renew, and the ability to differentiate into a variety of mesodermal lineages, making them an ideal source of replacement cells [15]. MSCs are multipotent and can differentiate into many phenotypes, including those that form bone, cartilage, muscle, fat, and other connective tissues [16, 17]. Apart from differentiating into distinctive and specialized cells such as osteoblasts, adipocytes, and chondroblasts, they also have the power to regulate the development of fibroblasts and endothelial cells [18]. Furthermore, MSCs are very accessible, being easily expanded from several tissue types namely skin, ligament, tendon, umbilical cord, and placenta among others. Thus, they are of great value for tissue engineering and regeneration treatments. There is a growing body of literature demonstrating the viability of MSC-based therapy in a variety of pre-clinical models, including acute lung injury [19], skeletal regeneration [20], and septic shock [21].

Heterogeneity is a critical aspect of MSCs that mainly arises from the nature and source of the donor tissue, harvesting and culture procedures, and the extent to which the cells have been previously expanded [22]. It is worth noting that heterogeneity causes disparities in differentiation capacity, proliferation rate, and protein expression profiles even between cell populations of similar origin. The extraordinary heterogeneity of MSC makes drawing reliable inferences about their therapeutic potency very challenging and frequently results in an unpredictable level of clinical efficacy [23, 24]. This issue emphasizes the necessity to meaningfully characterize and control the functionality of MSC products for widespread and successful clinical use. In addition, culture monitoring would not only tackle heterogeneity but also prevent the decline in cell efficacy as well as the expense incurred due to MSC batch failure.

1.3 Current MSC evaluation strategies and their shortcomings

The need to evaluate MSC's therapeutic potency has prompted the development and optimization of several techniques that can precisely correlate the critical quality attributes (CQA) of MSCs with their functionality [25]. Assays have been broadly adopted to predict immunomodulation, differentiation competency, clonogenicity, surface immunotype, and secretion of stromal factors for quality assessment of MSCs [26, 27, 28]. Proliferation, colony-forming unit, activation, and immune cell inhibition assays are some of the most researched in vitro assays that have been applied to determine MSC potency effectively [29, 30, 31]. Aside from assays, flow cytometry and metabolic labeling are also used to analyze gene expression profiling, surface markers, cell cycle, specialized differentiation, and biological properties of MSCs [32, 33, 34]. Generally, the methods mentioned above are performed on parallel cultures set up during or after the harvest of large-scale expansion cultures. While these assays are definitive in some cases, they fail to provide real-time, non-invasive data on expansion culture itself.

Regular visual examination under a microscope to morphologically profile cells is another qualitative approach that is used for routine quality assurance of cultures. Earlier studies have established that morphological traits can be highly predictive of MSC's downstream efficacy, and along with post hoc characterization assays, MSC cultures are typically subjected to morphological profiling during expansion [35, 36]. Functionally relevant morphological screening has previously shown that it can identify optimal priming conditions and accurately monitor immunosuppression capacity as well as transcriptomic and emergent phenotypes of MSCs [37, 38, 39]. Although widely used, the standard practice of morphological characterization via visual inspection lacks reliability due to subjectivity in interpretation by different scientists. Moreover, it is also tedious and time-consuming, making it incompatible with high-throughput analysis, quantification, or high temporal resolution. Therefore, a non-invasive, rapid, and objective method is essential in order to overcome these shortcomings for robust viability analysis of MSC cultures.

1.4 Exploiting computer vision to quantitatively assess MSCs

The aforementioned challenges can be resolved by adopting computer-aided techniques to analyze cultures using their microscopy images. The topic of computer vision and artificial intelligence-based cell analysis has received a lot of attention with increasing demands in bioinformatics and significant contributions in the domains of medical diagnostics and biomedical engineering [40, 41, 42]. It has been incorporated for understanding drug influences [43], image-guided therapy [44], and detection or diagnosis of medical conditions including COVID-19 [45], brain tumors [46], breast cancer [47] among others. In particular, morphological processing has become a standard theory for computerized cell image analysis and pattern recognition. This field encompasses a wide application area such as cell clump segmentation [48], feature extraction [49], and abnormal cell identification [50]. It has also been integrated with the investigation of histological tumor sections [51], boundary detection of epithelial cell nuclei [52, 53], localization of anatomical structures [54], cell tracking [55], and cell classification [56]. These studies expansively demonstrate and validate the potential of image-based analysis to analyze cells objectively and robustly. Moreover, it would be non-invasive and can be applied for automated evaluation of live cells to get quantitative measures of culture quality. This idea could be similarly extended to obtain MSC morphology as well as other culture parameters more efficiently as empirical values using computer vision. Thus, a computer-based examination would enable quantitative and rapid assessment of MSC culture quality and efficacy.

1.5 Proposed algorithms to analyze monolayer and three-dimensional MSC cultures

The main focus of existing analysis techniques has been on conventional thresholding techniques to segment MSCs [57]. However, this research was more driven towards identifying all cell regions rather than individual cells, making it unsuitable for culture quality monitoring through morphological profiling of each cell. In addition to segmentation, there has been significant research demonstrating the possibility of characterizing MSCs based on their shape [58]. Machine learning has been implemented previously to classify MSCs from other cell lines [59, 60], to pre-

dict immunosuppressive capacity using their functional subpopulations [39], and also to identify them based on their differentiation potential [61]. The presented work aims to expand on these approaches to segment individual MSCs from their monolayer culture images using morphology-based analysis. Morphometric and textural features are computed for all segmented cells as they serve as good indicators of the physiological state and functionality of MSCs. A machine learning classifier uses these extracted features to predict the morphological phenotype of cells. Based on phenotype estimation for cells, the algorithm measures culture parameters and viability. Thus, our morphological-based image analysis contributes to automated quantification of adherent MSC cultures and also validates the applicability of image analysis as a tool for characterizing MSCs.

The above-mentioned morphological analysis algorithm employs low-level pixel processing and mathematical modeling to construct a compound rule-based system that segments monolayer cultured MSCs in phase-contrast micrographs. The application of this algorithm is highly reliant on image quality and would only be suitable for images where cell boundaries, as well as their bodies, are distinctly perceptible in the image. As cells proliferate and the level of culture confluency increases, they form clusters. Localizing a higher number of single cells within each cluster with exact borders becomes a complex task for conventional techniques such as edge detection, thresholding, and morphological operations. Furthermore, such standard image processing tends to generalize inadequately on new culture images as they greatly rely on manually designed data representation, parameter tuning, and feature engineering. Thus, it is crucial to develop an image analysis pipeline that can provide correct segmentation to effectively measure morphological attributes of cells and accurately estimate culture efficacy.

Deep learning has the power to overcome these challenges and extract information not easily comprehended via visual investigation [62]. It can understand multi-modal information and implicitly capture intricate cellular features, leading to a process of learning far more superior than standard computational or machine learning models. To this end, we developed a dedicated deep learning-based model called 2U-Net in combination with standard image processing procedures for enhanced MSC segmentation. Similar to the morphological analysis algorithm, 2U-Net seg-

mentation is followed by feature computation and machine learning to determine the viability of every cell in the culture based on its phenotype. Our deep learning-based scheme significantly optimizes the segmentation of individual and clustered MSCs, enabling effective and reproducible estimation of culture's primary efficacy endpoints. This research shows great promise to bridge the gap between existing techniques and the need for a robust and streamlined process for examining MSCs. Also, our algorithm does not need staining and can work without inducing any changes to cell properties. Thus, it will be non-invasive and suitable for real-time monitoring for monolayer cultures.

It is also important to note that the clinical success of cell-based therapies to provide new remedies is highly reliant on the supply of cells in large and frequent doses. However, conventional monolayer culture systems impede the production of commercially viable quantities of cells as demand increases. This issue is mainly because the adherent surface area available for cell growth in two-dimensional (2D) cultures is limited. Three-dimensional (3D) cell culturing not only addresses this limitation by providing higher adherence area for cells but is also highly suitable for manufacturing clinical-grade cells that behave in a more physiologically relevant manner [63]. Similar to monolayer culture assessment, an analysis protocol is also required to ensure the quality of MSCs cultured in a 3D environment. We extend the conventional and deep learning-based algorithms developed for detecting cells grown in 2D to segment MSCs cultured on microspheres in 3D. Its pipeline utilizes combination of conventional and U-Net processing, and employs the segmentation results to calculate culture confluency and cell count for 3D cultures. This image processing scheme would improve reproducibility and characterization of 3D cultures, leading to a clinically significant therapeutic infusion of MSCs.

It is anticipated that the proposed 2D and 3D algorithms would enable biologists and cell manufacturers to draw conclusions about the functionality of recovered MSCs. These algorithms would potentially replace visual investigation, leading to rapid, objective, and standardized multidimensional MSC examination. Lastly, they could also be incorporated to scale up stem cell manufacturing, paving way for efficacious cell therapies to treat chronic diseases.

2. TECHNICAL BACKGROUND

2.1 Cytotherapy: a brief overview

Cell-based therapeutics, also known as cytotherapy is a biological process of injecting, grafting, or implanting viable cells into a patient to effectuate medicinal outcomes and cure illnesses. It aims to repair the mechanisms underlying disease initiation and progression via trophic effect or cell replacement [64]. Cytotherapy is considered one of the most promising disciplines in the fields of modern science and medicine because of its particular objective of repairing diseased cells, tissues, or organs and eventually retrieving normal function [65]. As a leading modality of regenerative medicine, it offers boundless possibilities to revolutionize treatments of deadly chronic conditions not adequately addressed by existing pharmaceuticals. Cell therapies have been broadly applied or are in clinical trials for prominent disorders such as cancer [66], cardiovascular [67], neurological [68], kidney [69], and liver [70].

With new technologies and innovative products, multiple cell types may be utilized as a part of therapies or treatments, including stem, progenitor, or primary cells. Stem cells have valuable potential for therapeutic uses in tissue engineering and regenerative medicine as they can build virtually every tissue in the human body. In recent years, a vast amount of research has been carried out on stem cell-based therapies as a potential new strategy for a wide range of treatments, specifically for degenerative diseases [71]. Stem cells offer the perfect solution for cytotherapies owing to their unique property of self-renewal in addition to their capacity to differentiate into specialized adult cell types. Long-term replacement of damaged tissue is achieved through stem cell differentiation into a specific cell type in the lab or after reaching the site of injury. These cells then integrate into the injury site, facilitating the improved function of organs or tissues. Furthermore, stem cells that release soluble factors such as cytokines, chemokines, and growth factors can promote self-healing of the organ by inducing cells to migrate towards the transplantation site. Apart from generating healthy cells, stem cells have also enabled increased understanding of dis-

ease occurrence and investigation of new drugs for safety and quality [72]. Thus, stem cell-based therapy is a tangible reality that is advancing rapidly and significantly impacting human health as well as the quality of life.

2.2 Mesenchymal stromal cells: current understanding, therapeutic potential, and clinical challenges

Adult bone marrow is the most well-characterized source for adult stem cells and it contains a heterogeneous population of cells such as hematopoietic cells, macrophages, erythrocytes, fibroblasts, adipocytes, and endothelial cells [73]. Aside from these cell types, bone marrow also contains a subset of non-hematopoietic stem cells known as marrow stromal stem cells, mesenchymal stem cells, or mesenchymal stromal cells (MSCs) [74]. MSCs are primitive cells that originate from the mesodermal germ layer and give rise to connective tissues, skeletal muscle cells, and cells of the vascular system [16, 17]. They are spindle-shaped plastic-adherent cells isolated from bone marrow, adipose, and other tissue sources [75]. MSCs can expand in culture for multiple population doublings using relatively simple culture conditions. MSCs possess multilineage potential with the capability to readily differentiate into cartilage cells (chondrocytes), bone cells (osteoblasts), and fat cells (adipocytes), both in vitro and in vivo [76].

MSCs, a heterogeneous group of stem cells has gained a lot of attention over the past few years for clinical applications in regenerative medicine due to a broad repertoire of attractive biological and clinical qualities. Their proliferative, immunomodulatory, and anti-inflammatory properties make them an excellent source of allogeneic cells for tissue repair in trauma and chronic diseases [77]. There is a substantial amount of early research exploring the therapeutic efficacy of MSCs in a variety of pre-clinical studies, including graft-versus-host-disease (GVHD) [78], acute myocardial infarction [79], autoimmune diseases [80], and liver diseases [81]. This vast number of published works proves that MSCs are an appealing tool for various cellular therapies. However, it is crucial to understand the biology, functionality, and role of MSCs in clinical trials to enhance therapeutic efficacy and effectively treat a variety of pathological conditions.

A key characteristic of MSCs is their heterogeneity that makes drawing reliable conclusions

about their therapeutic potential very complicated. Cellular heterogeneity arises due to different MSC origins as well as harvesting and culture procedures [82]. It has been demonstrated that MSC product quality greatly depends upon isolation and culture methods as well as traits and history of the donor [23]. Several in vitro studies have aimed at studying possible modifications of the ex vivo culture environment and MSCs themselves to increase their regenerative potential [83, 84]. The biggest challenge is standardizing these many variable factors during MSC expansion to obtain a viable number of therapeutically efficacious cells. Besides cell quantity, improved characterization of MSCs requires a well-defined correlation between phenotype and stromal/stem cell functions.

For decades, it has been appreciated that MSC morphology is highly predictive of downstream efficacy where efficacy is defined collectively as the potential for self-renewal, multipotentiality, and engraftment [85, 86, 87]. As such, MSC cultures are typically subjected to morphological profiling during expansion in addition to post hoc characterization assays [35, 36]. MSC phenotype has been categorized as RS (potentially functional and efficacious) or SR (potentially non-functional and deficient) in accordance with their morphology [88, 89, 90]. MSCs that rapidly self-replicate (RS-type) are relatively small, spindle-shaped, and fibroblastic, whereas cells that slowly replicate (SR-type) are large, flattened, and rhomboidal, illustrated in Fig. 2.1. In addition to slow replication, SR cells lose most of their ability to differentiate into multiple cell lineages and promote tissue repair, suggesting functional decline. These phenotypes have been predictive of proliferative and differentiation potential in vitro [88] and functional engraftment in vivo [89, 90]. Hence, relative contributions of RS and SR cell populations can be utilized to predict culture efficacy qualitatively and downstream performance of MSCs. Visual observation via a microscope is currently the standard methodology to estimate MSC phenotype as RS or SR. Although broadly employed, this technique lacks consistency and reliability due to subjectivity in interpretation by different scientists. It is tedious and labor-intensive, lacking compatibility with high-throughput analysis and robust quantification of MSC culture viability. Besides visual investigation, flow cytometry (forward and low side scatter) has been shown to resolve discrepancies in data obtained for similar preparation of MSCs by distinguishing cells into distinct categories, viz. RS and SR [91].

They isolate RS cells only based on size and granularity, not considering their crucial morphological features. Moreover, it is also an invasive and time-consuming process. Nevertheless, RS and SR morphological phenotypes of MSC represent a foundation on which more specific criteria to better identify MSC properties could be developed.

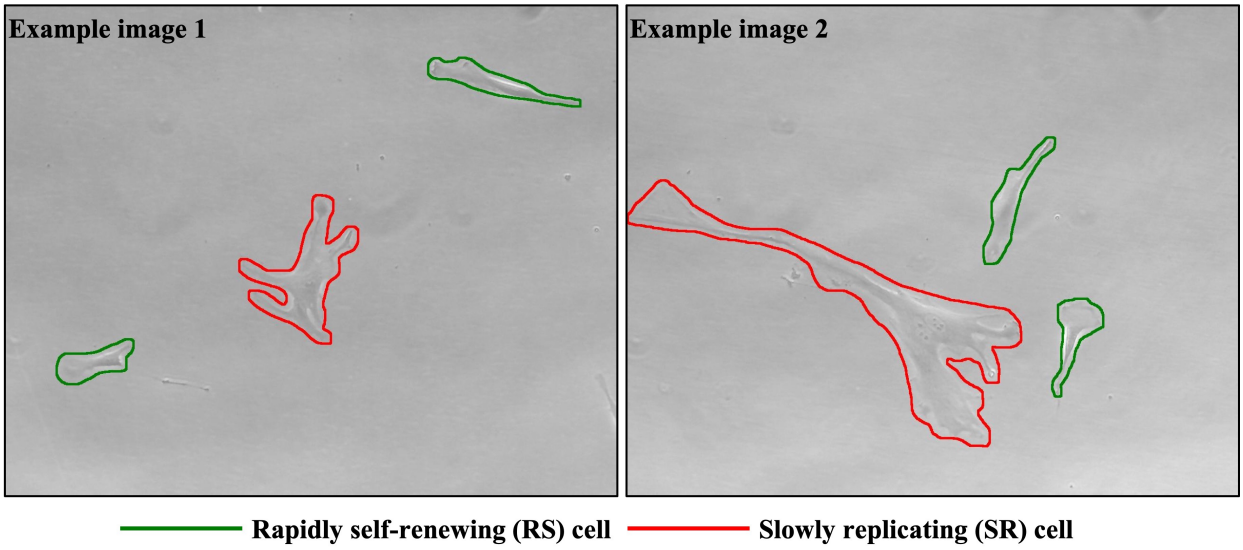

Figure 2.1: Example images showing distinct morphology of RS and SR MSCs.

2.3 Culture protocol, imaging, and dataset for MSCs grown on monolayer

The monolayer culturing protocol, imaging modality, and the acquired dataset described in this section are used for the development and evaluation of 2D image-based MSC analysis algorithms described in chapters 3-6.

2.3.1 Protocol for monolayer MSC cultures

Induced pluripotent stem cell-derived human MSCs were seeded at 100 cells/cm^2 under standard conditions of expansion and prepared as previously described [92, 93]. Cells were expanded in T175 flasks with 25 mL complete culture medium (CCM) containing 90% (v/v) α-minimum essential media (α-MEM), 10% (v/v) fetal bovine serum (FBS), 4 mM L-glutamine, and 100 U/mL

penicillin-streptomycin. For MSCs employed in this study, a density of approximately 1,000 cells/cm^2 (low cell density) was expected at day 2 and approximately 6,500 cells/cm^2 (medium cell density) at day 4.

2.3.2 Imaging of monolayer MSC cultures

Cell imaging is a crucial aspect of both qualitative and quantitative culture assessment. Few widely used methodologies for imaging monolayer cell cultures include time-lapse microscopy, interference contrast microscopy, and absorption microscopy [94, 95, 96, 97, 98, 99]. The two modalities that have gained importance to visualize cells with improved contrast are fluorescence imaging and phase-contrast microscopy. Fluorescence microscopy provides good image contrast; however, it suffers from certain limitations. For this method, cells need to be genetically engineered to generate fluorescent proteins or fluorescently labeled to enhance cell boundary information, which modifies cell physiological makeup and may cause an unknown change of cellular dynamics [100]. It also suffers from photobleaching, the process due to which the dyes used for staining undergo a photo-chemical alteration of their molecules such that they permanently can't fluoresce [101]. This issue limits the applicability of fluorescence microscopy for long-term cell monitoring. Also, it is an invasive method, making it inappropriate for real-time monitoring of live cell cultures.

Phase-contrast microscopy does not suffer from these disadvantages and provides relatively high image contrast without any biological modification to cells [102]. It is an optical imaging method that converts phase shifts caused due to refractive index differences between cells and substrate into intensity changes in the micrographs [103]. Phase shifts themselves are invisible but become visible when shown as brightness variations. This microscopy is particularly significant in biology as it can reveal many cellular structures that are invisible with a brightfield (BF) microscope. Optical layout of a phase-contrast imaging system is illustrated in Fig. 2.2 [1]. It is one of the few methods available to study living cells and their components without needing fluorescence or any staining that causes cell death. Therefore, a phase-contrast microscope can non-invasively image monolayer cultures with good contrast, making it a suitable choice for qualitative or quantitative evaluation of MSC morphology in real-time.

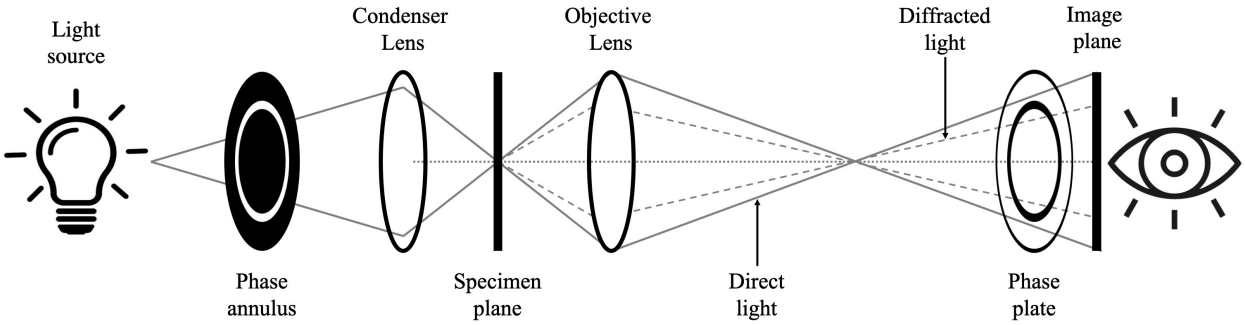

Figure 2.2: Basic schematic of a phase-contrast microscope adapted from [1].

A Motic AE31 phase-contrast microscope with a 10X objective and Moticam 1SP 1.0 MP camera was used to acquire culture images for this research. Cells were imaged on the second (low cell density) and fourth day (medium cell density) after the culture was prepared to capture variation in phenotype as cells proliferate. All the images collected for this study had a size of 1280×1024 pixels and a resolution of 1.56 pixels/μm.

2.3.3 Monolayer MSC culture dataset

Cell culture and image capture were repeated three times to generate the dataset for training and testing of the algorithms. A sample image from the acquired dataset is presented in Fig. 2.3. Using Adobe Photoshop and Microsoft Paint software, cells were manually segmented and labeled as RS or SR phenotype by an individual with more than 15 years of experience in culturing MSCs. Images from two cultures served as ground truth for training, and images from the third culture were used for independent testing as detailed in Table 2.1. The training dataset of 71 images consisted of 472 cells with 307 cells labeled as RS and 165 cells as SR. The algorithm was validated with 36 phase-contrast micrographs having 186 cells with 121 RS cells and 65 SR cells. Each cell from the segmented ground truth was also characterized as RS or SR by 20 people trained to visually identify MSC phenotype to further analyze the generalizability of the method.

Table 2.1: Monolayer MSC culture dataset used for the development and validation of image analysis algorithms.

Dataset	Culture day	No. of images	No. of cells	No. of RS cells	No. of SR cells
Culture 1 (Train)	2	15	76	49	27
	4	15	146	80	66
Culture 2 (Train)	2	17	96	73	23
	4	24	154	105	49
Culture 3 (Test)	2	15	64	46	18
	4	21	122	75	47

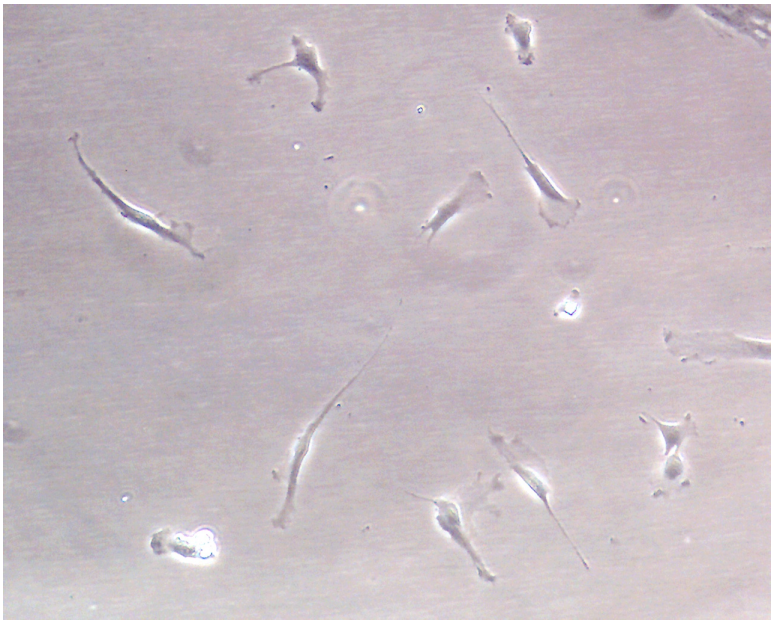

Figure 2.3: Example phase-contrast micrograph from the acquired monolayer MSC culture dataset.

2.4 Culture protocol, imaging, and dataset for MSCs grown on microcarriers

The 3D microcarrier-based culturing protocol, imaging modality, and acquired dataset outlined in this section are used for the development and evaluation of MSC assessment algorithms described in chapter 7.

2.4.1 Protocol for three-dimensional MSC cultures

MSCs were first cultured in monolayer conditions as described in section 2.3.1. 70% confluent cell monolayers were exposed to 0.25% trypsin and 0.1% ethylenediaminetetraacetic acid (EDTA) at 37°C for 5 minutes to passage the cells. CCM was used to deactivate trypsin, and centrifugation was performed at 500*g* for 5 minutes to collect cells. These cells were seeded onto synthesized gelatin methacryloyl (gelMA) microcarriers [104] in a 10 mL Rotating Wall Vessel (RWV) bioreactor. Approximately 110,000 gelMA microcarriers (90-120 μm diameter) with a combined growth area of 50 cm^2 and 5×10^4 cells (1,000 MSCs/cm^2) were incubated in 10 mL CCM in RWV. Cells were allowed to attach to microcarriers for 1 hour during on/off cycles of 24 RPM for 1 minute and 0 RPM for 20 minutes, respectively. The bioreactors were set at 24 RPM for the remainder of the culture duration. Half of the bioreactor media was replaced with fresh CCM every 2-3 days.

After expanding cells until day 3 and day 7, they were treated with 1 μM CellTracker Green (CTG) 5-chloromethylfluorescein diacetate (CMFDA) dye for 30 minutes, washed twice with 5 mL phosphate-buffered saline (PBS), and fixed in 4% paraformaldehyde (PFA) under gentle agitation for 1 hour. Following fixation, cell samples were treated with 20 μM DRAQ5 staining buffer at 37°C under agitation for 30 minutes, centrifuged at 50*g* for 30 seconds, and rinsed twice with 1.5 mL PBS. The sample was centrifuged again, PBS was removed, and 1 mL molten liquid agarose was added. The suspension was mixed to make it homogenous, and then cells in agarose were pipetted into the sample mount and allowed to cool and solidify for imaging.

2.4.2 Imaging of three-dimensional MSC cultures

In monolayer cultures, the intensity of transmitted illumination light is not affected because the cells are attached to transparent glass or plastic substrate tens of microns thick. During phase-

contrast imaging, cells generate contrast as they are aligned along a single plane that translates into the microscope's focal plane. However, for 3D cultures, this kind of BF imaging method would produce blurred images of cells adhered to spherical microcarriers. These microcarriers are thick (hundreds of microns in diameter) and made of materials that prevent light from interacting with cells without significant distortion. Moreover, cell themselves being do not produce sufficient contrast to facilitate single-cell resolution. These problems can be addressed by labeling the cell bodies and nuclei with appropriate fluorescent markers for high contrast visualization. Though destructive, fluorescent staining is currently the most effective option for robust morphological characterization and quality evaluation of 3D MSC cultures. Optical coherence tomography, multiphoton microscopy, confocal laser scanning microscopy, light-sheet fluorescence microscopy (LSFM), and structured illumination microscopy are a few popular imaging systems used to assess the structural and functional properties of 3D cell cultures [105, 106, 107].

Over recent years, LSFM has emerged as a promising imaging approach that overcomes the limitations of conventional fluorescence, confocal, structured illumination, and multiphoton microscopy [108]. The collection objective is placed perpendicularly to the excitation path in LSFM, unlike conventional fluorescence microscopes where both illumination and collection of the signal are performed on the same axis. Light-sheet illumination plane is created by either focusing laser in single dimension with cylindrical lens or scanning along a single axis and then concentrated by an illuminating objective to the focal plane of the detection objective. Moreover, LSFM microscopes can overcome the detrimental effects of sample absorption and scattering, providing increased penetration depth, resolution, and image contrast [109, 110]. The novel optical configuration of LSFMs, with uncoupled perpendicular illumination and detection axis, combined with fast camera-based detection, provides 3D representations of the samples with subcellular resolution at unprecedented speeds [111]. Additionally, it reduces phototoxicity and photobleaching by orders of magnitude compared with conventional microscopes by only illuminating the plane getting imaged, allowing efficient 3D imaging of live cell samples over hours and even days [112]. Hence, LSFM would serve as an ideal volumetric imaging approach to enable visualization, mor-

phological characterization, and quality assessment of 3D MSC cultures at the single-cell level [113, 114, 115, 116]. A simple schematic of LSFM is shown in Fig. 2.4 [2].

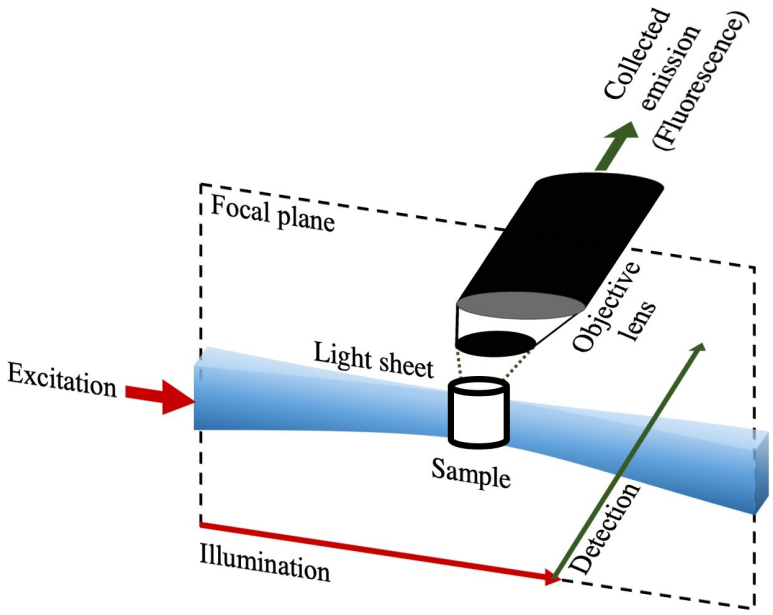

Figure 2.4: Basic optical configuration of a light-sheet fluorescence microscope adapted from [2].

In this study, volumetric images of MSCs attached to gelMA microcarriers were acquired using Zeiss Z.1 LSFM. BF mode trans-illuminated the sample with a white LED and was used to identify microcarriers in the sample mount and center them on the camera's field of view. Cell-bound CTG was excited at 488 nm, and nuclei-bound DRAQ5 was excited at 638 nm in fluorescence mode. Spectral filters were used to detect the 505-545 nm and 660+ nm emission bands for CTG and DRAQ5 channels, respectively. Dual objective illumination with pivot scanning and online fusion was used to improve illumination of the entire microcarrier surface and reduce acquisition time. The frame step size was optimized for a 20X, 1.0 NA water immersion objective lens, the light sheet thickness was set to 2.4 µm, and the detector size was changed to 1024 $\times$ 1024 pixels to reduce output file size. The illumination power and camera integration time were adjusted to minimize both laser power and imaging time while using the full dynamic range of the camera

detector. Z-stacks of the 3D sample were then acquired in both BF and fluorescence modes (CTG channel for cell body and DRAQ5 channel for nuclei).

2.4.3 Three-dimensional MSC culture dataset

Image volumes acquired of cells fixed on days 3 and 7 of culture expansion served as training, validation, and testing datasets for the presented algorithm [Table 2.2]. The algorithm was trained with volumetric images from these two days to capture changes in cells as they proliferate on microcarriers. Each image volume consisted of either a single microcarrier or multiple microcarriers clumped together. For both days, the number of cells ranged from 1 to 10 and 5 to 200 for single and multiple microcarrier volumes, respectively. Multiple microcarrier volumes consisted of clumps of 2 to 30 microcarriers. Figure 2.5 displays example volumetric image channels obtained through LSFM for microcarrier cultured MSCs.

Table 2.2: Three-dimensional MSC culture dataset used for the development and validation of image analysis algorithms.

Culture day	No. of microcarriers per image volume	Training		Validation	
		No. of image volumes	No. of cells	No. of image volumes	No. of cells
Day 3	Single	8	36	8	36
	Multiple	4	34	4	49
Day 7	Single	7	34	6	28
	Multiple	4	200	3	226

Imaris, a commercial microscopy image analysis software, was used to process the acquired volumetric images and prepare the ground truths for the algorithm. Gamma correction, layer normalization, and baseline subtraction were used to improve the quality of CTG and DRAQ5 Z-stacks

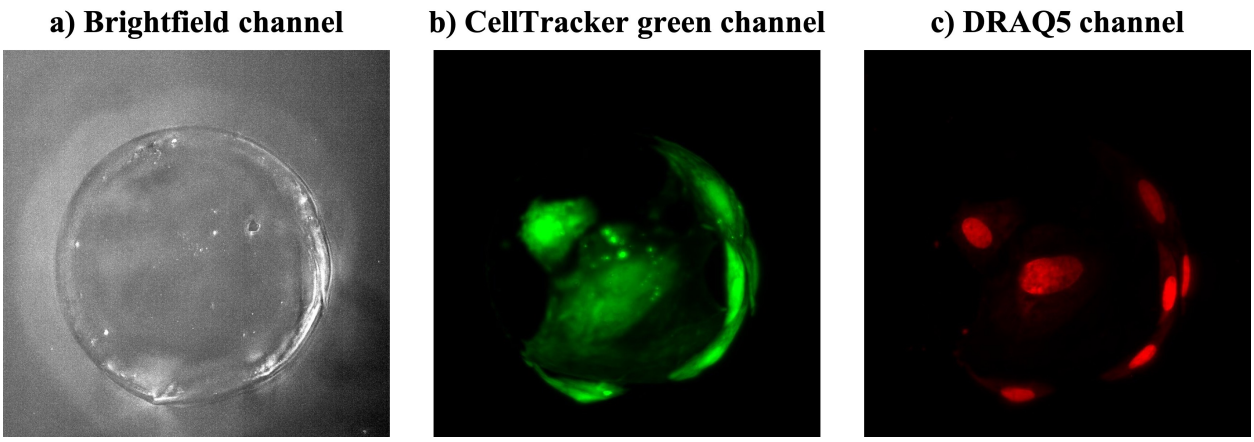

Figure 2.5: Example light-sheet fluorescence volumetric image channels from the acquired three-dimensional MSC culture dataset. a) Brightfield channel used to detect microcarrier locations in the sample mount and center them on the camera's field of view. b) CellTracker green channel showing fluorescently labeled cell bodies. c) DRAQ5 channel expressing fluorescence for nuclei.

before their analysis. Gamma correction adjusted the pixel brightness of a specific range of voxels, and layer normalization corrected the brightness and contrast of individual frames. Baseline subtraction removed the baseline intensity value from every voxel in the dataset to eliminate unwanted scattering artifacts. Cell nuclei regions on the microcarrier surface were detected using DRAQ5 datasets and spot counting functions of Imaris in a supervised manner. Nuclei region identification was followed by labeling the center of each detected object. The center points were then manually inspected and removed, added, or translated as necessary. The location and count of these points served as the reference for cell count and position in each volumetric image.

After generating nuclei center points, ground truth for segmentation of nuclei bodies was also obtained in a semi-automated approach using the DRAQ5 datasets and surface creation functions of Imaris. Object size filter and intensity threshold were applied to reduce unwanted signals from cell debris, scattering artifacts, or out-of-focus light. An additional filter on the number of voxels in the object was utilized to remove objects wrongly detected as nuclei. Finally, the segmented objects were inspected manually and validated via the nuclei center points from the cell location/count truth. The steps described for cell nuclei segmentation were similarly carried out for cell region ground truth using CTG datasets and built-in surface creation functions of Imaris. The segmented

cell regions were manually inspected and validated by identifying the corresponding cell nuclei from the segmented nuclei data.

2.5 Quantitative assessment of monolayer and 3D MSC cultures using image-based analysis

The imaging systems described above facilitate quick in-process visualization of MSCs and render images suitable for high-throughput image-based quantitative study of 2D and 3D cultures. The subject of cell image analysis has received much attention with the increasing demands in bioinformatics and biomedical applications [40, 41, 42]. Morphological cell analysis has become a standard theory for computerized cell image processing and pattern recognition. This field encompasses a vast application area, such as cell clump segmentation, morphological feature extraction, and abnormal cell identification [48, 49, 50]. Computer vision-based morphological evaluation of cells has been integrated with the study of histological tumor sections [51], boundary detection of cell nuclei [52, 53], characterization of cellular phenotypes [117], understanding drug influences [43], and many more. Thus, image analysis can be a powerful tool for the quantitative and automated study of multidimensional MSC cultures.

2.5.1 MSC segmentation from their monolayer culture images using conventional processing techniques

In digital image processing, cell segmentation is the first critical step to analyze and characterize cells from their microscopic images. Segmentation is the process of partitioning culture images into a set of pixels that belong to cells and the substrate. The goal of segmentation is to simplify cell images into something more meaningful and easier to examine. It aids with locating cells and their boundaries in images. An overview of the proposed cell segmentation steps is displayed in Fig 2.6. Automated cell segmentation includes two aspects cell localization and boundary detection [118]. Cell localization, a process of determining cell position in their images, is of great importance to carry out a reliable morphology-based analysis. Sobel filter, an edge detection technique, highlights pixels that have a sudden change in their intensity and locates cell boundaries in images. These detected edges on morphological dilation operation can provide candidate cell

regions. Morphological opening and closing after dilation will remove substrate pixels wrongly segmented as cells and enhance cell boundaries. Area-based thresholding can facilitate getting rid of false detections due to imaging artifacts and debris. The threshold values for these operations will be fixed during training based on their performance relative to the ground truth prepared via manual inspection by a culture expert.

Candidate cell regions will be processed further to detect markers for identifying single cells. Cell markers can be identified using local minima from cell intensity profile followed by morphological processing (dilation and erosion) [119]. A cell region can be distinguished as a single cell or a cluster of cells based on the number of markers it contains. Edge-based methods usually fail to precisely segment each cell as a cell cluster does not have well-defined edges. However, a region-based strategy can separate cells as it relies on the homogeneity of cell region pixels. Watershed segmentation [120], a popular region-growing technique, interprets an image as a topographic landscape with ridges and valleys. Their elevation values are defined using the gradient magnitude of respective pixels. Based on such a representation, the watershed algorithm decomposes an image into catchment basins. For each local minimum, a catchment basin comprises all points whose path of steepest descent terminates at this minimum. Water starts to rise from low gradient areas (interior of objects and background), and watershed borders are built at the maxima of the gradient magnitude, separating basins from each other. It decomposes an image completely and thus assigns each pixel to a region or a watershed. However, if watershed segmentation is applied directly to the gradient magnitude image, it generally results in over-segmentation due to intensity variations within objects and background. However, this problem is solved when water is allowed to rise only from places identified as markers. Cell markers segmented earlier will give the position of each cell within a cluster and the number of cells within each region. Thus, the marker-controlled watershed approach can segment clumped cells accurately, avoiding over or under segmentations [121].

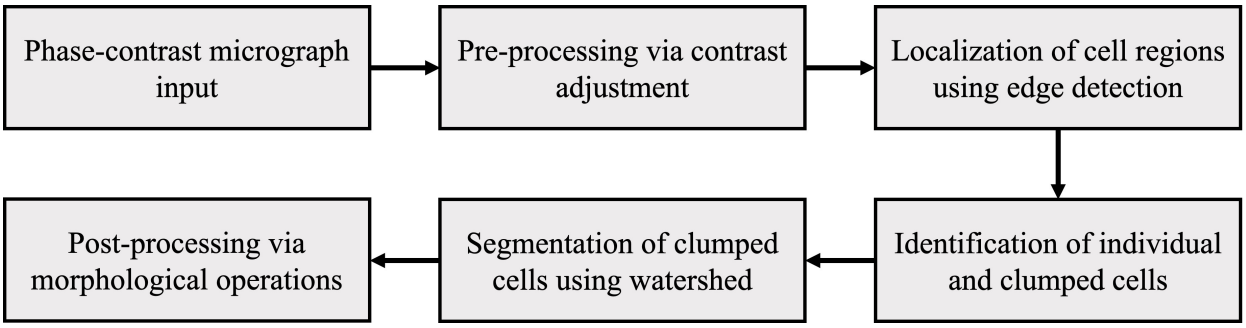

Figure 2.6: Layout of conventional image processing-based segmentation steps for monolayer cultured MSCs.

2.5.2 MSC segmentation from their monolayer and 3D culture images using deep learning

The image segmentation scheme described above performs sequential application of low-level pixel processing (edge and line detector filters, region growing) and mathematical modeling (fitting lines, circles, and ellipses) for automated MSC detection. It can validate the applicability of computer vision as a tool for objectively monitoring cells. However, such an implementation is only suitable for images distinctly recognizable cell boundaries and bodies. As the level of confluency increases, cells proliferate and form clusters. Segmenting a high number of cells within clusters with optimized boundaries is a complex task for conventional image processing algorithms. This complexity arises because pixel-level methods such as edge detection fail to work as expected when cell bodies overlap. Improper segmentation is undesirable as it leads to improper characterization of cells. It is necessary to overcome the dependency on manually engineered features and parameters for robust MSC segmentation. It is only logical to let computers learn features that optimally represent cells, which would help in quantifying their attributes. The concept of automated feature learning at various levels of abstraction forms the basis of deep learning algorithms [62]. It has dramatically improved the state-of-the-art visual object recognition, object detection, segmentation, and classification for biomedical applications [122, 123]. Deep learning-based segmentation extracts details that are difficult to understand visually and can effectively capture intricate cellular features [124, 125]. Therefore, deep learning can provide better precision in segmenting MSC

from their monolayer culture images.

U-Net model [3] shown in Fig. 2.7 is a state-of-the-art deep learning solution that has gained a lot of attention due to its effectiveness in training a convolutional neural network with a reduced number of annotated images. It has higher computational efficiency and can separate touching objects of the same class. Hence, it can be adopted for more efficient and optimized MSC segmentation. The U-Net architecture consists of three sections, namely contraction, bottleneck, and expansion. The contraction section is made of several blocks. Each block takes an input and applies two 3×3 convolution layers, followed by a rectified linear unit (ReLU) and a 2×2 max pooling. The number of feature maps doubles after each block to be able to learn complex structures effectively. The bottleneck layer is at the bottom and mediates between the contraction and expansion layers. It uses two 3×3 convolutional layers followed by ReLU and a 2×2 up convolution layer. The premise of this architecture lies in its expansion section. Similar to the contraction layer, it also consists of multiple blocks. Each block passes the input to two 3×3 convolutional layers followed by ReLU and a 2×2 upsampling layer. After each block, the number of feature maps used by the convolutional layer becomes half to maintain symmetry. The input also gets appended by feature maps of the corresponding contraction layer every time. This action ensures that features learned while contracting the image are used to reconstruct it. The number of expansion blocks is as same as the number of contraction blocks. After expansion, resultant mapping passes through another 1×1 convolutional layer to convert feature vectors into desired binary segmentation output.

Developing two separate U-Net models based on the above architecture for segmenting cell regions and markers will enhance the sensitivity and precision of detecting single MSCs. Outputs from both U-Nets can be integrated and processed using a marker-controlled watershed to delineate clustered cells similar to conventional processing algorithms. Another main advantage of the 2D U-Net model is that it can be translated easily to segment 3D volumes as the model architecture remains the same [126]. All the layers used in the 2D network can be designed similarly for the 3D network. Moreover, 3D U-Net simplifies the segmentation of different fluorescent channels used for labeling cell cytoplasm and nucleus owing to its ability to work with multiple channel

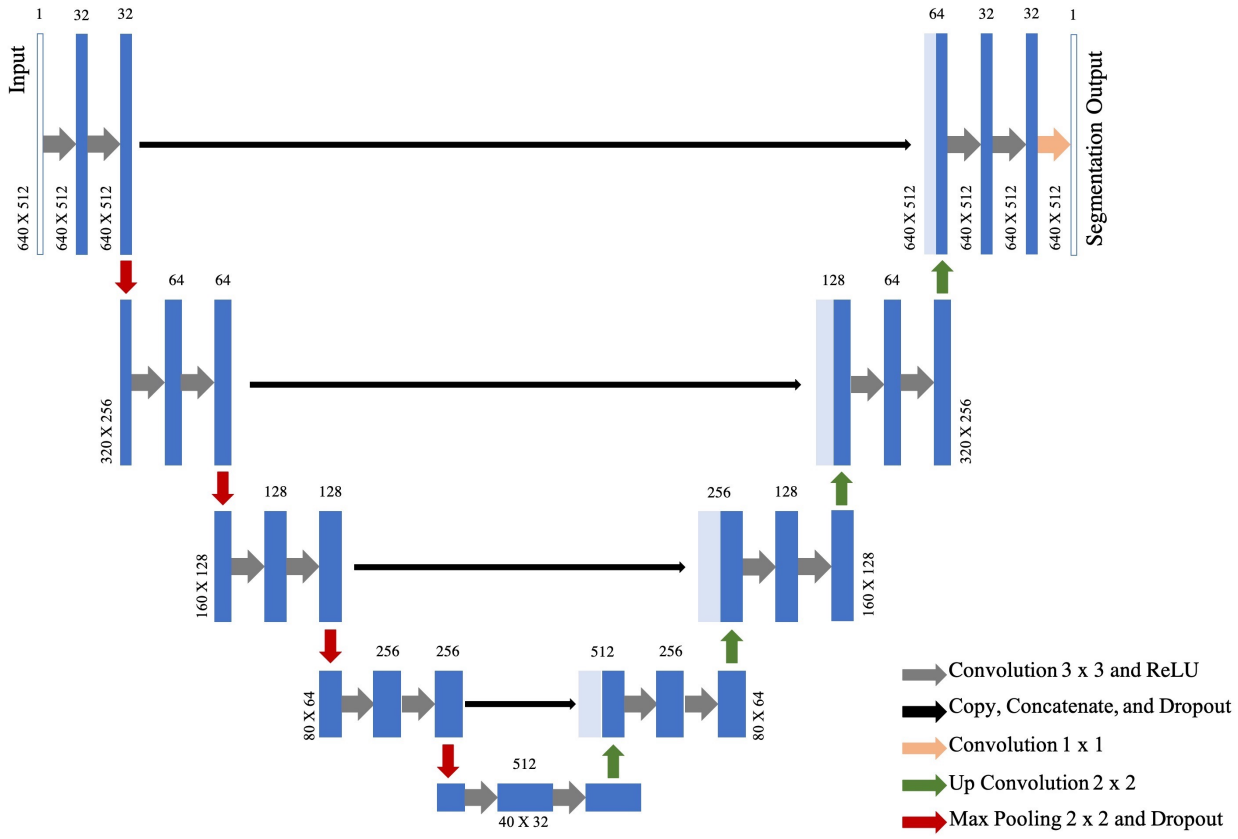

Figure 2.7: U-Net model architecture adapted from [3]. Each blue and white box denotes a multi-channel feature map and a copied feature map, respectively. The top of the box corresponds to the number of channels, and box side edges correspond to layer size. Arrows represent different operations listed in the figure.

inputs. It does not have high computational or memory requirements generally associated with 3D data processing as it needs fewer annotations and training samples to learn quickly. Thus, the U-Net-based deep learning algorithm can successfully generalize and segment MSCs cultured on monolayer and microcarriers. Lastly, reliable segmentation will aid estimation of total cell count and confluency for multidimensional MSCs cultures.

2.5.3 Extraction of human-engineered features for monolayer and 3D cultured MSCs

Feature engineering is the process of applying domain knowledge to compute features (characteristics, properties, attributes) from raw data using which exploration or prediction is carried out [127]. These features greatly influence the prediction outcome. Thus, measuring appropriate

features from segmented MSCs is essential to analyze them further. Human-engineered features with the potential to describe cell size and shape include area, perimeter, solidity, circularity, axis length, ellipticity, extent, and many others [128]. Some problem-specific features based on intensity profile and texture can also capture cell form [129]. A single feature would not be sufficient to retrieve information about cell characteristics from images, and a large number of features would be needed for further multivariate statistical analysis. A subset of the features that best describe MSC phenotypes with minimum redundancy would be used to separate RS and SR phenotypes. Similar features will be extracted in 3D to understand how they can serve as empirical descriptors of efficacy for MSCs grown on microcarriers.

2.5.4 Classification of monolayer cultured MSCs based on morphological phenotype

A classifier in machine learning is an algorithm that automatically orders or categorizes data into one or more sets of classes. For monolayer cultures, supervised machine learning techniques will suit the classification task best as cell labeling by culture experts is available. The classifiers will utilize features extracted during training to understand how they relate to MSC phenotype classes. Though there is a wide range of supervised machine learning methods, no single algorithm works best on all problems. Moreover, predictive modeling is highly dependent on the size and the structure of the dataset. As a result, it would be necessary to select a machine learning technique that is appropriate for distinguishing between RS and SR cells.

Initially, basic linear classifiers, viz. linear discriminant analysis (LDA), logistic regression (LR), and linear kernel support vector machine (LSVM), will be trained as binary classifiers to identify MSC phenotype. These methods make a classification decision based on a linear combination of feature vectors and reach accuracy levels comparable to nonlinear classifiers while taking less time to train and use. In cases where linear classifiers fail to perform as expected, other nonlinear classifiers such as K-nearest neighbor (KNN) and radial-basis kernel support vector machine (RSVM) would be modeled as they generalize better. Robust classification will provide the proportion of efficacious cells for monolayer cultures, a key parameter to monitor culture quality.

3. PRELIMINARY STUDY OF MORPHOLOGICAL CELL IMAGE ANALYSIS FOR MONITORING MONOLAYER CULTURED MESENCHYMAL STROMAL CELLS*

3.1 Introduction

The research presented in this chapter aimed to validate the applicability of image analysis for extracting relevant cellular features from phase-contrast images automatically and non-invasively [130]. The algorithm had the potential to segment the MSCs, and classify them based on their morphology to indicate their phenotype and the viability of the cell population [118].

3.2 Methods

Following the image acquisition of MSCs, the first step of the algorithm was pre-processing to adjust contrast and compensate for intensity non-uniformities that could not be avoided during image acquisition. Segmentation processes were used to locate and outline cells in images. This step was followed by the extraction of several human-engineered features to characterize cell morphology. The final step involved statistical analysis for classifying cells based on their phenotype [Fig. 3.1]. A subset of the dataset detailed in section 2.3.3 comprising 15 images with 67 cells was used in this preliminary study to train and evaluate the algorithm.

3.2.1 Image pre-processing and cell segmentation

The segmentation and classification steps of the algorithm are described in Fig. 3.2. Images were analyzed using a comprehensive set of reference-standard algorithms provided by the Image Processing Toolbox in MATLAB. Segmentation was simplified via pre-processing [Fig. 3.2(b)] by converting the input RGB image to gray-scale and increasing contrast to better distinguish cell regions from the culture substrate and marking them as foreground. Sharpening was performed using unsharp masking to strengthen small image features such as cell boundaries, to improve

*Reprinted with permission from "Morphological cell image analysis for real-time monitoring of stem cell culture" by S. M. Mota, R. E. Rogers, A. W. Haskell, E. P. McNeill, M. L. Giger, R. R. Kaunas, C. A. Gregory, K. C. Maitland, Proc. SPIE 10883, Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XXVI, 108831I, 2019, Copyright 2019 by International Society for Optics and Photonics.

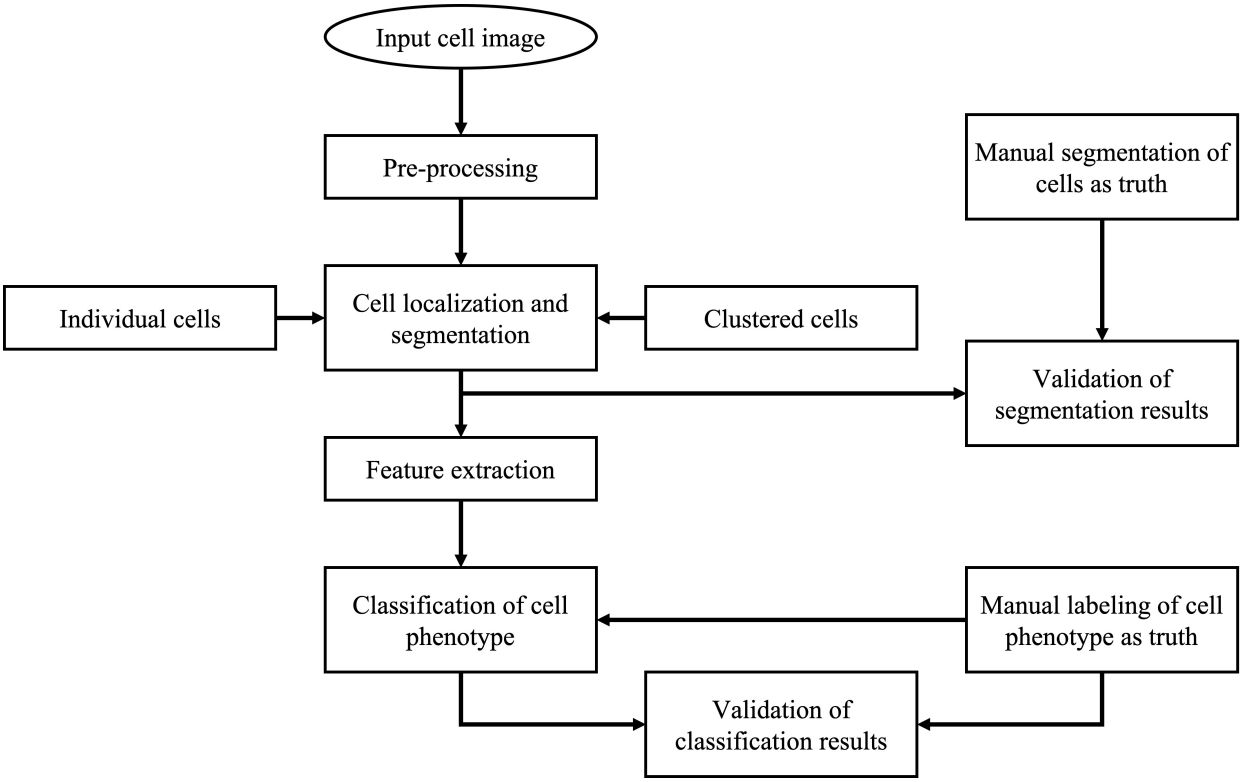

Figure 3.1: Flow diagram for the image analysis algorithm.

segmentation.

Automated cell image segmentation consisted of cell localization and boundary detection [131]. A Sobel filter was used on the pre-processed images for edge detection to obtain region proposal masks containing cells [Fig. 3.2(c-d)]. Seeds (candidate cell markers) were located within each mask region by detecting local minima [Fig. 3.2(e)]. Masks with one seed were marked as individual cell regions, while those with multiple nuclei were marked as clusters due to different segmentation strategies of the algorithm [119].

Sobel filter aided with the segmentation of individual cells by detecting edges at the points where the intensity gradient of the image was maximum. These edges were dilated to fill complete cell bodies, followed by morphological opening and closing to remove substrate pixels that were misidentified as cell objects. Finally, thresholding using the size of detected objects was performed to exclude objects in the image around the cell, such as debris. Thus, only the individual cell bodies

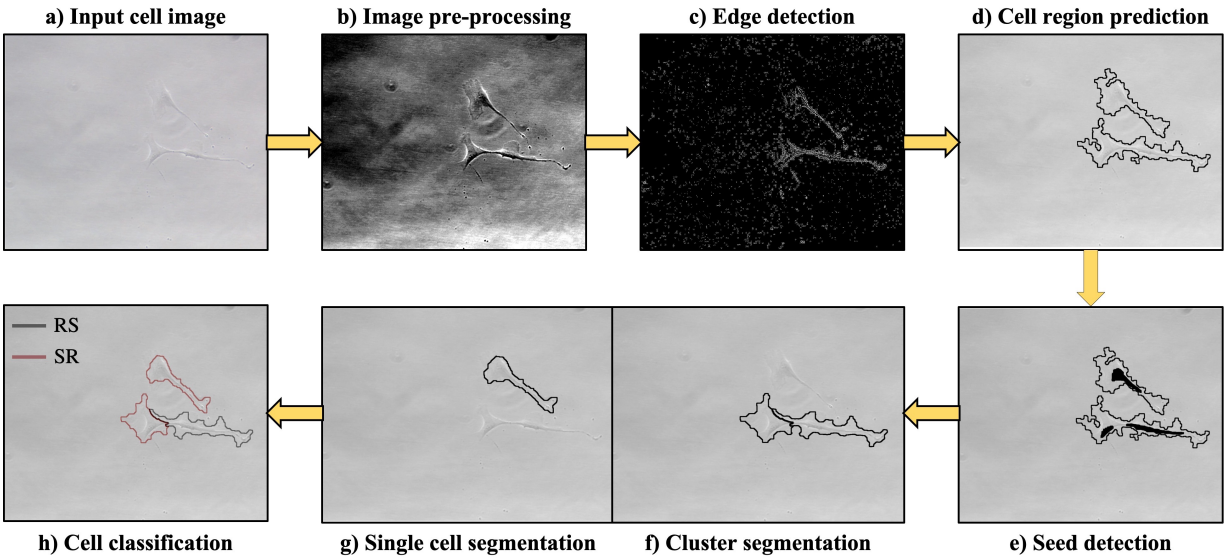

Figure 3.2: Figure showing pipeline of the image analysis algorithm. a) Input phase-contrast microscope image. b) Input image pre-processing by contrast adjustment and edge sharpening to make cell bodies more detectable. c) Cell edge detection using their gradient. d) Cell region prediction within the image using the detected edges. e) Seed (candidates cell markers) identification to distinguish each region proposal as either an individual cell or a cluster of cells. f) Clustered cell delineation using the marker-based watershed technique. g) Individual cell segmentation with the edge-based method. h) Classification of MSCs as either RS or SR phenotype using features extracted from segmented cells for logistic regression method.

were segmented [Fig. 3.2(f)].

The algorithm used a marker-based watershed segmentation technique to segment cells within clusters [132]. The watershed algorithm worked per intensity layer instead of per neighbor layer. Candidate markers identified in the previous step were used as markers. These seeds gave the position and number of cells within each region proposal for clustered cells. The intensity gradient magnitude of the proposed cluster regions was computed and modified, so regional minima only occurred at seed pixels. The watershed borders were then built at the maxima of the gradient magnitude to segment cells within the cluster [Fig. 3.2(g)].

The results of cell detection and segmentation were evaluated using accuracy and the Sorensen-Dice similarity coefficient [133], respectively. Accuracy, an object-based metric, was defined as the fraction of cells correctly detected relative to manual segmentation. DICE coefficient, a pixel-

based metric, measured the agreement between algorithm output (X) and manual interpretation (Y) via equation (3.1). Values of 0 and 1 signified no and complete overlap, respectively.

$$DICE = \frac{2\times \mid X \cap Y \mid}{\mid X \mid + \mid Y \mid}, \tag{3.1}$$

3.2.2 Feature extraction

Features with the potential to describe cell morphology such as circularity, rectangularity, minor axis length, major axis length, aspect ratio, and intensity profile were computed for all segmented objects [134]. Circularity and rectangularity features were computed using equation (3.2) and equation (3.3) with the area, perimeter, width, and height values obtained for segmented cells.

$$Circularity = \frac{4 \times \pi \times Area}{Perimeter^2}, \tag{3.2}$$

$$Rectangularity = \frac{Area}{Width \times Height}, \tag{3.3}$$

3.2.3 Classification of mesenchymal stromal cells based on their morphological phenotype

Cells segmented from two cultures were employed for training the classifier model, while cells from the third culture were used to validate classifier performance. The training was carried out only for cells that achieved a DICE score with ground truth of over 0.25 [135] to ensure that the classification model was not trained using features from objects that were not cells. Due to this cut-off, five incorrectly segmented user-defined cells were not utilized for training the classifier.

The extracted features were normalized and transformed for training a conventional LR classifier to distinguish RS and SR phenotypes. The algorithm classified the cell phenotype based on the probability output given by the sigmoid function of the classifier model.

The performance of the LR model for classifying phenotype was validated using three metrics. Classification accuracy was defined as the fraction of cells correctly classified based on their phenotype. AUC, the area under the receiver operating characteristic (ROC) curve [136], was used to measure the model's capability in distinguishing between RS and SR phenotypes. F-measure was

the weighted average of precision (P) and recall (R) as shown in equation (3.4). Precision was the fraction of cells classified as RS that were RS, and recall was the fraction of cells correctly classified as RS.

$$F - measure = \frac{2 \times P \times R}{P + R}, \tag{3.4}$$

3.3 Results

Cell detection accuracy and average DICE scores for MSCs segmented with the presented method are shown in Table. 3.1. This technique was able to correctly identify cells with over 90% accuracy and precisely segment over 70% of the manually-defined cell area.

Table 3.1: Cell detection accuracy and average DICE score for cell segmentation.

	RS phenotype	SR phenotype
No. of cells	37	30
Detection accuracy %	91.89	93.33
Average DICE score	0.727	0.715

Potential segmentation errors [Table. 3.2] included missed segmentation (number of user-defined cells not segmented), false segmentation (number of segmented regions that were not user-defined cells), under segmentation (number of cells not segmented within a cluster), and over segmentation (number of extra cells segmented within a cluster).

Table 3.2: Segmentation error rates.

No. of cells	Correct segmentation	Missed segmentation	False segmentation	Under segmentation	Over segmentation
67	62 (92.54%)	1 (1.49%)	1 (1.49%)	2 (2.99%)	0 (0%)

Table 3.3: Performance metrics of the algorithm for classification of MSCs based on their morphological phenotype.

	Training		Testing	
	RS Phenotype	SR Phenotype	RS phenotype	SR phenotype
No. of cells from culture 1	15	12	-	-
No. of cells from culture 2	8	7	-	-
No. of cells from culture 3	-	-	11	9
AUC $\pm$ s.e.	0.98 ± 0.02		0.87 ± 0.08	
Sensitivity	0.957		0.818	
Specificity	0.947		0.889	
F-measure	0.957		0.857	
Classification accuracy %	95.24		85.00	

Table. 3.3 shows the various performance metrics of the classifier, and ROC curves used for AUC analysis are shown in Fig. 3.3. Only 62 out of 67 cells were used for classification due to the previous segmentation errors. The classifier was able to identify phenotypes for over 90% of cells correctly with an AUC of 0.98 during training. Accuracy and AUC values were reduced for the

testing set by less than 10% and 0.1, respectively. This drop could be attributed to the low sample number in the testing cohort but may also have been due to the number of samples in the training cohort.

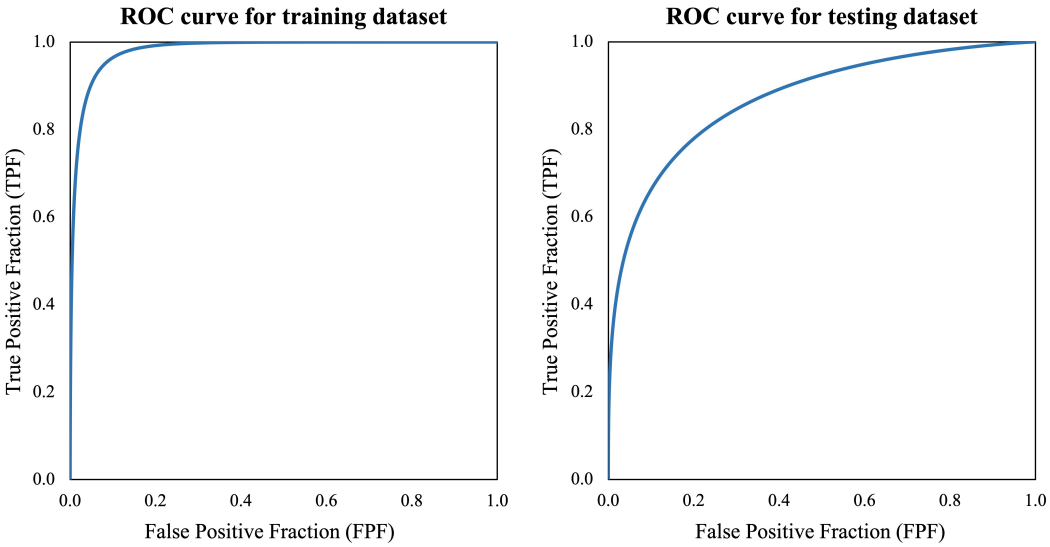

Figure 3.3: Performance of the logistic regression model for classifying morphological phenotype of MSCs was validated using AUC $\pm$ s.e. calculated as 0.98 ± 0.02 and 0.87 ± 0.08 for training and testing data sets, respectively, from the displayed ROC curves.

3.4 Discussion

The performance of the algorithm is dependent upon the quality of images used for training. It is crucial to consider challenges arising from physical and biological factors. To address imperfections in image quality and make the algorithm generalized and robust, it will need to be trained with a larger dataset consisting of images from several MSC cultures. Furthermore, images used for this analysis were acquired two days after expansion. Further research is required utilizing images taken at moderate (approximately 50%) and high (approximately 80%) confluence from the same cultures to extend the algorithm through the lifespan of monolayer cultures.

The main focus of the technique was to assess the viability of MSCs employing morphological phenotype for classification. Correct cell body segmentation was sufficient for accurate feature

extraction and classification; therefore, cell boundary optimization was not required for segmentation. The algorithm extracted morphological features that have been determined empirically to be viable indicators of the physiological state of MSCs from over 20 years of research. These features helped in classifying the phenotype of each cell and predicting their downstream characteristics. Based on the phenotype of the cells obtained from the image analysis, MSC manufacturers would be able to predict the efficacy and functionality of the recovered cells. The developed approach did not involve invasive processing, facilitating the implementation of live monolayer culture monitoring. With continuous monitoring and automation, modifications to the culture conditions could be employed in real-time to enhance reproducibility. Furthermore, automated image acquisition and analysis would replace the tedious exercise of manual visual inspection, making the process objective, facilitating standardization and quantification, and improving the definition of quality release criteria.

In conclusion, we present a computational method for the quantification of critical measures of MSC morphology. The system is automated, non-invasive, and quantitative, proving the capability of image analysis to replace the current standard of manual visual inspection and subjective evaluation for monolayer cultured MSCs.

4. AUTOMATED MESENCHYMAL STEM CELL SEGMENTATION AND MACHINE LEARNING-BASED PHENOTYPE CLASSIFICATION USING MORPHOMETRIC AND TEXTURAL ANALYSIS*

4.1 Introduction

Building on our preliminary work reported in chapter 3, the research reported here presents an integrated approach to segment and classify MSCs in phase micrographs, potentially providing automated analysis of monolayer culture viability. This was achieved by the development and evaluation of:

- an algorithm to localize and segment individual MSCs and MSCs in clusters from images of low and moderate cell density, and
- a machine learning model using morphological and textural features extracted from segmented cells to distinguish between RS and SR phenotypes of MSCs.

Here, segmentation of MSCs was handled as a three-step approach, where it first localized regions in the image that contain cells, then detected algorithm-defined markers, and finally integrated the regions with markers to segment individual cells inside clusters. The individual cells obtained with the algorithm could be analyzed further to draw conclusions about the culture population. In addition to segmenting MSCs, features that were potential indicators of the physiological state of the MSCs were also computed. These features were used by a machine learning model to classify the phenotype of each cell as RS or SR. Based on the phenotype of the cells from culture images, the algorithm would be able to provide the proportion of maximally efficacious cells in the culture. Thus, the developed image analysis protocol is novel in its contribution to automated and rapid image-based processing to objectively examine the efficacy of adherent MSCs cultures.

*Reprinted with permission from "Automated mesenchymal stem cell segmentation and machine learning-based phenotype classification using morphometric and textural analysis" by S. M. Mota, R. E. Rogers, A. W. Haskell, E. P. McNeill, R. R. Kaunas, C. A. Gregory, M. L. Giger, and K. C. Maitland, Journal of Medical Imaging, vol. 8, no. 1, pp. 1–20, 2021, Copyright 2021 by International Society for Optics and Photonics.

Another innovative aspect of this work is the implementation of a comprehensive top-to-bottom computer vision pipeline to identify MSCs and predict their relationship to RS or SR morphological phenotypes. Moreover, its potential to replace or augment visual inspection would make cell culture evaluation rapid, quantitative, and less tedious, rendering it beneficial for scale-up of cell manufacturing. Beyond validation of the applicability of this image analysis algorithm for cell quality control, an advantage of this work is its promise for streamlining culture processes for cell therapy development and manufacturing.

4.2 Methods

4.2.1 Overview

The image analysis approach was developed to classify mesenchymal stem cell (MSC) phenotype using phase-contrast micrographs of monolayer culture. The dataset used in this study is described in section 2.3.3. Figure 4.1 describes the overall flow of the method; each step is detailed in the subsequent sections. Following pre-processing, the algorithm estimated cell density. Morphological operations and thresholding detected regions of the image that contain cells. Candidate markers were localized within these regions to identify if the cell region was a single cell or cell cluster. A cell cluster was further segmented to identify individual cells. After the segmentation of each cell, the algorithm extracted several human-engineered morphometric as well as textural features. Cell segmentation and feature extraction algorithms were built using the comprehensive set of reference-standard algorithms provided by the Image Processing Toolbox in MATLAB 9.5 (R2018b). Machine learning classifiers were trained using these features to distinguish between the RS and SR phenotype. Classification models were developed in Jupyter Notebook 6.0.1 using Python 3.5.6 libraries.

4.2.2 Image pre-processing and area fraction estimation

The input RGB phase-contrast micrograph was converted to grayscale (I_{gray}), shown in Fig. 4.2(a), and pre-processed to reduce the effect of undesired imperfections introduced during imaging. Contrast of I_{gray} was adjusted to increase the intensity variation between the cells and the substrate,

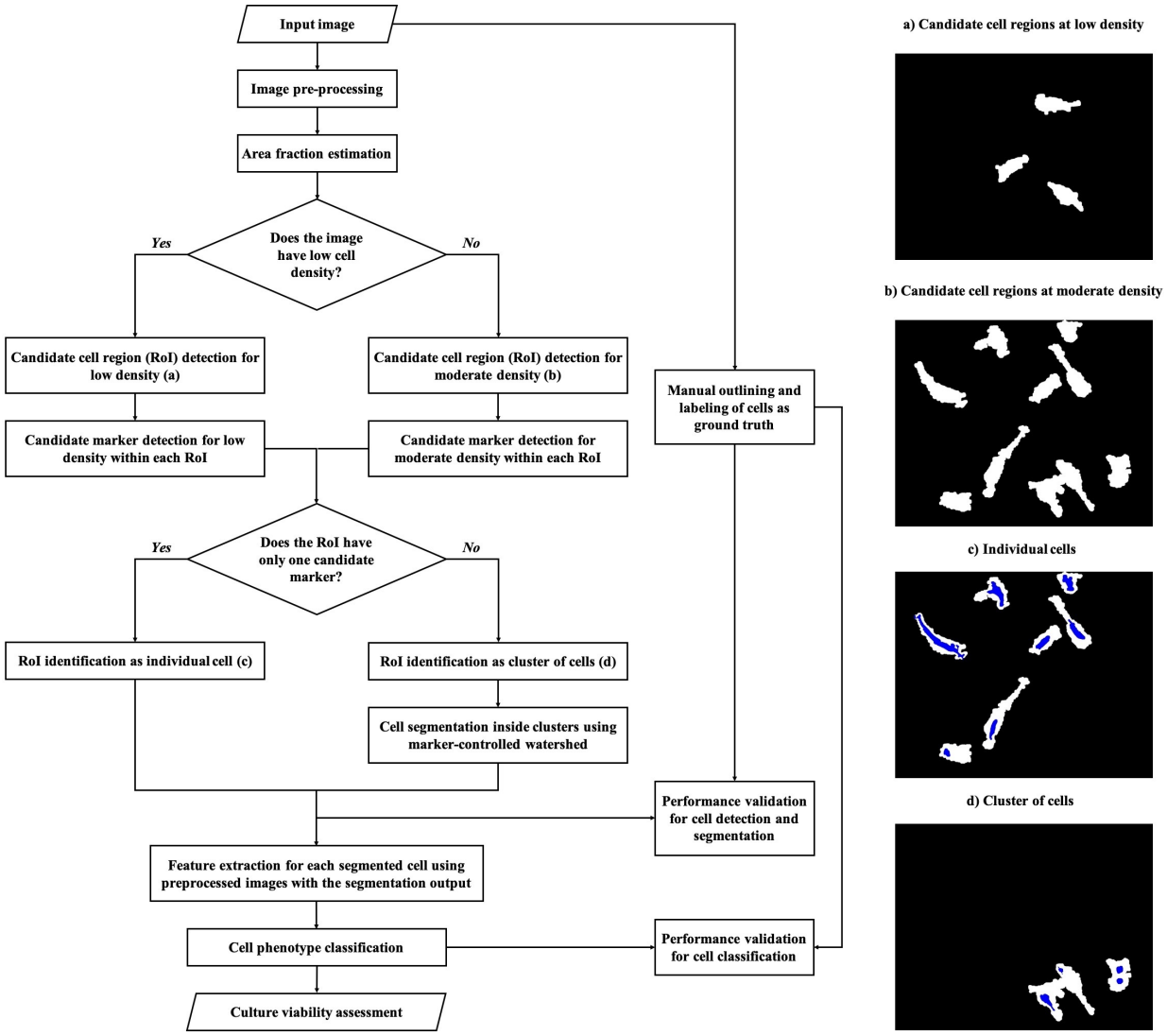

Figure 4.1: Pipeline of algorithm to classify mesenchymal stem cells in phase-contrast micrographs. Example images illustrate (a) low and (b) moderate density. Cell regions of interest (RoI) in image (b) include (c) individual cells and (d) clusters of cells, differentiated by the number of candidate markers (blue) inside RoIs (white).

making cell regions more detectable for segmentation. Cell edges were sharpened using unsharp masking and then filtered using anisotropic median-diffusion to remove unwanted artifacts while improving the signal-to-background ratio without distorting edges [Fig. 4.2(b)] [52]. The pre-processed image ($I_{pre\text{-}processed}$) was further processed by Sobel filter to identify cell boundaries in the images [Fig. 4.2(c)]. Sobel operator highlighted regions with maximum intensity change, detecting edges above a sensitivity threshold of 1. Once object outlines were obtained, dilation and closing were performed to connect the detected edges. This was followed by flood-fill operation to remove holes from the filled regions.

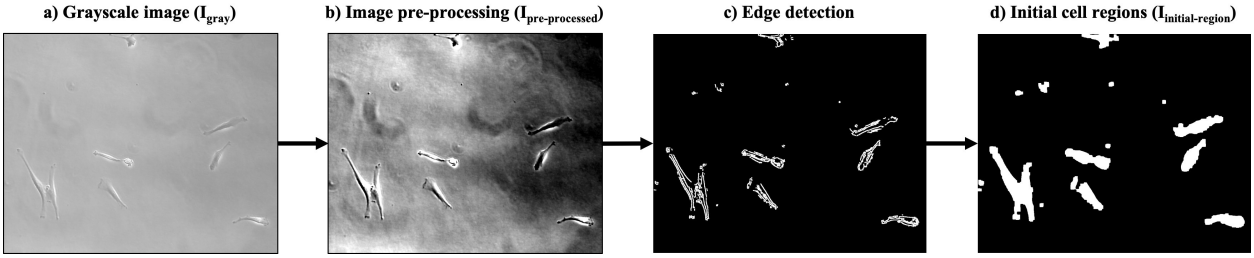

Figure 4.2: Example image of MSCs shows the steps involved in detecting initial cell regions. a) The input phase-contrast micrograph was converted to grayscale and b) pre-processed using contrast adjustment, sharpening, and anisotropic filtering. c) Edges were detected using Sobel operator. d) Edges were connected and filled using dilation, closing, and flood-fill operation to obtain initial cell regions ($I_{initial\text{-}region}$).

Using the same structuring element sizes for morphological operations to process images with a high number of cells and images with a low number of cells led to poor segmentation performance. Apart from the number of cells and clusters, the variances in their size and shape also become significant as the culture grew over time. Using area fraction as a deciding factor for cell density addressed this problem as different parameters were used for low and moderate density levels to identify cell regions as well as the markers inside them. Also, every image was evaluated based on their density estimate rather than through a potentially erroneous assumption that duration in culture was a robust predictor of density. The cell density-based criterion automatically triggered optimizations in the algorithm to ensure comparable performance at a range of cell density levels.

Area fraction (AF) of binary image ($I_{\text{initial-region}}$) was used to obtain an estimate of the input micrograph's cell density. It was calculated by the algorithm as the percentage of white pixels in the image as given by

$$AF = \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} I_{initial-region}(x, y)}{M \times N}, \tag{4.1}$$

where $I_{\text{initial-region}}(x, y)$ of size $M \times N$ pixels had a value of 1 for pixels belonging to the detected initial cell regions and a value of 0 for the background pixels. Based on the training dataset, an area fraction of 0.1 was selected as a threshold for the algorithm to decide if an image was less dense ($<$ 0.1) or moderately ($\geq$ 0.1) dense. In case an image had regions of both low and moderate density, the algorithm would estimate it to be less or moderately dense depending on the predominant region present. An image with a larger region of moderate density is more likely to have an increased AF and be handled as moderately dense. The less dense cells present would easily be detected as the thresholds for such images were set to facilitate complex cell segmentation. It should be noted that the thresholds used for moderately dense images would not be ideal for images with only less dense cells as it might lead to a greater number of false detections. On the other hand, regions of moderate density in images classified as less dense would likely be a small cluster of cells. The low density images have optimal thresholds for detecting markers in such clusters to separate individual cells.

4.2.3 Candidate cell region detection

Cell region detection was conducted as a semantic segmentation [137] step to identify pixels belonging to cells as defined in the truth, and hence, regions detected may contain more than one cell. Figure 4.3 demonstrates the process of candidate cell region detection. For less dense images, $I_{\text{initial-region}}$ was used directly to define regions with potential cell objects. Pre-processing was designed to detect objects with high sensitivity [Fig. 4.3(a)]; therefore, each object was evaluated to remove image artifacts and identify candidate cell regions using thresholding [Fig. 4.3(b)] and morphological operations [Fig. 4.3(c)]. For size thresholding, the detected object was removed

if the area was less than a threshold value determined by the minimum, maximum, mean, or the standard deviation of the area of all foreground objects in the image. This adaptive approach ensured that the threshold values were not overly biased towards the training set as it took relative object sizes in each image into account to understand if it was likely to be a cell or not. Similarly, the intensity-based thresholds were calculated using the maximum and minimum intensity values inside that object obtained from the pixel positions in $I_{\text{pre-processed}}$. Objects labeled as cells in the training set contained bright pixels in the cytoplasm and/or dark pixels inside the nucleus after contrast adjustment. Lack of both indicated that the object was not a cell as it had a relatively uniform intensity range similar to substrate. For shape thresholding, circularity and ellipticity features of the object were calculated. From the training data, circularity (mean $\pm$ s.d.) of MSCs and phase imaging artifacts were found to be 0.43 $\pm$ 0.18 and 0.82 $\pm$ 0.05, respectively, where 1 represented a perfect circle. Ellipticity was measured as

$$Ellipticity = \frac{M1 - M2}{M1}\,, \tag{4.2}$$

where $M1$ and $M2$ were the major and minor axis lengths, respectively, of an ellipse having the same normalized second central moment (variance) as the object. Ellipticity (mean $\pm$ s.d.) of MSCs and very thin artifacts such as fibers or strands have been found to be 0.46 $\pm$ 0.15 and 0.79 $\pm$ 0.04, respectively, where 1 represented a line segment and 0 represented a circle. Since MSCs are not as circular or elliptical as the artifacts, objects with high circularity and ellipticity were removed from the detected cell areas. Morphological operations such as opening and erosion were applied after thresholding to refine boundaries. Finally, objects with pixels connected to the image border were removed to avoid analysis of truncated cells. As shown in Fig. 4.3(d), these steps yielded the final image ($I_{\text{cell-region}}$) with detected candidate cell regions corresponding to RoI.

For moderately dense images, the edge detection step was performed with a reduced sensitivity threshold of 0.5 to cover all cell edges. Dilation and closing were performed with different structuring element sizes to get new potential cell regions. The same thresholding methods as less dense images were carried out on these objects with different threshold values to keep them relevant for

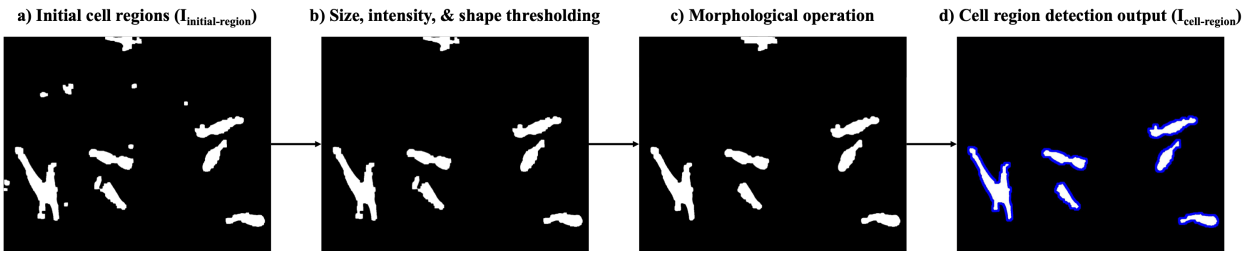

Figure 4.3: Illustration of the steps involved in identifying candidate cell regions. a) $I_{initial-region}$ obtained after edge detection and morphological operations were used as the input. b) Thresholding using size, intensity, and shape criteria removed detected objects that were not cells. c) Opening and erosion optimized the shape of cell areas. d) Clearing the image border removed incomplete cells resulting in the final candidate cell regions. (Borders of cell regions are highlighted in blue.)

images with more objects. Thresholding was also repeated more times compared to low cell density processing as a lower edge detection threshold can cause more false detections. Thresholding was followed by morphological opening, closing, and border clearing to obtain the $I_{cell-region}$ for images having moderate cell density.

4.2.4 Candidate marker detection

In phase-contrast micrographs of cells, intensity was brightest at the cell boundaries where the phase shift was maximum due to the optical path difference (refractive index and thickness) between cells and substrate, and was darker within the cell due to relative uniformity within the cell [103]. The darkest regions inside cells were taken as candidate cell markers to identify individual cells and to segment cells in clusters since each cell contains one prominent regional minimum. The image $I_{pre-processed}$ was further processed using Gaussian and median filtering to remove unnecessary noise and false local minima that don't belong to the marker. Then, contrast-limited adaptive histogram equalization (CLAHE) improved the contrast of the regional minimum. Finally, morphological reconstruction was performed using the histogram equalized image along with $I_{cell-region}$ as the mask to obtain $I_{marker-processed}$, shown in Fig. 4.4(a).

Markers were segmented using two different threshold values (A and B) for H-minima transform [138, 139, 119]. Used independently, the higher threshold value of A resulted in under-detection of markers, and the lower threshold value of B resulted in false positives; therefore, the

two were combined. First, a very high value was used as threshold A for H-minima transform to localize potential markers inside the RoIs from $I_{cell\text{-}region}$. Morphological opening and binary area opening were done to remove objects that weren't the regional minimum [Fig. 4.4(b)]. Then, minima were obtained with a lower threshold, value B. These minima outputs were dilated and closed to get the candidate markers [Fig. 4.4(c)]. In the case of zero markers from threshold A in any RoI, markers detected using B were added to that region. Markers from both the thresholds were also merged for potential cluster RoIs by using perimeter as a criterion [Fig. 4.4(d)].

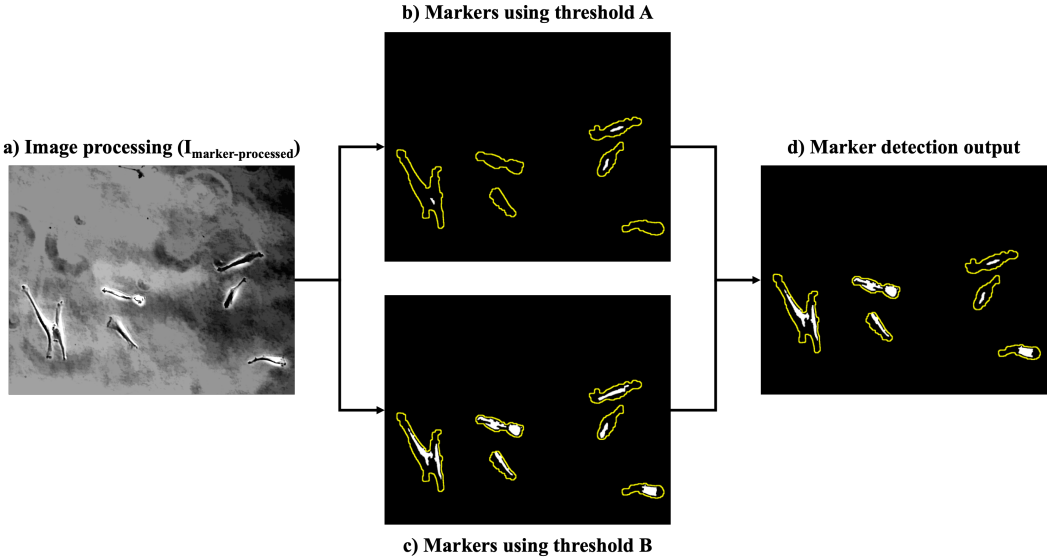

Figure 4.4: Pipeline for segmentation of markers inside the RoIs (shown in yellow) from $I_{cell\text{-}region}$. a) $I_{pre\text{-}processed}$ was processed to get an image ($I_{marker\text{-}processed}$) suitable for detecting markers. b) Markers were obtained from $I_{marker\text{-}processed}$ using H-minima transform with higher threshold A. c) Markers were obtained from $I_{marker\text{-}processed}$ using H-minima transform with lower threshold B. d) Markers from both the thresholds were merged for RoIs satisfying the perimeter threshold criterion and also for RoIs with zero markers from threshold A.

Over-detection error was managed for cell regions with more than one marker using distance thresholding along with area thresholding and morphological operations. The Euclidean distance between the centroid of the markers within a cell region was computed, and if the distance was too small, the regional maximum was over-segmented. Thus, one of the two close markers with

the smaller area was removed. Finally, dilation and erosion were performed to get rid of over-segmentations.

4.2.5 Cell segmentation and validation

RoIs and markers were combined as shown in Fig. 4.5(a) for the final instance segmentation [137] step to detect and delineate each cell in the image. A region with no marker was not considered a cell, only one marker was labeled as a single cell [Fig. 4.5(b)], and a region with more than one marker was treated as a cluster of cells [Fig. 4.5(c)]. The marker count within each cluster indicated the number of cells in the cluster. Each cluster region was segmented into individual cells [Fig. 4.5(d)] using marker-controlled watershed, which overcame the limitations of standard watershed technique, such as over and under-segmentation, by using markers [132, 140].

The algorithm's performance for cell detection and segmentation was validated using the expert-defined truth. For cell detection, the true positives (TP) were given by the number of individual cells and cells inside clusters that were correctly detected by the algorithm. The false negatives were of two types, namely missed detections (FN1), where individual cells were not detected, and under-detections (FN2), where the number of detected cells within a cluster was less than the actual number. Similarly, the two types of false positives were false detections (FP1) and over-detections (FP2). FP1 were debris or image artifacts that were erroneously detected as individual cell objects, and FP2 were the number of extra objects detected within a cluster by the algorithm. Sensitivity (S) and precision (P) were computed using

$$S = \frac{TP}{TP + FN1 + FN2}\,, \tag{4.3}$$

$$P = \frac{TP}{TP + FP1 + FP2}\,, \tag{4.4}$$

to estimate how well the algorithm was able to detect MSCs. For segmentation, Sorensen-Dice similarity coefficient (DICE) [133] was used as the metric via equation 3.1 to analyze the agreement between algorithm output and manual outlining for cell boundaries.

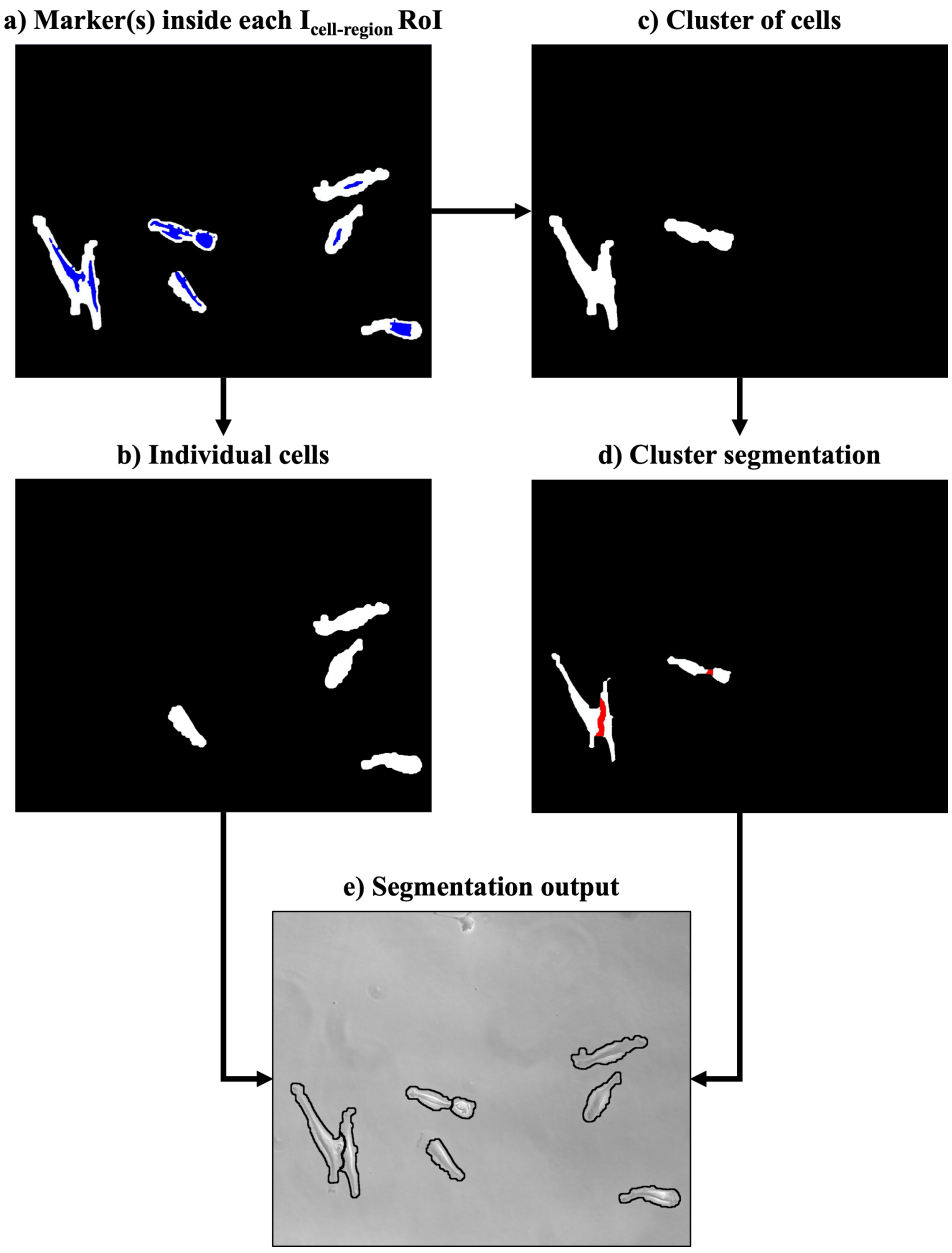

Figure 4.5: Representation of the cell segmentation steps. a) $I_{cell\text{-}region}$ was integrated with the results from marker detection (Markers shown in blue). b) RoIs with single markers were identified as individual cells, and c) RoIs with more than one marker were identified as clusters of cells. d) Marker-controlled watershed was carried out to segment individual cells within clusters (Watershed ridge lines shown in red). e) Results from the steps (b) and (d) were combined to get the segmentation output.

4.2.6 Feature extraction

False positive objects (FP1 and FP2) and all cells belonging to under-detected (FN2) clusters were excluded from the dataset to avoid training or validating classification models using incorrectly segmented objects. The final dataset used for feature extraction and classification is detailed in Table 4.1.

Table 4.1: Mesenchymal stem cell culture dataset used for feature extraction and classification.

Dataset	Culture day	No. of cells	No. of RS cells	No. of SR cells
Training (Culture 1 + Culture 2)	2	167	117	50
	4	296	182	114
Independent testing (Culture 3)	2	49	37	12
	4	91	54	37

Human engineered descriptors of the correctly detected cells were automatically extracted for the task of classifying each cell into RS or SR phenotype. A total of 30 features, consisting of a combination of size, shape, as well as first-order and second-order statistical texture measures, were computed [141, 142, 134]. The names of the features are listed in Table 4.2. The morphometric features were extracted to distinguish spindle-shaped RS cells from flattened SR cells, while texture-based features were extracted to use spatial distribution of intensity for differentiating between SR cells that are flattened and RS cells that have a more prominent phase-contrast halo around their cell body [88, 92]. Each first-order feature was calculated for segmented cell regions in I_{gray}, $I_{pre\text{-}processed}$, and $I_{reconstructed}$ giving 3 measures for each first-order feature. The second-order features were measured for the gray-level co-occurrence matrix (GLCM) of I_{gray}, $I_{pre\text{-}processed}$, and $I_{reconstructed}$. Also, each GLCM feature was computed in 24 different orientations yielding 72

measures for each second-order feature. Only one measure out of the 3 and 72 first and second-order feature measures, respectively, was selected by finding the measure with the highest AUC [136, 143] for distinguishing between RS and SR cells.

Table 4.2: Human-engineered features extracted for the segmented MSCs.

Feature type	Feature names
Size	Area, perimeter, minor axis, major axis, width, height
Shape	Elongation, compactness, circularity, ellipticity, solidity, extent
First-order features	Standard deviation, variance, intensity profile, skewness, mean intensity, balance, kurtosis, median, mode
Second-order features	Correlation, inertia, cluster prominence, energy, entropy, cluster shade, maximum probability, dissimilarity, homogeneity

The 30 features were sorted in descending order of their AUC value before computing the correlation matrix to ensure that features with higher AUCs were retained. Features with correlation greater than 0.8 were removed to reduce redundancy and optimize the computation for the classifier.

4.2.7 Cell classification and validation

The selected features were used to train linear and non-linear classifiers to find the most suitable model for our application. Features were transformed to have zero mean and unit variance before training, and the validation/test data were scaled using the training parameters. As seen in Table 2.1, the dataset was imbalanced with RS cells being more prevalent than the SR cells. To avoid poor classification of SR cells due to its low prevalence, synthetic minority oversampling technique (SMOTE) [144] was used during training.

LSVM, RSVM, LDA, KNN, and LR models were trained to classify MSCs as RS or SR phenotype. The models were trained using features from both day 2 and day 4 together, as well as day 2 and day 4 individually. As the feature correlation and relevance for data from day 2, day 4, and combination of day 2 and 4 was different, features were selected for each of them individually. It was observed that the classifier could learn better from features of cells whose images were acquired on the same day rather than the combination of features of cells from two different days. AUC values were obtained using ROCKIT software [145], and all the models were compared using the average AUC value from 5-fold cross-validation [146]. Further, ensemble classifiers based on soft-voting method were also trained with the top two classifiers for day 2 and day 4 separately [147]. Finally, the performance of the selected models for the independent test dataset was evaluated using AUC, sensitivity, and specificity. Sensitivity as well as specificity were determined by selecting a threshold that minimized $(1 - \text{sensitivity})^2 + (1 - \text{specificity})^2$ [147].

4.3 Results

4.3.1 Cell detection and segmentation

The algorithm's ability to accurately locate MSCs was evaluated using sensitivity and precision. Table 4.3 shows the breakdown of correctly detected cells and undetected cells for RS and SR cells in the training and independent testing sets. False negatives were further defined as missed detections (FN1) or under-detections (FN2) as described in Section 4.2.5. The algorithm correctly detected cells with sensitivity greater than 0.95 for both RS and SR cell phenotypes in the training set. A sensitivity over 0.8 in the independent testing set indicated the robustness of the algorithm.

Table 4.3: Cell detection sensitivity of the algorithm for training and independent testing.

Dataset	Cell phenotype	No. of cells	Correct detections (TP)	Missed detections (FN1)	Under detections (FN2)	Sensitivity (S)
Training	All	472	466	3	3	0.987
	RS	307	302	2	3	0.984
	SR	165	164	1	0	0.994
Independent testing	All	186	157	13	16	0.844
	RS	121	102	8	11	0.843
	SR	65	55	5	5	0.846

Precision was the fraction of objects detected by the algorithm that were identified as cells in the ground truth labeling of the dataset. Table 4.4 shows the number of incorrectly identified objects that were either false detections (FP1) or over-detections (FP2). The algorithm's precision for cultures used in training was above 0.95. A precision greater than 0.85 in the independent testing set demonstrated the generalizability of the algorithm to detect cells with a low false positive rate.

Table 4.4: Cell detection precision of the algorithm for training and independent testing.

Dataset	Correct detections (TP)	False detections (FP1)	Over detections (FP2)	Precision (P)
Training	466	5	8	0.973
Independent testing	157	13	13	0.858

The DICE metric was used to evaluate the overlap between the algorithm and the ground truth segmentation of MSCs. DICE scores (mean $\pm$ s.d.) for the training and testing dataset are summarized in Table 4.5 for all cells outlined in the truth as well as only for the cells that were correctly detected. The algorithm segmented over 85% and 80% of the manually defined cell areas for training and testing, respectively, regardless of cell phenotype.

Table 4.5: Cell segmentation DICE score of the algorithm for training and independent testing.

Dataset	Cell phenotype	DICE score (mean $\pm$ s.d.)	
		All cells (TP + FN1 + FN2)	Correctly detected cells (TP)
Training	All	0.875 ± 0.092	0.881 ± 0.067
	RS	0.878 ± 0.089	0.884 ± 0.061
	SR	0.870 ± 0.097	0.875 ± 0.077
Independent testing	All	0.803 ± 0.218	0.869 ± 0.082
	RS	0.814 ± 0.203	0.871 ± 0.080
	SR	0.783 ± 0.245	0.866 ± 0.085

The performance of the algorithm for cell detection and segmentation were analyzed per each cell in Tables 4.3-4.5. Sensitivity, precision, and DICE metrics of the algorithm were further examined per image from all cultures and days. The mean $\pm$ s.d. of these metrics for the training and testing images are given in Table 4.6, and it was confirmed that the algorithm could detect and segment cells consistently for each image.

Table 4.6: Cell detection and segmentation performance of the algorithm per image from training and testing datasets. Mean and standard deviations were calculated for all cells over all images.

Dataset	No. of images	Cell detection sensitivity (mean ± s.d.)	Cell detection precision (mean ± s.d.)	Cell segmentation DICE score (mean ± s.d.)
Training	71	0.991 ± 0.036	0.967 ± 0.107	0.896 ± 0.050
Independent testing	36	0.837 ± 0.207	0.861 ± 0.217	0.849 ± 0.106

Figure 4.6 shows the ground truth along with algorithm's segmentation results for two images of both low and moderate density from the test dataset. The difference in performance of the algorithm to localize cells in low and moderately dense images was also reviewed as it was trained with different parameters for these two levels of cell densities. Welch's t-test at 95% confidence level failed to show statistical difference in the values of sensitivity ($p = 0.375$), precision ($p = 0.191$), and DICE score ($p = 0.289$) for low and moderately dense cell images.

4.3.2 Cell phenotype classification

As mentioned previously, only correctly detected individual cells and cells in clusters with correct cell count were used to train and test the machine learning models for the task of classifying each cell phenotype as RS or SR. LSVM, RSVM, LDA, KNN, and LR were trained and compared using the AUC metric from 5-fold cross-validation. These five classifiers trained using object features from both day 2 and day 4 cultures were validated by testing their performance in cross-validation for a combination of cells from "day 2 + day 4", as well as day 2 and day 4 independently. It was observed that "day 2 + day 4" models did not perform as well for day 2 as they did for day 4 [Table 4.7]. This may be due to models' bias towards a higher prevalence of day 4 MSC features. Training the classifiers with features from day 2 and day 4 individually improved their performance for both days [Table 4.7]. The outputs from the top 2 classifiers (RSVM and

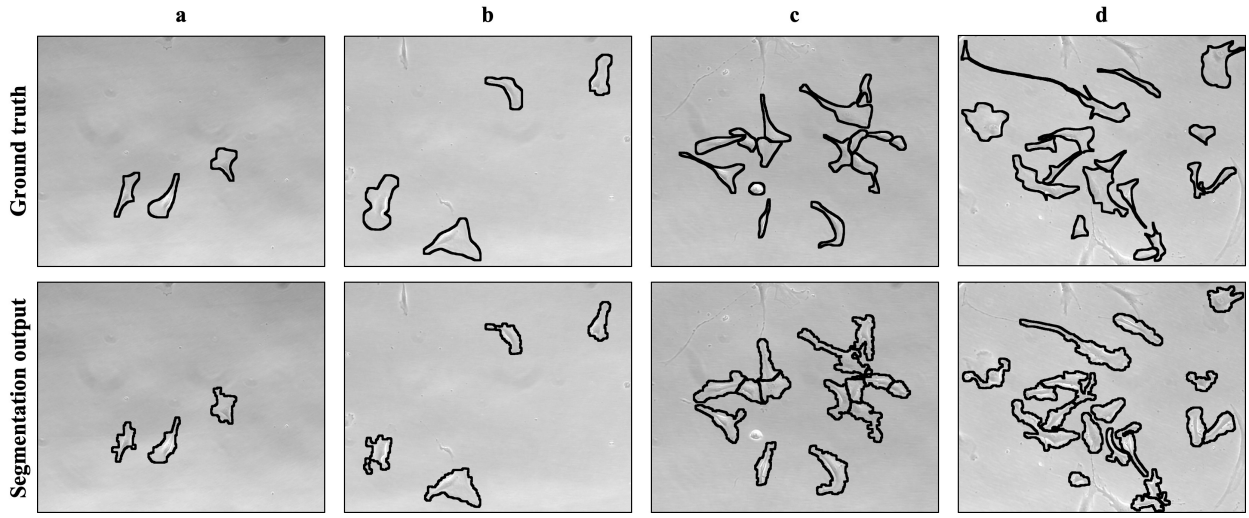

Figure 4.6: Sample images of MSCs from the independent test dataset comparing the ground truth cell outlines (top row) with the segmentation results of the algorithm (bottom row). Columns a and b are examples of low density images whose AF values were estimated as 0.04 and 0.08, respectively by the algorithm. Similarly, examples of moderate density images with AF values of 0.15 and 0.21 are shown in column c and d, respectively. The ground truth images also illustrate that the truncated cells on image borders were not considered for analysis.

LR for day 2; LDA and KNN for day 4) based on cross-validation AUC were combined for ensemble classification. Figure 4.7 shows the classifier agreement plot between the models used for ensemble classification. The x-axis in Fig. 4.7(b) was not continuous due to the discrete probability distribution of the KNN classifier. The disagreement between the top 2 classifiers for each day may be because they learn differently from the same features. "RSVM + LR" and "LDA + KNN" models being a combination of linear and non-linear approaches were able to harness the potential of both the models to make more accurate predictions. This likely explains why fusion classifiers performed better than the individual classifiers during 5-fold cross-validation, and hence, they were selected for the image analysis pipeline.

Table 4.7: 5-fold cross-validation to compare performance of models for distinguishing between RS and SR cells. AUC values are in bold for Day 2 and Day 4 classifiers that performed best during cross-validation.

	5-fold cross-validation AUC ± s.e. (Culture 1 + Culture 2)				
Training fold	Day 2 + Day 4			Day 2	Day 4
Testing fold	Day 2 + Day 4	Day 2	Day 4	Day 2	Day 4
LSVM	0.73 ± 0.01	0.67 ± 0.03	0.75 ± 0.02	0.72 ± 0.04	0.81 ± 0.03
RSVM	0.75 ± 0.01	0.68 ± 0.01	0.78 ± 0.03	0.734 ± 0.06	0.80 ± 0.02
LDA	0.73 ± 0.02	0.68 ± 0.04	0.75 ± 0.03	0.70 ± 0.04	0.83 ± 0.02
KNN	0.73 ± 0.01	0.68 ± 0.04	0.74 ± 0.02	0.72 ± 0.06	0.83 ± 0.02
LR	0.75 ± 0.01	0.70 ± 0.03	0.77 ± 0.03	0.73 ± 0.04	0.82 ± 0.02
RSVM + LR	-	-	-	**0.76 ± 0.04**	0.84 ± 0.04
LDA + KNN	-	-	-	0.75 ± 0.06	**0.86 ± 0.03**

The selected ensemble classifiers for day 2 and day 4 were further evaluated using the independent test dataset. The algorithm was able to correctly identify cell phenotypes with an AUC of 0.82 for day 2 and 0.79 for day 4. The classification models for both days were statistically proven to perform better than random chance using the 95% confidence interval for AUCs given in Table 4.8. These inferences weren't corrected for multiple comparisons as only a single statistical test was performed for each day. The fitted ROC curves obtained for the day 2 and day 4 ensemble classifiers are presented in Fig. 4.8. The algorithm achieved a sensitivity and specificity of over 0.75 for both days, further demonstrating its effectiveness in classifying MSCs based on their phenotype.

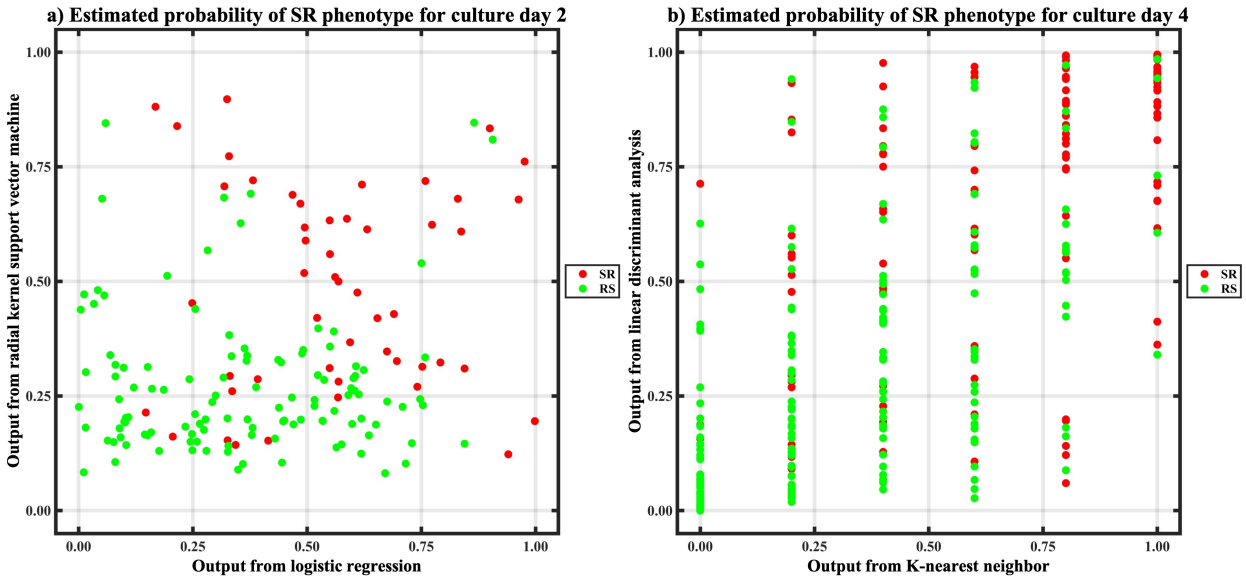

Figure 4.7: Diagonal classifier agreement plots between a) logistic regression classifier (x-axis) and radial kernel support vector machine (y-axis) for day 2 culture and b) K-nearest neighbor (x-axis) and linear discriminant analysis (y-axis) for day 4 culture. Each point represents a cell whose phenotype was predicted as RS (green) or SR (red) by each classifier.

Table 4.8: Performance of the algorithm for classifying MSCs as RS or SR for independent testing.

Culture day	Model	AUC $\pm$ s.e. (95% CI)	Sensitivity	Specificity
2	RSVM + LR	0.816 $\pm$ 0.060 (0.769, 0.886)	0.789	0.887
4	LDA + KNN	0.787 $\pm$ 0.04 (0.716, 0.851)	0.796	0.757

4.3.3 Generalizability of phenotype classification

The cells in the training and test dataset were also labeled by a group of 20 individuals with varied level of expertise in working with MSCs. The group comprised of nine people with one

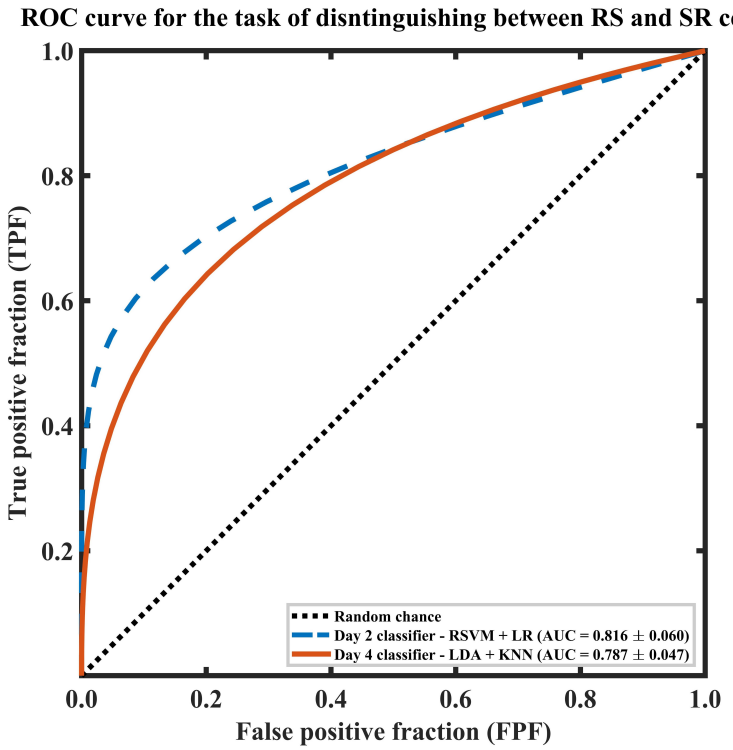

Figure 4.8: Fitted binormal ROC curves illustrating the performances of day 2 and day 4 ensemble classifiers for the task of distinguishing between RS and SR cells during independent testing. The dashed blue line represents the "RSVM + LR" day 2 classifier and the solid orange line represents the "LDA + KNN" day 4 classifier.

to three years of experience, six people with four to nine years of experience, and five people with ten or more years of experience. This analysis was performed to study the generalizability of the developed algorithm as phenotype assessment of MSCs is highly subjective. Each cell was classified as either RS (0) or SR (1) by all twenty individuals, and its average phenotype score (APS) was computed. Based on the range of average scores, the cells were split into categories such as highly RS, moderately RS, uncertain, moderately SR, and highly SR [Table 4.9]. The confidence of the group in identifying a cell as either RS or SR was represented by these five categories; where the highly RS/SR categories meant highest confidence in labeling that cell and uncertain category meant that the phenotype of that cell was almost indeterminate.

The output of the ensemble classifiers for day 2 and day 4 from training and testing set was studied to understand which category of the cells were most wrongly predicted by the algorithm. It

was noticed that the algorithm had the highest classification sensitivity (> 0.9) for highly RS and highly SR categories and the least sensitivity (as low as 0.4) for the uncertain category [Table 4.9]. This enabled us to understand that while our image analysis pipeline has the potential to predict the cell phenotype with a sensitivity very close to human interpretation, it did exhibit a similar weakness to trained humans in that it had lower sensitivity in categorizing marginal morphological characteristics. If these limitations are addressed in future iterations of the algorithm, this approach would have the potential to out-perform human observers with decades of experience.

Table 4.9: Classification sensitivity of the algorithm for the 5 cell phenotype categories.

Dataset	Culture day	No. of cells correctly classified / No. of cells as per truth (Sensitivity)				
		Highly RS	Moderately RS	Uncertain	Moderately SR	Highly SR
		APS ≤ 0.15	$0.15 <$ APS < 0.40	$0.40 \leq$ APS ≤ 0.60	$0.60 <$ APS < 0.85	APS ≥ 0.85
Training	2	10 / 10 (1.000)	9 / 11 (0.818)	5 / 12 (0.417)	11 / 13 (0.846)	3 / 3 (1.000)
	4	40 / 40 (1.000)	36 / 40 (0.900)	29 / 38 (0.763)	29 / 33 (0.879)	16 / 16 (1.000)
Independent testing	2	15 / 16 (0.937)	17 / 21 (0.809)	7 / 19 (0.368)	20 / 23 (0.869)	12 / 12 (1.000)
	4	72 / 74 (0.973)	66 / 77 (0.857)	26 / 46 (0.565)	57 / 65 (0.877)	31 / 34 (0.912)

4.4 Discussion

The image analysis method reported here is capable of segmenting and classifying MSCs based on their morphological phenotype. Segmentation results provide cell count per image, cell density (cells/cm^2), and percent confluency that indicates cell proliferation over the time course of the culture. Classification of segmented cells yields a count of undesirable SR cells (quiescent) and the ratio of these cells to viable RS cells (high potency) that serves as a vital indicator of culture quality. The promising sensitivity, precision, and DICE score for MSC localization in phase-contrast micrographs suggest that automated image-based analysis can be seamlessly integrated into the current cell culture workflow for their quantitative evaluation.

That being said, it is crucial to note that an overall evaluation of the developed image analysis approach has not been included in this study. It was only feasible to assess each stage of the algorithm separately using the available dataset and an additional independent dataset would be needed for the overall validation of the entire system. The cell detection and segmentation stages are greatly influenced by the intensity distribution, contrast, and clarity of the input phase-contrast micrographs. The classification stage in turn is dependent on the segmentation output. Upon resolution of these challenges, classifiers could be trained for more robust prediction by expanding the dataset as well as accounting for cells whose phenotype is uncertain. These challenges faced by different stages of the pipeline are discussed in detail subsequently.

The ability of the algorithm to detect and segment cells is greatly dependent upon the quality of the acquired images. In phase-contrast micrographs, high contrast is crucial to distinguish cells from the substrate. The majority of detection errors during training and testing were primarily due to poor image contrast and blur in some image regions. Fluorescence microscopy of labeled cells would provide higher contrast images and easier segmentation; however, phase contrast microscopy is the standard technique for the noninvasive evaluation of live cells. The low rate of false-positive and false-negative detections has minimal impact on the overall quality assessment as long as a sufficient number of images are acquired to capture a population of correctly segmented cells.

The classification results were studied to understand factors that affect the performance of proposed models. In this study, classifiers were trained using the truth defined by a biologist with 15+ years of experience working with MSCs. The majority of MSCs incorrectly classified were identified as cells undergoing differentiation from RS to SR phenotype. Additionally, phenotype labels were obtained for the entire dataset from 20 trained individuals with varying levels of expertise in culturing MSCs to assess the subjectivity of human classification and the generalizability of the algorithm. This labeling enabled an analysis of the impact of ambiguity in the morphology of differentiating cells on visual inspection and classification. The trained individuals had a minimal agreement about cell phenotype during differentiation, demonstrating the uncertainty in classifying cells in this transition both for human interpretation and the algorithm. The existing binary classifier could be trained as a multiclass problem with the task of distinguishing between RS, SR, and these indeterminate cells. Identifying cells with indeterminate phenotype would decrease the number of false classifications and increase certainty in the prediction of RS and SR phenotypes. However, more data would be necessary to train the machine learning model to predict the indeterminate class effectively. An alternative would be to calibrate the binary classifier that differentiates RS from SR cells for correlating its probability output with confidence in a cell's phenotype. The future endeavor would be fine-tuning the current approach by adopting the above-mentioned machine learning methods to improve classification robustness.

Classification performance was only validated for cells that were correctly detected and not evaluated for false-positive objects. Since this technique would be implemented in real-time as a pipeline where every segmented object would be classified, it is necessary to assess the effect of classifying incorrectly detected objects. Though not in the scope of the research discussed in this chapter, it may also be interesting to analyze if the algorithm tends to classify false positive objects as RS or SR cells. Images from three different cell cultures have been used for the dataset here. This algorithm can be refined and tested using a broader set of images from MSC cultures generated in other laboratories and obtained with different phase-contrast microscopes. The algorithm was trained to predict cell phenotype from day 2 and day 4 using two different classifiers. This was

done because a "day 2 + day 4" model was biased against cells from day 2 due to limited data as compared to day 4 when more cells were present. Additional data from more cultures may enable sufficient features for day 2 cells to train a common classifier with greater prediction capacity, removing the need for two separate classifiers. Apart from this, a larger dataset will not only increase data variability for algorithm training but also enable characterization of the capability to predict the quality of cultures during practical application. The algorithm's performance can also be compared to standard culture evaluation assays, many of which are time-consuming and labor-intensive, to evaluate quantitative phenotypic analysis as a measure of replication potential for MSC cultures. These considerations do not include the considerable amount of time, effort, and expense associated with training expert observers.

The focus of the research presented in this chapter was to prove the applicability of image-based analysis for non-invasive and objective determination of MSC phenotype in low and moderately dense cultures. The algorithm is not expected to perform well for highly dense cell images where there is a lot of cell overlap making even visual investigation complicated. Quantitative evaluation of the earlier stages of culture is more critical for monitoring the health of the culture. MSC cultures are typically harvested or passaged before high confluency. This methodology has the potential for being extended to:

- Forecast the percentage of non-functional MSCs in a culture for a future time based on the ratio computed for the present and past time points.
- Quantitatively represent the influence of change in culture protocol over the population of putatively efficacious and non-functional MSCs in culture.
- Monitor the ratio of non-functional to putatively efficacious MSCs as a function of confluency and cell density in addition to time.
- Estimate efficacy of other stem cell cultures for their prospective use in cytotherapies using image-based morphological analysis.

In summary, we have shown that the presented morphological image analysis can segment and classify MSCs to quantify the efficacy of monolayer cultures. As this computational pipeline is completely non-invasive, it enables continuous monitoring of culture conditions to enhance reproducibility. It is anticipated that this algorithm will facilitate biologists and cell manufacturers to draw conclusions about the functionality of recovered MSCs. The proposed solution with automated imaging leads to rapid, quantifiable, and standardized MSC quality control processes. It could be incorporated into high-volume stem cell manufacturing to pave way for efficient cell therapies to treat chronic diseases.

5. U-NET-BASED IMAGE SEGMENTATION OF MESENCHYMAL STROMAL CELLS*

5.1 Introduction

U-Net has been adopted for a variety of biomedical problems including, brain tumor detection, breast mass segmentation, and diabetic retinopathy [148, 149, 150]. Owing to its high computational efficiency and ability to learn from limited training data, we implemented a U-Net architecture-based pipeline to segment MSCs cultured in monolayer conditions. The research goal outlined in this chapter was to study and analyze how a standard U-Net can be adapted and optimized for more accurate and efficient segmentation of MSCs. The novelty of the work presented here is the application of deep learning for segmenting MSCs to facilitate their potency assessment. Thus, our presented algorithm bridges the gap between the current technology and the need for robust segmentation of MSCs.

5.2 Methods

The image dataset of monolayer MSC culture detailed in section 2.3.3 was applied to train and evaluate the presented U-Net deep learning model. The dataset consists of 71 images with 472 MSCs for training plus validation and 36 images with 186 MSCs for independent testing. Figure 5.1 shows a flow diagram of the U-Net-based image analysis method presented in this chapter. The input images of MSC monolayer cultures were acquired using a phase-contrast microscope. The ground truth was prepared manually for training and performance validation of the deep learning model via visual inspection by a culture expert with 15+ years of experience. The input images were pre-processed and then augmented to generate sufficient data for training the U-Net architecture. The trained U-Net model was used to obtain the cell prediction maps of the culture images. Cell prediction maps were post-processed to get the final MSC segmentation output.

*Reprinted with permission from "U-Net based image segmentation of mesenchymal stem cells" by S. M. Mota, R. E. Rogers, A. W. Haskell, E. P. McNeill, R. R. Kaunas, C. A. Gregory, M. L. Giger, and K. C. Maitland, Proc. SPIE 11647, Imaging, Manipulation, and Analysis of Biomolecules, Cells, and Tissues XIX, 116470V, 2021, Copyright 2021 by International Society for Optics and Photonics.

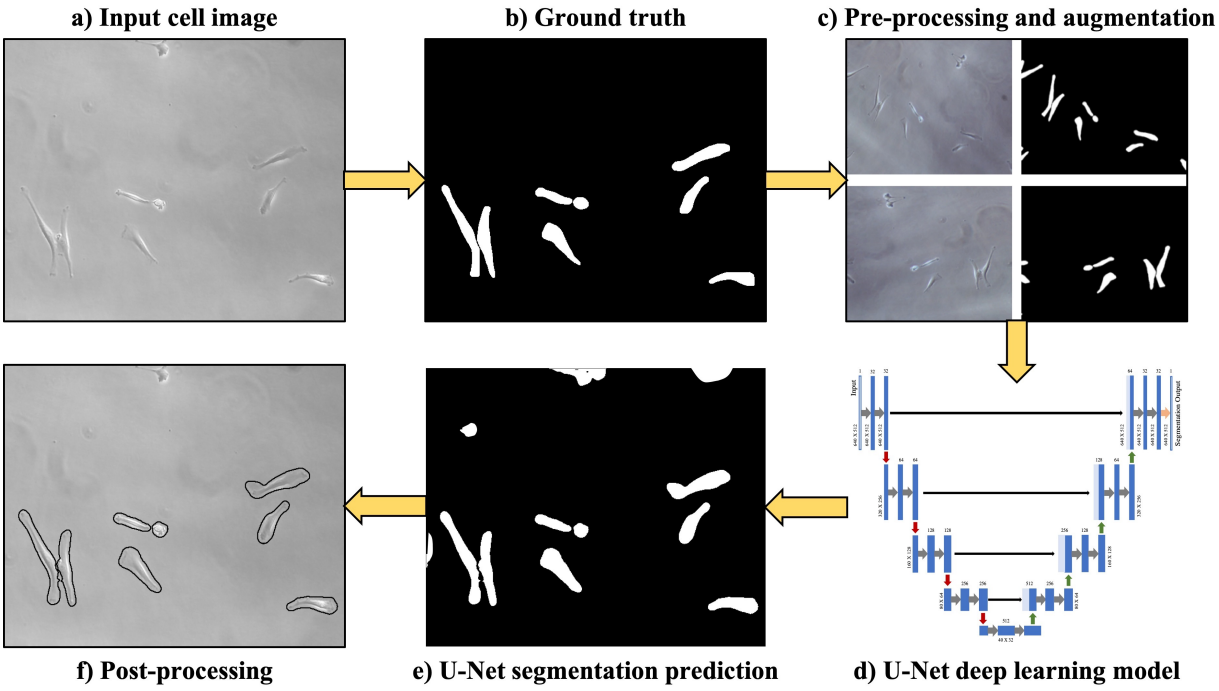

Figure 5.1: Pipeline of the proposed U-Net based mesenchymal stem cell segmentation method.

5.2.1 Pre-processing and augmentation

All the images in the training and testing dataset were pre-processed and standardized using pixel-wise centering and normalization. Pixel-wise centering scales all the image pixels to have a mean of zero. Normalization divides each pixel value by the standard deviation of all the pixel values of the image. Pre-processing leads to suppression of unwanted distortions and enhancement of relevant cell features so that the U-Net model can learn better from this improved data. Pre-processed images from cultures 1 and 2 were augmented to generate more data for training and validation. Data augmentation was essential to teach the network desired properties such as invariance and robustness when only a few training samples were available. Each image was augmented to generate 65 more images through basic image transformation operations such as rotation, width shifting, height shifting, zooming, shearing, flipping, and also by performing random elastic deformation.

5.2.2 U-Net model training

The U-Net architecture implemented for this study is described in section 2.5.2. A pixel-wise softmax was applied to the resultant image followed by a weighted binary cross-entropy loss function. The idea was that, even in segmentation, every pixel has to lie in one of the classes. Hence, this method converted the segmentation problem into a binary classification problem, performing very well compared to the traditional loss functions. The architecture used a novel weighted loss scheme for each pixel such that there was a higher weight at the border of segmented objects. This weight map was estimated using morphological operations, and it helped the network learn tiny separation regions between overlapping cells to delineate individual cells in the binary segmentation prediction.

The prepared training dataset was split randomly into training and validation to avoid overfitting the model. Adam optimizer was used to find proper network weights with a learning rate of 1×10^{-4}. The neural network was trained from scratch with initial weights drawn from a Gaussian distribution. It was run for 100 epochs, each iteration processing a mini-batch of five randomly picked training samples. The network was also assigned a callback function that reduced the learning rate when validation loss did not decrease for five consecutive epochs, saved the weights of the best performing model, and finally stopped the training when there was no improvement in the validation loss for ten consecutive epochs.

5.2.3 Post-processing and performance evaluation

It was observed from the training and the validation dataset results that there were many objects falsely predicted by U-Net as cells which greatly affected the cell detection performance. To overcome this, the segmentation map from the U-Net model was further post-processed with basic morphological operations. The objects below a particular area threshold were removed to get rid of speckles and small background artifacts. Objects at the border that were less than the size threshold were also removed to avoid truncated cells. These thresholds were optimized and fixed utilizing the training and validation images. Finally, morphological dilation and flood-fill operation were

carried out to optimize the detected cell boundaries and fill any holes present in the segmented objects.

The performance of the presented U-Net segmentation with post-processing was compared with morphological image analysis from chapter 4 as well as standard U-Net. Additionally, morphological analysis was also compared with standard U-Net to understand how it measures against state-of-the-art deep learning-based biomedical image segmentation technique. True positives (TP) were defined as the number of objects correctly detected as cells, false negatives (FN) by the number of cells not detected by the U-Net model, and false positives (FP) by the number of objects wrongly identified as cells. Sensitivity (S) and precision (P) metrics were used to assess cell detection, and DICE [133] was used to evaluate cell segmentation via equation 3.1. Values of sensitivity and precision were computed not only for all cells but also per image.

Morphological and textural features that serve as the descriptors of MSC phenotype were extracted for cells correctly segmented by morphological image analysis and the presented U-Net approach. AUC [136, 143] value was estimated from 10-fold cross-validation for each feature for the task of distinguishing between efficacious and non-efficacious cells. Finally, morphological and post-processed U-Net segmentations were compared using AUC values of their top three features for a comprehensive analysis.

5.3 Results

Detection and segmentation performances of morphological image analysis, standard U-Net, and the proposed method for training dataset are shown in Table 5.1. Performance was evaluated using sensitivity, precision, and DICE metrics. Sensitivity and precision values reported here were overall values for all cells as well per image. Agreement between segmentation output and ground truth was assessed for each image via the DICE score. Morphological image analysis performed better than standard U-Net in all of the evaluation metrics for this task. The low precision of the U-Net was determined to be due to the detection of several image artifacts as cells. Morphological analysis was able to overcome such false positive detections through various morphological thresholding steps. Although the U-Net model used weighted maps to learn the small separation

borders, it could not separate cells in clusters as effectively as the marker-controlled watershed technique used in the presented pipeline. An important observation from the training results was that post-processing effectively improved the precision of standard U-Net by almost 25%. It was also noticed that post-processing decreased the sensitivity by less than 0.3% compared to standard U-Net, which was negligible compared to the increase in precision.

Table 5.1: MSC detection and segmentation performance of morphological, standard U-Net, and presented U-Net algorithms for the training dataset.

Metrics	Morphological analysis	Standard U-Net analysis	Presented U-Net analysis
Sensitivity (S)	0.987	0.765	0.763
Precision (P)	0.973	0.653	0.818
Sensitivity per image (mean $\pm$ s.d.)	0.991 ± 0.036	0.817 ± 0.173	0.815 ± 0.173
Precision per image (mean $\pm$ s.d.)	0.967 ± 0.107	0.692 ± 0.213	0.860 ± 0.171
DICE score per image (mean $\pm$ s.d.)	0.896 ± 0.050	0.841 ± 0.056	0.852 ± 0.055

The cell detection and segmentation metrics of all the three approaches are described in Table 5.2 for the independent test dataset. It was confirmed from precision and DICE values that post-processing successfully optimized segmentation results from standard U-Net. Morphological analysis still performed best for segmenting and detecting MSCs. Additional work would be needed to modify and optimize U-Net's architecture for it to be able to localize MSCs more accurately.

Table 5.2: MSC detection and segmentation performance of morphological, standard U-Net, and presented U-Net algorithms for the independent test dataset.

Metrics	Morphological analysis	Standard U-Net analysis	Presented U-Net analysis
Sensitivity (S)	0.844	0.747	0.742
Precision (P)	0.858	0.556	0.789
Sensitivity per image (mean $\pm$ s.d.)	0.837 ± 0.207	0.772 ± 0.216	0.758 ± 0.217
Precision per image (mean $\pm$ s.d.)	0.861 ± 0.217	0.536 ± 0.210	0.806 ± 0.232
DICE score per image (mean $\pm$ s.d.)	0.849 ± 0.106	0.804 ± 0.050	0.823 ± 0.051

P-values from the paired two-sample t-test for independent testing are listed in Table 5.3 to evaluate the statistical significance of the work described in this chapter. Morphological analysis significantly outperformed standard U-Net in all of the evaluation metrics for detecting and segmenting MSCs. It was seen that U-Net with post-processing significantly performed better than standard U-Net for precision and DICE metrics. Moreover, the statistical test failed to show any difference in cell detection sensitivity of these two methods, making the decrease in sensitivity value after post-processing [Table 5.2] insignificant. The t-tests also failed to show a statistical difference between DICE scores of the developed algorithm and morphological image analysis. Though morphological image analysis performed significantly better than the post-processed U-Net in precision and sensitivity, p-values indicated only a small significance margin. It can be concluded from these statistical inferences that standard U-Net architecture with further modification, optimization, and post-processing has valuable potential to advance MSC segmentation and make it more robust than the existing techniques.

Table 5.3: P-values from two-sample t-tests for statistical comparison between morphological, standard U-Net, and presented U-Net algorithms.

Comparison of detection and segmentation performance	Sensitivity per image	Precision per image	DICE Score per image
Morphological vs. standard U-Net	2.5E-02	5.3E-10	8.0E-03
Morphological vs. presented U-Net	1.4E-02	3.1E-02	7.5E-02
Presented U-Net vs. standard U-Net	1.6E-01	6.3E-08	1.0E-07

Morphometric and textural features were extracted for the TP objects from both morphological and presented U-Net analysis outputs. AUC values were measured for all features to analyze how well the segmentation results can help predict MSC culture potency. Three features for both approaches that best-classified MSCs as efficacious or non-efficacious during 10-fold cross-validation are listed in Table 5.4 for low and moderate cell densities separately. The AUCs for the top three features were very close to each other for analyzing culture images acquired at low cell density. However, all the three features computed from the post-processed U-Net for medium cell density images appeared to be better descriptors of MSC viability than features from the morphological analysis. This study indicated that as confluency rises, U-Net would segment cells with higher robustness and make culture viability estimation more effective.

Table 5.4: 10-fold cross-validation AUC of top-performing features extracted from cells correctly segmented using morphological image analysis and presented U-Net for predicting phenotype of MSCs as viable or non-viable.

Culture density	Morphological analysis		Presented U-Net analysis	
	Feature name	AUC $\pm$ s.e.	Feature name	AUC $\pm$ s.e.
Low	Minor axis	0.773 ± 0.042	Ellipticity	0.762 ± 0.029
	Standard deviation	0.743 ± 0.040	Minor axis	0.754 ± 0.024
	Entropy	0.704 ± 0.029	Dissimilarity	0.698 ± 0.035
Medium	Inertia	0.728 ± 0.029	Minor axis	0.752 ± 0.029
	Standard deviation	0.705 ± 0.023	Standard deviation	0.735 ± 0.045
	Cluster prominence	0.697 ± 0.032	Inertia	0.713 ± 0.029

5.4 Discussion

The potential of a U-Net-based deep learning algorithm to be implemented for automated and non-invasive analysis of MSCs was substantiated by the independent test results. Our proposed method detected cell regions with sensitivity and precision greater than 75% and 80%, respectively. This study shows great promise for enhancing MSC segmentation by modifying the U-Net model to increase sensitivity and precision. The presented computational method would aid the rapid and effective quantification of useful MSC features. These features will be given to a machine learning model to classify the phenotype of each cell and predict its downstream characteristics. Based on the cell phenotype obtained, it is anticipated that cell biologists will be able to predict the viability and functionality of the recovered cells. Furthermore, the study of U-Net along with morphological post-processing for cell segmentation has led to the realization that integrating deep learning with standard techniques can improve performance to a great extent.

Over the last few years, a lot of variations have been developed to advance the performance of standard U-Net for medical image diagnosis, including nested U-Net, hybrid pyramid U-Net, cas-

cade U-Nets, among others [151, 152, 153]. The next step will be to adapt and transform the U-Net architecture and post-processing for significantly greater sensitivity and precision than morphological image analysis. To conclude, the U-Net-based approach does not involve invasive processing, facilitating the implementation of continuous monitoring. It has the ability to replace the tedious exercise of manual visual inspection, making the process objective, facilitating standardization, and advancing the definition of quality release criteria. Thus, it would make MSC segmentation consistent for various culture protocols to enable high-throughput stem cell analysis.

6. 2U-NET: DOUBLE U-NET SEGMENTATION TO OPTIMIZE MACHINE LEARNING-DRIVEN MORPHOLOGICAL PROFILING OF MESENCHYMAL STROMAL CELLS

6.1 Introduction

Our earlier work discussed in chapter 4 employed low-level pixel processing and mathematical modeling to construct a compound rule-based system that segments MSCs in phase-contrast micrographs. This algorithm is only suitable for images with distinct cell boundaries and bodies. As cells become confluent, segmentation becomes challenging for conventional techniques as they rely on manual data representation, parameter tuning, and feature engineering. Segmentation errors can affect the subsequent calculation of morphological attributes causing inaccurate estimations of culture efficacy. Thus, it is crucial to develop an image analysis pipeline unhindered by these limitations and capable of elucidating information not easily assessed even by visual investigation. Deep learning can meet these needs as computers directly learn features that represent cells optimally. Lately, deep learning-enabled segmentation of individual cells has been performed popularly using U-Net [3] that treats segmentation as a pixel-level classification problem, and Mask R-CNN [154] that works on object detection along with mask prediction for each bounding box. The proposed segmentation pipeline uses U-Net as it has demonstrated great success for semantic segmentation, especially where it was infeasible to obtain thousands of annotated data [149, 150].

Application of the original U-Net [3] with a weighted loss function failed to perform instance segmentation of MSCs robustly in comparison to our prior morphological analysis (Chapter 4). This is because identifying individual MSCs is more challenging as culture confluency increases. Our preliminary evaluation outlined in chapter 5 revealed that standard U-Net implementation for MSC segmentation is improved considerably through post-processing. Thus, a U-Net with further modification, optimization, and post-processing could potentially advance MSC analysis. Various adaptations of the original U-Net have furthered its performance for microscopy image

segmentation, including RIC-Unet [155], Dense-UNet [156], and Han-Net [157]. Specifically, algorithms reported by Lux et al. [158, 159] to segment densely clustered HeLa cells are relevant to our focus on delineating overlapping MSCs. These algorithms used two weighted U-Nets for detecting cellular areas and markers separately, followed by watershed transform to interpret localized areas of confluency. The presented pipeline used a similar approach where MSC regions and algorithm-defined markers were provided by two separate U-Nets. However, in contrast to strategy by Lux et al. [159], the following changes were incorporated in 2U-Net to increase accuracy of MSC segmentation: 1) cell region and marker U-Nets were trained using DICE loss rather than weighted cross-entropy or weighted mean square error, 2) H-minima transform output served as ground truth for cell marker detection U-Net instead of erosion-based annotation, 3) preprocessing and post-processing procedures were customized for MSC phase-contrast micrographs and cell region/marker predictions, respectively.

Beyond segmentation, the ultimate aim of this study was to profile MSCs based on their morphological phenotype to evaluate culture efficacy. To this end, we have developed binary machine learning models that can robustly classify segmented cells as RS or SR in a non-destructive and rapid manner. To summarize, the novelty and main contributions of research introduced in this chapter are 1) development of a dedicated deep learning-based model (2U-Net) in combination with standard image processing procedures for enhanced MSC segmentation, 2) extension of Double U-Net segmentation by integrating feature computation with machine learning to determine the functionality of every cell in the culture based on its morphology, and 3) assessment of proposed method's capability to standardize cell quality control by substituting visual inspection and rendering culture evaluation automated, reliable, and quantitative.

6.2 Methods

We have developed a 2U-Net scheme to optimize instance segmentation and phenotype prediction of MSCs from monolayer culture images. The dataset used for training and validating the proposed approach is tabulated in section 2.3.3. The flow of the presented scheme is illustrated in Fig. 6.1. The algorithm was trained and evaluated on a machine with an NVIDIA Titan V GPU.

It was written in Python 3.5.6 using Keras library for U-Nets, Scikit-learn library for machine learning, and MATLAB 9.5 (R2018b) for image processing.

6.2.1 Pre-processing, annotation, and augmentation

As the first pre-processing step for U-Net 1, input RGB phase-contrast micrographs were converted to gray-scale and resized from 1280×1024 to 640×512 pixels. Then each training image was standardized and normalized. Note that the validation and test images were scaled using the parameters computed during training. Annotations were prepared for U-Net 1 using the cell outline ground truth and all individual cells and cell clusters were annotated as cell regions. After pre-processing and annotating, training samples were augmented using geometric transformation operations such as rotation, horizontal shift, vertical shift, zoom, shear, and flip. A different strategy was followed to pre-process input images and annotate training data for U-Net 2. After gray-scale conversion and resizing, images were pre-processed using contrast adjustment in MATLAB. Standardization and normalization were not carried out for U-Net 2 as enhancing the contrast led to better segmentation of cell markers by U-Net 2. Binarization, erosion, and H-minima transform methods to generate cell marker annotations [Fig. 6.1(b)] were evaluated through cross-validation. Training U-Net 2 with H-minima annotation yielded the best cell marker detection sensitivity and precision (Appendix A.2). H-minima transform [138] consisted of detecting regional minima inside each expert outlined cell. In cases of more than one regional minima inside a cell, only the one with the maximum area was retained as a cell marker. The same augmentation strategy as U-Net 1 was followed for training U-Net 2.

6.2.2 Training procedure

U-Net 1 and U-Net 2 models developed for MSC segmentation followed the same convolutional neural network architecture as the original U-Net [3]. U-Net 1 used 16 feature channels in the first layer while U-Net 2 used 32 feature channels. The number of feature channels was doubled for U-Net 2 because segmenting cell markers was more complicated than segmenting cell regions. Dropout layers were added after each max pooling operation to avoid the overfitting problem. Also,

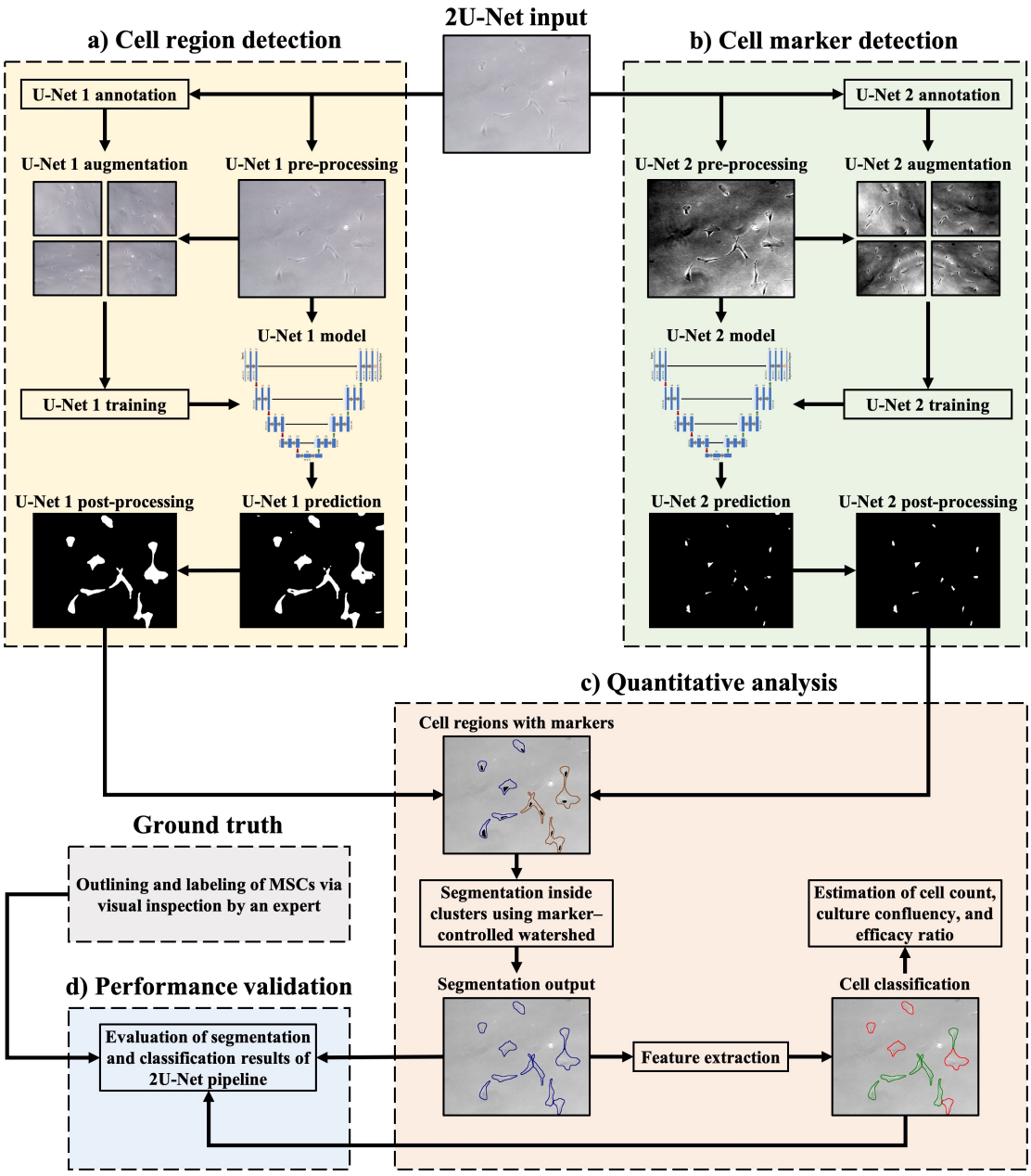

Figure 6.1: Flowchart illustrating the presented 2U-Net algorithm for image-based analysis of MSC cultures. a) U-Net 1 was trained with pre-processed, annotated, and augmented phase-contrast micrographs of MSCs. It was then used to predict cell regions that were further post-processed. Similarly, b) U-Net 2 was modeled to locate algorithm-defined cell markers in the images. These predictions were followed by c) identification of individual cells (blue) and cell clusters (brown); clusters were subsequently segmented using a marker-controlled watershed. After segmentation, object features were extracted and utilized by a machine learning classifier to categorize cells as either RS (green) or SR (red) phenotype. Quantitative metrics viz. culture confluency, cell count, and efficacy ratio were obtained using segmentation and classification results. d) Lastly, the algorithm's potential to estimate culture quality was verified using standard visual inspection as ground truth.

U-Net 1 and U-Net 2 included batch normalization after convolutions to make the networks stable and learn faster. Both U-Net models were cross-validated using DICE, weighted DICE, binary cross-entropy, and weighted binary cross-entropy loss functions [160, 3, 161]. DICE loss yielded the best cross-validation performance (Appendix A.1) and was used as the loss function for U-Net 1 and U-Net 2. Adam optimization method [162] with a learning rate of 1×10^{-4} was used by both networks to find weights. Both the networks were trained from scratch with initial weights obtained from a Gaussian distribution. During five-fold cross-validation, the networks were set to run for 100 epochs along with a callback function. The callback function reduced the learning rate when validation loss did not decrease for five consecutive epochs and ultimately stopped training when there was no improvement in validation loss for ten consecutive epochs. After setting all the hyperparameters as well as the loss function through cross-validation, the models were again trained from scratch using the pre-processed and augmented images from Culture 1 and 2 without setting aside any data for validation. The number of epochs in the final U-Net 1 and U-Net 2 training was set as the average of epochs that the models ran during cross-validation. These trained networks were saved as the final models and used by the algorithm for predicting cell regions and markers.

6.2.3 Post-processing and cell segmentation

U-Net 1 included post-processing steps of area thresholding, morphological dilation, flood-fill operation, and border clearing. Area thresholding removed artifacts too small to be a cell region. Dilation was performed to refine the shape and boundaries of the candidate cell regions, followed by a flood-fill operation to remove any holes inside the detected cell regions. Finally, the image borders were cleared of any objects via area thresholding, because truncated cells on the border were not considered by the algorithm for MSC culture analysis. The output of U-Net 2 was masked with the post-processed result from U-Net 1 so that any markers outside candidate cell regions were not considered. The markers were then processed using area thresholding, flood-fill, and morphological opening and closing. Area thresholding removed background noise, and flood filling completed the marker objects. The morphological opening was applied to detach markers with weak connec-

tions to ensure there was no clustering, and the closing operation properly connected the marker edges.

Post-processed images from U-Net 1 and U-Net 2 were combined, and the number of markers inside each region was counted. Every detected cell region was distinguished using its marker count as either a cell (1 marker) or a cluster ($>$1 marker). A region with zero markers was defined as an individual cell depending on its ellipticity, circularity, and area. The cell clusters were delineated into individual cells via marker-controlled watershed technique [140]. Performing watershed along with cell markers inhibited under or over-segmentation and separated the cluster into exactly as many cells as the markers. Cell regions labeled as individual cells together with watershed output provided final segmentation prediction for MSC culture images.

6.2.4 Feature extraction and phenotype classification

Human-engineered feature computation and machine-learning-based estimation of MSC efficacy were executed similar to section 4.2.6. After segmentation, morphometric (size and shape) and textural features (first and second-order) were extracted to predict the phenotype of cells as RS (efficacious cells) or SR (non-efficacious). The classifiers were only trained using features from segmented objects in the training dataset that were defined as cells by the ground truth. FP objects were omitted to prevent the classifiers from learning incorrect features. They were also removed from the test set to evaluate the performance of classifiers independent of segmentation results. The details of the number of cells used for feature extraction as well as classifier training and testing are listed in Table 6.1. Note that features from all segmented objects (TP and FP) were included during the assessment of the entire image analysis pipeline.

Table 6.1: Dataset used for feature extraction and classification of MSC morphological phenotype.

Dataset	Cell density	No. of cells	No. of RS cells	No. of SR cells
Training (Culture 1 + Culture 2)	Low	167	118	49
	Medium	279	176	103
Independent testing (Culture 3)	Low	62	44	18
	Medium	110	70	40

Features were sorted in descending order of their AUC [143] value for the task of distinguishing between RS and SR cells. Next, only features with a correlation less than 0.2 were selected to minimize redundancy. Each selected feature was transformed using centering and scaling by calculating the mean and variance. These statistics were stored during training and used for test data transformation later. Data augmentation was conducted for training folds during cross-validation using the synthetic minority oversampling technique [144] to deal with the lower prevalence of SR cells. Linear kernel support vector machine (LSVM), radial-basis kernel support vector machine (RSVM), linear discriminant analysis (LDA), K-nearest neighbor (KNN), and logistic regression (LR) were modeled to predict MSC phenotype. These classifiers were trained independently for features from low cell density, medium cell density, as well as low and medium combined. Features were selected for each of them separately, and these models were tuned and compared using their AUC from five-fold cross-validation. The best models were selected for the algorithm pipeline and saved for final evaluation using the independent testing data.

6.2.5 Performance evaluation

The performance of the image analysis methodology was validated for each of the three stages of cell detection, segmentation, and phenotype classification. Sensitivity (S) and precision (P) values measured using equations 3.3 and 3.4 respectively, indicated 2U-Net's ability to detect individual and clustered MSCs. Furthermore, DICE [133] was used to gauge overlap between truth and

segmentation prediction. Cell detection and segmentation metrics of the proposed technique were compared with morphological, standard U-Net, and post-processed standard U-Net techniques discussed in previous chapters 4 and 5. Statistical significance of the algorithm for independent test results was analyzed through paired two-sample t-tests [163]. P-values from t-tests were adjusted for multiple comparisons using the Bonferroni–Holm correction [164].

Receiver operating characteristic curves, AUC values, and their statistical analysis were obtained using pROC [165], an open-source package for the assessment of classification models. All linear and non-linear models were compared during cross-validation and the models with the best average AUC for distinguishing between RS and SR cells were chosen. The 95% confidence interval of the AUC metric on the test set was estimated to check if the algorithm performed significantly better than random chance. Inferences were also drawn from DeLong's test [166] to statistically verify if classifiers designed with features from 2U-Net segmentation predicted MSC phenotype more effectively than models trained with features from the morphological analysis. Sensitivity and specificity of the classifiers were found for test data by selecting a threshold that yielded a minimum for (1 – sensitivity)^2 + (1 – specificity)^2 [147]. Additionally, the manual labeling done by 20 individuals with MSC culturing experience was used to investigate the correlation between algorithm and human interpretation in evaluating culture efficacy.

Finally, the entire algorithm was applied to calculate culture attributes such as cell count, efficacy ratio, and pixel-wise confluency for each culture [Fig. 6.1(c)]. These parameters are crucial for monitoring cell health, proliferation rate, as well as culture viability. Segmented objects from all culture images were totaled to get cell count. Efficacy ratio was defined as the fraction of the number of cells predicted as RS to the total cell count (RS + SR). Pixel-wise confluency (PC) was calculated as

$$PC = \frac{\sum_{x=1}^{N} \sum_{i=1}^{P} \sum_{j=1}^{Q} I(i,j)}{N \times P \times Q} \tag{6.1}$$

wherein $I(i,j)$ represented the binary segmentation output of size $P \times Q$ having a value of one for cell pixels and zero for background pixels, and N was the number of culture images. Root

mean square errors were computed for these culture properties by means of algorithm responses vs. ground truth. As noted previously, this evaluation included FP and FN detections to account for the effect of segmentation on the algorithm's potential to determine culture quality, enabling the overall verification of the developed pipeline.

6.3 Results

6.3.1 Detection and segmentation of MSCs using 2U-Net

Results for the cell detection stage were obtained for four techniques namely morphological, standard U-Net, post-processed U-Net, and the presented Double U-Net, selected for comparison during training. The object-based sensitivity measures listed in Table 6.2 for training and independent testing provide an estimate of how well each algorithm could correctly identify objects defined as cells in the ground truth. The morphological method had high training sensitivity compared to others; however, it overfitted the training data and did not generalize as expected during testing. No U-Net-based approach suffered from this issue due to augmentation as well as U-Net's ability to learn well from limited data. Furthermore, it was verified through independent testing that the proposed Double U-Net achieved a sensitivity greater than 0.9, outperforming all other U-Net approaches. 2U-Net's lower count of FN2 could be attributed to having a separate convolutional network to localize cell markers that enhanced the segmentation of individual cells inside clusters. Additionally, applying DICE loss to identify MSC regions resulted in a reduction of FN1, thus improving the overall sensitivity of the Double U-Net algorithm.

In addition to sensitivity, precision was measured for the aforementioned techniques to assess their ability to avoid FP and detect objects that were actually cells. The object-based precision metric for training and independent testing is shown in Table 6.2. Similar to sensitivity, morphological analysis suffered from overfitting while U-Nets learned well from training data. The 2U-Net methodology had the best precision of more than 0.9, mainly due to the post-processing of outputs from U-Net 1 and U-Net 2. Post-processing removed truncated border cells, phase artifacts, and background noise that were detected as cells by standard U-Net. Post-processed U-

Net only processed standard U-Net output and performed poorly because it relied on a weighted loss function that caused over-detection of cells inside clusters (FP2). The method reported in this chapter avoided this problem as it used DICE loss with two U-nets, each separately post-processed, to identify regions and markers.

Table 6.2: Object-based cell detection sensitivity and precision of morphological, standard U-Net, post-processed U-Net, and 2U-Net algorithms for training and independent testing.

	Training (Culture 1 + Culture 2)				Independent testing (Culture 3)			
	Total no. of cells = 472				Total no. of cells = 186			
	Prior morphological analysis	Standard U-Net	Post-processed U-Net	2U-Net	Prior morphological analysis	Standard U-Net	Post-processed U-Net	2U-Net
TP	466	361	360	451	157	139	138	172
FN1	3	13	14	3	13	10	11	5
FN2	3	98	98	18	16	37	37	9
FP1	5	164	69	5	13	99	34	3
FP2	8	28	11	18	13	12	3	9
S	0.987	0.765	0.763	0.956	0.844	0.747	0.742	0.925
P	0.973	0.653	0.818	0.951	0.858	0.556	0.789	0.935

Apart from object-based analysis, segmentation results were also estimated per image using sensitivity, precision, and DICE for the testing dataset. The image-based metrics obtained for each method, shown in Table 6.3, demonstrated the Double U-Net's superior segmentation of MSCs from phase micrographs. A comparison of segmentation results with ground truth for four representative images is shown in Figure 6.2. 2U-Net had the least number of FP and FN followed

by the morphological method. Standard U-Net had the maximum number of false detections that were reduced by further processing as seen in segmentation images of post-processed U-Net. Additionally, p-values from one-tailed t-tests are outlined in Table 6.3 for independent testing data. The 2U-Net approach was proven to be significantly superior to all other techniques at p-value $<$ 0.01 across the sensitivity, precision, and DICE metrics for MSC detection and segmentation. This inference was drawn using p-values adjusted for multiple comparisons.

Table 6.3: Image-based cell detection and segmentation performance of morphological, standard U-Net, post-processed U-Net, and 2U-Net algorithms for training and independent testing. Mean and standard deviations (s.d.) were calculated across all images. Statistical comparison of 2U-Net with the other three methods illustrates its significant improvement across all metrics. P-values were adjusted for multiple comparisons using Bonferroni–Holm correction.

	Independent testing (Culture 3) with total no. of images = 36						
	Prior morphological analysis		Standard U-Net		Post-processed U-Net		2U-Net
	mean $\pm$ s.d.	p-value	mean $\pm$ s.d.	p-value	mean $\pm$ s.d.	p-value	mean $\pm$ s.d.
S	0.84 ± 0.21	5E-03	0.77 ± 0.22	5E-05	0.76 ± 0.22	3E-05	0.94 ± 0.12
P	0.86 ± 0.22	8E-03	0.54 ± 0.21	8E-12	0.81 ± 0.23	2E-03	0.95 ± 0.09
DICE	0.85 ± 0.11	2E-04	0.80 ± 0.05	6E-14	0.82 ± 0.05	1E-11	0.92 ± 0.05

6.3.2 Classification of MSC phenotype using machine learning

All the linear and nonlinear machine learning models were compared using the cross-validation AUC metric tabulated in Table 6.4. MSCs were classified more effectively when separate models were created for each level of culture density rather than having a common machine learning model. Logistic regression (LR) and linear discriminant analysis (LDA) classifiers had the highest AUCs for low and medium density, respectively, and were selected as the final models for the image

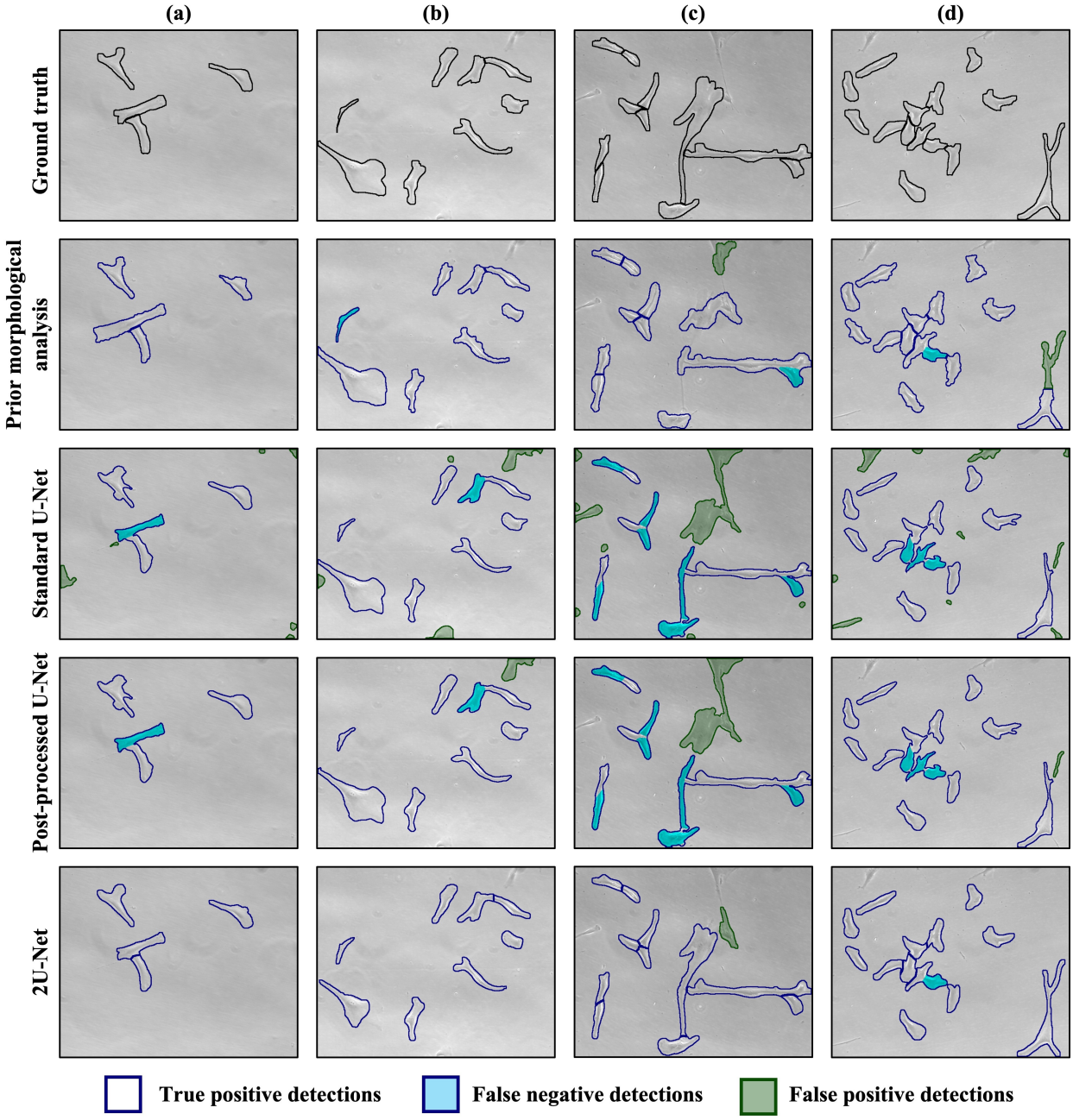

Figure 6.2: Example images (a – d) of MSCs from independent testing data comparing the ground truth cell outlines (first row) with segmentation results of previous morphological (second row), standard U-Net (third row), post-processed U-Net (fourth row), and reported 2U-Net (fifth row) methods. Cells in ground truth images are outlined in black and they also show that the truncated cells on borders are omitted for analysis. For all the algorithm outputs, true positives (correctly detected cells) are outlined in blue, false negatives (cells that were not detected) are outlined and shaded in blue, while false positives (objects incorrectly detected as cells) are outlined and shaded in green.

analysis algorithm described in this article. Independent testing was done on these classification models used in the Double U-Net pipeline.

Table 6.4: Comparison of machine learning models to distinguish between RS and SR phenotype of MSCs. AUC values of low and medium cell density classifiers that performed best during five-fold cross-validation are presented in bold.

Cell density		AUC ± S.E. from five-fold cross-validation of training data (Culture 1 + Culture 2)				
Training fold	Validation fold	LSVM	RSVM	LDA	KNN	LR
Low + Medium	Low + Medium	0.84 ± 0.01	0.85 ± 0.01	0.84 ± 0.01	0.77 ± 0.03	0.84 ± 0.01
	Low	0.79 ± 0.06	0.81 ± 0.05	0.78 ± 0.04	0.68 ± 0.04	0.77 ± 0.05
	Medium	0.85 ± 0.01	0.87 ± 0.01	0.85 ± 0.01	0.79 ± 0.03	0.85 ± 0.01
Low	Low	0.81 ± 0.03	0.80 ± 0.03	0.84 ± 0.02	0.73 ± 0.05	**0.86 ± 0.02**
Medium	Medium	0.87 ± 0.02	0.83 ± 0.01	**0.92 ± 0.01**	0.82 ± 0.02	0.87 ± 0.01

Classification performance metrics such as AUC, sensitivity, and specificity are detailed in Table 6.5 along with results from morphological image analysis previously reported in chapter 4. Classifiers were not modeled using features from standard or post-processed U-Net as their segmentation was not as robust. DeLong's test showed that 2U-Net classifiers predicted MSC phenotype significantly better than prior morphological analysis for low as well as medium density cultures. P-values were not adjusted because only a single statistical test was performed for each of the two cell density levels. Moreover, the 95% confidence intervals of these classifiers signified that they predicted cell phenotype significantly better than random chance. The fitted ROC curves for low and medium density are shown in Fig. 6.3. Lastly, Double U-Net achieved high sensitivity

and specificity over 0.80 for both density levels, exhibiting its ability to estimate culture efficacy based on MSC phenotype.

Table 6.5: Classification performance of morphological and 2U-Net algorithms for independent testing.

Metrics	Culture 3 - Low cell density		Culture 3 - Medium cell density	
	Prior morphological analysis	2U-Net	Prior morphological analysis	2U-Net
AUC $\pm$ s.e. (95% CI)	0.816 $\pm$ 0.060 (0.769, 0.886)	0.886 $\pm$ 0.042 (0.783, 0.948)	0.787 $\pm$ 0.047 (0.716, 0.851)	0.835 $\pm$ 0.037 (0.752, 0.898)
Sensitivity	0.789	0.841	0.796	0.843
Specificity	0.887	0.944	0.757	0.825
p-value	0.029		0.006	

6.3.3 Generalization of MSC phenotype prediction and assessment of culture attributes

Phenotype classification conducted by a group of 20 people (readers) with varying levels of expertise in culturing MSCs was used to understand the generalizability of the 2U-Net approach and compare it with prior morphological analysis. APS were used to distribute cells into five phenotype categories: Highly RS, Moderately RS, Uncertain, Moderately SR, and Highly SR, as defined in Table 6.6. These five categories were representative of the group's confidence in identifying a cell as either RS = 0 or SR = 1 with "Highly RS/SR" standing for very high agreement of the readers in marking a cell as RS or SR and "Uncertain" meaning that the readers had a high level of disagreement with each other (high inter-observer variability). Classifier outputs of prior morphological and 2U-Net methods were analyzed to understand the phenotype category of cells that had greater misclassification. For low and medium culture cell density, both algorithms achieved sen-

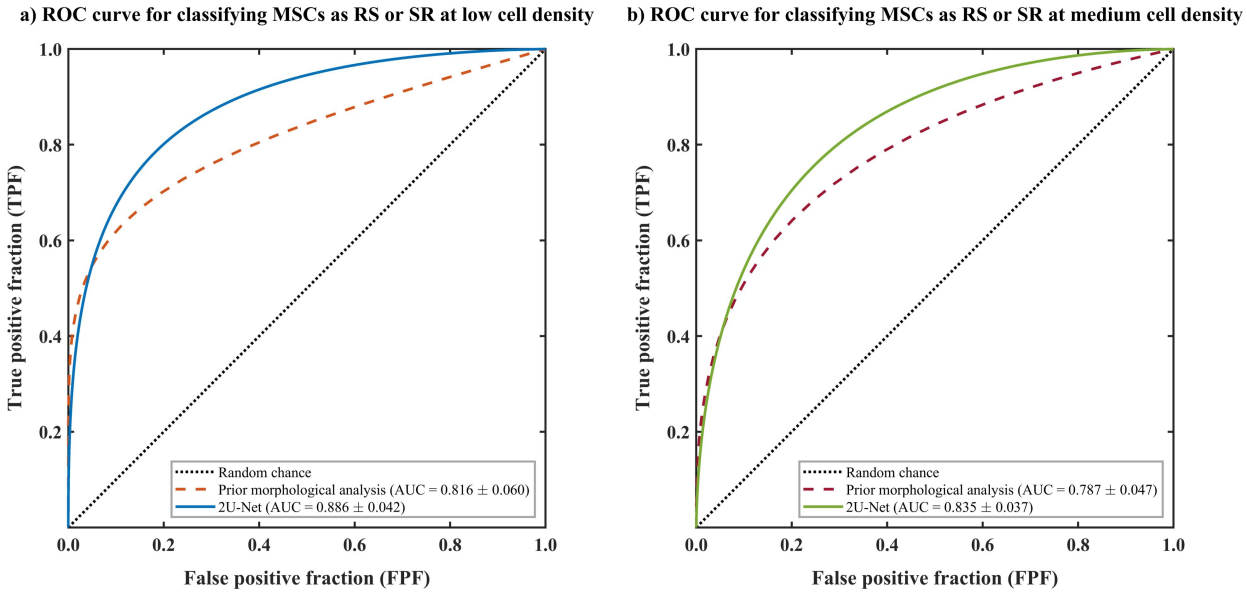

Figure 6.3: Fitted binormal ROC curves illustrating the performance of machine learning models to estimate cell phenotype as RS or SR during independent testing. a) Classifiers trained with features from MSCs segmented by morphological (dashed orange) and 2U-Net (solid blue) algorithms at low density. b) Classifiers trained with features from MSCs segmented by morphological (dashed red) and 2U-Net (solid green) algorithms at medium density.

sitivity of more than 0.9 for high confidence categories, but 2U-Net had slightly better sensitivity over the morphological method for moderate classes. Notably, Double U-Net enhanced sensitivity compared to prior morphological analysis, especially for the uncertain class. The sensitivity for this phenotype category was lower compared to the other four categories, but still improved over the morphological approach [Table 6.6]. Although 2U-Net has limitations classifying cells with marginal morphological traits, its increased sensitivity for uncertain cells shows great promise to address these constraints in future versions and estimate MSC phenotype at least as well as human inspection.

Table 6.6: Classification sensitivity of morphological and 2U-Net algorithms for five cell phenotype categories determined by average phenotype scores (APS) obtained from 20 readers for training and independent testing data. "Truth" here corresponds to labeling done by the single expert with 15+ years of experience in working with MSCs.

Phenotype class	Cell density	Sensitivity (No. of correctly classified cells / No. of truth-defined cells)			
		Training (Culture 1 + Culture 2)		Independent testing (Culture 3)	
		Prior morphological analysis	2U-Net	Prior morphological analysis	2U-Net
Highly RS	Low	1.000	1.000	0.973	1.000
(APS $\leq$ 0.2)	Medium	1.000	1.000	0.938	0.964
Moderately RS	Low	0.900	0.962	0.857	0.917
(0.2 $<$ APS $\leq$ 0.4)	Medium	0.727	0.877	0.810	0.833
Uncertain	Low	0.743	0.921	0.535	0.667
(0.4 $<$ APS $\leq$ 0.6)	Medium	0.500	0.837	0.313	0.654
Moderately SR	Low	0.886	0.939	0.882	0.929
(0.6 $<$ APS $\leq$ 0.8)	Medium	0.786	0.873	0.840	0.905
Highly SR	Low	1.000	1.000	0.912	1.000
(APS $\geq$ 1.0)	Medium	1.000	1.000	1.000	1.000

Overall testing of the image analysis pipeline was done by comparing culture attributes computed using 2U-Net's segmentation and classification results to the ground truth. Examination of the entire pipeline was carried out only for Double U-Net because it generalized MSC classification better than morphological technique. As the algorithm's estimates were very close to the truth, root mean square error (RMSE) was employed to statistically deduce how much it deviated from

the truth. The parameters were found for all three cultures in the dataset, and error values were very low for training and independent testing, indicating a good prediction accuracy and fit of the proposed methodology [Table 6.7].

Table 6.7: Comparison of culture attributes computed using 2U-Net algorithm and truth for training and independent testing.

Culture parameters	Cell density	Training					Independent testing		
		Culture 1		Culture 2		RMSE	Culture 3		RMSE
		Truth	2U-Net	Truth	2U-Net		Truth	2U-Net	
Cell count	Low	76	73	96	93	3.84	64	63	0.71
	Medium	146	141	154	158		122	122	
Pixel-wise confluency	Low	5.32	5.37	5.34	5.36	0.17	4.66	4.35	0.25
	Medium	10.64	10.90	7.60	7.38		6.76	6.58	
Efficacy ratio	Low	64.47	61.64	76.04	79.57	2.38	71.88	71.43	2.34
	Medium	54.80	54.61	68.18	69.62		61.48	64.75	

6.4 Discussion

In this chapter, we introduced customized 2U-Net architecture for deep learning-based segmentation of MSCs grown in monolayer. This approach exploited the power of deep learning with basic image processing techniques to enhance cell identification within clusters, facilitating more accurate segmentation of complex images. Efficient segmentation led to reliable and robust computation of downstream culture characteristics. The proposed analysis employed segmentation results to compute features for a machine learning model that predicts culture efficacy based on the morphological phenotype of cells. Optimized segmentation and classification could aid the robust computation of culture attributes such as cell count, pixel-wise confluency, and efficacy ratio. Validation of the algorithm at different stages proved that it outperformed state-of-the-art methods

for cell detection, segmentation, and classification. Thus, our methodology focused on not just improving segmentation but also providing a comprehensive analysis pipeline for rapid, precise, and objective culture examination at the single-cell level.

Despite very encouraging results, there are a few aspects of our algorithm that could be improved. As the current training was done with limited data, there is a need to carry out transfer learning using new MSC culture images that encompass variability due to different imaging systems and culturing conditions to help the algorithm generalize more efficiently. The efficacy ratio, a vital indicator for culture quality, is highly dependent on the correct identification of cells. The majority of cell detection errors arose due to images with poor contrast or blur. Although image contrast can be low in phase-contrast microscopy, it remains the standard technique for non-invasive and real-time monitoring of monolayer culture. Commercially-available sophisticated phase-contrast microscopes can automate imaging of live cells without manual adjustment of culture plates or wells. Moreover, false detections would not have a considerable impact on efficacy valuation if an ample number of images are analyzed to capture relevant cell populations. Similar to segmentation, classification procedures would also benefit from new and varied MSC images to gather more data and robustly learn features pertinent to cell phenotypes. Additionally, ground truth could be obtained from more experts for extending the existing classification to yield a probability that corresponds with the confidence of an expert in a cell being RS or SR. Generating a probability estimate for phenotype labels would address the concern of subjectivity in ground truth due to manual interpretation. Lastly, it was not feasible to use viability assays as ground truth for validating our research as they are population-based rather than single-cell level, costly, and labor-intensive. It would be exciting to compare the algorithm's capability for characterizing MSCs consistently to some of the popular culture assays.

In addition to U-Nets and other deep learning models, there are several open-source resources (CellProfiler [167], DeepImageJ [168]) that have demonstrated remarkable innovations in quantitative single-cell biology. They are user-friendly and have a collection of image analysis workflows specially designed for researchers/biologists without experience in computer vision or deep learn-

ing. Although they provide a generic platform trained using large datasets comprised of many types of cells, they still have some limitations that impede their usage for morphology-based examination of MSCs or other specific cell types. These solutions either require parameter adjustments, cannot be retrained with new data, or are biased towards fluorescently labeled microscopy data, making them function poorly on unstained and live-cell data acquired through brightfield imaging. 2U-Net devised here has demonstrated the ability to overcome these issues, enabling improved MSC segmentation in clusters and difficult images.

Our long-term goal to constantly enhance the algorithm by incorporating new data will aid in testing and updating the analysis pipeline to make it generalized and applicable at a larger scale. To this end, we have designed a Jupyter notebook interface for culture scientists without any prior knowledge in coding or image analysis. This interface would help biologists quickly and quantitatively evaluate MSC cultures and provide us their valuable insights on the practical usage of our method. Their feedback will ensure that our models learn continually to systematize the characterization of MSCs irrespective of their heterogeneity. Besides these advancements, the presented approach would have valuable potential to 1) design standard operating procedures for high-volume MSC manufacturing and efficacious clinical translation in treating various chronic conditions, 2) study the effect of changes in culture protocols on cell populations in a standardized manner, and 3) evaluate the viability of other stem cell lines for clinically important cell-based therapies.

In conclusion, our 2U-Net segmentation significantly optimized the segmentation of individual and clustered MSCs, enabling effective and reproducible estimation of primary efficacy endpoints. Our research shows great promise to bridge the gap between existing techniques and the need for an automated, rapid, easy, and streamlined process for monolayer culture monitoring. Double U-Net can be easily implemented in the present-day workflows for high-throughput morphological profiling of MSCs, paving the way for successful MSC-based therapies through quantifiable therapeutic potency of cells.

7. QUANTITATIVE ASSESSMENT OF THREE-DIMENSIONAL MESENCHYMAL STROMAL CELL CULTURE ATTRIBUTES USING A COMBINATION OF DEEP LEARNING AND CONVENTIONAL IMAGE ANALYSIS

7.1 Introduction

The application of MSCs for cytotherapeutic treatments in translational medicine is highly dependent on their high volume manufacture [169]. Currently, monolayer culturing techniques are used commonly in cell-based experiments mainly because they are well established with a lot of comparative literature, less expensive, and easier to analyze. Unfortunately, traditional 2D cultures provide limited surface area for cell adherence and fail to provide a physiologically relevant environment for their growth [170]. These problems cause an increase in the cost and failure rate of discoveries and clinical trials. Three-dimensional cell culture has valuable potential to overcome these limitations. It is an artificially created in vitro environment wherein cells grow and interact with their surroundings in all three dimensions. Thus, 3D culturing enhances the representation of human cells and tissues outside the body and offers a realistic way to translate study findings for in vivo applications [171]. Importantly, research has shown that stem cells cultured in 3D have greater viability, leading to the rapid scalability of functional cells [172]. Due to these benefits, the 3D cell culture industry is advancing and gaining prominence in stem cell research and cancer treatment [173, 174]. Other areas where 3D modeling has been highly advantageous include drug discovery, cytotoxicity, apoptosis, and survival [63]. Specific to MSCs, previous 3D culture experiments using bioreactors have produced a sizeable quantity of cells needed for clinical use and also established improvement in their differentiation potential [175, 176].

In addition to having a good quantity of MSCs, it is crucial to assess cell quality to ensure their functionality for various therapeutic applications. The widespread application of 3D cultures to drive translational biology necessitates a systemic approach to monitor and characterize them. Visual inspection, the most common way of analyzing cell phenotypes, is not possible with conven-

tional widefield and fluorescent microscopes generally available in cell culture labs. Such imaging systems are either suited for 2D visualization or lack the resolution needed for 3D representation of cells. Even with advanced 3D imaging, the manual investigation approach is not practical for the high-throughput setting because it is challenging and significantly time-consuming to interpret cellular details in 3D. This drawback calls for an automated method to morphologically characterize and empirically estimate the outcome for MSCs cultured in the 3D environment. Computer-aided image analysis has been employed successfully for several studies, such as high-content screening of 3D cultures, monitoring cell growth, phenotype quantification, and 3D nuclei segmentation [177, 178, 179, 180]. However, the capability of image processing to robustly analyze 3D data is highly dependent on the imaging modality used to acquire its input. LSFM allows parallel acquisition of millions of pixels, fast recording, and high-quality imaging for effective image processing. Thus, it would serve as an ideal volumetric imaging approach to assist morphological profiling and quantitative assessment of 3D MSC cultures at the single-cell level [112, 109, 113].

Our goal was to acquire images volumes of MSCs on microcarriers via LSFM and use them to develop an image analysis algorithm capable of measuring critical quality attributes of 3D culture. Segmentation of objects is an initial step required to carry out any automated image-based evaluation. Otsu thresholding [181] and U-Net deep learning model [3, 126] have been utilized vastly due to their simplicity, and there is a broad repertoire of literature demonstrating their efficiency for 3D segmentation [182, 183]. Otsu thresholding is a classical approach, and its adaptations have been used for several medical image segmentation tasks [184, 185, 186]. U-Net deep learning model has gained a lot of attention for biomedical segmentation tasks [187]. Both 2D and 3D versions of the U-Net have been applied for volumetric segmentation of brain tumors, lung nodules, ocular structures, and many more [188, 189, 190, 191]. Thus, 3D Conventional, 2D U-Net, and 3D U-Net segmentation techniques were examined in this research for volumetric segmentation of MSC cell regions and nuclei. As MSCs grow in 3D, they are likely to cluster on microcarrier surfaces, so segmenting both cell regions and nuclei was essential to delineate individual MSCs. Single-cell level segmentations were further utilized to compute cytomorphological features and critical cul-

ture parameters. These measures would be indicative of the efficacy and downstream functionality of MSCs cultured in 3D [86, 192].

In summary, the novel contribution of this work is the implementation and comparison of conventional and deep learning methodologies to localize each MSC in their volumetric images precisely. Segmentation results aid the estimation of critical parameters for monitoring the viability of 3D cultures. Furthermore, robust segmentation would provide insights into morphometric features that best describe MSC phenotypes for microcarrier-based cultures. Another noteworthy advantage is the potential of the presented work to bridge the gap for automated, rapid, high-throughput evaluation of MSCs grown in a 3D environment. Lastly, image-based processing of MSCs would serve as the front end for monitoring and streamlining 3D MSC culture protocols, aiding high volume manufacturing of efficacious cells for various therapies.

7.2 Methods

7.2.1 Overview

We present an image analysis pipeline for the objective assessment of MSCs cultured on microcarriers in a 3D environment. The algorithm was developed using MATLAB 9.5 (R2018b) for conventional image processing and Python 3.5.6 for deep learning networks. Dataset used for training and validating the 3D image analysis algorithm is detailed in section 2.4.3. The flow is demonstrated in Fig. 7.1. The input comprised two fluorescently labeled channels, namely CellTracker Green (CTG) labeling cell bodies, and DRAQ5 labeling nuclei. The CTG channel volume was pre-processed to improve the intensity, followed by segmentation. The segmentation output was post-processed to obtain the candidate cell regions. The performance of cell region segmentation was validated using the ground truth prepared with the commercial image analysis software Imaris 9.6.0 (Oxford Instruments). Similar to cell region segmentation, nuclei were identified by pre-processing the DRAQ5 channel, segmentation, and then post-processing the segmentation result. The algorithm's ability to detect and segment nuclei was also evaluated using Imaris generated ground truth. The candidate cell regions and nuclei were integrated to identify cell clusters.

Clumped cells were separated using marker-controlled watershed segmentation. Next, morphometric feature extraction was carried out for individual cells and their nuclei. Lastly, cell count and culture confluency were determined per image volume to obtain quantitative culture attributes.

7.2.2 Implementation of 2D U-Net and 3D U-Net deep learning models

2D U-Net model followed the same convolutional neural network topology as the original U-Net [3]. The proposed 2D U-Net architecture consisted of 32 feature channels in the first layer, and dropout layers were added after each max pooling operation to avoid overfitting. It included batch normalization after convolutions to stabilize the network and make it learn faster. DICE loss [160] was used as the loss function, and the Adam optimization method [162] with a 1×10^{-4} learning rate was adopted to find correct weights. The network was trained from scratch with initial weights obtained from a Gaussian distribution. It was set to run for 100 epochs and a batch size of 12 with a metric of DICE score [133]. A callback function was used for network training to reduce the learning rate when validation loss did not decrease for five consecutive epochs and ultimately stop running the model when there was no improvement in validation loss for 10 epochs consecutively.

2D U-Net can be translated easily to 3D U-Net model by adding a dimension to the network for segmenting complicated and highly variable 3D structures. The topology of the 3D U-Net remains the same comprising encoder and decoder paths, and all its layers can be designed similar to the 2D network [126]. The encoder of classical 3D U-Net was modified for this study to make the network more robust. VGG16 network [193] pre-trained for the large ImageNet [194] dataset was used as the encoder to optimize and improve the architecture. In addition to any previous pre-processing, inputs to 3D U-Net were also pre-processed similarly to the original VGG16 scheme to make them well suited for the encoder path. The network was initialized with weights from pre-trained VGG16 and used the same optimization method as 2D U-Net. A sum of DICE loss and focal loss [160] was used as the loss function for network training. The network was set to run for 250 epochs with a batch size of 8, and its performance was monitored using the DICE score [195] metric. 2D U-Net callback function was also utilized for the 3D model to monitor the validation loss, adjust the learning rate, and eventually stop network training.

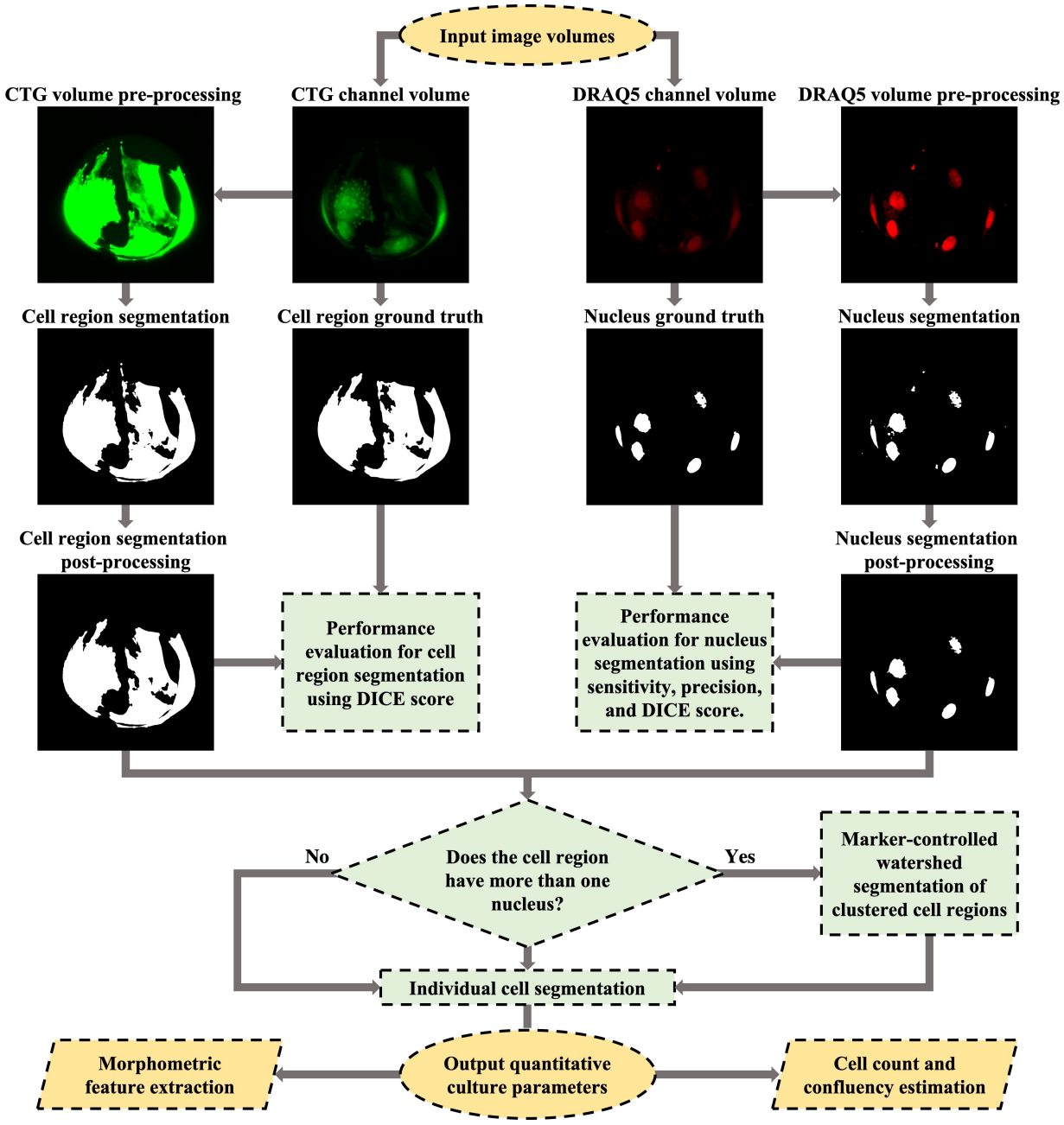

Figure 7.1: Schema describing the proposed image analysis framework.

7.2.3 Segmentation of cell regions from CellTracker Green channel image volumes

The CTG channel was used as the input and pre-processed to enhance fluorescent signal expressing cell regions. It was observed that standard contrast adjustment (saturating bottom 1% and top 1% of all voxel intensity values) was not well-suited because of inconsistent intensity levels within and between image volumes. So, limit-controlled contrast adjustment was employed to improve the signal-to-noise ratio (SNR) of image stacks. Limits were algorithm-defined, and they ensured that contrast was adjusted only for voxels whose intensity values were within those limits. These limits were very narrow and fixed such that they included only intensity values from regions of interest. Thus, cell regions were made prominent, and the background signal was diminished. Contrast limits were assigned by the algorithm depending on the mean and maximum values from the input volume's maximum intensity projection (MIP) image. These limits were set for the pre-processing algorithm during training based on what gave the best results.

For segmentation of cell regions using conventional image processing, the pre-processed volume was first binarized using Otsu's method with a globally determined intensity threshold. Binarization was followed by volume-based thresholding to remove small objects wrongly detected due to debris and noise in the image stack. Deep learning-based volumetric segmentation of cell regions was performed using 2D U-Net and 3D U-Net implementations described in section 7.2.2. Individual slices from each pre-processed CTG Z-stack and their corresponding Imaris truth were resized to 512×512 pixels and input into 2D U-Net. The 2D U-Net model processed one image slice at a time, and their 2D slice-by-slice segmentations were stacked to form a 3D volumetric segmentation. Inputs to the 3D U-Net were contrast-adjusted CTG volumes and 3D ground truth for cell regions from Imaris. Since image volumes were large, they were split into small blocks of $64 \times 64 \times 64$ voxels before being given to 3D U-Net to address computational complexity and memory need. The volumetric segmentation from 2D and 3D U-Nets were post-processed using volume-based thresholding similar to conventional image processing. Cell region segmentation was carried out using the same models and parameters irrespective of the culture day of the cells or the number of the microcarriers in the image volume. DICE score was used to evaluate how

well the cell regions were segmented relative to ground truth. Cell region segmentation results from conventional and deep learning-based methods were compared statistically using paired two-sample t-test for the validation data DICE score. The approach that performed significantly better was selected for the analysis pipeline.

7.2.4 Segmentation of cell nuclei from DRAQ5 channel image volumes

Since DRAQ5 staining of nuclei yielded non-specific binding in the rest of the cell bodies, standard contrast adjustment was inadequate. Therefore, pre-processing for nuclei segmentation was also done utilizing limit-controlled adjustment similar to CTG volume pre-processing. The pre-processed DRAQ5 volumetric image was binarized using Otsu's method for segmentation by conventional image processing. The intensity threshold for binarization was not estimated globally but defined by the algorithm depending upon the mean and maximum values found from the MIP image during pre-processing. Similar to contrast limits, these threshold values were also fixed during training. The binarized output was post-processed using volume-based thresholding and morphological operations. Volume-based thresholding facilitated the removal of small objects detected due to background noise, and morphological closing operation filled any holes in the segmented objects and refined their boundaries. Detected nuclei with large volumes were further processed using morphological erosion to avoid under-segmentation.

For nuclei segmentation with 2D U-Net, each image slice from pre-processed DRAQ5 Z-stack and their Imaris ground truth were resized to 512×512 pixels and inputted to the 2D model. 2D U-Net predictions for all image slices were stacked to create the volumetric segmentation. Pre-processed DRAQ5 volumes and their respective nuclei region ground truth from Imaris were divided into blocks of $64 \times 64 \times 64$ voxels and given as inputs to the 3D U-Net model. Note that 2D and 3D U-Nets were modeled separately for the tasks of detecting cell regions and nuclei. They were implemented as outlined in section 7.2.2. Nuclei prediction from 2D and 3D U-Nets were also post-processed through volume-based thresholding. The algorithm assigned the volume threshold for post-processing of all segmentation methods based upon the number of slices in the DRAQ5 volumetric image. This step accounted for variation in nuclei size when imaging mag-

nification and pixel array dimensions were modified to capture multiple microcarriers in a single image volume, making the Z-stack large. DICE score was used as the segmentation metric, and sensitivity and precision were used as the object detection metrics to evaluate all segmentation approaches. DICE score measured overlap between segmented results and ground truth, sensitivity metric measured how accurately all the nuclei were detected, and precision metric measured how accurately the algorithm did not detect objects that were not nuclei. Unlike cell region segmentation, object detection metrics were crucial for assessing nucleus segmentation to check if the algorithm localized nuclei without under or over-segmentation to perform single-cell analysis robustly. The image processing approach that segmented nuclei significantly better based on paired two-sample t-testing of validation data metrics was chosen for the analysis pipeline.

7.2.5 Computation of human-engineered features and culture attributes at the single-cell level

The cell region and nuclei segmentation results were combined to achieve individual cell segmentation. Detecting individual cells was required to perform any further analysis at the single-cell level. Segmented nuclei objects served as cell markers, and they were counted inside each detected cell region. A cell region with 1 cell marker was defined as an individual cell, and a region with $>$ 1 cell marker was identified as a cell cluster. If a region had no cell marker, meaning no segmented nucleus, it was removed and not considered a cell. Cells were separated inside each cluster using 3D marker-controlled watershed method [196]. Human-engineered features were computed for each 3D segmented cell and its nucleus. Volume, surface area, sphericity, elongation, ellipticity, solidity, and extent were extracted for all cells and nuclei [197, 198, 199]. Besides these features, cell surface area-volume ratio, nucleus surface area-volume ratio, nucleus-cell volume ratio, and nucleus-cell surface area ratio were also measured. These morphometric features express and quantify variations in size and shape of cells and nuclei and can act as vital diagnostic factors for cell proliferation and functional potency [200, 201]. Moreover, these features would serve as a front-end to a cell classification algorithm, making phenotype-based efficacy determination of 3D cultured cells rapid and reproducible.

Next, cell count and confluency were estimated as they can be used as quantitative parameters for non-destructive monitoring and quality control of microcarrier-based MSC cultures [202, 203, 204]. Segmented nuclei objects were totaled to get cell count per image volume. Cell count obtained using nuclei center points from Imaris was used as the reference (observed values). Deviation of the predicted cell count from the Imaris truth was estimated via RMSE. Confluence in monolayer cultures was defined as the percentage of area covered by adherent cells. Translating to 3D culture with an assumption that adherent cells on microcarriers have approximately half their cell surface area adhered to the microcarrier, confluency in 3D culture was defined as the ratio of half the surface area of cell regions to the total surface area of all microcarriers in the volume. Firstly, the radius was determined manually for each microcarrier using its line intensity profile from the MIP of the image volume. Next, it was used to get microcarrier surface area. The actual confluency value for each image volume was obtained using 1) half the surface area of all cell regions segmented by Imaris and 2) total surface area of all microcarriers in the volume. For predicted confluency values, algorithm segmented cell regions were used to get the cell surface area. Additionally, microcarrier segmentation was needed for automated calculation of their surface area to get algorithm predicted confluency. Therefore, single microcarriers were identified from the BF image volume as spherical objects. The BF volume [Fig. 7.2(a)] was processed through median filtering to reduce noise, and its intensity was enhanced with standard contrast adjustment [Fig. 7.2(b)]. This step was followed by 3D edge detection using the Sobel filter [Fig. 7.2(c)]. Background artifacts were removed using volume-based thresholding [Fig. 7.2(d)], and then detected edges were refined and connected using morphological erosion and closing [Fig. 7.2(e)]. A bounding box was created for the segmented object to get the radius and center of the microcarrier to generate the sphere [Fig. 7.2(f)] whose surface area was used in confluency measurement.

For BF image volumes with multiple microcarriers, individual microcarriers could not be detected due to clumping clumping of microcarriers and significant out of focus signal. However, microcarriers for such volumetric images were able to be identified using the DRAQ5 channel [Fig. 7.3(a)]. DRAQ5 channel was preferred over the CTG channel because it had a higher ratio

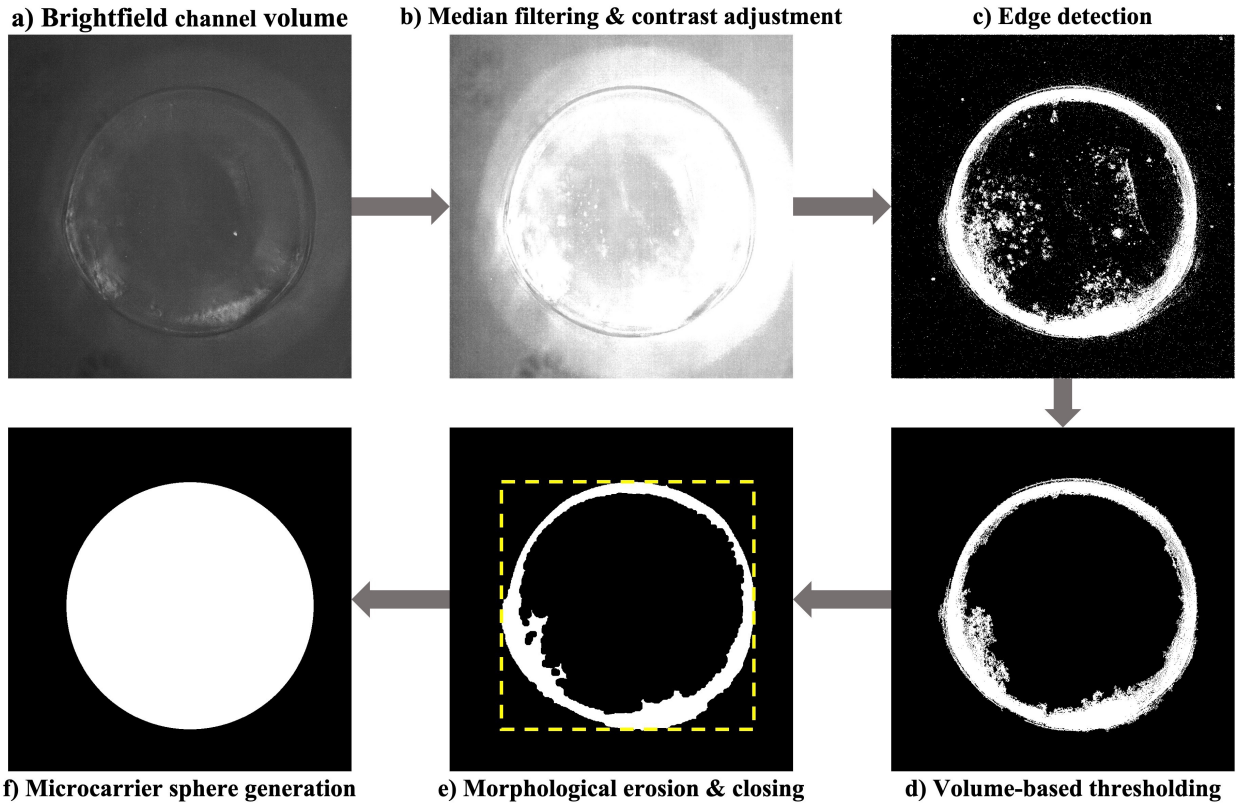

Figure 7.2: Illustration of steps involved in segmenting single microcarriers. a) Brightfield (BF) channel volume used as input. b) Reduction of noise and intensity enhancement for BF volume through median filtering and standard contrast adjustment. c) Microcarrier edge detection using Sobel filter. d) Removal of background noise and false positive objects through volume-based thresholding. e) Connection and refinement of detected microcarrier edges via morphological erosion followed by closing. f) Generation of microcarrier sphere using bounding box (shown in yellow) dimensions from the previous step.

of microcarrier area to fluorescent regions of interest than the latter. The image stack was first processed using limit-controlled contrast adjustment to increase the intensity of the backscattered signal from the microcarriers [Fig. 7.3(b)]. It was binarized using Otsu's method with an intensity threshold decided by the algorithm based on the mean value of volume's MIP [Fig. 7.3(c)]. After binarization, morphological opening was performed to fill and connect the areas of microcarriers [Fig. 7.3(d)]. Next, volume-based thresholding was carried out to eliminate falsely detected objects [Fig. 7.3(e)]. Lastly, the segmented volume was processed using morphological opening to refine boundaries [Fig. 7.3(f)]. The total surface area of this segmentation output was used to calculate

confluency for image volumes containing multiple microcarriers. In addition to RMSE, adjusted R^2 from linear regression was utilized to analyze how close the algorithm confluency predictions were to their observed values. Overall confluency of a 3D culture sample can be obtained by taking the average of confluency values from all its acquired image volumes.

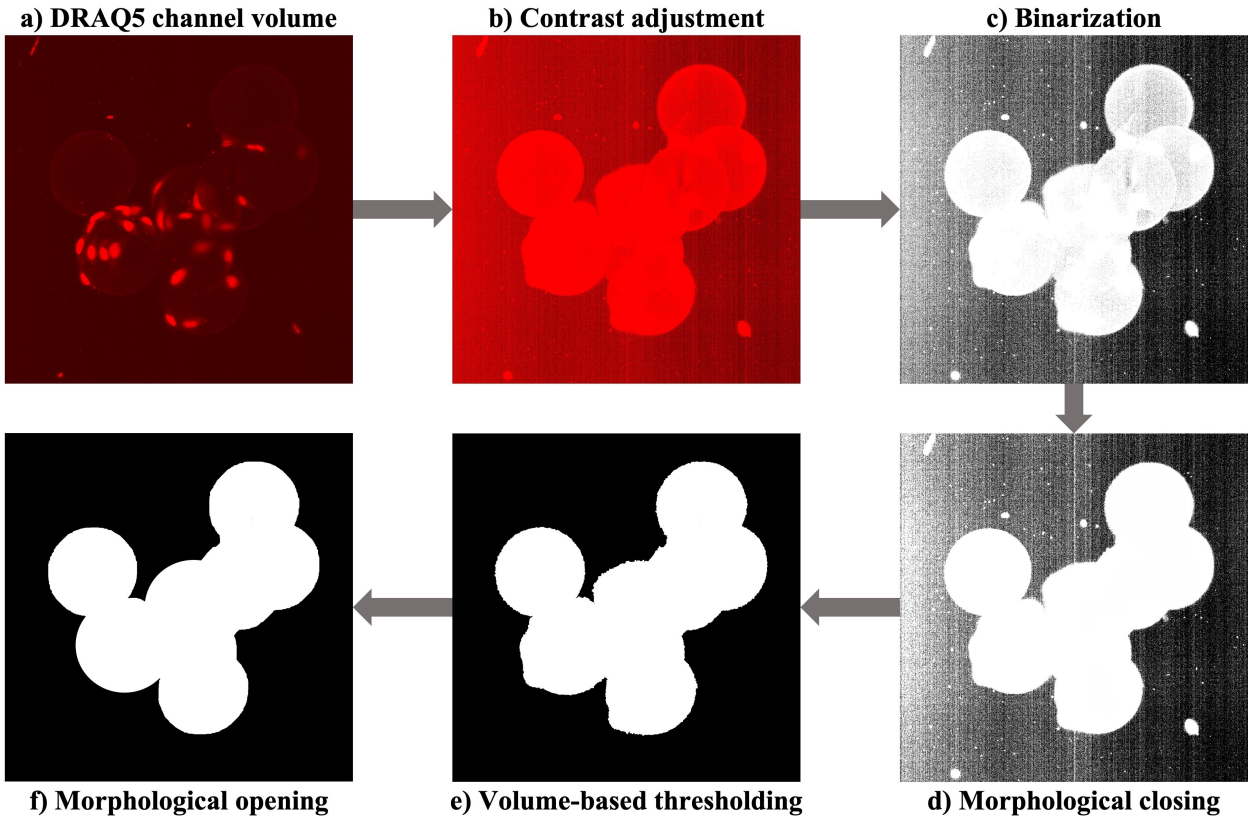

Figure 7.3: Representation of segmentation steps for multiple microcarriers. a) DRAQ5 channel volumetric image used as input. b) Improvement of intensity for microcarrier voxels applying limit-controlled contrast adjustment. c) Binarization using Otsu's method with algorithm-defined intensity threshold. d) Connection and filling of binarized voxels via morphological closing. e) Removal of background noise and false positive objects through volume-based thresholding. f) Fine-tuning segmented border of multiple microcarriers utilizing morphological opening.

7.3 Results

7.3.1 Cell region and nucleus segmentation from image volumes of microcarrier-based 3D MSC cultures

Segmentation performance for cell regions and nuclei was assessed for three image processing approaches, namely 3D conventional, 2D U-Net, and 3D U-Net. Identification of cell regions, a semantic segmentation task was evaluated using DICE score only to check the algorithm's overlap with ground truth from Imaris. Nuclei segmentation was treated as instance segmentation since every individual nucleus needed to be localized separately for single-cell level analysis. So, nuclei segmentation was validated using object detection metrics of sensitivity and precision along with the DICE score. Table 7.1 reports the above metrics for cell region and nuclei segmentation by all three methods during training and validation. It was confirmed using validation data that the 3D U-Net method segmented cell regions with the best DICE score having an overlap of more than 0.9 with the truth. 3D conventional (Otsu thresholding) and 2D U-Net (3D construction from stacking 2D predictions) approaches had comparatively low DICE scores for training and validation, showing that they were not robust enough to segment cell regions from MSC culture image volumes. Their low DICE score was most likely because they over-detected voxels around cell region boundaries. It was found that 3D conventional still performed better than the 2D U-Net approach. The significance of these inferences from comparing the three methods was statistically verified using one-tailed p-values from paired t-test reported in Table 7.2. Thus, the trained 3D U-Net was selected for the cell region segmentation part in the proposed image analysis pipeline. Figure 7.4 displays sample CTG image volumes with their ground truth and cell region segmentation by three methods compared in this study.

Table 7.1: Performance comparison of 3D conventional, 2D U-Net, and 3D U-Net methods for cell region and nucleus segmentation from image volumes of microcarrier cultured MSCs. The image analysis approach that performed best during training and validation is presented in bold for each metric.

Metrics	Segmentation method	Training (N = 23)	Validation (N = 21)
Cell region segmentation DICE score (mean ± s.d.)	3D conventional	0.789 ± 0.175	0.730 ± 0.210
	2D U-Net	0.831 ± 0.097	0.810 ± 0.079
	3D U-Net	**0.943 ± 0.023**	**0.927 ± 0.029**
Nucleus segmentation DICE score (mean ± s.d.)	**3D conventional**	**0.950 ± 0.018**	**0.933 ± 0.022**
	2D U-Net	0.608 ± 0.059	0.645 ± 0.081
	3D U-Net	0.760 ± 0.154	0.677 ± 0.184
Nucleus detection sensitivity (mean ± s.d.)	**3D conventional**	**0.975 ± 0.054**	**0.988 ± 0.030**
	2D U-Net	0.791 ± 0.219	0.830 ± 0.288
	3D U-Net	0.772 ± 0.255	0.546 ± 0.408
Nucleus detection precision (mean ± s.d.)	**3D conventional**	**0.996 ± 0.011**	**0.982 ± 0.073**
	2D U-Net	0.681 ± 0.271	0.746 ± 0.296
	3D U-Net	0.951 ± 0.122	0.976 ± 0.109

3D conventional image processing achieved the best sensitivity, precision, and DICE score in detecting and segmenting nuclei for training and validation data. It had high validation sensitivity and precision values of 0.98 and a 0.93 DICE score [Table 7.1]. Both the U-Net models (2D and 3D) had considerably lower sensitivity and DICE scores. The 2D U-Net method had lower precision too, but 3D U-Net precision was closer to 3D conventional processing. DRAQ5 channels had some background signal from cell regions around nuclei even after pre-processing. Although the intensity of this background noise was less than nuclei, it got predicted by 2D U-Net as nuclei, causing over-segmentations of voxels around nuclei. In the case of 3D U-Net, it was unable to

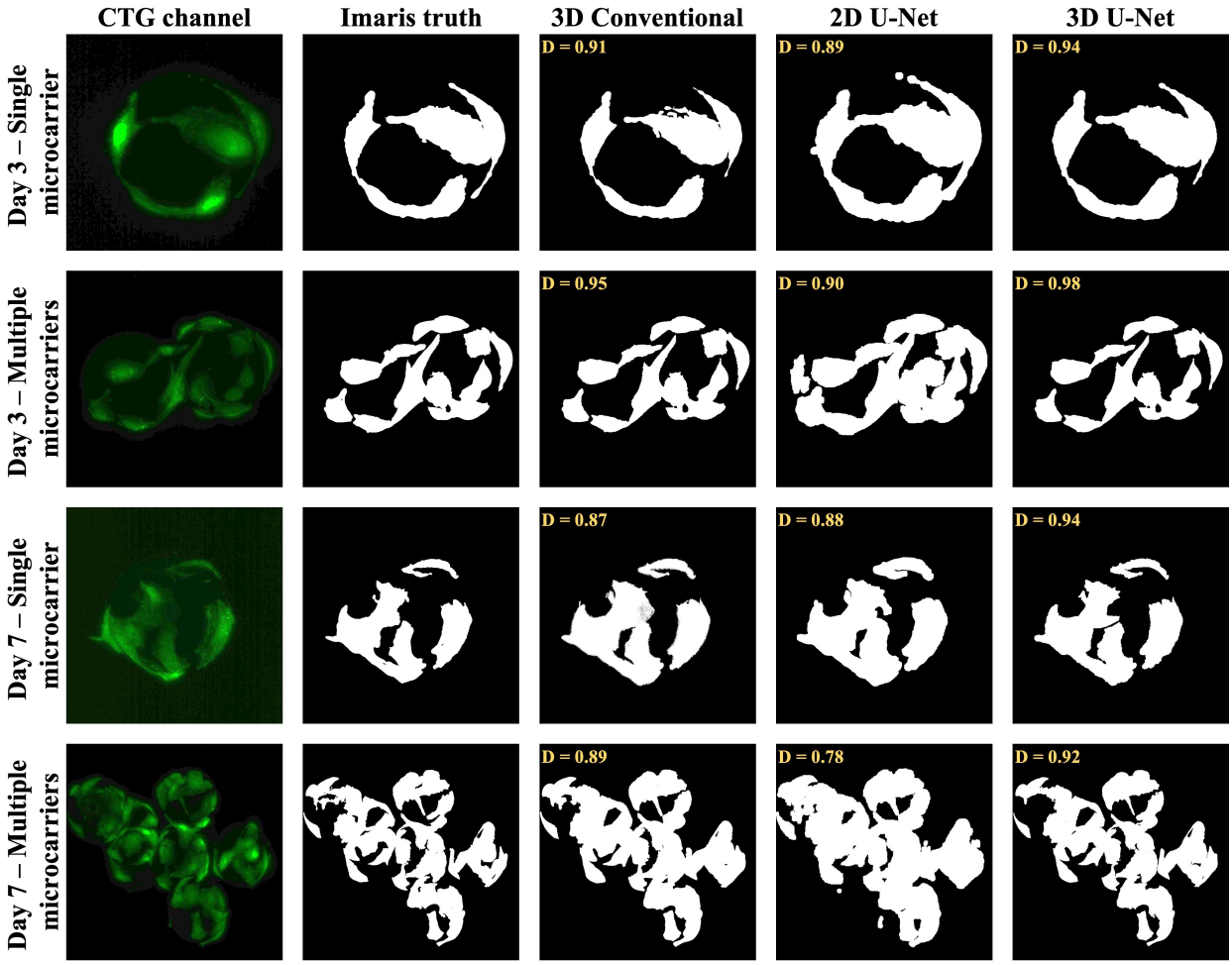

Figure 7.4: Sample image volumes of 3D cultured MSCs stained with CTG (first column) from validation data comparing the ground truth cell regions (second column) with the segmentation results of 3D conventional (third column), 2D U-Net (fourth column), and 3d U-Net (fifth column) methods. Each row is an example of a volumetric image from days 3 and 7 of culture expansion consisting of single and multiple microcarriers. DICE score (D) is shown in yellow for each segmentation result image on the top left corner.

generalize well due to highly variable intensity levels within and between DRAQ5 image volumes. This problem led to under-segmentation that contributed to poor sensitivity. 3D conventional processing did not suffer from these issues because of 1) adaptive pre-processing of DRAQ5 volume with limit-controlled contrast adjustment and 2) binarization of only nuclei voxels, ignoring background signal using algorithm-defined threshold values. Statistical significance of nuclei detection and segmentation results was also examined using paired two-sample t-tests [Table 7.2]. The 3D conventional method achieved a significantly higher DICE score and sensitivity than 2D and 3D U-Nets. The statistical test failed to show any difference between the DICE score of both U-Nets and 2D U-Net had significantly better sensitivity than 3D U-Net. For precision, 3D conventional and 3D U-Net methods outperformed 2D U-Net, and the t-test failed to show any statistical difference between the two 3D methods. Thus, 3D conventional processing could overall segment and detect nuclei most robustly and was chosen for the nuclei segmentation part of the image analysis pipeline. Note that all one-tailed p-values were adjusted for multiple comparisons using the Bonferroni-Holm technique before using them to investigate which cell region and nuclei segmentation method was significantly better than others. Example DRAQ5 image volumes with their ground truth and nuclei segmentation by three methods analyzed here are illustrated in Fig. 7.5. To summarize, 3D U-Net and 3D conventional methods demonstrated the best performance for segmenting cells regions and nuclei, respectively, enabling the identification of individual MSCs and their quantitative study at the single-cell level.

Table 7.2: Statistical comparisons between performance metrics of 3D conventional, 2D U-Net, and 3D U-Net methods for cell region and nucleus segmentation using Imaris ground truth (N = 21). P-values were adjusted for multiple comparisons using Bonferroni-Holm correction, and statistically significant results are indicated in bold. ">" in the table stands for "outperformed."

Metrics	3D conventional vs. 2D U-Net	3D conventional vs. 3D U-Net	2D U-Net vs. 3D U-Net
Cell region segmentation DICE score	**3D conventional** > 2D U-Net **p = 2.609E-02**	**3D U-Net** > 3D conventional **p = 1.833E-04**	**3D U-Net** > 2D U-Net **p = 4.781E-06**
Nucleus segmentation DICE score	**3D conventional** > 2D U-Net **p = 6.561E-12**	**3D conventional** > 3D U-Net **p = 4.506E-06**	Failed to show significance p = 2.552E-01
Nucleus detection sensitivity	**3D conventional** > 2D U-Net **p = 1.746E-02**	**3D conventional** > 3D U-Net **p = 7.246E-05**	**2D U-Net** > 3D U-Net **p = 5.623E-03**
Nucleus detection precision	**3D conventional** > 2D U-Net **p = 3.078E-03**	Failed to show significance p = 4.161E-01	**3D U-Net** > 2D U-Net **p = 1.450E-03**

7.3.2 Cell count and confluency estimation for microcarrier-based 3D MSC cultures

The ability of the presented image analysis pipeline to predict confluency and cell count parameters for 3D MSC cultures was also assessed against measurements obtained from Imaris in a semi-automated manner. Figure 7.6 exhibits scatter plots of observed confluency values from Imaris versus confluency values predicted by the algorithm. A linear regression line with its 95% confidence interval was used to visualize how close algorithm predictions were to the truth. All the points in the scatter plot for training data (Number of image volumes, N = 23) were very close

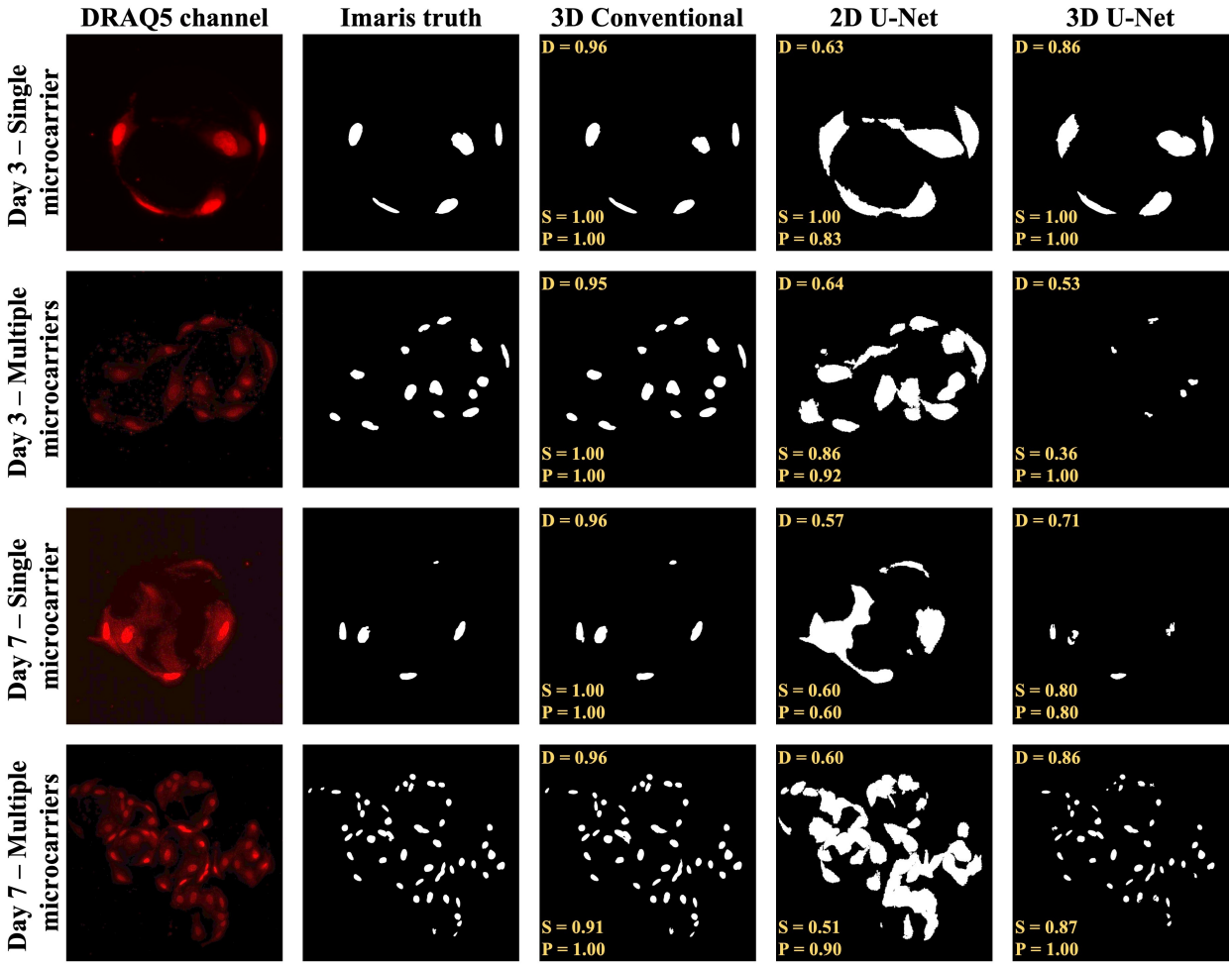

Figure 7.5: Sample image volumes of 3D cultured MSCs stained with DRAQ5 (first column) from validation data comparing the ground truth nuclei (second column) with the segmentation results of 3D conventional (third column), 2D U-Net (fourth column), and 3d U-Net (fifth column) methods. Each row is an example of a volumetric image from days 3 and 7 of culture expansion consisting of single and multiple microcarriers. DICE score (D) is shown on the top left corner, and sensitivity (S) and precision (P) are shown on the bottom left corner for each segmentation result image in yellow.

to the regression line, indicating a high R^2 value (> 0.9) and goodness of fit for the algorithm. Additionally, RMSE was less than 2.5%, exhibiting a negligible difference between truth and values predicted by the developed analysis scheme. These inferences were confirmed through the scatter plot for observed and predicted values from validation data (Number of image volumes, N = 21). The algorithm scored an R^2 value greater than 0.9 and an RMSE of 2.7% during validation, substantiating its potential to estimate culture confluency very close to the truth consistently. Similar to confluency, cell counts determined by the algorithm were compared with Imaris. The algorithm reported an RMSE of 1.504 and 4.180 for cell counts during training and validation, respectively. It was observed that image volumes with multiple clumped microcarriers having more cells (> 30) mainly contributed to these errors. Thus, cell count errors for such image volumes were inconsequential compared to the actual number of cells, verifying the algorithm's capacity to measure cell count with higher accuracy. The algorithm's effectiveness in determining culture parameters further proved that the proposed analysis pipeline could not only properly segment 3D cultured MSCs but also assess them quantitatively to characterize culture functionality.

7.4 Discussion

We developed a 3D image analysis pipeline that could localize individual MSCs with sensitivity and precision of over 0.9 via cell regions and nuclei segmentation from CTG and DRAQ5 fluorescence channels, respectively. This algorithm was established after comparing and validating the performance of three standard image processing methods including, 3D conventional, 2D U-Net, and 3D U-Net. From validation results and their statistical testing, it was inferred that a combination of conventional and deep learning methods was best suited for the proposed analysis. The presented experiments and final analysis scheme is impactful and innovative because such research has not been carried out previously for quantitative evaluation of multidimensional MSC cultures. Robust segmentation enabled the estimation of morphometric features that could serve as relevant descriptors for the cellular phenotype. In addition to feature extraction, it aided in predicting quantitative process parameters for 3D MSC culture, namely cell count and confluency. Cellular features along with these parameters would facilitate characterizing the efficacy of MSCs

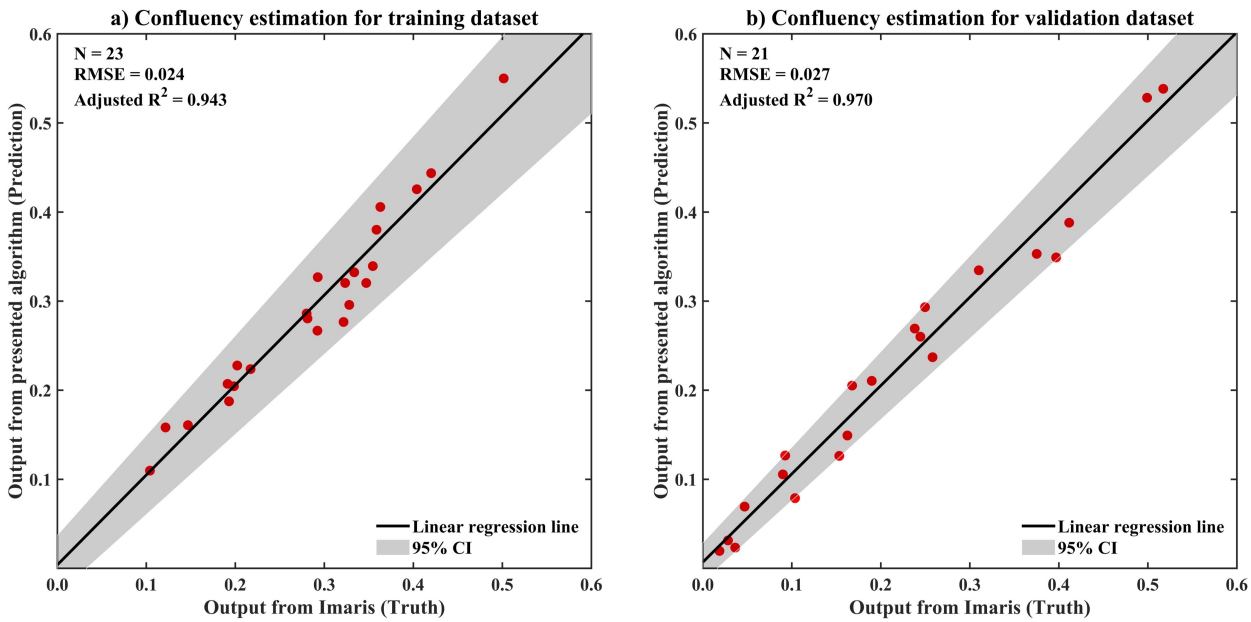

Figure 7.6: Scatter plots of observed (Imaris) versus predicted (algorithm) culture confluency values for training (N = 23) and validation (N=21) dataset. The solid black lines indicate computed linear regression lines, and the region shaded in gray color is a 95% confidence interval (CI) for the slope of the regression lines. Root mean square error (RMSE) and adjusted R^2 values are displayed on the top left corner of the plots. a) Plot showing a high linear correlation and close fit of the algorithm with the ground truth for training data. b) Validation data plot confirming high correlation and good fit of the presented algorithm in estimating confluency.

cultured in 3D on microcarriers.

Though this approach is quite promising, there are still some areas for improvement. Currently, algorithm training and validation were done using volumetric image samples from the same culture. It is essential to acquire image stacks from a different 3D culture sample to generate independent testing data for assessing the generalizability of the overall developed method. For culture parameters, confluency measurement assumed that half the surface area of a cell would be adhering to the microcarrier. In real-time experiments, this assumption might not always hold, especially when there are a lot of cells and their clusters, leading to erroneous confluency prediction. Cell attachment to multiple microcarriers was also not considered for image stacks with clumped microcarriers. These shortcomings need to be addressed in a future version of the algorithm to ensure consistent and accurate confluency estimation irrespective of the number of cells or microcarriers

in image volumes. Lastly, the algorithm relies on stained channels for segmenting cell regions and nuclei. This requirement would limit its application for real-time and non-invasive monitoring quality of 3D MSC cultures. Extending the presented technique to detect cell regions and nuclei from label-free image data would make it more impactful and suitable for adoption in current 3D culture workflows.

To summarize, we built an analysis pipeline that could segment MSCs grown on microcarriers and calculate critical 3D culture attributes with high temporal resolution. This study shows the capability to serve as a front-end for 3D MSC culture assessment, bridging the gap between existing techniques and the necessity for an objective, automated, and high-throughput approach. Progress in this research would streamline quality control processes for MSCs in a 3D environment, scaling up the manufacture of efficacious cells required to treat chronic diseases.

8. CONCLUSION AND FUTURE SCOPE

We developed a morphological analysis approach for monolayer cultures using conventional image processing techniques that segmented MSCs and classified them based on their morphological phenotype very close to human observation. It provided evidence supporting the applicability of image analysis as a tool to objectively monitor monolayer adherent MSCs. Moreover, its ability to morphologically profile MSCs would greatly aid the prediction of downstream functionality of cultures. It would replace or augment visual inspection to objectively assess MSCs in an automated and non-invasive manner. Thus, the promising performance of this approach for analyzing MSCs from their phase-contrast micrographs suggested that image-based analysis could be seamlessly integrated into the current monolayer culture workflows. The next important step for building upon this research would be fine-tuning the algorithm using new 2D MSC culture images that encompass variability due to different phase-contrast imaging systems and various culturing conditions. It would also be worthwhile to calibrate the machine learning model used in the current scheme to correlate its probability output with confidence in the phenotype of an MSC being RS or SR. Lastly, comparing the morphological image analysis with some commonly utilized culture evaluation assays would be interesting to estimate the effectiveness and generalizability of our research.

To further optimize the segmentation of MSCs, the power of deep learning was employed to replace human-engineered parameters used for the morphological analysis approach with features learned directly by the computer. Our customized deep learning pipeline significantly optimized analysis of monolayer cultures, outperforming state-of-the-art methods for image-based MSC detection and classification. Furthermore, enhanced MSC identification led to a meaningful improvement in the morphological profiling of cells. This comprehensive analysis pipeline resulted in accurate and reproducible computation of quantitative critical quality parameters and efficacy outcome for monolayer MSC cultures. Thus, the presented research on deep learning-based analysis has valuable potential to meet the need for streamlined and objective monolayer culture monitoring in real-time. For this dedicated deep learning algorithm, it is worthwhile to note there is scope

for further betterment in the morphological screening of MSCs cultured in monolayer conditions. This upgrade can be achieved by applying deep learning for also feature extraction and classification. The next step towards standardizing this image analysis solution would be to incorporate feedback from biologists to ensure the algorithm learns continually and systematizes MSC monitoring irrespective of their heterogeneity. It would also greatly benefit other cell-based therapies and treatments to expand this analysis for assessing the clinical relevance of more stem cell lines.

As the culturing paradigm for stem cell expansion is shifting from monolayer to 3D systems, we built an algorithm to extend our methodology for multidimensional analysis. It was concluded that a combination of conventional and deep learning-based methods was most suitable for segmenting MSC regions and their nuclei from volumetric images. It efficiently examined MSCs adhering to microcarriers in the 3D bioreactor environment at the single-cell level. Additionally, the 3D image analysis strategy robustly estimated empirical culture parameters for microcarrier-based MSC cultures similar to monolayer cultures. Therefore, this study would serve as a front-end algorithm for reproducible morphological phenotyping of MSC cultured in 3D. It would facilitate the manufacturing of MSCs at a large scale for clinically successful therapeutics. This 3D analysis would become more impactful by translating the current algorithm for real-time and non-invasive examination of 3D MSCs by segmenting them in label-free image volumes. Besides finding each MSC from image volumes, it is essential to characterize functionality and monitor viability as they grow on microcarriers in a 3D culture. Progress in this research would make it possible to study the effect of changes in the 3D culturing protocol on cell potency and design standard operating procedures required for reliable and high-volume manufacturing of efficacious cells.

This research on image-based analysis for evaluating 2D and 3D MSC cultures has facilitated the development of a system that can quantitatively describe cell density, distribution, and morphology. Estimation of these measures would enable the establishment of quality release criteria for cultures. Thus, further expansion and incorporation of this technology would significantly aid the manufacturing and monitoring of multidimensional cultures to ensure their successful application for cytotherapies.

REFERENCES

[1] F. Zernike, "Phase contrast, a new method for the microscopic observation of transparent objects," *Physica*, vol. 9, no. 7, pp. 686–698, 1942.

[2] O. E. Olarte, J. Andilla, E. J. Gualda, and P. Loza-Alvarez, "Light-sheet microscopy: a tutorial," *Advances in Optics and Photonics*, vol. 10, no. 1, pp. 111–179, 2018.

[3] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer Assisted Intervention (MICCAI)* (N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, eds.), vol. 9351, (Cham), pp. 234–241, Springer International Publishing, 2015.

[4] S. Bernell and S. W. Howard, "Use your words carefully: what is a chronic disease?," *Frontiers in Public Health*, vol. 4, p. 159, 2016.

[5] W. Raghupathi and V. Raghupathi, "An empirical study of chronic diseases in the United States: a visual analytics approach to public health," *International Journal of Environmental Research and Public Health*, vol. 15, no. 3, p. 431, 2018.

[6] National Center for Chronic Disease Prevention and Health Promotion (NCCDPHP), "Chronic diseases in America." `https://www.cdc.gov/chronicdisease/resources/infographic/chronic-diseases.htm`, 2019.

[7] Partnership to Fight Chronic Disease (PFCD), "The growing crisis of chronic sisease in the United States.." `https://www.fightchronicdisease.org/sites/default/files/docs/GrowingCrisisofChronicDiseaseintheUSfactsheet_81009.pdf`, 2007.

[8] A. B. Martin, M. Hartman, D. Lassman, and A. Catlin, "National health care spending in 2019: steady growth for the fourth consecutive year," *Health Affairs*, vol. 40, no. 1, pp. 14–24, 2021.

[9] R. Sanz-Ruiz, E. G. Ibañes, A. V. Arranz, M. E. F. Santos, P. L. S. Fernández, and F. Fernández-Avilés, "Phases I–III clinical trials using adult stem cells," *Stem Cells International*, vol. 2010, p. 579142, 2010.

[10] C. Agostini, "Stem cell therapy for chronic lung diseases: hope and reality," *Respiratory Medicine*, vol. 104, pp. S86–S91, 2010.

[11] T. Takami, S. Terai, and I. Sakaida, "Stem cell therapy in chronic liver disease," *Current Opinion in Gastroenterology*, vol. 28, no. 3, pp. 203–208, 2012.

[12] S. A. Fisher, C. Doree, A. Mathur, D. Taggart, and E. Martin-Rendon, "Stem cell therapy for chronic ischaemic heart disease and congestive heart failure," *Cochrane Database of Systematic Reviews*, vol. 12, p. CD007888, 2016.

[13] B. George, "Regulations and guidelines governing stem cell based products: clinical considerations," *Perspectives in Clinical Research*, vol. 2, no. 3, pp. 94–99, 2011.

[14] M. F. Pittenger, D. E. Discher, B. M. Péault, D. G. Phinney, J. M. Hare, and A. I. Caplan, "Mesenchymal stem cell perspective: cell biology to clinical progress," *npj Regenerative Medicine*, vol. 4, no. 1, p. 22, 2019.

[15] A. Caplan, "Why are MSCs therapeutic? New data: new insight," *The Journal of Pathology*, vol. 217, no. 2, pp. 318–324, 2009.

[16] D. Baksh, L. Song, and R. S. Tuan, "Adult mesenchymal stem cells: characterization, differentiation, and application in cell and gene therapy," *Journal of Cellular and Molecular Medicine*, vol. 8, no. 3, pp. 301–316, 2004.

[17] G. Chamberlain, J. Fox, B. Ashton, and J. Middleton, "Concise review: mesenchymal stem cells: their phenotype, differentiation capacity, immunological features, and potential for homing," *STEM CELLS*, vol. 25, no. 11, pp. 2739–2749, 2007.

[18] Y. Shi, G. Hu, J. Su, W. Li, Q. Chen, P. Shou, C. Xu, X. Chen, Y. Huang, Z. Zhu, X. Huang, X. Han, N. Xie, and G. Ren, "Mesenchymal stem cells: a new strategy for immunosuppression and tissue repair," *Cell Research*, vol. 20, no. 5, pp. 510–518, 2010.

[19] J. W. Lee, X. Fang, A. Krasnodembskaya, J. P. Howard, and M. A. Matthay, "Concise review: mesenchymal stem cells for acute lung injury: role of paracrine soluble factors," *STEM CELLS*, vol. 29, no. 6, pp. 913–919, 2011.

[20] S. Zeitouni, U. Krause, B. H. Clough, H. Halderman, A. Falster, D. T. Blalock, C. D. Chaput, H. W. Sampson, and C. A. Gregory, "Human mesenchymal stem cell-derived matrices for enhanced osteoregeneration," *Science Translational Medicine*, vol. 4, no. 132, p. 132ra55, 2012.

[21] K. Thavorn, S. van Katwyk, M. Krahn, S. H. J. Mei, D. J. Stewart, D. Fergusson, D. Coyle, and L. McIntyre, "Value of mesenchymal stem cell therapy for patients with septic shock: an early health economic evaluation," *International Journal of Technology Assessment in Health Care*, vol. 36, no. 5, pp. 525–532, 2020.

[22] D. G. Phinney, "Functional heterogeneity of mesenchymal stem cells: implications for cell therapy," *Journal of Cellular Biochemistry*, vol. 113, no. 9, pp. 2806–2812, 2012.

[23] B. Lukomska, L. Stanaszek, E. Zuba-Surma, P. Legosz, S. Sarzynska, and K. Drela, "Challenges and controversies in human mesenchymal stem cell therapy," *Stem Cells International*, vol. 2019, p. 9628536, 2019.

[24] O. Levy, R. Kuai, E. M. J. Siren, D. Bhere, Y. Milton, N. Nissar, M. D. Biasio, M. Heinelt, B. Reeve, R. Abdi, M. Alturki, M. Fallatah, A. Almalik, A. H. Alhasan, K. Shah, and J. M. Karp, "Shattering barriers toward clinically meaningful MSC therapies," *Science Advances*, vol. 6, no. 30, p. eaba6884, 2020.

[25] S. Torrents-Zapata, M. Codinach, L. Rodriguez, J. Fernandez-Sojo, M. Blanco, L. Alonso, G. Morón-Cazalilla, I. Elorza, G. Orti, R. Parody, A. Esquirol, M. Uría, M. Trabazo, M. Benitez, I. Badell, C. D. de Heredia, S. Querol, and J. Vives, "Finding meaning: potential relationship between critical quality attributes of MSC's and their clinical efficacy in GvHD," *Cytotherapy*, vol. 22, no. 5, Supplement, pp. S108–S109, 2020.

[26] U. Krause, A. Seckinger, and C. A. Gregory, "Assays of osteogenic differentiation by cultured human mesenchymal stem cells," in *Mesenchymal Stem Cell Assays and Applications* (M. Vemuri, L. G. Chase, and M. S. Rao, eds.), pp. 215–230, Totowa, NJ: Humana Press, 2011.

[27] J. L. Surdo and S. R. Bauer, "Quantitative approaches to detect donor and passage differences in adipogenic potential and clonogenicity in human bone marrow-derived mesenchymal stem cells," *Tissue Engineering Part C: Methods*, vol. 18, no. 11, pp. 877–889, 2012.

[28] J. Galipeau, M. Krampera, J. Barrett, F. Dazzi, R. J. Deans, J. DeBruijn, M. Dominici, W. E. Fibbe, A. P. Gee, J. M. Gimble, P. Hematti, M. B. Koh, K. LeBlanc, I. Martin, I. K. McNiece, M. Mendicino, S. Oh, L. Ortiz, D. G. Phinney, V. Planat, Y. Shi, D. F. Stroncek, S. Viswanathan, D. J. Weiss, and L. Sensebe, "International Society for Cellular Therapy perspective on immune functional assays for mesenchymal stromal cells as potency release criterion for advanced phase clinical trials," *Cytotherapy*, vol. 18, no. 2, pp. 151–159, 2016.

[29] P. Penfornis and R. Pochampally, "Colony forming unit assays," in *Mesenchymal Stem Cells: Methods and Protocols* (M. Gnecchi, ed.), pp. 159–169, New York, NY: Springer International Publishing, 2016.

[30] C. de Wolf, M. van de Bovenkamp, and M. Hoefnagel, "Regulatory perspective on in vitro potency assays for human mesenchymal stromal cells used in immunotherapy," *Cytotherapy*, vol. 19, no. 7, pp. 784–797, 2017.

[31] C. Wu, L. Chen, Y.-Z. Huang, Y. Huang, O. Parolini, Q. Zhong, X. Tian, and L. Deng, "Comparison of the proliferation and differentiation potential of human urine-, placenta decidua basalis-, and bone marrow-derived stem cells," *Stem Cells International*, vol. 2018, p. 7131532, 2018.

[32] B. Rocha, V. Calamia, J. Mateos, P. Fernández-Puente, F. J. Blanco, and C. Ruiz-Romero, "Metabolic labeling of human bone marrow mesenchymal stem cells for the quantitative

analysis of their chondrogenic differentiation," *Journal of Proteome Research*, vol. 11, no. 11, pp. 5350–5361, 2012.

[33] H. Jiang, T. Hong, T. Wang, X. Wang, L. Cao, X. Xu, and M. Zheng, "Gene expression profiling of human bone marrow mesenchymal stem cells during osteogenic differentiation," *Journal of Cellular Physiology*, vol. 234, no. 5, pp. 7070–7077, 2019.

[34] A. Allahverdi, E. Arefian, M. Soleimani, J. Ai, N. Nahanmoghaddam, A. Yousefi-Ahmadipour, and S. Ebrahimi-Barough, "MicroRNA-4731-5p delivered by AD-mesenchymal stem cells induces cell cycle arrest and apoptosis in glioblastoma," *Journal of Cellular Physiology*, vol. 235, no. 11, pp. 8167–8175, 2020.

[35] R. A. Marklein, J. L. L. Surdo, I. H. Bellayr, S. A. Godil, R. K. Puri, and S. R. Bauer, "High content imaging of early morphological signatures predicts long term mineralization capacity of human mesenchymal stem cells upon osteogenic induction," *STEM CELLS*, vol. 34, no. 4, pp. 935–947, 2016.

[36] M. W. Klinker, R. A. Marklein, J. L. L. Surdo, C.-H. Wei, and S. R. Bauer, "Morphological features of IFN-γ–stimulated mesenchymal stromal cells predict overall immunosuppressive capacity," *Proceedings of the National Academy of Sciences*, vol. 114, no. 13, pp. E2598–E2607, 2017.

[37] J. Lam, I. H. Bellayr, R. A. Marklein, S. R. Bauer, R. K. Puri, and K. E. Sung, "Functional profiling of chondrogenically induced multipotent stromal sell aggregates reveals transcriptomic and emergent morphological phenotypes predictive of differentiation capacity," *STEM CELLS Translational Medicine*, vol. 7, no. 9, pp. 664–675, 2018.

[38] R. A. Marklein, J. Lam, M. Guvendiren, K. E. Sung, and S. R. Bauer, "Functionally-relevant morphological profiling: a tool to assess cellular heterogeneity," *Trends in Biotechnology*, vol. 36, no. 1, pp. 105–118, 2018.

[39] R. A. Marklein, M. W. Klinker, K. A. Drake, H. G. Polikowsky, E. C. Lessey-Morillon, and S. R. Bauer, "Morphological profiling using machine learning reveals emergent subpopu-

lations of interferon-γ–stimulated mesenchymal stromal cells that predict immunosuppression," *Cytotherapy*, vol. 21, no. 1, pp. 17–31, 2019.

[40] H. Peng, "Bioimage informatics: a new area of engineering biology," *Bioinformatics*, vol. 24, no. 17, pp. 1827–1836, 2008.

[41] J. Rittscher, "Characterization of biological processes through automated image analysis," *Annual Review of Biomedical Engineering*, vol. 12, pp. 315–344, 2010.

[42] M. L. Giger, N. Karssemeijer, and J. A. Schnabel, "Breast image analysis for risk assessment, detection, diagnosis, and treatment of cancer," *Annual Review of Biomedical Engineering*, vol. 15, pp. 327–357, 2013.

[43] X. Zhou and S. T. Wong, "High content cellular imaging for drug development," *IEEE Signal Processing Magazine*, vol. 23, no. 2, pp. 170–174, 2006.

[44] P. Jannin, J. M. Fitzpatrick, D. J. Hawkes, X. Pennec, R. Shahidi, and M. W. Vannier, "Validation of medical image processing in image-guided therapy.," *IEEE Transactions on Medical Imaging*, vol. 21, no. 12, pp. 1445–1449, 2002.

[45] K. Li, Y. Fang, W. Li, C. Pan, P. Qin, Y. Zhong, X. Liu, M. Huang, Y. Liao, and S. Li, "CT image visual quantitative evaluation and clinical classification of coronavirus disease (COVID-19)," *European Radiology*, vol. 30, no. 8, pp. 4407–4416, 2020.

[46] G. Mohan and M. M. Subashini, "MRI based medical image analysis: survey on brain tumor grade classification," *Biomedical Signal Processing and Control*, vol. 39, pp. 139–161, 2018.

[47] M. Veta, J. P. W. Pluim, P. J. van Diest, and M. A. Viergever, "Breast cancer histopathology image analysis: a review," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 5, pp. 1400–1411, 2014.

[48] Y. Song, E.-L. Tan, X. Jiang, J.-Z. Cheng, D. Ni, S. Chen, B. Lei, and T. Wang, "Accurate cervical cell segmentation from overlapping clumps in pap smear images," *IEEE Transactions on Medical Imaging*, vol. 36, no. 1, pp. 288–300, 2017.

[49] J.-P. Thiran and B. Macq, "Morphological feature extraction for the classification of digital images of cancerous tissues," *IEEE Transactions on Biomedical Engineering*, vol. 43, no. 10, pp. 1011–1020, 1996.

[50] R. W. Veltri, K. Ashenayi, Y. Hu, and G. J. O'dowd, "Neural network for cell image analysis for identification of abnormal cells," Oct. 8 2002. US Patent 6,463,438.

[51] A. Sundstrom, E. Grabocka, D. Bar-Sagi, and B. Mishra, "Histological image processing features induce a quantitative characterization of chronic tumor hypoxia," *PLoS One*, vol. 11, no. 4, p. e0153623, 2016.

[52] B. L. Luck, K. D. Carlson, A. C. Bovik, and R. R. Richards-Kortum, "An image model and segmentation algorithm for reflectance confocal images of in vivo cervical tissue," *IEEE Transactions on Image Processing*, vol. 14, no. 9, pp. 1265–1276, 2005.

[53] M. A. Harris, A. N. Van, B. H. Malik, J. M. Jabbour, and K. C. Maitland, "A pulse coupled neural network segmentation algorithm for reflectance confocal images of epithelial tissue," *PLoS One*, vol. 10, no. 3, p. e0122368, 2015.

[54] J. Almotiri, K. Elleithy, and A. Elleithy, "A multi-anatomical retinal structure segmentation system for automatic eye screening using morphological adaptive fuzzy thresholding," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 6, pp. 1–23, 2018.

[55] D. Pastor, M. A. Luengo-Oroz, B. Lombardot, I. Gonzalvez, L. Duloquin, T. Savy, P. Bourgine, N. Peyrieras, and A. Santos, "Cell tracking in fluorescence images of embryogenesis processes with morphological reconstruction by 4D-tubular structuring elements," in *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 970–973, IEEE, 2009.

[56] V. Piuri and F. Scotti, "Morphological classification of blood leucocytes by microscope images," in *International Conference on Computational Intelligence for Measurement Systems and Applications (CIMSA)*, pp. 103–108, IEEE, 2004.

[57] M. J. Afridi, C. Liu, C. Chan, S. Baek, and X. Liu, "Image segmentation of mesenchymal stem cells in diverse culturing conditions," in *Winter Conference on Applications of Computer Vision (WACV)*, pp. 516–523, IEEE, 2014.

[58] F. Matsuoka, I. Takeuchi, H. Agata, H. Kagami, H. Shiono, Y. Kiyota, H. Honda, and R. Kato, "Morphology-based prediction of osteogenic differentiation potential of human mesenchymal stem cells," *PLoS One*, vol. 8, no. 2, p. e55082, 2013.

[59] K. Lim, S. H. Park, J. Kim, H. Seonwoo, P.-H. Choung, and J. H. Chung, "Cell image processing methods for automatic cell pattern recognition and morphological analysis of mesenchymal stem cells - an algorithm for cell classification and adaptive brightness correction," *Journal of Biosystems Engineering*, vol. 38, no. 1, pp. 55–63, 2013.

[60] V. K. Lam, T. C. Nguyen, V. Bui, B. M. Chung, L.-C. Chang, G. Nehmetallah, and C. B. Raub, "Quantitative scoring of epithelial and mesenchymal qualities of cancer cells using machine learning and quantitative phase imaging," *Journal of Biomedical Optics*, vol. 25, no. 2, pp. 1–17, 2020.

[61] C. Seiler, A. Gazdhar, M. Reyes, L. M. Benneker, T. Geiser, K. A. Siebenrock, and B. Gantenbein-Ritter, "Time-lapse microscopy and classification of 2D human mesenchymal stem cells based on cell shape picks up myogenic from osteogenic and adipogenic differentiation," *Journal of Tissue Engineering and Regenerative Medicine*, vol. 8, no. 9, pp. 737–746, 2014.

[62] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[63] M. Ravi, V. Paramesh, S. Kaviya, E. Anuradha, and F. P. Solomon, "3D cell culture systems: advantages and applications," *Journal of Cellular Physiology*, vol. 230, no. 1, pp. 16–26, 2015.

[64] E. Buzhor, L. Leshansky, J. Blumenthal, H. Barash, D. Warshawsky, Y. Mazor, and R. Shtrichman, "Cell-based therapy approaches: the hope for incurable diseases," *Regen-*

erative Medicine, vol. 9, no. 5, pp. 649–672, 2014.

[65] M. A. Fischbach, J. A. Bluestone, and W. A. Lim, "Cell-based therapeutics: the next pillar of medicine," *Science Translational Medicine*, vol. 5, no. 179, p. 179ps7, 2013.

[66] S. Guedan, M. Ruella, and C. H. June, "Emerging cellular therapies for cancer," *Annual Review of Immunology*, vol. 37, no. 1, pp. 145–171, 2019.

[67] N. H. Goradel, F. G. Hour, B. Negahdari, Z. V. Malekshahi, M. Hashemzehi, A. Masoudifar, and H. Mirzaei, "Stem cell therapy: a new therapeutic option for cardiovascular diseases," *Journal of Cellular Biochemistry*, vol. 119, no. 1, pp. 95–104, 2018.

[68] X. Hu, R. K. Leak, A. W. Thomson, F. Yu, Y. Xia, L. R. Wechsler, and J. Chen, "Promises and limitations of immune cell-based therapies in neurological disorders," *Nature Reviews Neurology*, vol. 14, no. 9, pp. 559–568, 2018.

[69] D. A. Papazova, N. R. Oosterhuis, H. Gremmels, A. van Koppen, J. A. Joles, and M. C. Verhaar, "Cell-based therapies for experimental chronic kidney disease: a systematic review and meta-analysis," *Disease Models & Mechanisms*, vol. 8, no. 3, pp. 281–293, 2015.

[70] I. V. Kholodenko and K. N. Yarygin, "Cellular mechanisms of liver regeneration and cell-based therapies of liver diseases," *BioMed Research International*, vol. 2017, p. 8910821, 2017.

[71] E. K. Zuba-Surma, W. Wojakowski, Z. Madeja, and M. Z. Ratajczak, "Stem cells as a novel tool for drug screening and treatment of degenerative diseases," *Current Pharmaceutical Design*, vol. 18, no. 18, pp. 2644–2656, 2012.

[72] J. D. McNeish, "Stem cells as screening tools in drug discovery," *Current Opinion in Pharmacology*, vol. 7, no. 5, pp. 515–520, 2007.

[73] S. M. Sarkaria, M. Decker, and L. Ding, "Bone marrow micro-environment in normal and deranged hematopoiesis: opportunities for regenerative medicine and therapies," *BioEssays*, vol. 40, no. 3, p. 1700190, 2018.

[74] F. Tekinturhan, L. Zimmerlin, V. S. Donnenberg, M. E. Pfeifer, D. A. Monlish, B. Guvenc, J. Bradley, and A. D. Donnenberg, "Characterization of hematopoietic and non-hematopoietic stem/progenitor cells in freshly isolated adult human bone marrow using an 8-color flow cytometric assay.," *Blood*, vol. 110, no. 11, p. 4045, 2007.

[75] E. M. Horwitz, M. Andreef, and F. Frassoni, "Mesenchymal stromal cells," *Current Opinion in Hematology*, vol. 13, no. 6, pp. 419–425, 2006.

[76] N. W. Marion and J. J. Mao, "Mesenchymal stem cells and tissue engineering," *Methods in Enzymology*, vol. 420, pp. 339–361, 2006.

[77] Y. Han, X. Li, Y. Zhang, Y. Han, F. Chang, and J. Ding, "Mesenchymal stem cells for regenerative medicine," *Cells*, vol. 8, no. 8, p. 886, 2019.

[78] B. Amorin, A. P. Alegretti, V. Valim, A. Pezzi, A. M. Laureano, M. A. L. da Silva, A. Wieck, and L. Silla, "Mesenchymal stem cell therapy and acute graft-versus-host disease: a review," *Human Cell*, vol. 27, no. 4, pp. 137–150, 2014.

[79] R. Thakker and P. Yang, "Mesenchymal stem cell therapy for cardiac repair," *Current Treatment Options in Cardiovascular Medicine*, vol. 16, no. 7, p. 323, 2014.

[80] F. Rad, M. Ghorbani, A. Mohammadi Roushandeh, and M. Habibi Roudkenar, "Mesenchymal stem cell-based therapy for autoimmune diseases: emerging roles of extracellular vesicles," *Molecular Biology Reports*, vol. 46, no. 1, pp. 1533–1549, 2019.

[81] Y. Yang, Y. Zhao, L. Zhang, F. Zhang, and L. Li, "The application of mesenchymal stem cells in the treatment of liver diseases: mechanism, efficacy, and safety issues," *Frontiers in Medicine*, vol. 8, p. 642, 2021.

[82] L. A. Costa, N. Eiro, M. Fraile, L. O. Gonzalez, J. Saá, P. Garcia-Portabella, B. Vega, J. Schneider, and F. J. Vizoso, "Functional heterogeneity of mesenchymal stem cells from natural niches to culture conditions: implications for further clinical uses," *Cellular and Molecular Life Sciences*, vol. 78, no. 2, pp. 447–467, 2021.

[83] D. K. W. Ocansey, B. Pei, Y. Yan, H. Qian, X. Zhang, W. Xu, and F. Mao, "Improved therapeutics of modified mesenchymal stem cells: an update," *Journal of Translational Medicine*, vol. 18, no. 1, p. 42, 2020.

[84] D. García-Bernal, M. García-Arranz, R. M. Yáñez, R. Hervás-Salcedo, A. Cortés, M. Fernández-García, M. Hernando-Rodríguez, Ó. Quintana-Bustamante, J. A. Bueren, D. García-Olmo, J. M. Moraleda, J. C. Segovia, and A. G. Zapata, "The current status of mesenchymal stromal cells: controversies, unresolved issues and some promising solutions to improve their therapeutic efficacy," *Frontiers in Cell and Developmental Biology*, vol. 9, p. 609, 2021.

[85] C. M. DiGirolamo, D. Stokes, D. Colter, D. G. Phinney, R. Class, and D. J. Prockop, "Propagation and senescence of human marrow stromal cells in culture: a simple colony-forming assay identifies samples with the greatest potential to propagate and differentiate," *British Journal of Haematology*, vol. 107, no. 2, pp. 275–281, 1999.

[86] D. C. Colter, I. Sekiya, and D. J. Prockop, "Identification of a subpopulation of rapidly self-renewing and multipotential adult stem cells in colonies of human marrow stromal cells," *Proceedings of the National Academy of Sciences*, vol. 98, no. 14, pp. 7841–7845, 2001.

[87] F. Haasters, W. C. Prall, D. Anz, C. Bourquin, C. Pautke, S. Endres, W. Mutschler, D. Docheva, and M. Schieker, "Morphological and immunocytochemical characteristics indicate the yield of early progenitors and represent a quality control for human mesenchymal stem cell culturing," *Journal of Anatomy*, vol. 214, no. 5, pp. 759–767, 2009.

[88] I. Sekiya, B. L. Larson, J. R. Smith, R. Pochampally, J.-G. Cui, and D. J. Prockop, "Expansion of human adult stem cells from bone marrow stroma: conditions that maximize the yields of early progenitors and evaluate their quality," *STEM CELLS*, vol. 20, no. 6, pp. 530–541, 2002.

[89] R. H. Lee, S. C. Hsu, J. Munoz, J. S. Jung, N. R. Lee, R. Pochampally, and D. J. Prockop, "A subset of human rapidly self-renewing marrow stromal cells preferentially engraft in mice,"

Blood, vol. 107, no. 5, pp. 2153–2161, 2006.

[90] R. H. Lee, M. J. Seo, A. A. Pulin, C. A. Gregory, J. Ylostalo, and D. J. Prockop, "The CD34-like protein PODXL and α6-integrin (CD49f) identify early progenitor MSCs with increased clonogenicity and migration to infarcted heart in mice," *Blood*, vol. 113, no. 4, pp. 816–826, 2009.

[91] J. R. Smith, R. Pochampally, A. Perry, S. C. Hsu, and D. J. Prockop, "Isolation of a highly clonogenic and multipotential subfraction of adult stem cells from bone marrow stroma," *STEM CELLS*, vol. 22, no. 5, pp. 823–831, 2004.

[92] C. A. Gregory and D. J. Prockop, "Fundamentals of culture and characterization of mesenchymal stem/progenitor cells (MSCs) from bone marrow stroma," in *Culture of Human Stem Cells* (R. I. Freshney, G. N. Stacey, and J. M. Auerbach, eds.), ch. 9, pp. 207–232, Hoboken: John Wiley & Sons, Ltd., 2007.

[93] Q. Zhao, C. A. Gregory, R. H. Lee, R. L. Reger, L. Qin, B. Hai, M. S. Park, N. Yoon, B. Clough, E. McNeill, D. J. Prockop, and F. Liu, "MSCs derived from iPSCs with a modified protocol are tumor-tropic but have much less potential to promote tumors than bone marrow MSCs," *Proceedings of the National Academy of Sciences*, vol. 112, no. 2, pp. 530–535, 2015.

[94] C. Zimmer, E. Labruyere, V. Meas-Yedid, N. Guillen, and J.-C. Olivo-Marin, "Segmentation and tracking of migrating cells in videomicroscopy with parametric active contours: a tool for cell-based drug testing," *IEEE Transactions on Medical Imaging*, vol. 21, no. 10, pp. 1212–1221, 2002.

[95] K. Li, E. D. Miller, M. Chen, T. Kanade, L. E. Weiss, and P. G. Campbell, "Cell population tracking and lineage construction with spatiotemporal context," *Medical Image Analysis*, vol. 12, no. 5, pp. 546–566, 2008.

[96] E. Mendoz and C. T. Lim, "Collective migration behaviors of human breast cancer cells in 2D," *Cellular and Molecular Bioengineering*, vol. 4, no. 3, p. 411, 2011.

[97] I. Seroussi, D. Veikherman, N. Ofer, S. Yehudai-Resheff, and K. Keren, "Segmentation and tracking of live cells in phase-contrast images using directional gradient vector flow for snakes," *Journal of Microscopy*, vol. 247, no. 2, pp. 137–146, 2012.

[98] J. Chalfoun, M. Kociolek, A. Dima, M. Halter, A. Cardone, A. Peskin, P. Bajcsy, and M. Brady, "Segmenting time-lapse phase contrast images of adjacent NIH 3T3 cells," *Journal of Microscopy*, vol. 249, no. 1, pp. 41–52, 2013.

[99] J. Chalfoun, M. Majurski, A. Dima, C. Stuelten, A. Peskin, and M. Brady, "FogBank: a single cell segmentation across multiple cell lines and image modalities," *BMC Bioinformatics*, vol. 15, no. 1, p. 431, 2014.

[100] R. Dixit and R. Cyr, "Cell damage and reactive oxygen species production induced by fluorescence microscopy: effect on mitosis and guidelines for non-invasive fluorescence microscopy," *The Plant Journal*, vol. 36, no. 2, pp. 280–290, 2003.

[101] R. M. Lasarow, R. R. Isseroff, and E. C. Gomez, "Quantitative in vitro assessment of phototoxicity by a fibroblast-neutral red assay," *Journal of Investigative Dermatology*, vol. 98, no. 5, pp. 725–729, 1992.

[102] N. Jaccard, L. D. Griffin, A. Keser, R. J. Macown, A. Super, F. S. Veraitch, and N. Szita, "Automated method for the rapid and precise estimation of adherent cell culture characteristics from phase contrast microscopy images," *Biotechnology and Bioengineering*, vol. 111, no. 3, pp. 504–517, 2014.

[103] S. A. Johnson, "Chapter 1 - phase contrast microscopy," in *Biomedical Optical Phase Microscopy and Nanoscopy* (N. T. Shaked, Z. Zalevsky, and L. L. Satterwhite, eds.), ch. 1, pp. 3–18, Oxford: Academic Press, 2013.

[104] R. E. Rogers, A. Haskell, B. P. White, S. Dalal, M. Lopez, D. Tahan, S. Pan, G. Kaur, H. Kim, H. Barreda, S. L. Woodard, O. R. Benavides, J. Dai, Q. Zhao, K. C. Maitland, A. Han, Z. L. Nikolov, F. Liu, R. H. Lee, C. A. Gregory, and R. Kaunas, "A scalable system

for generation of mesenchymal stem cells derived from induced pluripotent cells employing bioreactors and degradable microcarriers," *STEM CELLS Translational Medicine*, 20.

[105] B. W. Graf and S. A. Boppart, "Imaging and analysis of three-dimensional cell culture models," in *Live Cell Imaging: Methods and Protocols* (D. B. Papkovsky, ed.), pp. 211–227, Totowa, NJ: Humana Press, 2010.

[106] A. G. York, S. H. Parekh, D. D. Nogare, R. S. Fischer, K. Temprine, M. Mione, A. B. Chitnis, C. A. Combs, and H. Shroff, "Resolution doubling in live, multicellular organisms via multifocal structured illumination microscopy," *Nature Methods*, vol. 9, no. 7, pp. 749–754, 2012.

[107] F. Pampaloni, N. Ansari, and E. H. K. Stelzer, "High-resolution deep imaging of live cellular spheroids with light-sheet-based fluorescence microscopy," *Cell and Tissue Research*, vol. 352, no. 1, pp. 161–177, 2013.

[108] E. J. Gualda, H. Pereira, G. G. Martins, R. Gardner, and N. Moreno, "Three-dimensional imaging flow cytometry through light-sheet fluorescence microscopy," *Cytometry Part A*, vol. 91, no. 2, pp. 144–151, 2017.

[109] E. H. K. Stelzer, "Light-sheet fluorescence microscopy for quantitative biology," *Nature Methods*, vol. 12, no. 1, pp. 23–26, 2015.

[110] E. H. K. Stelzer, F. Strobl, B.-J. Chang, F. Preusser, S. Preibisch, K. McDole, and R. Fiolka, "Light sheet fluorescence microscopy," *Nature Reviews Methods Primers*, vol. 1, no. 1, p. 73, 2021.

[111] I. Albert-Smet, A. Marcos-Vidal, J. J. Vaquero, M. Desco, A. Muñoz-Barrutia, and J. Ripoll, "Applications of light-sheet microscopy in microdevices," *Frontiers in Neuroanatomy*, vol. 13, p. 1, 2019.

[112] F. Pampaloni, B.-J. Chang, and E. H. Stelzer, "Light sheet-based fluorescence microscopy (LSFM) for the quantitative imaging of cells and tissues," *Cell and Tissue Research*, vol. 360, no. 1, pp. 129–141, 2015.

[113] A. Schmitz, S. C. Fischer, C. Mattheyer, F. Pampaloni, and E. H. K. Stelzer, "Multiscale image analysis reveals structural heterogeneity of the cell microenvironment in homotypic spheroids," *Scientific Reports*, vol. 7, no. 1, p. 43693, 2017.

[114] B. Eismann, T. G. Krieger, J. Beneke, R. Bulkescher, L. Adam, H. Erfle, C. Herrmann, R. Eils, and C. Conrad, "Automated 3D light-sheet screening with high spatiotemporal resolution reveals mitotic phenotypes," *Journal of Cell Science*, vol. 133, no. 11, p. jcs245043, 2020.

[115] F. Sala, M. Castriotta, P. Paiè, A. Farina, S. D'Annunzio, A. Zippo, R. Osellame, F. Bragheri, and A. Bassi, "High-throughput 3D imaging of single cells with light-sheet fluorescence microscopy on chip," *Biomedical Optics Express*, vol. 11, no. 8, pp. 4397–4407, 2020.

[116] A. Diosdi, D. Hirling, M. Kovacs, T. Toth, M. Harmati, K. Koos, K. Buzas, F. Piccinini, and P. Horvath, "Cell lines and clearing approaches: a single-cell level 3D light-sheet fluorescence microscopy dataset of multicellular spheroids," *Data in Brief*, vol. 36, p. 107090, 2021.

[117] G. P. Way, M. Kost-Alimova, T. Shibue, W. F. Harrington, S. Gill, F. Piccioni, T. Becker, H. Shafqat-Abbasi, W. C. Hahn, A. E. Carpenter, F. Vazquez, and S. Singh, "Predicting cell health phenotypes using image-based morphology profiling," *Molecular Biology of the Cell*, vol. 32, no. 9, pp. 995–1005, 2021.

[118] S. Chen, M. Zhao, G. Wu, C. Yao, and J. Zhang, "Recent advances in morphological cell image analysis," *Computational and Mathematical Methods in Medicine*, vol. 2012, p. 101536, 2012.

[119] Y. Wang, Z. Zhang, H. Wang, and S. Bi, "Segmentation of the clustered cells with optimized boundary detection in negative phase contrast images," *PloS One*, vol. 10, no. 6, p. e0130178, 2015.

[120] N. Malpica, C. O. de Solórzano, J. J. Vaquero, A. Santos, I. Vallcorba, J. M. García-Sagredo, and F. del Pozo, "Applying watershed algorithms to the segmentation of clustered nuclei,"

Cytometry Part A, vol. 28, no. 4, pp. 289–297, 1997.

[121] K. Parvati, P. Rao, and M. Mariya Das, "Image segmentation using gray-scale morphology and marker-controlled watershed transformation," *Discrete Dynamics in Nature and Society*, vol. 2008, p. 384346, 2008.

[122] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: a brief review," *Computational Intelligence and Neuroscience*, vol. 2018, p. 7068349, 2018.

[123] R. Zemouri, N. Zerhouni, and D. Racoceanu, "Deep learning in the biomedical applications: recent and future status," *Applied Sciences*, vol. 9, no. 8, p. 1526, 2019.

[124] S. K. Sadanandan, P. Ranefall, S. L. Guyader, and C. Wählby, "Automated training of deep convolutional neural networks for cell segmentation," *Scientific Reports*, vol. 7, no. 1, p. 7860, 2017.

[125] E. Moen, D. Bannon, T. Kudo, W. Graf, M. Covert, and D. V. Valen, "Deep learning for cellular image analysis," *Nature Methods*, vol. 16, no. 12, pp. 1233–1246, 2019.

[126] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: learning dense volumetric segmentation from sparse annotation," in *Medical Image Computing and Computer Assisted Intervention (MICCAI)* (S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, eds.), (Cham), pp. 424–432, Springer International Publishing, 2016.

[127] F. Nargesian, H. Samulowitz, U. Khurana, E. B. Khalil, and D. S. Turaga, "Learning feature engineering for classification.," in *26th International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 2529–2535, IJCAI, 2017.

[128] J. Liu and Y. Shi, "Image feature extraction method based on shape characteristics and its application in medical image analysis," in *Applied Informatics and Communication* (D. Zeng, ed.), (Berlin, Heidelberg), pp. 172–178, Springer International Publishing, 2011.

[129] M. Arya, N. Mittal, and G. Singh, "Texture-based feature extraction of smear images for the detection of cervical cancer," *IET Computer Vision*, vol. 12, no. 8, pp. 1049–1059, 2018.

[130] F. Meyer and S. Beucher, "Morphological segmentation," *Journal of Visual Communication and Image Representation*, vol. 1, no. 1, pp. 21–46, 1990.

[131] E. Bengtsson, C. Wählby, and J. Lindblad, "Robust cell image segmentation methods," *Pattern Recognition and Image Analysis*, vol. 14, no. 2, pp. 157–167, 2004.

[132] N. Malpica, C. O. de Solorzano, J. J. Vaquero, A. Santos, I. Vallcorba, J. M. García-Sagredo, and F. Del Pozo, "Applying watershed algorithms to the segmentation of clustered nuclei," *Cytometry*, vol. 28, no. 4, pp. 289–297, 1997.

[133] V. Yeghiazaryan and I. Voiculescu, "Family of boundary overlap metrics for the evaluation of medical image segmentation," *Journal of Medical Imaging*, vol. 5, no. 01, p. 015006, 2018.

[134] Y. Mingqiang, K. Kidiyo, and R. Joseph, "A survey of shape feature extraction techniques," *Pattern Recognition*, vol. 15, no. 7, pp. 43–90, 2008.

[135] A. R. Sibley, M. L. Giger, V. Liarski, and M. Clark, "Simultaneous segmentation and classification of multichannel immuno-fluorescently labeled confocal microscopy images using deep convolutional neural networks," in *Medical Imaging 2018: Digital Pathology*, vol. 10581, p. 1058110, International Society for Optics and Photonics, 2018.

[136] C. E. Metz, "Roc analysis in medical imaging: a tutorial review of the literature," *Radiological Physics and Technology*, vol. 1, no. 1, pp. 2–12, 2008.

[137] A. M. Hafiz and G. M. Bhat, "A survey on instance segmentation: state of the art," *International Journal of Multimedia Information Retrieval*, vol. 9, no. 3, pp. 171–189, 2020.

[138] C. F. Koyuncu, E. Akhan, T. Ersahin, R. Cetin-Atalay, and C. Gunduz-Demir, "Iterative h-minima-based marker-controlled watershed for cell nucleus segmentation," *Cytometry Part A*, vol. 89, no. 4, pp. 338–349, 2016.

[139] C. Jung and C. Kim, "Segmenting clustered nuclei using H-minima transform-based marker extraction and contour parameterization," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 10, pp. 2600–2604, 2010.

[140] M. S. Rahman and M. R. Islam, "Counting objects in an image by marker controlled watershed segmentation and thresholding," in *3rd International Advance Computing Conference (IACC)*, pp. 1251–1256, IEEE, 2013.

[141] M. Veluchamy, K. Perumal, and T. Ponuchamy, "Feature extraction and classification of blood cells using artificial neural network," *American Journal of Applied Sciences*, vol. 9, no. 5, pp. 615–619, 2012.

[142] L.-K. Soh and C. Tsatsoulis, "Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices," *EEE Transactions on Geoscience and Remote Sensing*, vol. 37, no. 2, pp. 780–795, 1999.

[143] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern Recognition*, vol. 30, no. 7, pp. 1145–1159, 1997.

[144] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.

[145] "Metz ROC software." `http://metz-roc.uchicago.edu`, 2011.

[146] T. Fushiki, "Estimation of prediction error by using K-fold cross-validation," *Statistics and Computing*, vol. 21, no. 2, pp. 137–146, 2011.

[147] H. M. Whitney, H. Li, Y. Ji, P. Liu, and M. L. Giger, "Comparison of breast MRI tumor classification using human-engineered radiomics, transfer learning from deep convolutional neural networks, and fusion methods," *Proceedings of the IEEE*, vol. 108, no. 1, pp. 163–177, 2020.

[148] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic brain tumor detection and segmentation using U-Net based fully convolutional networks," in *Annual Conference on Medical Image Understanding and Analysis (MIUA)*, pp. 506–517, Springer International Publishing, 2017.

[149] S. Li, M. Dong, G. Du, and X. Mu, "Attention Dense-U-Net for automatic breast mass segmentation in digital mammogram," *IEEE Access*, vol. 7, pp. 59037–59047, 2019.

[150] Q. Li, S. Fan, and C. Chen, "An intelligent segmentation and diagnosis method for diabetic retinopathy based on improved U-Net network," *Journal of Medical Systems*, vol. 43, no. 9, p. 304, 2019.

[151] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (D. Stoyanov, Z. Taylor, G. Carneiro, T. Syeda-Mahmood, A. Martel, L. Maier-Hein, J. M. R. Tavares, A. Bradley, J. P. Papa, V. Belagiannis, J. C. Nascimento, Z. Lu, S. Conjeti, M. Moradi, H. Greenspan, and A. Madabhushi, eds.), (Cham), pp. 3–11, Springer International Publishing, 2018.

[152] X. Kong, G. Sun, Q. Wu, J. Liu, and F. Lin, "Hybrid pyramid U-Net model for brain tumor segmentation," in *Intelligent Information Processing IX* (Z. Shi, E. Mercier-Laurent, and J. Li, eds.), (Cham), pp. 346–355, Springer International Publishing, 2018.

[153] H. Liu, X. Shen, F. Shang, F. Ge, and F. Wang, "CU-Net: Cascaded U-Net with loss weighted sampling for brain tumor segmentation," in *Multimodal Brain Image Analysis and Mathematical Foundations of Computational Anatomy* (D. Zhu, J. Yan, H. Huang, L. Shen, P. M. Thompson, C.-F. Westin, X. Pennec, S. Joshi, M. Nielsen, T. Fletcher, S. Durrleman, and S. Sommer, eds.), (Cham), pp. 102–111, Springer International Publishing, 2019.

[154] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in *International Conference on Computer Vision (ICCV)*, pp. 2980–2988, IEEE, 2017.

[155] Z. Zeng, W. Xie, Y. Zhang, and Y. Lu, "RIC-Unet: an improved neural network based on Unet for nuclei segmentation in histology images," *IEEE Access*, vol. 7, pp. 21420–21428, 2019.

[156] S. Cai, Y. Tian, H. Lui, H. Zeng, Y. Wu, and G. Chen, "Dense-UNet: a novel multiphoton in vivo cellular image segmentation model based on a convolutional neural network,"

Quantitative Imaging in Medicine and Surgery, vol. 10, no. 6, pp. 1275–1285, 2020.

[157] H. He, C. Zhang, J. Chen, R. Geng, L. Chen, Y. Liang, Y. Lu, J. Wu, and Y. Xu, "A hybrid-attention nested UNet for nuclear segmentation in histopathological images," *Frontiers in Molecular Biosciences*, vol. 8, p. 614174, 2021.

[158] F. Lux and P. Matula, "DIC image segmentation of dense cell populations by combining deep learning and watershed," in *16th International Symposium on Biomedical Imaging (ISBI)*, pp. 236–239, IEEE, 2019.

[159] F. Lux and P. Matula, "Cell segmentation by combining marker-controlled watershed and deep learning," *arXiv preprint*, 2020.

[160] S. Jadon, "A survey of loss functions for semantic segmentation," in *Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, pp. 1–7, IEEE, 2020.

[161] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso, "Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (M. J. Cardoso, T. Arbel, G. Carneiro, T. Syeda-Mahmood, J. M. R. Tavares, M. Moradi, A. Bradley, H. Greenspan, J. P. Papa, A. Madabhushi, J. C. Nascimento, J. S. Cardoso, V. Belagiannis, and Z. Lu, eds.), (Cham), pp. 240–248, Springer International Publishing, 2017.

[162] D. P. Kingma and J. Ba, "Adam: a method for stochastic optimization," *arXiv preprint*, 2014.

[163] A. Ross and V. L. Willson, "Paired samples t-test," in *Basic and Advanced Statistical Tests*, ch. 4, pp. 17–19, Rotterdam: SensePublishers, 2017.

[164] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian Journal of Statistics*, vol. 6, no. 2, pp. 65–70, 1979.

[165] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J.-c. Sanchez, and M. Müller, "pROC: an open-source package for R and S+ to analyze and compare ROC curves," *BMC Bioinformatics*, vol. 12, no. 1, p. 77, 2011.

[166] E. R. DeLong, D. M. DeLong, and D. L. Clarke-Pearson, "Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach," *Biometrics*, vol. 44, no. 3, pp. 837–845, 1988.

[167] A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, and D. M. Sabatini, "CellProfiler: image analysis software for identifying and quantifying cell phenotypes," *Genome Biology*, vol. 7, no. 10, p. R100, 2006.

[168] E. Gómez-de-Mariscal, C. García-López-de-Haro, W. Ouyang, L. Donati, E. Lundberg, M. Unser, A. Muñoz-Barrutia, and D. Sage, "DeepImageJ: a user-friendly environment to run deep learning models in ImageJ," *bioRxiv preprint*, 2021.

[169] T. R. Heathman, A. W. Nienow, M. J. McCall, K. Coopman, B. Kara, and C. J. Hewitt, "The translation of cell-based therapies: clinical landscape and manufacturing challenges," *Regenerative Medicine*, vol. 10, no. 1, pp. 49–64, 2015.

[170] M. Kapałczyńska, T. Kolenda, W. Przybyła, M. Zajączkowska, A. Teresiak, V. Filas, M. Ibbs, R. Bliźniak, Ł. Łuczewski, and K. Lamperska, "2D and 3D cell cultures - a comparison of different types of cancer cell cultures," *Archives of Medical Science*, vol. 14, no. 4, pp. 910–919, 2018.

[171] D. Antoni, H. Burckel, E. Josset, and G. Noel, "Three-dimensional cell culture: a breakthrough in vivo," *International Journal of Molecular Sciences*, vol. 16, no. 3, pp. 5517–5527, 2015.

[172] J. E. Frith, B. Thomson, and P. G. Genever, "Dynamic three-dimensional culture methods enhance mesenchymal stem cell properties and increase therapeutic potential," *Tissue Engineering Part C: Methods*, vol. 16, no. 4, pp. 735–749, 2010.

[173] J. F. W. Greiner, B. Kaltschmidt, C. Kaltschmidt, and D. Widera, "Going 3D - cell culture approaches for stem cell research and therapy," *Current Tissue Engineering*, vol. 2, no. 1, pp. 8–19, 2013.

[174] N. Chaicharoenaudomrung, P. Kunhorm, and P. Noisa, "Three-dimensional cell culture systems as an in vitro platform for cancer and stem cell modeling," *World Journal of Stem Cells*, vol. 11, no. 12, pp. 1065–1083, 2019.

[175] A. Shekaran, E. Sim, K. Y. Tan, J. K. Y. Chan, M. Choolani, S. Reuveny, and S. Oh, "Enhanced in vitro osteogenic differentiation of human fetal MSCs attached to 3D microcarriers versus harvested from 2D monolayers," *BMC Biotechnology*, vol. 15, no. 1, pp. 1–13, 2015.

[176] L. Yan, B. Jiang, E. Li, X. Wang, Q. Ling, D. Zheng, J. W. Park, X. Chen, E. Cheung, X. Du, Y. Li, G. Cheng, E. He, and R.-H. Xu, "Scalable generation of mesenchymal stem cells from human embryonic stem cells in 3D," *International Journal of Biological Sciences*, vol. 14, no. 10, pp. 1196–1210, 2018.

[177] C. C. Bilgin, G. Fontenay, Q. Cheng, H. Chang, J. Han, and B. Parvin, "BioSig3D: high content screening of three-dimensional cell culture models," *PloS One*, vol. 11, no. 3, p. e0148379, 2016.

[178] M.-O. Baradez and D. Marshall, "The use of multidimensional image-based analysis to accurately monitor cell growth in 3D bioreactor culture," *PLoS One*, vol. 6, no. 10, p. e26104, 2011.

[179] Z. Di, M. J. D. Klop, V.-M. Rogkoti, S. E. L. Dévédec, B. van de Water, F. J. Verbeek, L. S. Price, and J. H. N. Meerman, "Ultra high content image analysis and phenotype profiling of 3D cultured micro-tissues," *PloS One*, vol. 9, no. 10, p. e109688, 2014.

[180] F. Piccinini, T. Balassa, A. Carbonaro, A. Diosdi, T. Toth, N. Moshkov, E. A. Tasnadi, and P. Horvath, "Software tools for 3D nuclei segmentation and quantitative analysis in multicellular aggregates," *Computational and Structural Biotechnology Journal*, vol. 18, pp. 1287–1300, 2020.

[181] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[182] Y. Feng, H. Zhao, X. Li, X. Zhang, and H. Li, "A multi-scale 3D Otsu thresholding algorithm for medical image segmentation," *Digital Signal Processing*, vol. 60, pp. 186–199, 2017.

[183] N. Siddique, P. Sidike, C. Elkin, and V. Devabhaktuni, "U-Net and its variants for medical image segmentation: theory and applications," *arXiv preprint*, 2020.

[184] S. Öz and Y. S. Doğrusöz, "Fast 3D reconstruction from medical image series based on thresholding method," in *15th National Biomedical Engineering Meeting (BIYOMUT)*, pp. 1–4, IEEE, 2010.

[185] P. Kalavathi, "Brain tissue segmentation in MR brain images using multiple Otsu's thresholding technique," in *8th International Conference on Computer Science Education*, pp. 639–642, IEEE, 2013.

[186] C. Kim, J. Yoon, and Y.-J. Lee, "Medical image segmentation by more sensitive adaptive thresholding," in *6th International Conference on IT Convergence and Security (ICITCS)*, pp. 1–3, 2016.

[187] G. Du, X. Cao, J. Liang, X. Chen, and Y. Zhan, "Medical image segmentation based on U-Net: a review," *Journal of Imaging Science and Technology*, vol. 64, no. 2, pp. 20508–1–20508–12, 2020.

[188] M. Noori, A. Bahri, and K. Mohammadi, "Attention-guided version of 2D UNet for automatic brain tumor segmentation," in *9th International Conference on Computer and Knowledge Engineering (ICCKE)*, pp. 269–275, 2019.

[189] Z. Xiao, B. Liu, L. Geng, F. Zhang, and Y. Liu, "Segmentation of lung nodules using improved 3D-UNet neural network," *Symmetry*, vol. 12, no. 11, p. 1787, 2020.

[190] H.-G. Nguyen, A. Pica, P. Maeder, A. Schalenbourg, M. Peroni, J. Hrbacek, D. C. Weber, M. B. Cuadra, and R. Sznitman, "Ocular structures segmentation from multi-sequences MRI using 3D Unet with fully connected CRFs," in *Computational Pathology and Ophthalmic*

Medical Image Analysis (D. Stoyanov, Z. Taylor, F. Ciompi, Y. Xu, A. Martel, L. Maier-Hein, N. Rajpoot, J. van der Laak, M. Veta, S. McKenna, D. Snead, E. Trucco, M. K. Garvin, X. J. Chen, and H. Bogunovic, eds.), (Cham), pp. 167–175, Springer International Publishing, 2018.

[191] N. Zettler and A. Mastmeyer, "Comparison of 2D vs. 3D U-Net organ segmentation in abdominal 3D CT images," *arXiv preprint*, 2021.

[192] C. Martin, É. Olmos, M.-L. Collignon, N. D. Isla, F. Blanchard, I. Chevalot, A. Marc, and E. Guedon, "Revisiting MSC expansion from critical quality attributes to critical culture process parameters," *Process Biochemistry*, vol. 59, pp. 231–243, 2017.

[193] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint*, 2014.

[194] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 248–255, IEEE, 2009.

[195] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," *BMC Medical Imaging*, vol. 15, no. 1, p. 29, 2015.

[196] T. Atta-Fosu, W. Guo, D. Jeter, C. M. Mizutani, N. Stopczynski, and R. Sousa-Neves, "3D clumped cell segmentation using curvature based seeded watershed," *Journal of Imaging*, vol. 2, no. 4, p. 31, 2016.

[197] H.-J. Choi and H.-K. Choi, "Grading of renal cell carcinoma by 3D morphological analysis of cell nuclei," *Computers in Biology and Medicine*, vol. 37, no. 9, pp. 1334–1341, 2007.

[198] H. Takko, C. Pajanoja, K. Kurtzeborn, J. Hsin, S. Kuure, and L. Kerosuo, "ShapeMetrics: a userfriendly pipeline for 3D cell segmentation and spatial tissue analysis," *Developmental Biology*, vol. 462, no. 1, pp. 7–19, 2020.

[199] S. Chakraborty, S. Aich, and H.-C. Kim, "3D textural, morphological and statistical analysis of voxel of interests in 3T MRI scans for the detection of Parkinson's disease using artificial neural networks," *Healthcare*, vol. 8, no. 1, p. 34, 2020.

[200] P. A. Kenny, G. Y. Lee, C. A. Myers, R. M. Neve, J. R. Semeiks, P. T. Spellman, K. Lorenz, E. H. Lee, M. H. Barcellos-Hoff, O. W. Petersen, J. W. Gray, and M. J. Bissell, "The morphologies of breast cancer cell lines in three-dimensional assays correlate with their profiles of gene expression," *Molecular Oncology*, vol. 1, no. 1, pp. 84–96, 2007.

[201] M. Werner, S. B. G. Blanquer, S. P. Haimi, G. Korus, J. W. C. Dunlop, G. N. Duda, D. W. Grijpma, and A. Petersen, "Surface curvature differentially regulates stem cell migration and differentiation via altered attachment morphology and nuclear deformation," *Advanced Science*, vol. 4, no. 2, p. 1600347, 2017.

[202] D. L. Deskins, D. Bastakoty, S. Saraswati, A. Shinar, G. E. Holt, and P. P. Young, "Human mesenchymal stromal cells: identifying assays to predict potency for therapeutic selection," *STEM CELLS Translational Medicine*, vol. 2, no. 2, pp. 151–158, 2013.

[203] A. Campbell, T. Brieva, L. Raviv, J. Rowley, K. Niss, H. Brandwein, S. Oh, and O. Karnieli, "Concise review: process development considerations for cell therapy," *STEM CELLS Translational Medicine*, vol. 4, no. 10, pp. 1155–1163, 2015.

[204] A. O. O. Odeleye, S. Castillo-Avila, M. Boon, H. Martin, and K. Coopman, "Development of an optical system for the non-invasive tracking of stem cell growth on microcarriers," *Biotechnology and Bioengineering*, vol. 114, no. 9, pp. 2032–2042, 2017.

APPENDIX A

2U-NET TRAINING PROCEDURE

A.1 Determination of loss function for U-Net 1

It was important to determine the most suitable loss function as it is a critical attribute for neural network optimization during training as well as for validating the model's performance. U-Net 1 loss function to segment cell regions was identified using the metrics sensitivity, precision, and DICE from five-fold cross-validation (Table A.1). Sensitivity and precision values served as indicators of the U-Net 1's ability to correctly detect MSC regions, while DICE gave a measure of how well the segmented cell regions and their borders overlapped with the ground truth. It was concluded that the DICE loss function gave better results, and hence, it was chosen for the overall training of the network. Though other loss functions had comparable sensitivity and DICE, they had lower precision values mainly due to regions being over-segmented.

Table A.1: Comparison of loss functions for training U-Net 1 model to segment cell regions. The metrics for the loss function that performed best during five-fold cross-validation are presented in bold.

Loss function	U-Net 1 five-fold cross validation metrics for segmenting cell regions		
	Sensitivity (S) (mean $\pm$ s.d.)	Precision (P) (mean $\pm$ s.d.)	DICE Score (mean $\pm$ s.d.)
DICE	**0.988 $\pm$ 0.012**	**0.847 $\pm$ 0.059**	**0.899 $\pm$ 0.014**
Weighted DICE	0.974 $\pm$ 0.015	0.801 $\pm$ 0.064	0.875 $\pm$ 0.013
Binary cross-entropy	0.968 $\pm$ 0.018	0.764 $\pm$ 0.048	0.859 $\pm$ 0.015
Weighted binary cross-entropy	0.959 $\pm$ 0.021	0.742 $\pm$ 0.044	0.820 $\pm$ 0.017

A.2 Determination of annotation method and loss function for U-Net 2

For U-Net 2, along with loss functions, marker annotation methods were also evaluated via five-fold cross-validation (Table A.2). Annotation methods (binarization, erosion, and H-minima transform) for training U-Net 2 were human-defined and generated by the algorithm, making it necessary to pick the appropriate method for localizing cell markers. In the binarization method, the gray-scale input image was inverted followed by thresholding with a high value. Then, only the object with the maximum area was retained inside each ground truth-defined cell. The erosion method involved morphologically eroding each cell in truth so that it shrank and became a cell marker. It led to the separation of the touching/overlapping cells. Also, it was ensured using the ground truth that each cell had only one object with the maximum area as the marker by removing all other eroded objects. Annotations using H-minima transform method were obtained as described in section 6.2.1.

Sensitivity and precision metrics were used as they were more relevant to assess cell marker localization rather than the overlap between algorithm-defined annotation and U-Net 2 segmentation (DICE score). A combination of H-minima transform and DICE loss was selected for training U-Net 2, owing to higher sensitivity and precision than other methods and loss functions. This also agreed with previous work on morphological image analysis of MSCs detailed in chapter 4, wherein H-minima transform was used as an initial step in detecting cell markers.

Table A.2: Comparison of marker annotation methods and loss functions for training U-Net 2 model to locate cell markers. The metrics for the combination of marker annotation method and loss function that performed best during five-fold cross-validation are presented in bold.

Marker annotation method	Loss function	U-Net 2 five-fold cross validation metrics for locating cell markers	
		Sensitivity (S) (mean $\pm$ s.d.)	Precision (P) (mean $\pm$ s.d.)
Binarization	DICE	0.857 ± 0.033	0.703 ± 0.051
	Weighted DICE	0.866 ± 0.029	0.721 ± 0.055
	Binary cross-entropy	0.819 ± 0.041	0.619 ± 0.049
	Weighted binary cross-entropy	0.823 ± 0.035	0.758 ± 0.046
Erosion	DICE	0.877 ± 0.026	0.885 ± 0.041
	Weighted DICE	0.817 ± 0.040	0.860 ± 0.045
	Binary cross-entropy	0.819 ± 0.038	0.843 ± 0.047
	Weighted binary cross-entropy	0.867 ± 0.029	0.734 ± 0.057
H-minima transform	DICE	$\mathbf{0.933 \pm 0.021}$	$\mathbf{0.900 \pm 0.042}$
	Weighted DICE	0.876 ± 0.024	0.844 ± 0.045
	Binary cross-entropy	0.857 ± 0.034	0.763 ± 0.049
	Weighted binary cross-entropy	0.914 ± 0.022	0.716 ± 0.053